Proceedings of VIAC in October 2026

October 23, 2026

Teaching, Learning and E-learning (IAC-TLEl)
Management, Economics and Marketing (IAC-MEM)
Transport, Logistics, Tourism and Sport Science (IAC-TLTS)



ISBN 978-80-88203-48-3

CTU
CZECH TECHNICAL UNIVERSITY IN PRAGUE

CIAE
Czech Institute of Academic Education z.s.

Impressum

Title:

"Proceedings of VIAC in October 2026", ISBN 978-80-88203-48-3

October 2026 in Prague, *1st edition*

Publisher / Creator of publication / Copyright holder:

Czech Institute of Academic Education z.s.

Address of Publisher:

Vodnicka 309/20, 149 00 - Prague 4, Czech Republic

Email: info@conferences-scientific.cz

Web: www.conferences-scientific.cz

Technical editorial staff:

Helena Kratochvílová

Radek Kratochvíl, Ph.D.

Název:

"Proceedings of VIAC in October 2026", ISBN 978-80-88203-48-3

Říjen 2026 v Praze, *První vydání*

Vydavatel / Tvůrce publikace / Nositel autorských práv:

Czech Institute of Academic Education z.s.

Adresa vydavatele:

Vodnická 309/20, 149 00 - Praha 4, Česká Republika

Email: info@conferences-scientific.cz

Web: www.conferences-scientific.cz

Technická redakce tohoto vydání:

Mgr. Helena Kratochvílová, DiS.

Ing. Radek Kratochvíl, Ph.D.

Conference Scientific Committee / Reviewers

Assoc. Prof. PhDr. Maria Janesova - Slovakia
- *Czech Technical University in Prague*

Mgr. Helena Kratochvilova - Czech Republic
- *Czech Institute of Academic Education*

Piotr Jednaszewski, Ph.D. - Scotland
- *St Mary`s Academy Ltd*

Assoc. Prof. Mario Konecki, Ph.D. - Croatia
- *University of Zagreb, Faculty of Organization and Informatics*

Assoc. Prof. Souvik Pal - India
- *Sister Nivedita University, Depart. of Computer Science & Eng,*

Prof. Emeritus Dr. Habil. Ioan-Gheorghe Rotaru, Ph.D. - Romania
- *Timotheus' Brethren Theological Institute of Bucharest*

Assoc. Prof. Dr. Ömer Beyhan - Turkey
- *Konya NE.University, A.K. Education Faculty*

Assoc. Prof. Anjali - India
- *Department of Computer Science & Engg, GNIT, Mullana*

Assoc. Prof. Jitender Grover - India
- *Department of Computer Science & Engg., M.M. University*

Radek Kratochvil, Ph.D. - Czech Republic
- *Czech Institute of Academic Education, Czech Technical University in Prague*

Assist. Prof. Dr. Mohamad H. Atyeh - Kuwait
- *School of Business, Australian College of Kuwait*

Assoc. Prof. Dr. Ramazan Sak - Turkey
- *Yüzüncü Yıl University, Faculty of Education*

Assoc. Prof. Dr. İkbal Tuba Şahin-Sak - Turkey
- *Yüzüncü Yıl University, Faculty of Education*

Assist. Prof. Dr. Ayşegül Derman - Turkey
- *Konya NE.University, A.K. Education Faculty*

Assist. Prof. Dr. Serdar Derman - Turkey
- *Konya NE.University, A.K. Education Faculty*

Assoc. Prof. Dr. Fethi Kayalar - Turkey
- *Erzincan University, Faculty of Education*

Prof. Dr. Thomas Rachfall - Germany
- *Hochschule Merseburg*

Dr. Dirk Förster-Trallo - Germany
- *Hochschule für Technik und Wirtschaft Berlin*

Luboš Fleischmann, Ph.D. - Czech Republic
- *Prague University of Economics and Business*

Editors and Technical publishing board:

Mgr. Helena Kratochvilova - Czech Republic
Radek Kratochvil, Ph.D. - Czech Republic
Assoc. Prof. Irena Konecki, Ph.D. - Croatia

Organizing Committee:

Mgr. Helena Kratochvilova - Czech Republic
Assoc. Prof. PhDr. Maria Janesova - Slovakia
Radek Kratochvil, Ph.D. - Czech Republic

Table of Contents

Human Capital Efficiency in Postgraduate Management Education: A Comparative Framework for Professional and Doctoral Study Pathways **07 - 11**
Róbert CZAKÓ

Barriers to Learning the State Language and Ways to Overcome Them for Azerbaijanis and Armenians Living in Georgia **12 - 18**
Magda CHIKABERIDZE, Tamar SHAISHMELASHVILI

A Literature Review of the Methodological Challenges of Using ChatGPT to Teach Writing **19 - 24**
Gábor GUTAI

Cost-Volume-Profit analysis: The case of Lithuanian manufacturing company **25 - 31**
Indrė KNYVIENĖ

The Relationship Between Self-efficacy and Emotional Intelligence in Pre-service Teachers **32 - 38**
Monika BOBAKOVA, Nikoleta IZDENCZYOVA

VAT Harmonisation and EU Integration in Albania **39 - 50**
Mikel ALLA

In Contemporary Society, Man is Increasingly Manipulated by Education, Media, and Subjected to the Economic Factor: The Timeliness of the Critique of the Philosophers of the Frankfurt School **51 - 57**
Jak SIMONI

Matching Governance to Cross-Government Problems: A Framework for Center of Government Involvement **58 - 70**
Elina PETROVSKA

Financial Distress Classification among Albanian Treasury Payment Recipients Using Explainable Machine Learning **71 - 84**
Mikel ALLA

Digital Marketing Opportunities and Risks in Spices and Herbs Market **85 - 103**
Andriy POPOVYCH

From Digital Crack Images to Mathematical Models: A Box-Counting Activity for Secondary Mathematics Education **104 - 113**
Tetiana ZHYLENKO

How Digitalization Translates into Operational Performance: The Mediating Role of Lean Management in Manufacturing Enterprises **114 - 120**
Klaus DÄNNER

Airports as the First Impression of a Destination: Examples from the United Arab Emirates **121 - 130**
Iva BULATOVIC, Gamal KHALIFA, Mahrha Hamad Saeed Saif ALHASSANI, Maitha Saif Juma Ali ALKAABI, Mariam Humaid Naser Humaid ALHOSANI

Developing Strategies for Cultivating a Culture of Learning in Educational Institutions **131 - 137**
Fethi KAYALAR

Knowing and Self-Knowing in the Teaching-Learning Process **138 - 142**
Pavol BURCL

Human Capital Efficiency in Postgraduate Management Education: A Comparative Framework for Professional and Doctoral Study Pathways

Róbert CZAKÓ[a*]

[a]*DTI University, PhD study programme: Economics and Management, Sládkovičova 533/20, 018 41 Dubnica nad Váhom, Slovak Republic, e-mail: ing.robertczako@elektron.sk*
ORCID: 0009-0004-6683-8495; Paper No.: IAC202610001

Abstract

The paper proposes a comparative framework for evaluating how professional and doctoral postgraduate study pathways contribute to human capital efficiency in management and economics. The starting point is the observation that practice-oriented management education and doctoral education create value through different mechanisms. Professional management programmes are usually connected with career acceleration, managerial competence and return on educational investment, whereas doctoral study develops research capacity, methodological rigor and long-term expert capital. The paper synthesizes human capital theory, management education literature and doctoral education indicators into a practical evaluation model. The proposed framework integrates economic, career, competence, knowledge-output, network and time-efficiency dimensions. It is intended for universities, employers and postgraduate learners who need a transparent basis for comparing educational investments. The contribution of the paper is conceptual and methodological: it offers an operational structure that can later be tested through survey data, career tracking and institutional records.

Keywords: human capital, postgraduate education, management education, doctoral study, educational investment

1. INTRODUCTION

Human capital has become a decisive factor of competitiveness in contemporary knowledge-based economies. Organizations no longer compete only through physical assets, financial resources or technological infrastructure. They increasingly compete through the knowledge, competences, judgement and adaptability of people. In this context, postgraduate education represents an important mechanism for developing the productive and strategic capacity of individuals and organizations.

Postgraduate management education is, however, not a homogeneous category. A professionally oriented management pathway, such as a Master of Business Administration programme, is typically linked to practical managerial development, leadership, networking and a faster transfer of knowledge into organizational practice. A doctoral pathway, by contrast, is based on scientific inquiry, methodological discipline, independent research and the production of new knowledge. Both forms of education may increase human capital, but they do so through different mechanisms and over different time horizons.

* Corresponding author.

This difference creates a practical evaluation problem. If postgraduate education is measured only through salary growth, the long-term research and expert value of doctoral study may be underestimated. If it is measured only through publications and citations, the managerial and organizational impact of professional management education may be ignored. A more balanced approach is therefore needed, especially for universities, employers and learners who must decide how to allocate time, money and effort in education.

The aim of this paper is to propose a comparative framework for assessing the contribution of professional and doctoral postgraduate study pathways to human capital efficiency. The paper does not claim to provide final empirical results. Instead, it develops a methodological model that can be applied in future empirical research and in practical decision-making.

2. HUMAN CAPITAL AND POSTGRADUATE EDUCATION

The concept of human capital is rooted in economic thinking about education, productivity and income. Education can be understood as an investment that increases the productive capacity of individuals and contributes to economic growth [1-3]. From this perspective, postgraduate education can be understood as an investment with direct costs, opportunity costs and expected future benefits.

Nevertheless, the value of postgraduate education is wider than immediate income growth. Human capital also includes knowledge, skills, health, motivation, adaptability and the capacity to create or apply ideas. Skills and competences are important not only for individual employability but also for social and economic development [4, 5]. In management and economics, these skills include analytical reasoning, strategic decision-making, communication, leadership and the ability to work with uncertainty.

A narrow financial interpretation of human capital may be insufficient for postgraduate education [6, 7]. A professional management programme may generate value through faster promotion, improved decision quality or better ability to manage teams and projects. A doctoral programme may generate value through research outputs, methodological competence, academic credibility and the ability to solve complex problems. These outputs are not directly comparable unless a broader set of indicators is used.

3. PROFESSIONAL AND DOCTORAL STUDY PATHWAYS

3.1. Professional management pathway

Professional management education is usually designed for learners who already have practical experience and who seek to strengthen their managerial effectiveness. Its value is connected with the application of knowledge in organizations. It may improve decision-making, financial understanding, leadership, negotiation, people management and strategic thinking [8, 9, 12].

The economic effect of this pathway is often measured through salary growth, promotion, career mobility or return on investment [8, 10]. However, these indicators should be interpreted carefully. Career growth may also be influenced by labour market conditions, prior experience, sectoral dynamics or personal networks. Therefore, a useful evaluation model should combine objective indicators with an attribution question: to what extent did the study itself contribute to the observed career change? [11, 12]

3.2. Doctoral pathway

Doctoral study develops a different form of human capital. It focuses on independent research, methodological rigor, critical thinking, academic writing and the ability to contribute to knowledge. Its outputs are usually visible in dissertations, publications, conference participation, citations and research collaboration [13, 14].

The economic effect of doctoral education is often indirect and delayed. Doctoral graduates may become academics, researchers, consultants, expert analysts or senior professionals whose value is based on specialized knowledge. This means that doctoral human capital should not be measured only by short-term income change. It should also be assessed through research competence, knowledge production, expert recognition and the ability to handle complex analytical tasks.

4. METHODS AND CONCEPTUAL DESIGN

The paper uses a conceptual methodology design. The purpose is to construct an evaluation framework by synthesizing existing theory and by translating it into measurable dimensions. This design is appropriate when the research objective is to clarify a problem, organize indicators and prepare a model for later empirical testing.

The framework was developed in four steps. First, the human capital literature was used to define education as an investment. Second, management education literature was used to identify indicators relevant to professional management programmes. Third, doctoral education literature was used to identify indicators relevant to research-oriented study. Fourth, the two sets of indicators were integrated into one comparative framework.

The framework can be used in future quantitative or mixed-method research. Quantitative data may include salary bands, changes in job position, study duration, publication outputs, number of implemented projects and professional collaborations. Qualitative data may include self-assessment, employer or supervisor assessment and short career narratives. The combination of these sources may reduce the risk of evaluating postgraduate education through one incomplete measure.

5. PROPOSED COMPARATIVE FRAMEWORK

The proposed framework is based on six dimensions: economic effect, career mobility, competence transfer, knowledge output, network capital and time efficiency. These dimensions allow the two pathways to be compared without forcing them into the same definition of success. The six dimensions and their suggested evidence are summarized in Table 1.

Table 1. Comparative indicator framework for postgraduate management education

Dimension	Professional management pathway	Doctoral pathway	Suggested evidence
Economic effect	Salary growth, promotion, return on educational investment	Expert premium, grant access, analytical/research roles	Baseline and follow-up career records, income bands, employer confirmation
Career mobility	Movement into managerial, project or strategic positions	Movement into research, expert, academic or methodological roles	Change in position level, responsibility scope and sector mobility
Competence transfer	Decision-making, leadership, financial and people-management competences	Methodological rigor, research design, academic writing and evidence evaluation	Graduate self-assessment, supervisor/employer assessment, documented outputs
Knowledge output	Business projects, process improvements, organizational innovation	Dissertation, publications, citations, conference outputs	Implemented projects, publications, bibliographic records and impact evidence
Network capital	Professional network, business contacts, alumni and managerial communities	Research network, co-authorship, scientific collaboration	Number and relevance of professional contacts and collaborations
Time efficiency	Time to observable career benefit and payback period	Time-to-degree and milestone completion	Study duration, workload, completion data and benefit timing

Source: author's own conceptual processing, 2026.

5.1. Human capital efficiency score

For practical use, the indicators may be summarized into a composite Human Capital Efficiency Score. The score should not be interpreted as a universal ranking of one pathway over another. It should instead be used as a decision-support tool that reflects the purpose of the evaluation.

A simple expression of the model is: HCES = w1(EE) + w2(CM) + w3(CT) + w4(KO) + w5(NC) + w6(TE), where EE means economic effect, CM career mobility, CT competence transfer, KO knowledge output, NC network capital and TE time efficiency. The weights should be defined before data collection. For example, an employer may assign higher weight to competence transfer and economic effect, while a university may assign higher weight to knowledge output and completion efficiency.

This scoring logic is useful because it makes hidden assumptions explicit. If an organization values immediate managerial performance, the professional pathway may receive a higher score. If it values research capacity, methodological expertise and long-term expert capital, the doctoral pathway may become more relevant. The model therefore supports contextual evaluation rather than a simple universal answer. The five-stage application process is presented in Table 2.

Table 2. Proposed evaluation process

Step	Purpose	Practical output
1. Baseline mapping	Describe the learner's position before entering postgraduate study.	Initial career, income, competence and motivation profile
2. Programme classification	Distinguish whether the programme is practice-oriented or research-oriented.	Selection of relevant indicator weights
3. Indicator measurement	Collect quantitative and qualitative indicators.	Comparable data matrix for both pathways
4. Attribution check	Separate the effect of education from natural career progression.	Graduate and employer/supervisor attribution score
5. Human capital interpretation	Translate indicators into managerial and economic meaning.	Recommendation for individuals, employers and institutions

Source: author's own conceptual processing, 2026.

6. DISCUSSION

The proposed framework has implications for three groups of stakeholders. First, universities can use it to describe and improve the value proposition of their postgraduate programmes. A business school can show how its programme supports managerial mobility and organizational performance, while a doctoral school can demonstrate how it develops research capability and expert competence.

Second, employers can use the framework when deciding whether to support employees in postgraduate education. If the organization needs immediate managerial improvement, a professional management programme may be more appropriate. If the organization needs analytical depth, innovation capacity or methodological expertise, doctoral study may be more useful. In both cases, the decision should be based on strategic need rather than on the prestige of a title alone.

Third, learners can use the model when choosing their own educational pathway. The decision between professional and doctoral study should be connected with career goals, preferred type of work, expected time horizon and personal motivation. A person seeking managerial acceleration may expect different outcomes than a person seeking research independence or expert recognition.

The framework also has limitations. It is conceptual and therefore requires empirical validation. Some indicators, such as salary growth or publication output, are relatively easy to measure. Others, such as competence transfer or network quality, require carefully designed questionnaires and qualitative interpretation. Future research should test the framework on graduates from different sectors and compare results across countries, institutions and study formats.

7. CONCLUSION

This paper proposed a comparative framework for evaluating human capital efficiency in professional and doctoral postgraduate management education. The main contribution is the integration of economic, career, competence, knowledge-output, network and time-efficiency indicators into one methodological structure.

The paper argues that professional and doctoral pathways create different but valuable forms of human capital. Professional management education is primarily connected with managerial application, career mobility and organizational value. Doctoral education is primarily connected with research capability, methodological rigor and long-term expert capital. These differences should be respected in any evaluation of educational investment.

The proposed framework can be used as a basis for future empirical research and for practical decision-making by universities, employers and learners. Its value lies in making comparison more transparent and in preventing the

reduction of postgraduate education to one narrow indicator. Future research should operationalize the model through longitudinal data, graduate surveys and employer or supervisor assessments.

Acknowledgements

The author declares that no external funding was received for the preparation of this paper.

References

[1] Schultz, T. W. (1961). Investment in human capital. The American Economic Review, 51(1), 1-17.
[2] Mincer, J. (1974). Schooling, experience, and earnings. National Bureau of Economic Research.
[3] Becker, G. S. (1993). Human capital: A theoretical and empirical analysis, with special reference to education. University of Chicago Press.
[4] OECD. (2001). The well-being of nations: The role of human and social capital. OECD Publishing.
[5] OECD. (2019). OECD skills strategy 2019: Skills to shape a better future. OECD Publishing.
[6] Wolfe, B. L., & Zuvekas, S. (1997). Nonmarket outcomes of schooling. International Journal of Educational Research, 27(6), 491-502.
[7] Marginson, S. (2019). Limitations of human capital theory. Studies in Higher Education, 44(2), 287-301.
[8] Baruch, Y. (2009). To MBA or not to MBA? Exploring value of the MBA programme. International Journal of Management Reviews, 11(4), 421-440.
[9] Boyatzis, R. E., & Saatcioglu, A. (2008). A 20-year view of trying to develop emotional, social and cognitive intelligence competencies in graduate management education. Journal of Management Development, 27(1), 92-108.
[10] Dau, L. A., & Richter, K. D. (2016). The MBA degree and global career success. Journal of World Business, 51(3), 420-432.
[11] Pfeffer, J., & Fong, C. T. (2002). The end of business schools? Less success than meets the eye. Academy of Management Learning & Education, 1(1), 78-95.
[12] Rubin, R. S., & Dierdorff, E. C. (2009). How relevant is the MBA? Assessing the alignment of required curricula and required managerial competencies. Academy of Management Learning & Education, 8(2), 208-224.
[13] Council of Graduate Schools. (2020). Graduate education and career outcomes: Data, metrics, and strategic planning. Council of Graduate Schools.
[14] Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America, 102(46), 16569-16572.

Barriers to Learning the State Language and Ways to Overcome Them for Azerbaijanis and Armenians Living in Georgia

Magda CHIKABERIDZE[a], Tamar SHAISHMELASHVILI [b*]

[a] *Associate Professor, Technical University of Georgia, Faculty of Social Sciences, Department of Georgian Philology and Media Technologies, 77 Kostava St., Tbilisi, Georgia, mchikaberidze@gmail.com*

[b] *Assistant Professor, Technical University of Georgia, Faculty of Social Sciences, Department of Georgian Philology and Media Technologies, 77 Kostava St., Tbilisi, Georgia, tamartamriko830@gmail.com*

Abstract

Georgia's education system faces significant challenges in improving ethnic minorities' proficiency in the state language and their overall level of general education. For more than twenty years, programs for teaching Georgian to ethnic minorities in Georgia have been funded. Both the state and international organizations devote substantial material and human resources to this purpose: training programs are financed, bilingual education is being introduced, and considerable attention is paid to the professional development of Georgian-language teachers and the improvement of Georgian-language textbooks. Nevertheless, the results achieved remain insufficient. The difficulties involved in teaching Georgian to ethnic minorities have several causes. These include low levels of community engagement, a shortage of qualified teachers proficient in Georgian in schools, an insufficient number of Georgian-language lessons (only three hours per week), a shortage of school textbooks aligned with contemporary standards, and outdated teaching methodologies. It is evident that both proficiency in the state language and access to quality education constitute the foundation of civic integration. An adequate command of the state language provides a solid basis for the civic integration of all citizens of Georgia. Fundamental changes are therefore required in this area.

Keywords: State language, Integration, Ethnic minorities,Bilingual education, 1+4 educational program

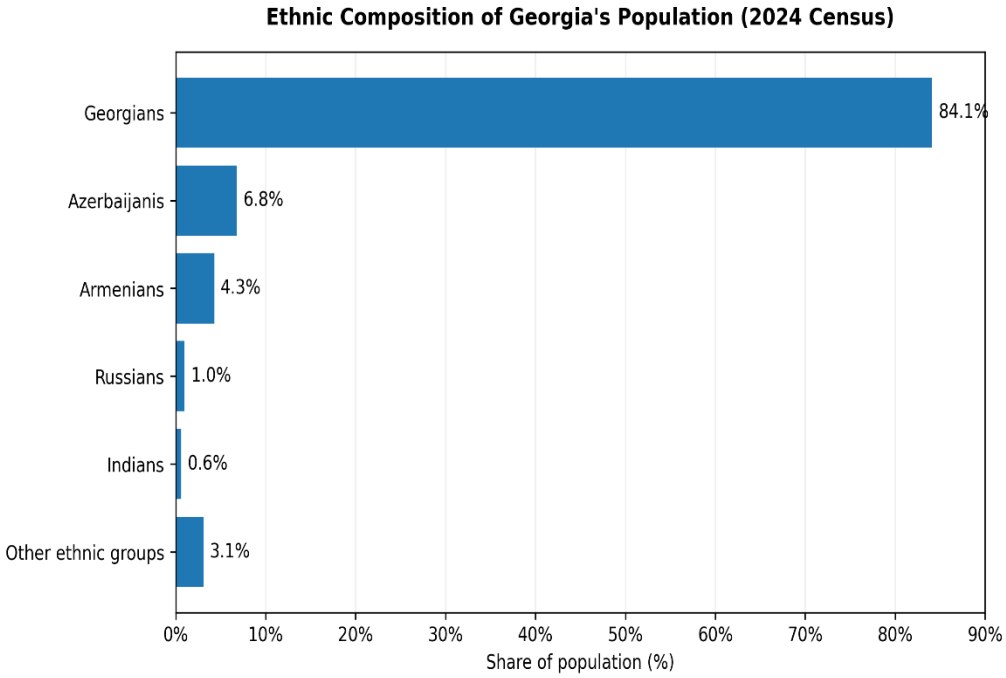
1. INTRODUCTION

The integration of ethnic minorities in Georgia remains a serious problem. A large proportion of the country's linguistic minorities do not have command of the state language and consequently face significant barriers. Lack of language proficiency restricts their social, political, and economic participation. The situation is particularly unfavorable in two regions of Georgia with compact ethnic-minority settlements: Kvemo Kartli, where Azerbaijanis live, and Samtskhe-Javakheti, where Armenians live. A large proportion of the Azerbaijani and Armenian populations in these regions live in relative isolation from Georgian society.

* Corresponding author.

According to the 2024 census data of the National Statistics Office of Georgia, ethnic minorities account for approximately 13% of Georgia's population, with Azerbaijanis (6.8%) and Armenians (4.3%) constituting the largest groups by percentage.

Table 1 - Ethnic Composition



Source: National Statistics Office of Georgia (Geostat), 2024 Population Census.

Full data: Georgians: 84.1%; Azerbaijanis: 6.8%; Armenians: 4.3%; Russians: 1%; Indians: 0.6%; other ethnic groups: 3.1%.

One of the most powerful means of fostering the civic integration of ethnic minorities living in Georgia is education, including education received at Georgian higher education institutions. Accordingly, the Government of Georgia and international organizations are taking significant steps in this direction.

Since **2005**, Georgia has had a National Concept for Tolerance and Civic Integration of Ethnic Minorities. According to this concept, alongside other factors, access to quality education and improved proficiency in the state language are of paramount importance. In **2009**, the State Concept for Tolerance and Civic Integration was also developed.

Thus, for more than twenty years, programs for teaching Georgian to ethnic minorities in Georgia have been funded. To this end, both the state and international organizations expend substantial material and human resources. Alongside general reforms in Georgia's education sector, training programs are financed, bilingual education is being introduced, and considerable attention is paid to teacher retraining and the improvement of textbooks.Nevertheless, the results remain insufficient.

2. MAIN TEXT

2.1. General Overview and the Purpose of the 1+4 Educational Program

According to 2020 data, approximately 75% of Azerbaijanis living in Georgia report that they either do not know Georgian at all or have only a basic command of it. This figure is particularly high in the Kvemo Kartli region, where Azerbaijanis live in compact settlements and use only Azerbaijani in everyday life.

Under the Law of Georgia on General Education, citizens of Georgia for whom Georgian is not their native language have the right to receive complete general (school) education in their native language (Azerbaijani, Armenian, Russian, etc.). At the same time, the study and teaching of Georgian is compulsory in non-Georgian-language general education schools/sectors. In addition, since 2011, pupils in non-Georgian-language schools have

been required to take a test in Georgian as a second language in order to receive a school-leaving certificate. Nevertheless, the quality of Georgian-language teaching in non-Georgian-language schools remains low. After graduating from school, applicants often continue their studies abroad (in Russia, Azerbaijan, or Armenia), which distances them even further from the Georgian state.

To address this problem, a preferential policy for national minorities was introduced from the year in which the Unified National Examinations began **(2005)**. This policy remains in force and will continue until the level of Georgian-language proficiency and general education in non-Georgian-language schools improves.

Since 2008, non-Georgian-speaking applicants have been permitted to take the Unified National Examinations in their native language, which has significantly increased the number of students. In particular: "In 2008, 250 Azerbaijani-speaking and 113 Armenian-speaking applicants registered for the national examinations for admission to higher education. By 2011, the number of Azerbaijani-speaking applicants had increased to 335 (by 34%), while the number of Armenian-speaking applicants had increased to 253 (by 124%)" (1).

Most importantly, **since 2010**, the 1+4 Georgian Language Preparatory Educational Program has operated within Georgia's higher education system. Its purpose is to address the deficiencies that non-Georgian-speaking students carry over from school. During one year, students study Georgian and acquire the skills and knowledge required for higher education. After accumulating 60 credits, they become eligible to pursue a bachelor's degree program in Georgian.

Tuition for the Georgian Language Preparatory Educational Program is fee-based (GEL 2,250). Despite the cost, the number of applicants increases from year to year.

According to data from the Office of the State Minister of Georgia for Reconciliation and Civic Equality, whereas the total number of students enrolled in this program (ethnic Armenian and Azerbaijani students) was 301 in **2010**, by **2023** it had risen to 1,545, and the number continues to increase annually.

Work on state-language teaching programs began at the end of the twentieth century, in the first years following the restoration of Georgia's independence, while from the twenty-first century onward the policy of intercultural education became more refined and consistent.

2.2. Historical Overview

In the nineteenth century, Georgia formed part of the Russian Empire and was divided between two Russian governorates (The Tiflis and Kutaisi Governorates). The education of the non-Russian population, including ethnic minorities, was influenced by imperial policy. From the outset, the imperial authorities actively introduced Russian into schools. Georgian was subjected to particularly severe suppression. Georgian public figures had to struggle actively to preserve their native language.

The situation was comparatively better in the first quarter of the nineteenth century, when the Russian Empire was consolidating its position in the Caucasus. By 1820, various languages were taught at the Tiflis School for the Nobility (Present-day Tbilisi Classical Gymnasium No. 1), namely Russian, Georgian, Azerbaijani, Armenian, Latin, and German. Knowledge of Georgian was compulsory even for Russian teachers and pupils. In 1830, the school was reorganized as a gymnasium.

By 1855, there were approximately 150 educational institutions in Georgia.

In 1876, the Gori Teachers' Seminary (Transcaucasian Teachers' Seminary) was opened. It played an important role in training teaching personnel not only for Georgia but for the entire Transcaucasus.

Instruction at the Gori Teachers' Seminary was conducted in Russian. Georgian, Armenian, and Azerbaijani were taught as separate subjects. Georgian, Russian, Armenian, and Azerbaijani primary schools operated alongside the seminary.

At various times, individuals who later became renowned public and cultural figures studied there: the great Georgian poet and humanist Vazha-Pshavela and his brothers, including the writers Bachana and Tedo Razikashvili; Georgian educator, folklorist, and public figure Lado Aghniashvili; the Azerbaijani writer, politician, and publicist Nariman Narimanov; the Azerbaijani composer and conductor Muslim Magomayev, grandfather of the famous Azerbaijani singer and baritone Muslim Magomayev; and the Azerbaijani composer and public figure Uzeyir Hajibeyov, among others.

Georgian writers and public figures were actively involved in teaching the Georgian language. In this regard, the contributions of **Ekaterine Gabashvili** (1851-1938) and **Soprom Mgaloblishvili** (1851-1925) are particularly noteworthy.

From 1868, Ekaterine Gabashvili headed a **girls' school** founded by the clergy of the **Armenian Church** in Gori. Georgian and Armenian children studied together in one group and received their primary education in Georgian.

"At the time I knew nothing about pedagogical theories," Ekaterine Gabashvili recalls, "but I clearly understood that the children of Gori residents had to begin their education in Georgian, because they did not know a word of Armenian. And that is exactly what I did: irrespective of nationality, I had everyone begin studying in Georgian, using Gogebashvili's alphabet and a method that was entirely new to me, since I myself had been taught according to the old method" (2).

In his memoirs, Soprom Mgaloblishvili describes the environment that prevailed at the **Gori Theological School** when he was a pupil. According to the memoirist, "the instruction there was neither Russian nor Georgian; it was some kind of mixture of languages. None of us understood Russian, so they were compelled to speak to us in Georgian; translation was done from Russian into Georgian and from Georgian language into Russian. Georgian itself was not regarded as a subject" (3).

It should be noted that such a "mixture of languages" is a serious methodological error.

In 1873, Soprom Mgaloblishvili returned to the Gori Theological School as a teacher and worked there for twenty years. "There were eight Georgian-language lessons at the school; no money had been allocated from the treasury for this, and the teachers did not particularly trouble themselves with teaching Georgian. I undertook the teaching without remuneration" (4). As a result, "the young teacher greatly increased the pupils' interest in Georgian-language lessons, raised the standard of its teaching, and restored it to its rightful place" (5).

In the 1880s and 1890s, when a deliberate policy of Russification was being pursued, the Georgian language was left dependent on a handful of such enthusiasts. The native language was being driven out of Georgian schools. Under the curriculum devised by **Kirill Yanovsky**, Trustee of the Caucasian Educational District, the Russian imperial government intensified the marginalization of Georgian in Georgian schools and introduced Russian as the language of instruction for all subjects.

"The native language (During the period of reaction) in Georgian schools did not enjoy rights equal not only to those of any foreign language (Classical or modern), but even to the languages of other peoples living in the territory of Georgia (Armenians, Ossetians, Tatars, and others)" (6).

From the very first years of its independence (1918 - 1921), Georgia set itself the goal of addressing important issues concerning ethnic minorities. The Constitution of Georgia of 1921 granted national minorities broad rights. Today, numerous state and private programs for teaching the state language (Georgian) operate in Georgia.

With the support of the Government of Georgia, the State Language Strategy for 2021-2030 (a unified program) was developed. It provides for increased access to state-language learning for ethnic minorities. Particular attention will be paid to regions densely populated by ethnic minorities. New programs/courses will be developed within the framework of the State Language Teaching and Integration Program.

2.3. Difficulties in Teaching Georgian and Useful Recommendations

Low levels of community engagement constitute a significant barrier to state-language acquisition among linguistic minorities. Unfortunately, some members of ethnic minorities do not have a clear understanding of how knowledge of Georgian can assist them in employment and career advancement. The reasons for this can, to some extent, be sought in the historical past.

The dissolution of the Soviet Union meant that Russian ceased to function as a means of interethnic communication not only in Georgia but across the post-Soviet space generally. Russian – the sole means of such communication – "has significantly lost ground over the past fifteen years" (7).

We believe that mistakes made in the education system over the years are responsible for the fact that representatives of ethnic minorities often know only their native language and, quite apart from Georgian, can no longer communicate with other ethnic groups even in Russian. Naturally, all of this leads to isolation.

Azerbaijanis living in compact settlements (mainly in Kvemo Kartli) and Armenians (mainly in Samtskhe-Javakheti) use their native languages or Russian (particularly the older generation) in everyday communication, because local communities are detached from the Georgian-language information environment, which in turn creates an information vacuum.

Students themselves confirm this. "Because of the language barrier, there is an information vacuum in Ninotsminda and Akhalkalaki; the population does not have information about developments taking place in the country. Therefore, I decided to enroll in the preparatory program in order to learn Georgian and obtain higher education", – say students from Meskheti-Javakheti, an area densely populated by Armenians. The same point is made by students from Marneuli, an area densely populated by Azerbaijanis, who study in the 1+4 Georgian Language Preparatory Educational Program at Georgian Technical University.

In multicultural regions, the lack of everyday communication with Georgian-speaking residents hinders the natural acquisition of the language. Pupils have no opportunity for live communication and language practice outside the classroom.

In regions populated by minorities, Georgian is studied as just one ordinary school subject. At the same time, very little time – only three hours per week – is allocated to teaching Georgian language and literature in non-Georgian-language schools.

Outdated methodologies for teaching **Georgian as a second language** in schools, including mechanical learning and rote memorization, do not meet contemporary educational standards. Instruction often emphasizes grammar rather than the development of communicative skills.

This problem is compounded by another difficulty: the insufficient number of Georgian-language school textbooks, as well as the failure of existing Georgian-language textbooks to correspond to contemporary standards and pupils' actual needs.

We surveyed students enrolled in the 1+4 program. Both Armenian and Azerbaijani students who had graduated from schools in the regions noted that high-quality and interactive learning materials, applications, and literature are less accessible there. In some villages in the regions, Georgian is taught by teachers who lack appropriate professional qualifications. The shortage of qualified teachers proficient in Georgian weakens the quality of instruction.

Table 2 - Characteristics of a Survey

Characteristics of a Survey	
What factors hinder learning the Georgian language?	
Lack of textbooks designed to meet modern standards	20%
Shortage of qualified teachers	40%
Outdated teaching methodology	20%
Lack of Georgian language lessons	15%
Other factors	5%

Kakha Gabunia, Professor at the Faculty of Humanities of Tbilisi State University and Head of the Multilingual Education Program at the Faculty of Psychology and Educational Sciences, regards the noticeable increase in non-Georgian-speaking students at Georgian-language higher education institutions as a positive development. The 1+4 program operates in more than 30 higher education institutions, and students spend one year preparing in Georgian. However, he also notes that there is a very large disparity in knowledge between Georgian-speaking and non-Georgian-speaking students. In other words, as the years pass, the situation in terms of educational provision is deteriorating –specifically in non-Georgian-language schools.

"In approximately 300 schools (214 independent schools and around 80 sectors), there are teachers who are well beyond retirement age. They cannot be retired, but this means that we are lowering the quality of knowledge. This has resulted in a catastrophic shortage of teachers in these schools" (8).

With regard to methodology, the most serious errors in teaching Georgian to ethnic minorities are the mixing of languages and the insufficient and unsystematic use of practical-functional grammar, whereas practical-functional grammar should constitute the principal foundation of teaching methodology. This applies to the teaching of both second and foreign languages, as well as to the teaching of native-language grammar at the age of figurative-associative thinking (up to the age of 9-10) (9).

In all three cases, attempts to teach theoretical grammar and attempts to teach a language through grammar are as ineffective as, for example, trying to teach swimming in a classroom by giving instructions, demonstrating movements, and even practicing those movements.

Another methodological error is teaching reading and, even more so, writing when a child has not yet acquired even the most elementary level of speaking and listening. Both non-Georgian children learning Georgian and Georgian children learning a foreign language become victims of this error.

The Ministry of Education of Georgia was compelled to introduce the 1+4 program and devote the first year of higher education to learning the state language. This was necessary for young school graduates who had, at least formally, received Georgian-language lessons at school. Primary-school pupils likewise require preparation before they begin learning to read and write in Georgian and studying practical functional grammar.

From a methodological perspective, our state language functions as a **second language** for Azerbaijani children living in close proximity to Georgians; however, in densely populated Azerbaijani communities, Azerbaijani children have virtually no contact with Georgian. Therefore, from a methodological perspective, Georgian is a **foreign language** for them. Accordingly, neither second-language teaching methodologies nor integrated teaching methodologies (for example, teaching mathematics together with Georgian) prove effective.

A solution: In our view, in densely populated Azerbaijani and Armenian communities, children aged 5 - 6 should spend an entire year learning Georgian at a basic level of speaking and listening (A1 or A2). During this period, other subjects should not be taught; Georgian-language instruction should be intensive, otherwise one year will not be sufficient. Thereafter, in the following year, it will be possible to teach reading and writing in Georgian as well as to introduce integrated instruction. Naturally, both the teaching methodology and textbooks must be methodologically sound.

The problems mentioned above are compounded by a shortage of specialists in **bilingual education** in regional schools. Yet bilingual/multilingual education is regarded as "a powerful instrument of educational policy for ethnic minorities" (10).

The purpose of bilingual education, after all, is "on the one hand, to ensure an adequate command of the state language and, on the other, to preserve and develop minority languages" (11).

It is true that the number of ethnic-minority applicants and students at Georgian higher education institutions increases from year to year, and this is undoubtedly a result of successful higher education reform. Nevertheless, it must also be acknowledged that bilingual education in schools continues to be implemented with shortcomings.

Most teachers in schools who have been trained in the relevant bilingual methodology do not themselves have an adequate command of Georgian, while Georgian-language lessons are often merely formal in character.

Research confirms that a professional and certified bilingual teacher is one of the most important factors in bilingual education. Therefore, we believe that priority should be given, above all, to ensuring a high level of teacher qualification.

For students enrolled in the Georgian Language Preparatory Educational Program, acquiring a second language, in our view, primarily entails the gradual expansion of vocabulary. A lecture-based format is inappropriate. The teacher should not do most of the talking. Students should be actively involved at every stage of the lesson.

The most effective form of integrated instruction is independent work by students in pairs or groups. Appropriately and thoughtfully selected activities are a prerequisite for the successful outcome of an integrated lesson.

Role-playing, situational exercises, question-and-answer activities, and purpose-designed educational crosswords help language learners develop their thinking skills.

It is noteworthy that last year the Department of Georgian Philology at Georgian Technical University published a two-volume textbook entitled **Linguistics and Geography**, for students in the 1+4 program. The books were edited by **Professor Temur Jagodnishvil**i, Doctor of Philological Sciences.

The course Linguistics and Geography is being introduced into the Georgian educational sphere for the first time. Within the framework of the Georgian Language Preparatory Educational Program, it is designed to support the teaching of Georgian and represents the first step of its kind in the Georgian context. It combines not only linguistic components of language learning but also cultural components.

Linguistics and Geography (Linguoculturology) represents a symbiosis of two traditional fields of knowledge – it is interdisciplinary, that is, an educational course created through the combination of two fields of knowledge, and it offers a new method for learning a language (a second language or a foreign language). It does not regard language as an isolated phenomenon, nor does it consider grammar alone to be the path and means of learning it; rather, it approaches language within the context of the culture of which language itself is one of the most important value-bearing elements", – Professor Temur Jagodnishvili writes in the preface (12).

3. CONCLUSION

Alongside the development of effective mechanisms by the state, the existing situation can indeed be improved through educational programs in order to achieve full integration.

It is noteworthy that ethnic minorities have a desire to learn Georgian, which offers an important means of realizing their potential and advancing their careers. However, reality presents a different picture. State mechanisms for teaching Georgian have not yet been fully studied and analyzed.

The principal problems arising from a lack of knowledge of Georgian as the state language can be resolved only through effective reform. For the full integration of ethnic minorities, properly implemented reforms and a sound analysis of existing shortcomings are required. It is important that the state and public authorities develop the necessary recommendations.

It is evident that both proficiency in the state language and access to quality education form the foundation of civic integration. Radical changes are therefore needed in this area, since minor adjustments will not be sufficient to improve the situation.

Georgia, for which European orientation constitutes a major political priority, should implement integration-oriented programs specifically in relation to ethnic minorities that will significantly change the existing situation.

References

[1] Malazonia, D. (2017). Intercultural Education: Problems, Their Analysis and Prospects for Development in Georgia. Tbilisi: Ilia State University. (pp. 2 -3).
[2] Gabashvili, E. (1960). Selected Writings, Vol. II. Tbilisi: (p. 234).
[3] Mgaloblishvili, S. (1938). Memoirs. Tiflis. (p. 38).
[4] Mgaloblishvili, S. (1938). Memoirs. Tiflis. (p. 71).
[5] Minashvili, L. (2017). Essays on the History of Georgian Literature. Tbilisi. (p. 208).
[6] Sarishvili, T. (1951). School and Public Education in Pre-Revolutionary Georgia (1901-1921). A Short Essay. Tbilisi. (p.9).
[7] Dumbadze, N. (2019). The Ethnic Composition of the Population of Georgia .Tbilisi (p.71). Retrieved from https://openscience.ge/bitstream/1/786/1/samagistro%20dumbadze%20nino.pdf
[8] Toradze, M. (2023). Why Is It Difficult to Find Teaching Staff for Approximately 300 Non-Georgian-Language Schools? Retrieved from https://gazeti.tsu.ge/araqartulenovani-skolebistvis
[9] Vakhania, Z. (2009). Georgian Philology, Vol. I.. Tbilisi. (pp. 35-40).
[10] UNESCO (2003). Education in a Multilingual World. UNESCO Education Position Paper. Paris.
[11] Gabunia, K. (2010). Peculiarities of Integrated Language and Content Teaching in a Bilingual Classroom. Bilingual Education, 1, 29-35.
[12] (2025). Linguistics and Geography. Vol. 1. Tbilisi: Universal. (p.3).

A Literature Review of the Methodological Challenges of Using ChatGPT to Teach Writing

Gábor GUTAI[a*]

[a] *Constantine the Philosopher University in Nitra, Faculty of Education, Department of English Language and Culture, Dražovská 4, 949 74, Nitra, Slovakia, ORCID: 0009-0000-2913-1360, ggutai@ukf.sk*

Abstract

ChatGPT has moved rapidly from experimental technology to a routine, though unevenly regulated, component of university writing. Research reports benefits in brainstorming, language support, formative feedback, and revision, especially for students writing in an additional language. Yet claims about its instructional effectiveness remain difficult to interpret because the technology, the writing tasks, and the conditions of use change faster than conventional educational research designs can accommodate. This literature review shows methodological challenges in research and teaching practice concerning ChatGPT-supported writing in higher education. The review presents seven connected challenges: unstable treatment definitions, weak causal designs and inadequate comparison conditions, incomplete observation of human-AI interaction, uncertain measurement of writing quality and learning transfer, authorship and academic integrity ambiguity, inequitable access and uneven AI literacy, and limited ecological, disciplinary, linguistic, and longitudinal validity. The review argues that these are not merely technical limitations. They affect what educators can claim students have learned and whether assessment results represent students´ independent or AI-mediated capabilities. A stronger methodology treats ChatGPT use as a documented writing process rather than a binary exposure, separates product quality from learning, records model and prompt conditions, and uses multiple human and performance measures. The paper concludes with a research and teaching framework centered on process evidence, calibrated human oversight, transparent permitted-use rules, and program-level assessment design.

Keywords: ChatGPT, academic writing, methodology, feedback, generative artificial intelligence

1. INTRODUCTION

The public release of ChatGPT in November 2022 altered the practical conditions under which university writing is taught and assessed. Students can now request outlines, explanations, examples, sentence-level revisions, simulated peer review, and complete drafts within one interface. These functions overlap with activities traditionally used to teach planning, drafting, feedback, revision, and editing. Their availability has encouraged experimentation in English for academic purposes, first-year composition, postgraduate writing, and writing-intensive disciplinary courses. It has also made an old methodological problem more urgent: a polished text does not necessarily show who performed the intellectual work or what the writer learned.

* Corresponding author.

Early reviews described a mixture of instructional opportunities and risks. ChatGPT may offer immediate individualized support, reduce the delay between drafting and feedback, and help multilingual writers notice language problems (Lo, 2023; Tlili et al.,2023). Studies of tertiary English writing have reported scores of positive perceptions when students use ChatGPT as a feedback or revision tool (Ding et al., 2025; Mahapatra, 2024; Marzuli et al., 2023). These findings justify continued investigation, but they do not establish that unrestricted ChatGPT use improves durable writing competence. In the aforementioned studies, the outcome is the quality of a text produced with access to the same tool that constitutes the intervention. Such a design can show assisted performance; it cannot by itself show that students can plan, argue, revise, or edit more effectively after the assistance is removed.

The distinction between assisted performance and learning is central to writing pedagogy. Writing is both a product and a cognitive, social, and rhetorical process (Lo, 2023). A writer interprets a task, develops a position, selects evidence, anticipates readers, drafts, evaluates feedback, and decides what to revise. ChatGPT can participate in each stage, but its contribution varies with the prompt, model version, interface, language, subscription tier, and the user's capacity to judge its output (Tlili et al., 2023).

2. OBJECTIVES

Higher education institutions face a related assessment problem. If generative AI can produce acceptable responses to conventional assignments and its use cannot be detected reliably, educators cannot infer learning from the final product alone. Guidance from UNESCO (Miao & Holmes, 2023) and the Australian Tertiary Education Quality and Standards Agency (Lodge et al., 2023) consequently emphasizes human agency, AI literacy, transparent rules, attention to the learning process, and multiple sources of assessment evidence. These principles are especially relevant to writing because written products are often used both to develop learning and to certify individual attainment.

This article reviews the methodological challenges of using ChatGPT to teach writing in higher education. Its focus is narrower than a general review of benefits, ethics, or educational applications. It asks three questions:

1. What methodological problems limit current claims about ChatGPT-supported writing instruction?

2. How do these problems affect the design of writing activities and assessment?

3. What research practices would produce more credible and transferable evidence?

The central argument is that ChatGPT should be studied as a variable, recorded form of participation in the writing process. Credible research must specify what the system did, what the student did, and what evidence demonstrates learning beyond the jointly produced text. This framing permits constructive instructional use while preserving the distinction between assistance, achievement, and authorship.

3. REVIEW METHOD

This article uses a structured narrative review because the literature is heterogeneous in purpose, design, participants, writing genres, AI systems, and outcomes. A meta-analysis would require commensurable interventions and measures that the present evidence base does not consistently provide.

Targeted searches were conducted for English-language literature published between November 2022, when ChatGPT became publicly available and September 2026. Search terms combined *ChatGPT* with *writing instruction, academic writing, EFL writing, feedback, revision, higher education, assessment, authorship, academic integrity, critical thinking,* and *methodology.* Priority was given to peer-reviewed studies of postsecondary writing, critical reviews of generative AI in education, and authoritative policy documents with direct implications for university teaching and assessment. Conceptual papers were retained when they clarified a methodological issue not yet adequately examined in empirical work.

Sources were included when they addressed at least one of four areas: the effect of ChatGPT or a closely related large language model on writing performance; student or teacher interaction with AI-generated writing or feedback; assessment and authorship in higher education; or research-design problems relevant to evaluating generative AI. Sources centered solely on technical model performance, noneducational professional writing, or school education without a transferable higher-education implication were excluded. Because the purpose was analytical rather than exhaustive, the review does not claim database-level completeness. The rapidly changing publication record and inconsistent indexing of early online articles are further limitations (Ding et al., 2025; Mahapatra, 2024).

The selected literature was read for reported instructional use, comparison condition, outcome measure, evidence of the writing process, model documentation, duration, educational context, and stated limitations. Themes were developed iteratively by grouping problems that changed the interpretation of instructional effectiveness. The synthesis was also checked against recent empirical studies of AI feedback and student revision, since these studies test a prominent claim about the value of ChatGPT for writing (Lo, 2023).

3.1 The Problem of an Unstable Instructional Treatment

Model and interface change

Educational interventions are usually described well enough to be reproduced. ChatGPT challenges that expectation (Lodge et al., 2023). A study may identify the product name but omit the model version, date of access, subscription tier, system settings, enabled tools, memory state, or whether students used custom instructions. Yet these details affect output quality and interaction. The same prompt can produce different answers across sessions, and a platform update can change behavior during data collection. Research published months later may evaluate a system that users can no longer access in the same form (Tlili et al., 2023).

This instability creates construct and replication problems. If the independent variable is named only as "ChatGPT use," readers cannot determine whether an observed outcome resulted from the general availability of a conversational model, a particular model's capability, the prompts provided by researchers, or students' improvised strategies. Findings may also be attributed to ChatGPT when the real intervention includes extensive teacher scaffolding, a detailed rubric, prompt templates, peer discussion, and repeated opportunities to revise. Those components are educationally valuable, but they must be separated analytically from the model (Ding et al., 2025; Mahapatra, 2024).

Researchers should report the model and interface, access dates, account conditions, language, temperature or other settings when available, the full task instructions, and all researcher-provided prompts. If students develop prompts independently, studies should preserve interaction logs with appropriate consent and redact personal data. Repeated benchmark prompts can document whether output changes over the study period. Reporting should distinguish platform-level affordances from the pedagogical design wrapped around them.

Variation in the meaning of use

The phrase "using ChatGPT for writing" covers interventions with different learning mechanisms. Idea generation may reduce the difficulty of starting a task but can anchor students to generic positions. Explanatory dialogue may build knowledge, whereas requesting a complete draft may replace knowledge construction. Grammar feedback may support noticing, but direct rewriting can conceal whether the student understood the change. Rubric-based feedback may direct revision, yet its effect depends on whether students evaluate and act on it.

This heterogeneity appears across the literature. Mahapatra (2024) examined tertiary ESL academic writing through a mixed-methods intervention and found improvement alongside positive student perceptions. Marzuki et al. (2023) reported that university EFL teachers perceived several AI writing tools as improving content and organization. Studies focused on automated writing evaluation examine a different mechanism: students receive feedback and decide how to revise (Ding et al., 2025). Treating these activities as equivalent obscures which part of writing development is supported.

A useful taxonomy should describe the AI role at each stage: task interpretation, knowledge building, planning, drafting, feedback, revision, language editing, source work, or metacognitive reflection. It should also record the level of AI contribution, from critique of student-authored text to generation of text incorporated into the submission. This process-based description makes studies comparable without pretending that all uses are the same treatment.

3.2 Weak Casual Designs and Inadequate Comparisons

Self-selection and novelty

Much early research relies on voluntary adoption, student surveys, short classroom trials, or comparisons between pre- and post-task scores without a credible control group. Students who use ChatGPT skillfully may already have stronger digital literacy, higher motivation, better English proficiency, or more confidence in experimentation. Conversely, students who seek extensive assistance may begin with greater writing difficulty. Without measuring these differences, an association between AI use and writing quality is hard to interpret (Lo, 2023).

Novelty also affects attitudes and engagement. Students may report high satisfaction because immediate conversational feedback is new, fast, and less socially threatening than asking a teacher. Positive perceptions matter for adoption, but they are not equivalent to learning. Perceived usefulness can coexist with inaccurate feedback, shallow revision, or reduced independent effort. Reviews show broad variation in students' reactions to ChatGPT, indicating that attitudes are shaped by context and experience rather than by the technology alone (Lo, 2023; Tlili et al., 2023).

Stronger studies should use random assignment when feasible or measure baseline writing ability, language proficiency, AI literacy, prior AI use, motivation, and epistemic beliefs. A pretest should be equivalent in genre and difficulty but not identical to the intervention task. Analyses should report attrition, contamination between groups, and whether control participants independently accessed generative AI. Qualitative evidence can explain why an intervention worked, but it should complement rather than substitute for performance evidence (Mahapatra, 2024).

The choice of comparison conditions

No-tool control is easy to interpret but may answer a limited question. University writers already use dictionaries, search engines, grammar checkers, peers, writing centres, and teacher feedback. Comparing ChatGPT with no support may exaggerate its practical advantage. Comparing it with teacher feedback alone can also be misleading if the AI produces much more feedback, responds instantly, or receives a detailed prompt while teachers work under ordinary time constraints (Lo, 2023).

Farrokhnia et al. (2026) illustrate why comparison design matters. In a randomized study of 70 graduate students, a prompting condition produced feedback rated more highly than both a simpler AI condition and teacher feedback, but the higher-rated feedback did not yield significantly better revisions. Teacher feedback led to comparable gains. The result separates feedback supply from feedback uptake: producing more or apparently better comments does not ensure that students can interpret and apply them.

Future trials should compare pedagogically plausible alternatives: teacher feedback, trained peer feedback, established automated writing evaluation, ChatGPT with minimal prompting, ChatGPT with explicit evaluation scaffolds, and hybrid teacher–AI feedback. Time-on-task and quantity of feedback should be measured. If an AI condition receives unlimited interaction while another condition receives one short comment, the estimated effect combines the feedback source with unequal instructional dosage (Tlili et al., 2023).

3.3 Incomplete Observation of Human AI Interaction

Final texts hide the process

The final essay is an inadequate record of an AI-mediated writing process. It does not reveal whether the student generated ideas, accepted suggestions uncritically, corrected false information, rejected generic prose, or used AI only for language editing. Two essays with similar scores may represent very different levels of learning and agency. This creates a methodological blind spot in both research and classroom assessment (Tlili et al., 2023).

Interaction logs offer a partial solution. They can show prompt sequences, pasted text, model responses, and iterative decisions. Version histories can reveal how a draft changed. Stimulated-recall interviews can ask students why they accepted or rejected a suggestion. Screen recordings and think-aloud protocols provide richer evidence, although they can alter normal behaviour and create substantial coding burdens. No single method is sufficient. Logs show what occurred in the interface but not necessarily what the writer understood; retrospective explanations may be incomplete or self-protective (Farrokhnia et.al., 2026).

Research should therefore triangulate traces, texts, and explanations. A practical design could combine timestamped drafts, exported chat logs, a short decision commentary, and a post-task interview for a subsample. Revision coding should distinguish surface changes from meaning-level changes, such as strengthening a claim, qualifying evidence, reorganizing an argument, or correcting a conceptual error. These data make it possible to analyse student agency instead of inferring it from the polished product (Ding et al., 2025).

Prompting is a learner variable

Prompt quality is often treated as a technical detail, yet it is part of the competence under study. Students who specify audience, genre, criteria, and desired feedback are likely to receive more useful responses. They also need enough subject knowledge to detect plausible errors and enough writing knowledge to request appropriate assistance.

Prompting skill can therefore mediate the relationship between ChatGPT access and writing outcomes (Ding et al., 2025).

Providing identical prompts increases experimental control but reduces ecological validity. Allowing unrestricted prompting reflects authentic use but introduces large variation. Both designs are defensible if their purpose is explicit. A staged design can reconcile them: first compare standardized interactions, then examine independent use after prompt and AI-literacy instruction. Researchers should analyse prompts rather than merely append a few examples. Relevant features include purpose, context, constraints, use of criteria, requests for explanation, follow-up questions, and verification behaviour (Lo, 2023; Tlili et al., 2023).

4. IMPLICATIONS

Current evidence supports cautious, structured use rather than unrestricted generation or blanket prohibition. The clearest instructional role is formative: ChatGPT can provide an additional source of questions, examples, and feedback while the student remains responsible for claims and revisions. Teachers should begin with a learning outcome and decide which AI functions support it. If the goal is argument development, students can ask for counterarguments and then evaluate them against course readings. If the goal is revision, students can compare AI, peer, and teacher feedback and justify which suggestions they use. If the goal is sentence-level control, students can request explanations of errors before producing their own revision (Farrokhnia et.al., 2026).

Assignments should require visible decisions. A brief process appendix can include the tool and model, the purposes for which it was used, representative prompts, significant suggestions accepted or rejected, and verification steps. Version histories and annotated drafts can supplement this record. The requirement should remain proportionate; exhaustive transcript submission can invade privacy and create unmanageable marking.

Rubrics should reward judgement, source accuracy, and revision quality as well as fluency. Teachers need examples of permitted and prohibited use tied to each task. A general institutional policy cannot resolve every disciplinary question. Students should know whether they may brainstorm, receive feedback, edit language, or generate prose, and how to disclose each activity (Ding et al., 2025; Mahapatra, 2024).

Finally, course-level design should connect with program-level assurance. Students need opportunities to learn responsible AI-supported writing because such tools are likely to remain part of professional work. Programs also need meaningful points at which students demonstrate independent capability. A combination of portfolios, supervised writing, interactive oral assessment, and discussion of submitted work provides stronger evidence than an AI detector or a single take-home essay (Tlili et al., 2023; Farrokhnia et.al., 2026).

5. LIMITATIONS

This review is limited by the speed and heterogeneity of the field. Studies published under the name ChatGPT may concern different models and interfaces. Some recent articles were available online ahead of issue assignment, and indexing was inconsistent. The structured narrative method supports thematic interpretation but does not provide the exhaustive coverage or study-level risk-of-bias assessment of a full systematic review. The review includes English-language sources and therefore may underrepresent research and institutional practice published in other languages. It also concentrates on higher education; findings may not transfer to younger learners whose cognitive development, consent conditions, and classroom supervision differ (Tlili et al., 2023; Mahapatra, 2024).

These limitations reinforce the study's main conclusion. Claims about ChatGPT and writing should remain local to the documented model, task, learners, instructional design, and outcome. Accumulation of evidence will depend less on the number of new classroom trials than on whether studies describe these conditions consistently and distinguish better AI-assisted texts from better writers (Farrokhnia et.al., 2026).

6. CONCLUSION

ChatGPT can expand access to feedback and support several stages of university writing, but present evidence does not justify treating improved AI-assisted products as proof of durable writing development. The main methodological difficulty is that the intervention is unstable and distributed: outcomes emerge from a particular model, prompt, task, student, teacher, and policy context. Final texts hide that interaction.

Research should define the AI role, preserve process evidence, use realistic comparisons, and measure delayed unaided transfer. Teaching should make acceptable use explicit, require students to evaluate and disclose consequential AI contributions, and combine formative AI-supported work with trustworthy evidence of individual

capability. These practices do not resolve every ethical or technical question. They do allow educators to make more defensive claims about what ChatGPT contributed and what students learned.

Acknowledgements

The current study is part of the UGA V/4/2026 project titled Implementation of Artificial Intelligence Tools in Teaching English as a Foreign Language.

References

[1] Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

[2] Ding, L., Zou, D., & Kohnke, L. (2025). ChatGPT as an automated writing evaluation tool: How students perceive it and how it affects their writing. Education and Information Technologies, 30, 26777–26799. https://doi.org/10.1007/s10639-025-13775-3

[3] Farrokhnia, M., Latifi, S., Papadopoulos, P. M., Hogenkamp, L., Gijlers, H., Khosravi, H., & Noroozi, O. (2026). Generative AI offers more, but students revise less: Comparing the effects of teacher and AI feedback on student essay revisions. International Journal of Educational Technology in Higher Education, 23, Article 6. https://doi.org/10.1186/s41239-026-00579-9

[4] Jovic, M., Papakonstantinidis, S., & Kirkpatrick, R. (2025). From red ink to algorithms: Investigating the use of large language models in academic writing feedback. Language Testing in Asia, 15, Article 59. https://doi.org/10.1186/s40468-025-00389-2

[5] Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences, 13(4), Article 410. https://doi.org/10.3390/educsci13040410

[6] Lodge, J. M., Howard, S., Bearman, M., Dawson, P., & Associates. (2023). Assessment reform for the age of artificial intelligence. Tertiary Education Quality and Standards Agency. https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf

[7] Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. Smart Learning Environments, 11, Article 9. https://doi.org/10.1186/s40561-024-00295-9

[8] Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. Cogent Education, 10(2), Article 2236469. https://doi.org/10.1080/2331186X.2023.2236469

[9] Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

[10] Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10, Article 15. https://doi.org/10.1186/s40561-023-00237-x

Cost-Volume-Profit analysis: The Case of Lithuanian Manufacturing Company

Indrė KNYVIENĖ[a*]

[a]*Kauno kolegija Higher Education Insitution, Faculty of Bussiness, Pramonės pr. 20, Kaunas, Lithuania, indre.knyviene@go.kauko.lt*

ORCID: orcid.org/0000- 0003-0246-2199

Abstract

Companies constantly face challenges when trying to increase their sales and profits while managing the complex relationships between quantity, price, unit variable costs, and unit fixed costs. Price, costs, and quantities are interrelated and influence each other, depending on competition, customer response, and the need to better utilize existing production capacity. This article examines a solution to the complex problem of the interaction between these variables in order to increase overall operating profit. The main objective is to introduce and compare some different financial indicators, such as the fixed cost ratio, sales ratio, safety margin, net profit margin, and operating profit. The indicators can be used to reveal which indicators determine financial results. The study will be based on an analysis of existing scientific literature and analysing Lithuanian manufacturing company. This will seek to understand the reasons why the company is interested in increasing competition. The limitations of the article are related to the fact that only one Lithuanian manufacturing company is analysed. The study showed that company need to run an aggressive marketing campaign in the coming years to maintain its current growth. Will the marketing campaign planned for next year be successful when the company's employees provided data on variable, direct production, variable overhead costs, as well as fixed costs. By evaluating various costs, gross profit and the break-even point are estimated. This study stands out because it analyses price, costs, and quantities are interrelated and influence each other, depending on competition, customer response, and the need to make better use of existing production capacity. The results of the research will be useful for manufacturing company in order to maximize profits.

Keywords: Cost- volume-profit analysis, financial indicators, costs.

1. INTRODUCTION

Effective profit planning and control are essential for maintaining financial stability and achieving strategic goals. Cost, volume, and profit (CVP) analysis offers a systematic approach to understanding the relationship between costs, sales volume, and profit, enabling business leaders to make informed financial decisions. By understanding the relationship between costs, volume, and profit, managers can set product prices to maximize profitability. It is possible to propose the development of dynamic product pricing strategies so that prices are adjusted in response to fluctuations in demand.

Therefore, management control is crucial for establishing pricing policies, controlling costs, and analyzing operational profitability, thereby providing the necessary information in a timely manner and tailored to the needs of

* Corresponding author.

decision-makers. This information includes both financial and operational indicators and essential for all stages of the management cycle, promoting continuous monitoring of organizational performance. Since businesses exist in a dynamic and complex environment, the operation of decisions taken in science through tools such as CVP minimizes the business risk and maximizes economic efficiency (Almansour et al. (2024); Joshi, Saphota, (2026)).

Classifying costs into fixed and variable costs, depending on their relationship to scale of production. Is particularly important when analyzing profitability and selecting appropriate cost calculation methods (Setyawani et al. (2022). A thorough understanding ao costs is essential for effective financial management. Detailed information on various cost components, such as raw materials, direct labor costs, and overhead costs, allows for an accurate assessment of the company's financial condition and helps in making informed strategic decisions. Cost accounting systems designed for calculating, recording, and analyzing costs differ primarily in how they handle fixed production costs (Datar&Rajan (2021); Asani &Veliu (2025)). CVP analysis is useful tool for decision-making because it allows one to understand the interaction between fixed costs, variable costs, selling price, and production volume to determine profitability (Blocher et al. (2021)).

This analysis can identify loss-making products, support the definition of pricing policies and guide the composition of the production mix (Carvalho et al., 2020; Silva et al., 2021)

Using tools such as break-even analysis, contribution margin calculation, and scenario modeling, it provides support for decision-making aimed at optimizing pricing, cost management, and targeted strategic planning (Mishra, Mishra (2024); Martens, (2025)).

Research problem. CVP analysis is based on various assumptions (e.g., constant selling price, fixed costs, constant production volume, etc.). Research often examines how deviations from these assumptions affect the accuracy and validity of break-even analysis results. In companies with particularly complex cost structures, it can be difficult to accurately determine variable costs. Research may focus on methods to improve the allocation or estimation of variable costs in order to perform a more reliable CVP analysis. CVP analysis is straightforward for a single product but becomes complicated for companies selling multiple products with different prices, costs, and sales volumes. Research objectives include the development of effective models for CVP analysis involving multiple products.

Research aim. The aim of this paper is to analyse the relationship between costs, volume, and profit that managers can evaluate to make informed business decisions.

Research objectives:

1. To analyse scientific literature on the CVP analysis indicators.
2. To conduct a survey of manufacturing companies to calculate and analyse indicators.
3. To provide conclusions and proposed future calculation of indicators.

Research methods. The paper is based on an analysis and review of the available scientific literature and an analysis of the company'data, focusing on CVP indicators. It seeks to understand the CVP in a manufacturing company. The limitations of the study are related to the manufacturing company.

Structure of the paper. The paper is structured as follows: the first section presents an analysis of the scientific literature on the application of CVP indicators, the second section analyzes the data results, and the final section presents conclusions.

2. LITERATURE REVIEW

Cost-Volume-Profit analysis is an analysis of the relationship between the three key performance indicators of a company: costs, production (sales/operating) volume, and profit. The main objective of this analysis is to determine what the company's financial results will be if the projected production level changes (Sun, (2024); Nararya & Aisyah (2024). (This information is particularly important for company management, which, by examining the relationship between costs, production volume, and profit, can assess the potential profit or loss at various levels of activity and obtain data to make certain management decisions related to operations.

According to Bastomi et al. (2023) who emphasise "decisions regarding the main objective in the company are to maximise profit, so that to obtain maximum profit the company must be able to target how much profit will be achieved and the company must also think about how to achieve the expected profit as much as possible. Managers must be able to make a good plan in order to make use of resources that are arranged in an organised and controlled manner (Trivani et al. (2025). According to Carvalho et al. (2020) One of the analytical tools used is using cost, volume and profit analysis or commonly referred to as Cost Volume Profit (CVP) analysis".

According to Clark (2024) and Datar and Rjan (2021) "profit planning or profitability optimisation is the development of a work plan for company operations expressed in a calculation. Profit planning is useful to assist management in planning, budgeting and decision making both short and long term in order to achieve company goals".

There are several tools for calculating profit, namely (Ihenyen, Michael (2025; Joshi, Sapkota (2026) the Contribution margin, Break event point, Margin of safety, Operating laverage

In companies is very important identify which costs are fixed and which are variable which costs are fixed and which costs are variable and then determines contribution margin (Purwoko et al. (2022)).

When different quantities of goods are sold, only total revenue and total variable costs change. The difference between these two indicators is called contribution margin Paletti, (2025).

CM=TR – TVC,

where CM means Contribution margin, TR - Total revenues, TVC - Total variable costs

Contribution margin per unit is an excellent tool for calculating operating income and contribution margin (Paletti, 2025).

CMU=SP-VCU,

where CMU means Contribution margin per unit, SP – Selling price, VCU - Variable costs per unit

Contribution margin per unit reflects the relationship between the selling price and variable costs per unit.

Contribution margin per unit is another way to calculate contribution margin (Burritt et al. (2025).

CM=CMU*NUS,

where NUS mean Numbers of units sold.

Contribution income statement because it groups costs into variable costs and fixed costs to highlight contribution margin.

Income statement in which costs are classified into variable and fixed costs to determine the contribution margin. The increase in the contribution margin corresponds to the increase in operating profit (o decrease in operating loss) (Laba et.al, (2024)).

OI = CM- FC,

where OP means Operating income, FC – Fixed costs

In place of expressing the markup in euros per unit, companies state it as a percentage, which is referred to as the contribution margin percentage (or contribution margin ratio). (Burritt et al. (2025)).

CMP (CMR)=CM/R,

where CMP means Contribution margin percentage, CMR - Contribution margin ratio, R - Revenues

The contribution margin percentage is a practical way to calculate the contribution margin when revenues are uneven. By substituting values into the equation that defines the contribution margin percentage (Paletti, 2025).

CM=CMP*R

To determine the ratio of operating revenue to operating contribution margin percentage, it is important to note that it can be calculated in this way.

OI = CM – FC or

OI=CMP*R-FX

When there in only one product, we can divide both the numberator and the denominator of the contribution margin percentage equation by number of units sold and calculate the Contribution percentage as follows (Hidayat, 2024).

CMP=CM/R/US=CMU/SP

The essence of this information lies in the equations that express the relationship between cost, profit volume and influence the structure of contribution income statement. There are three related approaches (methods) for describing CVP methods:

1. The equation.
2. The contribution margin.
3. The graph.

Different methods are useful for different decisions.

The equation and contribution margin methods are most appropriate when management want to determine operating income at a few specific sales levels (for example, 5, 15, 25, and 40 units sold). The graph method helps

managers visualize the relationship between units sold and operating income over a wide range of quantities (Lamichhane et al. (2024)).

Equation Method. It can be expressed as an equation.

R-VC-FC=OI

How are revenues in each column calculated?

Revenues = Selling price (SP) * Quantity of units sold (Q).

How are variable costs in calculated?

Variable costs = Variable cost per unit (VCU) * Quantity of units sold (Q)

So,

[(Selling price) * (Quantity of units sold)- (Variable cost per unit) * (Quantity of units sold)]- Fixed costs = Operating income (Equation 1)

Taking in account the varying number of units sold, Equation 1 serves as the basis for calculating operating income.

Contribution Margin Method

Rearranging equation 1, (Hidayat, 2024).

[(Selling price- Variable cost a per unit) * (Quantity of units sold)]- Fixed costs = Operating income

(Contribution margin per unit*Quantity of units sold) - Fixed costs = Operating income * (Equation 2)

The graphical method helps managers understand dependency between total revenues and total costs vary linearly with the number of units sold (which certainly does not happen in real life), this relationship is represented by a straight line on the graph (Moroz, 2026).

The breakeven point (BEP) according to Sprague, Lee (2025) **"**is that quantity of output sold at which total revenues equal total costs—that is, the quantity of output sold that results in $0 of operating income".

Success rather than trying out different quantities and checking when operating income equals 0 Eur. Recall the equation method (Equation 1):

We need to find a solution instead of trying out different values and checking when operating income equals 0 euros (Moroz, 2026).

[(Selling price * Quantity of units sold)- (Variable cost per unit * Quantity of units sold)] - Fixed costs = Operating income

At the breakeven point, operating income is by definition $0, and so,

Contribution margin per unit * Breakeven quantity of units = Fixed costs (Equation 3)

Sensitivity analysis – a indicator showing by how much sales volume can decrease while the company still remains profitable. This type of analysis helps answer questions such us "What will operating income be if the quantity of units sold decreases by 5% from the original prediction?" or "What will operating income be if variable cost per unit increases by 10%?" (Guo, 2022).

An important aspect of sensitivity analysis is margin of safety:

Margin of safety = Budgeted (or actual) revenues- Breakeven revenues

Margin of safety (in units) = Budgeted (or actual) sales quantity- Breakeven quantity

The margin of safety is often used as a simple risk assessment tool. If a company's safety margin is relatively large compared to sales revenue, it is likely that lower-planned sales revenue or higher-than-planned expenses will have much smaller impact on net income. Sales may decline as a result of factors such as a poor marketing campaign or better product offered by competitor (Guo, 2022).

Operating leverage measures the risk-reward ratio, taking in account different cost structures. Operating leverage describes the impact of fixed costs on changes in operating income, considering the number of units sold and the contribution margin. In companies where fixed costs structure, even a small increase in sales volume results in a significant increase in operating income. Meanwhile, a slight decrease in sales volume lead to significant decrease in operating income and as increase in risk at any sales level (Hidayat, 2024).

Degree of operating leverage=Contribution margin/Operating income

3. RESEARCH METHODOLOGY AND RESULTS

The financial data was collected from a manufacturing company that manufactures linen textiles products for children such as linen toys. The company engages in retail and online sales of lines products for children. The selection isn't large at the moment, but it's popular because of its organic products. It is a small family company with seven employees. Information from 2025 income statement was used to perform the data analysis.

Limitations of the study. Since only one manufacturing company and one year were selected, it was not possible to compare how the indicators change with different production volumes.

Research results

Manufacturing company sells the toys at uniform price €48 per unit. The company's fixed annual costs amount to €36,000. The calculated break-even point for 2025 was €90,000.

Calculation the contribution margin ratio:

€90,000=€360,000/x

x=0,4 (CM)

Calculation the contribution margin per unit:

0,4 = (€48-x)/€48

x = €19,2 (VC)

€48+= €19,2 (CMU)

The company decided to count operating income for 2025, assuming that sales revenue for that year was €120,000.

€120,000 = (€120,000/€48) * €19,2+€36,000+x

x=€36,000 (OI)

In 2026, the company plans to manufacture its products using higher quality fabrics, which would increase the variable cost per toy by €1,90.

So, there is a question, what should the selling price of the toy be that contribution margin remains unchanged?

0,4 = (x-€19,2+1,9)/x

x=€35,17 (SP)

The selling price has changed from €48 to €35,17.

What would be the contribution margin for a single unit at this price?

€35,17-€21,10=€14,07

Due to an increase in the variable costs per unit, company management plans to rise the selling price of each toy to €52 per unit.

Company managers calculated the break-even point, assuming the company plans to spend an additional €1,950 on advertising higher-quality product.

BEP=(€36,000+€1,950)/0,59=€64,322

CM=(€52-€21,10)/ €52=0,59

Last question was to decide whether it is appropriate to approve this plan, given that 6 500 new toys are expected be sold in 2026.

This can be verified by preparing contribution margin statement for 2026.

Table 1. Contribution margin statement (Source: research results)

Sales revenue	€338,000
Variable costs	€137,150
Contribution margin	€200,850
Fixed costs	€37,950
Operating income	**€162,900**

Calculation have confirmed that, when both fixed and variable costs change, CVP analysis enables the identification of changes and the making of management decisions. In the case under consideration, it can be concluded that changes in prices have a significant impact on operating income, and it is important for company to make right decisions.

4. CONCLUSIONS

By applying CVP analysis, we can conduct a comprehensive and in-depth analysis of a company's current production and operational status, and plan management decisions. CVP model analysis broadens the scope of CVP analysis, taking into account not only the impact of variable costs, fixed costs, production, sales volume, and unit price on costs and profit, but also by expanding certain decisions related to corporate management issues. Therefore, when CVP is used for business decision-making and corporate financial management, calculating various indicators can

increase the relevance and accuracy of the analysis, which is more practical. The article presents examples from two companies on how to plan sales and prices and select certain business alternatives when organizing sales.

in which they are mentioned, and separated from the main text by a short line extending at the foot of the column. The 'footnote' style is available in this template for the text of the footnote.

.

References

[1] Almansour A.Y., Almansour B.Y., (2024). Optimizing Break-Even Point: A New Methodology for Identifying Critical Components and Managing Financial Stability. Montenegrin Journal of Economics 20(4): 51-63. https://doi.org/10.14254/1800-5845/2024.20-4.5

[2] Asani, B., & Veliu, N. A. (2025). The Impact of Accounting Information on Managerial Decision-Making. International Journal of Business and Economics, 10(1), 151–156. https://doi.org/10.2139/ssrn.5377604

[3] Bastomi M., Sudarmiatin, Hermawan A. (2023). Cost Volume Profit (CVP) Analysis As A Profit Planning Tool At Lay Cang MSMEs. Asian Journal of Management Entrepreneurship and Social Science ISSN: 2808 7399 https://ajmesc.com/index.php/ajmesc Volume. https://doi.org/10.63922/ajmesc.v6i02.1549

[4] Blocher, J., Juras, E. and Smith, S. (2021), 'Cost management: A strategic emphasis,' 9th ed., McGraw-Hill Education.

[5] Burritt M., Souza S.V., Peterson Ch. (2025). When Will Controlled Environment Agriculture in Its Vertical Form Fulfill Its Potential? *Sustainability* 2025, *17*(7), 2957; https://doi.org/10.3390/su17072957

[6] Carvalho, F. M., Fernandes, C. H. and Bandeira-de-Mello, R. (2020), 'Gerenciamento de custos em ambientes de incerteza: uma abordagem baseada na análise CVR,' Revista de Administração da UFSM, 13(4), 885–902. https://doi.org/10.5902/1983465934231

[7] Clark J.M. (2024). Overhead Costs in Modern Industry, I-III. Classics in Institutional Economics, Part II (6): 120-203. Retrieved from https://ideas.repec.org/a/ucp/jpolec/v31y1913p606.html

[8] Datar, S. and Rajan, M. (2021), 'Horngren's cost accounting: A managerial emphasis,' (17 th ed.). Pearson.

[9] Edström, K., Malmqvist, J., & Roslöf, J. (2020). Scholarly development of engineering education – the CDIO approach. European Journal of Engineering Education, 45(1), 3–12. https://doi.org/10.1080/03043797.2020.1704361

[10] Guo Y. (2022). Enterprise Management Decision and Financial Management Based on Dynamic Cost Volume Profit Model. Journal of Function Spaces, Volume 2023. https://doi.org?10.1155/2022/9016060

[11] Hidayat R. (2024). Analisis Cost Volume Profit (CVP) sebagai Metode Manajemen Laba pada UMKM Madu Asy-Syifa. MAMEN (Jurnal Manajemen) journal.literasisains.id/index.php/MAMEN e-ISSN 2809-8099 | p-ISSN 2810-0484 Vol. 3 No. 2 (April 2024) 134-145 DOI: 10.55123/mamen.v3i2.3549

[12] Ihenyen, J. C., & Michael, O. A. (2025). Cost Volume Profit (CVP) Analysis: Application: Application and Limitation in Business Planning. Journal of Public Administration and Social Welfare Research, 10(3), 103–109. https://doi.org/10.56201/jpaswr.v10.no3.2025.pg103.110

[13] Joshi T. R., Sapkota S.P. (2026). The Role of CVP Analysis in Strategic Decision-Making for Small and Medium Enterprises. AMC Juornal (Dhandhagi). Vol. 7; Pp. 127-138. https://doi.org/10.3126/amcjd.v7i1.91976

[14] Laba M. M., Hidyat R., Asy-Syifa M. (2024). Analisis Cost Volume Profit (CVP). MAMEN (Jurnal Manajemen) journal.literasisains.id/index.php/MAMEN e-ISSN 2809-8099 | p-ISSN 2810-0484 Vol. 3 No. 2 P.p. 134-145 DOI: 10.55123/mamen.v3i2.3549

[15] Lamichhane B.D., Bahadur P., Lophan L. (2024). Cost-Volume Analysis and its Impact on Profitability: An Evidence from Nepal. International Journal of Religion 2024 Volume: 5| Number 12 | pp. 2007 – 2018 ISSN: 2633-352X (Print) | ISSN: 2633-3538 (Online) ijor.co.uk DOI: https://doi.org/10.61707/xtpk0c12

[16] Martens W. (2025). Breakeven Referencing, Scenario Modeling, and Margin Strategies: An Analysis of CVP Tool Adoption in Managerial Decisions. Asian Journal of Technology & Management Research (AJTMR) ISSN: 2249 –0892 Vol. 15. P.p. 1-9

[17] Mishra A. Mishra M. (2024). A case study on Cost-Volume-Profit Analysis. SAMAGRA GYAN JOURNAL. E-ISSN 2582–726X Vol. 07, Issue 7. DOI:10.61707/xtpk0c12

[18] Moroz K. (2026). Break-even pont modelling as a tool for strategic planning and business risk assessment. European Journal of Economic and Financial Risk. Volume 10. Pp. 19-28. DOI: 10.46827/ejefr.v10i2.2144

[19] Nararya, Naser, A., & Aisyah, S. (2024). Implementasi Metode Activity Based Costing (Abc) Dan Economic Order Quantity (Eoq) Dalam Pengendalian Bahan Baku Di Konfeksi Erstore Medan. Jurnal Manajemen Terapan Dan Keuangan, 13(02), 350–362. https://doi.org/10.22437/jmk.v13i02.33002

[20] Palleti P. (2025). Conversational CPQ Copilots: Margin-Aware Generative Configuration on Salesforce. Journal of artificial intelligence, machine learning and data science 2025, 3(2): 1-3. **DOI:** https://doi.org/10.51219/JAIMLD/pavan-palleti/599

[21] Purwoko, B., Hadi, M. A., Gamal, A., & Prihartanti, W. (2022). The Effect Production Costs and Sales Volumes on Profits of Registered Food and Beverage Sub-Sector Manufacturing Companies on the Indonesia Stock Exchange for the 2018-2020 Period. Enrichment: Journal of Management, 12(3), 2459–2468.

[22] Setyawati D. M., Usman S., Fidyah F. (2023). Cost-volume-profit calculation analysis as a profit planning approach in "Markisa Manis Mampang" micro business Depok city, west java province. http://ejournal.seaninstitute.or.id/index.php/Ekonomi Jurnal Ekonomi, Volume 12, No 03, 2023 ISSN: 2301-6280 (print) ISSN: 2721-9879 (online). DOI**:** https://doi.org/10.54209/ekonomi.v15i01 Volume 12

[23] Sprague S. M., Lee W.J., (2025). Breakeven Costs: An Industry Review Since 2009. SPE Annual Technical Conference and Exhibition: D011S007R003. https://doi.org/10.2118/228149-MS

[24] Sun L. (2024). Management research of big data technology in financial decision making of enterprise cloud accounting. Journal of Information Knowledge Management 23(01). https://doi.org/10.1142/S0219649223500673

[25] Triyani, A., Astuti, E. B., Setyahuni, S. W., Oktoriza, L. A., & Maghfiroh, A. R. (2025). Optimalisasi Strategi Pemasaran BLUE untuk Meningkatkan Keberlanjutan Produk Perikanan UMKM Mami Sera. Jurnal Abdimas (Journal of Community Service): Sasambo, 7(4), 683

[26] Van der Geer, J., Hanraads, J. A. J., & Lupton, R. A. (2000). The art of writing a scientific article. *Journal of Scientific Communications*, *163*, 51 - 59.
[27] Strunk, W., Jr., & White, E. B. (1979). *The elements of style* (3rd ed.). New York: Macmillan.
[28] Mettam, G. R., & Adams, L. B. (1994). How to prepare an electronic version of your article. In B. S. Jones, & R. Z. Smith (Eds.), *Introduction to the electronic age* (pp. 281-304). New York: E-Publishing Inc.

The Relationship Between Self-efficacy and Emotional Intelligence in Pre-service Teachers

Monika BOBAKOVA[a*], Nikoleta IZDENCZYOVA [b]

[a] *University of Presov, Faculty of Education, Presov, Slovakia, monika.bobakova@unipo.sk*
[b] *University of Presov, Faculty of Education, Presov, Slovakia, nikoleta.izdenczyova@unipo.sk*

Abstract

This study examined the relationship between trait emotional intelligence and teacher self-efficacy among pre-service teachers and assessed the predictive contribution of emotional intelligence to self-efficacy. The sample consisted of 181 students of the Faculty of Education at the University of Presov (178 women and 3 men) aged 18–25 years. Participants completed the Slovak adaptations of the Ohio State Teacher Efficacy Scale and the Trait Emotional Intelligence Questionnaire–Short Form. Participants reported above-average teacher self-efficacy and slightly above-average trait emotional intelligence. Teacher self-efficacy was positively associated with overall emotional intelligence and its four dimensions: well-being, self-control, emotionality, and sociability. Although the correlations were significant, they were relatively weak. Overall trait emotional intelligence explained 15% of the variance in teacher self-efficacy. The findings highlight the importance of developing emotional competencies in pre-service teacher education while considering additional psychological and contextual predictors of teacher self-efficacy.

Keywords: teacher self-efficacy, emotional intelligence, teachers

1. INTRODUCTION

The teaching profession is considered one of the most stressful occupations, with the high prevalence of stress among teachers highlighting the growing demands associated with their work (Travers, 2017). Prolonged exposure to occupational stress may adversely affect not only teachers' psychological well-being but also the quality of their teaching. In this context, it is important to focus on emotional intelligence and self-efficacy as personal resources that may help teachers cope with stressful situations. A deeper understanding of the role these factors play in the experience and regulation of teacher stress may provide a foundation for developing preventive intervention programmes aimed at enhancing teachers' emotional competencies and strengthening their self-efficacy. Interventions designed in this way could improve their ability to cope more effectively with stress related to the demands of academic and professional environments (Wapaño, 2021).

* Corresponding author

1.1. Emotional intelligence

Emotional intelligence is generally defined as the ability to perceive, identify, understand, and regulate one's own emotions as well as those of others. It also encompasses the ability to use emotional information in cognitive processing, logical reasoning, decision-making, and problem-solving. It includes the verbal and non-verbal expression and appraisal of emotions, as well as their use in motivating and guiding behaviour (Mayer et al., 1990; Salovey & Mayer, 1997; Mayer, Caruso, & Salovey, 1999; Mayer, Salovey, & Caruso, 2004).

The ability to work with emotions is among the fundamental prerequisites for successful performance in the teaching profession. Teaching is an emotionally demanding activity during which teachers continuously perceive and regulate their own emotions, respond to students' emotional expressions, and strive to create a positive learning environment. The appropriate management of the emotional demands of teaching therefore contributes significantly to teachers' pedagogical effectiveness (Gates, 2000; Wu et al., 2019). Emotional intelligence can thus be regarded as an important factor in the quality of teachers' work.

Teachers with high levels of social and emotional competence are better able to use emotional expression and verbal support effectively to enhance students' enthusiasm, motivation, and enjoyment of learning. They are also more successful in guiding student behaviour and managing the classroom. Furthermore, they have a better understanding of the dynamics of conflict situations, enabling them to respond appropriately to emerging problems and reduce disruptive behaviour (Jennings & Greenberg, 2009).

The importance of emotional intelligence is also supported by empirical findings indicating that teachers with higher levels of emotional intelligence generally achieve better teaching outcomes (Yoke & Panatik, 2015). Emotionally competent teachers demonstrate greater overall teaching effectiveness, are more sensitive in identifying students' specific needs, and create stimulating, supportive, and caring learning environments (Gay, 2010; Wapaño, 2021).

Emotional intelligence is important not only for the quality of teaching and classroom relationships but also for teachers' psychological resilience. Individuals with higher levels of emotional intelligence employ more effective stress-regulation techniques and more adaptive coping strategies, which may contribute to reducing the risk of occupational burnout (Mayer et al., 1990; Montes-Berges & Augusto-Landa, 2014). Similarly, research focusing on teachers identifies emotional intelligence as a factor that facilitates the effective management of stress associated with the teaching profession (Chan, 2008).

Emotional intelligence therefore represents a comprehensive professional resource that helps teachers manage the emotional demands of their work, build high-quality relationships with students, effectively manage classroom dynamics, and create favourable conditions for learning. At the same time, it may serve as a protective factor in managing occupational demands and preventing burnout.

1.2. Teacher self-efficacy

Self-efficacy represents an important psychological resource that influences an individual's ability to cope with the demands of specific situations. It is defined as a person's belief in their own ability to successfully perform tasks and achieve established goals (Bandura, 1997; Shubayr & Dailah, 2025). Individuals with high levels of self-efficacy tend to perceive demanding tasks as challenges to be mastered rather than as threats to be avoided. Even when faced with stressful or threatening situations, they approach them with the belief that they can cope successfully and maintain control. Such a mindset promotes perseverance and personal achievement, helps alleviate stress, and reduces vulnerability to depression (Wapaño, 2021).

In the context of the teaching profession, self-efficacy refers to the extent to which teachers believe that their pedagogical practice can influence students' behaviour and learning outcomes (Tschannen-Moran, Hoy, & Hoy, 1998). Teacher self-efficacy is associated with the sense of professional competence experienced by practising or prospective teachers (Sahin, 2017). Bandura (1994) stated that teachers with high levels of perceived self-efficacy are better able to motivate and encourage students and support their cognitive development. He therefore regarded teachers' belief in their ability to perform their pedagogical duties successfully as one of the fundamental prerequisites for creating an effective learning environment.

The importance of teacher self-efficacy is evident across several areas of pedagogical practice. Research has demonstrated its association with students' educational outcomes, the classroom climate, and the overall functioning and organisational health of schools (Wapaño, 2021). Teachers with higher levels of self-efficacy also exhibit stronger professional engagement and commitment to the teaching profession and are less likely to consider leaving it (Chesnut & Burley, 2015).

Higher levels of self-efficacy are also reflected in how teachers prepare and deliver instruction. Teachers with a strong belief in their own efficacy devote greater attention to planning and organising the instructional process and to classroom management, while also displaying greater enthusiasm in their work. They demonstrate greater persistence when confronted with obstacles and greater resilience in coping with setbacks (Tschannen-Moran & Hoy, 2001). They are also more open to adopting new teaching methods and more willing to apply them in their own pedagogical practice (Ross, 1994).

Teacher self-efficacy is also reflected in teachers' approaches to students. Teachers with higher levels of self-efficacy adopt a more humane approach, present subject matter more creatively, and use questions more effectively to encourage students' active participation and thinking (Emmer & Hickman, 1991). Such teachers also tend to be less critical of students who make mistakes (Tschannen-Moran & Hoy, 2001), provide greater support to students with poorer educational outcomes, and persist in their pedagogical efforts despite students' occasional lack of success (Ross, 1994).

Self-efficacy is also important in coping with the psychological demands associated with the teaching profession. It may serve as a protective factor against stress (Chan, 2002) and has also been examined as a predictor of teacher stress (Wapaño, 2021). Conversely, low levels of teacher self-efficacy may be associated with an increased risk of burnout syndrome (Schwarzer & Hallum, 2008).

Based on the above, teacher self-efficacy may be regarded as an important personal and professional resource. It influences the quality of pedagogical practice, teachers' relationships with students, their willingness to innovate their teaching and persevere in the face of obstacles, as well as their ability to manage occupational demands and remain in the teaching profession over the long term.

1.3. Teacher self-efficacy and emotional intelligence

Several empirical studies have identified a positive relationship between emotional intelligence and teachers' perceived self-efficacy (Sarkhosh & Rezaee, 2014; Wu et al., 2019; Wapaño, 2021; Sun & Lyu, 2022). Although this relationship is statistically significant, its strength generally ranges from weak to moderate. Sarkhosh & Rezaee (2014) further suggest that emotional intelligence may positively influence teachers' job performance, thereby contributing to stronger beliefs in their own self-efficacy.

Drawing on self-efficacy theory, Sun & Lyu (2022) emphasise that individuals assess their ability to manage specific tasks partly on the basis of their current emotional state. The ability to regulate emotions appropriately promotes positive psychological experiences and more adaptive forms of behaviour. Individuals are therefore more likely to approach challenging situations constructively, which may strengthen their self-efficacy in areas such as learning, teaching, professional practice, and social interactions in the workplace.

According to Sun & Lyu (2022), emotional intelligence may positively predict individual self-efficacy because individuals with higher emotional intelligence are better able to analyse situations requiring decision-making, strengthen their self-confidence, and influence others appropriately through suitable emotional expressions. These abilities enable them to solve problems more effectively and make appropriate decisions, thereby reinforcing their belief in their ability to manage various demands and situations successfully.

Emotional intelligence was also identified as a predictor of teacher self-efficacy by Wang & Wang (2022). Teachers with higher levels of emotional awareness and more developed emotional competencies tend to demonstrate greater confidence in their ability to perform teaching activities. The ability to accurately identify students' emotions and foster positive emotional experiences may help teachers manage the classroom more effectively and motivate students who experience boredom or anxiety during learning (Moafian & Ghanizadeh, 2009).

Predictive significance can be attributed not only to overall emotional intelligence but also to its individual components (Moafian & Ghanizadeh, 2009; Sarkhosh & Rezaee, 2014). Chan (2004), for example, found that the positive regulation of emotions significantly predicted teachers' general self-efficacy. Teachers' ability to recognise and appropriately regulate their emotions and use them when addressing work-related situations may therefore enhance their confidence in managing pedagogical tasks. Similarly, other studies have associated better regulation of one's own emotions, as a dimension of emotional intelligence, with stronger teacher self-efficacy beliefs (Koçoğlu, 2011) and higher-quality relationships with students (Moafian & Ghanizadeh, 2009).

Higher self-efficacy associated with well-developed emotional competencies may also have a protective role against the adverse consequences of the teaching profession. Wang & Wang (2022) report that teachers' positive beliefs about their own abilities help them resist emotional exhaustion more effectively. Similar findings were reported by Geraci et al. (2023), who examined teachers' emotional intelligence, self-efficacy, work engagement,

and burnout during the lockdown associated with the COVID-19 pandemic. Their findings showed that levels of self-efficacy, burnout, and work engagement differed according to the level of emotional intelligence. In this context, emotional intelligence served as a protective factor against burnout while simultaneously promoting higher levels of self-efficacy and work engagement.

The existing evidence therefore suggests that emotional intelligence and teacher self-efficacy are interrelated constructs. Teachers who are better able to perceive, understand, and regulate their own emotions, while responding sensitively to students' emotional experiences, generally have greater confidence in their professional abilities, manage the demands of teaching more effectively, and develop higher-quality relationships with students. In this way, higher emotional intelligence may not only strengthen teachers' self-efficacy but also promote their work engagement and reduce the risk of emotional exhaustion and burnout.

2. METHODOLOGY

The main objective of the study was to investigate the relationship between trait emotional intelligence and pre-service teachers' self-efficacy beliefs. The study also aimed to determine the extent to which teachers' emotional intelligence contributes to predicting their sense of self-efficacy.

The following two research hypotheses were formulated and examined:

H1: There is a significant positive relationship between emotional intelligence and self-efficacy among pre-service teachers.

H2: There is a significant positive relationship between the dimensions of emotional intelligence and self-efficacy among pre-service teachers.

2.1. Participants

The sample consisted of 181 pre-service teachers studying at the Faculty of Education of the University of Presov. A total of 178 women and 3 men between the ages of 18 and 25 participated in the study (M = 20.26, SD = 1.02).

2.2. Instruments

The Slovak adaptation of the Ohio State Teacher Efficacy Scale (OSTES; Tschannen-Moran & Hoy, 2001), an instrument designed to measure teacher self-efficacy, was translated and developed by Gavora (2011). The scale is unidimensional and consists of 24 items. The items focus on what a teacher or pre-service teacher is capable of doing in a given situation. Respondents indicate their answers on a nine-point scale ranging from 1 ("nothing at all") to 9 ("very much").

The Trait Emotional Intelligence Questionnaire–Short Form (TEIQue-SF) is designed to measure trait emotional intelligence. It is based on Petrides's model of trait emotional intelligence, which conceptualises emotional intelligence as a set of personality-related or behavioural dispositions associated with emotional functioning. The questionnaire assesses global trait emotional intelligence and provides approximate scores for its four factors: well-being, emotionality, self-control, and sociability. The short form was derived from the original full version of the TEIQue (Petrides & Furnham, 2001, 2006). The Slovak adaptation of the questionnaire was developed by Kaliská, Heinzová, & Nábělková (2019). The questionnaire comprises 30 items rated on a seven-point scale, with 1 ("complete disagreement") and 7 ("complete agreement").

3. RESULTS AND DISCUSSION

In a sample of 181 pre-service teachers, we examined the total teacher self-efficacy score, as well as the global score and individual dimension scores for trait emotional intelligence (EI). The relationships between these variables were subsequently analysed.

Table 1 presents the descriptive statistics for the variables examined. The pre-service teachers in our study achieved above-average scores on the teacher self-efficacy scale, indicating that they had sufficient confidence in their professional capabilities. They perceived their abilities to manage the classroom, use effective teaching strategies, and engage students as very good. This may be considered a favourable prerequisite for the effective completion of teaching practice. However, as Gavora (2011) points out, these scores reflect teachers' self-reported rather than actual competence, which becomes evident only during teaching practice.

Table 1. Descriptive analysis of the examined variables (N = 181)

	Mean	Std. Deviation	Min.	Max.	α
Self-efficacy	7.54	0.89	3.04	9.00	0,94
EI	4.78	0.70	2.50	6.40	0.86
EI-well-being	5.28	1.00	1.83	7.00	0.80
EI-self-control	4.22	0.98	1.50	6.83	0.76
EI-emotionality	5.11	0.87	2.50	7.00	0.74
EI-sociability	4.50	0.88	2.17	7.00	0.70

Note: EI= emotional intelligence, Min.= minimum, Max.= maximum, α= Cronbach's alpha

Overall, the students achieved slightly above-average emotional intelligence scores. No substantial differences were found among the mean scores for the individual dimensions of emotional intelligence. The students obtained lower mean scores on the self-control (M=4.22) and sociability (M=4.50) subscales, while higher scores were recorded for emotionality (M=5.11) and well-being (M=5.28). Thus, the students reported relatively above-average perceived abilities to regulate emotions, impulses, and stress, as well as relatively high levels of assertiveness and social awareness. They also rated their abilities to perceive and express emotions effectively, along with their own well-being and sense of happiness, as above average.

Table 2. Intercorrelation matrix of teacher's self-efficacy, trait emotional intelligence and its factors (N = 181)

	1	**2**	**3**	**4**	**5**
1. Self-efficacy	-				
2. EI	.392***	-			
3. EI-well-being	.296***	.746***	-		
4. EI-self-control	.267***	.806***	.508***	-	
5. EI-emotionality	.334***	.778***	.427***	.510***	-
6. EI-sociability	.263***	.609***	.255***	.385***	.342***

Note: EI= emotional intelligence, ***p < 0.001 (2-tailed)

Table 2 summarises the relationships between emotional intelligence, its dimensions, and self-efficacy, as assessed using Pearson's correlation coefficient. Emotional intelligence exhibited strong positive correlations with all four of its dimensions. The correlation between the total emotional intelligence and self-efficacy scores was significant and positive, although relatively weak. Other variables are therefore also likely to play a role in the relationship between these constructs. Similar findings have been reported in several studies (e.g., Sarkhosh & Rezaee, 2014; Wu et al., 2019; Wapaño, 2021; Sun & Lyu, 2022). Both hypotheses were therefore supported.

Table 3. Regression analysis for emotional intelligence as the predictor of teachers' self-efficacy

	B	SE	β	t	p
Self-efficacy (R^2=0.15, F (1,181) = 32.50, p < 0.001)					
EI	0.50	0.09	0.39	5.701	0.000
(Constant)	5.14	0.43		12.058	0.000

Note: EI= emotional intelligence

The relationship between emotional intelligence and self-efficacy was examined in greater detail using regression analysis, which assessed the relative contribution of global trait emotional intelligence to teacher self-efficacy. The results indicated that global trait emotional intelligence accounted for only 15% (R^2 = .15) of the variance in teacher self-efficacy (see Table 3). Thus, higher emotional intelligence significantly contributes to a stronger sense of self-efficacy among pre-service teachers, a finding also supported by previous studies (Moafian & Ghanizadeh, 2009; Sarkhosh & Rezaee, 2014; Sun & Lyu, 2022; Wapaño, 2021). However, other variables also appear to play a role in this relationship and warrant further investigation. Studies examining mediators of the relationship between emotional intelligence and self-efficacy have identified several potentially relevant factors, including stress-coping

strategies and teaching performance (Wu et al., 2019; Wapaño, 2021; Sun & Lyu, 2022), as well as aspects of social intelligence (Moafian & Ghanizadeh, 2009; Kostić-Bobanović, 2020).

Previous research has also shown that despite the predictive power of emotional intelligence as a global construct, not all dimensions of trait emotional intelligence contribute significantly to the prediction of self-efficacy (e.g., Moafian & Ghanizadeh, 2009; Sahin, 2017).

Future research should therefore examine the relationships between self-efficacy and emotional intelligence at the level of their individual dimensions. It would also be appropriate to investigate the variables mediating the relationship between emotional intelligence and self-efficacy, as well as predictors of teacher self-efficacy other than emotional intelligence.

4. CONCLUSION

The findings obtained from a sample of 181 pre-service teachers indicated an above-average level of teacher self-efficacy and a slightly above-average level of trait emotional intelligence. Statistically significant positive relationships were found between emotional intelligence, its individual dimensions, and self-efficacy. Global emotional intelligence accounted for 15% of the variance in teacher self-efficacy and may therefore be regarded as a significant, although not the sole, predictor of pre-service teachers' confidence in their professional capabilities.

These findings support the importance of developing emotional competencies in initial teacher education. Further research should examine the interrelationships among the dimensions of these variables and the role of potential mediators—particularly coping strategies, the quality of teaching performance, and social intelligence—in the relationship between emotional intelligence and self-efficacy.

Acknowledgements

The current study is part of the project VEGA 1/0421/26, titled Self-efficacy of Pre-service Teachers and its Psychological Predictors.

References

[1] Bandura A. (1997). Self-efficacy: the exercise of control. New York: W.H. Freeman & Co.
[2] Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), Encyclopedia of human behavior (Vol. 4, pp. 71-81). New York: Academic Press.
[3] Emmer, E. T. & Hickman, J. (1991). Teacher efficacy in classroom management and discipline. *Educational and Psychological Measurement,* 51, 755-765.
[4] Gates, G. S. (2000). The socialization of feelings in undergraduate education: A study of emotional management. *College Student Journal*, *34*, 485–504.
[5] Gavora, P. (2011). Zisťovanie profesijnej zdatnosti učiteľa pomocou dotazníka OSTES. *Pedagogika. sk*, 2, 88-107.
[6] Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice*. New York: Teachers College Press.
[7] Geraci, A., Di Domenico, L., Inguglia, C., & D'Amico, A. (2023). Teachers' Emotional Intelligence, Burnout, Work Engagement, and Self-Efficacy during COVID-19 Lockdown. *Behavioral Sciences*, *13*, 4, 296.
[8] Chan, D. W. (2002). Stress, self-efficacy, social support, and psychological distress among prospective Chinese teachers in Hong Kong. *Educational Psychology, 22*, 5, 557–569.
[9] Chan, D. W. (2004). Perceived emotional intelligence and self-efficacy among Chinese secondary school teachers in Hong Kong. *Personality and individual differences*, *36*, 8, 1781-1795.
[10] Chan, D. W. (2008). Emotional intelligence, self-efficacy and coping among Chinese prospective and in-service teachers in Hongkong. *Educational Psychology, 28,* 397-498.
[11] Chesnut, S. R., & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. *Educational Research Review*, *15*, 1–16.
[12] Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, *79*, 491–525.
[13] Kaliská, L., Heinzová, Z., & Nábělková, E. (2019). *Dotazníky črtovej emocionálnej inteligencie TEIQue-SF/TEIQue-CSF: manuál k skráteným formám (2. časť).* Banská Bystrica: Belianum.
[14] Koçoğlu, Z. (2011). Emotional intelligence and teacher efficacy: A study of Turkish EFL pre-service teachers. *Teacher Development*, *15*, 4, 471-484.
[15] Kostić-Bobanović, M. (2020). Perceived emotional intelligence and self-efficacy among novice and experienced foreign language teachers. *Economic research-Ekonomska istraživanja*, *33*, 1, 1200-1213.
[16] Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence,27*, 4, 267-298.

[17] Mayer, J.D., DiPaolo, M., Salovey, P. (1990). Perceiving affective content in ambigu-ous visual stimuli: a component of emotional intelligence. Journal of Personality Assessment, 54, 3–4, 772–81.
[18] Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Target Articles: "Emotional intelligence: Theory, findings, and implications." *Psychological Inquiry,15*, 3, 197-215.
[19] Moafian, F., & Ghanizadeh, A. (2009). The relationship between Iranian EFL teachers' emotional intelligence and their self-efficacy in Language Institutes. *System*, *37*, 4, 708-718.
[20] Montes-Berges, B., Augusto-Landa, J.M. (2014). Emotional intelligence and affective intensity as life satisfaction and psychological well-being predictors on nurs-ing professionals. Journal of Profesional Nursing, 30, 1, 80–8.
[21] Petrides, K. V. & Furnham, A. (2001). Trait emotional intelligence: Psychometric investigation with reference to established trait taxonomies. *European Journal of Personality, 15,* 425-448.
[22] Petrides, K. V. & Furnham, A. (2006). The role of trait emotional intelligence in a gender-specific model of organizational variables. *Journal of Applied Social Psychology*, *36,* 552-569.
[23] Ross, J.A. (1994). *Beliefs that Make a Difference. The Origins and Impacts of Teacher Efficacy*. Proceedings of the Annual Meeting of the Canadian Society for the Study of Education, Calgary, June 1994.
[24] Sahin, H. (2017). Emotional intelligence and self-esteem as predictors of teacher self-efficacy. *Educational Research and Reviews*, 12, 22, 1107-1111.
[25] Salovey, P., Mayer, J.D. (1997). Emotional development and emotional intelligence. *Educational Psychological Review*, 9, 3, 251–69.
[26] Sarkhosh, M., & Rezaee, A. A. (2014). How does university teachers' emotional intelligence relate to their self-efficacy beliefs? *Porta Linguarum*, *21*, 85–100.
[27] Shubayr, N., & Dailah, H. (2025). Assessment of emotional intelligence, self-efficacy, and perceived stress among nursing students in clinical practice: a cross-sectional study. *BMC nursing*, *24*, 1, 505.
[28] Schwarzer, R., & Hallum, S. (2008). Perceived Teacher Self-Efficacy as a Predictor of Job Stress and Burnout: Mediation Analyses. *Applied Psychology: An International Review*, *57*.
[29] Sun, G., & Lyu, B. (2022). Relationship between emotional intelligence and self-efficacy among college students: the mediating role of coping styles. *Discover Psychology*, *2*, 1, 42
[30] Travers, C. (2017). Current knowledge on the nature, prevalence, sources and potential impact of teacher stress. *Educator stress: An occupational health perspective*, 23-54.
[31] Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*, 202–248.
[32] Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and teacher education*, *17*, 7, 783-805.
[33] Wang, Y., & Wang, Y. (2022). The interrelationship between emotional intelligence, self-efficacy, and burnout among foreign language teachers: a meta-analytic review. *Frontiers in Psychology*, *13*, 913638.
[34] Wapaño, M. R. R. (2021). Emotional intelligence, self-efficacy and occupational stress of academic personnel. *International Journal of Research and Innovation in Social Science*, *5*, 5, 264-75.
[35] Wu, Y., Lian, K., Hong, P., Liu, S., Lin, R. M., & Lian, R. (2019). Teachers' emotional intelligence and self-efficacy: Mediating role of teaching performance. *Social Behavior and Personality: an international journal*, *47*, 3, 1-10.
[36] Yoke, L. B., & Panatik, S. A. (2015). Emotional intelligence and job performance among school teachers. *Asian Social Science*, *11*, 227–234.

VAT Harmonisation and EU Integration in Albania

Mikel ALLA[a,b,c*]

[a] *Accounting & Fiscal Expertise Studio, Elbasan, Albania*
[b] *University of Elbasan "Aleksandër Xhuvani", Elbasan, Albania*
[c] *University of Durrës "Aleksandër Moisiu", Durrës, Albania*
Correspondence: Lagjja Çlirimi, Rr. Andon Beça, Elbasan, Albania; mikel.alla@uniel.edu.al

Abstract

Albania's approximation to the European Union's value-added tax framework raises a conceptual problem: legislative alignment, administrative performance and digital deployment provide different evidence of convergence. This study develops a framework connecting these dimensions. A purposive integrative literature review is combined with doctrinal analysis of selected EU provisions and judgments and documentary analysis of Albania's 2025 accession assessments. Sources are compared by analytical function, evidentiary scope and temporal applicability. The synthesis yields four propositions: legal compatibility concerns the full treatment of a transaction; operational neutrality requires procedures that enforce liabilities and give effect to lawful recovery; the contribution of digital reporting depends on the interpretation and administrative use of records; and comparison requires a specified jurisdiction and period. The Albanian application distinguishes exemption alignment, fiscalisation deployment and interoperability preparation, while recognising reported progress in refunds. The contribution is a transaction-level framework that identifies the evidence needed to connect legal rules, administrative decisions and digital information. It provides a basis for targeted legal comparison and operational research, without estimating fiscal effects or establishing national compliance.

Keywords: VAT harmonisation; Albania; European integration; digital tax administration.

1. INTRODUCTION

VAT harmonisation raises a problem of institutional correspondence. A national system may reproduce the terminology of the European Union's common framework while applying different conditions to exemptions or deduction. It may also record invoices electronically without ensuring that the records support legally correct decisions. Comparing tax rates, legislative adoption or platform deployment in isolation therefore leaves the relationship between formal approximation and operational convergence unresolved.

Several bodies of research illuminate this problem, but they ask different questions. VAT studies examine the design and fiscal consequences of the tax [1, 2]. Enforcement research investigates how information changes reporting and collection [3, 4]. Institutional scholarship separates the adoption of formal rules from their implementation and intended outcomes [5]. Their intersection suggests an analytical gap: a common basis is needed for interpreting what evidence from each domain establishes about the same transaction.

Albania supplies a specific setting for this inquiry. The General Directorate of Taxation identifies Law No. 92/2014 as the central VAT statute and situates it within approximation to the European framework [6]. The

* Corresponding author

European Commission's 2025 report nevertheless identifies remaining exemptions outside the acquis. The same report describes fiscalisation digitisation as complete while distinguishing that development from the further use of information and preparation for interoperability with EU systems [7, pp. 72–73]. The coexistence of these findings makes it necessary to specify what each form of evidence establishes.

The European reference framework is itself changing. Directive (EU) 2025/516 introduces phased changes under the VAT in the Digital Age package, or ViDA. A comparison undertaken during this transition must distinguish rules already applicable from adopted requirements with future implementation dates [8]. For a candidate country, it must also separate the obligations of membership from arrangements that depend on the terms of accession. Neither a future accession date nor the outcome of a requested transition can be presumed.

The objective is to develop and apply a conceptual framework for assessing Albania's VAT convergence without conflating legal alignment, administrative performance and digital integration. The central research question asks how convergence should be assessed when these dimensions develop at different speeds. Three subsidiary questions guide the analysis: What distinguishes permissible national differentiation from legal divergence? How do enforcement procedures and the exercise of deduction and refund rights interact? Under what legal, organisational and temporal conditions can digital records support interoperable VAT administration?

The contribution is a transaction-level account of the connections between legal treatment, administrative procedure and information requirements. Functional convergence is defined here as the capacity to give effect to the relevant EU VAT treatment through legally appropriate procedures and usable information. The definition supplies an assessment criterion, not an observed country score. The paper develops the institutional strand of the author's earlier research on Albanian VAT [9] through a separate conceptual design. The literature establishes the mechanisms, the methods explain source selection and synthesis, and the findings apply the framework to selected legal and documentary issues before the discussion and conclusions assess its implications.

2. THEORY AND LITERATURE SYTHESIS

2.1. VAT architecture, revenue performance and the choice of benchmark

The economic rationale of VAT depends on the interaction between output taxation and recovery of eligible input tax. A comparison that observes only the rate charged on sales omits the mechanism through which VAT is intended to tax final consumption rather than accumulate through production. The OECD's International VAT/GST Guidelines connect this architecture with neutrality and the coordination of cross-border taxation, particularly for services and intangibles [10]. Their principles provide an analytical reference, although they do not replace the specific obligations of EU law.

VAT adoption and VAT harmonisation are different objects of explanation. Keen and Lockwood [1] associate adoption with revenue gains that vary across national circumstances. Alavuotunki et al. [11], using alternative cross-country specifications and instrumental-variable analysis, do not find a general increase in government revenue attributable to VAT adoption. The contrast concerns identification, sample composition and the policy intervention being studied; it cannot be resolved by assuming that one universal effect applies. Neither study directly estimates the consequences of bringing an existing VAT into conformity with EU law.

Keen [2] separates components of VAT performance through the analysis of rates and collection relative to a consumption benchmark. Acosta-Ormaechea and Morozumi [12] likewise distinguish changes associated with the standard rate from changes in VAT design in their analysis of growth. Neither approach identifies EU legal conformity from revenue statistics. Their relevance is methodological: a fiscal outcome becomes interpretable only after the mechanism and comparison benchmark are specified.

This distinction is particularly clear in Cnossen's [13] analysis of the EU VAT's departures from a broad consumption-tax benchmark. A legal framework can itself contain differentiated treatment. Consequently, deviation from a theoretical revenue benchmark and deviation from the acquis are not synonymous. The present paper treats revenue performance, economic neutrality and legal compatibility as connected but non-equivalent concepts. It does not assume that eliminating a measured policy gap is a legal condition of accession.

2.2. Information, enforcement and the limits of digital reporting

Information-based enforcement research explains why the invoice-credit mechanism can create incentives for transactions to be recorded. Pomeranz's [3] randomised experiments in Chile identify deterrence and spillover mechanisms associated with the VAT paper trail. Naritomi's [14] study of consumer participation in São Paulo

examines a different margin: incentives for final consumers to request receipts and report transactions. Together, these studies distinguish information generated by commercial relationships from information generated through consumer participation. Their results are not evidence that every stage of a VAT chain is automatically self-enforcing.

The information mechanism has limits. Carrillo et al. [4] examine firms notified of revenue discrepancies in Ecuador and show that responses can be constrained by enforcement capacity and offset through other reporting margins. Their setting concerns business income taxation rather than a direct test of Albanian VAT. Its use here is therefore conceptual: observing one component of the tax base does not necessarily determine the final liability or payment. Slemrod's [15] review places information reporting within a wider system of audit, remittance arrangements and taxpayer responses rather than treating information as a sufficient enforcement instrument.

Electronic invoicing changes the speed and structure of information without eliminating these distinctions. Bellon et al. [16] exploit Peru's phased introduction of VAT e-invoicing and find changes in reported activity and liabilities. They also identify the importance of existing VAT credits for the translation of liabilities into collections. The study therefore separates digital adoption, reporting responses and cash revenue. Its findings support examining the mechanisms connecting those stages, not transferring a Peruvian effect size to Albania.

Mascagni et al. [17] add a different perspective through mixed-method evidence from Rwanda. Their analysis links inconsistencies in VAT implementation with complexity, enforcement and unequal effective burdens. The contribution is not a universal description of lower-income administrations. It demonstrates that formal proportionality and an invoice-credit structure can coexist with operational differences across businesses. Besley and Persson's [18] account of fiscal capacity provides the broader institutional context for considering why the same statutory instrument can function differently across settings.

The more recent literature makes the administrative complementarity explicit. Kotsogiannis et al. [19] examine Rwandan tax filings and associate e-invoicing with higher net VAT payments and more effective audits, attributing the response among audited firms to audit efficiency rather than simply to greater voluntary compliance after adoption. Read alongside Bellon et al. [16], this suggests that information affects different margins depending on credit positions and administrative use. The studies support a conditional mechanism, not an invariant fiscal return to digitalisation.

Okunogbe and Santoro [20] distinguish the technological capacity to identify and monitor a tax base from adoption, institutional integration and regulatory conditions that determine its use. Okunogbe and Tourek [21] further connect technology with tax officials' incentives and political support for capacity building. These reviews qualify an exclusively technical interpretation of e-invoicing: more accessible information can assist enforcement, but organisational decisions determine whether and how it informs action. For the present framework, that complementarity also creates a point of contact with taxpayer rights, because information must be assessed through an authorised decision procedure.

2.3. European rule transfer and institutional implementation

Research on Europeanisation offers an explanation for studying adoption separately from domestic operation. Schimmelfennig and Sedelmeier [22] examine EU rule transfer in Central and Eastern European candidate countries through an external-incentives framework. Their account relates adoption to the conditions of rule transfer and domestic adoption costs. It is relevant to the structure of an accession inquiry, but it neither determines Albania's current institutional position nor predicts its accession trajectory.

Bromley and Powell [5] distinguish two forms of institutional decoupling. One concerns a gap between formal policy and practice; the other concerns implemented practices whose connection with intended outcomes remains uncertain. Applied as an analytical lens, this distinction prevents two different situations from being collapsed: legislation that is not operationalised, and operational procedures that do not establish the outcome attributed to them. The framework does not require an inference of symbolic intent or bad faith by any Albanian institution.

For VAT, the distinction can be expressed through a sequence of questions. Has the relevant treatment been enacted? Can the required procedure be carried out? Does that procedure establish the correct liability or entitlement? Can the relevant information retain its meaning when exchanged? Each question requires different evidence. A formal rule, an operational system and a correctly resolved transaction may therefore represent different stages of institutional correspondence rather than alternative measures of the same outcome.

The European Interoperability Framework strengthens this interpretation by distinguishing legal, organisational, semantic and technical aspects of public-service interoperability [23]. Its value for the present study is conceptual, not as a separate VAT accession test. It explains why compatible software interfaces cannot, on their own, resolve

differences in legal categories or organisational responsibilities. Digital integration is consequently treated as an institutional relationship involving information, not merely an information-technology installation.

2.4. Fiscal neutrality as a legal and administrative relationship

Economic neutrality and the legal principle of fiscal neutrality overlap but should not be used interchangeably. The former concerns the effects of tax treatment on economic decisions; the latter also guides interpretation of particular EU VAT obligations and entitlements. Their relationship becomes concrete in deduction and refund disputes. In Enel Maritsa Iztok 3 AD, the Court connected refund conditions with the taxpayer's ability to recover excess VAT without bearing an inappropriate financial burden [24].

The evidentiary dimension is developed in Barlis 06. The judgment distinguishes substantive conditions for deduction from shortcomings in invoice particulars where the administration has the information necessary to establish the entitlement [25, paras. 42–44]. Mahagében and Dávid address a related boundary between the prevention of fraud and the conditions under which a purchaser's deduction may be refused [26]. These cases connect tax enforcement with the allocation of evidentiary responsibilities rather than treating taxpayer protection as external to the operation of VAT.

The resulting theoretical tension is productive. More information can assist the administration, yet the legal significance of an inconsistency cannot be determined solely by its detection. A comparison of harmonisation must therefore consider the procedure linking an observation to a decision. That procedure includes the opportunity to establish substantive facts and the legal basis for accepting or denying an entitlement. This relationship is especially relevant when electronic records are used to initiate administrative action.

2.5. The analytical gap addressed by the study

The reviewed literature explains distinct parts of the problem: VAT design establishes the tax mechanism; enforcement research explains the conditional role of information; institutional research separates adoption from operation; and legal doctrine identifies the conditions governing liabilities and rights. Reading these strands together exposes a specific analytical problem: evidence about one part is often insufficient to establish correspondence across the whole transaction.

The organising concept is legal–procedural–information consistency: a traceable relationship between the treatment applicable to a transaction, the procedure used to determine it and the information supporting that procedure. Internal consistency alone is insufficient: a well-administered domestic rule may still differ from the EU benchmark. Assessment therefore requires both an appropriate legal reference and evidence that procedures give it effect. This distinction connects the literature on implementation with a specific inquiry into what administrative information and decisions establish about VAT convergence.

3. METHODS

3.1. Research design and epistemic purpose

The study combines theory synthesis, doctrinal legal analysis and documentary analysis. Theory synthesis connects mechanisms developed in separate research traditions: VAT architecture explains liabilities and input-tax recovery; information-based enforcement explains transaction visibility and behavioural responses; institutional theory distinguishes formal adoption, practical implementation and outcomes. This use of complementary theoretical roles follows the design principles for conceptual articles discussed by Jaakkola [27]. The purpose is analytical explanation, rather than estimation or theory testing with original observations.

Doctrinal analysis examines selected EU provisions and judgments to identify the conditions governing tax treatment, entitlements and evidence. The analysis reads the relevant provisions together, distinguishes substantive conditions from procedural requirements and confines case-law inferences to the issues addressed by the judgments. Documentary analysis then uses official accession material to identify Albania-specific applications. Following Hutchinson and Duncan [28], the legal reasoning is made explicit rather than treated as an unspecified review of legislation.

The analytical unit is the connection between a VAT rule, the procedure that gives it effect and the information required in that procedure. The starting point is the treatment of a supply or acquisition; deduction and refund analysis also follows the relevant tax period or claim where several transactions are aggregated. Five areas were

selected for their relevance to these connections: rates and exemptions, deduction and refunds, invoice evidence, cross-border classification and electronic reporting. They provide theoretically informative illustrations. No original firm-level dataset, survey, experiment or econometric analysis is used.

3.2. Literature identification and source selection

The review is purposive and integrative. It uses the existing VAT, enforcement and institutional literature as a starting set, followed by targeted title and thematic searches and verification against publisher pages, scholarly repositories and Crossref bibliographic records. Thematic searches focused on VAT, electronic invoicing, tax audits, information technology and tax mobilisation; related methodological sources clarify conceptual synthesis and doctrinal analysis. EU materials were consulted through EUR-Lex, the European Commission and the Council; Albanian statutory context was checked through the General Directorate of Taxation. Source verification and targeted updating were completed on 27 September 2026.

A source was retained when it addressed a mechanism required by the research questions, had identifiable authorship or institutional authority, and provided accessible material supporting the claim for which it was used. Foundational studies were retained for their explanatory relevance; recent peer-reviewed work extends the digital-administration discussion through 2025. Published journal versions take precedence over working-paper versions. Search snippets and secondary descriptions served as discovery aids, not independent evidence. Bibliographic verification establishes publication details; substantive use is limited to findings supported by the accessible source text.

The design is an integrative review, consistent with Snyder's [29] distinction between review approaches. It does not provide the exhaustive search and screening record required for a systematic review. Scopus and Web of Science were not searched directly. Crossref was used for bibliographic verification rather than as evidence of study quality. The reference list identifies the sources used; the resulting synthesis remains subject to selection and accessibility bias.

3.3. Unit of analysis, source hierarchy and temporal scope

Sources were classified by analytical function: economic mechanisms, institutional explanations, binding EU legal rules and their judicial interpretation, official country assessments, and domestic legislative context. These classes are not interchangeable evidence. An empirical study can identify a mechanism within its research design; a judgment can clarify a legal condition; an accession report can document an institutional assessment. Table 1 summarises the permissible inference and boundary for each source class.

The legal comparison uses Directive 2006/112/EC as consolidated on 14 April 2025, read together with the staged application provisions of Directive (EU) 2025/516. The principal Albania-specific evidence comprises the EU common position of 8 May 2025 and the Commission's 2025 report, which covers 1 September 2024 to 1 September 2025 [7, p. 3]. The September 2026 source-verification date does not extend that reporting period. Domestic repositories establish the statutory setting; a comprehensive article-by-article concordance of Albanian and EU legislation is outside the design.

Three temporal statuses are distinguished: rules applicable within the EU at the comparison date, adopted provisions with later application dates, and arrangements dependent on accession negotiations. Country-report findings retain their reporting period. Requests for derogations remain requests unless an authoritative instrument establishes otherwise. EU judgments are used as benchmarks for interpretation of the acquis, without assuming that candidacy alone makes every membership obligation directly applicable in Albania.

3.4. Analytical procedure and treatment of rival interpretations

Synthesis proceeded in four steps. First, each source was assigned an evidentiary role and connected to a research question. Second, its relevant rule, mechanism or institutional observation was compared with the procedure and information needed for the transaction concerned. Third, contrasting explanations were examined: a permitted rate difference can coexist with legal compatibility; formal invoice defects can coexist with a substantiated deduction; reporting increases can coexist with limited cash collection because credits intervene. Fourth, the recurring relationships were expressed as four bounded analytical propositions, presented in Table 2.

This procedure distinguishes source findings from the paper's interpretation. The propositions express relationships derived from the literature and doctrinal examples; their explanatory and practical value remains open

to subsequent examination. Documentary application to Albania identifies whether a source establishes a legal difference, an administrative development or a digital capability, and what evidence would connect these observations. The analysis uses qualitative comparison and reasoned interpretation, without numerical aggregation or statistical hypothesis testing.

The synthesis also considers rival interpretations. Better detection may account for increased assessments without a corresponding change in voluntary compliance. Completed deployment may describe information capture without establishing effective use. Faster refunds may indicate procedural progress without resolving every dispute about entitlement. These alternatives prevent a single administrative output from being treated as proof of national convergence.

Table 1. Source roles and limits of inference

Source class	Analytical use	Boundary
EU legislation and judgments [8, 24, 25, 26, 30]	Identify legal treatment, entitlements, evidence and application dates.	Do not establish the frequency of correct application in Albania.
Accession documents [7, 31]	Identify attributed findings and the status of negotiating requests.	Findings retain their reporting period; requests are not granted derogations.
Empirical VAT studies [3, 14, 16, 17, 19]	Compare information, reporting, audit and implementation mechanisms.	Foreign estimates do not measure Albanian effects or EU legal compatibility.
Institutional and digital frameworks [5, 20, 21, 23]	Distinguish adoption, use, outcomes and shared information meaning.	Conceptual categories do not establish a national administrative defect.
Albanian statutory repositories [6, 32]	Identify the VAT and fiscalisation legal setting.	Repository review is not a comprehensive legal concordance or a system test.

Source: Author's classification of the cited sources.

3.5. Validity and limits of inference

Analytical validity depends on maintaining distinctions between legal compatibility, effective implementation, interoperability and fiscal outcomes. Legal inferences are tied to primary provisions and judgments; country-specific statements retain institutional attribution and date. Empirical studies are compared with attention to research design, jurisdiction and outcome definition. Their mechanisms inform the framework, but foreign effect sizes are not transferred to Albania.

The main limitations are purposive source selection, uneven access to published material and the absence of original administrative observations. The analysis cannot establish the frequency of correctly resolved Albanian transactions, test the framework's predictive validity or certify comprehensive legal conformity. Its contribution is a reasoned framework whose applications can be examined against transaction records, legal decisions and system specifications in later research.

4. FINDINGS FROM THEORITICAL AND DOCUMENTARY ANALYSIS

4.1. Rates, exemptions and legally relevant comparability

Articles 97 and 98 of the VAT Directive distinguish the minimum standard rate from the conditions governing reduced rates. Those conditions concern scope as well as percentages. A national treatment must therefore be compared with the relevant category, entitlement and conditions, not only with another country's rate [30, Arts. 97–98 and Annex III].

The Commission's 2025 assessment records Albania's standard rate of 20% and reduced rates of 6% and 10%, alongside the remaining exemption differences [7, p. 72]. These findings identify the subject of comparison but do not establish that every application of a listed rate is either permitted or prohibited. The transaction-level question is whether the precise supply falls within a common rule, an available national option or an agreed exception.

The distinction between exemptions with and without deduction is equally consequential. Two supplies may both carry no output VAT while producing different recovery rights for the supplier. Numerical similarity at the invoice

level would then conceal a difference elsewhere in the tax mechanism. Conversely, different permitted rates need not indicate a conflict with the acquis. The proposed framework therefore examines the combination of classification, output treatment and input recovery rather than treating any one element as a sufficient measure of legal alignment.

This also explains why the C-efficiency literature cannot be used as a legal compliance test. A legally permitted exemption may contribute to a difference from a broad consumption-tax benchmark. The converse is also possible: an incompatibility affecting a narrow transaction class might have a limited effect on aggregate revenue. The conclusion follows from the different objects being compared, not from a claim that revenue considerations are irrelevant [2, 13].

4.2. National options and accession-dependent arrangements

The EU common position of 8 May 2025 provides two instructive examples. It records Albania's request to maintain an exemption without deduction for specified agricultural inputs until the end of 2035 and invites reconsideration. For international passenger transport, it identifies an existing option under Article 98(2)(a) and Annex III, point 5, rather than a need for the requested derogation. These are the positions documented at that negotiating stage, not final accession arrangements [31, pp. 6–7].

The examples support a three-way classification: treatment already possible within the common rules, treatment requiring a change in legal form or conditions, and treatment dependent on a separately agreed exception. Classification matters because it determines what evidence can resolve the issue. A domestic objective does not by itself establish compatibility, while the existence of differentiation does not by itself establish incompatibility.

The analytical lesson is narrower than an appraisal of the merits of either measure. It is that legal convergence requires identifying the source of permission. A complete comparison must therefore record the scope of a provision and the basis on which it is applicable. Copying the rate or terminology of a Member State is insufficient where the underlying authorisation differs.

4.3. Deduction and refunds: the procedural dimension of neutrality

The Directive links deduction to qualifying uses of inputs and permits nationally specified arrangements for excess deductions [30, Arts. 168 and 183]. The entitlement and its administration are consequently parts of the same tax mechanism. A formal right that cannot be exercised on the applicable conditions differs from a right whose conditions can be established through a functioning procedure.

In Enel Maritsa Iztok 3 AD, the Court related repayment procedures to recovery within a reasonable period and protection against financial risk; it also addressed interest during verification and accepted that set-off is not inherently excluded [24, paras. 33, 61 and 64–67]. The interpretive point is that procedural arrangements can affect the substance of recovery. This EU benchmark informs the convergence inquiry without establishing how Albanian refunds are administered in individual cases.

This connection has an economic interpretation. Where excess deductible VAT arises, the timing of recovery affects the period during which funds remain unavailable to the business. Consequently, collection and refund procedures cannot be evaluated as entirely separate administrative outputs. Lawful net collection depends on recognising eligible deductions as well as enforcing liabilities. The argument does not assume that all refund claims are valid or that verification is unnecessary. It requires a distinction between establishing an entitlement, preventing an unsupported claim and delaying a substantiated claim.

A refund claim therefore links the legal grounds for recovery, the procedure determining the amount and timing, and the records supporting the decision across the relevant transactions and tax periods. An operational assessment would trace these links through the claim and its resolution. Processing counts or retained revenue alone cannot show whether a substantiated entitlement was correctly determined and given effect.

The Albania application also contains evidence of progress. The Commission reports improved speed in VAT refunds for businesses, while identifying further implementation needs for refunds connected to EU-funded NGO activities [7, p. 90]. This qualifies any account based solely on administrative shortfalls. It establishes a reported procedural improvement, while transaction-level evidence would still be required to evaluate timeliness, eligibility decisions and remedies across cases.

4.4. Invoice evidence, substantive conditions and the allocation of responsibility

Barlis 06 distinguishes substantive entitlement from deficiencies in specified invoice particulars. Where the administration possesses the information needed to establish the substantive requirements, those deficiencies alone cannot justify denial in the circumstances examined by the Court. The taxable person nevertheless bears the burden of establishing eligibility and may be required to produce supporting evidence [25, paras. 42–46]. The case thus supports correction and evidence assessment, not a general exemption from invoice requirements.

Mahagében and Dávid address refusal of a purchaser's deduction because of irregularities elsewhere in the supply chain. Where the relevant deduction conditions are fulfilled, refusal on the ground of a supplier's fraud requires objective evidence that the purchaser knew or ought to have known of the connection. The judgment also limits general obligations to verify suppliers' compliance [26, paras. 45–50 and 60–62]. Read alongside Barlis 06, it distinguishes substantiation of an entitlement from the grounds for denying deduction because of participation in fraud. This distinction does not determine criminal responsibility.

The legal relevance of a digital discrepancy depends on what it establishes. A missing field, a timing mismatch, an incorrect classification and an unsupported transaction are not interchangeable legal events. As a conceptual consequence of the cases, the procedure connecting detection to decision must be capable of distinguishing them. This observation concerns the design of analysis, not an allegation about automated decision-making in Albania.

4.5. Cross-border classification and jurisdictional scope

Cross-border comparison must distinguish goods from services and membership-based arrangements from transactions involving a third country. The VAT Directive allocates transactions through rules concerning supplies, intra-Community acquisitions, imports and the place of supply. The appropriate classification therefore depends on more than the physical presence of an invoice or the currency of payment [30, Titles IV–V].

The OECD Guidelines identify coordination problems arising when national systems interact, particularly for services and intangibles [10]. Their conceptual relevance is that consistent taxation requires compatible allocation rules. In the present framework, jurisdiction is thus part of the legal meaning that information must preserve when exchanged. A successful transmission between systems does not establish that both administrations are interpreting the same transaction category.

For Albania, a prospective comparison with membership arrangements must not be written as though those arrangements already apply domestically by virtue of candidacy. At the same time, business obligations arising from particular cross-border supplies need to be distinguished from state-level accession obligations. This distinction is maintained as a question of legal scope rather than as a prediction about accession.

4.6. From fiscalisation deployment to the use of evidence

Albania's fiscalisation repository identifies Law No. 87/2019 and implementing instruments concerning electronic invoices, technical requirements and the central invoice platform [32]. The Commission's 2025 distinction between completed digitisation and further requirements concerning data use provides a documented example of deployment and utilisation being assessed separately [7, p. 73].

The distinction corresponds to different points in an information process. A transaction must first be recorded; the record must be linked to the relevant person and period; its tax meaning must be determined; and any inconsistency must be addressed through an authorised procedure. Successful completion of the first stage provides evidence of visibility, not necessarily of correct interpretation at the last stage. This sequence specifies the links that an operational assessment would need to examine.

The economic literature supports separating reporting, liability and collection. Bellon et al. [16] emphasise the intervening role of credits; Kotsogiannis et al. [19] identify the importance of audit use; and Mascagni et al. [17] show how complexity affects implementation. The OECD's comparison of electronic invoicing arrangements likewise documents varied national configurations [33]. Consequently, platform coverage should be connected to evidence on classification, correction and follow-up before it is interpreted as functional convergence.

4.7. ViDA and the temporal organisation of convergence

The VAT in the Digital Age package combines digital reporting with reforms for certain platform-facilitated supplies and VAT registration arrangements [34]. Directive (EU) 2025/516 schedules major cross-border digital

reporting changes for 1 July 2030. The relevant platform deemed-supplier measures may apply between 1 July 2028 and 1 January 2030. The 1 January 2035 alignment deadline concerns specified pre-existing domestic reporting systems and authorisations; it is not a general transition automatically available to every future Member State [8, Art. 6].

A temporally coherent comparison therefore asks which version of a requirement governs the transaction at issue. Legislative adoption, national implementation and the date when a provision must be applied are distinct events. An analysis that merges them could identify a present incompatibility by reference to a rule not yet applicable, or overlook an adopted change relevant to future system design.

For a candidate country, the same distinction prevents an assumed accession timetable from determining the legal conclusion. The study compares the documented current framework and adopted European developments without assigning Albania an accession date or an unnegotiated transition. The substantive finding is that temporal applicability is part of the correspondence between a rule and its administration, rather than an ancillary chronological detail.

4.8. Semantic consistency, correction and institutional responsibility

The European Interoperability Framework distinguishes the ability to exchange data from the ability to preserve shared meaning and coordinate organisational responsibilities [23]. Applied to VAT, semantic consistency means that categories such as an exemption, a reverse-charge transaction or a correction retain the legally relevant meaning required by the receiving procedure. Identical field names do not necessarily establish such consistency; different labels may nevertheless be compatible through an accurate mapping.

Consider an illustrative transaction recorded with no output VAT. That record alone does not identify whether the underlying supply is exempt with deduction, exempt without deduction, outside the relevant territorial scope or subject to another treatment. Its administrative meaning depends on the classification and associated conditions. This example is constructed to examine the framework; it is not a reported case or an empirical finding.

Correction creates a further interface. If a record changes, the procedure needs to identify which transaction and period are affected and whether the change concerns form, tax classification or amount. The legal analysis then determines the consequences; technical acceptance cannot decide them by itself. Barlis 06 makes the distinction between invoice particulars and substantive entitlement relevant to that reasoning, without establishing a general rule that every electronic validation requirement is impermissible [25].

Institutional responsibility is the remaining link. A discrepancy becomes usable evidence only through an account of who can review it, what supporting material is relevant and how a decision can be corrected or challenged. These questions join enforcement and taxpayer protection within the same process. They also identify the evidence that would be necessary for a subsequent operational assessment: specifications, decision procedures and transaction histories, not simply confirmation that a platform exists.

5. DISCUSSION

5.1. The contribution of the conceptual framework

The findings support a framework in which the transaction provides a common reference point for three distinct inquiries. Legal alignment identifies the authorised treatment. Administrative implementation establishes the procedure for determining liability or entitlement. Digital interoperability concerns the information required by that procedure and the preservation of its meaning when exchanged. The framework is relational: its value lies in tracing a specific connection, rather than adding separate indicators into a readiness score.

Table 2 links each proposition to its supporting sources and the evidence needed for application. The first three propositions concern legal treatment, administration and information; the fourth fixes their jurisdictional and temporal scope. In a subsequent operational assessment, evidence could support alignment for the issue examined, identify a specific discrepancy or leave the relationship unresolved. Missing evidence belongs to the third category and is not itself proof of administrative failure. These distinctions give the framework a diagnostic purpose without implying a national readiness score.

Table 2. Analytical propositions and evidence for their application

Analytical proposition	**Basis in the synthesis**	**Evidence required for application**
P1. Compatibility concerns the full legal treatment of a transaction.	VAT architecture, authorised rate options and deduction consequences [2, 13, 30, 31].	Provision, scope, eligibility, input recovery and legal basis of any exception.
P2. Operational neutrality includes enforcement and lawful recovery.	Invoice-credit mechanism and deduction and refund judgments [24, 25, 26, 30].	Evidence considered, reasons for the decision, recovery timing and remedies.
P3. Digital reporting supports convergence through usable evidence.	Enforcement mechanisms, information meaning and administrative use [16, 19, 20, 23, 25].	Classification rules, correction histories, data meaning and follow-up procedures.
P4. Comparison requires a specified jurisdiction and period.	Staged ViDA requirements and accession-dependent arrangements [8, 31].	Applicable legal version, commencement date and status of any transition.

Source: Author's conceptual synthesis of the cited literature, legal provisions and accession documents.

5.2. Interpretation and scope of the propositions

P1 follows from the difference between numerical uniformity and legally authorised differentiation. A reduced rate permitted by the common framework can differ from another jurisdiction's rate without creating a conflict. Conversely, an apparently similar output treatment can conceal different deduction consequences. The proposition directs comparison towards the legal basis and complete treatment of the supply, while leaving unexamined transactions unresolved.

P2 connects collection and taxpayer entitlements within the same VAT mechanism. The legal cases do not make every refund payable without verification; they require the decision to address the applicable conditions. Administrative effectiveness therefore includes the capacity to reject unsupported claims, recognise substantiated deductions and provide the recovery required by law. Gross collections or processing counts cannot independently establish that capacity.

P3 treats digital reporting as a potential input to functional convergence. The empirical literature supplies mechanisms through which information influences detection and reporting; the legal analysis identifies what an ensuing decision must establish. Their integration makes classification, evidentiary sufficiency, correction and administrative responsibility central to evaluation. It also explains why a reporting gain, an audit gain and a cash-revenue gain are different outcomes rather than interchangeable measures of digital success.

P4 defines the temporal boundary of the first three propositions. A comparison is interpretable only when its rule, jurisdiction and period are specified. ViDA makes this requirement visible through staged application dates; accession requests make it visible through the difference between a proposed exception and a legally available one. Preparation for an adopted future rule is relevant to planning but does not establish a present violation or a negotiated entitlement.

5.3. Theoretical implications and a bounded research agenda

The synthesis refines the use of institutional decoupling in VAT analysis. Policy–practice decoupling directs attention to whether an enacted treatment can be administered. Means–ends decoupling asks whether an implemented procedure establishes the outcome attributed to it [5]. Connecting each question to a transaction's legal conditions makes it possible to distinguish incomplete implementation from a mistaken interpretation of a functioning administrative output. The framework does not diagnose either form of decoupling across Albania as a whole.

A further implication concerns policy sequencing. A legal change may require revised invoice categories and staff guidance; a reporting change may require procedures for resolving mismatches before data can support decisions. The framework therefore suggests analysing dependencies between reforms. It does not imply that investment must follow one universal sequence, since existing capacities, transaction risks and legally fixed deadlines can justify parallel implementation.

The framework also has limits of explanatory reach. Political incentives, taxpayer trust, compliance costs and organisational capacity may influence how the identified connections operate [18, 21]. Legal–procedural–

information consistency specifies what must be examined for a convergence claim; it does not by itself explain all determinants of compliance or establish the optimal allocation of administrative resources.

6. CONCLUSIONS AND IMPLICATIONS

The analysis answers the central research question by locating VAT convergence in the relationship between the applicable rule, the procedure giving it effect and the information supporting the decision. These elements should be assessed against the same legal benchmark, jurisdiction and period. The Albanian documents illustrate why this matters: exemption alignment, fiscalisation deployment and refund performance describe different dimensions of reform. Their coexistence calls for a differentiated assessment rather than a single inference from legislative adoption or platform coverage.

The first subsidiary question is answered by distinguishing authorised differentiation from incompatible treatment. Scope, eligibility conditions and input-tax recovery matter alongside rates. The agricultural-input and passenger-transport examples show why a requested derogation and an option already available under EU rules require different legal responses. Harmonisation concerns the legal basis and conditions of treatment; numerical uniformity across jurisdictions is not an adequate benchmark.

The second answer connects enforcement with lawful recovery. Deduction and refund procedures contribute to VAT neutrality when they distinguish substantiated entitlements from unsupported claims and give effect to the applicable decision. The EU judgments define relevant interpretive boundaries. The Commission's account of faster Albanian refunds provides a documented example of procedural progress, while its remaining implementation concerns require examination at the level of particular claims and decisions.

The third answer concerns the conditions under which digital information can support convergence: records must preserve legally relevant meaning, allow correction and inform decisions for which institutional responsibility is clear. Fiscalisation deployment provides an infrastructure for these functions, but its operation and interoperability require separate evidence. The phased ViDA rules add a temporal requirement: present obligations, adopted future measures and accession-dependent arrangements must remain distinct in assessment and reform planning.

The theoretical contribution is to connect VAT architecture, information-based enforcement and institutional implementation through four bounded analytical propositions. This specifies how legal and administrative evidence can be compared without conflating their meanings. For Albanian policymakers, the proposed application is to connect legal concordance work with procedure design and data definitions. An assessment should record the applicable legal basis, any authorised national option, deduction consequences, evidence requirements, responsible administrative function and application date. It should distinguish a demonstrated discrepancy from a relationship for which evidence remains insufficient.

For tax administrators and businesses, the analysis supports attention to correction histories, reasons for decisions and the treatment of input-tax entitlements alongside invoice transmission. Such records would make it possible to distinguish technical completion from legally effective processing. The study does not estimate the costs or revenue benefits of these measures, so their prioritisation requires further operational evidence.

The conclusions are limited by purposive literature selection, reliance on official documentary assessments and the absence of original administrative observations. Future research should compare selected Albanian and EU transaction rules provision by provision; trace refund claims through eligibility decisions, payment and review; and examine whether invoice corrections preserve the legal classification and liability across systems. Such studies should record both conformity and discrepancies, distinguishing unresolved evidence from demonstrated failure. Comparison across candidate countries could assess the framework's transferability. Its practical value ultimately depends on whether these distinctions improve the identification and explanation of specific convergence problems.

References

[1] Keen, M., & Lockwood, B. (2010). The value added tax: Its causes and consequences. *Journal of Development Economics, 92*(2), 138–151. https://doi.org/10.1016/j.jdeveco.2009.01.012

[2] Keen, M. (2013). The anatomy of the VAT. *National Tax Journal, 66*(2), 423–446. https://doi.org/10.17310/ntj.2013.2.06

[3] Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review, 105*(8), 2539–2569. https://doi.org/10.1257/aer.20130393

[4] Carrillo, P., Pomeranz, D., & Singhal, M. (2017). Dodging the taxman: Firm misreporting and limits to tax enforcement. *American Economic Journal: Applied Economics, 9*(2), 144–164. https://doi.org/10.1257/app.20140495

[5] Bromley, P., & Powell, W. W. (2012). From smoke and mirrors to walking the talk: Decoupling in the contemporary world. *Academy of Management Annals, 6*(1), 483–530. https://doi.org/10.5465/19416520.2012.684462
[6] General Directorate of Taxation of Albania. (n.d.). Tatimi mbi vlerën e shtuar [Official legislative repository]. Retrieved September 27, 2026, from https://www.tatime.gov.al/c/6/71/tatimi-mbi-vleren-e-shtuar
[7] European Commission. (2025). Albania 2025 report (SWD(2025) 750 final). https://enlargement.ec.europa.eu/document/download/fe9138b7-90fe-4277-a12c-3a03f6d1957f_en?filename=albania-report-2025.pdf
[8] Council of the European Union. (2025). Council Directive (EU) 2025/516 of 11 March 2025 amending Directive 2006/112/EC as regards VAT rules for the digital age. https://eur-lex.europa.eu/eli/dir/2025/516/oj/eng
[9] Alla, M. (2016). Tatimi mbi vlerën e shtuar në Shqipëri e në vendet e tjera të Ballkanit dhe ndikimi i tij te të ardhurat [Doctoral dissertation, Universiteti Europian i Tiranës]. https://www.uet.edu.al/wp-content/uploads/2021/11/MIKEL_ALLA.pdf
[10] OECD. (2017). International VAT/GST guidelines. OECD Publishing. https://doi.org/10.1787/9789264271401-en
[11] Alavuotunki, K., Haapanen, M., & Pirttilä, J. (2019). The effects of the value-added tax on revenue and inequality. *The Journal of Development Studies, 55*(4), 490–508. https://doi.org/10.1080/00220388.2017.1400015
[12] Acosta-Ormaechea, S., & Morozumi, A. (2021). The value-added tax and growth: Design matters. *International Tax and Public Finance, 28*(5), 1211–1241. https://doi.org/10.1007/s10797-021-09681-2
[13] Cnossen, S. (2022). The C-inefficiency of the EU-VAT and what can be done about it. *International Tax and Public Finance, 29*(1), 215–236. https://doi.org/10.1007/s10797-021-09683-0
[14] Naritomi, J. (2019). Consumers as tax auditors. *American Economic Review, 109*(9), 3031–3072. https://doi.org/10.1257/aer.20160658
[15] Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature, 57*(4), 904–954. https://doi.org/10.1257/jel.20181437
[16] Bellon, M., Dabla-Norris, E., Khalid, S., & Lima, F. (2022). Digitalization to improve tax compliance: Evidence from VAT e-Invoicing in Peru. *Journal of Public Economics, 210*, Article 104661. https://doi.org/10.1016/j.jpubeco.2022.104661
[17] Mascagni, G., Dom, R., Santoro, F., & Mukama, D. (2023). The VAT in practice: Equity, enforcement, and complexity. *International Tax and Public Finance, 30*(2), 525–563. https://doi.org/10.1007/s10797-022-09743-z
[18] Besley, T., & Persson, T. (2014). Why do developing countries tax so little? *Journal of Economic Perspectives, 28*(4), 99–120. https://doi.org/10.1257/jep.28.4.99
[19] Kotsogiannis, C., Salvadori, L., Karangwa, J., & Murasi, I. (2025). E-invoicing, tax audits and VAT compliance. *Journal of Development Economics, 172*, Article 103403. https://doi.org/10.1016/j.jdeveco.2024.103403
[20] Okunogbe, O., & Santoro, F. (2023). The promise and limitations of information technology for tax mobilization. *The World Bank Research Observer, 38*(2), 295–324. https://doi.org/10.1093/wbro/lkac008
[21] Okunogbe, O., & Tourek, G. (2024). How can lower-income countries collect more taxes? The role of technology, tax agents, and politics. *Journal of Economic Perspectives, 38*(1), 81–106. https://doi.org/10.1257/jep.38.1.81
[22] Schimmelfennig, F., & Sedelmeier, U. (2004). Governance by conditionality: EU rule transfer to the candidate countries of Central and Eastern Europe. *Journal of European Public Policy, 11*(4), 661–679. https://doi.org/10.1080/1350176042000248089
[23] European Commission. (2017). European Interoperability Framework—Implementation strategy (COM(2017) 134 final, including Annex 2). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52017DC0134
[24] Court of Justice of the European Union. (2011, May 12). Enel Maritsa Iztok 3 AD (C-107/10, ECLI:EU:C:2011:298). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62010CJ0107
[25] Court of Justice of the European Union. (2016, September 15). Barlis 06—Investimentos Imobiliários e Turísticos SA (C-516/14, ECLI:EU:C:2016:690). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62014CJ0516
[26] Court of Justice of the European Union. (2012, June 21). Mahagében kft and Péter Dávid (Joined Cases C-80/11 and C-142/11, ECLI:EU:C:2012:373). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62011CJ0080
[27] Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review, 10*(1–2), 18–26. https://doi.org/10.1007/s13162-020-00161-0
[28] Hutchinson, T., & Duncan, N. (2012). Defining and describing what we do: Doctrinal legal research. *Deakin Law Review, 17*(1), 83–119. https://doi.org/10.21153/dlr2012vol17no1art70
[29] Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039
[30] Council of the European Union. (2006). Council Directive 2006/112/EC of 28 November 2006 on the common system of value added tax (consolidated version of 14 April 2025). https://eur-lex.europa.eu/eli/dir/2006/112/2025-04-14/eng
[31] Council of the European Union. (2025, May 8). European Union common position: Cluster 3—Competitiveness and inclusive growth (AD 6/25, CONF-ALB 6/25). https://data.consilium.europa.eu/doc/document/AD-6-2025-INIT/en/pdf
[32] General Directorate of Taxation of Albania. (n.d.). Fiskalizimi [Official legislative repository]. Retrieved September 27, 2026, from https://www.tatime.gov.al/c/6/452/fiskalizimi
[33] OECD. (2022). Tax Administration 3.0 and electronic invoicing: Initial findings. OECD Publishing. https://doi.org/10.1787/2ffc88ed-en
[34] European Commission. (n.d.). VAT in the Digital Age (ViDA). Retrieved September 27, 2026, from https://taxation-customs.ec.europa.eu/taxation/vat/vat-digital-age-vida_en

In Contemporary Society, Man is Increasingly Manipulated by Education, Media, and Subjected to the Economic Factor: The Timeliness of the Critique of the Philosophers of the Frankfurt School

Jak SIMONI [a*]

[a] *"Aleksandër Moisiu" University,Facutly of Education, Spitalle Durrës, Albania. jaksimoni@yahoo.com*

Abstract

Although more than a century has passed since the initial rise of the thinkers of the Frankfurt School, the themes they addressed remain profoundly relevant today. While their thought has found reception in numerous countries worldwide, it has failed to yield substantial effects on society. The economic factor, of which they spoke, has become even more decisive in shaping human life. Those who hold political and economic power exercise increasingly authority. Artificial needs progressively eclipse genuine ones, and consumption is continuously steered by advertising. Educational systems aim at preparing professionals and specialists for the market, setting human values aside. Time to engage in contemplative thought is lacking; instead, other ways of spending free time are invented, like gaming,traveling etc. Advertising dictates the manner in which free time is spent. Ideas for escaping this condition, hitherto submerged, must now emerge. I think that insufficient debate has been devoted to the fact that producing the superfluous wastes the valuable time needed for reflection. These are the core issues addressed in this article.

Keywords: Frankfurt School, Critical Theory, consumerism, sobriety.

1. INTRODUCTION

Although capitalist social organization by producing immense development has triumphed over alternative social models, particularly the socialist model which brought terror and immense suffering to every country where it was implemented, critiques from within the system, as well as from the outside, have never been lacking. I consider the critique articulated by the thinkers of the Frankfurt School to be the most significant. This system has produced a materially comfortable lifestyle, yet in many respects, it has rendered man a kind of modern slave. Given the misery caused by socialism, this critique is often minimized through a binary logic that conceives of only two social setups without any middle ground. Indeed, this represents an error that limits the pursuit of other possibilities or the potential to reform and improve the system without awaiting revolutionary movements. The greatest revolutions are those carried out daily, above all by changing one's way of thinking; and this is precisely what the thinkers of this school maintained.

I shall not dwell on the history of this school, as it spans several generations and would require endless pages; rather, I shall address certain themes that appear particularly timely today. The article is structured into several

* Corresponding author

sections dealing with the reception of these thinkers, the role of education and the media, the submission of man to the economic factor, consumerism, and waste. The space accorded to these issues in literature, especially Italian literature, will be consistently highlighted. Finally, a prominent place will be given to contemporary proposals for change capable of liberating man from his current state of alienation.

1.1. The Reception of Frankfurt School Thinkers Up to the Present Day

The thinkers of this school found reception across various countries. Although they initially conducted their research in Frankfurt, dramatic political events forced its representatives to relocate, primarily to the United States, where they continued their work and thereby contributed to disseminating their critique in that country. Particularly after the Second World War, their works were translated into various languages, accompanied by the publication of numerous books, journal articles, essays, press contributions etc. The reception of the Frankfurt School has been studied by the Italian professor of Political and Social Philosophy, Marina Calloni, who, in an article published in the journal *Quaderni di Teoria Sociale*, demonstrates the complex dialectic characterizing the history and global reception of the Frankfurt School over time[1]. According to her, the development of such a current of thought is carried out in comparison with current issues and with competing currents of thought[2].

Also in Italy, an important study on this school was conducted by the philosopher Giuseppe D'Alterio, who authored a dedicated chapter on the Frankfurt School in his *Compendio del pensiero filosofico occidentale del XX secolo*. According to D'Alterio, the critique formulated by the representatives of this school serves a precise purpose, expressed as follows: "The objective is to investigate the organization of contemporary capitalist society (family, cultural institutions, mass media) in order to unmask its deep, hidden contradictions through a 'utopian' model capable of acting as a revolutionary ideal whose sole concern is a development that leads to an unexploited society, aiming at the realization of a future, free, and disalienated humanity"[3].

Instead, the Italian professor of Political philosophy at La Sapienza University, Stefano Petrucciani, writes about the importance and relevance of the thought of this school's representatives regarding social justice. Referring to Horkheimer, he writes that critical thought implies a concept of humanity that stands in contradiction with itself so long as this identity remains unfulfilled: "the interest in the abolition of social injustice"[4]. According to Petrucciani, Horkheimer maintains that the goal for which critical theory strives is a "rational society," wherein production is consciously planned and oriented toward the satisfaction of human needs[5]. As is evident, they advocate regulating the relationship between production and consumption; society must produce what is necessary, opposing production as an end in itself. Regarding productivity as an end in itself, Luca Baldassarre, lecturer in Philosophy and History, referencing the Frankfurt critique, observes that in a system where the goal is self-serving productivity, freedoms—including free choice and free initiative, the foundations of liberal politics and economics—lose their critical-negative character[6].

On the theme of human emancipation, which for these thinkers constitutes a prerequisite for the emancipation of society, Associate Professor of Political Philosophy Giorgio Fazio emphasizes in an article in the journal *SCIENZA & POLITICA*. Referring to Axel Honneth he writes: "The irrationality of society is thus a source of suffering for individuals, insofar as it hinders self-determination and self-realization, which would be achievable within a rational social organization based on mutual cooperation. Frankfurt Critical Theory nevertheless assumes that those who suffer socially, whether consciously or unconsciously, likewise desire to be liberated"[7]. The Italian professor contends that these thinkers propose a rational society in which individuals can fulfill themselves, a state achieved through human emancipation.

Recently, Martin Jay published the volume: *Immanent Critiques The Frankfurt School Under Pressure*. It is a book rich in historical arguments focusing notably on the 1968 movements in Western countries as well as those in Czechoslovakia, Poland etc. Movements that were inspired in part by Frankfurt School thought. According to Jay, man remains subjugated even in Western nations that self-identify as free. Referring to the critical perspective of the

---

[1]https://www.academia.edu/54072395/Marina_Calloni_La_divergente_unit%C3%A0_della_cosiddetta_Scuola_di_Francoforte_in_Quaderni_di_Teoria_Sociale_1_2_2020_pp_197_214

[2] Same p. 210

[3] La scuola di Francoforte: Horkheimer, Adorno, Marcuse, Habermas, Fromm: Collana "Compendio del pensiero filosofico occidentale del XX secolo", volume IV p.5

[4] Petrucciani, S. (2023) *La Scuola di Francoforte Storia e attualità.* Roma: Carocci editore. p. 19

[5] Same p.75

[6] Baldassarre, L. (2021) *La scuola di Francoforte. Una introduzione.* Firenze: Editrice Clinamen p.57

[7] https://scienzaepolitica.unibo.it/article/download/16377/15560 vol. XXXIV, no. 67, 2022, pp. 55-74 p.60

Frankfurt thinkers, he observes: “Immanent critique, understood in this way, counterposes a laudable objective to its realization, criticizing, for instance, countries that call themselves the ‘free world’ yet remain afflicted by tangible forms of servitude”[8].

As noted, the range of topics addressed by these thinkers is vast. In this regard, Luciano Gallino, in an article titled *Sociologia e teoria critica della società*, wrote: "These were themes such as the denunciation of the totalitarian character of capitalist society; the internalization of repression within the personality of its members; the family as an institution in which parents and children are educated and educate themselves to passively endure, as well as brutally exercise, authority in social life; the intrinsic connection between technological rationality and the logic of domination; and the diminishing possibility of an emerging collective subject capable of moving established society toward a historical alternative”[9]. From among these extensive themes, I shall subsequently focus on specific issues—such as the roles of education, the mass media, the economic factor, consumerism, and waste—which I regard as acutely relevant today.

1.2. In Contemporary Society, Man Is Increasingly Manipulated by Education and Mass Media

The educational system has diverted its primary function toward the service of the economy. This is apparent from early childhood structures, such as nurseries and preschools, which ‘remove’ children from their mothers to free the latter to pursue work and careers. Little is said regarding the trauma experienced by young children who feel abandoned. Ivan Illich was justified in viewing the school as an institution for safeguarding children while parents dedicate themselves to labor. Naturally, children must integrate into educational institutions, but not in a standardized, massified way. Integration should proceed gradually, respecting individual differences in psycho-physical development. At school, as well as within the family, children learn obedience and social conformity. On this matter, referring to Horkheimer, the Italian philosopher D'Alterio wrote: “The family is the privileged locus for the social reproduction of consent; within it, individuals are conditioned to believe in the principle of paternal authority as well as in possessive egoism”[10]. He then continues: “Such an education renders them fit to accept and internalize the rules of social domination at all levels”[11]. According to Adorno and Horkheimer, humanity is manipulated by the mass media and advertising. On this point, D'Alterio continued: “For the two philosophers, even the democratic West has transitioned from the totalitarian dictatorships of various political ideologies to the dictatorship of mass media, advertising, and the world of commodities, which, in order to develop according to its own logic, needs to invent more and more, manipulating needs and even ideas”[12].

Regarding the manipulation of needs by dominant economic interests, D'Alterio writes with reference to Marcuse: “The term totalitarian applies not only to a terroristic organization of society, but also to a non-terroristic, economic-technical organization that operates through the manipulation of needs by vested interests”[13]. In the dictatorship of the mass media, television occupies a very important place. In an article published in the journal *In Circolo*, the researcher Francesco Garbelli brings us a passage from the Italian intellectual Pier Paolo Pasolini who, regarding the critique of consumer society and the mass media, had much in common with the thinkers of the Frankfurt School: “The model was that of television which proposed, by exhibiting them, certain modes of existence that citizens passively absorbed through the consumption of images; the hedonistic society it had foreshadowed forced them to accept and reproduce the ever-same, creating consumers anxious to achieve the standard permitted by power”[14]. Indeed, Pasolini too is against consumerism which satisfies the needs of the market, rather than those of human beings.

Let us return to Marcuse, citing a passage where he wrote that never before has an “ideological domination” occurred to such an extent, exercised both by the ruling classes and by the pervasive and relentless apparatus of persuasion, which makes us slaves of the system and incapable of organizing our lives in an autonomous and

---

[8] Jay, M. (2025) *Critiche immanenti. La scuola di francoforte sotto pressione*. A cura di Roberto Mordacci. Traduzione di Cecilia Vergani. Milano: Mimesis p.26

[9] https://journals.openedition.org/qds/792 pp. 229-246

[10] La scuola di Francoforte: Horkheimer, Adorno, Marcuse, Habermas, Fromm: Collana "Compendio del pensiero filosofico occidentale del XX secolo", volume IV p.14

[11] Same p. 14

[12] Same p. 20

[13] Same p. 31

[14] https://www.researchgate.net/publication/355889890_IL_MARXISMO_FRANCOFORTESE_DI_PIER_PAOLO_PASOLINI p.12

different manner[15]. This passage is highly relevant today, because the media apparatus of persuasion has been perfected. Beyond traditional media, the Internet is a powerful instrument of persuasion. The philosophers of this school dreamed of a different society, where values predominate over vices. Apparently, an utopian society. D'Alterio, referring to Fromm, brings us a passage where he describes the new man of the future society: "The path to follow is that of encouraging the emergence of a new man, whose character structure possesses a series of qualities in which the virtues, merits, and values accumulated by humanity over time are condensed,' and which can be summarized as follows: loyalty, solidarity, sharing, participation, love and respect for life, for nature, for differences in all their manifestations, comprehensive and critical formation of the human being, rejection of every form of oppression, deception, destructiveness, and exploitation, promotion of freedom, justice, autonomy of thought, radical detachment from material things and power; ultimately, moving away from the existential mode of having to foster that of being"[16]. Fromm resolved the dilemma between being and having in favor of being.

1.3. Man Increasingly Subordinated to the Economic Factor

The economic factor remains dominant in human life. Not only the great mass of people, but also the ruling class has submitted to this factor, becoming alienated. All psycho-physical energies are dedicated to enrichment, career, and the economy. As we have mentioned, even the educational system prepares professionals, specialists, and functionaries for the market. Regarding this issue, the Italian philosopher Carlo Sini wrote: "We have a reduction of the graduate to a blinded person who must become a slave to the labor market as quickly as possible, naturally in the interest of the general growth of financial capital"[17]. The entire formation of man is aimed at production. On this issue, the Italian philosopher already mentioned, D'Alterio, referring to Marcuse, wrote: "In this new dimension, individuals are dominated by criteria external to them and are therefore 'other-directed' by the new social order imposed by the politico-economic mechanism, which demands that everything be standardized: it dictates to us what to buy, which film to watch, how to entertain ourselves, and what to think"[18].

Regarding the submission of man to the economic factor, D'Alterio, again referring to Marcuse, brings us a concept from Smith: "Rationality has by now identified itself with organized structures, with machines; thought has become 'instrumental': *homo sapiens* has transformed into *homo economicus*"[19]. For Marcuse, the entire economy is subjected to political control. This is the "State Capitalism" that established itself in the United States and in dictatorial regimes, which led to a monopolistic concentration of enterprises and bound the ruling classes to industrial management, whose sole objective is to serve its own interests, imposing on society false needs and tastes, as well as induced consensus and lifestyles, such as "the need to relax, to have fun, to behave and consume in accordance with advertisements, to love and hate what others love and hate"[20]. Still according to Marcuse, the overbearing dominance of labor productivity has introduced us into the "age of anxiety," where everyone lives "in a state of general anesthesia that makes the individual happy"[21].

Man lacks the time to reason or to question the condition in which he finds himself; hence, he considers it normal to concern himself solely with economic matters: "Reason has become a mere appendage of the all-encompassing economic apparatus"[22]. In this regard, Adorno wrote: "Here among us, there is no difference between economic fate and man himself. No one is anything other than his wealth, his income, his position, his prospects. ... Each person is worth as much as he earns, each person earns as much as he is worth. He learns what he is from the vicissitudes of his economic existence, and knows himself as nothing different"[23]. According to Marcuse, the individual has been forced to employ all psycho-physical energies for working and productive purposes, thereby repressing his human demands for happiness and pleasure[24]. Thus, he is de-eroticized. The philosophers of the Frankfurt School imagine a future man in a different society without human alienation. On this matter, the Italian professor of political

[15] La scuola di Francoforte: Horkheimer, Adorno, Marcuse, Habermas, Fromm: Collana "Compendio del pensiero filosofico occidentale del XX secolo", volume IV p. 31
[16] Same p.51
[17] Sini, C. (2003) *Le arti dinamiche Filozofia e Pedagogia* Milano: CUEM p.194
[18] La scuola di Francoforte: Horkheimer, Adorno, Marcuse, Habermas, Fromm: Collana "Compendio del pensiero filosofico occidentale del XX secolo", volume IV p. 24
[19] Same p. 26
[20] Same p.26
[21] Same p.28
[22] Gennari, M. (2001) *Filosofia della formazione dell'uomo*. Milano: Studi Bompiani p.492
[23] Horkheimer, M & Adorno, Th. W. (2008) *Sulla genesi della stupidità e altri scriti*. Napoli: La scuola di Pitagora editrice p.15
[24] La scuola di Francoforte: Horkheimer, Adorno, Marcuse, Habermas, Fromm: Collana "Compendio del pensiero filosofico occidentale del XX secolo", volume IV p. 28

philosophy at La Sapienza University in Rome, Stefano Petrucciani, writes in his book *La Scuola di Francoforte: Storia e attualità*: "philosophy must orient itself toward the concrete historical realization of the most authentic possibilities of human Dasein, which are denied in the world of capitalist alienation, where man is dominated by the object"[25]. In this way, he retrieves the Marxist thesis that philosophy must not interpret the world as it is, but as it ought to be.

1.4. From Consumption to Waste

As I have mentioned, the critique of the Frankfurt School philosophers did not impact society, where production for its own sake predominates and consumption has increased to the extreme. Regarding production as an end in itself, Stefano Petrucciani, referring to Adorno, writes: "Production as an end in itself and unlimited accumulation translate into the infinite creation of new needs"[26]. There are several factors that drive consumption. Real needs are naturally satisfied according to the buyers' means. But after real needs are satisfied, new needs are created that become indispensable because advertising, which stimulates the desire to consume, along with fashion, mass behavior, etc., exerts its influence. Clothes, furniture, toys, useless gifts, trips, surgical procedures to alter the body, etc. Beyond mass behavior, standardized lifestyles, and everything that surrounds man, even bodies and faces are standardized, flattening differences. All of this feeds production and, conversely, production feeds consumption. On this point, the Italian philosopher Elio Franzini states that man is reduced merely to reacting to external stimuli, dominated by the idea of consumption accompanied by fashion[27].

Radical in his critique of excessive consumption and waste is Jean Baudrillard. Here is how he defines civilization in his book *The Consumer Society*: "The garbage-can civilization and the corresponding garbage-can sociology: tell me what you throw away, and I will tell you who you are"[28]. According to him, "massive and useless expenditure is considered a kind of madness, insanity, a dysfunction of instinct that leads man to burn his reserves and harm his survival conditions through an irrational practice"[29]. We have moved from a consumer society to a waste society, where the motto is: use and throw away. Consumption and waste also concern the food industry, which develops in synchrony with the pharmaceutical industry. Countless diseases are the effect of that industry, which the pharmaceutical industry must then treat. Paradoxically, it may coincide that the very same corporation owns both the industry that causes the diseases and the one that treats them. Man has become a slave to consumption, and on this point Habermas wrote: "Man believes he is obeying his own will, but he is a slave to money, career, leisure time, and consumption"[30].

As we have said, the philosophers of the Frankfurt School do not limit themselves to describing the situation in which man finds himself, but point to a political solution; an alternative model. In this regard, Marcuse envisioned a future society: "it is a society without war, without exploitation, without oppression, without poverty, and without waste"[31]. To achieve such a state, educational institutions must play an important role, particularly philosophy, which must transform human beings and make them conscious. Furthermore, as Marcuse argues, politics must engage men in democratic life: "Let there be a shift toward an ethical refoundation of politics that leads to a social 're-politicization' capable of making people more aware and active in the decisions governing democratic life, transforming them from 'consumer-citizens' and 'hypertrophied' clients into self-determined citizens, free from every form of the pathology of 'modernity' as well as from the distortion and disruption of communication operated by power and ideological domination"[32]. Fromm, on the other hand, thinks of curing man of these pathologies of modernism. In this regard, he wrote: "To build a sane economy for sane people" capable of transforming current patterns of "pathological and undifferentiated consumption" into rational consumption for the health of all, "since consumers' desires will no longer be manufactured by producers"[33]. But Stefano Petrucciani maintains that before

[25] Petrucciani, S. (2023) *La Scuola di Francoforte Storia e attualità.* Roma: Carocci editore p.76
[26] Petrucciani, S. (2007) *Introduzione a Adorno* Roma: Editori Laterza p.96
[27] Franzini, E. (2018) *Moderno e postmodero* Milano: Raffaello Cortina Editore p.86
[28] Baudrillard J. (2020) *Shoqëria e konsumit.* Shkup: Logos-A, p.53
[29] Same p.53
[30] La scuola di Francoforte: Horkheimer, Adorno, Marcuse, Habermas, Fromm: Collana "Compendio del pensiero filosofico occidentale del XX secolo", volume IV p. 38
[31] Same p. 26
[32] Same p. 42
[33] Same p. 52

society can be transformed, the individual must be transformed: "no social transformation is possible without a transformation of individuals[34]".

The critique of the philosophers of this school is directed at capitalism in full development in Western countries, but I would add that Eastern countries also display the same "symptoms": consumption, waste, mass behavior predominate, and even more so corruption, especially in the Balkan countries. Corruption feeds authoritarianism and even the arrogance of those who hold economic and political power, which often coincide. An uncontrollable frenzy to accumulate wealth by any means prevails. A critique of exaggerated standards of living and universal consumption was advanced by Christopher Lasch. The Italian philosopher Mario Gennari brings us a passage where he writes: "The idea of universal consumption exported through the expansion of the market across the entire globe appears futile and shortsighted. The ecosystem gradually collapses under the unlimited spread of a progress that increasingly resembles a dangerous regression"[35]. As I have mentioned, contemporary man is in a situation that demands change.

1.5. The Courage to Be Sober

The capitalist economy functions on the basis of optimism and confidence in growth. Production and consumption sustain this economy. The more that is produced, the more that is consumed, and vice versa. All of this, in addition to the positive aspects it entails, has produced the alienation of man, making him a sort of modern slave. Thus, reflective thought on the entire system has never been lacking. Capitalism can be compared to the mechanism of a bicycle, which loses its balance if it stops. But since uncontrolled speed involves risks, bicycles also have brakes. For all the reasons listed by the philosophers of the Frankfurt School and other critics, one can attempt to change one's lifestyle, with the goal of making man a bit freer, not in a constant hurry, and with more time to reflect. A sort of solution could be a more restrained lifestyle regarding the exaggerated consumption that characterizes modern society; thus, a more sober life. The widespread adoption of a sober life in society could have effects that might entail stagnation and even economic recession; precisely for this reason, I have called it a courageous thesis. But on the other hand, unchecked development also carries risks that could lead, for example, to crises similar to those of the past, which had dramatic effects. The thesis of applying a brake to development is not new. Friedrich Nietzsche, in a section titled *A Brake for Civilization* in his book *Human, All Too Human*, wrote: "sometimes a brake is necessary for civilization: when it runs too fast downhill, or perhaps, as in this case, uphill"[36].

Regarding metaphors, even more surprising metaphors than that of the bicycle have been made. Ivan Illich proposed imitating the snail. The snail stops the growth of the number of coils on its shell in time, because one more coil would progressively increase its size and thus make it difficult to move. According to him, society too must limit its activity. Illich calls for frugality and the limitation of means of transport: "For me, the alternative for society must be found in the conscious limitation of technology to those services that are truly beneficial. I mean the limitation of vehicle speeds"[37]. It seems like a paradoxical idea, but it is no more paradoxical than birth control, which has been applied in some countries. Indeed, philosophy has taught us sobriety since antiquity: Socrates with his famous exclamation at the Athenian market: "How many things I do not need!"; and then the famous Cynics and many other philosophers. Indeed, in life everything could be simpler: clothes, food, houses, furniture, cars, cities. Decor in cities is exaggerated, especially for celebrations. Night has been swapped with day. All of this carries an enormous cost, not only energetic but also in human effort. To produce the superfluous, men waste the time necessary to enjoy and to reflect.

2. CONCLUSION

As I have mentioned, the philosophers of the Frankfurt School do not limit themselves merely to describing in a pessimistic manner the condition of man in capitalist society, but also propose alternatives, certainly not easy to realize, but not impossible. In addition to a more sober lifestyle, of which I have spoken, which depends primarily on individual choice, the alternative could be the liberation of the labor market, which depends on other factors, especially economic ones. Society has reached this state of development thanks to work, cooperation, and the advancement of science and technology; but regarding work and the use of science and technology, everything can

[34] Petrucciani, S. (2023) *La Scuola di Francoforte Storia e attualità* Roma: Carocci editore p.98
[35] Gennari, M. (2001) *Filosofia della formazione dell'uomo*. Milano: Studi Bompiani p.458
[36] Nietzsche F. tra.it. (1990) *Umano, troppo umano* Roma: Grandi Tascabili Economici Newton p.402
[37] I.B.E-UNESKO *Thinkers on education*, perkth. Alb. *Mendimtarët për edukimin* 2008. p. 43

also function differently. Throughout history, man has not been accustomed to working at the same pace in all seasons of the year, but the development of industry has alienated and transformed the worker almost into a mechanism. Work can also be organized in a different way. One can work by the hour, and hours can be flexibly distributed over fewer days of the week, month, or year; one could even work 16 hours and not work the following day; one can work from home, which is already being applied, etc. All of this can also contribute to reducing traffic. Naturally, one must overcome the fear of the consequences that might arise.

The reason for this proposal is human well-being, which is not necessarily linked to an abundance of material goods. Many philosophers have said this, from Socrates in antiquity to Adorno in the last century, and many others today. To live in a simpler manner, to work less and in a different way in order to recover time, the time for reflection, to understand what man truly is.

References

[1] https://www.academia.edu/54072395/Marina_Calloni_La_divergente_unit%C3%A0_della_cosiddetta_Scuola_di_Francoforte_in_Quaderni_di_Teoria_Sociale_1_2_2020_pp_197_214
[2] La scuola di Francoforte: Horkheimer, Adorno, Marcuse, Habermas, Fromm: Collana "Compendio del pensiero filosofico occidentale del XX secolo", volume IV
[3] Petrucciani, S. (2023) *La Scuola di Francoforte Storia e attualità.* Roma: Carocci editore.
[4] Baldassarre, L. (2021) *La scuola di Francoforte. Una introduzione.* Firenze: Editrice Clinamen
[5] https://scienzaepolitica.unibo.it/article/download/16377/15560 vol. XXXIV, no. 67, 2022, pp. 55-74
[6] Jay, M. (2025) *Critiche immanenti. La scuola di francoforte sotto pressione*. A cura di Roberto Mordacci. Traduzione di Cecilia Vergani. Milano: Mimesis
[7] https://journals.openedition.org/qds/792 pp. 229-246
[8] https://www.researchgate.net/publication/355889890_IL_MARXISMO_FRANCOFORTESE_DI_PIER_PAOLO_PASOLINI
[9] Sini, C. (2003) Le arti dinamiche Filozofia e Pedagogia Milano: CUEM
[10] Gennari, M. (2001) *Filosofia della formazione dell'uomo*. Milano: Studi Bompiani
[11] Horkheimer, M & Adorno, Th. W. (2008) *Sulla genesi della stupidità e altri scriti*. Napoli: La scuola di Pitagora editrice
[12] Petrucciani, S. (2007) *Introduzione a Adorno* Roma: Editori Laterza
[13] Franzini, E. (2018) *Moderno e postmodero* Milano: Raffaello Cortina Editore
[14] Baudrillard J. (2020) *Shoqëria e konsumit.* Shkup: Logos-A,
[15] Nietzsche F. tra.it. (1990) *Umano, troppo umano* Roma: Grandi Tascabili Economici Newton
[16] I.B.E-UNESKO (2008) *Thinkers on education*, perkth. Alb. *Mendimtarët për edukimin*. Tiranë: Plejad

Matching Governance to Cross-Government Problems: A Framework for Center of Government Involvement

Elina PETROVSKA[a*]

[a] *University of Latvia, Faculty of Economics and Social Sciences, Department of Management and Business, Aspazijas blvd. 5, Riga, Latvia, petrovska.research@inbox.lv*

Abstract

Cross-government problems are often met with calls for stronger coordination, but this offers little guidance on what kind of form of coordination is needed. Drawing on structural contingency theory, this paper asks how governance mechanisms can be suited to different cross-government problems and when center of government (CoG) involvement adds value. The analysis is based on a theory guided integrative synthesis of four literature streams: mission-oriented innovation, collaborative governance and mediation, expert advice and knowledge integration, and CoG organization. It identifies three recurring governance needs: recurrent portfolio choices among multiple initiatives or solution pathways; negotiable disagreement among interdependent actors who still need a joint decision; and a specific knowledge gap that requires specialized experience. These needs are linked to three governance archetypes. The paper then separates the choice of mechanism from the choice of institutional locus. A mechanism may fit the problem without needing to be owned by the CoG; central involvement is warranted only where the CoG contributes a capability that cannot readily be provided elsewhere. Both mechanism and locus are treated as contingent and open to revision as conditions change. The framework developed in the paper helps clarify which governance arrangements can be matched to different cross-government problems and when CoG involvement adds a distinct institutional advantage.

Keywords: Structural Contingency Theory, Centre of Government, Governance Mechanisms, Institutional Design

1. INTRODUCTION

Cross-cutting policy problems do not fit easily within administrative systems organized around ministries and sectoral responsibilities. Alessandro *et al.* note that an increasing number of such issues require whole-of-government responses and describe the center of government (CoG) as supporting the chief executive through political and technical coordination, strategic planning, performance monitoring, and communication [1, p. 56]. OECD also identifies strategic direction, prioritization and alignment across ministries as important CoG functions [20, p. 56]. At the same time, it warns that coordination arrangements can multiply, duplicate work, blur responsibilities and place additional demands on central capacity [20, p. 56]. The implication is not simply that governments need more coordination, but that coordination arrangements need to be fit for purpose.

* Corresponding author

The difficulty is that cross-government problems are not alike. Research on mission-oriented innovation, for example, emphasizes experimentation, learning and exploring multiple solution paths when solutions are uncertain, rather than prescribing a single route in advance [17, pp. 807, 810-812; 28, pp. 479, 485]. Collaborative governance research addresses a different kind of problem: actors may depend on one another while holding different interests, values or political positions and still need to reach a joint decision [2, pp. 159, 165]. Expert advice research, in turn, addresses situations in which the available research remains insufficient for timely, informed decision-making and expert judgement can provide additional support [19, p. 7176]. These literatures address different governance needs, even though all of them may appear in practice as a need for "better coordination".

A cross-government problem also does not determine where the response should be located. Comparative CoG research shows considerable institutional variation. James and Ben-Gera argue that there is no single correct way to organize the center and caution that institutional models must be adapted to the constitutional, political and administrative setting in which they operate [14, p. 5]. Brown *et al.* assess CoG tools partly in terms of what they call the "comparative advantage of central deployment" [p. 10] – that is, the extent to which a tool is better suited to use by the center than by line ministries [7, pp. 10-11]. The design question is therefore not only which response is appropriate, but also whether the CoG adds something that cannot be provided adequately elsewhere.

Structural contingency theory offers a way to connect these questions. Donaldson argues that effective organizational design depends on fit between organizational structure and the contingencies it faces, rather than on any single structure being universally effective [9, p. 19]. Fit is also dynamic: changes in contingencies can create misfit and prompt structural adaptation to regain fit [9, pp. 20-21]. The paper applies this logic to cross-government governance without assuming that structural contingency theory itself specifies the mechanisms or CoG arrangements examined here.

Against this background, the study asks: *How can structural contingency theory inform the selection, design, and institutional location of governance mechanisms for cross-government problems, and under what conditions does CoG involvement provide institutional advantage?*

In order to answer the question, the paper brings together several literature streams. It uses them to distinguish the conditions that create different governance needs, develop three governance archetypes, and then consider where those mechanisms should sit within government. The CoG is treated as a possible institutional locus rather than as a mechanism in its own right. This leads to a distinction between mechanism fit – whether an arrangement addresses the governance problem – and locus fit – whether that arrangement is appropriately situated and supported. Both can change as the underlying conditions change.

2. THEORY

2.1. Structural Contingency Theory and Institutional Fit

Structural contingency theory provides the starting point for the analysis. Building on Donaldson's fit logic, the analysis treats the appropriateness of organizational structure as contingent on the circumstances faced by the organization [9, pp. 19, 37].

Fit is not necessarily permanent. Donaldson's Structural Adaptation to Regain Fit (SARFIT) model treats fit and misfit as temporary states that can alternate over time [9, pp. 20-21]. Donaldson also discusses the concept of quasi-fit, in which an organization moves towards fit without necessarily achieving a full fit [9, p. 23].

The full SARFIT model is not applied here to the government. In particular, this study does not assume that Donaldson's [9] organizational growth sequence applies directly to public sector coordination. For this paper, the more limited implication is that institutional arrangements may need to be reconsidered when relevant conditions change, rather than treated as fixed once they have been established.

Recent organization design research extends this reasoning to coordination. Joseph and Sengul define organization design broadly in terms of "formal and informal organizational attributes, including structures, processes and systems" [p. 249], and review contemporary research through four major approaches: configuration, control, channelization, and coordination [15, pp. 249, 252]. Their coordination perspective is especially relevant because it treats organizations as systems whose decisions and efforts must be coordinated and integrated through organizational design [15, p. 252].

For this paper, the value of these works lies in the questions they make possible. Neither Donaldson [9] nor Joseph and Sengul [15] specifies the three governance mechanisms developed below or prescribes a particular CoG arrangement. They provide the underlying fit and coordination logic. The properties of the mechanisms themselves are drawn from the specialized literatures discussed in the next sections.

2.2. From Governance Problems to Governance Mechanisms

Choosing a governance mechanism requires more than labelling a problem as complex. Organization design research indicates that the appropriate means of coordinating interdependent work also depend on characteristics of the knowledge involved. Joseph and Sengul note, based on prior research, that greater knowledge separability can allow work to be divided more vertically [15, p. 256], while shared understanding can reduce explicit communication demands and facilitate coordination under complex interdependence [15, p. 277].

Evidence from public organizations points in the same direction. Comparing two U.S. federal agencies, Lee and Rhee found that the effects of inter-unit collaboration differed markedly between organizational settings [16]. The authors link this difference partly to task characteristics: in the science research center, highly specialized technical work was organized into relatively compartmentalized tasks, whereas the law enforcement agency faced less programmable and emergency-driven work that required more real-time coordination across units [16, p. 12]. The comparison matters here because technical complexity alone did not determine the need for stronger horizontal coordination. How the work was divided and how strongly units depended on one another also mattered.

Uncertainty adds another distinction. Wanzenböck *et al.* analyze societal challenges in terms of complexity, contestation, and uncertainty on both the problem and solution sides [28, pp. 476-478]. They argue against applying a uniform approach to mission-oriented innovation policy, as different problem-solution constellations may require different policy strategies [28, p. 481]. Uncertainty is therefore relevant to diagnosis, but it does not by itself identify a governance mechanism. The question then becomes how such uncertainty should be handled in decision-making. Morgan addresses one part of this problem through expert elicitation [19]. He argues that expert elicitation should draw on the best available evidence and analysis and be used only where the existing knowledge does not provide an adequate basis for timely and informed assessment and decision-making [19, p. 7176]. This gives expert-based arrangements a clear boundary: the issue is a remaining epistemic deficit, not simply the technical difficulty of the subject.

A different problem arises when actors depend on one another but disagree about interests, values or preferred outcomes rather than insufficient knowledge. Ansell *et al.*, in their discussion of political negotiation, draw on earlier literature to identify bargaining, alliance formation and compromise as practices involved in reaching joint decisions despite such differences [2, pp. 165-166]. This problem is distinct from one in which the main difficulty is insufficient knowledge.

The reviewed literatures do not provide a single common classification of cross-government problems. The classification used here is therefore a synthesis. Epistemic uncertainty and temporal horizon or recurrence are treated as cross-cutting characteristics, while task decomposability serves as a boundary condition. The main distinction is the form of interdependence that has to be managed: among initiatives and resources, among actors with different preferences and interests, or among holders of specialized knowledge.

From these distinctions, the paper develops three governance archetypes: *Strategic Innovation Portfolio Governance, Mediated Coordination,* and *a Project-specific Expert Working Group*. They are analytical constructs rather than categories taken from any one of the reviewed literatures. Their relationship to the diagnostic conditions is set out in the Findings section.

2.3. Center of Government as a Contingent Institutional Locus

The question of institutional location begins with a distinction between functions and organizations. Dunleavy and Rhodes, writing about the core executive, use a functional definition that includes organizations and structures responsible for integrating central government policies or resolving conflicts within the executive [10, p. 4]. Alessandro *et al*. make a related point in the CoG literature: the functions performed may be more informative than formal structural location alone [1, p. 10].

That distinction matters because the CoG can perform several roles. As noted earlier, Alessandro *et al.* identify a number of functions performed by the CoG [1]. OECD notes that CoGs can use “fit-for-purpose mechanisms such as taskforces, working groups, advisory bodies and expert groups”, and stresses that these should have “clear purposes, processes, timeframes and objectives” [20, p. 56]. This variation is also reflected in the institutional design of CoGs. OECD defines the CoG as a “body, or group of bodies” [20, p. 10]. It also emphasizes that the institutional configuration of CoGs differs substantially across countries: in some cases, relevant bodies are concentrated within one institution, while in other responsibilities are distributed across several institutions or units reporting directly to the highest level of the executive [20, p. 10]. Although CoGs differ in their institutional structure, they perform several

common functions [20, p. 10]. This suggests that CoG functions and the mechanisms used to perform them do not correspond to a single institutional model.

The paper therefore keeps two concepts separate. A governance mechanism is the way a particular governance problem is addressed. Institutional locus is the organizational level and setting in which that mechanism is located, sponsored, coordinated or governed. The CoG is treated as one possible locus, not as a fourth governance mechanism. Whether the CoG should be involved is considered separately from the question of which mechanism fits the problem.

Latvia provides a useful illustration of a distributed central setting. Petrovska and Berzins identify several institutions as forming part of the Latvian CoG [21]. In their analysis, the State Chancellery is associated with national development planning and policy alignment, the Ministry of Justice with legal standards, and the Ministry of Finance with financial policy and oversight of state-budget implementation [21, p. 112]. The example is used only to show that central functions can be distributed across institutions; it is not presented as a model for other administrative systems.

Mechanism selection and locus selection are therefore treated as separate design questions. The first asks what governance capacity the problem requires. The second asks where that capacity can be provided most appropriately. The conditions under which CoG involvement may add value are examined later in the analysis.

3. METHODS

3.1. Research Design

This is a conceptual study based on a theory-guided integrative synthesis of literature and framework development. It does not seek to estimate an aggregate empirical effect or provide exhaustive coverage of a single field. Instead, it brings together several bodies of literature that address different aspects of cross-government coordination and uses them to develop a common institutional-design framework.

This approach is consistent with established methodological work on literature-based and conceptual research. Snyder treats literature review as a research methodology and distinguishes systematic, semi-systematic and integrative approaches, noting that different review designs serve different purposes [25]. Torraco views the integrative literature review as a way of critically combining existing research to produce new insights into a topic [27]. Jaakkola emphasizes that conceptual research should follow a clearly articulated research design [13, p. 18]. She also argues that authors need to explain why particular theories are selected and clarify how those theories contribute to the analysis [13, p. 18].

Structural contingency theory is used as the analytical lens. It helps examine whether different problem conditions call for different institutional responses, but it is not treated as the source of the three mechanisms developed in the paper. Those mechanisms emerge from comparison of the specialized literatures and are then brought together within a common contingency-based framework.

The synthesis draws on four bodies of literature: mission-oriented innovation and portfolio governance; collaborative governance, negotiation and mediation; expert advice, temporary expert arrangements and knowledge integration; and CoG organization and governance. The first three are used to examine when different mechanisms may be needed and what capacities they provide. The CoG literature is considered separately because it addresses a different question: where a selected mechanism should be located or supported.

3.2. Literature Selection

The literature was assembled through an iterative and purposive search rather than a systematic review protocol. For the three mechanism-specific streams, relevant publications were identified through searches in ScienceDirect and Google Scholar. Searches were refined as the conceptual framework developed and new analytical questions arose.

The search developed alongside the analysis. Initial searches covered mission-oriented innovation and portfolio governance; collaborative governance, negotiation, mediation and conflict management; and expert advice, expert groups, temporary organization and knowledge integration. Later searches addressed issues that became important during comparison, including task interdependence and decomposability, organizational learning, knowledge transfer and institutional memory.

The CoG literature was assembled partly through a different route. It drew on a body of CoG literature previously compiled for the author's ongoing doctoral research and re-examined for the research question of this paper. More

recent publications were added where they provided relevant evidence on CoG definitions and functions, institutional arrangements, coordination, strategic management, monitoring, and relations between the CoG and line ministries.

Sources were retained when they helped answer the analytical questions used in the study: when an additional governance response may be needed; what capacity a mechanism provides; how mandate and authority are defined; who participates; how the process works; what output is expected; and what safeguards, adaptation or termination arrangements are relevant. The aim was to assemble a sufficiently varied basis for comparison across the four literature streams, not to achieve exhaustive coverage of any one of them.

3.3. Analytical and Synthesis Procedure

The three mechanism-specific streams were examined with the same broad set of questions. The analysis considered when a mechanism becomes relevant, what governance capacity it provides, how its mandate and authority are defined, who participates, how the process operates, what output it is expected to produce, and what safeguards or arrangements for adaptation and termination are required. Using the same questions across streams made comparison possible despite their different disciplinary origins.

The analytical framework changed during that comparison. An initial set of seven candidate variables was reconsidered because the variables did not operate at the same analytical level. The revised framework treats temporal horizon or recurrence and epistemic uncertainty as cross-cutting characteristics, the dominant form of interdependence as the main differentiator between mechanisms, and task decomposability as a boundary condition. Mechanism-specific conditions – recurrent portfolio choice, the need for a joint decision, and the integration of specialized knowledge – are considered separately.

The CoG literature was handled in a separate analytical step. It was used to examine the institutional setting in which a mechanism might operate, not to identify another governance mechanism. This separation made it possible to ask two different questions: which mechanism fits the governance problem, and where should that mechanism be located or supported?

During the analysis, substantive claims were checked against the underlying sources. Evidence reported directly by individual publications was kept distinct from conclusions drawn across several sources and from propositions developed in this paper. Claims that could not be supported were removed or reformulated. The final framework links problem diagnosis required governance capacity, mechanism choice, institutional locus and later reassessment as conditions change.

4. FINDINGS

4.1. Diagnostic Conditions for Mechanism Selection

The synthesis did not support treating all candidate conditions as equivalent. Instead, they fall into a sequence: a preliminary gate, cross-cutting characteristics, a boundary condition, the dominant form of interdependence, and a final mechanism-specific test (see Table 1). This ordering is an outcome of the synthesis rather than a decision rule taken from any one source.

Table 1. Diagnostic architecture for governance mechanism selection

Diagnostic stage	Core question	Role in the framework
Preliminary gate	Can existing administrative responsibilities, specialization, and ordinary coordination handle the problem adequately?	Determines whether an additional mechanism is needed
Cross-cutting diagnosis	What is the level of epistemic uncertainty, and what is the temporal horizon or recurrence of governance need?	Characterizes the problem context
Boundary condition	Can the relevant work be decomposed into sufficiently independent specialist tasks?	Tests on whether existing specialization may remain adequate
Dominant interdependence	What primarily needs to be integrated or reconciled:	Main mechanism differentiator

	initiatives/resources, preferences/interests, or specialized knowledge?	
Mechanism-specific test	Are the defining conditions of a particular mechanism present?	Determines the mechanism indicated by the framework

The first question is whether existing administrative responsibilities, specialization and ordinary coordination can handle the problem adequately. If they cannot, epistemic uncertainty and temporal horizon or recurrence are considered as cross-cutting characteristics, with task decomposability used as a boundary condition. These features help describe the problem and test whether existing specialization may still be sufficient; none of them selects a mechanism on its own.

The main distinction comes from dominant interdependence. From a coordination perspective, organization design addresses the coordination problems created by the division of labor among interdependent actors and the design arrangements used to manage those problems [15, p. 272]. Building on this logic, the synthesis developed here distinguishes three cases according to what primarily has to be integrated or reconciled: initiatives and resources, preferences and interests, or specialized knowledge. This three-part distinction is developed in this paper rather than adopted from Joseph and Sengul [15].

The final step asks whether the defining condition of a particular mechanism is present. Strategic Innovation Portfolio Governance is relevant where multiple initiatives or solution pathways require recurrent portfolio-level prioritization and resource choice; Mazzucato provides the underlying portfolio-based experimentation and learning logic [17], while Smith *et al.* show why innovation governance may need to extend beyond individual projects [24]. In this study, Mediated Coordination refers to situations in which interdependent actors have different interests but still need to negotiate and reach a collective decision. This understanding draws on Ansell *et al.*, particularly their discussion of political negotiation in collaborative governance [2, pp. 165-166]. A Project-specific Expert Working Group applies where a defined knowledge gap requires specialized expertise to be incorporated into an advisory output; Christensen and Hesstvedt provide a related institutional perspective through their analysis of temporary expert groups [8]. The way these conditions are combined into selection tests is the synthesis developed here.

4.2. Strategic Innovation Portfolio Governance

Strategic Innovation Portfolio Governance is concerned with governing a set of initiatives over time rather than managing each project in isolation. Mazzucato provides a key conceptual basis for this logic [17]. She links mission-oriented innovation to learning by doing, trial and error, progress assessment and adaptation, and argues for a portfolio approach that can accommodate different possible solutions [17, pp. 810-812]. Taken together, these elements support a standing capacity to compare and review initiatives as a portfolio. The concept of Strategic Innovation Portfolio Governance is the synthesis developed in this paper, not an institutional approach proposed by Mazzucato [17].

Smith *et al.* add an important institutional dimension [24]. Their analysis of responsible innovation in research funding argues that project-centered governance can leave broader institutional conditions largely unexamined, and it identifies several "inflection points" for intervention: agendas, spaces, funding calls, and evaluation [24, pp. 390-391]. Examples of existing governance instruments at these points include funding calls, evaluation panels, reporting requirements, portfolio analyses, and strategic visions and delivery plans [24, p. 391]. This suggests that portfolio governance concerns not only periodic project assessment but also how strategic direction is set, how initiatives enter the portfolio, and how their development is subsequently reviewed.

Portfolio discretion also creates a need for safeguards. Meijer and Thaens argue that public innovation can produce adverse effects associated with low public control and low public value and propose dynamic assessment as a response [18]. Their approach builds recurring reflection into the innovation process, considers both desired public value outcomes and possible perverse effects, and allows assessment to lead to further requirements or, where adverse effects dominate, termination [18, p. 149]. This provides a basis for recurrent review and for treating continuation as a decision rather than an automatic consequence of earlier selection.

Authority and accountability matter as well. Bommert, discussing collaborative public sector innovation, argues that sharing decision rights with actors outside government raises questions about the distribution of accountability [5, pp. 28-29]. Bommert emphasizes the need to connect decision rights with accountability in public sector innovation through collaboration [5, pp. 28–29]. He also discusses the risk that participating interests may influence decision-making in ways that undermine public value objectives [5, pp. 26-27]. For the archetype developed here, this supports

clear decision rights, accountability arrangements and criteria for reviewing portfolio decisions. These safeguards are drawn together from literature rather than taken from a single portfolio governance approach.

On this basis, Strategic Innovation Portfolio Governance is defined here as a standing governance capacity that combines strategic direction, portfolio-level decision-making, experimentation and learning with recurrent assessment and accountability. Its purpose is not to manage individual projects directly, but to govern choices and relationships across an evolving portfolio. Its precise authority, composition, resource powers and institutional location remain contingent design choices.

4.3. Mediated Coordination

Mediated Coordination presents a different design problem. It needs to help interdependent actors work through disagreement without transferring the substantive decision to the mediator. Ansell and Gash describe the mediator or broker as a collaborative leadership role that helps manage conflict, facilitates constructive exchanges between stakeholders, mediates between different positions, and supports trust-building [3, p. 11]. Mediation is therefore treated here as a process capacity: it structures negotiation while leaving responsibility for the substantive outcome with the actors themselves.

The distinction also suggests that effective mediation need not depend on complete organizational detachment. Ansell and Gash distinguish professional facilitators, whose legitimacy derives from neutrality and independence, from "organic" leaders who come from within the stakeholder community and draw on local knowledge, relationships, and social capital [3, pp. 6, 18]. Both may act as honest brokers, although their strengths and weaknesses differ [3, p. 6]. For this paper, the relevant requirement is credible procedural impartiality or honest broker legitimacy rather than complete organizational detachment. This formulation is a synthesis, not a threshold specified by Ansell and Gash [3].

The process itself matters as much as the identity of the mediator. Bianchi *et al.* highlight a distinction in the collaborative governance literature between formal organizational factors, such as rules, structures, roles and protocols, and informal factors, including commitment, facilitative leadership, shared understanding, and trust [4, p. 1583]. They also note that carefully structured collaborative initiatives may fail when suitable models and methods are lacking to support leadership, manage conflict and build trust [4, p. 1583]. This supports a mediated process with a defined mandate, clear roles and ground rules, together with active facilitation and conflict management. The exact institutional form remains open to design.

Mediation also has limits. Torfing *et al.*, in their systematic review of conflict within collaborative governance, find that collaborative arrangements have capacities for dealing with more limited disagreements "through negotiation, brokering and consensus formation", while the "literature has less to say about how fundamental conflicts" rooted in different basic beliefs and world views can be resolved [26, p. 14]. They also note that deeply entrenched ideological differences may go beyond what consensus building or negotiation can readily resolve [26, p. 16].

That boundary should not be drawn too sharply, however. Building on political negotiation literature, Ansell *et al.* argue that political conflicts rooted in fundamental differences can sometimes be handled constructively through negotiation dynamics, while stressing that the effects of these dynamics are shaped by the type of conflict, the participating actors and the setting in which collaboration occurs [2, pp. 165-166]. They also note that further research is needed to determine the precise impact of these conditions [2, p. 166]. Deep or rigid conflict is therefore treated here as a boundary condition rather than an automatic reason to exclude mediation. Mediation remains plausible where there is still a credible basis for negotiation or compromise.

Taken together, these points define Mediated Coordination as an institutionalized mediation and brokerage capacity for actors who remain interdependent despite substantive disagreement. Its design rests on a credible honest-broker role, clear process rules, facilitative leadership, conflict management and enough room for negotiation. The mediator supports the process without taking over the substantive decision. Questions of organizational location, formality and the relationship to the CoG are dealt with separately as locus questions.

4.4. Project-specific Expert Working Group

The expert group archetype raises another design problem: how to assemble, assess and transfer expert knowledge into the permanent decision-making structure. It is treated here as a temporary knowledge-integration arrangement, not as an additional decision-making authority.

Morgan [19] provides an important starting point for considering how expert judgement should be handled within a Project-specific Expert Working Group. Individual assessments may reveal a wider range of uncertainty than

consensus processes, and Morgan suggests that they can usefully inform later group deliberation [19, p. 7181]. He also cautions against automatically combining divergent expert judgements, arguing that where experts differ substantially in their views of the underlying science, separate analyses may be more appropriate [19, p. 7182]. This supports a process that keeps uncertainty and materially different expert judgements visible rather than treating consensus as an end in itself.

Christensen and Hesstvedt [8] provide a useful institutional analogue for the temporary organizational form. Their study of Norwegian Official Commissions describes bodies appointed by government, functioning outside the permanent administrative structure and working under terms of reference that specify the questions to be addressed and the advice sought [8, p. 1269]. The commissions have formally appointed members and secretariat support, with membership spanning public administration and external organizations [8, p. 1269]; they have an average duration of about two years and conclude their work with a public report setting out their findings and policy advice [8, p. 1270]. These features show how a temporary expert arrangement can have a bounded mandate and a defined output. They are used here as an analogue, not as a universal model for expert working groups.

Composition also deserves attention. Eriksson questions the idea that expert advice is a neutral part of decision-making and discusses how knowledge production can be influenced by the questions selected, disciplinary purposes, scientific values and assumptions [11, pp. 1-2]. She also discusses research suggesting that representational diversity may contribute to knowledge attainment [11, p. 2]. In her empirical study of Swedish governmental policy investigations, academic experts were less demographically representative than several comparison groups, including the academic community from which they were drawn, particularly with respect to social background and first-generation immigrant status [11, pp. 8-10]. These findings do not demonstrate that demographic diversity automatically produces better expert advice. However, Eriksson's discussion highlights the relevance of considering which social and disciplinary perspectives are represented in expert groups [11, pp. 2-3].

A temporary expert body needs to be clearly connected to the permanent organization that is expected to take its work forward. Drawing on earlier work on organizational learning and knowledge creation, Brix distinguishes knowledge creation at the individual and team levels from organizational learning [6, pp. 114–115]. He then develops this distinction further by arguing that organizational learning should involve a formal organizational-level decision to use, rework, or reject the newly created knowledge [6, p. 117]. In their systematic review, Rashman *et al.* identify knowledge sharing, transfer, diffusion, application and embedding as important processes in organizational and inter-organizational learning, including learning between organizations [22, pp. 475-478].

This temporary-permanent interface is especially important in public-sector settings. Godenhjelm *et al.*, drawing on earlier research, highlight the challenges of transferring project knowledge and results to permanent organizations and the need for context-sensitive mechanisms that link temporary projects with permanent organizational structures [12, pp. 8, 19-20]. For the expert group archetype, this supports identifying the intended recipient of the group's output before the group finishes its work, rather than assuming that its findings will enter organizational practice automatically.

The Project-specific Expert Working Group is therefore defined as a temporary, task-specific mechanism with a clear mandate, task-relevant composition, appropriate analytical support, bounded advisory authority, a specified output, and an identifiable permanent recipient. The proposed lifecycle is simple: the group produces its deliverable, transfers it to the designated recipient, the recipient decides whether to use, revise or reject the knowledge, and the group is dissolved unless a new mandate is explicitly established. This sequence is a design synthesis developed in the paper rather than a procedure prescribed by any single source.

4.5. Central Comparative Advantage and Institutional Locus

Once a mechanism has been selected, a separate question remains: where should it be located or supported? The distinction introduced in Section 2.3 is treated here as a second fit problem. A mechanism may suit a cross-government problem without needing to be owned or hosted by the CoG.

The division of responsibility between the CoG and line ministries is another important issue. OECD stresses the need to balance the CoG's stewardship role with the responsibility of line ministries for driving policy, especially where implementation and coordination can overlap [20, p. 56]. Clear roles and mandates are important for maintaining this balance and for preserving ministerial responsibility for policy development and implementation.

Building on this distinction, the concept of comparative advantage of central deployment is used here as an analytical starting point rather than as a validated measurement instrument. Brown *et al.* note that their assessments were not based on a formal methodology but rather drew on practitioner experience gained across several governments [7, p. 11]. The literature also points to limits on central intervention. In their discussion of central delivery units, the

authors emphasize that these units should assist rather than replace line ministries, and that responsibility for day-to-day implementation should remain with the relevant ministries [7, pp. 23-24]. Shostak *et al. l*ikewise emphasize selectivity, arguing that CoG attention should focus on identified gaps and on issues within government priorities that genuinely merit central involvement, or "priorities within the priorities" [23, p. 22]. They also stress the need to take account of available resources and to avoid overload and overstretch [23, p. 22].

On this basis, central comparative advantage is assessed through the four dimensions presented in Table 2. These dimensions are synthesized from the literature for the purposes of this paper and do not constitute a validated scale or a decision rule proposed by Brown *et al.* [7], OECD [20], or Shostak *et al.* [23].

Table 2. Central comparative advantage test for institutional locus

Dimension	Diagnostic question	Condition favoring CoG involvement
Cross-ministerial interdependence	Can a single ministry or agency adequately manage the relevant interdependencies?	Effective action requires coordination across several ministries or agencies that cannot be provided adequately from one sectoral locus
Strategic and whole-of-government salience	Does the problem require alignment of government-wide priorities, policies or resources and merit central attention?	The issue has cross-government strategic significance beyond the mandate of an individual ministry or agency
Executive-level integration	Would higher level convening, integration or decision support add a capacity unavailable at sectoral level?	Effective coordination requires access to a cross-government convening or integrating capacity
Net institutional value	Does central involvement add sufficient value to justify its organizational demands and risks?	Expected central added value outweighs duplication, capacity and overload risks

Where central comparative advantage is limited, the mechanism can remain within a ministry, agency or another suitable institutional setting. Where the test identifies a distinct contribution from the center, some form of CoG involvement can be considered. This is a decision about locus, not the selection of a fourth governance mechanism. The next section considers how intensive that involvement needs to be.

4.6. Forms and Intensity of Center of Government Involvement

A comparative advantage at the center does not necessarily mean that the CoG should own the mechanism. The literature points to several possible roles for the CoG, including guidance, convening, facilitation, dedicated support and, in more limited cases, deeper involvement in coordination or delivery. This range of roles suggests that the degree of CoG involvement need not be uniform across coordination mechanisms.

On this basis, CoG involvement can be represented as a graduated continuum: no additional involvement; sponsorship or mandate; convening or brokering; hosting or administrative support; and ownership or leadership. The continuum is a conceptual synthesis rather than an existing CoG typology. It distinguishes the capacity supplied by the center from the degree of institutional responsibility it assumes for the mechanism. A mechanism may require political authorization or central convening without being permanently located at the CoG.

This suggests a design principle: where a distinct central comparative advantage has been identified, CoG involvement should be matched to the level needed to realize that advantage. This is consistent with the literature's emphasis on selectivity. Shostak *et al.* note that some instruments initially developed at the CoG later may be delegated to ministries or agencies as capabilities mature [23, pp. 140-141]. Brown *et al.* likewise note that central units may be discontinued, or their responsibilities transferred to the relevant line ministries, once the rationale for maintaining them has diminished [7, p. 24].

4.7. Integrated Conceptual Framework and Propositions

The findings point to two different kinds of fit. Mechanism fit asks whether a governance arrangement addresses the problem that has to be managed. Locus fit asks whether that arrangement gains a meaningful advantage from its institutional setting. The distinction is developed in this paper by applying Donaldson's general contingency logic to the present governance setting [9, pp. 19, 37].

The framework brings together the staged diagnosis, the design of the three mechanisms, and the subsequent locus test. Joseph and Sengul indicate that greater interdependence can increase information processing demands, potentially creating a need for greater coordination capacity [15, pp. 251], while Brown *et al.* locate a corresponding coordination role at the CoG when policy problems depend on coordinated action across multiple ministries [7, p. 6]. The categories and tests that connect these literatures – the three forms of interdependence, their link to the archetypes, the central comparative advantage test, and the continuum of CoG involvement – are products of the synthesis rather than established scales or taxonomies.

The developed framework treats fit as something that requires reassessment over time. Donaldson distinguishes between environmental contingencies that are outside managerial control and internal contingencies that managers can influence, while OECD and Shostak *et al.* emphasize feedback, ongoing adjustment and adapting CoG instruments to specific circumstances [9, p. 26; 20; 23]. In this framework, reassessment may lead to modifying the mechanism, changing the level or locus of CoG involvement, or ending an arrangement once it is no longer needed.

The synthesis leads to four propositions:

1. *Mechanism-fit proposition.* Where ordinary administrative arrangements are insufficient, mechanism choice should be based on the dominant form of interdependence together with the relevant mechanism-specific conditions, rather than on broad problem labels such as complexity alone.

2. *Mechanism – locus separation proposition.* Selection of an appropriate governance mechanism does not by itself determine its institutional locus; CoG involvement constitutes a separate design decision.

3. *Central comparative advantage proposition.* CoG involvement is warranted where a cross-government problem is sufficiently strategically salient and the selected mechanism requires executive-level convening and/or integration that cannot be provided adequately at sectoral level; its intensity should be limited to the level needed to supply that advantage.

4. *Dynamic-fit proposition.* Because the conditions underlying mechanism fit and locus fit may change, both the governance mechanism and the form of CoG involvement should be subject to periodic reassessment and adaptation.

Taken together, the propositions describe a sequence without making it static. Not every cross-government problem calls for a specialized mechanism, and a specialized mechanism does not automatically call for central ownership.

5. DISCUSSION

5.1. From Center of Government Functions to Mechanism Selection

The first implication concerns diagnosis. A functional approach requires CoG analysis to look beyond formally designated central institutions and identify the organizations and units that actually perform relevant CoG functions. This is particularly important in comparative analysis, where focusing only on the President's or Prime Minister's Office may exclude institutions that actually perform CoG functions [1, p. 9]. A functional perspective is useful for identifying where and how CoG functions are performed, but it does not by itself show which arrangement fits a particular problem.

The findings insert problem diagnosis between the general coordination function and the choice of mechanism. This interpretation is consistent with prior research emphasizing that policy and coordination arrangements need to be adapted to differences in problem structures, tasks and organizational contexts [28; 16]. Problem diagnosis helps connect the general need for coordination with the choice of an appropriate coordination mechanism.

The contribution here is therefore not another list of CoG functions. It is the link between the character of a governance problem and the capacity that a mechanism needs to provide. Broad labels such as complexity are too general for that purpose. The framework instead asks what form of interdependence dominates and whether the conditions associated with a particular mechanism are present. This is how the first part of the research question is addressed.

5.2. From a Stronger Center to Contingent Central Involvement

A similar argument applies to central involvement. Cross-government problems do not automatically justify stronger centralization. James and Ben-Gera distinguish coordination from command: while some central capacity is necessary to maintain policy coherence, they also caution against excessive control [14, p. 13]. In most cases, CoG's role remains limited, although more active involvement may be needed when issues are particularly complex or involve multiple actors. The question, then, is not simply whether stronger central involvement is needed, but how much involvement the problem requires and what form it should take.

Taken together, these considerations point to a practical limit to CoG involvement. CoG coordination capacity and senior level attention are limited, so intensive engagement cannot be sustained across a large number of issues at once [23, p. 79; 20, p. 56]. This makes selectivity a practical requirement for CoG involvement rather than simply an administrative preference.

The question is less about choosing a generally stronger or weaker CoG and more about identifying what the CoG needs to contribute in a particular case. This paper addresses that question through a locus test and the principle of minimum sufficient CoG involvement. Both are conceptual developments: central involvement is justified where it supplies a cross-government capacity that cannot be provided adequately elsewhere, and only to the extent needed to supply that capacity.

5.3. Separating Mechanism from Locus

The separation of mechanism from institutional location is another part of the argument. The relevant coordinating arrangements need not be concentrated in a single formal institution. Dunleavy and Rhodes describe the British CoG as involving a complex system of institutions, relationships, and practices, while Alessandro *et al.* emphasize that the institutions supporting the chief executive vary across countries and may also change from one administration to another [10, p. 3; 1, p. 8]. The analytical focus is therefore on how coordination is organized in practice, rather than on assuming a single institutional form.

The framework carries that functional reasoning into the design problem by distinguishing mechanism fit from locus fit. Mechanism fit concerns whether the arrangement supplies the capacity required by the diagnosed problem. Locus fit concerns whether the institutional setting can provide the authority, cross-government reach, convening capacity or other support that the mechanism needs.

This distinction also clarifies two different ways in which design can fail. The wrong mechanism may be chosen for the problem, or the right mechanism may be placed in a setting that cannot support it. Mechanism misfit and locus misfit are terms introduced here for those two cases. Recognizing the need for cross-government coordination therefore does not, on its own, determine either the arrangement or whether the CoG should own it.

5.4. Dynamic Fit and Adaptation

Fit also has a temporal dimension. As conditions change, mechanisms may need to be reassessed and adjusted to maintain their fit with the organizational context. Adaptation can take place within the mechanism itself, but it can also involve changes in where the mechanism is located institutionally.

Earlier studies suggest that adaptation may involve modifying the mechanism itself [18; 7] or changing its institutional location [7; 23]. In some cases, this may also mean that arrangements initially introduced as temporary become incorporated into more permanent structures. Reassessment can therefore consider whether outputs from temporary arrangements have simply been transferred elsewhere or have become integrated into permanent organizational structures [6; 12].

In the framework developed here, reassessment can lead to continuation, modification, reduced central involvement, relocation, or termination. These are not intended as stages of a standard lifecycle established in literature. Instead, they reflect the different ways in which a mechanism may respond to changing conditions, either through changes to the mechanism itself, its institutional location, or both.

5.5. Limitations and Future Research

The framework has several limitations. The first follows from the research design: this is a purposive, theory-guided integrative synthesis rather than a systematic review. The search was iterative and non-exhaustive because the aim was to connect several bodies of literature, not to map any one of them comprehensively.

The evidence also draws on studies from different research traditions, institutional settings and source types, for example, Christensen and Hesstvedt [8] and Brown *et al.* [7]. These sources can inform conceptual comparison, but they should not be treated as methodologically equivalent or universally generalizable.

A further limitation is that the framework has not been empirically validated as an integrated decision model. The staged diagnostic logic, the three forms of dominant interdependence, their connection to the governance archetypes, the distinction between mechanism fit and locus fit, the central comparative advantage test, and the continuum of CoG involvement are all conceptual results of the synthesis. The three mechanisms are therefore best understood as analytical archetypes, not as an exhaustive taxonomy of cross-government governance arrangements.

Future research could examine the propositions through comparative cases of mechanism choice and institutional location. Cases in which several forms of interdependence coexist would be particularly useful for assessing whether mechanisms need to be combined or sequenced. Comparative work could also ask whether similar mechanisms perform differently when located centrally or within sectoral institutions. Longitudinal studies could then trace how governments recognize emerging misfit and adapt, transfer or discontinue arrangements over time. Such research would provide the empirical basis needed to refine the framework and judge its usefulness across administrative settings.

6. CONCLUSIONS

The research question asked how structural contingency theory can inform the selection, design and institutional location of governance mechanisms for cross-government problems, and when CoG involvement provides institutional advantage. The answer developed here rests on a simple contingency principle: the response should follow the problem rather than a preferred organizational form. Where ordinary administrative responsibilities and coordination are insufficient, mechanism choice turns on the dominant form of interdependence and the relevant mechanism-specific conditions. A second decision then concerns locus – whether the chosen mechanism needs capacities that are better supplied from the center.

The analysis distinguishes three governance archetypes – Strategic Innovation Portfolio Governance, Mediated Coordination, and a Project-specific Expert Working Group – but none of them implies CoG ownership by default. In the framework developed here, CoG involvement is justified where a problem is cross-government and strategically salient and where the selected mechanism requires executive level convening and/or integration that cannot be provided adequately at sectoral level. Even then, the center should do no more than is needed to supply that advantage.

Neither mechanism fit nor locus fit is assumed to be permanent. Changes in the problem, interdependencies or administrative capacities may call for a different mechanism, more or less central involvement, a change of locus, or termination of an arrangement that is no longer useful. Comparative and longitudinal research is needed before its propositions, and boundary conditions can be assessed across different administrative settings.

References

[1] Alessandro, M., Lafuente, M., & Santiso, C. (2013). *The role of the center of government: A literature review.* Inter-American Development Bank. https://dx.doi.org/10.18235/0009130

[2] Ansell, C., Sørensen, E., & Torfing, J. (2025). Theorizing the political dimension of collaborative governance. *Perspectives on Public Management and Governance, 8*(3), 158-171. https://doi.org/10.1093/ppmgov/gvaf007

[3] Ansell, C., & Gash, A. (2012). Stewards, mediators, and catalysts: Toward a model of collaborative leadership. *The Innovation Journal, 17*(1), 2-20.

[4] Bianchi, C., Nasi, G., & Rivenbark, W. C. (2021). Implementing collaborative governance: models, experiences, and challenges. *Public Management Review, 23*(11), 1581-1589. https://doi.org/10.1080/14719037.2021.1878777

[5] Bommert, B. (2010). Collaborative innovation in the public sector. *International Public Management Review, 11*(1), 15-33.

[6] Brix, J. (2017). Exploring knowledge creation processes as a source of organizational learning: A longitudinal case study of a public innovation project. *Scandinavian Journal of Management, 33*(2), 113-127. http://dx.doi.org/10.1016/j.scaman.2017.05.001

[7] Brown, D., Kohli, J., & Mignotte, S. (2021). *Tools at the Centre of Government: Research and Practitioners' Insight*. Blavatnik School of Government, University of Oxford. https://doi.org/10.35489/bsg-peoplegov-ri_2021/002

[8] Christensen, J., & Hesstvedt, S. (2024). The influence of expert groups: a citation analysis. *Journal of European Public Policy, 31*(5), 1259-1294. https://doi.org/10.1080/13501763.2023.2174168

[9] Donaldson, L. (2006). *The Contingency Theory of Organizational Design: Challenges and Opportunities*. In: Burton, R.M., Håkonsson, D.D., Eriksen, B., Snow, C.C. (eds) Organization Design. Information and Organization Design Series, vol 6. Springer, Boston, MA, pp. 19-40. https://doi.org/10.1007/0-387-34173-0_2

[10] Dunleavy, P., & Rhodes, R. A. W. (1990). Core executive studies in Britain. *Public Administration, 68*(1), 3-28. https://doi.org/10.1111/j.1467-9299.1990.tb00744.x
[11] Eriksson, R. (2025). Who gives expert advice? Empirical evidence from Swedish policy-making. *Governance, 38*(4), e70046. https://doi.org/10.1111/gove.70046
[12] Godenhjelm, S., Sjöblom, S. M., & Jensen, C. (2019). *Project Governance in an Embedded State: Possibilities and Challenges.* In Hodgson, D., Fred, M., Bailey, S., & Hall, P. (eds.), The Projectification of the Public Sector. Routledge Critical Studies in Public Management, Routledge, New York, pp. 149-168. https://doi.org/10.4324/9781315098586-9
[13] Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS Review, 10*(1), 18-26. https://doi.org/10.1007/s13162-020-00161-0
[14] James, S., & Ben-Gera, M. (2004). *A comparative analysis of government offices in OECD countries* (GOV/PGC/MPM/RD(2004)1). OECD.
[15] Joseph, J., & Sengul, M. (2025). Organization design: Current insights and future research directions. *Journal of Management, 51*(1), 249-308. https://doi.org/10.1177/01492063241271242
[16] Lee, H. W., & Rhee, D. Y. (2026). Organizational Context Matters: Interdepartmental Collaboration and Performance in Law Enforcement vs. Scientific Research Agencies. *Behavioral Sciences, 16*(5), 787. https://doi.org/10.3390/bs16050787
[17] Mazzucato, M. (2018). Mission-oriented innovation policies: Challenges and opportunities. *Industrial and Corporate Change, 27*(5), 803-815. https://doi.org/10.1093/icc/dty034
[18] Meijer, A., & Thaens, M. (2021). The dark side of public innovation. *Public Performance & Management Review, 44*(1), 136-154. https://doi.org/10.1080/15309576.2020.1782954
[19] Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National academy of Sciences, 111*(20), 7176-7184. https://doi.org/10.1073/pnas.1319946111
[20] OECD (2024). *Steering from the Centre of Government in Times of Complexity: Compendium of Practices.* OECD Publishing, Paris. https://doi.org/10.1787/69b1f129-en
[21] Petrovska, E., & Berzins, G. (2024). Governance Framework of the Centre of Government. *Economics and Culture, 21*(2), 107-122. https://doi.org/10.2478/jec-2024-0022
[22] Rashman, L., Withers, E., & Hartley, J. (2009). Organizational learning and knowledge in public service organizations: A systematic review of the literature. *International Journal of Management Reviews, 11*(4), 463-494. https://doi.org/10.1111/j.1468-2370.2009.00257.x
[23] Shostak, R., Alessandro, M., Diamond, P., Mosqueira, E., & Lafuente, M. (2023). *The center of government, revisited: a decade of global reforms* (IDB Monograph No. 1113). Inter-American Development Bank. https://doi.org/10.18235/0004994
[24] Smith, R. D., Schäfer, S., & Bernstein, M. J. (2024). Governing beyond the project: Refocusing innovation governance in emerging science and technology funding. *Social Studies of Science, 54*(3), 377-404. https://doi.org/10.1177/03063127231205043
[25] Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039
[26] Torfing, J., Payandeh, R., Jalili, S. M., & Banafi, M. (2026). A systematic review of conflict within collaborative governance. *Policy & Politics, 54*(1), 2-25. https://doi.org/10.1332/03055736Y2025D000000070
[27] Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review, 4*(3), 356-367. https://doi.org/10.1177/1534484305278283
[28] Wanzenböck, I., Wesseling, J. H., Frenken, K., Hekkert, M. P., & Weber, K. M. (2020). A framework for mission-oriented innovation policy: Alternative pathways through the problem-solution space. *Science and Public Policy, 47*(4), 474-489. https://doi.org/10.1093/scipol/scaa027

Declaration of AI-Assisted Language Editing

During the preparation of this paper, the author used ChatGPT (OpenAI) solely to improve English language, grammar, syntax, and readability.

Financial Distress Classification among Albanian Treasury Payment Recipients Using Explainable Machine Learning

Mikel ALLA[a,b,c*]

[a] *Accounting & Fiscal Expertise Studio, Elbasan, Albania*
[b] *University of Elbasan "Aleksandër Xhuvani", Elbasan, Albania*
[c] *University "Aleksandër Moisiu", Durrës, Albania*
Correspondence: mikel_alla@yahoo.com

Abstract

This study compares machine-learning classification of an Altman Z″ accounting benchmark among major Albanian State Treasury payment recipients in 2024. Of 100 records, 96 have a computable target, including 18 distress labels. Penalized Logistic Regression, Random Forest, Gradient Boosting, and Support Vector Machine are evaluated using stratified five-fold cross-validation repeated three times, with preprocessing fitted within training folds. Twelve predictors retain accounting information related to the target. Area under the precision–recall curve (PR-AUC) is the primary metric; held-out permutation importance describes Gradient Boosting model reliance. Random Forest and Gradient Boosting both achieve mean PR-AUC of 0.943. Gradient Boosting attains balanced accuracy of 0.918 and recall of 0.883; Random Forest records precision of 0.880 and specificity of 0.966. Support Vector Machine has recall of 0.944 but precision of 0.375. Current Ratio dominates permutation importance, with mean PR-AUC reduction of 0.541; Treasury-payment value and transaction frequency have near-zero importance. The ensembles closely reproduce the contemporaneous benchmark with different error profiles. Accounting overlap, proxy inputs, and the small positive class limit interpretation. The findings do not validate future-default prediction, causal explanations, or replacement of direct Z″ computation.

Keywords*:* financial distress classification; explainable machine learning; Altman Z″; Treasury payments

1. INTRODUCTION

Corporate financial distress is a deterioration in a firm's capacity to generate adequate returns, maintain liquidity, and satisfy contractual obligations. It may precede formal insolvency by a considerable period, so accounting information is often used for early screening even when no legal failure event has occurred. This distinction matters because accounting vulnerability, payment default, restructuring, and bankruptcy are related but non-equivalent outcomes (Altman et al., 2019; Beaver, 1966).

The issue has a public-finance dimension when firms receive substantial payments from the State Treasury. Financial weakness among suppliers, contractors, concessionaires, or other beneficiaries can affect project continuity and counterparties, yet the value of Treasury receipts is not itself a measure of financial resilience.

* Corresponding author

Payment totals describe the scale of recorded public disbursements; they do not identify the underlying transaction type, settlement delay, contract concentration, profitability, or liability maturity.

Accounting-based distress models remain influential because they translate observable financial ratios into transparent screening rules. Beaver (1966), Altman (1968), Ohlson (1980), and Zmijewski (1984) established profitability, liquidity, leverage, accumulated earnings, and operating efficiency as central dimensions of financial condition. Their interpretability is an advantage, but fixed functional forms can miss nonlinear thresholds and interactions.

Machine-learning methods can represent nonlinear relationships among financial indicators. Evidence from bankruptcy and distress studies frequently favours tree ensembles over conventional benchmarks, although rankings depend on the sample and the evaluation design (Barboza et al., 2017; Dasilas & Rigani, 2024). Recent studies extend this comparison to time-aware validation, cross-market stability, and imbalance-sensitive learning (Hussain & Sun, 2026; Liu et al., 2026; Nam & Huy, 2026). These advances make the definition of the predicted outcome and the quality of validation central to assessing reported performance.

Explainability poses a related question: which information sustains a model's classifications? Finance-focused reviews distinguish model transparency, post-hoc explanation, and the needs of decision-makers (Černevičienė & Kabašinskas, 2024; Khan et al., 2025; Weber et al., 2024; Yeo et al., 2025). An importance ranking describes the fitted model's use of information. It does not, by itself, identify the causes of financial distress.

Against this background, the study evaluates four fixed machine-learning configurations in a cross-section of 100 Albanian enterprises receiving the highest aggregate State Treasury payments in 2024. A computable binary Altman Z″ benchmark is available for 96 firms, including 18 distress labels and 78 grey/safe labels.

The empirical objective is to compare four fixed configurations under a common internal evaluation protocol. Repeated stratified five-fold cross-validation generates 15 held-out evaluations per model. PR-AUC is the primary criterion because distress labels are a minority; ranking and threshold-dependent metrics provide complementary information about false alerts and missed labels. The interpretation analysis then examines global reliance in Gradient Boosting, the configuration with the highest unrounded mean PR-AUC.

Three research questions organise the analysis. RQ1 asks how accurately the four fixed model configurations classify the accounting benchmark in held-out observations. RQ2 asks how model rankings change across ROC-AUC, PR-AUC, precision, sensitivity, and specificity. RQ3 asks which predictors the explained model relies on most strongly in maintaining held-out PR-AUC.

The contribution is an application to a payment-selected Albanian sample that links accounting-target construction, imbalance-sensitive model comparison, and explanation of model reliance. The study examines how well alternative ratios reproduce an established screening rule and how the choice of metric changes the apparent ranking of classifiers. It also makes explicit why strong benchmark agreement cannot be treated as evidence of independent bankruptcy prediction.

The research design is therefore contemporaneous and descriptive. The 2024 accounts and Treasury aggregates characterise the same fiscal year; there is no forecast horizon or independently observed subsequent failure outcome.

2. THEORY AND LITERATURE REVIEW

2.1 Financial Distress as a Measurement Problem

Financial distress is not a single universally observed event. Empirical studies operationalise it through legal bankruptcy, payment default, restructuring, negative equity, persistent losses, special regulatory status, or accounting-based scores. Because these targets represent different stages of deterioration, performance statistics are not directly comparable across studies without considering how the outcome is defined (Dasilas & Rigani, 2024).

The Altman Z-score family remains influential because it combines several accounting dimensions into an interpretable screening score. International evidence confirms broad usefulness but also shows that classification performance varies across countries, sectors, and periods (Altman et al., 2017). In the present setting, Z″ is used as an accounting benchmark. Its conventional cut-offs are screening thresholds, not locally calibrated insolvency probabilities, and a model trained on those labels classifies the benchmark rather than independently observing failure.

A constructed accounting target creates information overlap even when its named components are excluded from the predictor matrix. Related ratios may retain identical information through algebraic transformations. Leakage research shows why separation of training and test data is only one part of assessing the validity of a predictive claim (Kapoor & Narayanan, 2023). In this study, classification against Z″ is the stated task, so accounting overlap

is intrinsic to the estimand. Agreement with that benchmark does not, by itself, provide independent evidence of future failure.

2.2 Accounting Dimensions of Corporate Risk

Profitability indicators summarise different relationships between earnings and the resources used to generate them. Return on assets (ROA) relates net profit to total assets, return on equity (ROE) to book equity, and Net Margin to revenue. Their role in distress screening follows the emphasis on earnings capacity in classical accounting models (Beaver, 1966; Ohlson, 1980). In a single-year cross-section, these ratios describe current financial condition rather than persistence of losses.

Liquidity ratios compare current resources with short-term obligations. Current Ratio includes all current assets, Quick Ratio removes inventories and deferred expenses, and Cash Ratio focuses on recorded cash. These measures have different economic meanings because receivable recoverability and the timing of liabilities affect the availability of resources for payment. The present data do not identify those timing differences directly.

Leverage describes the share of recorded assets financed by liabilities and the size of liabilities relative to book equity. Here, the variables conventionally named Debt-to-Assets and Debt-to-Equity use total liabilities, including non-interest-bearing obligations. They therefore measure balance-sheet leverage rather than bank borrowing alone. This distinction matters when comparing contractors, service providers, and other Treasury recipients.

Asset Turnover approximates operating efficiency and Interest Coverage captures the ability to service financing costs. Treasury-payment value and transaction frequency add a public-payment dimension. These variables may proxy scale or recurring public-sector activity, but they do not directly measure payment delays, contract concentration, the share of public receipts in total revenue, or underlying profitability. Their role is therefore treated as empirical rather than assumed.

2.3 Recent Machine-Learning Evidence

Comparative studies establish that flexible models can improve distress classification, but do not identify a universally superior algorithm. Barboza et al. (2017) provide an influential comparison of machine-learning and traditional approaches. More recent evidence covers Nordic enterprises (Abrahamsen et al., 2024), Chinese listed construction firms (Rahman & Zhu, 2024), and TSX-listed firms (Lokanan & Ramzan, 2024). These settings differ in the target event, sector coverage, observation period, and available information, limiting direct comparison of headline accuracy measures.

Credit-scoring benchmarks provide a complementary evaluation perspective. Lessmann et al. (2015) compare 41 classifiers across eight retail-credit datasets using six performance measures and examine the relationship between statistical accuracy and business value. Their design supports comparing several criteria here, while the difference between borrower outcomes and the present accounting proxy precludes direct comparison of performance levels.

The closest Albanian study is Dhamo et al. (2025), which analyses 16 financial ratios for 187 large non-financial businesses over 2011–2014, selected by 2014 revenue. It reports favourable performance for Random Forest and AdaBoost using classification accuracy, confusion matrices, and AUC. The present study changes the selection frame to Treasury-payment recipients in 2024 and focuses explicitly on minority-class retrieval and model reliance. It extends a particular empirical setting and reporting approach; it does not establish priority for applying machine learning to Albanian financial distress.

Class-imbalance research explains why evaluation must extend beyond ordinary accuracy. Song et al. (2024) broaden the information set, while Amirshahi and Lahmiri (2024) and Gao et al. (2025) examine ensemble approaches under imbalance. Huang and Wang (2026) combine sampling, cost-sensitive learning, and heterogeneous ensembles across China, Poland, and the United States, linking classification improvements to misclassification costs. Their results support examining error trade-offs explicitly; they do not establish the value of any particular cost ratio for this Albanian sample.

Recent work also sharpens the distinction between discrimination and validation. Nam and Huy (2026) use time-aware training and a later-period evaluation in Vietnam. Hussain and Sun (2026) examine explainable models across markets and structural regimes, while Liu et al. (2026) address interpretable learning under class imbalance. The present cross-section cannot reproduce those temporal or cross-market tests. Its restricted sample instead supports a comparison of constrained tabular configurations, with uncertainty assessed descriptively across partitions.

2.4 Explainability, Stability, and Governance

Post-hoc explanations depend on the fitted function, the reference data, and the way predictors are perturbed. Shapley additive explanations (SHAP) allocate predictions among input features, whereas permutation importance measures the deterioration in a chosen performance statistic when a feature is disrupted (Lundberg & Lee, 2017; Molnar, 2022). These methods answer different questions. The present analysis uses permutation importance only; it provides neither individual-company SHAP explanations nor estimated causal effects.

Feature dependence and model choice affect the interpretation of importance. Fisher et al. (2019) show that similarly accurate models can rely on different variables. Strobl et al. (2008) demonstrate for random forests that predictor correlation and the permutation scheme affect importance, distinguishing marginal from conditional assessments. This distinction informs interpretation here: unconditional permutation does not isolate a predictor's unique contribution given the remaining ratios. With correlated liquidity and leverage measures, small importance can reflect substitution, while disruption can create implausible accounting combinations. Finance reviews likewise identify explanation stability and stakeholder relevance as issues alongside discrimination (Khan et al., 2025; Yeo et al., 2025).

Held-out permutation importance is used here as a global diagnostic of Gradient Boosting. A positive value means that disrupting a predictor reduces the reported held-out PR-AUC while the fitted model remains unchanged. The measure does not show the direction of a relationship, allocate distress causally, or estimate performance after retraining without the variable.

2.5 Research Gap and Analytical Framework

The empirical gap concerns the combination of a Treasury-payment selection frame, a contemporaneous accounting benchmark, and explicit evaluation of both classification errors and model reliance. Revenue-ranked Albanian studies and international failure-prediction studies do not directly answer how these elements behave together in this sample.

RQ1 is addressed through mean held-out discrimination and classification metrics. RQ2 is addressed by comparing the models' precision, recall, specificity, and repeated confusion counts. RQ3 is addressed through the Gradient Boosting permutation ranking. This structure separates descriptive model comparison from explanations of financial condition.

The analysis does not formulate causal hypotheses or inferential tests of algorithmic superiority. Ensemble flexibility provides a reason to compare nonlinear and linear specifications, but any observed advantage is conditional on the target, predictors, fixed configurations, and resampling design.

3. DATA AND METHODOLOGY

3.1 Sample Definition and Data Sources

The fixed research sample consists of 100 enterprise records in Open Data Albania's ranking of the businesses receiving the largest aggregate payments from the Albanian State Treasury in 2024, published on 12 May 2025 (Open Data Albania, 2025). Selection reflects Treasury-payment value rather than revenue, assets, employment, or market capitalisation. The sample is therefore policy-relevant but is not a conventional ranking of the largest Albanian firms.

Financial information was extracted from the enterprises' 2024 statements and mapped to common fields for assets, liabilities, equity, revenue, profit, and financing. The source log records statement filenames and business identifiers. The processing record documents one name-change mapping and nine reporting-year label normalisations accepted after the researcher confirmed that the underlying information concerned 2024. Missing fields remain identifiable, and source values are retained without retrospective correction based on model performance.

The database contains 100 distinct business identifiers. Four records lack the information required to calculate the target, leaving 96 observations for supervised estimation; missing targets are not imputed. Inclusion is determined by target computability rather than by complete observation of every predictor. This distinction avoids confusing the full research database with the estimation sample.

3.2 Target Construction

The study uses the four-component Altman Z″ specification without an additive constant (Altman et al., 2017, 2019):

$$Z'' = 6.56\,\frac{WC}{TA} + 3.26\,\frac{RE}{TA} + 6.72\,\frac{EBIT}{TA} + 1.05\,\frac{E}{TL}$$

WC is current assets minus current liabilities; TA is total assets; RE is retained earnings or accumulated losses; E is book equity; and TL is total liabilities. EBIT is the database's approximate operating-earnings measure, using operating profit where disclosed or profit before tax adjusted for financial and participation results. Thus, the earnings component reflects statement-template harmonisation rather than an independently standardised EBIT series.

Conventional Z″ thresholds are applied: distress for values below 1.10, grey zone for values from 1.10 to 2.60, and safe for values above 2.60. The binary target equals 1 for distress and 0 for the combined grey/safe class. The 96 labelled firms comprise 18 distress, 25 grey, and 53 safe observations. Combining grey and safe into class 0 is a modelling convention and does not certify that every zero-labelled firm is financially sound.

Retained earnings are separately available for 63 labelled observations. For the other 33, RE is set to zero in the benchmark, without implying that accumulated earnings were actually zero. The complete score would equal the recorded proxy plus 3.26(RE/TA): positive undisclosed retained earnings would raise the complete score, while accumulated losses would lower it. The direction and extent of label error therefore depend on the missing value and proximity to the cut-off. The resulting target is an accounting proxy, not a locally calibrated probability of insolvency.

3.3 Predictor Construction

Table 1. Predictor set used in the machine-learning models

Dimension	**Variables**	**Definition**
Profitability	ROA; ROE; Net Margin	ROA = net profit / assets; ROE = net profit / equity; Net Margin = net profit / revenue
Liquidity	Current Ratio; Quick Ratio; Cash Ratio	Current = current assets / current liabilities; Quick = (current assets − inventories − deferred expenses) / current liabilities; Cash = cash / current liabilities
Leverage	Debt-to-Assets; Debt-to-Equity	Debt-to-Assets = total liabilities / assets; Debt-to-Equity = total liabilities / equity
Efficiency and servicing	Asset Turnover; Interest Coverage	Asset Turnover = revenue / assets; Interest Coverage = approximate EBIT / absolute interest expense
Treasury exposure	Log Treasury Payments; Log Treasury Transactions	ln(1 + annual Treasury payments in ALL); ln(1 + annual transaction count)

Note. All financial ratios use the database definitions; denominators for assets and equity are year-end values. ALL = Albanian lek. Zero and missing inputs are treated as described in Section 3.3.

The four named Z″ components are excluded as explicit predictors, but exact and partial accounting overlap remain. Debt-to-Equity is TL/E, so its reciprocal equals E/TL whenever both denominators are non-zero. If TA = TL + E, Debt-to-Assets also implies E/TL = 1/(TL/TA) − 1. Current Ratio is related to WC/TA through current liabilities relative to assets. Consequently, excluding component names does not create an information set independent of the target; the experiment measures benchmark classification from related accounting representations.

Table 1 defines the 12 predictors. Balance-sheet denominators are the recorded year-end amounts rather than annual averages. Revenue excludes separately identified other operating income. The Treasury variables use natural logarithms, ln(1 + payment value in Albanian lek [ALL]) and ln(1 + transaction count), to retain zero-valued inputs if present and compress differences in scale.

Missing or zero denominators produce missing ratios. Before model fitting, missing inventory and deferred-expense inputs are treated as zero in Quick Ratio; 63 of its 95 available observations use at least one such substitution. If an omitted amount is positive, this rule overstates the resulting liquidity ratio. Statement-unit metadata are incomplete for part of the sample. Ratios within a consistently scaled statement are scale-invariant, but cross-statement ratios such as ROA and EBIT/TA depend on compatible scales between the income statement and balance sheet. Model-stage imputation cannot correct those source-level assumptions.

3.4 Missing Data and Preprocessing

The processed database preserves model-stage missing values rather than replacing them before validation. Median imputation is fitted separately within each training fold and then applied to the corresponding held-out fold. Missingness indicators are added for incomplete predictors so that the models can distinguish observed from imputed values.

Robust scaling for Penalized Logistic Regression and Support Vector Machine uses training-fold location and scale parameters. The tree ensembles are unscaled. Balanced class weights are used for Logistic Regression and Support Vector Machine, balanced subsample weights for Random Forest, and balanced training-sample weights for Gradient Boosting. Weighting addresses class frequency in estimation; it does not identify the economic cost of a missed label or false alert.

Interest Coverage has the highest model-stage missingness, with 48 valid and 48 missing observations. This feature is retained, with a missingness indicator, but its substantive interpretation is weaker than that of fully observed predictors. Fold-contained imputation does not correct source-level construction assumptions; those affect the underlying ratios before cross-validation.

3.5 Model Specifications

Table 2. Fixed model configurations

Model	**Specification**	**Analytical role**
Penalized Logistic Regression	C = 0.5; liblinear solver; balanced class weights; robust scaling	Linear classification benchmark
Random Forest	120 trees; maximum depth 4; minimum leaf 3; sqrt features; balanced subsample weights	Constrained nonlinear ensemble
Gradient Boosting	60 estimators; learning rate 0.05; depth 2; minimum leaf 4; subsample 0.80; balanced sample weights	Sequential shallow-tree ensemble
Support Vector Machine	RBF kernel; C = 1; gamma = scale; balanced class weights; robust scaling	Nonlinear margin classifier

Note. C is inverse regularisation strength; RBF = radial basis function; sqrt indicates square-root feature sampling. The specifications are fixed model configurations, not the result of a reported nested tuning comparison.

The configurations follow established model families (Breiman, 2001; Cortes & Vapnik, 1995; Friedman, 2001) and standard statistical-learning practice (Hastie et al., 2009; James et al., 2023). They compare constrained specifications in a small sample rather than exhaustively tuned algorithm families. Gradient Boosting is selected for explanation because its unrounded mean PR-AUC is the highest. That selection uses the same internal evaluation as the model comparison, so it provides no independent post-selection performance estimate (Cawley & Talbot, 2010).

3.6 Model Testing Protocol

A permanent train-test split is not used because only 18 positive labels are available. A single split would place only a few distress cases in the test set and would make the reported performance unusually sensitive to the random partition.

Instead, repeated stratified five-fold cross-validation is used. Each test fold contains approximately 19 or 20 firms, including three or four positive labels, while the remaining observations form the corresponding training fold. Stratification preserves the class proportion as closely as possible.

The five-fold procedure is repeated three times, yielding 15 held-out evaluations for each model. Each firm is tested once in each repeat, so every model produces 288 held-out classification instances from 96 firms: 54 repeated positive instances and 234 repeated negative instances. These are repeated predictions of the same 96 firms, not 288 independent observations.

Model-stage preprocessing is fitted within the training portion of each fold. Median imputation, missingness indicators, scaling, and class or sample weighting use no held-out observations. This separation limits preprocessing leakage. The fixed accounting definitions and source-level substitutions described in Sections 3.2–3.3 precede validation, and their measurement limitations and predictor–target overlap remain.

The primary statistic is PR-AUC, the precision–recall area measure recorded in the analytical outputs. Positive labels account for 18.75% of the sample. Precision–recall analysis is sensitive to positive-class retrieval and prevalence, whereas ROC-AUC describes ranking through true-positive and false-positive rates (Davis & Goadrich, 2006; Fawcett, 2006; Saito & Rehmsmeier, 2015). ROC-AUC and PR-AUC need not rank models identically. Accuracy is the proportion classified correctly; balanced accuracy is the mean of recall and specificity; precision is the proportion of alerts carrying a positive label; and F1 is the harmonic mean of precision and recall.

Metrics are calculated separately for each of the 15 held-out folds and summarised using unweighted arithmetic means and sample standard deviations. The confusion counts are also summed across folds, producing a distinct aggregation of repeated classification instances. Overlapping training samples and repeated evaluation of the same firms mean that fold standard deviations are descriptive measures of partition sensitivity, not confidence intervals (Bengio & Grandvalet, 2004). No paired significance test, probability-calibration assessment, decision-cost optimisation, or independent temporal evaluation is reported.

3.7 Model Interpretation and Analytical Record

Gradient Boosting is examined with held-out permutation importance. Each predictor is disrupted while the fitted model is held fixed, and the decrease in PR-AUC measures reliance on that predictor. The reported mean importance and standard deviation are retained as descriptive outputs. The dispersion is not a confidence interval or a significance test; explanations apply to the selected model rather than all four classifiers.

The analytical workbook retains 15 fold-level metric and confusion-count records for each model. Its separate out-of-fold table contains 96 company-level mean scores, score standard deviations, and classifications. For these aggregated scores, stored classes correspond to a 0.5 boundary for Logistic Regression and the tree ensembles, and a zero boundary for Support Vector Machine decision scores. These company-level summaries are distinct from the 288 repeated fold classifications and cannot replace the individual fold scores when reconstructing AUC. Scores are not interpreted as calibrated default probabilities.

The workbook does not retain individual scores for every fold and repeat, random seeds, software versions, the precise PR-AUC integration convention, or permutation repetition and aggregation settings. It supports checks of the reported summaries and confusion counts, but is insufficient for independent reconstruction of AUC calculations and permutation dispersion.

4. FINDINGS

4.1 Sample Composition and Descriptive Statistics

The modelling sample contains 96 firms: 18 distress labels and 78 grey/safe labels. A classifier assigning every observation to the majority class would achieve 81.25% accuracy and 50% balanced accuracy while identifying no distress cases. This arithmetic baseline illustrates why ordinary accuracy cannot serve as the primary criterion.

Table 3. Descriptive statistics for the modelling sample

Predictor	Valid N	Missing	Mean	SD	Median	Minimum	Maximum
ROA	96	0	0.083	0.086	0.058	-0.026	0.491
ROE	96	0	0.223	0.210	0.159	-0.189	0.994
Net Margin	94	2	0.119	0.150	0.074	-0.071	1.107

Predictor	Valid N	Missing	Mean	SD	Median	Minimum	Maximum
Current Ratio	95	1	3.095	8.735	1.431	0.171	83.120
Quick Ratio	95	1	2.546	6.705	1.301	0.131	62.403
Cash Ratio	95	1	0.517	2.019	0.073	-0.070	18.841
Debt-to-Assets	96	0	0.596	0.228	0.614	0.009	0.985
Debt-to-Equity	96	0	3.971	9.187	1.591	0.009	65.517
Asset Turnover	94	2	0.802	0.514	0.676	0.055	2.634
Interest Coverage	48	48	1,080.757	3,836.384	21.663	-0.165	22,305.056
Log Treasury Payments	96	0	20.399	0.624	20.215	19.633	22.102
Log Treasury Transactions	96	0	3.782	1.343	3.850	1.099	6.795

Note. N = 96 labelled observations. Statistics describe observed values before training-fold imputation. SD is the sample standard deviation. Financial ratios are dimensionless; Treasury predictors use natural logarithms.

Profitability indicators are generally positive, but several distributions are right-skewed. Liquidity and leverage display especially wide dispersion. Current Ratio has a mean of 3.095, a median of 1.431, and a maximum of 83.120, showing that a small number of observations strongly influence the mean.

The median Debt-to-Assets ratio is 0.614, corresponding to liabilities of 61.4% of assets at the midpoint of the ratio distribution. Debt-to-Equity displays greater dispersion; small equity denominators can amplify this ratio.

Interest Coverage is highly skewed and incomplete: its median is 21.663, its mean is 1,080.757, and 48 observations are available. Cash Ratio has a negative minimum (−0.070) because one recorded cash numerator is negative. That value is retained as recorded and represents a source-measurement limitation; it should not be interpreted as evidence that negative cash is an ordinary liquidity condition. Robust scaling limits the influence of scale differences but does not validate anomalous inputs.

4.2 Comparative Model Performance

Table 4. Mean performance across 15 held-out test folds

Model	ROC-AUC	PR-AUC	Balanced accuracy	Accuracy	Precision	Recall	Specificity	F1
Logistic regression	0.937	0.826	0.844	0.826	0.543	0.872	0.816	0.658
Random Forest	0.986	0.943	0.911	0.948	0.880	0.856	0.966	0.846
Gradient Boosting	0.973	0.943	0.918	0.942	0.841	0.883	0.953	0.845
SVM	0.912	0.756	0.769	0.661	0.375	0.944	0.594	0.527

Note. Values are unweighted means of 15 held-out fold metrics, rounded to three decimals. PR-AUC retains the metric label used in the analytical record; its exact integration convention is not documented. SVM = Support Vector Machine. Logistic regression denotes the penalized specification in Table 2.

Random Forest and Gradient Boosting form the leading observed group. Gradient Boosting records the highest mean balanced accuracy (0.918) and an unrounded PR-AUC only 0.000304 above Random Forest; both round to 0.943. Its mean recall of 0.883 exceeds Random Forest's 0.856, while its specificity remains high at 0.953.

Random Forest records the highest mean ROC-AUC (0.986), overall accuracy (0.948), precision (0.880), specificity (0.966), and F1-score (0.846). At the applied decision rule, these values indicate more selective positive alerts and fewer false-positive classifications than the other configurations.

The PR-AUC gap has no established statistical significance. Ranking measures and threshold-dependent error profiles describe distinct aspects of performance: AUC comparisons concern score rankings, whereas precision, recall, and confusion counts also depend on the applied classification rule.

Penalized Logistic Regression remains an informative transparent benchmark. Its ROC-AUC of 0.937 and recall of 0.872 indicate substantial discrimination with a linear decision function in the constructed predictors. Its precision of 0.543, however, indicates more false-positive alerts than the leading ensembles.

Support Vector Machine records the highest recall (0.944) but much lower specificity (0.594) and precision (0.375). Its accuracy of 0.661 is below the always-negative arithmetic benchmark because the model generates many false positives. Whether such a sensitivity-oriented trade-off is useful depends on decision costs and threshold selection, neither of which is optimised in this study.

4.3 Aggregate Held-Out Classification Counts

Table 5. Aggregate confusion counts across all 15 held-out folds

Model	TN	FP	FN	TP	Repeated classifications
Logistic regression	191	43	7	47	288
Random Forest	226	8	7	47	288
Gradient Boosting	223	11	6	48	288
SVM	139	95	3	51	288

Note. TN = true negatives; FP = false positives; FN = false negatives; TP = true positives. Totals cover three held-out evaluations per firm, comprising 54 repeated positive and 234 repeated negative instances. They are not counts of unique firms.

The pooled counts make the classification trade-offs more concrete. Because each of the 96 firms appears in one held-out fold during each of three repeats, every model generates 288 repeated classification instances. These are not counts of 288 unique firms.

Random Forest produces 8 false-positive and 7 false-negative instances across the pooled repeated predictions, compared with 11 and 6 for Gradient Boosting, 43 and 7 for Penalized Logistic Regression, and 95 and 3 for Support Vector Machine. The figures confirm that Random Forest is more selective at the applied decision rule, while Gradient Boosting misses one fewer repeated positive instance.

Pooled counts and mean fold metrics describe different aggregation levels. For example, Random Forest's pooled recall is 47/54 = 0.870, whereas its unweighted mean fold recall is 0.856; Gradient Boosting's pooled precision is 48/59 = 0.814, whereas its mean fold precision is 0.841. Such differences are expected when fold denominators vary and metrics are nonlinear.

4.4 Stability Across Test Folds

Table 6. Standard deviation across held-out folds

Model	ROC-AUC	PR-AUC	Balanced accuracy	Accuracy	Precision	Recall	Specificity	F1
Logistic regression	0.053	0.161	0.095	0.077	0.148	0.172	0.085	0.136
Random Forest	0.019	0.081	0.112	0.048	0.141	0.226	0.041	0.160
Gradient Boosting	0.066	0.114	0.113	0.060	0.180	0.218	0.056	0.174
SVM	0.057	0.175	0.096	0.147	0.104	0.116	0.190	0.103

Note. Sample SDs use 15 fold observations and a denominator of 14. Overlapping folds preclude treating these values as independent-sample uncertainty estimates. Column headings identify the metric whose SD is reported.

Table 7. Observed fold ranges for the principal imbalance-sensitive metrics

Model	PR-AUC range	Balanced accuracy range
Logistic regression	0.533-1.000	0.650-0.969
Random Forest	0.804-1.000	0.667-1.000
Gradient Boosting	0.627-1.000	0.635-1.000
SVM	0.458-1.000	0.608-0.906

Note. Each interval gives the minimum and maximum across the 15 held-out evaluations. The intervals are observed ranges, not confidence intervals.

Random Forest has the lowest standard deviation for ROC-AUC (0.019), PR-AUC (0.081), accuracy (0.048), and specificity (0.041). Gradient Boosting has slightly lower recall variability than Random Forest (0.218 versus 0.226), whereas Random Forest has lower precision variability (0.141 versus 0.180). Stability is therefore metric-specific rather than uniformly dominated by one model.

Recall varies materially across folds because each held-out fold contains only three or four positive labels. One additional false negative can change fold-level recall by approximately one-third or one-quarter. The wide ranges therefore reflect the small positive class as well as partition sensitivity.

Repeated testing makes partition sensitivity visible but does not add independent distress cases. The particularly small number of positives per fold explains why an apparently stable ranking statistic can coexist with variable recall. These patterns support cautious comparisons rather than claims of uniformly stable operational performance.

4.5 Explainability Results

Table 8. Held-out permutation importance for Gradient Boosting

Predictor	Mean PR-AUC reduction	SD
Current Ratio	0.541	0.057
Net Margin	0.025	0.055
Quick Ratio	0.013	0.030
Debt-to-Equity	0.006	0.009
Debt-to-Assets	0.006	0.014
ROE	0.005	0.012
ROA	0.005	0.007
Log Treasury Transactions	0.005	0.011
Interest Coverage	0.004	0.009
Cash Ratio	0.000	0.000
Asset Turnover	0.000	0.000
Log Treasury Payments	-0.006	0.014

Note. Importance is an absolute reduction in the recorded held-out PR-AUC, with no refitting after permutation. SD is the dispersion retained in the importance output; the detailed repetition and aggregation scheme is unavailable. Values are descriptive, not tests of significance. A negative mean indicates that performance increased slightly after disruption.

Current Ratio has the largest reported permutation importance. Disrupting it reduces Gradient Boosting's held-out PR-AUC by 0.541 on average. This is an absolute change in the metric, not a percentage of distress explained.

The result identifies a strong dependence of the fitted specification on this predictor, without establishing how a particular change in Current Ratio affects a firm's score.

The finding is compatible with the importance of short-term financial capacity in accounting screening. An economic explanation involving expenditure, receivables, and payment timing is plausible for some Treasury recipients, but those mechanisms are not measured here. The evidence concerns model reliance on the recorded ratio.

The result also has a direct methodological interpretation. Current Ratio is conceptually related to the working-capital component of Z″, and the leverage predictors can algebraically encode E/TL under the standard balance-sheet identity. The large importance value should therefore not be presented as a wholly independent discovery about the causes of distress.

Net Margin and Quick Ratio have smaller mean reductions of 0.025 and 0.013. Several other predictors have values close to zero, and their reported dispersion is often larger than their mean importance. Correlated predictors can substitute for one another, while unconditional permutation can disrupt normal accounting relationships (Fisher et al., 2019; Molnar, 2022; Strobl et al., 2008). The ranking should therefore not be read as a complete ordering of the economic determinants of distress.

Log Treasury Payments has mean importance of −0.006 and Log Treasury Transactions 0.005. These near-zero values indicate little apparent reliance on the two Treasury features in the explained Gradient Boosting model. They do not establish that the variables are economically irrelevant or that a model retrained without them would perform equally well; that stronger claim would require paired ablation or retraining analysis.

5. DISCUSSION

The first research question is answered by the ensembles' high internal agreement with the Z″ proxy. Random Forest and Gradient Boosting outperform the logistic configuration descriptively on mean PR-AUC and balanced accuracy. Their ability to represent nonlinear boundaries is consistent with this result, but the experiment does not isolate the interactions responsible for the difference. Logistic Regression's strong ROC-AUC also shows that a linear decision function provides substantial discrimination within this predictor set.

For RQ2, the ranking depends on the criterion. Random Forest is more selective, Gradient Boosting has slightly higher mean balanced accuracy and recall among the leading pair, and Support Vector Machine detects the largest share of positive labels while generating many false alerts. This metric-dependent comparison follows the distinction between statistical accuracy and decision value in Lessmann et al. (2015). Huang and Wang (2026) further show why cost-sensitive evaluation matters. Here, no costs are estimated, so the trade-offs do not establish which model minimises expected economic loss.

For RQ3, Current Ratio dominates the Gradient Boosting permutation results, while the Treasury variables have near-zero importance. This answers a question about reliance in the selected model. It leaves open whether another well-performing model would use the same information, and whether excluding Treasury variables and retraining would change performance (Fisher et al., 2019).

The ensemble advantage is directionally consistent with Albanian evidence from Dhamo et al. (2025) and comparisons in other markets (Abrahamsen et al., 2024; Rahman & Zhu, 2024; Nam & Huy, 2026). The emphasis on liquidity is less directly comparable with studies that identify profitability as the strongest signal, such as Nam and Huy (2026). Differences in target construction, selection frame, predictors, and validation timing offer methodological explanations for this contrast; the present design does not test whether the difference is caused by Albania's institutional setting.

The deterministic accounting structure of Z″ explains why high held-out discrimination can arise from recovering an existing boundary. Related liquidity and leverage ratios retain information used to construct the target. As Kapoor and Narayanan (2023) emphasise, predictive validity depends on the information and outcome definition; cross-validation cannot convert a contemporaneous proxy into evidence of subsequent failure.

This comparison clarifies the study's practical contribution: it documents which related indicators reproduce the accounting screen and how classifiers distribute errors. When its components are available, Z″ can be calculated directly without classification error relative to its own rule. The results establish neither an improvement over direct calculation nor validated labels for the four excluded records.

For professional use, the results support examining liquidity and competing alert profiles. Random Forest's lower false-positive count suggests fewer cases for review at the implemented rule, although workloads and economic benefits are unmeasured. Any alert requires verification against the underlying statements, cash-flow capacity, receivable recoverability, and liability maturity. Operational use would require independent validation, explicit

threshold selection, and calibration before interpreting scores probabilistically; an alert alone establishes neither insolvency nor inability to perform a public contract.

The Treasury-payment findings should be interpreted narrowly. The explained model shows little reliance on annual payment value and transaction count once the other predictors are present, but restricted variation within a top-payment sample and coarse annual measurement may contribute to that result. More informative public-finance variables could include payment delays, Treasury receipts relative to revenue, unpaid public receivables, contracting-authority concentration, and dependence on the largest public contract.

The conceptual implication is that explainability can help diagnose overlap between an accounting benchmark and the inputs used to reproduce it. This is useful even when it reveals redundancy rather than a new risk determinant. The present global explanation does not establish individual-case validity, causal direction, fairness, or reliability across sectors.

5.1 Study Limitations

The study is limited to a 2024 cross-section of 96 payment-selected firms, including 18 distress labels. This restricts generalisation beyond major Treasury recipients and leaves temporal stability untested. The outcome is a constructed Z″ benchmark, with partly algebraic predictor–target overlap and zero retained-earnings substitutions for 33 firms. Consequently, the results assess contemporaneous benchmark classification rather than independently observed default. Repeated cross-validation cannot compensate for the small positive class or validate the benchmark's local thresholds.

Missing and approximated accounting inputs, incomplete unit metadata, and annual Treasury aggregates constrain measurement; their effects have not been quantified through sensitivity analysis. Fixed model configurations and internal selection of the explained model also limit comparisons across algorithms. The incomplete computational record described in Section 3.7 prevents full reconstruction of AUC and permutation dispersion. Without external validation, calibration, decision-cost assessment, or retraining with and without Treasury variables, the findings do not establish operational benefits, causal effects, or the incremental predictive value of public-payment information.

6. CONCLUSIONS AND IMPLICATIONS

Among 96 labelled Albanian enterprises drawn from the 100 major State Treasury payment recipients in 2024, Random Forest and Gradient Boosting achieve the strongest observed agreement with an Altman Z″ accounting proxy. Both have mean PR-AUC of 0.943. Gradient Boosting records higher mean balanced accuracy and recall within this pair, while Random Forest records higher ROC-AUC, precision, specificity, accuracy, and F1. The negligible PR-AUC difference provides no basis for declaring one algorithm generally superior.

The comparison shows why minority-class performance requires several complementary measures. Across 288 repeated held-out classifications of the same 96 firms, Random Forest generates 8 false positives and 7 false negatives, compared with 11 and 6 for Gradient Boosting. Support Vector Machine has the highest mean recall but produces substantially more false alerts. These are alternative error profiles whose usefulness depends on costs and decision thresholds not estimated in the study.

Gradient Boosting shows its largest measured reliance on Current Ratio, with mean held-out PR-AUC reduction of 0.541 after permutation. This interpretation remains tied to the liquidity content of Z″ and dependence among financial ratios. The near-zero importance of Treasury-payment value and transaction frequency indicates limited reliance in the fitted model, without establishing their incremental value after retraining.

The study's contribution is to connect accounting-target construction, minority-class evaluation, and model explanation in a payment-selected Albanian sample. It identifies the strongest observed configurations, distinguishes their error profiles, and shows how a dominant importance measure relates to the accounting benchmark. These findings clarify what is learned from reproducing an accounting screen; direct Z″ calculation remains the transparent reference when its inputs are available.

For accounting and risk-review practice, the findings support examination of liquidity and verification of the accounts behind an alert. They do not warrant automatic conclusions about solvency or contract performance. Any operational application requires evidence from independently observed outcomes and an assessment of how errors translate into decisions.

The next empirical step is a longitudinal design with a defined information date and subsequent default, restructuring, or insolvency outcomes. Full retention of code and fold-level scores would permit computational

replication. Sensitivity to the retained-earnings proxy and source-unit assumptions, followed by paired retraining with richer public-payment variables, would distinguish benchmark reconstruction from incremental risk information. Independent temporal evaluation would then assess whether the observed relationships persist beyond 2024.

References

[1] Abrahamsen, N.-G. B., Nylén-Forthun, E., Møller, M., de Lange, P. E., & Risstad, M. (2024). Financial distress prediction in the Nordics: Early warnings from machine learning models. *Journal of Risk and Financial Management, 17*(10), Article 432. https://doi.org/10.3390/jrfm17100432

[2] Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance, 23*(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

[3] Altman, E. I., Hotchkiss, E., & Wang, W. (2019). *Corporate financial distress, restructuring, and bankruptcy: Analyze leveraged finance, distressed debt, and bankruptcy* (4th ed.). Wiley.

[4] Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management & Accounting, 28*(2), 131–171. https://doi.org/10.1111/jifm.12053

[5] Amirshahi, B., & Lahmiri, S. (2024). Bankruptcy prediction using optimal ensemble models under balanced and imbalanced data. *Expert Systems, 41*(8), Article e13599. https://doi.org/10.1111/exsy.13599

[6] Barboza, F., Kimura, H., & Altman, E. I. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications, 83*, 405–417. https://doi.org/10.1016/j.eswa.2017.04.006

[7] Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*, 71–111. https://doi.org/10.2307/2490171

[8] Bengio, Y., & Grandvalet, Y. (2004). No unbiased estimator of the variance of K-fold cross-validation. *Journal of Machine Learning Research, 5*, 1089–1105.

[9] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

[10] Cawley, G. C., & Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research, 11*, 2079–2107.

[11] Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: A systematic literature review. *Artificial Intelligence Review, 57*, Article 216. https://doi.org/10.1007/s10462-024-10854-8

[12] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273–297. https://doi.org/10.1007/BF00994018

[13] Dasilas, A., & Rigani, A. (2024). Machine learning techniques in bankruptcy prediction: A systematic literature review. *Expert Systems with Applications, 255*, Article 124761. https://doi.org/10.1016/j.eswa.2024.124761

[14] Davis, J., & Goadrich, M. (2006). The relationship between precision-recall and ROC curves. In Proceedings of the 23rd International Conference on *Machine Learning* (pp. 233–240). Association for Computing Machinery. https://doi.org/10.1145/1143844.1143874

[15] Dhamo, Z., Gjeçi, A., Zibri, A., & Prendi, X. (2025). Business distress prediction in Albania: An analysis of classification methods. *Journal of Risk and Financial Management, 18*(3), Article 118. https://doi.org/10.3390/jrfm18030118

[16] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

[17] Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research, 20*(177), 1–81. https://jmlr.org/papers/v20/18-760.html

[18] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

[19] Gao, R., Cui, S., Wang, Y., & Xu, W. (2025). Predicting financial distress in high-dimensional imbalanced datasets: A multi-heterogeneous self-paced ensemble learning framework. *Financial Innovation, 11*, Article 50. https://doi.org/10.1186/s40854-024-00745-w

[20] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

[21] Huang, W.-Q., & Wang, L. (2026). Combining sampling methods, cost-sensitive learning, and ensemble techniques for highly class-imbalanced financial distress prediction. *Journal of Forecasting, 45*(6), 2861–2889. https://doi.org/10.1002/for.70162

[22] Hussain, A. B., & Sun, J. (2026). Predicting financial distress: A comparative analysis of explainable machine learning models for cross-market bankruptcy prediction. *Journal of Forecasting*. Advance online publication. https://doi.org/10.1002/for.70193

[23] James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). *An introduction to statistical learning: With applications in Python*. Springer. https://doi.org/10.1007/978-3-031-38747-0

[24] Kapoor, S., & Narayanan, A. (2023). Leakage and the reproducibility crisis in machine-learning-based science. *Patterns, 4*(9), Article 100804. https://doi.org/10.1016/j.patter.2023.100804

[25] Khan, F. S., Mazhar, S. S., Mazhar, K., AlSaleh, D. A., & Mazhar, A. (2025). Model-agnostic explainable artificial intelligence methods in finance: A systematic review, recent developments, limitations, challenges and future directions. *Artificial Intelligence Review, 58*, Article 232. https://doi.org/10.1007/s10462-025-11215-9

[26] Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research, 247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

[27] Liu, W., Lan, X., Xia, M., Zou, Y., Pang, C., & Song, G. (2026). Confidence-scaled margin adaptation boosting for interpretable financial distress prediction. *International Journal of Forecasting, 42*(3), 1008–1032. https://doi.org/10.1016/j.ijforecast.2026.01.003

[28] Lokanan, M. E., & Ramzan, S. (2024). Predicting financial distress in TSX-listed firms using machine learning algorithms. *Frontiers in Artificial Intelligence, 7*, Article 1466321. https://doi.org/10.3389/frai.2024.1466321

[29] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30).
[30] Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable* (2nd ed.). Leanpub.
[31] Nam, T. L., & Huy, T. P. (2026). Explainable machine learning for corporate financial distress prediction in Vietnam. *Cogent Business & Management, 13*(1), Article 2670659. https://doi.org/10.1080/23311975.2026.2670659
[32] Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109–131. https://doi.org/10.2307/2490395
[33] Open Data Albania. (2025, May 12). Lista e 100 bizneseve me më shumë arkëtime nga Thesari i Shtetit 2024. https://ndiqparate.al/?p=24407
[34] Rahman, M. J., & Zhu, H. (2024). Predicting financial distress using machine learning approaches: Evidence China. *Journal of Contemporary Accounting & Economics, 20*(1), Article 100403. https://doi.org/10.1016/j.jcae.2024.100403
[35] Saito, T., & Rehmsmeier, M. (2015). The precision–recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE, 10*(3), Article e0118432. https://doi.org/10.1371/journal.pone.0118432
[36] Song, Y., Jiang, M., Li, S., & Zhao, S. (2024). Class-imbalanced financial distress prediction with machine learning: Incorporating financial, management, textual, and social responsibility features into index system. *Journal of Forecasting, 43*(3), 593–614. https://doi.org/10.1002/for.3050
[37] Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics, 9*, Article 307. https://doi.org/10.1186/1471-2105-9-307
[38] Weber, P., Carl, K. V., & Hinz, O. (2024). Applications of explainable artificial intelligence in finance—A systematic review of finance, information systems, and computer science literature. *Management Review Quarterly, 74*, 867–907. https://doi.org/10.1007/s11301-023-00320-0
[39] Yeo, W. J., Van Der Heever, W., Mao, R., Cambria, E., Satapathy, R., & Mengaldo, G. (2025). A comprehensive review on financial explainable AI. *Artificial Intelligence Review, 58*, Article 189. https://doi.org/10.1007/s10462-024-11077-7
[40] Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial-distress prediction models. *Journal of Accounting Research, 22*, 59–82. https://doi.org/10.2307/2490859

Digital Marketing Opportunities and Risks in Spices and Herbs Market

Andriy POPOVYCH[a*]

[a] *Independent Researcher (Almere, Netherlands, apopovych@kse.org.ua)*

Abstract

The need for proper nutrition and a healthy lifestyle, urbanization, tourism development, and increased migration have influenced the increasing number of people incorporating spices and herbs into their diets. The paper examines the specifics of the spices and herbs market, consumer preferences, and the marketing communications specifics. The article examines the micro- and macro-environment of the trading and consumption of spices and herbs, identifies the main risks and their impact on companies in the industry, and conducts practical analysis of the current market situation and ways of the development of business in this sector. The research methods include collecting data from open sources, such as expert publications, industry reviews, and company activity reports. The main emphasis of the study is the digital marketing as an effective tool in promoting product, placing particular emphasis on social media and influencers. The data obtained allowed us to formulate recommendations for spice and herb producers and sellers, taking into account the needs of the target audience and their preferences regarding purchase location, content, and communication channels.

Keywords: spices and herbs market, digital marketing, target audience, consumer research

1. INTRODUCTION

The growth in purchasing power in Europe, and particularly in Eastern Europe, over the past decades has increased the demand for spices and herbs, while the target consumer audience is still constantly expanding. With growing interest in culinary experimentation, natural products, and ethnic cuisine, comes the interest and necessity in analyzing the spice and herb market.

Spices are defined as various aromatic and flavoring additives used in cooking to enhance the flavor of dishes or alter their texture. Medicinal and aromatic herbs are whole, ground dried plants, or their leafy parts, possessing a persistent, distinctive aroma and flavor, used to flavor dishes and prepare medicinal infusions. An essential and mandatory requirement for spices is that they must be used only in pre-processed and dried form, although herbs can be used both fresh and dried.

In recent decades, the demand for spices and herbs is steadily growing worldwide due to changing gastronomic habits, the development of international trade, and the popularization of various cuisines. These factors are driving producers and suppliers to seek ways to meet the diverse needs of consumers and develop products that meet their expectations.

* Corresponding author

Given that market of spices and herbs is a complex undertaking incorporating three sectors: agricultural, industrial, and marketing, the greatest emphasis of market and scientific research is placed on specific agricultural issues. Primary attention in publications is given to cultivation based on the application of fertilizers and organic matter, increasing yields and oil content, enriching fruits with valuable components, but insufficient attention is paid to the marketing component while losing focus on the global market.

Since changing socioeconomic conditions and the introduction of scientific and technological advances are reflected in various markets, including the spices and herbs market, the purpose of this study is to identify development patterns in the global spice trade. The research program includes three areas: studying market dynamics, the competitive environment, and the target audience; investigating key promotion methods; and analyzing existing marketing reports.

In the spice and herbs market, priority is shifting from traditional marketing communications channels to digital ones: contextual and targeted advertising, viral advertising, social media promotion, and relationship marketing. Under these conditions, the ability to identify current trends and predict changes in demand, adapt product ranges, and develop effective marketing strategies is key to successful business in this industry. The goal was to understand key consumer preferences in the spices and herbs market and identify the role of digital marketing communications in spice and herb sales, using market of Eastern Europe as an example.

The market's size and dependence on external suppliers make Europe an attractive target market for suppliers from developing countries. Europe is a mature and diverse spices and herbs market, home to two of the world's ten largest importing countries. European countries with the greatest potential include the Netherlands, Germany, Spain, the UK, Poland, and France. Overall demand in Europe is relatively stable, and supply volumes are large. With a population of nearly 530 million consumers, Europe imports more spices and herbs than the Americas combined. According to the Harmonized Index of Consumer Prices (HICP) for European Union (EU) markets, the price index for salt, spices, and culinary herbs was 134.21 points in August 2024 (Eurostat), compared to 115.18 points in October 2022. This means that consumer prices for these goods have increased by 16.5% in less than two years.

Europe relies on imports from developing countries, as the majority of spice imports into Europe come from Asia. The situation for herbs is different, as most herbs consumed in Europe are also produced in Europe. In the short and medium term, European import volumes are expected to remain relatively stable, with possible limited growth due to the role of spices and herbs as salt substitutes and flavor enhancers in plant-based foods. Eastern Europe accounts for 8% of European imports, with Poland accounting for over 50% of the region's imports from developing countries, followed by the Czech Republic and Bulgaria. However, the majority of the region's imports come from the rest of Europe. Imports from outside Europe are primarily from East Asia (dried chili peppers and paprika, as well as ginger from China) and Southeast Asia (primarily peppers from Vietnam)[1].

2. CURRENT STATE AND DYNAMICS OF THE GLOBAL SPICES AND HERBS MARKET

2.1. Spices and Herbs Trade: History and Today

In recent decades, the spices and herbs market has become a truly important element of the modern food industry. Demand for spices is steadily growing due to rising affluence, changing household eating habits, the development of international trade, increased tourism and migration, and the popularization of diverse global cuisines. These factors are driving producers and suppliers to seek ways to meet the diverse needs of consumers and develop products that meet their expectations.

The origins of the spice trade date back to the third millennium BC in ancient China, India, and Egypt. By 2000 BC, it had spread through South Asia and the Middle East, forming one of the first global food trade networks. During the Middle Ages, the spice trade was a driving force behind the development of international commerce and colonization. Since the 1960s, the global trade in spices and herbs has grown exponentially, with the value of international trade (in nominal terms) increasing 41-fold, outpacing the overall growth of food trade [1].

In 2024, the global market for spices and herbs reached $24.51 billion and is projected to reach $34.17 billion by 2030 (see Fig. 1), growing at a CAGR of 5.69% [1]. The most important spice-producing countries, among the top ten exporters, include Bangladesh, Pakistan, Iran, Nepal, Colombia, Ethiopia, and Sri Lanka. In terms of both individual and mass use in the food industry, the most popular spices are black pepper, cinnamon, ginger, and basil. This is due to their high nutrient content, affordability, and widespread availability worldwide. The Asia-Pacific region (APAC) is the fastest-growing market and holds the largest share of the global spices and herbs market, with

[1] GloballyCool based on UN Comtrade, September 2024

India being the world's largest producer, consumer, and exporter of spices. The region is one of the fastest-growing markets for spices and herbs (see Fig. 2), fueled by a rich culinary heritage and long-standing traditions of spice production. The main drivers of growth in the APAC region are the expanding middle class and rising demand for convenience and processed foods [2, 3].

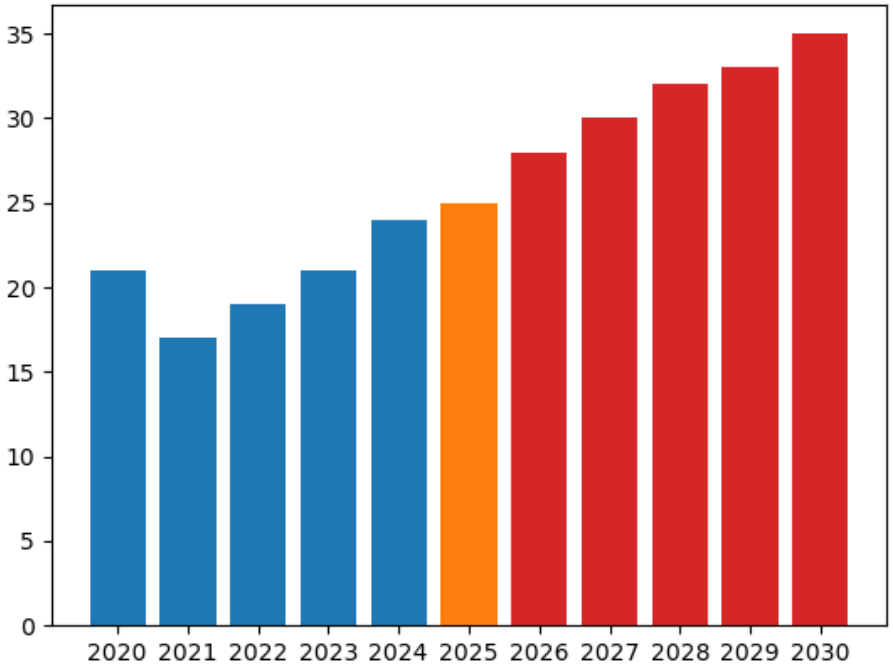


Fig. 1. Dynamics of Global Market for Herbs and Spices

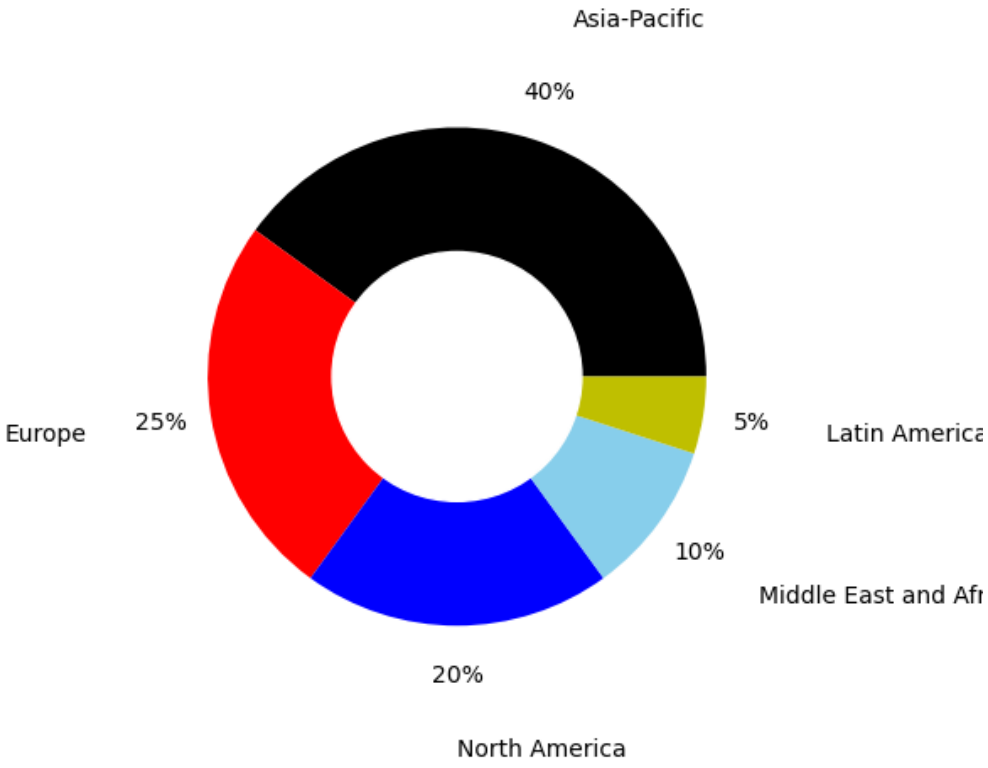


Fig. 2. Geography of Herbs and Spices Consumption

Black pepper is the most widely consumed and popular spice in the world, followed by other versatile spices and herbs such as chilies, turmeric, garlic, cinnamon, etc. (see Fig. 3)

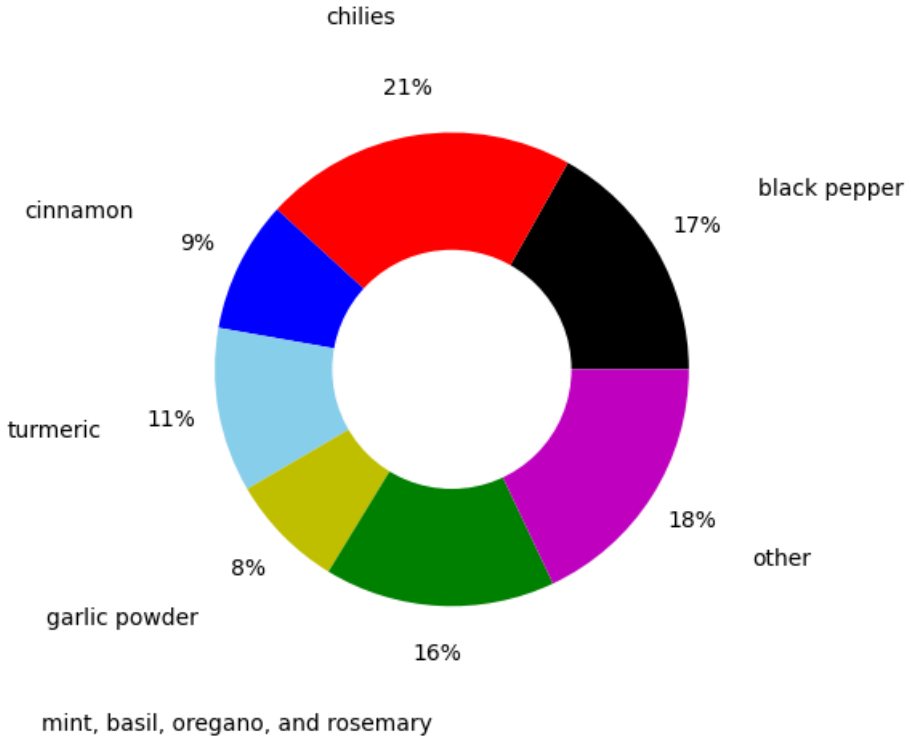


Fig. 3. Global Trade in Herbs and Spices

2.2. Spices and Herbs Trade: Drivers of Development

Several drivers are identified as key trend indicators for market development in terms of consumer demand. They are modern trends in health and wellness and the health benefits associated with spices, urbanization and a growing preference for organic food consumption, demand for authentic and exotic cuisine, continued consumer interest in national dishes due to the expansion of the hospitality industry, and growing demand for convenient food products.

The health and wellness trend is driving the natural ingredients market overall by increased production of healthy foods and growing demand for them. Products claiming to contain natural ingredients, such as spices, help manufacturers attract more customers. Consumers are turning to natural alternatives as lifestyle-related diseases such as cardiovascular disease, obesity, osteoporosis, and diabetes rise all over the world, and they increasingly view spices as functional ingredients that offer health benefits beyond simply altering the flavor, aroma, and enhancing the culinary quality of foods.

By creating new flavor properties in foods, spices enhance the physiological effects of food on the digestive system, promoting its better absorption, as they contain significant amounts of vitamins, minerals, and other beneficial substances. Furthermore, spices have a distinctive property of inhibiting bacteria (i.e., they are bactericidal), primarily putrefactive bacteria, and thus help extend the shelf life of the product (acting as a preservative). Spices also have a beneficial effect on metabolism, the nervous system, and the cardiovascular system. Most have the ability to stimulate the elimination of various types of toxins from the body, acting as filters that cleanse the human body of mechanical and biological contaminants. For this reason, a significant number of spices are, and were, used in the past primarily in medicine as pharmaceutical substances.

The greatest growth in the market for such products is observed in large cities, which is explained by several factors. First, a higher standard of living, which ensures greater purchasing power and, consequently, consumer loyalty to more expensive products. Second, a higher consumer culture and environmental awareness among the population. Third, the higher popularity of a healthy lifestyle among urban residents, amid the challenging environmental situation in large cities.

Today, natural ingredients are preferred worldwide because the growing awareness of harmful ingredients and additives in processed foods is pushing consumers to use organic products propelling demand for the organic spices

to grow significantly faster than for conventional spices. In response, manufacturers are expanding their product ranges to include "natural," "eco-friendly," and "organic" products, actively embracing the healthy lifestyle trend by highlighting naturalness on the front of the packaging. The implementation of European Union Regulation (EU 2018/848) has led to the implementation of comprehensive certification protocols, creating a favorable environment for trusted suppliers who have invested in sophisticated traceability systems [4].

Increasing migration and the expanding hospitality industry are transforming consumer gastronomic preferences. Inspired by social media content and international travel experiences, consumers are increasingly eager to try new cuisines and flavorful dishes prepared with a variety of spices and herbs.

Consumers driven by busy lifestyles are seeking faster, easier-to-prepare and more convenient meal options. The convenience, prepared and processed foods industry is growing as consumers increasingly seek restaurant-quality dishes in a home-cooked format. Additionally, consumer spending has increased due to the higher income and lifestyle changes, the growing number of women in the workforce, and the penetration of savory snacks, soups, noodles, beverages, and ready-to-eat foods into retail channels, particularly in developed countries. The growing demand for convenient foods is driving the consumption of spices and herbs. Spices and herbs play a critical role in enhancing the flavor of convenience foods, including ready-to-eat meals, frozen foods, and snacks. Manufacturers are capitalizing on this opportunity because spices not only enhance flavor but also simplify cooking, reducing preparation time [5].

Overall, the future development of the spices and herbs market will be increasingly focusing on cost-effective strategies that combine value propositions for the target audience in terms of quality, price, and service convenience. To strengthen their position among their key competitors, leading market players should be maximizing efforts on the adaptation of their management activities to consumer preferences. This underscores the need for various marketing researches, particularly identifying the macro environment and consumer choice priorities.

2.3. Porter's Five Forces in Spices and Herbs Market

Analyzing the macro environment of the spices and herbs market using Porter's model allows identifying the main risks and threats facing companies in this segment and understanding its structure. Currently, this methodology is the primary tool for assessing the level of competition in the industry and developing an effective competitive strategy. The model evaluates five key competitive forces that influence a company's profitability: the intensity of competition among existing players, the threat of new entrants, the threat of substitutes, the power of suppliers, and the buyers' power [6].

The factor of competition among existing competitors assesses the intensity of competition among existing companies in the industry, for example, through price reductions or marketing campaigns. The factor of competition in the spices and herbs market is intense and increasingly influential, as companies strive to differentiate themselves through unique advantages and strengthen their positions while high marketing expenditures increase the intensity of competition.

The threat of new entrants' factor assesses the ease with which new companies can enter the market and compete with existing players. The threat of new competitors in the spices and herbs market is low, as supply is sufficient. Barriers to entry include branding and establishing relationships with large customers, making it difficult for new companies to adapt to the new environment and cope with the high pressure from existing companies. On the other hand, unique advantages are easily imitated, requiring constant monitoring of market competition and maintaining and developing the superior qualities of one's own product. However, this threat increases over time, as the market can be attractive to new companies due to growing demand.

The impact of the substitute product factor, which assesses consumers' ability to find alternative products or services that can satisfy the same needs, is low. The low threat from substitute products is because the market generally offers a wide range of spices, so substitute products are few. The main substitute products are traditional additives, flavorings, and salt, but they are not complete substitutes for spices and herbs, either in terms of functionality or in terms of environmental friendliness. On the other hand, in certain geographically localized markets, certain types of spices are interchangeable and can compete with each other.

The suppliers' power, that is, the ability of suppliers to influence prices for resources or services required by companies in the industry, is high, as most companies in the market rely on imported raw materials, allowing them to dictate terms. Suppliers have advantages over purchasing companies due to the following market factors: supply instability due to import and export restrictions, the sensitivity of logistics prices to changes in raw material prices in the country of origin, and companies' high dependence on imported spices and herbs. Supplier influence is persistent, as companies have already adapted to unfavorable market conditions by entering into long-term, mutually beneficial agreements and maintaining good relationships.

The buyers' power defines consumers' ability to influence prices and conditions in the industry and its influence is high and growing, as consumers have a wide choice among sellers. Despite all the socioeconomic difficulties in recent times, market supply has not diminished, allowing consumers to choose from a wide range of options for purchasing spices and herbs, and the presence of intense competition in the market increases buyers' opportunities.

Table 1. Porter's Five Forces in Spices and Herbs Market

FORCES	STATE
Existing Competition	high
Threat of New Entrants	low
Threat of Substitutes	low
Suppliers' Power	high
Buyers' Power	high

Source: Author

A market threat analysis using Porter's model revealed that the main risks in the spices and herbs market are related to intense competition within the industry, strong supplier power, and changing consumer preferences (see Table 1). Consumer and competitor influence continues to grow, which may ultimately lead companies to rethink their internal strategies, while supplier influence remains constant.

A positive aspect of intense competition in the global spices and herbs market is that it fosters innovation, forcing companies to develop new, unique blends that meet varying consumer preferences. Technological advances are driving this trend, making it possible and feasible to develop new products with more favorable taste profiles, targeting health-conscious consumers seeking convenience.

2.4. Consumer as a Center Point of Spices and Herbs Market

To understand the trends in the spices and herbs market it is crucial to consider current consumer behavior, sentiment, interests, and expectations, as well as the characteristics of the target audience. Despite differences in gender, age, and profession, the target audience for spices and herbs sellers shares some similarities. Key characteristics and interests of the target audience include an interest in a variety of culinary recipes, a desire to use high-quality products, a preference for organic and natural ingredients, and a desire to add unique flavors and aromas to their dishes.

Consumers in the spices and herbs market can be classified according to whether they are B2C (business-to-consumer) or B2B (business-to-business). When considering the B2C market, four main categories of consumers are distinguished:

1. Family households account for the majority of consumers in the B2C market. Spice buyers are primarily women, who are responsible for cooking at home. They prefer to buy traditional herbs that have been tested by personal experience. However, they also occasionally buy something unusual in small quantities, allowing themselves to experiment with different spices to diversify their home menu.

2. Professional chefs, a rarer category of customers who know a lot about spices, purchase spices and herbs mostly for work-related purposes.

3. Experimental gourmets: this category are prone to experimentation and "take risks" with new culinary creations, as they are usually preparing a dish for themselves. These consumers actively use social media, follow popular culinary bloggers, and frequently visit restaurants and cafes, where they gain knowledge and ideas for their dishes. Consumers in this category often pay attention to the cuisines of different nations.

4. Connoisseurs of national dishes: many representatives of different nations and cultures who immigrate continue to preserve their traditions and customs, which manifests itself in culinary aspects and preferences. Different nations and cultures have their own national dishes with unique recipes, which can motivate people to purchase spices and herbs common in their country of origin.

Consumer segmentation in the B2B market identifies the following customer categories:

1. Retail stores and chain stores: This category represents the largest share of consumers in the B2B market. This category also includes wholesale customers (grocery stores, grocery departments, hypermarkets, etc.)

2. Catering establishments: these include cafes, restaurants, restaurants specializing in national cuisine, etc. For these establishments, products are packaged in larger volumes – in bags and boxes.

3. Prepared food producers: These prefer traditional spices and herbs, as their product range represents familiar and everyday dishes for the middle and lower segments of the population, as well as owners of small confectioneries, etc.

The trade in spices and herbs has many unique aspects compared to the trade in most agricultural products, as there are significant geographic distances between the most important suppliers and buyers. Therefore, international trade in these products requires long-distance transportation and a complex logistics system. Furthermore, transportation networks are underdeveloped in many countries, and many countries require warehouses and storage facilities for all the spices produced. The organic spice market continues to face significant supply constraints worldwide, as does the storage of all the spices produced [3, 7].

Some spices and herbs are characterized by seasonal consumption because each spice type having its own characteristics, but there are spices that are in demand regardless of season. In addition, there is also seasonal harvesting, which in turn influences the price of spices and herbs. Therefore, during peak delivery periods, the market experiences a surplus of raw materials, while before the next harvest there can be a shortage. Consequently, price fluctuations occur during certain periods due to changes in supply.

Overall, the spices and herbs market continues to face significant supply constraints worldwide due to several factors. Among them, there are the strict and complex regulatory requirements in individual countries and regions, limited availability of organic and sustainable raw materials as certified organic since farmland accounts for only a small proportion of global spice cultivation area, technical challenges in developing new spice blends and issues with shelf life, stability and sensory quality [7].

Currently, the main trend in the spice, condiment, and herb market is the growth of the mid-price and premium segments. Rising incomes are driving consumers to higher-priced products. Demand for higher quality, more expensive products has grown significantly, and consumers have become interested in premium products. Meanwhile, manufacturers previously operating in the lower price segment have begun to expand into the mid-price and premium segments, in particular by improving design and packaging quality, but most importantly, by improving the quality of raw materials. Therefore, selecting reliable suppliers is becoming increasingly important and crucial for building a business in this category [7].

3. MARKETING AT THE SERVICE OF SPICES AND HERBS

3.1. Spices and Herbs as Marketable Product

Marketing is particularly important in the food distribution industry, of which the spices and herbs market is an important component. The food industry experiences quick market response to marketing efforts, as it is directly linked to consumers and their needs. Therefore, it is crucial for marketing in this area to study consumer preferences for functional foods, based on the patterns that define the specifics of the wholesale and retail trade of spices and herbs [8]. The main factors that influence consumer behavior in the food market are pricing, labeling and packaging, marketing strategies, innovation in agribusiness, market segments, social aspects, labeling, and branding [9].

The specific characteristics of spice include its small market supply volumes (the rational consumption level for the population is no more than 6 grams per day per adult), its high logistical expenditures while adhering to packaging requirements, standard warehouse space requirements, and low delivery costs due to the product's small size. Therefore, regional consumption volumes are primarily determined by the size and age composition of the population.

From a marketing perspective, all spices and herbs are complex commodities that are difficult to sell because of impulse demand and low turnover. The primary promotional method is to offer products to as many retail outlets as possible, ensuring a consistently high product presence on store shelves in each region. The widespread belief among the spices and herbs traders is that other marketing efforts are only worthwhile once the majority of their retail outlets are stocked.

Marketing mix is a set of elements that form the core of a company's marketing concept and that must be developed in order to enter the market and promote its products and services. The choice of each element depends on external conditions, the market situation, and the company's internal situation. In other words, marketing helps businesses deliver the right product to the right place at the right price at the right time. This system was developed in the second half of the 20th century, is abbreviated as the 4P model (from the first letters of each key element), and includes product, place, price, and promotion. From a 4Ps perspective, spices and herbs as a product have both strengths and weaknesses [10].

The strength lies in the fact that spices are a trendy commodity. The use of spices and herbs is very popular today, both at the household level and in the food industry and catering, and experimenting with them is trendy. Spices have a high retail markup of 70 to 100%, but some stores charge much higher markups, targeting the premium segment [11].

3.2. Spices and Herbs Business Specifics of Trade and Promotion

The spice and herb trade is a highly specialized business that typically has few direct competitors and few entrepreneurs dare to launch. The breadth and uniqueness of its product range makes it easy to differentiate itself from stores, markets, and retail outlets selling regular spices.

The downsides and challenges of selling spices and seasonings stem from the fact that they are ancillary products, purchased as an add-on, rather than a necessity. With diminished purchasing power, consumers may easily forego unusual ingredients, focusing on basic options. Even without direct competition, stores and supermarkets selling the most popular spices will always be in abundance. Dealing with suppliers can be challenging, as not all spice suppliers are reliable and capable of meeting all needs.

Business format and location are crucial in the spice trade, although it is generally a space-saving business, and large stores are rare. Therefore, the primary permanent options are a retail pavilion or a market outlet. The last one is the simplest retail option, with the lowest costs for rent, equipment, and advertising, and it can be opened in a grocery store or a section of a supermarket. Another location option is a department, kiosk, or island in a shopping center. Ideally, market stalls should be located near the entrances of large anchor grocery hypermarkets, as spices complement the main purchases. It is also possible to set up shop through a small, stand-alone store, usually located close to grocery stores. Since most people do not specifically travel for spices, the location should be connected to grocery shopping with a high footfall [12, 13].

Price levels influence the formation of sales value and volume. Undoubtedly, the global price of a commodity is a significant indicator of growth in production and export, reflecting demand both in the global market as a whole and in the market of a specific country or segment. The main factors determining spice prices are quality and compliance with consumer preferences. It should be noted that the quality and quantity of the products under consideration are significantly influenced by natural climatic and annually changing weather conditions, which are sometimes impossible to predict or mitigate.

Spice and herb prices are influenced by factors such as the rarity of the plant (narrow habitat, specific climatic conditions), labor-intensive harvesting (manual labor, short flowering or fruiting periods), complexity of processing (lengthy fermentation, drying, sorting), and limited supplies, as some countries strictly control exports. The high price of spices significantly limits producers' acceptance of natural ingredients. While the health and wellness trend encourages the use of these natural ingredients, the cost associated with many spices can be prohibitive for both producers and consumers.

The choice of advertising and promotional methods primarily depends on the target audience, where they live, shop, and their preference for purchasing spices therefore anything related to food or a new culinary experience is best for advertising. A must-have option for a store is outdoor advertising or signage, as well as signage in nearby high-traffic areas. Handouts such as brochures and flyers—can be distributed for any occasion, such as openings, discounts, or sales [12].

Spice stores can distribute samples, use loyalty programs, use partnerships or cross marketing. This involves establishing partnerships with related businesses: restaurants, cafes, culinary schools and courses, grocery stores, and fair and food festival organizers. The owner can exchange advertising links and materials with the partners, create a joint promotion, or host a joint event such as fairs and gastronomic festivals and tastings, as well as master classes for adults and children [12].

The common saying "the product sells itself" does not quite apply to spices. From a distance, all spices, although colorful, look similar in some ways. A salesperson should not simply present spices from the shelf and allow customers to appreciate their aromas but they should also act as a professional consultant on their use, a kind of guide to the world's cuisines [13].

Economic and political factors, in addition to social and demographic factors, have the most profound impact on the behavior of modern consumers and food buyers. To facilitate food product selection, buyers pay attention to individual product attributes (or indicators). For example, information about the place or country of production and processing, as well as the manufacturer, replaces information about product quality and social acceptability.

Information about the place of production is of great importance for imported products. On the other hand, the reputation of a place or a country of production can have a negative impact on importers. Therefore, when entering a market or any other country, foreign manufacturers must consider the reputation of their country, which has already been formed in the minds of consumers and is characterized either positively or negatively. This is also important when analyzing and evaluating an organization's position in comparison with its main competitors in the market, since, for example, most products from domestic manufacturers have a significant advantage in the minds of buyers, but in less developed countries, foreign manufacturers are more popular, and their products are in greater demand.

4. DIGITALIZATION OF THE SPICE AND HERB TRADE

4.1. Information Technologies for Marketing of Spices and Herbs

With rapidly evolving technologies, different perspectives on the scientific and creative components of marketing are emerging while marketing tools and methods are undergoing profound changes today. With the advent of new technologies, marketing, previously dominated by advertising, is now defined by digital channels of interaction and measured by both modern strategies and technological solutions, which testifies to the development of the scientific aspects of marketing [14].

The modern consumer prefers communication in the digital space, while the global internet environment is open and adaptive. Consequently, traditional marketing communications are partially losing their relevance, while digital channels are becoming increasingly popular. This necessitates a company's constant adaptation to changing environmental conditions and consumer preferences when using marketing communication channels to communicate with and influence consumers.

The digital environment offers new ways to communicate with target audiences. Digital communications, also known as "new media," are becoming increasingly prominent. The key characteristics of "new media" are the presence of an online environment, interactive interaction, and equal participation in marketing communications [15]. These include websites, email marketing, online advertising, online communities, mobile apps, electronic kiosks, digital cameras, AR and VR experiences, and more. With the development of the internet, the communications model has evolved into a fundamentally different form of communication between companies and consumers. A distinctive feature of the digital communications model is the selectivity of choice and the ability to maintain relationships between parties involved in the communication process.

Integrated marketing communications in "new media" offer several advantages, including global reach, multi-effective impact on the target audience, clear audience segmentation, personalized communications, targeted messaging, dialogue and feedback with the target audience, the availability of tools for analyzing consumer behavior, and the expansion of communication channels and formats, which impacts the promotion of all areas of activity and products. Consequently, the aforementioned traditional marketing tools are actively migrating into the digital environment, following changing consumer preferences [15].

Consumers themselves determine which channels they prefer to receive information through and which ones to use when obtaining the information they need. The form of the message itself is also changing becoming more customized and personalized. The new generation of consumers, using digital communications, can independently produce and distribute content about the company.

For businesses, the use of new technologies offers the opportunity to sell spices and herbs online, especially since the experience of the coronavirus pandemic has forced many people to increasingly order online. Although at present most buyers still prefer to experience spices in person at brick-and-mortar stores, but the popularity of e-commerce is growing every year.

4.2. New Media Tools for Spices and Herbs Trade

Social media, direct marketing, websites, along with the development of mobile apps and e-commerce platforms, are considered the most promising areas of digital marketing in the spice and herb industry. A key requirement for this is a well-balanced spice and herb online store website filled with unique and engaging content for its customers, ranging from an effective design to a convenient checkout process. Website elements include a striking sales video, a user manual, a detailed description of all the company's services and benefits, a brand history, and a platform for a

corporate blog where all the ingredients of a gourmet presentation that simply would not fit into a traditional commercial proposal can be included. Spices are also a good topic for filling a website or blog with useful, engaging information such as recipes, master classes, life hacks, shoot videos, post photo carousels, and much more, while simultaneously selling the products and creating user activities.

Online retail options include online spice stores, spice stores on marketplaces, and social media marketplaces. In the first case, many people learn about spices online and are ready to make a purchase there. There are numerous marketplaces where sellers can present their products. Furthermore, a great option for small shops is to display their product range by creating business accounts on popular social media. It is important to note that these options are not mutually exclusive and can, and ideally should be used in combination [13].

The key stages of transitioning to online spice retail include: market research, defining the business's specifics, identifying the target audience, developing a product range, ensuring inventory warehousing, selecting the necessary equipment, developing a promotional strategy, and preparing for store opening [15]. Preparing to open an online spice and herb store includes choosing a platform for the store, registering a domain and hosting, creating a design and logo, setting up a payment and delivery system, and creating informational pages and contact information [13]. Creating an online store does not require significant investment, and the cost depends on the service package. It will be more expensive if ordering a turnkey website, with development, content, and subsequent SEO promotion to the top of search engine results.

Promoting an online spice store requires developing proprietary marketing strategies to attract customers through online promotion and active use of social media. To promote an online store selling spices and herbs, it is recommended using contextual advertising, social media advertising, creating high quality and unique content, cross marketing, and offline promotion through partnerships or participation in events. Interactivity and customer engagement through surveys, contests, and promotions are important to encourage customers to be active and interact with the online store.

Online advertising has significant advantages over other marketing communications tools because it allows for a quick response to market conditions and 24/7 access to information. The relevance of using online advertising as one of the basic marketing communications tools for an online spice store is determined by the dynamics of the digital economy [17].

For social media advertising, thematic communities on food and cooking are best suited for posting ads there but ensuring, that it is useful and interesting to the audience since users usually reject straightforward advertising [17]. Targeted advertising should be directed at an audience of a specific age, gender, and interests while contextual advertising should be aimed at promoting the web resource. Direct marketing requires selecting keywords and creating ads to display the searched combination in the top search results.

Customers value expertise that can be verified. Simply calling someone an expert has become so fashionable that it is no longer taken for granted, but only a competent person can write a meaningful article with a minimum of fluff [18]. There are culinary public pages and communities where culinary topics are actively discussed, and advertising is possible. It is important to collaborate with popular bloggers who have a large audience and influence their followers' opinions. However, it is important to remember that advertising should not just be a link, but interesting and useful content that will attract attention and generate interest among followers [18]. Collaboration with food bloggers to advertise either the entire store or a specific product by offering their audience personalized discount codes for purchases. A blog also offers unobtrusive care: posts about corporate gifts offered with a promo code or for an action and can be maintained not only on a website but also on social media platforms like Instagram or Facebook.

Successful promotion of an online store cannot be limited to online promotions alone. Offline promotion, using colorful flyers with recipes, appearing on a local culinary program, writing articles for specialized media, hosting food tasting events, and distributing small samples of spices and herbs so customers could try them out proved very effective in attracting new customers and increasing brand awareness. Collaborating with other entrepreneurs and conducting joint marketing events is a great way to reach new audiences. Teaming up with chefs, cafes, restaurants, and grocery stores to host joint events such as master classes and food festivals. Even simply sharing information and links on social media has already helped attract new customers who have not yet heard of the store.

5. META-ANALYSIS OF EAST EUROPEAN SPICE AND HERBS MARKETING RESEARCH

5.1. Methods and Materials

The methodology of this study includes an analysis of consumer surveys, open-source data and publications by industry experts. The materials include analytical reports from marketing agencies, consulting companies, and

marketplaces. Data on changes in supply and demand, consumer preferences, and the main factors influencing the development of this segment were also used as materials for the analysis. These reports contain information on spice and herb sales, changes in product range, and demand for various product categories.

Several marketing studies related to mainly Eastern European region were selected to identify key characteristics of the current stage of development of the spices and herbs market, as well as those influencing decision-making in traditional marketing methods and the development of new digital product promotion channels. Based on available reports, the results of marketing research on consumer preferences for spices and herbs were examined, their naturalness and organic origin, competitive advantages in this market segment, and public attitudes toward using spices and herbs. The focus is on the domestic spices and herbs markets of Poland, Romania, and other Central and East European countries. Furthermore, one of the analyzed materials is a report on the results of a project to digitalize the marketing activities of a particular company.

As a rule, the analytical method in these studies was quantitative research based on surveys, collecting primary information by directly asking consumers the questions about their level of knowledge, attitudes toward the product, preferences and purchasing behavior, as well as market performance and development characteristics. Each study is presented as a separate case due to the unique characteristics of each report, and conclusions are drawn from the aggregated information.

The meta-analysis East European spice and herbs marketing research includes four studies, here investigated as cases.

5.2. Case #1: The Herb and Spice Market in Poland: Development Prospects

In 2024, a marketing study of the Polish spices and herbs market was conducted [19] (see Fig. 4) using a formalized online questionnaire, with 190 respondents participating. Of the total number of end consumers, men and women, aged 20 to 65 were represented in approximately equal proportions. The most common choices were ground black pepper, paprika, ginger, basil, and English herb.

Respondents were primarily divided between those who purchase spices less than once a month (59%) and those who purchase them once or more than once a month (32%). Furthermore, some respondents (9%) said they do not purchase spices at all. Among the places they use to purchase spices, convenience stores were the most frequently cited (62%). Chain supermarkets were the second most popular choice (57%), while marketplaces were the least popular (26%). The most preferred type of spice, based on its geographical origin, was Mediterranean spices (39%). Indian (21%) and European (20%) were also common choices.

Ground spices were the favorite (73%), followed by blends (67%) and dried herbs (55%), while extracts and essences were the least popular (15%). Among spice preferences, the most common use was for cooking meat dishes and baking (44%) while less common were marinades and drinks (34%). However, a significant portion of respondents (19%0 indicated that they add spices to almost every dish. Regarding the packaging of purchased spices, the most preferred volumes were 100-200 g (49%) and 10-100 g (19%), with slightly less than a third (32%) purchasing spices in 200-500 g volumes. The most popular packaging formats were glass jars and foil-lined kraft paper bags, with more than half of respondents preferring these (53%). Glass bottles with dispensers came in second (31%), while kraft paper bags with zip-locks were less popular (13%). Paper bags and plastic jars with dispensers were also mentioned (3%).

Respondents place greater trust in sources such as the opinions of family and friends (45%) and information on packaging (35%). The survey also found that many obtain information about spices and herbs from point-of-sale advertising (22%) and online (15%).

Country of origin is becoming important when choosing spices and herbs (70%) and consumers prefer domestic producers to foreign ones (64%). The importance of the country and region of production, manufacture, and origin of food products for Polish consumers when choosing and purchasing spices and herbs is similar to that for the food market as a whole. For the majority of respondents (63%), the choice of manufacturer is directly dependent on the type of food product. It is important to note that some respondents choose food products from domestic producers in order to support them (20%).

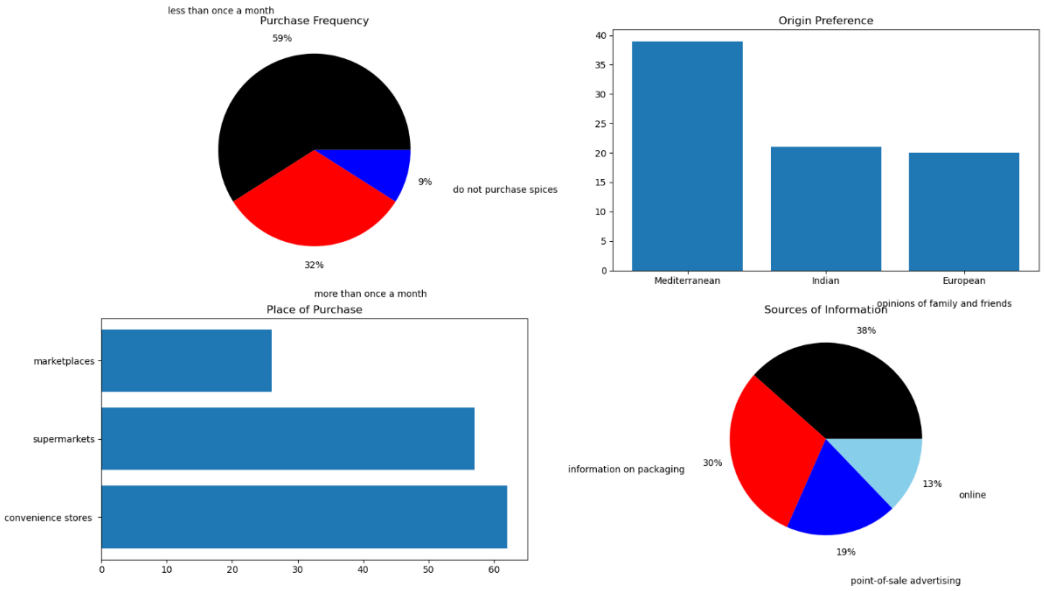


Fig. 4. The Herb and Spice Market in Poland: Development Prospects

5.3. Case # 2: Market study on the spices and aromatic herbs market in Romania

According to a marketing study of the spices and herbs market in Romania [20] (see Fig. 5, Fig. 6), respondents' responses were distributed as follows: 36% buy them from time to time, 9% buy them several times a week, 19% buy them several times a month, and 11% almost never buy them. Spice mixes account for the largest share of total spices, herbs, seasonings, and condiments purchased, accounting for 46% of the total, black pepper ranks second with 44%, and ginger third with 22%.

More than two-thirds of survey respondents (70%) cited health benefits as the primary reason for purchasing spices and herbs. Respondents also see their purchase of spices and herbs as a concern not only for their health but also for the environment (40%). Respondents prefer to purchase healthy food products in grocery stores (80%), specialty stores (30%), pharmacies (20%), markets (20%), online stores (10%), and other retailers (5%). Consequently, most consumers stick to traditional shopping methods and prefer mass-market products.

According to the majority of respondents, online and television advertising, as well as reviews from friends and acquaintances, are their top priorities for attracting attention to spices and herbs (33%, 26%, and 18%, respectively). Outdoor and print advertising are the least important to the target audience, despite their widespread use by businesses (7% and 4%, respectively). Therefore, traditional advertising campaigns should be reoriented toward integrated marketing communications, implemented primarily online.

Analyzing the results, it can be concluded that when choosing spices and herbs, Romanian consumers primarily consider price (48%) and product quality (39%). Fifteen percent of respondents are interested in the product's eco-friendliness. Ten percent of respondents rely on the opinions of friends and acquaintances. Consumers are least interested in the presence of a safety seal on the packaging of spices and herbs and brand recognition (5% of respondents each noted both factors).

Next, respondents selected spices from a list provided and also indicated others as free responses. Spice selection by geographic region: African spices (e.g. harissa, guinea pepper seeds (grains of paradise), korarima, berbere, dukka, njangsa seeds), American spices (e.g. paprika, vanilla, cocoa, jalapeño chili, etc.) 14%, European spices (e.g. dill, parsley, horseradish, savory, etc.) 20%, cumin, nutmeg, etc.) 6%, Asian spices (e.g. star anise, ginger, chili, coriander, etc.) 12%, Indian spices (e.g. cardamom, turmeric, curry, black pepper, etc.) 19%, Mediterranean spices (e.g. oregano, basil, rosemary, bay leaf, etc.) 29%.

Data for marketing research on the importance of factors when choosing spices was collected through an online survey among visitors to the vom-achterhof.de website [21]. There are 1,419 respondents participated in the survey, 65% of whom were women and 35% men. 56% of respondents resided in cities, while 34% resided in rural areas. Respondents were grouped into five age categories, along with education and monthly income.

Respondents were asked to rate the importance of these factors on a scale of 1 to 5, with 1 meaning no influence and 5 meaning a significant influence. Respondents were asked to rate the importance of various factors in their spice purchasing decisions, including price, packaging design, packaging type, brand, appearance, naturalness, complete manufacturer information, packaging method, and quality certificates.

The survey revealed that the main factors influencing spice selection include ingredients, packaging convenience, and brand, with average scores of 4.5, 3.9, and 3.1, respectively. Price, however, remains an important but not decisive factor, with a score of 2.9. Country of origin has an average impact of 2.3 points for most respondents. If the price of a foreign product is lower than that of a domestic one (assuming the same quality), 75% of respondents would choose the foreign product, meaning country of origin is a less significant factor than price.

Weight and quality certificates have a weak influence on consumers when purchasing spices and herbs compared to other factors. However, respondents cited additional criteria such as freshness, aroma, taste, and naturalness as factors determining their choice of spices and herbs.

A chi-square test evaluating the relationship between variables revealed no differences in consumer expectations based on gender and age. The observed percentage of respondents for whom individual attributes of spices and herbs are unimportant is typical for the market introduction stage, when the consumer experience of a product is still unknown or unclear to consumers.

The final question in the survey asked respondents about their understanding of "high-quality spices." Respondents were also given a five-point ranking system for the importance of various factors characterizing this concept. Freshness, aroma, and taste ranked first, with over 80% rating these factors at 4 and 5. Naturalness and packaging methods ranked second and third, respectively. Complete information about the manufacturer was found to be the least important, with only a third of respondents asking about this information.

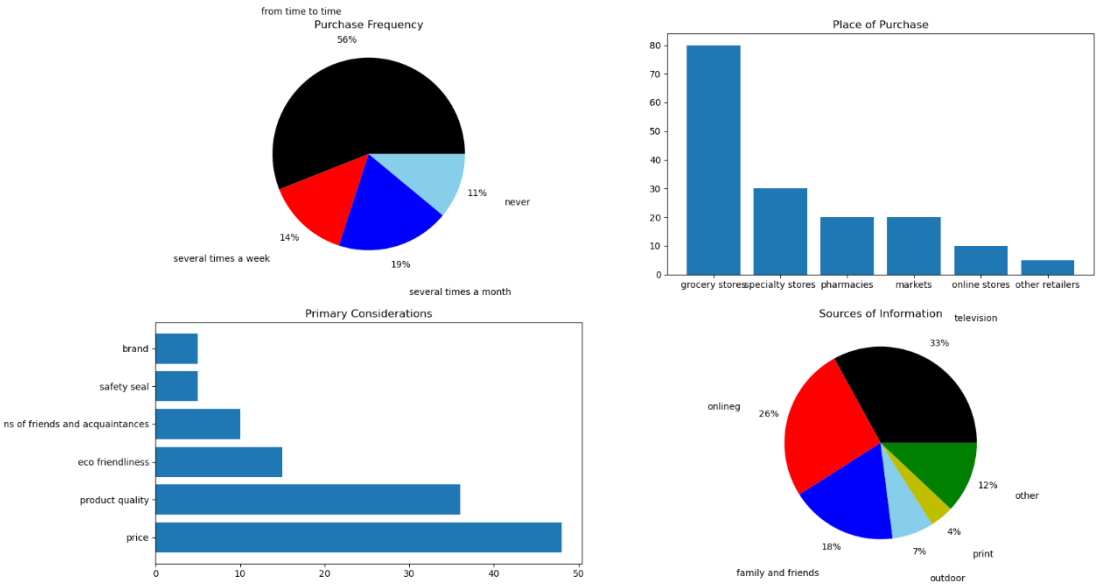


Fig. 5. Market study on the spices and aromatic herbs market in Romania

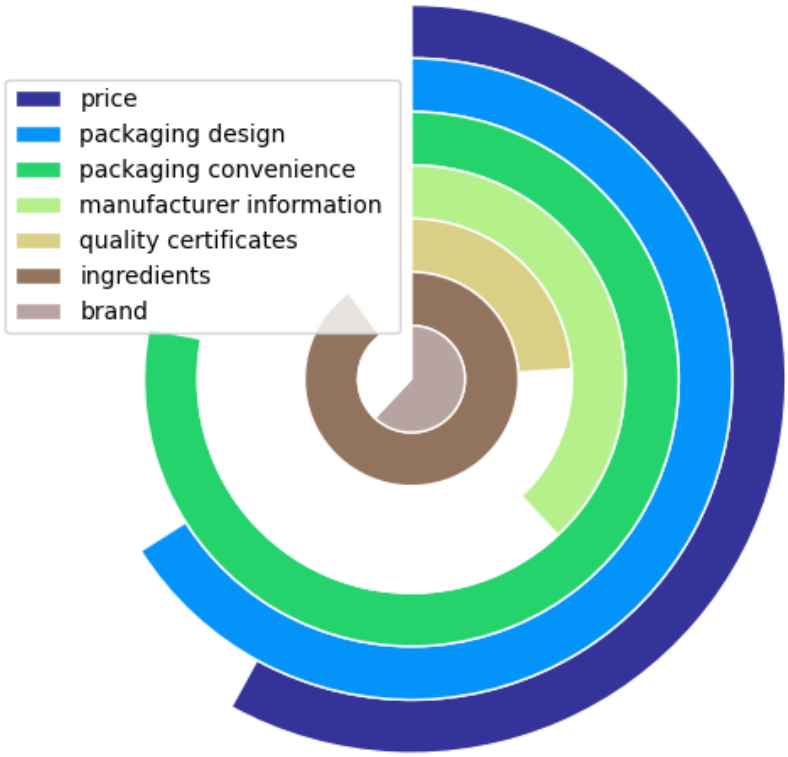


Fig. 6. Importance of Factors When Choosing Herbs and Spices

5.4. Case # 3: Key Characteristics of the Largest Companies in the Spices and Herbs Market

A marketing agency conducted a study to identify the key drivers of competitive advantage by monitoring offerings at retailers with online product catalogs [22]. In 2023, the initial sample of stores offered 65 spices and herbs across 21 brands. When comparing the results with the 2020 study data, it was noted that 11 previously available brands had disappeared from the surveyed retailers' product ranges, while 8 new brands had appeared. Therefore, the sample showed a reduction in the number of product names and brands in 2022 compared to 2021.

The study included an analysis of the key characteristics of the largest companies in the spices and herbs market. The results indicate that the main advantages of these manufacturers are expressed in the following: a large number of partners, long-term presence in the market, a large number of employees, and a global regional presence. Among the main competitors in the spice and herbs market, Kamis is particularly strong in terms of partner count, with over 2,000 partnerships. Other leading companies also have a wide network of partners, typically over a thousand.

Leading companies in the spice and herbs market typically have over 20 years of experience. While the oldest company in the market has been present for over 100 years, a record of accomplishment of over 30 years is not particularly significant.

The number of registered employees is an indicator of a company's size and an important competitive advantage. The most successful companies have staff numbers ranging from several thousand to tens of thousands.

Regional presence (primary regions) is also an important competitive factor, and leading global companies are represented in markets ranging from fifty to hundreds of countries. Knorr and Maggi, whose products are distributed in 90 and over 100 countries, respectively, demonstrated the greatest regional advantage.

Based on the comparative analysis, KOTANYI, Kamis (McCormick & Co), and Maggi (Nestle) emerged as the most significant competitors in the market. They boast extensive experience and longevity in the European market, a broad partnership base across the market, which inspires confidence in their campaigns, and a wide product range.

However, Maggi strengthens its advantage with delivery, a loyalty program, and various purchasing methods. KOTANYI and Knorr do not offer delivery services, which is a significant disadvantage in the current conditions of doing business in this market. However, KOTANYI's lack of delivery is a significant drawback. While this doesn't prevent the company from maintaining a competitive advantage, it is not particularly welcome given current trends in the spices and herbs market. The company also has significant potential for geographic growth. KOTANYI is

currently present in 53 countries worldwide, which is unimpressive compared to Maggi's regional presence (over 100 countries). The same applies to Kamis, whose operations extend to over 50 countries. A loyalty program, delivery options, and various purchasing methods also make a company attractive. Generally, all large companies offer purchasing methods such as wholesale and bulk purchasing.

A study of the market positioning process revealed that, depending on the chosen target segment, manufacturers position their brands as organic (40% of the products available at the studied retail outlets) or as non-organic (60% of the products).

Rarely does a company stand out with its loyalty program. Most global companies limit themselves to standard discounts and promotions, where loyalty programs include discounts for orders over a certain amount, discounts for bulk purchases, and promotions for retail purchases, such as a percentage discount for orders over a certain amount. For example, Podravka attracts customers with the following "scheme": 5% off orders over €50 and 10% off orders over €100.

Regional companies are not as popular as their global competitors are. The primary regional presence of European companies is focused on the spices and herbs market of a single country or several countries in the region. Prymat stands out, operating in over 10 countries despite its relatively short history (>10 years), relatively few partners (>300), and relatively small headcount (±100). The other regional companies are not significantly different from each other, and in many respects, they are almost identical.

At the same time, regional companies have greater potential for growth; in a relatively short period of existence, they have expanded their operations to over 20 countries. While most companies focus on large-volume deliveries, as evidenced by their loyalty system aimed at economies of scale, companies with smaller performance indicators can focus on niche market segments or improve their performance across all key criteria to compete more successfully.

Thus, based on the competitor analysis presented above, it can be concluded that the main competitors currently are KOTANYI, Kamis (McCormick & Co), and Maggi (Nestle), thanks to their extensive experience and market popularity. Knorr and Prymat have opportunities to develop competitiveness, but only in the premium market segment. Smaller regional companies can focus on niche segments for further development and increase their potential for competitive growth.

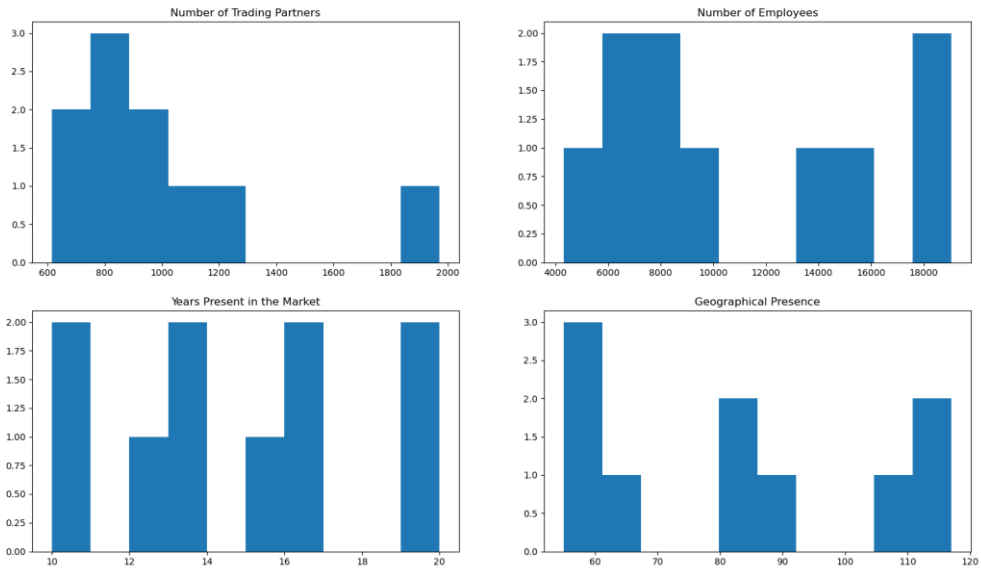


Fig. 7. Key Characteristics of the Largest Companies in the Spices and Herbs Market

5.5. Case # 4: Strategic Marketing Plan for Spices and Herbs Trading Firm

The one of the leading trading company on Polish spices and herbs market ordered the complex investigation of its marketing actions and development and support of a strategic marketing plan. The marketing agency developed

the brand's strategy and visual concept for social media, with a focus on Instagram. The marketing agency also assumed full responsibility for managing the company's accounts, which share recipes with users, announce new products, host contests and themed marathons, implement engaging mechanics, communicate with subscribers via comments and private messages, and mitigate negative feedback.

The marketing agency developed a social media strategy that, over the course of a year, attracted over 12 million unique users to all projects. In addition to social media, the agency is responsible for all the brand's promotional campaigns, which are held seasonally in collaboration with major retailers and include prize draws. Over the course of a year, the agency organized five promotional campaigns in major retail stores and oversaw all processes, from the landing page to processing participant receipts and prize delivery.

The agency implemented influencer marketing and integrates video and photo recipes from both million-follower bloggers and micro-influencers specializing in cooking and travel. As a result, the company collaborates monthly with YouTube and Instagram bloggers as part of an SMM strategy, which helps increase consumer loyalty and generates additional traffic to the website and social media.

The marketing agency developed and launched the brand's website, continuously monitors and updates the content, and offers relevant technical solutions, improvements, and integrations. The agency also runs BTL campaigns in collaboration with food retailers, organizing and conducting samplings, highlighting the benefits and unique properties of the company's products.

In preparation for the exhibition, the agency designed and assembled an eco-style stand, striving for a combination of modernity, eco-friendliness, and comfort. The area included coffee zones, meeting spaces, product display stands, a utility room, and a branded reception desk. Promoters were invited to conduct an advertising campaign not only at the stand but also throughout the exhibition.

The agency developed and launched a mobile app so users could learn more about the brand, dishes that can be prepared with the company's products, select favorite recipes, and add their own. The mobile app also makes it easy to search for dishes by ingredients, share recipes with friends, and find groups based on their interests. All recipes included integrated brand tips and a "Where to buy?" feature, allowing users to find the nearest store that sells the company's spices.

The agency also implemented numerous other services, including email newsletters to databases, social media and forum analytics for specific queries, direct advertising through banner networks, targeted advertising, and video ad placement via pre-rolls on ad networks. Additionally, the agency implemented special projects including placement on culinary portals, user contests and cashback promotions, and focus groups to research consumer opinion before launching new products. The agency also developed the design and branding of souvenirs, promotional uniforms, and a product catalog.

6. RESULTS AND DISCUSSION

An analysis of the conducted research and the identified trends led to the following key conclusions. Primarily, they suggest that modern consumers prefer high quality, natural products. European, Mediterranean, and Indian spices are particularly popular among European consumers, but the popularity of spices and herbs from other regions is gradually growing, driven by the popularization of various global cuisines.

The analysis revealed key trends in consumer preferences in the spices and herbs market. The primary criteria for consumers choosing spices are quality, naturalness, aroma, and convenient packaging. In the retail market, packaging plays a significant role, directly influencing price. The price of the same product in different packaging can vary significantly. Brand plays a secondary role, indicating the importance of a product's organoleptic properties over its marketing positioning.

The results of the study also confirm that the spices and herbs market is a dynamically developing segment in which consumers demonstrate clear preferences when choosing products. Thus, it was established that the overwhelming majority of respondents purchase spices and herbs with varying frequency, from several times a week to once in a month. Furthermore, the survey found that three-quarters of respondents who do not purchase spices and herbs would like to do so. This suggests that the spices and herbs market has significant potential for sales growth. Furthermore, an analysis of purchase preferences by frequency revealed that most consumers purchase spices once a month or less. This suggests that spices are not essential products, but their choice is significantly important to consumers.

It is interesting to note that ground spices and dried herbs are the most popular forms, while extracts and essences occupy a less significant market share. This may indicate that consumers prefer traditional methods of using spices in cooking.

Another important trend is the choice of packaging formats: glass jars and kraft paper bags with a foil layer proved to be the most popular. Research has shown that when choosing spices and herbs, modern European consumers pay attention to the country of origin, but the location and country of origin are not particularly significant when choosing a particular product or product group. Currently, European consumers place greater trust in domestic brands and trademarks. When choosing foreign-made products, European consumers generally prefer EU producers due to their high product quality, as well as the brand's reputation and trustworthiness.

Products from Asia, Africa, and the Americas are little known to European consumers, but informing European buyers and consumers about the composition and quality of products produced in these regions can help change purchasing behavior and incorporate products from these countries into their purchase patterns.

The European spices and herbs market is seeing both the emergence of new brands and the exit of existing brands. However, the overall number of brands remains relatively stable.

Macro-environmental factors, such as inflation, also influence consumer choice and overall behavior. Price is a determining factor, and a study of consumer expectations regarding individual attributes revealed that the majority of respondents consider it necessary or desirable for spices to be cheaper. Therefore, when promoting the sale of spices and herbs online, the emphasis should be on maintaining reasonable prices, high-quality service, and convenient operating hours.

This is why it is essential to continually research consumer preferences, both by the marketing system to achieve the manufacturer's primary goals and by the government to regulate the economy and the market. Similarly, when developing new products, given the dynamism of innovative markets, marketing research must be conducted regularly. Identified changes in the expectations and behavior of key market participants should be taken into account when substantiating marketing decisions related to positioning, pricing, and product differentiation to successfully launch a new product.

Today, marketers have a wide variety of promotional tools at their disposal. The primary tools are advertising, sales promotion, PR, and direct marketing. The effectiveness of traditional tools is declining, while their costs are rising. The high costs of traditional marketing tools reduce their appeal to managers. Social media platforms such as Instagram, Facebook, and Telegram have become particularly popular, as have opinion leaders—those who are listened to by their audience. In today's competitive environment, a trend toward creative promotional tools, such as viral and mobile marketing, scent marketing, and others can be observed. When developing and planning a promotion strategy, several marketing communications tools, both basic and creative, are used in combination.

Competitors quickly identify and copy high-quality content that garners a lot of likes and shares. As a result, investments in creating high-quality content will ultimately benefit the competition. A solution to this problem is to meaningfully tie the content created to the promoted product, service, or brand. However, this approach is difficult to implement, as it narrows the range of possible publications. For each specific company, brand, and market situation, a spice promotion campaign takes on specific characteristics based on numerous factors, including targeting target groups through contextual and targeted advertising, social media, and point-of-sale promotions such as promotional materials, merchandising, and expert advice from a salesperson or specialist, which can be delivered through direct marketing.

The decision to purchase spices and herbs depends on expert advice from a salesperson or other authoritative source of information, as a lack of information can lead to improper use of spices and herbs, resulting in dissatisfaction and even health problems. Therefore, the use of direct marketing in the digital space is a relevant promotional tool. Advertising on a specialized website for chefs is of little interest to consumers. At the same time, for most consumers, lack of direct contact with the product is the main reason they choose not to purchase it online. In this case, each manufacturer needs to combine digital marketing and retail, paying close attention to marketing channels that combine these two areas.

7. CONCLUSION

The spices and herbs market continues to grow and attract new players, particularly in the premium, natural, and organic segments. With the right market knowledge and sufficient resources, companies focused on quality and innovative solutions can gain a competitive advantage and strengthen their position in this segment.

The growing popularity of healthy lifestyle trends, environmental concerns, and the increased use of chemicals in agriculture are increasing public concern about environmental issues and the rising cost of medical treatment. As a result, the population is increasingly demanding healthy foods, which is driving demand for spices and herbs.

Saturation in the spice and herbs market is leading to a decline in the number of small producers and retailers and increased competition, primarily through increased advertising. Even small firms have realized that brand building requires investment in product promotion. This trend will continue in the future, and marketing expenditures will

continue to rise. In a competitive and constantly changing market environment, marketing research is a key focus. The results of research subsequently form the basis for developing sales estimates, and, based on these, for projected revenue and profit levels. Market research is essential for the success of products offered on the market as it enables companies to monitor changing consumer needs.

Proposals to mitigate competitive and consumer risks include brand development and increased brand loyalty. Other possible risk mitigation options include optimizing logistics processes to minimize costs by allowing companies to offer a more competitive price compared to competitors by minimizing costs and expanding the company's product catalog to better attract consumer attention. Thus, by systematizing existing approaches taking into account the specifics of innovative markets, it is necessary to regularly conduct marketing research in three key areas during the development of new products: market research, key market participants (both competitors and consumers), and process research (positioning, pricing, distribution, etc.). Manufacturers can use the research results and retailers to optimize product ranges, improve packaging quality, and develop marketing strategies.

Furthermore, the specifics of consumer demand largely depend on the gastronomic habits and cultural preferences of the population. This can form the basis for further research in the field of personalized marketing targeted at different consumer groups. In modern marketing practice, given the active integration of digital tools into the research process, the most common research methods are consumer, market, competitor, and individual marketing mix tools.

The digital environment gives modern consumers the opportunity to interact with manufacturers and sellers at any convenient time, offers personalized communication, and allows consumers to become equal participants in the communication process.

Best practices for promoting online spice and herb retail include contextual advertising, which attracts traffic; social media advertising, which provides the opportunity to reach a target audience and attract new customers; engaging content that attracts attention and generates interest; cross-marketing, which attracts new audiences; and offline promotion, which helps raise awareness of spices and herbs and increase sales. It is important to pay attention to survey results, analyze feedback, and optimize advertising campaigns accordingly. Placing simple advertising links without engaging content or publishing trivial content should be avoided. Image-building activities are an important promotional channel: sponsorship of culinary projects on video services and in electronic and traditional media, point-of-sale promotions, and sampling (distribution of free or trial product samples). All this helps build a customer base and increase consumer loyalty to your products through ongoing advertising.

References

[1] Food And Beverages Global Market Report 2025, Research and Markets, https://www.researchandmarkets.com/report/food-beverage#product

[2] Food and Beverages Market Size and Forecast 2025 to 2034 Precedence Research https://www.precedenceresearch.com/press-release/food-and-beverages-market-size

[3] Food & Beverage Market Report 2025 (Global Edition) Cognitive Market Research https://www.cognitivemarketresearch.com/food-beverage-market-report

[4] European Union, "Regulation (EU) 2018/848 Of The European Parliament And Of The Council," eur-lex.europa.eu]

[5] NACS, "Global Convenience Markets," convenience.org.

[6] Bruijl, D. G. H. T. (2018). The relevance of Porter's five forces in today's innovative and changing business environment. Available at SSRN 3192207.

[7] Europe Seasoning And Spices Market Size & Share Analysis - Growth Trends and Forecast (2025 - 2030) https://www.mordorintelligence.com/industry-reports/europe-seasoning-and-spices-market

[8] Flores, P. J., & Jansson, J. (2022). SPICe—Determinants of consumer green innovation adoption across domains: A systematic review of marketing journals and suggestions for a research agenda. International journal of consumer studies, 46(5), 1761-1784.

[9] Martinho, V. J. P. D. (2020). Food marketing as a special ingredient in consumer choices: the main insights from existing literature. Foods, 9(11), 1651.

[10] Hanifah, P., Nuringwahyu, S., & Krisdianto, D. (2022). Effect of 4P'S Marketing Mix on Purchase Decision. Marg. J. Manag. Accounting, Gen. Financ. Int. Econ. Issues, 1(4), 207-218.

[11] What is the demand for spices and herbs on the European market? (n.d.), CBI Europe, https://www.cbi.eu/market-information/spices-herbs/what-demand

[12] Your Guide to Starting a Spice Business Successfully (n.d.), Levapack, https://www.levapack.com/starting-a-spice-business-guide

[13] Opening a Spice Shop (n.d.), Spiceandtea, https://www.spiceandtea.com/blogs/franchising/starting-spice-company

[14] Scott Brinker (2013). What do you mean by marketing as a science? Chiefmartec https://chiefmartec.com/2013/03/what-do-you-mean-by-marketing-as-a-science

[15] Adiyono, N. G., Rahmat, T. Y., & Anindita, R. (2021). Digital marketing strategies to increase online business sales through social media. Journal of Humanities, Social Science, Public Administration and Management (HUSOCPUMENT), 1(2), 31-37.

[16] Vitsentzatou, E., Tsoulfas, G. T., & Mihiotis, A. N. (2022). The digital transformation of the marketing mix in the food and beverage service supply chain: A grey DEMATEL approach. Sustainability, 14(22), 15228.

[17] Alamsyah, D. P., Ratnapuri, C. I., Aryanto, R., & Othman, N. A. (2021). Digital marketing: Implementation of digital advertising preference to support brand awareness. Academy of Strategic Management Journal, 20(2), 1-10.
[18] Andriana, R. (2025). The Influence of Digital Marketing and Brand Trust on Customer Loyalty Through Customer Satisfaction of Food and Beverage Companies. International Journal Of Education, Social Studies, And Management (IJESSM), 5(1), 214-224.
[19] Rynek Ziół i Przypraw w Polsce: Perspektywy Rozwoju (2024) SW Research Agencja badań Rynku i Opinii https://swresearch.pl
[20] Studiu de piață al pieței condimentelor și ierburilor aromatice din România (2023) The Markers Consult, https://www.themarkers.ro
[21] Gewürze und Kräuter in Europa: Eine Online-Verbraucherumfrage (2024), Achterhof, https://vom-achterhof.de
[22] Agencja KS, https://ks.pl
[23] 4822, https://4822.pl

From Digital Crack Images to Mathematical Models: A Box-Counting Activity for Secondary Mathematics Education

Tetiana ZHYLENKO[a*]

[a] *Institute of Mathematics Education, University of Cologne, Cologne, Germany,*

tetiana.zhylenko@uni-koeln.de

Abstract

Digital image analysis offers a context for connecting mathematical modelling, computational thinking and interdisciplinary STEM learning. This paper presents an activity in which an irregular crack image is treated as a real-world object to be mathematised rather than as a materials-science result. Students move from the original image to a binary representation, perform box counting at several scales, organise the data and, as an extension, interpret the slope of a log-log representation as an estimate of fractal dimension. Physics provides the real phenomenon, mathematics provides scale, functional relationships and quantitative interpretation, and informatics provides image representation, thresholding, iteration and automation. The activity was preliminarily introduced to a regular mathematics-club group of 15 students in Grades 9–10 in a German university outreach context. The program and its mathematical basis were demonstrated, after which the students received a task based on the same principle. The session was exploratory rather than a controlled empirical study, so no claims about learning gains are made. The design therefore focuses on making the transition from manual exploration to software-supported analysis transparent and on avoiding black-box use of digital tools.

Keywords: mathematics education; mathematical modelling; box counting; digital images

1. INTRODUCTION

Mathematical modelling offers a natural framework for connecting school mathematics with phenomena that are initially encountered outside mathematics. Learners have to identify relevant features, simplify a situation, construct and work with a mathematical representation, and interpret the result in relation to the original context (Blum et al., 2007; Niss & Blum, 2020). In interdisciplinary STEM settings, these transitions are especially important because the same phenomenon can be viewed through different disciplinary lenses: physics explains and questions the phenomenon, mathematics provides quantitative structures for describing it, and informatics enables representation, algorithmisation and repeated computation (Maass et al., 2019).

Digital images are particularly suitable for such interdisciplinary work. They originate from observable objects, but they are already computational representations composed of pixels and numerical values. This makes otherwise hidden modelling decisions visible: an object must be selected, classified, discretised, measured and interpreted.

* Corresponding author

Digital tools can support exploration, visualisation and repeated calculation, yet they can also conceal the mathematical operations that generate an output. The resulting black-box problem is well documented in technology-supported modelling and motivates learning designs in which students reconstruct essential operations before relying on automation (Geiger et al., 2010; Cevikbas et al., 2023; Engelbrecht & Borba, 2024).

The starting point for the activity is a digital image of an irregular crack. It is useful precisely because the same image can be discussed from three perspectives—physics, mathematics and informatics. The crack provides an authentic physical phenomenon and raises questions about shape, branching and complexity, but the physical fracture mechanism is not taught or modelled in detail. Instead, students investigate how the visible structure can be translated into a binary representation, measured at different scales and described quantitatively. Box counting provides the central mathematical procedure: it links geometry and scale with numerical data, functional dependence and, as an enrichment component, logarithmic transformation and the interpretation of slope. The corresponding computational workflow introduces thresholding, connected structures, iteration and automation.

The activity was developed for an extracurricular mathematics club with students in Grades 9–10 in Germany. The contribution of the proposed design is not the box-counting method itself, but its didactical reconstruction as an interdisciplinary learning sequence. A research-related image-analysis procedure is opened up into steps that secondary-school students can inspect, reproduce and discuss, connecting a physical phenomenon with mathematical modelling and computational processing.

The activity was first demonstrated with students in Grades 9–10 in the mathematics-club setting described above. The demonstration was exploratory rather than a controlled empirical intervention; consequently, no claims about learning gains are made. The contribution is therefore a design-oriented paper. It asks how an existing computational image-analysis workflow can be reconstructed as a transparent learning sequence, how the three disciplinary perspectives can be coordinated without turning the activity into a specialised materials-science lesson, and which aspects are suitable for a later systematic classroom study.

2. LITERATURE BACKGROUND AND DIDACTICAL RATIONALE

Mathematical modelling has developed into a substantial field of mathematics-education research. The ICMI study edited by Blum et al. (2007) and the synthesis by Niss and Blum (2020) describe modelling as a cyclic activity involving transitions between an extra-mathematical situation and a mathematical model. Blomhøj and Jensen (2003) conceptualise modelling competence as the ability to carry through a whole modelling process and distinguish several necessary sub-competences, while Niss and Højgaard (2019) place modelling within a broader framework of mathematical competencies. International perspectives also differ in the purposes assigned to modelling, ranging from pragmatic applications to educational and epistemological goals (Kaiser & Sriraman, 2006). Empirical research has shown that transitions within the modelling process are demanding for le arners (Schukajlow et al., 2018). Consequently, modelling competence is not reducible to successful calculation; it includes interrelated cognitive and metacognitive activities. A systematic review by Cevikbas, Kaiser and Schukajlow (2022) confirms the diversity of approaches used to conceptualise, measure and foster these competencies.

For classroom design, this means that students need opportunities to experience the modelling process rather than only its final mathematical form. Scaffolding can be useful when it structures the process without predetermining every decision. For example, Schukajlow, Kolter and Blum (2015) showed how a solution plan can support students in modelling tasks. In the activity proposed here, the sequence image - representation - grid - count - graph - interpretation plays a related orienting role, but the representations are generated from digital image data rather than from a conventional word problem.

The representation of the real-world situation is itself didactically significant. Jablonski (2023) demonstrated that students' modelling activities can differ depending on whether they work with real objects or representations of those objects. A digital image is therefore not a neutral carrier of information: it is already a representation in which spatial reality has been converted into pixels. Converting the image again into a binary matrix makes another modelling decision visible. For learners, this creates an opportunity to discuss what is preserved, what is discarded and why different representations may lead to different measurements.

Digital technology adds both possibilities and risks to this process. Cevikbas, Greefrath and Siller (2023), in a systematic review of 38 studies, report benefits including support for exploration, conceptual understanding, feedback and the investigation of authentic problems, while also identifying technological competence and black-box use as challenges. Earlier classroom research similarly showed that digital technologies can provoke productive mathematical interaction in modelling tasks rather than merely serve as calculation devices (Geiger et al., 2010). Research on modelling with digital tools also shows that technological actions can enter several phases of a

modelling cycle, including mathematising, calculation, visualisation and validation (Greefrath et al., 2018; Koyunkaya & Dede, 2024). A recent systematic review further emphasises the need to connect particular digital-tool activities explicitly with modelling phases and sub-competencies (Riemer & Greefrath, 2026). From a broader perspective, Borba and Villarreal (2005) conceptualise mathematical thinking as reorganised through interaction with media, and Engelbrecht and Borba (2024) describe recent shifts in mathematics education associated with digital learning spaces, STEAM and technology-mediated activity.

For this reason, the software is introduced only after the students have worked with the basic idea themselves. Students should not begin with a displayed value of fractal dimension. They should first encounter the image, propose possible measures of complexity, inspect a binary representation, count occupied boxes for selected scales and organise the resulting data. Automation can then extend the number of scales and enable comparison. In this design, software is introduced after key mathematical decisions have become discussable, reducing the risk that the final numerical result is treated as an unexplained property produced by a program.

The box-counting method is especially suitable for such a transition because it links an intuitive visual property - irregularity across scales - with a reproducible numerical procedure. Fractal geometry provides the theoretical background for scale-dependent complexity (Mandelbrot, 1982), while computational box-counting methods make the idea applicable to digital images (Li et al., 2009). For Grades 9–10, the educational goal is not a formal study of fractal geometry. Instead, box counting serves as a context in which students can examine how N(s) changes with scale s, organise pairs of values, reason about functional dependence and, where appropriate, explore a logarithmic transformation and the slope of a log-log graph.

The activity also creates a natural connection between mathematics and informatics. Thresholding can be discussed as a classification rule, a binary image as a matrix of discrete values, connected-component selection as an algorithmic procedure, and repeated box counting as iteration over scales. At the same time, these computational steps retain mathematical meanings that students can inspect. This is consistent with the broader view that digital technology in mathematics education is most productive when representation, experimentation and interpretation remain intertwined rather than being separated into 'mathematics' and 'software' phases (Borba & Villarreal, 2005; Engelbrecht & Borba, 2024).

The present design also draws on earlier work with digital environments for mathematics learning, where cloud-based tools and automated feedback were used to support mathematical activity (Zhylenko & Zakharchenko, 2013; Zhylenko et al., 2020). The choice of an irregular branching structure also has a scientific precedent in studies of hierarchical and statistically analysed branching structures (Olemskoi et al., 2012a, 2012b). These earlier studies are used here only to motivate the digital and visual context; the didactical rationale of the activity is grounded in mathematics-education research.

A further rationale comes from interdisciplinary STEM education. Mathematics can function as a connecting language between scientific phenomena and computational representations, particularly when modelling is used to cross disciplinary boundaries (Maass et al., 2019). In the present activity, physics supplies the real-world phenomenon and questions about crack formation and branching; mathematics supplies scale, functional dependence, logarithmic transformation and fractal dimension; and informatics supplies the representation of the image as data, thresholding, connected-component analysis, iteration and automation. Research on computational thinking in K–12 mathematics similarly describes productive mathematics–computing integration as a reciprocal process in which mathematical reasoning is used to construct and interpret computational outputs (Ye et al., 2023).

3. COMPUTIONAL BACKGROUND USED FOR THE LEARNING ACTIVITY

The learning design was derived from a digital image-analysis workflow developed for quantitative crack analysis. In the underlying computational procedure, an input image is converted to grayscale, thresholded to obtain a binary representation, and cleaned by retaining the relevant connected component. The selected structure is then analysed by box counting. For a grid scale s, the number N(s) of boxes intersecting the structure is determined; the operation is repeated for several scales. A log-log representation is subsequently used to estimate a box-counting dimension from the slope of an approximately linear region.

Only this image-analysis part is used in the educational activity. Mechanical parameters, fracture models and the physical interpretation of the crack are not included. This separation is intentional: the same computational sequence could be applied to another irregular visual structure, while the mathematical questions about representation, scale and functional dependence would remain.

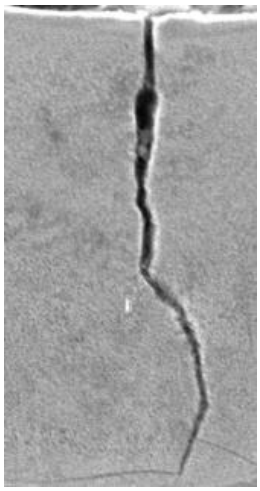


Fig. 1. Original digital image of the crack used as the starting point of the modelling activity

4. COMPUTIONAL BACKGROUND USED FOR THE LEARNING ACTIVITY

The first didactical step is to postpone the formula and begin with the image. Students are asked how the complexity of the visible line could be described by numbers. Typical mathematical possibilities can be discussed without deciding immediately which is "correct": length, occupied area, maximum width, number of turns, number of branches, or a measure that depends on observation scale. The purpose of this opening is to create a need for modelling decisions.

The image is then represented as a binary array B(x,y) with values 0 and 1. At this point the image can be discussed as a discrete mathematical object. A threshold is not merely a technical setting; it is a rule that decides which pixels belong to the object. Connected-component selection raises a second question: if isolated dark pixels are present, should they be counted as part of the crack? Thus a seemingly simple image-processing step introduces classification, neighbourhood and the distinction between data and a chosen representation.



Fig. 2. Binary representation of the same crack after image segmentation

5. BOX COUNTING AS SEQUENCE OF STUDENT ACTIONS

After the binary representation has been discussed, a square grid is placed over the structure. Students count the boxes that contain at least one part of the object. The same operation is repeated with a finer grid. Before using software, two or three scales can be counted manually or semi-manually. This makes the dependence on scale visible and gives meaning to the notation N(s).

Table 1 illustrates how the worksheet can be used during the activity. The grid sizes and the structure of the table are prepared by the teacher, whereas the entries shown in red are completed by the students. Students count the occupied boxes and enter their results; in an extended version, they also calculate the logarithmic values. The numerical entries are illustrative values used to demonstrate a completed worksheet and are not reported as empirical data from the preliminary mathematics-club demonstration.

Table 1. Illustrative student worksheet for recording box-counting data. The table structure and grid sizes are provided by the teacher; entries shown in red are completed by the students. Numerical values are illustrative and are not empirical classroom data

Grid size s (px)	Occupied boxes N(**s**)	Log(1/s*)	Log N(s)
64	6	1.386	1.792
32	15	2.079	2.708
16	36	2.773	3.584
8	88	3.466	4.477
4	210	4.159	5.347

The transition from the first two columns to the last two is a mathematically important step. The original relation between scale and occupied-box count is not linear. A logarithmic transformation makes it possible to examine whether a range of scales can be represented approximately by a straight line. In that range, the slope can be interpreted as an estimate of the box-counting dimension D:

$$D \approx \Delta \log N(s) / \Delta \log(1/s^*).$$

To obtain a linear representation of the scale-dependent relationship, both variables are logarithmically transformed. In the numerical example, Log denotes the natural logarithm. The transformed data are represented as Log(1/s*) versus Log N(s). For a power-law relationship, this transformation produces an approximately linear dependence, and the slope of the fitted line is used to estimate the box-counting dimension D. The logarithm base does not change the estimated slope when the same base is used consistently for both coordinates (Fig. 3).

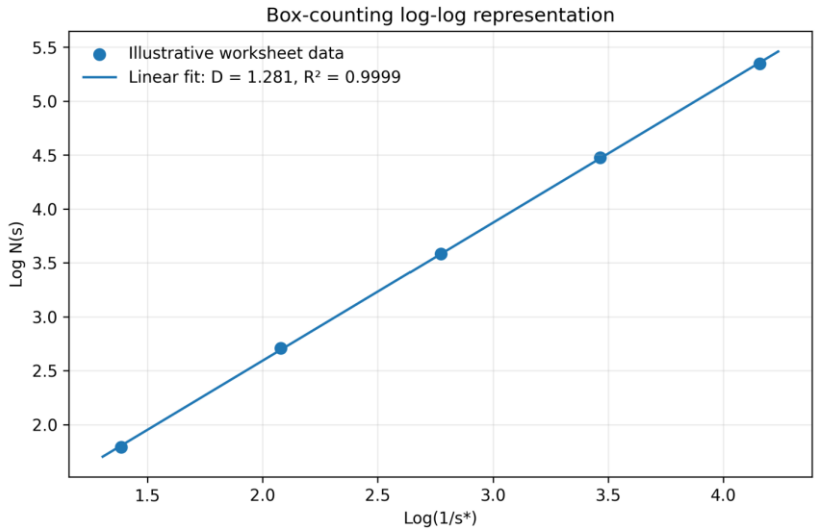


Fig. 3. Log-Log representation of the illustrative box-counting data from Table 1. The slope of the fitted line provides an estimate of the box-counting dimension (D = 1.281, R^2 = 0.9999). Here, Log denotes the natural logarithm

6. DIDACTICAL STRUCTURE OF THE ACTIVITY

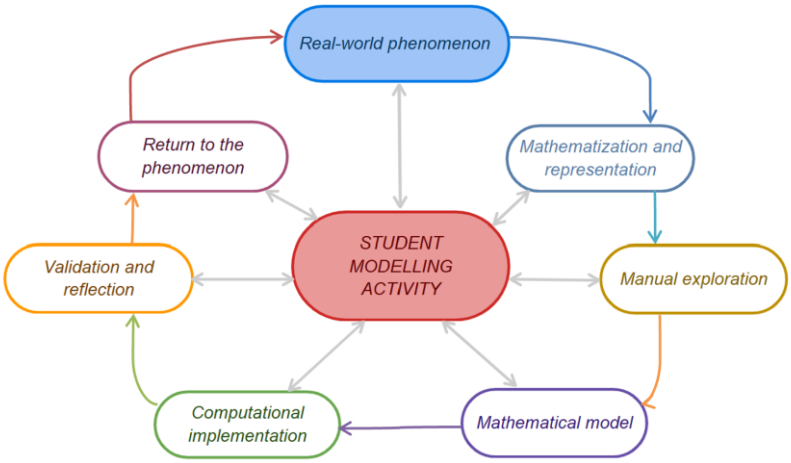


Fig. 4. Schematic didactical modelling cycle connecting physics, mathematics and informatics

The simplified diagram in Figure 4 deliberately shows only the main stages of the learning cycle. The detailed disciplinary and student actions are described here rather than inside the figure. The real-world phenomenon provides the physical context: students consider the material, the visible crack, its formation and branching, and later return to the original structure to interpret the result. During mathematization and manual exploration, attention shifts to geometry, scale and the counting function N(s). Students choose a box size s, count occupied boxes, organise the values in a table, compare different scales and make predictions. The mathematical-model stage connects these data with functional dependence and, as an optional extension, logarithmic transformation, a log-log representation and the interpretation of the slope as an estimate of fractal dimension D. The computational stage makes the procedure executable through image processing, thresholding, binary representation, automated box counting and visualisation of the results. Validation then requires students to compare manual and digital results, discuss scale and accuracy, consider what the program actually calculates and identify information that may be lost when the original image is converted to a binary representation. Finally, students return to the real phenomenon and ask what the numerical result means for the original crack, how the physical context may help to interpret the observed structure, what the limits of the model are and which further questions arise. Throughout the cycle, the central student modelling activity consists of exploring, calculating, using digital tools, discussing, interpreting and reflecting.

Table 2. Proposed learning sequence and its mathematical and algorithmic components

Phase	Student activity	Mathematical focus	Digital/algorithmic focus
1. Explore	Describe the visible structure and propose measurable features.	Geometry; measurement; modelling question	Image as data
2. Represent	Distinguish object and background; discuss a binary representation.	Sets; discretisation; coordinates	Threshold rule; pixels
3. Measure	Count occupied boxes for several grid sizes.	Scale; counting; ratios	Iteration
4. Transform	Create a table and transform the scale-dependent data.	Functions; logarithms	Data processing
5. Model	Plot transformed data and examine a linear region.	Slope; linear model; approximation	Regression/visualisation

7. CONNECTING MATHEMATICS, PHYSICS AND INFORMATICS

The crack image can be treated as a boundary object between three school subjects. From a physics perspective, it is evidence of a real process in which forces, stresses and material properties may lead to the formation and branching of cracks. The present learning activity does not require a quantitative fracture-mechanics model; the physical perspective is used to generate questions about why the observed structure is irregular and why two structures of similar size may differ in geometric complexity. This keeps the physical context authentic without shifting the learning goal away from mathematics.

The mathematical perspective begins when the visual phenomenon is simplified and quantified. Students consider which characteristics could describe complexity, examine the effect of scale, record the number of occupied boxes N(s), and move between a table, a graph and a functional relation. For advanced students, the log-log representation provides an extension in which the slope can be interpreted as an estimate of fractal dimension. The informatics perspective makes the same process executable: the image becomes pixel data, a threshold produces a binary representation, connected pixels are identified, and a loop repeats the counting procedure for several scales. Thus, the three subjects are not placed side by side; they meet in the same modelling decisions.

This structure is consistent with interdisciplinary STEM approaches in which mathematical modelling helps connect disciplinary knowledge rather than reducing mathematics to a calculation service (Maass et al., 2019). It also provides a concrete setting for computational thinking. Decomposition appears when the analysis is separated into image preparation, segmentation, counting and interpretation; abstraction appears when the crack is represented as a binary structure; and algorithmic thinking appears when the box-counting procedure is formulated as a repeatable sequence. A systematic review of K–12 mathematics education found that mathematical and computational reasoning can develop in an interactive, cyclical relationship (Ye et al., 2023).

Table 3. Interdisciplinary structure of the proposed learning activity

Perspective	Guiding question	Conceptual focus
Physics	Why does the observed structure form and branch?	real phenomenon; qualitative ideas of stress, material response and crack formation
Mathematics	How can the geometric complexity be described quantitatively?	scale; geometry; N(s); functions; logarithms; slope; fractal dimension
Informatics	How can the procedure be represented and automated?	pixels; binary data; threshold; connected components; iteration; algorithm
Integrated modelling	What does the computed result mean for the original image?	interpretation; validation; assumptions; limits of the digital model
Perspective	Guiding question	Conceptual focus

8. PRELIMINARY DEMONSTRATION WITH STUDENTS IN GRADES 9-10

A preliminary demonstration of the approach was carried out in an extracurricular mathematics-club setting with students in Grades 9–10 in a German university outreach context. The regular group consisted of 15 students. The session was intended to show a concrete example of how mathematics can be used together with physics and informatics in a real-world context; it was not organised as an empirical study. The students were shown the original crack image, the image-analysis program and the main mathematical ideas underlying its operation. After the

demonstration, they received a task based on the same principle, directing attention from the software output back to the mathematical steps on which it depends.

Grades 9–10 are a suitable target group for the core version of the activity because the task can build on mathematical competences expected by the end of lower secondary education in North Rhine-Westphalia. The NRW mathematics curriculum emphasises translating real situations into mathematical models, interpreting and validating model results, working with tables and graphs, understanding functional relationships, and using digital tools such as spreadsheets and function plotters to explore mathematical problems (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2023). The crack task brings these process competences together in a single non-routine context.

The activity should nevertheless distinguish between a core and an extension level. At the core level, students in Grades 9–10 can compare grids, determine N(s), organise values in a table, construct graphical representations and discuss how the count changes with scale. The logarithmic transformation and the formal interpretation of the slope as a fractal dimension should be treated as an enrichment element rather than assumed prior knowledge. In NRW, logarithmic functions belong more naturally to upper-secondary mathematics, whereas modelling, functional thinking and the use of digital tools already provide an appropriate lower-secondary foundation. This differentiation makes the activity usable with heterogeneous groups without weakening its mathematical coherence.

The German school context is also supported by recent empirical work. Bayer et al. (2024) investigated a four-day mathematical modelling project with 14 students in Grades 9–11 at a German secondary school and observed computational-thinking aspects such as data collection, pattern recognition and abstraction emerging within modelling activity. Their findings support the use of project-like and extracurricular settings for tasks in which mathematical and computational reasoning are deliberately connected. The present crack-image activity is narrower in duration and content, but it follows a compatible principle: students should understand and discuss the mathematical structure before treating the computer output as an answer.

Because the demonstration was exploratory, no claims about effectiveness, conceptual change or achievement are made here. A subsequent empirical implementation should document student worksheets, group discussions and interactions with the software. Of particular interest would be how students move between physical interpretation, visual representation, discrete pixel data, numerical tables and graphs, and whether manual box counting helps them explain what the automated algorithm actually computes.

9. DIFFERENTIATION AND POSSIBLE EXTENSION

The same visual context can be used at different levels without requiring the same formal mathematics. For Grades 7–8, the focus can remain on grids, counting, scale and comparison of visual complexity. Students can work with several artificial lines and decide which requires more boxes at progressively finer scales. The term fractal dimension is not necessary at this stage.

For Grades 9–10, the task can include tables, coordinate representations, functional dependence and repeated algorithmic procedures. Logarithmic transformation can be introduced as an enrichment component, especially in a mathematics-club setting where the activity is not constrained by a single lesson or by the exact sequence of the regular curriculum. The central idea is that a change of representation may reveal a simpler mathematical structure.

At upper-secondary or introductory university level, the full log-log analysis can be used. Students can compare different scale intervals, discuss the quality of a linear approximation, examine the influence of image resolution and threshold selection, and consider why an estimated dimension depends on modelling choices. This creates a natural route from calculation to model criticism.

A useful extension is to compare several structures of increasing visual complexity: a straight segment, a zigzag line, a branching artificial curve and an irregular real-world image. Students first predict an ordering, then perform box counting and finally compare their prediction with the calculated behaviour. Such a task introduces validation and prevents the software result from being treated as an unquestionable answer.

10. DISCUSSION

The main didactical point is not the crack itself, but the way one object can be approached through mathematics, physics and informatics. Physics supplies an authentic phenomenon and questions about the origin and meaning of visible structures; mathematics provides the language of scale, counting, functions, logarithms and approximation; informatics provides digital representation, classification, iteration and automation. The didactical value lies in coordinating these perspectives without allowing any one of them to dominate the task. The crack is therefore not

used to teach fracture mechanics, and the software is not used as an opaque calculator. Both serve as contexts in which students can make and discuss modelling decisions.

The crack image is particularly suitable as a motivating example because its shape is recognisably real but not easily reduced to a standard geometric figure. Nevertheless, the didactical design does not depend on the physical origin of the crack. Similar activities could be constructed from coastlines, leaf contours, branching patterns or other irregular digital structures. This transferability is important if the approach is to be developed beyond a single demonstration.

The current version has one clear limitation: no systematic empirical data were collected during the mathematics-club demonstration. Accordingly, the session is used here to document feasibility and to motivate the refinement of the learning design, not to claim learning gains. A later study should collect student worksheets and discussion data and compare manually obtained values with computational reference values. This would allow students' modelling processes, sources of deviation and understanding of data reliability to be examined systematically .

11. CONCLUSION

The crack-image activity shows one way of turning a research-related digital object into a task that is accessible to secondary-school students. The important step is to delay automation: students first work with the image, the grid and the counts, and only afterwards compare this work with the software procedure. In this way, box counting is not presented as a ready-made formula, and the digital tool does not replace the mathematical discussion.

For Grades 9–10, the most appropriate core consists of scale, tables, graphs, functional dependence and algorithmic reasoning. Logarithms and fractal dimension can remain an optional extension. The preliminary demonstration in the mathematics club was sufficient to test the practicability of the sequence, but not to evaluate learning effects. A later study should therefore collect student worksheets and discussion data and compare students' manual results with computational reference values. This would also open a useful discussion of measurement error, data reliability and the way quantitative image data can enter a broader technical or engineering workflow.

References

[1] Blomhøj, M., & Jensen, T. H. (2003). Developing mathematical modelling competence: Conceptual clarification and educational planning. *Teaching Mathematics and its Applications*, *22*(3), 123–139. https://doi.org/10.1093/teamat/22.3.123

[2] Blum, W., Galbraith, P. L., Henn, H.-W., & Niss, M. (Eds.). (2007). *Modelling and Applications in Mathematics Education: The 14th ICMI Study*. Springer. https://doi.org/10.1007/978-0-387-29822-1

[3] Cevikbas, M., Greefrath, G., & Siller, H.-S. (2023). Advantages and challenges of using digital technologies in mathematical modelling education – a descriptive systematic literature review. *Frontiers in Education*, *8*, 1142556. https://doi.org/10.3389/feduc.2023.1142556

[4] Cevikbas, M., Kaiser, G., & Schukajlow, S. (2022). A systematic literature review of the current discussion on mathematical modelling competencies: State-of-the-art developments in conceptualizing, measuring, and fostering. *Educational Studies in Mathematics*, *109*, 205–236. https://doi.org/10.1007/s10649-021-10104-6

[5] Engelbrecht, J., & Borba, M. C. (2024). Recent developments in using digital technology in mathematics education. *ZDM – Mathematics Education*, *56*, 281–292. https://doi.org/10.1007/s11858-023-01530-2

[6] Geiger, V., Faragher, R., & Goos, M. (2010). CAS-enabled technologies as 'agents provocateurs' in teaching and learning mathematical modelling in secondary school classrooms. *Mathematics Education Research Journal*, *22*, 48–68. https://doi.org/10.1007/BF03217565

[7] Greefrath, G., Hertleif, C., & Siller, H.-S. (2018). Mathematical modelling with digital tools – a quantitative study on mathematising with dynamic geometry software. *ZDM – Mathematics Education*, *50*, 233–244. https://doi.org/10.1007/s11858-018-0924-6

[8] Jablonski, S. (2023). Is it all about the setting? A comparison of mathematical modelling with real objects and their representation. *Educational Studies in Mathematics*, *113*, 307–330. https://doi.org/10.1007/s10649-023-10215-2

[9] Kaiser, G., & Sriraman, B. (2006). A global survey of international perspectives on modelling in mathematics education. *ZDM*, *38*(3), 302–310. https://doi.org/10.1007/BF02652813

[10] Koyunkaya, M. Y., & Dede, A. T. (2024). Using different digital tools in designing and solving mathematical modelling problems. *Education and Information Technologies*, *29*, 19035–19065. https://doi.org/10.1007/s10639-024-12577-3

[11] Li, J., Du, Q., & Sun, C. (2009). An improved box-counting method for image fractal dimension estimation. *Pattern Recognition*, *42*(11), 2460–2469. https://doi.org/10.1016/j.patcog.2009.03.001

[12] Maass, K., Geiger, V., Romero Ariza, M., & Goos, M. (2019). The role of mathematics in interdisciplinary STEM education. *ZDM – Mathematics Education*, *51*, 869–884. https://doi.org/10.1007/s11858-019-01100-5

[13] Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. W. H. Freeman.

[14] Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. (2023). *Kernlehrplan Mathematik für die Sekundarstufe I*. Schulentwicklung NRW.

[15] Niss, M., & Blum, W. (2020). *The Learning and Teaching of Mathematical Modelling*. Routledge. https://doi.org/10.4324/9781315189314

[16] Schukajlow, S., Kaiser, G., & Stillman, G. (2018). Empirical research on teaching and learning of mathematical modelling: A survey on the current state-of-the-art. *ZDM – Mathematics Education*, *50*, 5–18. https://doi.org/10.1007/s11858-018-0933-5

[17] Ye, H., Liang, B., Ng, O.-L., & Chai, C. S. (2023). Integration of computational thinking in K–12 mathematics education: A systematic review on CT-based mathematics instruction and student learning. *International Journal of STEM Education*, *10*, 3. https://doi.org/10.1186/s40594-023-00396-w

[18] Bayer, S., Greefrath, G., & Siller, H.-S. (2024). Computational thinking in mathematical modelling projects with secondary-school students. *Quadrante*, *33*(2). https://doi.org/10.48489/quadrante.36870

[19] Olemskoi, A. I., Yushchenko, O. V., Borysiuk, V. N., Perekrestov, V. I., Kosminska, Y. O., & Zhylenko, T. I. (2012a). Hierarchical condensation near phase equilibrium. *Physica A: Statistical Mechanics and its Applications*, *391*(11), 3277–3284. https://doi.org/10.1016/j.physa.2011.10.027

[20] Olemskoi, O. I., Borysiuk, V. M., Yushchenko, O. V., & Zhylenko, T. I. (2012b). The statistical analysis of structure of the metal condensates obtained by a sputtering at phase equilibrium. *Metallofizika i Noveishie Tekhnologii*, *34*(1), 87–97.

[21] Zhylenko, T. I., & Zakharchenko, N. M. (2013). Application of cloud technologies in the study of higher mathematics. *CTE Workshop Proceedings*, *1*, 118. https://doi.org/10.55056/cte.158

[22] Zhylenko, T. I., Martynova, N. S., Shuda, I. A., Chykalov, Y. A., & Kuzmuk, D. A. (2020). Auto Checker of Higher Mathematics – an element of mobile cloud education. *CTE Workshop Proceedings*, *7*, 662–673. https://doi.org/10.55056/cte.422

How Digitalization Translates into Operational Performance: The Mediating Role of Lean Management in Manufacturing Enterprises

Klaus DÄNNER[a*]

[a]*University of Latvia, Faculty of Economics and Social Sciences, Department of Economics, Riga, Latvia;*
Email: [insert email]
** Corresponding author*

Abstract

This paper examines how digitalization and lean management are jointly associated with operational performance in manufacturing enterprises, with particular attention to whether lean management acts as a moderator or a mediator. The study follows an exploratory sequential mixed-methods design. Fourteen expert interviews were used to refine the empirical framework, followed by an online survey of 252 respondents from manufacturing enterprises. Exploratory factor analysis was applied to consolidate literature-derived indicators into reliable composite measures for digitalization, lean management and operational performance. Direct relationships and moderation were examined with regression-based ANCOVA models using robust standard errors, while mediation was tested with the PROCESS macro and bootstrap estimation. Digitalization is positively associated with operational performance ($B = 0.179$, $p = 0.002$) and with lean management implementation ($B = 0.540$, $p < 0.001$). Lean management is positively associated with operational performance ($B = 0.335$, $p < 0.001$). The interaction between digitalization and lean management is not significant ($B = 0.018$, $p = 0.646$). In contrast, the indirect relationship through lean management is significant ($B = 0.333$, 95% bootstrap CI [0.250, 0.426]), while the direct relationship becomes non-significant after the mediator is included. The results indicate that lean management functions primarily as an organizational transmission mechanism through which digital capabilities are associated with operational improvements rather than as a contextual amplifier of the digitalization-performance relationship.

Keywords: digitalization, lean management, operational performance, manufacturing enterprises, mediation

1. INTRODUCTION

Manufacturing enterprises are under continuous pressure to improve productivity, quality, flexibility and cost performance while simultaneously adapting their production systems to rapid technological change. Digitalization and lean management are two of the most influential approaches used to address this challenge. Digitalization expands the availability of real-time information, connectivity, analytics and automation, whereas lean management provides process-oriented principles for reducing waste, stabilizing workflows and sustaining continuous improvement. The

* Corresponding author

practical question is therefore no longer whether either approach matters in isolation, but how both approaches interact inside manufacturing systems.

Research increasingly describes lean management and digitalization as complementary capabilities rather than independent transformation programmes. Buer et al. map the growing connection between Industry 4.0 and lean manufacturing [1], while later empirical work identifies a complementary effect on operational performance [2]. Recent studies additionally emphasize organizational culture, structured management systems and the integration of advanced digital technologies into existing lean production environments [3-5]. A dynamic-capabilities perspective likewise suggests that integrated lean and digital transformation depends on the interaction of technological and organizational capabilities [7].

Despite this development, an important conceptual ambiguity remains. Lean management may influence the digitalization-performance relationship in at least two fundamentally different ways. First, it may act as a moderator, meaning that digitalization is more strongly associated with operational performance when lean management is implemented at a higher level. Second, lean management may act as a mediator, meaning that digitalization is associated with stronger lean management implementation, which in turn is associated with higher operational performance. These explanations are theoretically different and require different statistical tests.

This paper addresses that distinction by empirically examining digitalization, lean management and operational performance within one integrated model. The study focuses on manufacturing enterprises because operational outcomes in this context depend on the coordinated interaction of technologies, equipment, employees, material flows, standardized processes and improvement routines. The central research question is: How and to what extent is digitalization associated with operational performance in manufacturing enterprises, and what role does lean management play as a mediator and moderator in this relationship?

The contribution of the paper is threefold. First, it empirically distinguishes moderation from mediation instead of treating the lean-digital relationship as a generic interaction. Second, it examines digitalization, lean management and operational performance as higher-level constructs derived from multiple observable indicators rather than from one isolated technology or method. Third, it derives direct managerial implications for the design of integrated digital and lean transformation in manufacturing.

2. THEORETICAL BACKGROUND AND HYPOTHESES

2.1. Digitalization and operational performance

Digitalization in manufacturing encompasses the operational integration of technologies for connectivity, data generation, data processing, visualization and automated decision support. The specific technologies used by firms may change over time, but the broader capability remains relevant because it reflects the degree to which production systems are digitally enabled. Existing research associates digitally supported manufacturing with improved information availability, faster decision-making, higher flexibility and process transparency. At the same time, recent studies caution that standalone technology adoption does not automatically generate operational benefits when technologies are not embedded in appropriate production and management systems [4,5].

H1: A higher degree of digitalization is positively associated with operational performance in manufacturing enterprises.

2.2. Digitalization and lean management

Lean management is a process- and organizational capability based on stable processes, standardization, flow, waste reduction, employee involvement and continuous improvement. Digital technologies can strengthen these capabilities by increasing transparency, shortening feedback loops and improving the availability of production information. Conversely, lean structures can provide a disciplined operational foundation for the implementation of digital technologies. This reciprocal logic is reflected in research on Lean 4.0, digitally supported lean production and integrated transformation frameworks [1,3,6-8].

H2: A higher degree of digitalization is positively associated with a higher level of lean management implementation in manufacturing enterprises.

2.3. Lean management and operational performance

Lean management is widely associated with improvements in cost, quality, lead time, inventory, flexibility and process reliability. Buer et al. report complementary effects of lean manufacturing and digitalisation on operational performance [2], while Ghaithan et al. show that lean manufacturing contributes to multidimensional performance in combination with Industry 4.0 and other operational capabilities [9]. In the present study, operational performance is therefore treated as a multidimensional outcome rather than as a single financial or production indicator.

H3: A higher level of lean management implementation is positively associated with operational performance in manufacturing enterprises.

2.4. Moderation versus mediation

Moderation and mediation answer different questions. A moderation model tests whether the strength of the relationship between a predictor and an outcome changes depending on a third variable. A mediation model instead tests whether the predictor is associated with the outcome indirectly through an intervening variable [13]. Translating this distinction to the lean-digital context leads to two competing mechanisms. If lean management is a moderator, higher lean implementation should amplify the digitalization-performance relationship. If lean management is a mediator, digitalization should be associated with stronger lean implementation and lean implementation should subsequently be associated with operational performance.

H4: Lean management moderates the relationship between digitalization and operational performance, such that the positive relationship is stronger at higher levels of lean management implementation.

H5: Lean management mediates the relationship between digitalization and operational performance, such that higher digitalization is associated with stronger lean management implementation, which in turn is associated with higher operational performance.

3. RESEARCH METHODOLOGY

3.1. Research design and data collection

The study follows an exploratory sequential mixed-methods design [10]. The theoretical framework was developed from the literature and refined through a qualitative phase consisting of 14 semi-standardized expert interviews with executives and specialists experienced in manufacturing, digitalization and lean management. The interviews were used to explore practical interrelations, implementation challenges and relevant technologies and lean practices, and to support the preparation of the subsequent quantitative instrument.

The quantitative phase consisted of an online survey conducted from May to July 2023. Participants were recruited through targeted professional-network searches and direct outreach to professionals in manufacturing-related roles. Across the recruitment activities, 6,103 potential participants were contacted and 252 responses were obtained, corresponding to a response rate of 4.13% in relation to the contacted sample. Of the 252 responses, 228 were submitted through the German-language questionnaire and 24 through the English-language version. The questionnaire used seven-point response scales and included items for digitalization, lean management, operational performance and contextual variables such as company size, revenue, industry affiliation, year of birth and gender.

3.2. Measurement and construct formation

The questionnaire initially contained 11 items representing digitalization technologies, 20 items representing lean management methods and 10 items representing operational performance. The items were derived from previous scientific studies and were selected to represent multiple facets of each construct. Exploratory factor analysis was applied because the empirical structure of these broader constructs was not assumed to be fully known in advance. Principal axis factoring with Oblimin rotation was used in line with established recommendations for exploratory factor analysis [11,12].

The analyses identified empirically coherent item groups that were subsequently used to calculate composite variables. The resulting scales show high internal consistency. Digitalization is represented by six items related to digitally integrated production, connectivity and data processing. Lean management is represented by eight items capturing flow, synchronization, flexible resource use and related production practices. Operational performance is

represented by seven items covering cost, profit, energy, scrap and rework, flexibility, inventory and lead time. Table 1 summarizes the final construct measures.

Table 1. Empirically consolidated constructs

Construct	Initial items	Retained items	Cronbach's alpha
Digitalization	11	6	0.896
Lean management	20	8	0.946
Operational performance	10	7	0.908

3.3. Statistical analysis

Direct relationships and the moderation hypothesis were examined in IBM SPSS Statistics using regression-based ANCOVA models with HC3 robust standard errors. Contextual variables were included in the specified models to reduce the risk that observed associations merely reflected structural differences between firms and respondents. H4 was evaluated through the interaction term between digitalization and lean management. Mediation was tested with the PROCESS macro for SPSS using 5,000 bootstrap samples [13]. Because the data are cross-sectional, all statistical relationships are interpreted as associations. A significant indirect effect is consistent with the proposed mediation mechanism but does not establish a causal sequence.

4. RESULTS

4.1. Direct relationships

The analysis of H1 shows a positive and statistically significant association between digitalization and operational performance. The unstandardized parameter estimate is B = 0.179 with a robust standard error of 0.056 (t = 3.185, p = 0.002, 95% CI [0.068, 0.291]). The result indicates that manufacturing enterprises with higher digitalization scores tend to report higher operational performance after the additional variables included in the model are considered. H1 is therefore supported.

H2 is also supported. Digitalization is positively associated with lean management implementation (B = 0.540, robust SE = 0.075, t = 7.170, p < 0.001, 95% CI [0.392, 0.689]). The magnitude and significance of this relationship indicate that firms with higher levels of digitalization also tend to report stronger implementation of the lean management construct.

For H3, lean management is positively and significantly associated with operational performance (B = 0.335, robust SE = 0.064, t = 5.214, p < 0.001, 95% CI [0.209, 0.462]). In the corresponding model, the coefficient for digitalization is not statistically significant (B = -0.004, p = 0.942), whereas lean management remains significant. The overall H3 model explains 57.8% of the variance in operational performance (R^2 = 0.578; adjusted R^2 = 0.558). H3 is therefore supported.

4.2. Moderation analysis

H4 proposes that lean management changes the strength of the digitalization-performance relationship. The interaction term between digitalization and lean management is positive but not statistically significant (B = 0.018, robust SE = 0.038, t = 0.459, p = 0.646, 95% CI [-0.058, 0.093]). The corresponding partial eta squared value is 0.001. The analysis therefore provides no statistical evidence that the association between digitalization and operational performance becomes stronger or weaker depending on the level of lean management implementation. H4 is not supported.

4.3. Mediation analysis

The mediation model provides a different result. The total association between digitalization and operational performance is significant (B = 0.338, HC3 SE = 0.050, t = 6.821, p < 0.001, 95% CI [0.240, 0.435]). After lean management is included as the mediator, the direct association between digitalization and operational performance becomes non-significant (B = 0.005, HC3 SE = 0.058, t = 0.086, p = 0.932, 95% CI [-0.109, 0.119]).

In contrast, the indirect association through lean management is positive and statistically significant (B = 0.333, BootSE = 0.045, 95% bootstrap CI [0.250, 0.426]). The completely standardized indirect effect is 0.461 with a 95% bootstrap confidence interval of [0.362, 0.566]. Because the bootstrap confidence interval for the indirect effect does not include zero, H5 is supported. Within the specified statistical model, the results are therefore consistent with an indirect pathway in which higher digitalization is associated with stronger lean management implementation, which in turn is associated with higher operational performance.

Table 2. Summary of hypothesis tests

Hypothesis	Relationship	Estimate	Significance	Result
H1	Digitalization -> Operational performance	B = 0.179	p = 0.002	Supported
H2	Digitalization -> Lean management	B = 0.540	p < 0.001	Supported
H3	Lean management -> Operational performance	B = 0.335	p < 0.001	Supported
H4	Digitalization x Lean management -> Operational performance	B = 0.018	p = 0.646	Not supported
H5	Indirect: Digitalization -> Lean management -> Operational performance	B = 0.333	95% BootCI [0.250, 0.426]	Supported

5. DISCUSSION

The findings provide a differentiated explanation of the relationship between digitalization, lean management and operational performance. At first sight, the direct H1 result appears to suggest that digitalization itself is associated with better operational performance. However, the mediation analysis changes the interpretation. Once lean management is included as the mediator, the direct digitalization-performance association becomes statistically non-significant, while the indirect association through lean management remains significant. This pattern indicates that the relationship is largely represented by the digitalization-lean-performance pathway within the specified model.

This result is theoretically important because it separates two explanations that are often combined under the broad idea of "integration." The non-significant interaction term provides no evidence that lean management simply amplifies the effect of digitalization when lean implementation is high. Instead, lean management appears to function as an organizational mechanism through which digital capabilities are connected with operational improvement. This interpretation is consistent with research arguing that digital technologies create value when they are embedded in structured management systems and existing lean-based production environments [3-5]. It also aligns with the dynamic-capabilities view that integrated transformation depends on the coordination of technological and organizational capabilities rather than on technology adoption alone [7].

The results also support a construct-level understanding of digitalization. The empirical digitalization scale is not tied to one specific technology but combines six technologies associated with connectivity, data processing and integrated production. This is useful because individual technologies evolve rapidly. The research model therefore remains conceptually relevant even when specific digital technologies change, although new technologies would need to be empirically validated before being incorporated into the measurement scale.

For lean management, the retained indicators emphasize production flow, synchronization, flexible layout and workforce deployment. The factor structure suggests that the empirically strongest lean dimension in this sample is closely related to the organization and coordination of manufacturing flow. This may help explain the mediation result: digitalization can increase the availability and speed of information, but operational benefits depend on the existence of processes and routines capable of using that information to improve production. Lean management provides such routines through standardized work, flow orientation, problem solving and continuous improvement.

6. MANAGERIAL IMPLICATIONS

The primary managerial implication is that digitalization and lean management should not be implemented as parallel or competing transformation programmes. Technology investments should be connected to existing process-

improvement structures, standardized workflows and operational objectives. A technology-first approach can increase data availability without necessarily improving the production system. The empirical mediation result instead suggests that digital capability should be translated into performance through disciplined operational management.

Second, managers should align digital investment decisions with the maturity of lean capabilities. Before selecting technologies, enterprises should identify the operational problem to be solved and determine whether the necessary process stability, standards, responsibilities and improvement routines are in place. Digital tools can then be selected according to their contribution to transparency, process stability, flexibility, quality, lead time, inventory or other defined performance targets.

Third, performance management should integrate the three dimensions examined in this study. Digital progress should not be measured only through installed technologies, and lean progress should not be measured only through the existence of methods. Enterprises should simultaneously monitor digital implementation, lean management capability and operational outcomes. This integrated measurement approach makes it easier to evaluate whether digital investments are actually translated into operational improvement.

Finally, leadership and employee capability remain central. Digital transformation requires technical competence, while lean management requires process understanding, problem-solving capability and continuous learning. An integrated Lean Digital mindset should therefore combine technological skills with operational excellence rather than replacing human-centered improvement practices with technology.

7. CONCLUSION

This paper examined the relationships between digitalization, lean management and operational performance in manufacturing enterprises and directly compared moderation and mediation as competing explanations of the role of lean management. The empirical results support positive associations between digitalization and operational performance, digitalization and lean management, and lean management and operational performance. However, the interaction between digitalization and lean management is not significant, providing no evidence of moderation. In contrast, the indirect relationship through lean management is significant, while the direct digitalization-performance relationship becomes non-significant after the mediator is included.

The central conclusion is therefore that lean management appears primarily as an organizational transmission mechanism rather than as a contextual amplifier. Digitalization is associated with operational performance to the extent that digital capabilities are connected with structured process management and lean implementation. For manufacturing practice, this means that the value of digital technologies should be evaluated in relation to the production system in which they are embedded.

The study is subject to several limitations. The quantitative data are cross-sectional and based on respondents' assessments, so the statistical relationships do not establish causality. The targeted professional-network recruitment strategy and the manufacturing-specific sample also limit generalizability. Future research should therefore apply longitudinal designs, incorporate objective operational performance measures and examine whether the identified mediation pattern differs across industries, production types, company sizes and levels of technological maturity. Further studies could also test which specific combinations of digital technologies and lean practices generate the strongest operational benefits.

References

[1] Buer, S. V., Strandhagen, J. O., & Chan, F. (2018). The link between Industry 4.0 and lean manufacturing: Mapping current research and establishing a research agenda. International Journal of Production Research, 56(8), 2924-2940.

[2] Buer, S. V., Semini, M., Strandhagen, J. O., & Sgarbossa, F. (2021). The complementary effect of lean manufacturing and digitalisation on operational performance. International Journal of Production Research, 59(7), 1976-1992.

[3] Gianni, M., & Gotzamani, K. (2024). A lean management perspective on digital transformation: The role of culture in technology adoption. International Journal of Quality & Reliability Management.

[4] Johansson, P. E., Bruch, J., Chirumalla, K., Ostermann, C., & Stalberg, L. (2024). Integrating advanced digital technologies in existing lean-based production systems: analysis of paradoxes, imbalances and management strategies. International Journal of Operations & Production Management, 44(6), 1158-1191.

[5] Powell, D. (2024). Artificial intelligence in lean manufacturing: digitalization with a human touch? International Journal of Lean Six Sigma, 15(3), 719-729.

[6] Gomaa, A. H. (2025). Lean 4.0: A Strategic Roadmap for Operational Excellence and Innovation in Smart Manufacturing. International Journal of Emerging Science and Engineering, 13(5).

[7] Komkowski, T., Antony, J., Lizarelli, F. L., Gutierrez, L., Sony, M., Powell, D., Pongboonchai-Empl, T., Garza-Reyes, J. A., & Araos Acharan, L. I. (2026). Lean Management and Industry 4.0: a dynamic capabilities framework for an integrated transformation. Production Planning & Control.
[8] Dillinger, F., Tropschuh, B., Dervis, M. Y., & Reinhart, G. (2022). A Systematic Approach to Identify the Interdependencies of Lean Production and Industry 4.0 Elements. Procedia CIRP, 112, 85-90.
[9] Ghaithan, A. M., Alshammakhi, Y., Mohammed, A., & Mazher, K. M. (2023). Integrated Impact of Circular Economy, Industry 4.0, and Lean Manufacturing on Sustainability Performance of Manufacturing Firms. International Journal of Environmental Research and Public Health, 20, 5119.
[10] Creswell, J. W. (2015). A Concise Introduction to Mixed Methods Research. London: Sage Publications.
[11] Costello, A. B., & Osborne, J. W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. Practical Assessment, Research & Evaluation, 10(7).
[12] Field, A. (2018). Discovering Statistics Using IBM SPSS Statistics (5th ed.). London: Sage Publications.
[13] Hayes, A. F. (2018). Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach (2nd ed.). New York: The Guilford Press.

Airports as the First Impression of a Destination: Examples from the United Arab Emirates

Iva BULATOVIC[a*], Gamal KHALIFA[b], Mahrha Hamad Saeed Saif ALHASSANI[c], Maitha Saif Juma Ali ALKAABI[c], Mariam Humaid Naser Humaid ALHOSANI[c]

[a] *Higher Colleges of Technology, Faculty of Business, Abu Dhabi, United Arab Emirates, ibulatovic@hct.ac.ae*
[b] *Higher Colleges of Technology, Faculty of Business, Abu Dhabi, United Arab Emirates, gkhalifa@hct.ac.ae*
[c] *Higher Colleges of Technology, Student, Abu Dhabi, United Arab Emirates, H00447350@hct.ac.ae; H00467524@hct.ac.ae; H00491043@hct.ac.ae*

Abstract

Airports are often the first and last place a visitor sees in a country. For this reason, they strongly influence how travellers see a destination. This paper looks at how airports in the United Arab Emirates present the country to visitors, using Dubai International Airport and Zayed International Airport in Abu Dhabi. The study uses a simple two-step method. First, the authors visit the terminals and use an observation checklist to record how each terminal presents the UAE. Second, a short online questionnaire asks travellers who have recently used UAE airports about their first impressions and whether the airport changed how they see the country. The paper identifies which airport features travellers notice most, such as terminal architecture and interior design, bilingual Arabic and English signage, local art and cultural displays, Emirati food and souvenir shops, and the friendliness of staff at immigration and arrival. It then explains how these features shape travellers' image of the UAE as a destination. It also compares a large hub airport with a smaller regional airport. The study ends with practical recommendations for airport operators and tourism authorities on how airports can better welcome visitors and present the UAE.

Keywords: airports, destination image, first impressions, United Arab Emirates

1. INTRODUCTION

For most international visitors, the airport is the first physical place they experience in a new country and the last place they see before they leave. Within the first hour after landing, a traveller meets the destination's architecture, its signage, its border officers, its shops and its level of organisation. Tourism and air transport research increasingly recognises that this encounter is not neutral. Airports are part of the tourism experience itself, and they can act as representatives of the destinations they serve (Wattanacharoensil et al., 2016; Goetz, 2019). A smooth, welcoming and distinctive arrival can confirm what a destination promises in its marketing, while queues, confusion or unfriendly service can weaken that promise before the visitor has even left the terminal.

* Corresponding author

The United Arab Emirates (UAE) is a particularly useful setting for studying this relationship. Aviation, tourism and place branding are closely connected in the country, and the national carriers and airports are central to its international image (Al Saed et al., 2020). Dubai International Airport (DXB) welcomed a record 95.2 million passengers in 2025, 3.1% more than in the previous year, which makes it one of the busiest international airports in the world (Dubai Media Office, 2026). Abu Dhabi's five airports recorded more than 33 million passengers in 2025, and Zayed International Airport (AUH) handled about 98% of this traffic (Khaleej Times, 2026). Zayed International opened its new Terminal A in 2023 and took its current name in 2024. The two airports therefore represent a very large global hub and a smaller hub serving the capital region. They also serve two cities with different destination identities: Dubai is usually framed as global, energetic and entertainment-led, while Abu Dhabi is associated more with culture, heritage and prestige (Slak Valek & Williams, 2018; Williams, 2022).

Despite the importance of airports to the UAE's tourism strategy, evidence on what travellers actually notice inside UAE terminals, and how those features relate to their image of the country, remains limited. Existing studies focus mainly on destination branding, airline promotion and national tourism competitiveness (Reisinger et al., 2019; Al Saed et al., 2020). This paper addresses that gap. Its aim is to examine how airports in the UAE present the country to visitors and how airport features shape travellers' first impressions of the UAE as a destination. The study pursues three research questions:

- RQ1: Which airport features do travellers notice most when they use Dubai International Airport and Zayed International Airport?
- RQ2: How does the airport experience relate to travellers' image of the UAE as a destination?
- RQ3: How do the first impressions created by the larger hub (DXB) and the smaller hub (AUH) differ?

The paper is organised as follows. Section 2 reviews the literature on destination image, the airport experience and the role of airports in destination image. Section 3 describes the two-step method. Section 4 presents the results, Section 5 discusses them, and Section 6 concludes with recommendations for airport operators and tourism authorities.

2. LITERATURE REVIEW

2.1. Destination image and first impressions

Destination image refers to the beliefs, ideas and impressions that a person holds about a place. It is commonly understood to have a cognitive component, which concerns what people know or believe about a destination's attributes, and an affective component, which concerns how they feel about it (Baloglu & McCleary, 1999). Image is formed from many sources, including advertising, media, word of mouth and, most powerfully, direct personal experience. Once formed, destination image influences tourist satisfaction and behavioural intentions such as revisiting and recommending the destination (Chi & Qu, 2008). Because the airport is usually the first point of direct experience, it is well placed to confirm, strengthen or challenge the image that a visitor brings with them.

2.2. The airport experience and service quality

Passengers evaluate airports on several dimensions at once. Fodness and Murray (2007) identified function, interaction and diversion as the core dimensions of airport service quality, and later work added cleanliness, signage, comfort and the physical environment (Bezerra & Gomes, 2016; Usman et al., 2022). Wattanacharoensil et al. (2016, 2017) moved beyond service quality to propose an airport experience framework from a tourism perspective. They showed that airports increasingly offer cultural, leisure and sensory experiences that extend their role from a transport utility to a place in its own right. Studies based on online reviews and social media confirm that passengers talk about airports in terms of both operational issues, such as queues and wayfinding, and the atmosphere of the terminal (Martin-Domingo et al., 2019; Booranakittipinyo et al., 2024).

2.3. Airports as part of destination image

A growing body of research links the airport experience directly to destination image. Wattanacharoensil et al. (2022) found that the airport environment, or airportscape, creates a sense of place that influences how travellers

perceive the destination. Fakfare (2023) showed that the airport environment affects perceived image and memorable experiences, and Saut and Song (2022) found that airport service quality and satisfaction shape travellers' intention to visit the destination country. Abouseada et al. (2023) argued that airport branding plays an active role in shaping tourist destination image. From a broader perspective, Goetz (2019) described how airports have developed into attractions and airport cities, and Ziakas et al. (2026) highlighted the triangular relationship between airlines, airports and destinations in shaping the overall travel experience. In the Gulf, Elmoghazy et al. (2025) showed that identity indicators in Saudi airport terminals, such as traditional patterns, palm-inspired columns and calligraphy, can create attachment and turn terminals into memorable cultural places.

2.4. The arrival experience and wayfinding

Within the airport journey, arrival deserves particular attention. Arriving passengers, and especially first-time visitors, must pass through immigration, collect baggage and find onward transport in an unfamiliar environment, often after a long flight. Wattanacharoensil et al. (2017) found that air travellers experience the airport as a sequence of stages, each with its own emotions, and that uncertainty and stress are concentrated in the processing stages. Wayfinding and signage are therefore central to the experience of unfamiliar passengers, and service quality reviews repeatedly identify information and signage as core dimensions of airport performance (Usman et al., 2022). Social media analyses also show that clear information, waiting times and the availability of help are among the topics that passengers raise most often when they evaluate airports (Martin-Domingo et al., 2019; Booranakittipinyo et al., 2024). For a destination, this means that the first impression is formed not only by what the terminal looks like but also by how easily a newcomer can move through it.

2.5. The UAE context

In the UAE, airports and airlines are closely tied to national tourism strategy. Reisinger et al. (2019) found that infrastructure and support services are among the strongest drivers of perceived tourism competitiveness in the UAE. Al Saed et al. (2020) showed that the promotional activities of Emirates Airline strengthen Dubai's destination brand, and Pike et al. (2021) found that the perceived ambience of Dubai can increase its attractiveness as a stopover destination. At the level of destination branding, studies consistently show that the UAE combines heritage and modernity. Heritage is used strategically in Dubai's tourism branding alongside its modern image (Haq et al., 2021), while Abu Dhabi is promoted more strongly through culture and heritage (Slak Valek & Williams, 2018; Williams, 2022).

Taken together, the literature establishes that airports can shape destination image and that the UAE builds its image on a balance of heritage and modernity. However, few studies directly examine which features inside UAE airports travellers notice, how the airport experience relates to their image of the country, or whether the two main hubs create different first impressions. This study responds to that gap by combining terminal observation with a traveller survey.

3. METHODOLOGY

3.1. Research design and study sites

The study uses a simple two-step design. Structured observation records what the airports offer to visitors (the supply side), while an online questionnaire captures how travellers perceive these features and how they relate to their image of the UAE (the demand side). The two study sites are Dubai International Airport and Zayed International Airport. Both are major international gateways, but they differ in scale and serve two cities with different destination identities, which makes them suitable for comparison.

3.2. Step 1: Structured observation

The authors visited the public and arrival areas of each terminal and completed a structured observation checklist. Observation followed the passenger journey from the arrival hall to ground transport and covered the twelve areas listed in Table 1. Each area was rated on a five-point scale (1 = not present or very poor, 5 = very strong) and supported by written field notes. In line with airport regulations, no photographs were taken inside the terminals.

Table 1. Areas recorded in the observation checklist

Area	What was recorded
Architecture and interior	Design elements reflecting the UAE: shapes, materials, colours and patterns inspired by local heritage
Overall atmosphere	Light, space, cleanliness, comfort, noise and crowding
Signage and wayfinding	Arabic and English signs, directions, screens and information desks
Welcome messages	Greetings and UAE or emirate branding visible on arrival
Art and decoration	Arabic calligraphy, local artworks and cultural displays
Food and drink	Emirati dishes, Arabic coffee, dates and local cafés
Shopping	Local products, Emirati souvenirs, perfumes and crafts
Staff and hospitality	Greetings, helpfulness and traditional hospitality gestures
Immigration area	Queue length, smart gates and staff manner (as visible)
Baggage claim	Waiting time, organisation and information screens
Facilities	Prayer rooms, seating, Wi-Fi, accessibility and family areas
Ground transport	Taxi, metro, bus and car rental, and how easy they are to find

3.3. Step 2: Online questionnaire

A short online questionnaire was distributed through Microsoft Forms. The target population was adults aged 18 and over who had used DXB or AUH within the previous 12 months, including UAE nationals, UAE residents and international visitors. A convenience and snowball sampling approach was used, with the link shared through the authors' professional and student networks and social media. Participation was voluntary and anonymous, and respondents gave consent before starting.

This paper reports four parts of the questionnaire. The traveller profile recorded age, gender, respondent type, region of origin, airport used, type of trip, purpose of travel and whether it was the respondent's first time at the airport. The airport experience section contained ten statements about the terminal (B1 to B10), and the destination image section contained nine statements about the UAE (D1 to D9). Both used a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). Items were adapted from established airport experience and destination image studies (Fodness & Murray, 2007; Wattanacharoensil et al., 2022; Saut & Song, 2022). Four open-ended questions asked what respondents noticed first, what they liked most, what the airport could improve, and how they would describe the airport in three words. Questions on cultural representation in the terminal were also collected but are reported in a separate study.

3.4. Data analysis

Survey data were analysed using descriptive statistics: frequencies, means, standard deviations and the share of respondents who agreed or strongly agreed with each statement. Differences between the two airports, and between first-time and repeat users, were tested with the Mann–Whitney U test, and differences between respondent types and types of trip were tested with the Kruskal–Wallis test. Both tests are suitable for ordinal Likert data. A significance level of .05 was used. Answers to the open-ended questions were analysed through keyword-assisted thematic coding, in which similar answers were grouped into themes. A single answer could be coded under more than one theme.

4. RESULTS

4.1. Observation findings

The observation ratings are summarised in Table 2. Both terminals made a strong overall impression, with an average rating of 4.42 for Dubai International and 4.25 for Zayed International. The two airports performed equally well on architecture and interior design, staff and hospitality, the immigration area, baggage claim, facilities and

ground transport, all of which received the highest rating of 5. These are the core physical and operational parts of the arrival journey, and the ratings confirm that both airports deliver them to a very high standard.

The weakest areas were the same at both airports: welcome messages and art and decoration each received a rating of 3. In other words, the cues that would tell visitors in a direct and symbolic way that they have arrived in the UAE, such as greetings, emirate branding, local artworks and cultural displays, were less visible than the quality of the building and the services. The airports differed in three areas. Dubai International was rated higher for signage and wayfinding (5 compared with 3) and for shopping (5 compared with 4), reflecting its scale and its strong commercial offer. Zayed International was rated higher for food and drink (4 compared with 3). The overall atmosphere was rated 4 at both airports.

Table 2. Observation ratings by airport (1 = not present or very poor, 5 = very strong)

Area observed	DXB	AUH
Architecture and interior	5	5
Overall atmosphere	4	4
Signage and wayfinding	5	3
Welcome messages	3	3
Art and decoration	3	3
Food and drink	3	4
Shopping	5	4
Staff and hospitality	5	5
Immigration area	5	5
Baggage claim	5	5
Facilities	5	5
Ground transport	5	5
Average	**4.42**	**4.25**

4.2. Respondent profile

A total of 100 valid responses were collected, divided equally between the two airports (50 for DXB and 50 for AUH). Table 3 summarises the sample. Almost half of the respondents were visitors to the UAE (48%), while 30% were non-national residents and 22% were UAE nationals. Visitors came mainly from Europe (18), the Americas (9), Oceania (6) and Asia (5). Just over half of the respondents were arriving in the UAE (53%), and 39% were using the airport for the first time.

Table 3. Respondent profile (n = 100)

Characteristic	Category	n (%)
Airport used	Dubai International (DXB) / Zayed International (AUH)	50 / 50
Respondent type	Visitor to the UAE / UAE resident (non-national) / UAE national	48 / 30 / 22
Age group	18–24 / 25–34 / 35–44 / 45–54 / 55+	16 / 28 / 30 / 16 / 10
Gender	Female / Male / Prefer not to say	47 / 46 / 7
Type of trip	Arriving in the UAE / Departing the UAE / Transit only	53 / 36 / 11
Purpose of travel	Visiting friends or family / Leisure / Business / Study / Transit or stopover / Other	28 / 21 / 15 / 12 / 9 / 15
First time at this airport	Yes / No	39 / 61

Note: with n = 100, counts and percentages are identical.

4.3. Airport features travellers notice most (RQ1)

Table 4 presents the ratings of the ten airport experience statements, ranked by the overall mean . Respondents evaluated both airports very positively: every statement received a mean close to or above 4.0, and 85% agreed that they had a positive experience overall (B10). The most highly rated features were cleanliness and comfort (M = 4.34), terminal architecture and interior design (M = 4.29) and staff friendliness (M = 4.23). Architecture and design was the most consistent result: 92% of respondents agreed that it was impressive, and it had the lowest standard deviation of all items (SD = 0.61).

The lowest-rated features were all functional parts of the journey: transport from the airport to the city (M = 3.97), ease of finding the way (M = 4.03) and bilingual Arabic and English signage (M = 4.05). These items also had larger standard deviations, which indicates that experiences of navigation varied more between respondents than experiences of the building itself.

Table 4. Airport experience ratings (1 = strongly disagree, 5 = strongly agree)

Statement	Mean	SD	Agree (%)	DXB	AUH
B2 The airport was clean and comfortable	4.34	0.87	83	4.24	4.44
B1 Terminal architecture and interior design were impressive	4.29	0.61	92	4.32	4.26
B7 Airport staff were friendly and welcoming	4.23	0.83	86	4.30	4.16
B10 Overall, I had a positive experience at the airport	4.22	0.82	85	4.24	4.20
B8 There was a good choice of shops and food outlets	4.18	0.85	84	4.20	4.16
B5 Immigration and passport control were quick and smooth	4.12	1.07	81	4.04	4.20
B6 Baggage claim was fast and well organised	4.08	0.90	85	3.96	4.20
B3 Signs in Arabic and English were clear and easy to follow	4.05	1.00	79	4.08	4.02
B4 It was easy to find my way through the airport	4.03	0.95	75	4.00	4.06
B9 It was easy to find transport from the airport to the city	3.97	0.89	75	3.88	4.06

Note: Agree (%) = share answering 4 (agree) or 5 (strongly agree). DXB and AUH columns show means for each airport (n = 50 each). No airport difference was significant (Mann–Whitney U, p > .05).

The open-ended answers confirm what travellers notice first (Table 5). The size and scale of the terminal was the most frequent first impression (23 mentions), followed by the architecture, design and lighting (17) and the crowds and international mix of people (17). Typical answers were “the size of the airport”, “the ceiling and the open space” and “the architecture caught my attention first”. Signage, screens and technology (13) and shops and food (10) came next, while staff were mentioned first by only five respondents.

First-time users experienced the airport differently from repeat users. They rated staff friendliness significantly higher (M = 4.46 compared with 4.08; p = .016), which suggests that a friendly welcome matters most when the place is unfamiliar. They also rated bilingual signage lower (M = 3.74 compared with 4.25), although this difference was not statistically significant (p = .11).

Ratings were also compared by type of respondent and type of trip. None of the differences between UAE nationals, non-national residents and visitors was statistically significant. Visitors gave the highest ratings to the UAE as well organised and efficient (M = 4.40) and were the group most likely to recommend the country (M = 4.44), while UAE nationals rated bilingual signage more highly (M = 4.41) than residents (3.90) and visitors (3.98). By type of trip, staff friendliness differed significantly (Kruskal–Wallis, p = .021): it was rated highest by transit passengers (M = 4.55) and arriving passengers (M = 4.36) and lowest by departing passengers (M = 3.94). Transit passengers also gave the lowest rating to signage (M = 3.45), although this group was small (n = 11). Together with the first-time results, this suggests that staff and information matter most to the travellers who are least familiar with the terminal.

4.4. Airport experience and image of the UAE (RQ2)

Respondents held a very positive image of the UAE after their airport experience. The UAE was seen above all as modern and advanced (M = 4.32), welcoming and hospitable (M = 4.32) and well organised and efficient (M =

4.27). Rich in culture and heritage was the lowest-rated image attribute (M = 4.04). Most respondents (79%) agreed that their airport experience improved their overall impression of the UAE (D6, M = 4.15), and 86% intended to visit or return. The strongest result was recommendation: 91% would recommend the UAE to friends and family (D9, M = 4.34). Figure 1 shows the image ratings for each airport.

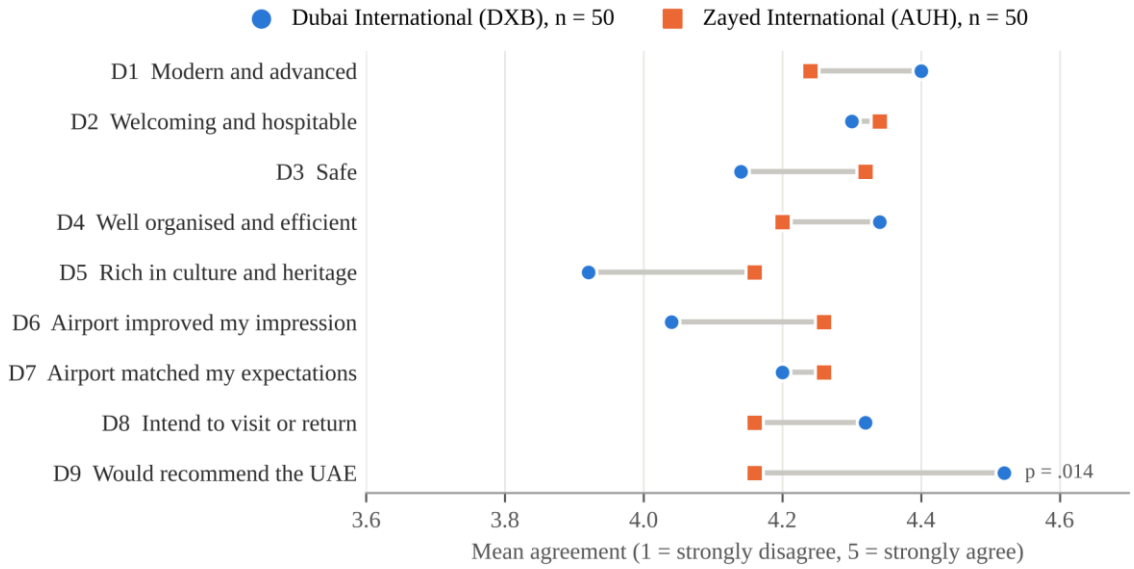


Fig. 1. Image of the UAE after the airport experience, mean agreement by airport (n = 50 per airport). The axis is cut at 3.6 to show the differences; p is reported only for the significant difference (Mann–Whitney U).

The open-ended answers show how airport features were translated into this image. When asked what they liked most, respondents most often named the design, space and atmosphere of the terminal (34 mentions), followed by shops, food and duty free (24), organisation and efficiency (18), clear information and wayfinding (15) and friendly, helpful staff (15). One respondent wrote: “It felt spacious and modern, and the lighting made the interior look attractive.” Another said: “The staff were polite and willing to help … I appreciated being able to ask questions when I wasn't sure where to go.” The three words used most often to describe the airports were *modern* (49 mentions), *spacious* (29), *clean* (24), *busy* (23) and *organised* (18). These words closely match the image attributes rated highest in the survey, namely modern, efficient and well organised, which suggests that the terminal acts as a compressed version of the national image.

4.5. Dubai International compared with Zayed International (RQ3)

The two airports created broadly similar first impressions. None of the ten airport experience ratings differed significantly between them (Table 4). However, the pattern of means points to two different profiles. Dubai International scored slightly higher on architecture and design (4.32 compared with 4.26) and staff friendliness (4.30 compared with 4.16), while Zayed International scored higher on cleanliness and comfort (4.44 compared with 4.24), baggage claim (4.20 compared with 3.96), immigration (4.20 compared with 4.04) and transport to the city (4.06 compared with 3.88).

The image results follow the two cities' brand identities (Figure 1). Travellers who used Dubai International rated the UAE higher as modern and advanced (4.40 compared with 4.24) and as well organised (4.34 compared with 4.20), and they were significantly more likely to recommend the UAE (4.52 compared with 4.16; p = .014). Travellers who used Zayed International rated the UAE higher as rich in culture and heritage (4.16 compared with 3.92) and as safe (4.32 compared with 4.14), and they were more likely to say that the airport improved their impression of the country (4.26 compared with 4.04), although these differences were not statistically significant. The three-word descriptions show the same contrast: DXB was more often called *stylish* and *big*, while AUH was more often described as *efficient*, *lively*, *bright* and *welcoming*.

4.6. What the airports could improve

When asked what the airport could improve, respondents focused on the needs of people who do not know the terminal (Table 5). The most common suggestion was clearer signage and more information for first-time visitors (34 mentions), for example "clearer instructions for first time visitors, especially around connections and terminal transfers". Respondents also asked for more seating and rest areas, especially near gates and along long walking routes (27), more affordable food (14) and more visible staff and assistance points (10). Seven respondents asked for more local culture and tourist information, such as "more information about local attractions" or a welcome ceremony for arrivals.

Table 5. Main themes in the open-ended answers (number of mentions)

What travellers noticed first	n	What the airport could improve	n
Size and scale of the terminal	23	Clearer signage and information for first-time visitors	34
Architecture, design and lighting	17	More seating and rest areas	27
Crowds and international mix of people	17	More affordable food and shops	14
Signage, screens and technology	13	More visible staff and assistance points	10
Shops and food	10	Shorter queues and waiting times	7
Cleanliness	7	Easier transport and shorter walking distances	7
Aircraft and runway	6	More local culture and tourist information	7
Staff and welcome	5		

Note: n = 100 respondents; one answer can belong to more than one theme.

5. DISCUSSION

The findings support the view that airports act as the first impression of a destination. Regarding RQ1, travellers noticed the physical environment of the terminal first and remembered it most positively. Size, architecture, lighting and cleanliness dominated both the ratings and the spontaneous answers. This is consistent with research showing that the airportscape creates a sense of place and that the physical environment is a key dimension of the airport experience (Bezerra & Gomes, 2016; Wattanacharoensil et al., 2022). At the same time, the lowest ratings and most improvement suggestions concerned functional parts of the journey: signage, wayfinding and transport to the city. In line with Fodness and Murray (2007), passengers judge airports on function as well as on atmosphere. The observation points in the same direction. Both terminals received the highest rating for architecture, facilities and processing, while signage at Zayed International was rated only 3, although travellers rated signage similarly at both airports. The results suggest that UAE airports already succeed in impressing visitors, but they could do more to guide them.

Regarding RQ2, the airport experience was closely aligned with the image of the UAE. The attributes that travellers associated with the country, namely modern, welcoming and well organised, were the same qualities they used to describe the terminals. Most respondents said that the airport improved their impression of the UAE, and intentions to return and recommend were high. This supports earlier findings that airport service quality, airport image and the airport environment influence destination image and intention to visit (Saut & Song, 2022; Fakfare, 2023; Abouseada et al., 2023). It also reflects the UAE's heritage-modernity image (Haq et al., 2021; Williams, 2022), but the balance leans clearly towards modernity: culture and heritage was the weakest image attribute, and only a few respondents mentioned culture spontaneously. The observation confirms this: welcome messages and art and decoration were the weakest areas at both airports, each rated 3 out of 5. This suggests that the terminals communicate the modern side of the UAE much more strongly than its heritage, a gap also observed in Saudi airports, where globalised design can hide local identity (Elmoghazy et al., 2025).

Regarding RQ3, the two airports created similar overall impressions, but each reinforced the identity of its own city. Dubai International, the larger hub, was associated more with modernity, organisation and recommendation, which matches Dubai's image as a global and dynamic city. Zayed International, the smaller hub, performed better on comfort and processing and was associated more with culture, heritage and safety, which matches Abu Dhabi's image as a cultural capital (Slak Valek & Williams, 2018; Williams, 2022). The observation adds that Dubai International offers stronger signage and shopping, while Zayed International offers a stronger food and drink

experience. Only the difference in recommendation was statistically significant in the survey, so these patterns should be read as tendencies. Even so, they suggest that each airport acts as a gateway not only to the UAE but to a specific emirate and its brand.

Finally, the experience of first-time visitors deserves attention. They valued friendly staff more highly than repeat users and found signage harder to follow. This is consistent with the view that arrival is the stage of greatest uncertainty for unfamiliar passengers (Wattanacharoensil et al., 2017). It also suggests that a first impression depends on a combination of human welcome and clear information.

From a theoretical perspective, the results extend the airport experience framework of Wattanacharoensil et al. (2016) to the Gulf context. They show that the airport is not only a place where the destination image is confirmed but also a place where it is selectively edited: the terminal amplifies some parts of the national brand, such as modernity and efficiency, more than others, such as heritage. The words that travellers used to describe the terminals closely matched the attributes they later attributed to the country. This suggests that image transfer from the airport to the destination happens through a small number of strong, easily noticed cues, such as scale, light, cleanliness and a friendly greeting, rather than through every element of the service chain. Future models of destination image formation could therefore treat the airport as a distinct stage of direct experience, located between the image formed before travel and the image formed during the stay.

6. CONCLUSION AND RECOMMENDATIONS

This paper examined how Dubai International Airport and Zayed International Airport act as the first impression of the UAE. Based on a two-step method of terminal observation and a survey of 100 travellers, it shows that both airports create a very positive first impression, driven mainly by impressive architecture, cleanliness, efficient processing and friendly staff. Both the observation and the survey show that symbolic cues of national identity, such as welcome messages, local art and cultural displays, are the weakest part of this first impression. This impression is closely aligned with the UAE's image as modern, welcoming and well organised, and most travellers said that the airport improved their view of the country. The larger hub, Dubai International, projected a more modern image and generated stronger recommendation, while the smaller hub, Zayed International, was rated slightly better for comfort, processing and cultural image.

The study offers several practical recommendations for airport operators and tourism authorities:

- Treat the arrival journey, from the aircraft door to ground transport, as part of the destination product rather than only as an operational process.
- Improve wayfinding for first-time visitors through clearer bilingual signage, information points near immigration and baggage claim, and clear guidance on transport to the city.
- Make staff visible and approachable in arrival areas, since a friendly welcome matters most to first-time visitors.
- Add seating and rest areas near gates and along long walking routes.
- Strengthen the heritage side of the national image by making Emirati culture more visible in the terminal through art, calligraphy, local food, Arabic coffee and tourist information about the city.
- Allow each airport to reflect the identity of its own emirate, so that Dubai International and Zayed International offer distinct but complementary first impressions of the UAE.

The study contributes to the literature by providing UAE-specific evidence on what travellers notice in airport terminals and how the airport experience relates to destination image, an area in which direct evidence has been limited. It also has limitations. The sample of 100 respondents was collected through convenience and snowball sampling and is not representative of all travellers using UAE airports. The analysis was descriptive and based on individual survey items, so it does not establish causal relationships. Observation was limited to public and arrival areas and to a small number of visits. Future research could collect larger samples at the airports themselves, include other Gulf hubs for comparison, and test the relationships between airport features and destination image with multivariate methods such as structural equation modelling.

Declaration of generative AI use

The authors used Claude (Anthropic) to assist with editing the text, formatting the manuscript, and preparing figures and tables. The authors reviewed all content and take full responsibility for the paper.

References

[1] Abouseada, A. A. A. H., Hassan, T. H., Saleh, M. I., & Radwan, S. H. (2023). The power of airport branding in shaping tourist destination image: Passenger commitment perspective. *GeoJournal of Tourism and Geosites*, *47*(2), 440–449. https://doi.org/10.30892/gtg.47210-1042

[2] Al Saed, R., Upadhya, A., & Abu Saleh, M. (2020). Role of airline promotion activities in destination branding: Case of Dubai vis-à-vis Emirates Airline. *European Research on Management and Business Economics*, *26*(3), 121–126. https://doi.org/10.1016/j.iedeen.2020.07.001

[3] Baloglu, S., & McCleary, K. W. (1999). A model of destination image formation. *Annals of Tourism Research*, *26*(4), 868–897. https://doi.org/10.1016/S0160-7383(99)00030-4

[4] Bezerra, G. C. L., & Gomes, C. F. (2016). Measuring airport service quality: A multidimensional approach. *Journal of Air Transport Management*, *53*, 85–93. https://doi.org/10.1016/j.jairtraman.2016.02.001

[5] Booranakittipinyo, A., Li, R. Y. M., & Phakdeephirot, N. (2024). Travelers' perception of smart airport facilities: An X (Twitter) sentiment analysis. *Journal of Air Transport Management*, *118*, 102600. https://doi.org/10.1016/j.jairtraman.2024.102600

[6] Chi, C. G.-Q., & Qu, H. (2008). Examining the structural relationships of destination image, tourist satisfaction and destination loyalty: An integrated approach. *Tourism Management*, *29*(4), 624–636. https://doi.org/10.1016/j.tourman.2007.06.007

[7] Dubai Media Office. (2026, February 11). *DXB sets new global benchmark as record traffic volumes become the norm*. https://mediaoffice.ae/en/news/2026/february/11-02/dxb-sets-new-global

[8] Elmoghazy, Z. A. A. E., Afify, H. M. N., & Alzenifeer, B. M. (2025). Redefining airport terminal design through identity indicators: Case studies from airport terminals in Saudi Arabia. *Buildings*, *15*(8), 1261. https://doi.org/10.3390/buildings15081261

[9] Fakfare, P. (2023). Examining the effect of airport environment on perceived image, memorable experiences and passenger relational behaviours. *Advances in Hospitality and Tourism Research*, *11*(1), 72–96. https://doi.org/10.30519/ahtr.1051578

[10] Fodness, D., & Murray, B. (2007). Passengers' expectations of airport service quality. *Journal of Services Marketing*, *21*(7), 492–506. https://doi.org/10.1108/08876040710824852

[11] Goetz, A. R. (2019). The airport as an attraction: The airport city and aerotropolis concept. In A. Graham & F. Dobruszkes (Eds.), *Air transport – A tourism perspective* (pp. 217–232). Elsevier. https://doi.org/10.1016/B978-0-12-812857-2.00016-6

[12] Haq, F., Seraphim, J., & Medhekar, A. (2021). Branding heritage tourism in Dubai: A qualitative study. *Advances in Hospitality and Tourism Research*, *9*(2), 243–267. https://doi.org/10.30519/ahtr.782679

[13] Khaleej Times. (2026, January 29). *Abu Dhabi Airports closes 2025 with record traffic*. https://www.khaleejtimes.com/business/aviation/abu-dhabi-airports-closes-2025-with-record-traffic

[14] Martin-Domingo, L., Martín, J. C., & Mandsberg, G. (2019). Social media as a resource for sentiment analysis of Airport Service Quality (ASQ). *Journal of Air Transport Management*, *78*, 106–115. https://doi.org/10.1016/j.jairtraman.2019.01.004

[15] Pike, S., Pontes, N., & Kotsi, F. (2021). Stopover destination attractiveness: A quasi-experimental approach. *Journal of Destination Marketing & Management*, *19*, 100514. https://doi.org/10.1016/j.jdmm.2020.100514

[16] Reisinger, Y., Michael, N., & Hayes, J. P. (2019). Destination competitiveness from a tourist perspective: A case of the United Arab Emirates. *International Journal of Tourism Research*, *21*(2), 259–279. https://doi.org/10.1002/jtr.2259

[17] Saut, M., & Song, V. (2022). Influences of airport service quality, satisfaction, and image on behavioral intention towards destination visit. *Urban, Planning and Transport Research*, *10*(1), 82–109. https://doi.org/10.1080/21650020.2022.2054857

[18] Slak Valek, N., & Williams, R. B. (2018). One place, two perspectives: Destination image for tourists and nationals in Abu Dhabi. *Tourism Management Perspectives*, *27*, 152–161. https://doi.org/10.1016/j.tmp.2018.06.004

[19] Usman, A., Azis, Y., Harsanto, B., & Azis, A. M. (2022). Airport service quality dimension and measurement: A systematic literature review and future research agenda. *International Journal of Quality & Reliability Management*, *39*(10), 2302–2322. https://doi.org/10.1108/IJQRM-07-2021-0198

[20] Wattanacharoensil, W., Fakfare, P., & Graham, A. (2022). Airportscape and its effect on airport sense of place and destination image perception. *Tourism Review*, *77*(2), 549–569. https://doi.org/10.1108/TR-01-2021-0036

[21] Wattanacharoensil, W., Schuckert, M., & Graham, A. (2016). An airport experience framework from a tourism perspective. *Transport Reviews*, *36*(3), 318–340. https://doi.org/10.1080/01441647.2015.1077287

[22] Wattanacharoensil, W., Schuckert, M., Graham, A., & Dean, A. (2017). An analysis of the airport experience from an air traveler perspective. *Journal of Hospitality and Tourism Management*, *32*, 124–135. https://doi.org/10.1016/j.jhtm.2017.06.003

[23] Williams, R. B. (2022). Same same but different: The destination image of Dubai and Abu Dhabi on YouTube. *International Journal of Technology Marketing*, *16*(4), 332–348. https://doi.org/10.1504/ijtmkt.2022.126279

[24] Ziakas, A., Vlassi, E., Pagoni, I., & Papatheodorou, A. (2026). Air transport, tourism and culture: Exploring the triangular relationship between airlines, airports, and tourism destinations. In *Reference module in social sciences*. Elsevier. https://doi.org/10.1016/B978-0-443-27698-9.00080-5

Developing Strategies for Cultivating a Culture of Learning in Educational Institutions

Fethi KAYALAR[a*]

[a]*Prof. Dr., (Erzincan B. Y. University), Faculty of Education,TURKEY* fkayalar@erzincan.edu.tr

Abstract

Learning is a process undertaken to acquire knowledge that brings about change in the learner. It should become a habit—a part of the culture—for students. Strategies to cultivate a learning culture within educational institutions include promoting literacy, fostering learning motivation, employing repetition, ensuring focused attention and effective participation, facilitating stages of attitudinal and behavioral transformation, and recognizing the vital role of teachers and mentors. The formation of an educational culture will provide a strong foundation for the learning culture within the institution, as the educational culture in an educational institution will also be reflected in the learning culture. The advancement of educational institutions is firmly supported by a culture of high values—reflected in traditional behaviors such as solidarity, discipline, exemplary conduct, strong commitment, consistency, perseverance, unity in building quality, and the preservation of academic values. In the study, developing strategies for cultivating a culture of learning in educational institutions were examined under the light of literature. It has been concluded that promoting literacy, enhancing learning motivation, building focused attention, effective participation, transformation of attitudes and behaviors, and presence of teachers and mentors are of great importance for cultivating a culture of learning.

Keywords: Learning Culture, Educational Management, Quality of Education, Cultivation of Learning

1. INTRODUCTION

Establishing a learning culture in educational institutions is considered a vital necessity for enhancing educational quality, supporting teachers' professional development, and ensuring students' active participation in learning processes. A learning culture signifies that institutional members embrace continuous learning, knowledge sharing, collaboration, innovation, and development as shared values. The cultivation of this culture depends not merely on individual efforts but is closely linked to the institution's leadership approach, organizational structure, communication patterns, and evaluation processes. This study aims to elucidate the conceptual foundations of learning culture in educational institutions and to discuss key strategies for fostering it. Accordingly, strategies such as creating a shared vision, developing professional learning communities among teachers, fostering an environment of psychological safety, strengthening instructional leadership, utilizing data-driven decision-making, increasing student and parent engagement, and institutionalizing continuous improvement processes are examined. Ultimately, it is argued that

* Corresponding author

establishing a sustainable learning culture relies on the holistic implementation of participatory leadership, trust-based relationships, shared responsibility, and systematic evaluation processes.

Rapid advancements in knowledge and technology necessitate that educational institutions evolve from entities that merely transmit existing knowledge and skills into organizations that engage in continuous learning, self-renewal, and adaptation to changing needs. In this context, the quality of educational institutions is determined not only by the implementation of curricula or students' academic achievements but also by the extent to which values, relationships, and practices that foster learning are embedded within the institution. A learning culture encompasses the shared understandings and practices that support individuals and groups in acquiring new knowledge, leveraging their experiences, sharing information, and transforming the learning gained into organizational development. Fostering this culture within educational institutions is crucial for enhancing teachers' professional competencies, encouraging students to take responsibility for their own learning, and enabling administrators to focus on improving teaching and learning processes.

However, establishing a learning culture is not limited merely to organizing in-service training activities or implementing new teaching methods. This process involves a multidimensional structure that encompasses the attitudes toward learning, communication styles, decision-making processes, and responsibilities regarding development held by all institutional stakeholders. In institutions where learning is encouraged, mistakes are viewed as opportunities for growth, and diverse viewpoints can be openly expressed, a strengthening of collective learning capacity is to be expected.

This study outlines the conceptual framework of learning culture in educational institutions, explains the key factors influencing the development of learning culture, and discusses strategies that can be implemented to establish a sustainable learning culture. In this context, answers to the following questions were sought:

1. What are the fundamental components of the learning culture in educational institutions?
2. What roles do school administrators and teachers play in the development of the learning culture?
3. What strategies can be implemented to establish a sustainable learning culture in educational institutions?
4. Through which indicators can the development of the learning culture be evaluated?

2. THE CONCEPT OF LEARNING CULTURE

Learning culture can be defined as a set of shared values, norms, and practices within an institution that prioritizes learning, supports knowledge sharing, encourages individual and collective development, and ensures the systematic utilization of experiences. This culture aims to ensure that learning is not limited to training activities conducted at specific times but becomes an integral part of the institution's daily operations and decision-making processes.

In the "learning organization" approach, Senge (1990) emphasizes the importance of individuals developing their capacity to learn together and acting in alignment with a shared vision of the future. This approach indicates that, in order to adapt to changing conditions, educational institutions must enhance not only individual competencies but also organizational learning processes.

Watkins and Marsick (1993) address the development of learning organizations through dimensions such as continuous learning, inquiry and dialogue, team learning, systems connectivity, and leadership that supports learning. This framework demonstrates that learning culture cannot be explained solely by individuals' desire to learn; rather, it must be evaluated in conjunction with organizational processes and leadership practices.

For educational institutions, learning culture is grounded in a holistic approach that does not separate teachers' professional development from students' learning experiences. In this approach, the school is viewed as a community where not only students but also teachers and administrators engage in continuous learning.

2.1. Key components of a learning culture

The key components of a strong and sustainable learning culture are as follows:

- Psychological Safety and the Courage to Make Mistakes: A safe environment where individuals can ask questions, generate ideas, and experiment without fear of judgment. Mistakes are viewed as opportunities for learning rather than grounds for punishment.

- Shared Vision and Prioritization of Learning: Positioning learning not merely as a periodic obligation such as mandatory training sessions, but as an indispensable means of achieving individual and organizational goals.

- Open, Two-Way Communication and Feedback: The active implementation of mutual, constructive, and growth-oriented feedback mechanisms, regardless of hierarchy (e.g., superior-subordinate or teacher-student distinctions).

- Easily Accessible Learning Resources: Convenient access to information, training materials, mentorship programs, and digital resources such as LinkedIn Learning, independent of time or location.

- Leadership Role and Coaching Approach: Managers or educators acting as role models who not only demand results but also engage in continuous learning, champion development, and guide their teams through coaching.

- Knowledge Sharing and Collaboration: Transparently disseminating acquired knowledge, experiences, and best practices across the team or organization, rather than hoarding them within silos.

- Growth Mindset: Adopting a collective belief and mindset that abilities and intelligence are not fixed but can be continuously developed through effort, experience, and feedback.

The development of a learning culture relies on various interconnected components. Foremost among these are a shared vision, collaboration, psychological safety, continuous professional development, knowledge sharing, and processes for monitoring progress.

A shared vision enables institution members to develop a common understanding regarding the purpose and priorities of learning. Collaboration contributes to transforming individual experiences into collective knowledge. Psychological safety encourages employees to ask questions, express diverse viewpoints, and share their mistakes. While continuous professional development supports teachers in keeping their pedagogical knowledge and skills up to date, systematic evaluation allows for the analysis of the impact of practices and the implementation of necessary improvements.

The mutual reinforcement of these components facilitates the adoption of learning as an institutional value and ensures its continuity.

3. STRATEGIES FOR DEVELOPING LEARNING IN EDUCATIONAL INSTITUTIONS

How is a learning culture developed? It must align with the principles of learning. Through the consistent and continuous application of learning principles, these practices become habits and subsequently evolve into strategies for developing a learning culture. Strategies for developing a learning culture in primary education will differ somewhat from those implemented in higher education.

As stated by Adz-Dzakiey in his book Prophetic Psychology (2008), strategies for developing a learning culture in primary education include fostering learning motivation, repetition, focused attention, effective participation, the stages of transforming attitudes and behaviors, and the importance of teachers and mentors. One of these six elements is fostering literacy.

According to the author, these seven elements are strategies for developing a learning culture in educational institutions:

- Establishing a shared vision and common goals

The first step in developing a learning culture is for the institution to formulate a shared vision regarding learning. This shared vision should not consist solely of goals set by administrators; rather, it should be shaped by the contributions of teachers, students, parents, and other relevant stakeholders.

When institution members embrace common goals, learning activities are integrated toward a specific developmental objective rather than being conducted as isolated, independent practices. In this context, educational institutions should define annual development goals, identify those responsible for these goals, and regularly evaluate progress. For a shared vision to be effective, it is important that it does not remain limited to abstract statements. For

instance, observable goals can be established, such as increasing student participation, fostering a culture of collaborative planning among teachers, or alleviating difficulties associated with specific learning outcomes.

- Promoting Literacy

In the Great Dictionary, literacy is defined as the ability to read and write; knowledge or skills in a specific field or activity; an individual's capacity to process information and apply it to life skills; and the use of letters to represent sounds or words.

Simply put, literacy can be defined as the ability to read and write. However, in modern times, literacy has taken on a broader meaning; it encompasses a variety of concepts rather than just a single definition. For instance, there are various types of literacy, such as computer literacy, media literacy, technology literacy, economic literacy, information literacy, and even moral literacy.

In general, several purposes for reading activities has been outlined, including: (1) reading for pleasure, which does not involve a complex thought process. This activity is typically undertaken to fill leisure time; examples include reading novels, newspapers, magazines, or comic books; (2) reading to enhance knowledge and insight, such as reading textbooks or scholarly works; and (3) reading to support the performance of a job or profession—for example, reading books on practical technical skills or general knowledge (popular science).

Promoting literacy can be achieved by providing reading materials in educational institutions—such as general history books, stories or short stories, illustrated tales, and other books suitable for students at the basic education level or other educational stages. The approach to literacy must be tailored to the specific needs of different levels, distinguishing between primary (basic) education and secondary education.

Reading activities need not be confined to textbooks; rather, they should focus on the children's own preferences regarding how they wish to read within the school or out-of-school setting. Furthermore, to foster reading habits among children, literacy resources should be made available in every corner of the classroom or school, and at least 30 minutes should be set aside each morning before lessons begin for children to read available literature based on their preferences or interests. Students are also encouraged to use their free time at school to engage with available literature.

Regarding literacy, the governments all over the world have launched a movement to promote literacy. One such initiative, the school literacy movement, is implemented by integrating it into curricular, co-curricular, and extracurricular activities. This movement can be carried out either inside or outside the classroom with the support of parents and the community.

- Developing professional learning communities

Teacher professional development is a fundamental aspect of the learning culture. Traditional professional development practices often rely on individual participation and short-term training activities. In contrast, professional learning communities encourage teachers to meet regularly around shared goals, examine their instructional practices, and make collaborative decisions aimed at enhancing student learning.

DuFour (2004) emphasizes that professional learning communities should focus on student learning rather than merely on instructional processes. Accordingly, it is important for teachers to evaluate not only the topics they cover but also what students are actually learning, the underlying causes of learning difficulties, and which students require additional support.

To ensure the effective operation of professional learning communities in educational institutions, regular meeting times must be established, meetings should be grounded in shared data and student work, and the outcomes of implemented decisions must be monitored. Furthermore, the scope of knowledge sharing can be broadened through interdisciplinary initiatives that bring together teachers from different subject areas.

- Enhancing Learning Motivation

Learning facilitated by educators in educational institutions must foster or encourage student motivation. Motivation to learn is crucial for students' enthusiasm for learning and acquiring knowledge. The concept of motivation stems from the idea of a "force"—defined as the driving power within an individual that compels them to undertake specific activities to achieve a goal. Motivation is defined as an energy shift within an individual, characterized by the emergence of an "emotion" and initiated by a response to a goal (Zainal and Bahar, 2013). Based on this understanding, learning motivation can be defined as the general driving force within students that initiates, sustains, and directs learning activities, thereby leading to the achievement of learning objectives. Motivation is vital in learning activities, particularly within educational institutions. Students—especially children—require motivation to learn. Without motivation, effective learning activities that achieve educational goals would be impossible.

Motivation falls into two categories: intrinsic motivation and extrinsic motivation. Intrinsic motivation originates from within the individual, arising from their own volition without external coercion or encouragement. Extrinsic motivation, on the other hand, arises from external influences—such as invitations, commands, or pressure from others—that encourage students to perform a task or engage in learning. The ability to provide motivation and implement strategies that effectively foster it will significantly impact student learning. In this context, the key task for teachers is to harness students' motivation in order to accelerate the development of their abilities and achieve desired learning objectives.

- Creating a learning environment based on psychological safety

The sustainability of a learning culture is closely linked to the extent to which organizational members feel safe. If employees believe they will be penalized for making mistakes, asking questions, or expressing dissenting views, knowledge sharing and innovative behaviors may be stifled.

Edmondson (1999) defines psychological safety as a shared belief that it is safe to take interpersonal risks within a team. When this approach is adapted to educational institutions, it becomes crucial for teachers to be able to share the challenges they face in their lessons, critically evaluate their practices, and seek support from their colleagues.

To foster psychological safety, administrators must encourage open communication, ensure that feedback processes focus on understanding problems rather than assigning blame, and respect diverse viewpoints. However, an environment of trust should not imply the elimination of accountability. The primary goal is to create an environment that enables learning from mistakes and fosters development while maintaining performance expectations.

- Establishing Focused Attention

Attention is crucial in learning activities. Sugihartono et al. stated that attention is a vital factor influencing learning because it captures students' interest and enables them to easily comprehend the material presented by the teacher. Gage and Berliner expressed a similar view, noting that attention plays a critical role in learning activities, as the learning process is impossible without it. Therefore, students must maintain their attention throughout the learning process (Irachmat, 2015).

Learning experiences that are both effective and aligned with planned educational goals will enhance students' cognitive, affective, and psychomotor abilities. However, to develop these abilities, students must remain highly focused on the subject matter and be diligent in their learning efforts.

- Strengthening instructional leadership and administrative support

School administrators play a pivotal role in shaping the learning culture. If administration focuses solely on operational matters, efforts to improve the quality of instruction may be relegated to the background. In contrast, the instructional leadership approach highlights the responsibilities of school administrators to define learning goals, support instructional processes, and monitor student development.

Hallinger (2003) examines the relationship between instructional leadership and a school's teaching and learning processes within the context of school leadership research. In this context, to foster a learning culture, administrators need to provide teachers with professional development opportunities, allocate time for collaborative planning, direct resources according to learning needs, and offer constructive feedback on instructional practices.

Furthermore, ensuring teacher participation in decision-making processes can strengthen institutional ownership. Participatory leadership can help employees move beyond merely implementing decisions to becoming active stakeholders in institutional development.

- Effective Participation

Effective learning is learning that provides students with opportunities to participate in activities. Modern learning requires students to work and practice as they learn. Through such work, students acquire knowledge and develop skills. Meaningful learning occurs when students actively participate in the learning process; they do not merely memorize concepts presented by the teacher but also engage in practical activities. Consequently, teachers need to create situations that encourage active student participation, enabling students to explore and expand upon the foundational knowledge imparted by the teacher.

The direct participation of students and educators in the learning process can foster a shared experience for every student. Furthermore, students feel appreciated by their teachers because the teacher engages with each of them individually—greeting them and inquiring about their level of understanding regarding the material. This interaction between teachers and students generates a renewed enthusiasm for learning.

Educators consistently involve students, respect their opinions, and view them as responsible, valued individuals capable of solving problems encountered in the classroom (Sagala, 2015). Direct student participation is understood as active and effective engagement in learning activities.

- Enhancing data-driven decision-making and feedback processes

To institutionalize a culture of learning, it is crucial that decisions are not based solely on personal assumptions or past experiences but rather on a holistic evaluation of diverse data sources. Student achievement, formative assessment results, attendance records, feedback from students and parents, and teacher observations can all contribute to institutional development processes.

Data-driven decision-making requires interpreting information within its specific context rather than merely applying the data mechanically. For instance, a failure to master a specific learning objective in a classroom should not be viewed simply as student failure; the situation should also be examined in terms of teaching methods, learning materials, prior knowledge, and student needs.

In this context, it is important to hold regular evaluation meetings, prepare intervention plans for identified issues, and monitor the outcomes of these interventions. In this way, assessment can evolve from a process that merely measures past performance into a mechanism that actively enhances learning.

- The Presence of Teachers and Mentors

The concept of teachers and mentors signifies that teachers are not merely ordinary individuals; rather, they are keys to understanding knowledge and science, sources of role models, and parental figures who foster and nurture the mental, spiritual, moral, and social development of every individual.

The presence of teachers and mentors enhances and boosts children's enthusiasm for learning. This heightened attention from teachers and mentors supports children in every respect, including their mental, emotional, and spiritual well-being, as well as other aspects of their growth and development. Therefore, the presence of teachers and mentors is a crucial strategy for cultivating a learning culture among students in educational institutions.

4. CONCLUSION

Strategies for fostering a learning culture in educational institutions point to an organizational development approach that extends beyond individual learning activities. Transforming learning into an institutional value depends on the mutual reinforcement of leadership, trust, collaboration, professional development, and evaluation processes.

The "learning organization" approach offers a significant conceptual foundation for educational institutions to adapt to changing conditions and transform individual learning into collective capacity. Similarly, professional learning communities enable teachers to enhance their individual experiences through processes of collaborative reflection and problem-solving. However, establishing a learning culture may encounter various challenges. These include a lack of time, heavy workloads, centralized decision-making processes, resistance to change, resource limitations, and the use of evaluations in a punitive manner. Therefore, strategies must be designed in alignment with the institution's specific context, the needs of its staff, and available resources.

Furthermore, the development of a learning culture should not be assessed based solely on short-term outcomes. Strengthening institutional trust, embedding collaboration among teachers, and fostering students' sense of responsibility for their own learning are processes that require time. Continuity, stakeholder engagement, and a commitment to learning from practices are crucial factors in these processes.

In this context, establishing a learning culture should be viewed as a transformation of the institution's way of working rather than merely the implementation of a specific program. To ensure the effectiveness of strategies, consistency must be maintained among institutional goals, daily practices, and evaluation mechanisms.

As conclusion, establishing a learning culture within educational institutions is a crucial organizational requirement for the sustainable improvement of educational quality. A learning culture develops through the integration of components such as a shared vision, collaboration, psychological safety, continuous professional development, participatory leadership, data-driven decision-making, and continuous improvement.

The strategies addressed in this study demonstrate that learning should be structured as a process that encompasses all institutional stakeholders and is supported by daily practices. In this regard, it is essential for school administrators to adopt a leadership approach that fosters learning, for teachers to engage in collaborative learning processes, and for students to take responsibility for their own learning.

References

[1] Adz-Dzakiey H. B.(2008). *Psikologi Kenabian: Prophetic Psychology,* Yogyakarta: Al-Manar, 2008.

[2] DuFour, R. (2004). What is a "professional learning community"? Educational Leadership, 61(8), 6–11.

[3] Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly,* 44(2), 350–383. https://doi.org/10.2307/2666999

[4] Gage, N. L.and Berliner, D. C. (1992). *Educational psychology,* fifth edition, Boston : Houghton Mifflin

[5] Hallinger, P. (2003). Leading educational change: Reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education,* 33(3), 329–352. https://doi.org/10.1080/0305764032000122005

[6] Irachmat, M.R. (2015). Increasing Grade 3 Students' Attention During the Learning Process Through Ice-Breaking Games Di SDN Gembongan, *Jurnal Pendidikan Guru Sekolah Dasar* (Edisi 2 Tahun Ke IV.

[7] Sagala, R. (2015). The Effectiveness of STEM-Based on Gender Differences: The Impact of Physics Concept Understanding. *European Journal of Educational Research,* Volume 8, Issue 3, 753 - 761.

[8] Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization.* Doubleday. London

[9] Sugihartono, T., Anandri, A. R. F., Dian P., Nurur F. M., Ningrum, Arrendi A.& Nopiyanto, Y. O. (2026). Strategies for Managing Students with Physical Disabilities in Physical Education, Sports, and Health Classes at Public Junior High Schools in Bengkulu City. Sport Gymnatics: *Jurnal Ilmiah Pendidikan Jasmani.* Vol. 7 No. 1, April 2026, pp. 105 - 132

[10] Watkins, K. E., & Marsick, V. J. (1993). *Sculpting the learning organization: Lessons in the art and science of systemic change*. Jossey-Bass. New York.

[11] Zainal, V. R. and Bahar, F. (2013). *Islamic Education Management: Dari Teori ke Praktik*, Jakarta: Rajawali Pers

Knowing and Self-Knowing in the Teaching-Learning Process

Pavol BURCL[a*]

[a]*Constantine the Philosopher University in Nitra, Faculty of Education, The Department of English Language and Culture, Dražovská cesta 4, 94901 Nitra, Slovakia, pburcl@ukf.sk*

Abstract

The author focuses on certain overlooked issues in the teaching-learning process. Teachers must know themselves first, so they are able to know their learners. The process of self-knowing is discussed. In this way, a learner-centred approach means that learners themselves can also be left to self-know, strengthening their learner profile and understanding their own limits and capabilities. Not to be oppressed by questionnaires or interviews that may make them anxious and uncomfortable. The author guesses that it is similar for teachers. There are various ways to use narrative reflexive qualitative research with both learners and teachers. The purpose is to obtain relevant, authentic data that can help teachers develop precise methodological approaches in the classroom. Autoethnography or journaling can be these ways.

Keywords: knowing, self-knowing, reflective narrative, journaling, autoethnography, the teaching-learning process

1. INTRODUCTION

The educational field in Slovakia undergoes rapid changes. There is a reform that may be challenging because it transforms not only the content but the system's entire structure. But the essential condition remains the same – to have a learner-centred teaching-learning process that leaves no one behind. So curricula, standards, methodological manuals, internal directives, and other guidance should be refocused on the aforementioned goal.

Being innovative does not mean that everything must be new. The term "fashionable" in terms of the methodological approaches seems very confusing. We would rather use the term "personalisation", "individualisation", or "differentiation" of the teaching-learning process. Robinson and Aronica (2029, p. 238) underline that *"the future of education lies not in standardization, but in customization; not in promoting groupthink, but in cultivating the true depth and dynamism of human capacity of all kinds."* Tomlinson (2017, p. 7) says: *"Differentiated instruction is not a single strategy or a recipe; it is a way of thinking about teaching and learning."*

Even in the new national curriculum reform (Ministry of Education, Research, Development, and Youth of the Slovak Republic, 2023), it is mentioned that personalisation and individualisation of education are listed as core pillars. It emphasises adapting instructional pace, teaching methods, and assessment strategies to the point that they meet the individual needs and developmental specifics of every learner.

In addition, as it is written in our previous article – see Burcl (2025, p. 229), to do so, "teachers must be prepared, motivated, and willing to get "closer" to their learners, and vice versa. Clear, bias-free, and respectful communication

* Corresponding author

is the only way to professional success and goal fulfilment. All that creates a bias- and anxiety-free atmosphere of mutual trust and positive creativity, enabling the complex cognitive and emotional development of learners."

Something is more important than the proclaimed control of output, timing, and results measuring. All these technocratic approaches bring pressure, stress, and anxiety. Unlike in the corporate sphere, in an educational context, this can be counterproductive. Deeper, more valuable aspects of education are vanishing under these trends. Those which are more personal and character-related are overlooked. That is why it is crucial to bring more light into the teacher-learner relationship. For this reason, *knowing* and *self-knowing* are absolutely essential steps toward qualitative change in this segment because they are the rudiments of any pedagogical relations.

As Ngan, Hertz and Thong (2021) note about the professional identity of a teacher, besides expertise and competences, there are other agents. These include personal self-awareness, value systems, and perception and experience. Hupková and Petlák (2004) emphasise that self-reflection in a teacher's practice enables personal professional growth while also safeguarding against teacher burnout. Most importantly, however, regular and informal self-reflection enhances the quality of the educational process and its outcomes. Moreover, traditional linear transmission of information from teacher to learner is overcome.

Therefore, for a valuable transformational pedagogical dialogue, both partners, teacher and learner, must enter the interaction as self-aware subjects who mutually reflect on their own positions and connections. According to Žydžiūnaitė and Daugėla (2020, p. 160), "*teacher's professional self-awareness is a developmental process which requires from the person reflection on his/her own experiences. ... Through working with students, teachers step into the "unknown" and learn within togetherness. Being in togetherness brings bilateral interchange between teachers and students, which motivates both sides to be self-aware. These reciprocal interactions invite participants to grow and seek mutual interchange through different experiences and contexts.*

2. IMPORTANCE OF KNOWING, SELF-KNOWING, AND SELF-RECOGNITION

The basic premise of this article is that the teacher cannot autonomously and authentically know, understand, and appropriately lead learners methodologically until he/she first goes through the process of deep reflective self-knowing. A teacher's self-knowing involves identifying own mental models, emotional triggers, and cognitive limits, as well as hidden prejudices that may shape or distort the class climate and interactions with learners. That is why learning self-reflection and self-awareness of own limits is crucial for personal development – of both teacher and learner as well, e.g. see Derla and Baguio (2024), who claim that there is a significant relationship between self-awareness learning and educational development.

Bradová (2013) shows that self-knowing is the first necessary step to power decentralising in the class, shifting away from strictly directive management in the classroom, and creating conditions for a real learner-centred teaching-learning process; The teacher takes the role of a facilitator and consciously regulates overbearing displays of dominance. We think that the process of self-knowing or self-recognition is beneficial to a collaborative approach, where the boundaries between teacher and learner blur, fostering mutual understanding and creating a safe educational space. Olin, Almqvist, and Hamza (2021) claim, based on their research, that *"through well-structured dialogues, the participants transformed their understanding and valuing of both themselves and others in relation to the task of producing new didactical knowledge. This is interpreted as transformed self-recognition for the teachers, who started to acknowledge themselves as knowledge producers. This transformation was crucial for developing the mutual recognition through which new didactical knowledge emerged as a result of the collaboration."*

Pascual (2026) states that teacher recognition significantly boosts teaching quality and professional motivation. Teachers who feel respected and supported show higher effectiveness in classroom management and lesson planning. According to Novocký (2017), professional reflection becomes the primary source of professional/personal development and educational practice improvement.

If a learner-centred approach is perceived as defended and privileged, we must critically select the instruments of assessment and evaluation that are used to measure and diagnose a learner's personality, abilities, and limits. Traditional diagnostic and research methods, such as rigid structured questionnaires, tests, or formal interviews, often fail in practice. They tend to generate a high level of anxiety, discomfort, and external pressure.

Learners under constant pressure lose natural intrinsic motivation and the ability to perceive their own limits and competences autonomously. According to Flavian (2016), to support their authentic self-recognition, there must be an educational place disposed of permanent external assessment and judgement. Mistakes must be seen not as a failure, but as a legitimate tool of knowledge. This stance is supported by the State School Inspectorate (2021), declaring in its methodological materials that self-evaluation within the framework of formative assessment minimises the fear of failure and considerably strengthens the autonomy of learners.

Pandey (2013) assumes that learners must also be given space to freely self-discover. Common quantitative testing reaches its limits because the approach should be personalised, reflective, and evocative, e.g. journaling. He (ibid., p. 75) sees it as a *"cross-cultural approach to transcend individual limitations among learners."* Hiemstra (2001, p. 20) defines journaling as *"a means for recording personal thoughts, daily experiences, and evolving insights."* All efforts should end up in self-directed learning when the learner is empowered to lead own learning trajectory, identifying needs and goals, taking effective strategies, and self-evaluating (self-reflecting).

Smith (2019) specifies: "Self-Directed Learning (SDL) is an instructional model where learners take initiative in their educational processes, allowing them to learn at their own pace and based on pre-determined content set by instructors. This approach emphasizes personal autonomy, self-management, and learner control, enabling individuals to diagnose their learning needs, set goals, and evaluate their outcomes. SDL is particularly effective in adult education, as adults often possess the motivation and responsibility to engage in autonomous learning."

3. PERCEIVING REFLECTION AND RETROSPECTIVE IN WRITING

As we said earlier, self-knowledge and self-recognition keep the quality of individual outputs. Teachers and learners are unique human beings, and therefore, finding and detecting deep personal details, emotionality, and lived experiences is like gold mining. You have to go into the depths of a soul, but still give the most delicate treatment. If research is to be productive, a traditional set of methods will not be useful. In this situation, a specific kind of qualitative research must be taken into consideration – ***reflective narratives***.

Alam, Asmawi and Fatema (2025, p. 57) note: "Narrative data or reflective narratives (RNs) represent a vital qualitative data source. This kind of data captures the subjective, self-reflexive, and contextually embedded experiences of individuals on any specific phenomenon. RNs have widespread use in the field of education, nursing, social work, and counselling, and almost all qualitative methodologies. Yet, RNs are under-explained, especially for novice researchers."

Bolton (2001, p. 13), in her book bestseller 'Reflective Practice: Writing and Professional Development', claims that using writing as a means of reflective practice allows us to *"express and explore our own and others' stories: crafting and shaping them to help us understand and develop."* She (ibid., p. 3) defines reflection: *"Reflection is a state of mind, an ongoing constituent of practice, not a technique, or curriculum element. Reflective Practice can enable practitioners to learn from experience about themselves, their work, and the way they relate to home and work, significant others, and wider society and culture."* The main idea of using reflective writing was originally psychotherapeutic, with the main focus on one's own values and subjective perception of inner and outer reality. Her own method of writing, called *writing-through-the-mirror*, is based on free writing of real or fictional stories, where the purpose is the writing itself, not producing a specific text. It is the authentic process of using the text to explain one's own acting and behaving. On the other hand, for some people, it can be frustrating to open up about sensitive issues via the process of reflection; they can feel uncertainty – a big obstacle in critical self-knowing.,

Another type of writing activity is ***journaling***. It can also serve therapeutic purposes. It means writing a diary, but in a specific, structured way. It focuses on answering self-posed questions. As Starfire (2023) notes: *"Reflective journaling is a therapeutic practice that can help you identify, grow from, and find meaning in your past experiences. And the new perspective you gain along the way can change the way you respond to people and events in the future."*

Some schools or universities require using journals for various reasons, e.g., Hong Kong University. According to their Assessment Resources, see Chan (2009), *"a reflective journal is a means of recording ideas, personal thoughts and experiences, as well as reflections and insights a student has in a course of learning process. ... The benefits of the reflective learning process are usually accumulated over a period of time, in which the students usually show a series of developmental changes, personal growth and changes in perspectives during the process."* They can do structured or unstructured/free-form journals. Here are the attributes of journaling (Chan, ibid., 2009):

- to discuss or argue a journal paper/report/an issue from a private standpoint or from various perspectives;
- to synthesise or analyse some materials or resources for building up an argument;
- to compare and contrast a particular issue with prior or new knowledge;
- to generate questions and think deeply as a result of perception of the resources;
- to describe personal experiences and integrate them into the issues in concern;

- to express freely for or against the specific questions given by teachers;
- to develop the ability and critical attitude to integrate learning into real-world experiences.

The third type of reflective writing is **autoethnography**. According to Ellis and Bochner (2000, p. 739), "autoethnography is an autobiographical genre of writing and research that displays multiple layers of consciousness, connecting the personal to the cultural. Back and forth autoethnographers gaze, first through an ethnographic wide-angle lens, focusing outward on social and cultural aspects of their personal experience; then, they look inward, exposing a vulnerable self that is moved by and may move through, refract, and resist cultural interpretations."

Except using autoethnography as a methodological approach with teachers and learners, Barr (2019) goes beyond the current perception, analysing the whole course, specifically his MA-level class on autoethnography in international relations. He (ibid.) comments: *"Students report that the course had the "potential to revolutionize" learning and teaching as a means of questioning themselves in relation to the social and political worlds they studied. Autoethnography has transformed the way we approach ourselves and our research; it is now time we let it change how we teach."*

The benefit of these reflective writing variants is that they apply to teachers and learners as well. Traditional ways of gathering information are mostly direct (questionnaires, interviews), which can be personally complicated, where respondents may feel frustrated or anxious. With reflective writing, even though they are given more comfort and freedom to choose what to bring in (it is not so rigid), they deliver qualitatively better outputs. It is due to conditions and atmosphere that they can go deeper. They can be analytical and evocative at the same time.

4. CONCLUSION

The quality level of the teaching-learning process stands on the relationship between teachers and learners. To preserve it, both must know. Only a teacher who knows his/her own personality is prepared to know learners – to be aware of values, emotions, assumptions, limits, potential biases, and their individual needs; moreover, to create a respectful, creative, supportive, and safe educational space. Self-knowledge and self-reflection should be perceived and accepted as essential features of teachers' professional and personal development.

Also, learners should be given sufficient time, space, and opportunities for self-recognition – to test their abilities, limits, emotions, and learning needs. Generally, a learner-centred approach should be applied throughout the school system, preferring individualisation and personalisation of education, reduction of anxiety and pressure, and giving autonomy, motivation, and support to learners. Mistakes can be confronted via reflection.

Reflective narratives, journaling, and autoethnography are those valuable qualitative approaches that can support these processes.

References

[1] ALAM, Saiful; Adelina ASMAWI & Sayeda FATEMA, (2025). Reflective Narratives as Qualitative Data: A Fundamental Understanding. *Integrated Journal for Research in Arts and Humanities*, 5(4), 57–63. https://doi.org/10.55544/ijrah.5.4.9.

[2] BARR, Michael, (2019). Autoethnography as Pedagogy: Writing the "I" in IR. *Qualitative Inquiry*, 25(9-10), 1106-1114. *https://doi.org/10.1177/1077800418792940*

[3] BOLTON, Gillie, (2001). *Reflective Practice: Writing and Professional Development*. London: Paul Chapman Publishing. 240pp. ISBN 0-7619-6729-X.

[4] BRADOVÁ, Gabriela, (2013). Interakčný model učiteľa etickej výchovy: Teoretické východiská problematiky (Časť 1) [The Interaction Model of an Ethical Education Teacher: Theoretical Foundations of the Issue (Part 1)]. Online. Banská Bystrica: UMB. Available from: *https://www.pdf.umb.sk/app/cmsFile.php?disposition=a&ID=6425*. [viewed 2026-10-12].

[5] BURCL, Pavol, (2025). Contemplation as a Complementary Force in the Teaching-Learning Process. In: *Proceedings of IAC 2025 in Prague: Praha, 14.12.2025-15.12.2025 / ed. Helena Kratochvílová, Radek Kratochvíl.* Praha: Czech Institute of Academic Education. pp. 229–234. ISBN 978-80-88203-45-2.

[6] CHAN, Cecilia K. Y., (2009). *Assessment: Reflective Journal*. Online. Available from: *https://ar.hku.hk/pdf/ReflectiveJournal.pdf*. [viewed 2026-10-12].

[7] DERLA, Honey M.C. & Josephine B. BAGUIO, (2024). Self-Awareness Learning in Relation to Educational Development of Students in Public Elementary Schools. *Asian Journal of Education and Social Studies*, 50(11), 146-154. *https://doi.org/10.9734/ajess/2024/v50i111642*.

[8] ELLIS, Carolyn & Arthur P. BOCHNER, (2000). *Autoethnography, personal narrative, reflexivity: Researcher as subject.* In N. K. Denzin & Y. S. Lincoln (Eds.), The handbook of qualitative research (2nd ed., pp. 733–768). Sage.

[9] FLAVIAN, Heidi, (2016). Towards teaching and beyond: Strengthening education by understanding students' self-awareness development. *Power and Education*, 8(1), 88–100. *https://doi.org/10.1177/1757743815624118*

[10] HIEMSTRA, Roger, (2001). Uses and benefits of journal writing. In Leona M. English & Marie A. Gillen (Eds.), Promoting Journal Writing in Adult Education: New Directions for Adult and Continuing Education, 90, pp. 19–26. San Francisco: Jossey-Bass.

[11] HUPKOVÁ, Marianna & Erich PETLÁK, (2004). Sebareflexia a kompetencie v práci učiteľa [Self-reflection and Competencies in the Teacher's Work]. Bratislava: IRIS. 135p. ISBN 80-89018-77-7.

[12] MINISTRY of EDUCATION, RESEARCH, DEVELOPMENT, and YOUTH of the SLOVAK REPUBLIC, (2023). *Štátny vzdelávací program pre základné vzdelávanie* [State Educational Programme for Primary Education]. Bratislava: National Institute for Education and Youth (NIVaM).

[13] NGAN, Duong T. N.; Maria HERCZ & Nguyen V. THONG, (2021). Teachers' Professional Knowledge and Their Professional Identity: If they Relate or not? *International Journal of Childhood Education*, *2*(4), 17–31. *https://doi.org/10.33422/ijce.v2i4.93*

[14] NOVOCKÝ, Michal, (2017). Reflexívna výučba ako prostriedok rozvoja učiteľov. Pedagogické rozhľady, 5, 20-21. ISSN 1335-0404.

[15] OLIN, Anette; Jonas ALMQVIST & Karim HAMZA, (2021). To recognize oneself and others in teacher-researcher collaboration. *Educational Action Research*, 31(2), 248–264. *https://doi.org/10.1080/09650792.2021.1897949*

[16] PANDEY, Kashi R., (2013). Journaling: A Cross-cultural Approach to Transcend Individual Limitations Among Learners. Journal of Education and Research, 3(1), 75–91. *http://dx.doi.org/10.3126/jer.v3i0.7854*

[17] PASCUAL, Anne G. D., (2026). Teacher Recognition and Instructional Effectiveness in Public Schools: Implications for Teacher Professional Development. *Contemporary Research Analysis Journal*, 3(6), 480–488. *https://doi.org/10.55677/CRAJ/11-2026-Vol03I06*

[18] ROBINSON, Ken & Lou ARONICA, (2009). The Element: How Finding Your Passion Changes Everything. New York: Viking/Penguin Group. p. 238. ISBN 978-0-670-02047-8.

[19] SMITH, Tricia, (2019). Self-Directed Learning. Online. Available from: *https://www.ebsco.com/research-starters/education/self-directed-learning/#full-article*. [viewed 2026-10-12].

[20] STARFIRE, Amber L., (2023). *What Is Reflective Journaling?* Online. Available from: *https://www.dummies.com/article/academics-the-arts/language-language-arts/writing/what-is-reflective-journaling-296242/*. [viewed 2026-10-12].

[21] STATE SCHOOL INSPECTORATE (SSI), (2021). *Hodnotenie a sebahodnotenie vlastných výkonov žiakov vo vyučovaní na 1. stupni ZŠ* [Assessment and Self-assessment of Pupils' Performance in Lessons at the Primary Level of Elementary School]. Bratislava: ŠŠI. 26p.

[22] TOMLINSON, Carol Ann, (2017). How to Differentiate Instruction in Academically Diverse Classrooms. 3rd ed. Alexandria, Virginia: ASCD. p. 7. ISBN 978-1416623304.

[23] ŽYDŽIŪNAITĖ, Vilma & Marius DAUGĖLA, (2020). Teacher's Professional Self-Awareness Within The Interactions With Students In Higher Education: Temporality and Relationality. In *Acta Pedagogica Vilnensia*, 45. Vilnius: Vilniaus Universitetas. pp. 160–170. https://doi.org/10.15388/ActPaed.45.10

CIAE
Czech Institute of Academic Education z.s.

Title:

"Proceedings of VIAC in October 2026", ISBN 978-80-88203-48-3

October 2026 in Prague, *1st edition*

Publisher / Creator / Copyright holder:

Czech Institute of Academic Education z.s.

Address of Publisher:

Vodnicka 309/20, 149 00 - Prague 4, Czech Republic

Email: info@conferences-scientific.cz

Web: www.conferences-scientific.cz

ISBN 978-80-88203-48-3

Czech Institute of Academic Education z.s.