Proceedings of IAC 2025 in Prague

April 25 - 26, 2025

Teaching, Learning and E-learning (IAC-TLEl 2025)
Management, Economics and Marketing (IAC-MEM 2025)
Transport, Logistics and Tourism (IAC-TLT 2025)

ISBN 978-80-88203-42-1

CTU
CZECH TECHNICAL UNIVERSITY IN PRAGUE

CIAE
Czech Institute of Academic Education z.s.

Impressum

Title:

"Proceedings of IAC 2025 in Prague", ISBN 978-80-88203-42-1

April 2025 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 IAC 2025 in Prague", ISBN 978-80-88203-42-1

Duben 2025 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*

Assist. 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*

Table of Contents

Management, Economics and Marketing

Teaching, Learning and E-learning

Transport, Logistics and Tourism

Management, Economics and Marketing (IAC-MEM)

Research Approaches of Business Ethics of Different Age Groups

Temur SHENGELIA [a*], Khatuna BERISHVILI [b]

[1] *Doctor of Economic Sciences, Professor, Ivane Javakhishvili Tbilisi State University, Faculty of Economics and Business, Tbilisi, Georgia, temur.shengelia@tsu.ge*

2 Doctor of Economics, Associate Professor, Ivane Javakhishvili Tbilisi State University, Faculty of Economics and Business, Tbilisi, Georgia, khatuna.berishvili@tsu.ge

Abstract

The study and identification of different approaches to ethical values among different generations of employees in the Georgian business sector is a topical scientific problem. Due to its novelty, this problem has been less analyzed and studied by scientists. The relevance of the problem is determined by the fact that its study makes it possible to identify ethical values among different age groups of people employed in business and manage conflicts on this basis. In Georgian companies, value incompatibility of different age groups is manifested, the contradictions arising on this basis can become the basis for the emergence of conflict situations. One of the reasons for the emergence of labor conflicts is different approaches to business ethics values among different age groups. Such differences are manifested in different perceptions and interpretations of business ethical norms according to national, religious, and sexual characteristics. The process of socialization of the younger generation took place in the conditions of post-Soviet Georgia. Based on the above, they differ significantly from the older generation in terms of their value system, which is why different generations have developed different approaches to business ethical norms. Using various methodologies for researching cultural values, the paper identifies intergenerational differences in the priorities of business ethical values. Based on the analysis of different approaches to business ethics in different age groups, a value system was identified that influences the formation of stereotypes in their cognition. The research results showed that intergenerational differences in business ethical values determine their approaches to both labor in general and business processes in general.

Keywords: ethical values, cultural values, intergenerational differences, identification, research methodology.

1. INTRODUCTION

The modern business world is dynamic and changing. In this dynamic world, values, norms, and attitudes are changing. The labor market is undergoing transformation, it is being rejuvenated, society is making new demands on relativistic approaches to business ethics, and they are becoming universal. It is also clear that in Georgian companies, there is a value incompatibility between different age groups, and the contradictions arising on this basis lead to conflict situations in labor collectives. Inother words, one of the reasons for the conflict is different approaches to the values of business ethics. Such differences are manifested in different perceptions and interpretations of business ethical norms at the national, racial, religious, and gender levels. Young people were socialized in the conditions of the new Georgia and have a different value system than the previous generation, which gives rise to different approaches to business ethical norms in different generations.

* Corresponding author.

2. LITERATURE REVIEW

Over the past two decades, business ethics, as a field of scientific and practical knowledge, has attracted great interest from scholars. The famous British scientist R. Lewis, in his book "Business Cultures in International Business: From Collision to Understanding", writes that "business ethics is a set of standards, rules and principles that are the basis of honesty and moral behavior in business relationships" [Richard, D. Lewis. 2018]. L. Nash believes that business ethics is not a simple, specific set of moral standards, but a tool for solving and analyzing the problems that people face in the process of doing business [Sison, A., Beabout, R., Ferrero, I. 2017]. In the scientific literature, business ethics is also defined as a discipline that studies "the interaction of corporate business activities, their moral adequacy, and the specific application of moral standards in business, politics, and institutions" [Goodpaster, K. 1997]. It "includes formal and identifiable activities that are carried out between individuals, organizations, or other economic entities related to business and involve its relationship to ethical (i.e. moral) norms" [Norman, W. 2013].

It is also clear that in economic and social sciences there are many definitions of culture, which characterize various aspects of this very complex, multidimensional concept. In this regard, the definition of the American scientist D. Matsumoto is not of interest, according to which: "Culture is a dynamic system of explicit and implicit rules established by individual groups for their survival and it includes the common values, norms and models of behavior of these groups, the realization of which is carried out among its members individually, is transmitted from generation to generation and changes over time" [Matsumoto, D. 2001. P.40]. D. Matsumoto's definition is interesting for us insofar as it most of all concerns ideas, attitudes and values. Culture is realized through values, which in turn are its basis. Values are conscious or unconscious perceptions of what is desirable, characteristic of an individual or group of individuals, which determine the possible means and methods of realizing individual or group goals.

The scientific literature is widely presented with methodological approaches to measuring values in different cultures. In this regard, the works of G. Hopstede, Strodbeck, E. Hall, S. Schwartz, F. Trompernaars, S. Hampden-Turner and other scientists are interesting. In our study, we use the modernized methodology for measuring values proposed by S. Schwartz. The emphasis in the study is only on individual-level values, since the subject of analysis is an individual, not a social group. S. Schwartz divided individual values into the following parts: power, achievement, hedonism, stimulation, autonomy, universalism, care, tradition, conformity, security. In later studies, he created a theory of dynamic relations, according to which all values are represented on several bipolar value-motivational axes: openness to change (values of independence, stimulation and hedonism); preservation (security, conformity and tradition); self-assertion (power, achievement, hedonism) and self-determination (universalism and benevolence) [Schwartz, M., Weber, J. 2006].

Today, there are practically no scientific studies that substantiate the existence of intergenerational differences in ethical approaches to business, however, there is a fairly extensive body of research that examines intergenerational approaches to values [Velasquez, M. 2016]. In this context, the research we present is not only important, but also innovative and relevant.

3. RESEARCH METHODOLOGY

The aim of the research is to study the approaches to business ethical values among different generations of Georgian business representatives and to identify the contradictions that arise on this basis.

Several hypotheses were defined during the research process:

H1- Approaches to ethical values are different among different age groups employed in the business sector.
H2- There are differences between different generations in ideas about the ethics of business relations.

114 respondents were selected in the research process. Young people (19-30 years old) and older generations (30-70 years old) employed in the business sector of Georgia participated in it. Gender balance was maintained in the process of collecting and processing data. The gender and age structure of the participants in the research is presented in the diagram (see Fig 1).

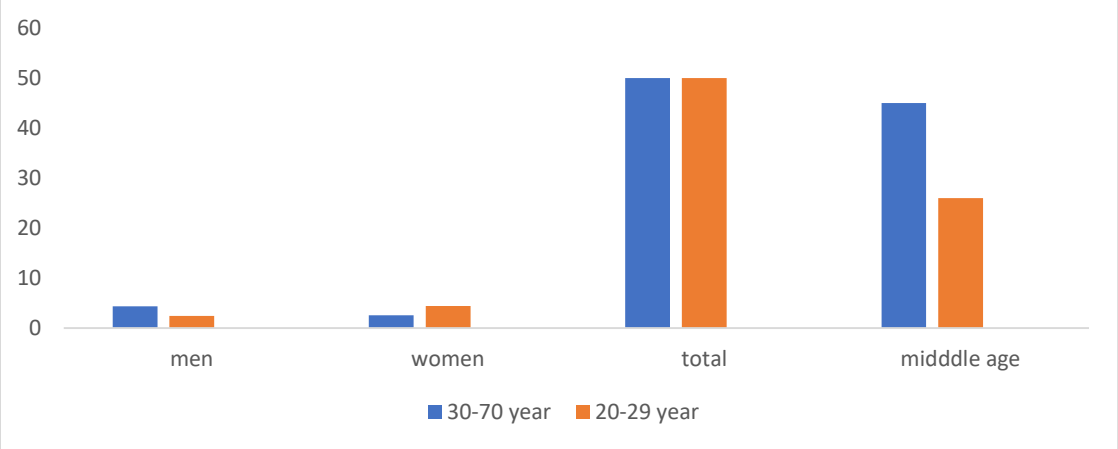

Fig. 1 Gender and age structure of the study participants

(According to S. Schwartz). Statistical data processing was carried out using the SPSS Statistics software package. To determine the correspondence of cultural values and established practices of business ethics and to identify intergenerational differences, statistical research methods were used: mean values were compared using the t-Student test. The multiple regression analysis method was used in the research process. To determine significant differences between indicators in groups, the Kolmogorov-Smirnov Z-test was used.

4. RESEARCH RESULTS

In the first stage, the differences in individual values of employees of different generations in the business field were determined, which was assessed according to ten indicators. The t-Student test was used in the research process (see Table 1).

Table 1. Intergenerational differences in individual values (according to Student-t-test)

year	30-70 year			20-29 year		
Indicators	Min- Max. value.	Average values	Standard Deviation	Min- Max. value.	Average values	Standard Deviation
1. Conformism	3,3-5,4	4.72	0.54	2.1-4	3.74	0,63
2. tradition	1.6-4.2	3.11	0,70	0,2-5,1	2,36	0,72
3. Universalism	2-5,2	3,52	0,51	2-3	3,32	0,47
4. Independence	2,4-6	3.72	0.50	2,0-3,5	4,31	0,46
5. Stimulation	0,5-4,2	3,05	0,74	2-6	3,87	0,82
6. Gedism	1,3-5-3	3,62	0,82	2-6,1	4,86	0,81
7. Achievement	2,8-4	3,89	0,54	3-5,5	4,30	0,52
8. Attitude	1,8-5,3	3,25	0,68	2,9-6,3	3,81	0,71
9. Openness to change	2,3-4,5	3,4	0,53	0,4-5,3	4,41	0,52
10. Self-establishment	2,1-4,5	3,5	0,47	3,0-5,6	4,38	0,51

As can be seen from these data, there are significant differences in such key indicators as: conformism, tradition, hedonism, achievement, openness to change, self-assertion, independence. There are similarities in the parameters of: universalism, stimulation, dependence, universalism. In fact, according to seven out of ten indicators, there is a significant difference in intergenerational approaches. Closeness is observed only in terms of three indicators, which means that the approaches to business ethical values are different in different age groups employed in the business sector, thus confirming the H1 hypothesis.

To deepen the validity of the obtained data, the weight and variance of business ethics factors were analyzed at the next stage of the study (see: Fig 2).

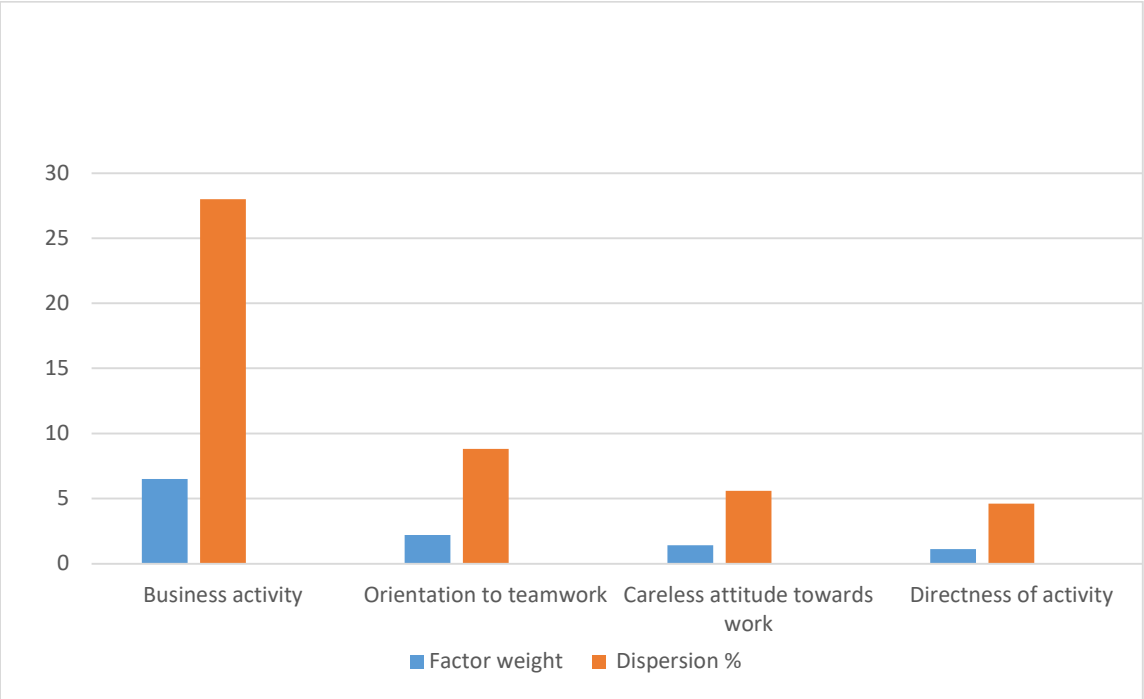

Fig. 2 Factor weight and dispersion of individual values of value differences

Cronbach's α indicators for all identified factors have high values (see: Fig 3), therefore, the variables included in the factors correspond to each other and the scales are valid.

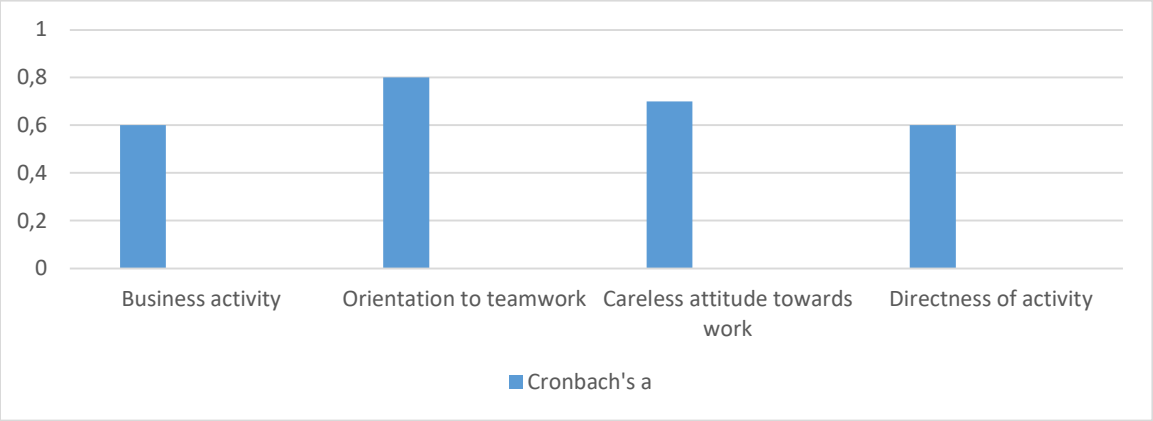

Fig. 3 Cronbach α by identified factors

Intergenerational differences in business ethics according to the Kolmogorov-Smirnov criterion are presented in the following table (see: Table 2)

Table 2. Intergenerational differences in business ethics according to the Kolmogorov-Smirnov criterion

	30-70 years		20-29 years	
Factors	Average values	Standard deviation	Average values	Standard deviation
Business activity	2.7	0.3	3.5	0.3
Orientation to teamwork	3.4	0.3	2.9	0.4
Careless attitude towards work	3.5	03	3.6	0.5

Directness in business	2.0	0.5	2.7	0.4

These data show that the younger generation in business prioritizes values such as: business activity, and directness in business, while the older generation values teamwork.

Correlation analysis was used to determine the differences in the relationship between values and ideas about business ethics between different generations. The results of the correlation analysis are presented in the diagram. (See: Fig 4).

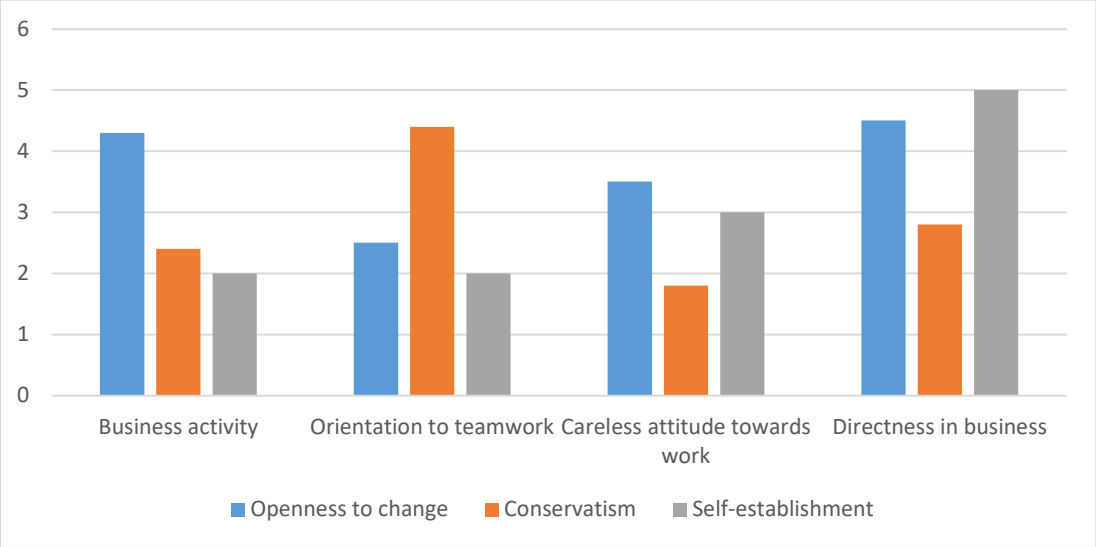

Fig. 4 Results of correlation analysis of differences in the relationship between values and ideas about business ethics among different generations

Correlation analysis revealed that in the older generation, business activity and teamwork orientation are associated with openness to change, while in the younger generation, directness in business is associated with openness to change and self-assertion, thus confirming hypothesis H2.

5. CONCLUSIONS

As a result of the conducted research, the following conclusions can be made: among business entities employed in the Georgian business sphere, there are intergenerational differences in value priorities at the individual level. Young people give preference to: independence, Gedoisism, achievement, openness to change, self-establishment. Representatives of the older generation prefer conformism, traditionalism, which is logical, since the younger generation was born in post-communist Georgia, which gave rise to their approaches to business, which are oriented towards progress and development.

The hypothesis about intergenerational differences in business ethics was confirmed. Intergenerational differences in business ethics according to the Kolmogorov-Smirnov criterion indicate that there are significant differences between the younger and older generation respondents in the factors: “business activity”, “teamwork orientation”, “relationship orientation” and “directness in business”.

Correlation analysis revealed that in the older generation, business activity and teamwork orientation are related to openness to change, while in young people, directness in business is related to openness to change and self-assertion.

Thus, the conducted research confirmed the essential differences that dominate the value systems of the older generation and the younger groups, which often give rise to different approaches to the labor process between them. The study and analysis of this value system is essential for the process of managing intergenerational conflicts among employees in the business sphere.

References

[1] Brenkert, G., Beauchamp, T. (2018). The Oxford Handbook of Business Ethics. OUP USA, 2010. https://books.google.ge/books/about/The_Oxford_Handbook_of_Business_Ethics.html?id=Qn1BDj7SRUsC&redir_esc=y

[2] Kluckhohn, C., Strodtbeck, F. (2006). Variations in value orientations. Evanston, IL.
[3] Matsumoto,D.(2001)Psychology and culture. Oxford University Press. https://books.google.ge/books?hl=en&lr=&id=xNeM09f87cwC&oi=fnd&pg=PR3&dq=Matsumoto+D.+Psychology+and+culture&ots=fPbMd-oeJn&sig=pCrthZSRpxs
[4] Norman,W.(2013)BusinessEthics.https://www.hbs.edu/faculty/Shared%20Documents/conferences/2016-newe/Norman%2C%20Business%20Ethics%2C%20IntEncycEthics.pdf
[5] Richard, D. Lewis. (2018) When Cultures Collide: Leading Across Cultures. 4th Edition. https://www.amazon.com/When-Cultures-Collide-Leading-Across/dp/147368482X/ref=sr
[6] Shengelia, T. (2018). Influence of Cultural Determinants on the Process of Business Innovations Management. Ecoforum Journal. Volume 7, Issue 1
[7] Shengelia, T. (2017). Interpersonal and Organizational Trust, as a Factor of Social Capital, and its Influence on the Motivation of International Company Employees. Ecoforum Journal. Volume 6, Issue 2
[8] Shengelia, T., Berishvili Kh. (2020). Methodology of social entrepreneurship research and its influence on countries with small economy. International Journal of Social and Humanities Sciences. Volume, Issue 3
[9] Shengelia, T., Berishvili Kh (2023). The European experience of corporate social responsibility and the reality of Georgia. Academics and Science Reviews Materials. Issue 5
[10] Sison, A., Beabout, R., Ferrero, I. (2017). Handbook of Virtue Ethics in Business and Management.springerhttps://link.springer.com/referencework/10.1007/978-94-007-6510-8
[11] Schwartz, M., Weber, J. (2006) A business ethics national index (BENI): Measuring business ethics activity around the world // Business a. society. L., V. 45. No. 3.
[12] Velasquez, M. (2016) Business ethics: Concepts and cases. Upper Saddle River (N.J.)

Financial Development and Economic Growth Nexus in a Developing Economy: Nigeria Evidence

Chioma MOGBO [a*]

[a] *Texila American University, School of Business and Management, Plantation Providenc, East Bank Demerara, Guyana, chiwils@yahoo.co.uk*

Abstract

Motivated by the need to evaluate the extent to which the Nigerian economy interdepends with financial development indicators as outlined by World Bank (2018;2022) standards. Secondary data originating from the financial development indicators database of World Bank from 1994 to 2023 and the Statistical Bulletin of Central Bank of Nigeria were utilized. The results of Auto Regressive Distributive Lag Bound Test applied indicated lending disbursed to the private sector (financial institutions depth proxy) and the number of operating bank accounts per 100,000 adult populations (financial access proxy) were significant in predicting Nigeria's gross domestic product (economic growth proxy). However, the Lending-deposit spread (financial institutions' efficiency proxy) and asset quality ratio (financial institutions' stability proxy) failed the significance test. The Granger causality test results revealed no significant causalities between Nigeria's gross domestic product and the explanatory variables. Further, it was concluded that Nigeria's economy and the four explanatory variables could not promote each other as they all operate independently. Consequently, the study recommended, among other things, that CBN should raise the statutory maximum loan/deposit ratios for operating banks to enable them lend more to customers as its multiplier effects will grow the economy faster.

Keywords: Financial Development, Economic Growth, Financial depth, Financial Efficiency, Lending-deposit Spread.

1. INTRODUCTION

Financial intermediation theory asserts that the financial sector theoretically plays a weighty part in instigating economic performance through quality fiscal operations improvement. Status implies the extent to which financial development exerts a vital part in economic growth. Levine's (1997; 2005) observation hinged on the notion that financial development helps to simplify information about possible investments and monitoring, as well as risk diversification and management. Further strengthening pooling of money saved in the financial system and economic transactions. Hence, the progress outcome of financial expansion is centered on saving mobilization and efficient resource allocation.

Over the years, scholarly and policy interest centred on the type of predicting interrelationships among financial development and economic growth of nations notably emergent nations. World Bank (2018) specifically provided two broad classes of financial development indicators, which were derived from financial organizations and money markets. While financial institution's correlates include financial institutions' depth, access, efficiency, and stability each with its proxies, the financial market's correlates derive from proxies some of which include stock market capitalization, ratio of public debt securities to gross domestic product, and volatility of stock prices. Further, Nigeria's

* Corresponding author.

economy underwent significant deregulation (privatization) since the year 1986, as well as consolidation of the finance sector in 2005 aimed at introducing significant market discipline. These developments have largely altered the dynamics of Nigeria's financial institutions as well as the economy. With deregulation of Nigerian economy in 1986 and consolidation of financial industry in 2005, there is an urgent need to evaluate true nature of interrelationships within Nigeria's financial development and economic growth utilizing latest information. The focus of this research is to address the above.

2. LITERATURE REVIEW

2.1. Concept of Financial Development

According to Bilir, Chor, and Manova (2019), financial development is process of enhancing availability of financial services, including so many options and touches every part of the economy. Financial development relates to progress and improvement of economic systems and organizations that enhance efficiency and effectiveness of finance markets. It covers several spheres, including the advancement of financial intermediaries, enlargement of financial markets, along with complexity of financial products. Usually, when nations grow, their financial systems change to enable better distribution of money, more access to financial resources, and more financial stability.

The notion behind financial development is based on the knowledge that properly operating financial systems have a major influence on economic growth. Savings mobilization, resource allocation, risk management, and provision of necessary services to individuals and organizations depend critically on financial institutions including banks, insurance companies and capital markets (Levine, 1997). A robust financial system enables efficient intermediation between savers and borrowers, ultimately contributing to higher levels of investment and economic output.

2.2. Dimensions of Financial Development

Careful evaluation of financial development is significant in evaluation of progress made in financial services to realize the importance of financial development on economic expansion. Practically, financial advancement remains an idea with special multiplicity. Largely, empirical studies so far are based on standard quantitative indicators, which data prevail in relevant countries (Rewilak, 2017). Still, these policies provide approximations and do not fully reflect all facets of financial development since the investment sector of nation includes several financial institutions, markets, and products.

Financial depth—particularly, stock of private sector bank credit and market capitalization as a percentage of gross domestic product—measures financial development most of the time. This kind of indicator emphasizes the quantity side of financial growth. Still, financial development is a complex idea (Fashina, Asaleye, Ogunjobi, & Lawal, 2018). In this sense, Ibrahim and Alagidede (2018) contended that rather than concentrating just on improvement of banking system, there is also a need to consider several kinds of financial markets including equities, bond, and insurance markets to properly depict delicate complexity of financial development. Moreover, we should evaluate the quality features of financial development, including cost performance, industry breadth, and market efficiency, in addition to only noting the size and activity of the financial sector. Besides, claims by Hasan, Koetter, and Wedow (2009) is quality of financial intermediation, not quantity. Research stressed that, regarding profit and cost effectiveness, quality of financial intermediation should take front stage. To evaluate the quality of financial development, this study urged for a larger perspective on the multi-dimensional features of it. Given this posture, the investigation took a wider perspective of financial development with the dimensions of market diversity or breadth; market liquidity; market efficiency as well as institutional environment, thereby reflecting a composite index of legal and institutional development. The Global Financial Development Database (2018) of World Bank thoroughly created simple concept in capturing financial development world over. This background notes four sets of proxies representing functional financial systems. They include depth, access, efficiency, and stability

2.3. Concept of Economic Growth

Economic growth - procedure whereby economic welfare, standard of living of people are improved according to targeted goals and objectives. The growth in an economy is generally appraised relative to performance of financial wellness goals. This can be long term, like viable growth and development, or short term, for instance economic stability with respect to unexpected occurrences, termed financial crisis. To figure out to what extent an economy is doing versus these targets, business analysts utilized variety of finance measures. Economic indicators measure broad

economic metrics that explicitly or implicitly help business analysts determine if state of economy has gotten much better or retrograded. Monitoring these criterions are specifically helpful for policy makers, in terms of both evaluating when to get involved and if intervention has worked or not (Ghirmay, 2004).

2.4. Theoretical Framework

2.4.1. Theory of financial Intermediation

Financial intermediation is amongst the most vital responsibilities of the financial management in any economy. It is regarded as the procedure of efficient mobilization of economic resources by formal financial organizations from surplus savings units in the economy at a price called interest rate and allocation of the same resources to efficient deficit spending units for the funding of optimally appraised projects/investments in the economy which will ensure the return of the disbursed funds to the savings sources.

2.4.2. Financial Development Theory

Theoretical models tend to suggest that advancement of the financial service sector is essential for economic growth. However, lack of consensus on the direction of causality, as economists do not agree about responsibility of the financial industry in economic expansion as evidenced in the studies of Mckinnon (1973), Shaw (1973), Fry (1988), Pagano and Jappelli (1993) and Levine (2004) Robinson (1952), Greenwood and Jovanovic (1990) and Arestis and Demetriades (1991). Alternatively, Levine (2004) built wide range of theoretical and empirical studies. The study argued that advanced financing systems alleviate the external financing constraints which businesses encounter. Some of the above studies favor deregulation of the financial sector. They affirmed that financial entities do act as mediators and influence the grades of savings and allocation of investment funds positively. Hence, they promote economic expansion. This finding premised on competition. Intense competition among financial entities theoretically results in a well-organized intermediation enabling high investment activities caused by lowering of interest rate premium. This sort of approach will improve loaning to high yielding sectors and ensure optimal allocation of funds by distributing funds to mostly economic investments. Total effect on economic development and welfare would tend to be positive.

2.4.3. Demand Following and Supply Leading Theory of Financial Development and Sectoral Performance

Analyzing the function of finance from many theoretical angles is essential to understand how financial development supports national progress. Two different kinds of financial development were distinguished by Patrick (1966) as "demand following" and "supply leading". Basically, the preceding describes phenomena wherein demand for financial services results from increase in real economic activities; the latter describes the situation whereby development of the financial system promotes actual economic expansion. Several theoretical and empirical theories abound to explain the connection among economic expansion and financial development. This results from different points of view among economists about the function of finance and growth with regards to stimulating the other. Aside from the forms of financial growth above, the literature has revealed numerous theories describing the function of financial advancement in the economy.

2.5. Empirical Review

Emphasizing Nigeria's service, industrial, and agricultural sectors, Adeyemi and Yusuf (2023) looked at how sectoral growth related to financial development. With a fixed-effects model and a panel data approach, the writers assess the years 1995–2021. Variables in financial development include total banking assets, stock market turnover, and domestic private sector loans. Research finds that financial development significantly affects expansion of the agricultural and services industries but has a less pronounced effect on manufacturing growth. The authors attribute the fact that manufacturing sector encounters significant structural problems, such as poor infrastructure and high production costs, which limit the potential benefits of financial development. They suggest that targeted financial policies focusing on credit availability and interest rate management could help unlock the manufacturing sector's growth potential.

Lawal and Ibrahim (2023) looks at relationships among Nigeria's inflation, economic progress, and financial development. Utilizing a cointegration approach and Granger causality tests on data sourced from 1986 - 2022, this work focuses on inflationary pressures, financial development, and GDP growth. Important variables researchers

consider are inflation rate, private business credit, and general money supply. According to their studies, financial development and inflation are inversely linked; that is, inflation reduces financial development, so slowing down economic growth. The paper emphasizes inflation control as essential to ensure that financial development supports long-term expansion of Nigeria's economy. It is advised that conservative fiscal and monetary policies help to moderate inflation thereby promoting financial innovation and economic development.

Amadi and Edeh (2021) investigates relationship among financial development and economic growth utilizing macroeconomic methods based upon data sourced from 1980 to 2020. The study performs a Granger causality test employing a Vector Autoregression (VAR) model to ascertain whether financial development prompts economic expansion, or the contrary is true. Indices of financial development are domestic credit to the private sector, broad money supply, and banking industry efficiency ratios. Results revealed direct link among financial development and economic progress, meaning the former causes latter to happen rather than the other way round. The writers argue that Nigeria's somewhat underdeveloped financial sector must be expanded by strengthened policy measures emphasizing on deepening financial markets and promoting innovation in financial services.

Balogun and Nwankwo (2023) study, the writers examine how institutional quality influences financial development-growth nexus in Nigeria. Covering the years 1986 to 2022, study utilizes an autoregressive distributed lag (ARDL) model. Along with indicators of financial development like private sector lending and stock market capitalization, the main factors include measures of institutional quality including corruption perceptions indexes, regulatory quality, and rule of law. Findings showed effectiveness of financial development in advancing economic growth is much influenced by institutional quality. Weak institutional structures including inadequate control of regulations and high degrees of corruption reduce the favorable effect of financial development on expansion. The writers come to conclusion that unlocking the maximum potential with Nigeria's financial industry in promoting economic development depends on changes in institutional quality and government.

Agbo and Okeke's (2022) main research article investigate relationship between financial development and economic progress. Examining how external debt affects financial development and economic growth in Nigeria, immediately and over long term, research employs a VECM to evaluate data from 1985–2021. One should consider general level of outside debt, debt service payments, home loans to private businesses, and GDP increase. Two elements that help to explain the negative consequences of external debt servicing on financial development and economic growth over long run are high debt payment obligations limiting access to credit for private sector and discouragement of domestic investments. Conversely, making money more accessible by means of outside debt can provide a temporary increase in financial growth. The writers believe that if Nigeria wishes to continue expanding over time, it should be more cautious with its foreign debt and give policies supporting domestic investment top priority. Furthermore, underlined are the possibilities of better debt restructuring and financial management to mitigate the negative outcomes of outside debt on economic progress.

Examining the expansion nexus in connection to financial development, Ojo and Adeyemi (2023) sectoral analyze how financial development affects Nigeria's agricultural, industrial, and service sectors. Using a panel data methodology spanning the period from 1995 to 2022, the study employs a generalized least squares (GLS) method to assess different effects of financial development across various industries. Variables in financial development are privatization of credit, market capitalization, and banking sector efficiency. The results show that, among the service and agriculture sectors, manufacturing is less impacted by financial development than others. Two structural problems in the manufacturing sector, the writers claim limit the benefits of financial growth are high production costs and poor infrastructure. They advise interest rate control and sector-specific lending programs to maximize the favorable effect of financial development on improvement in Nigeria's manufacturing and agricultural sectors.

3. MATERIALS AND METHODS

This study employed the time series longitudinal research design since series of data points over an interval of time were collected. The subject under consideration is purely a secondary data topic. The related data were collected from established financial institutions including World Bank and Nigerian Apex Statistical Bulletin that are responsible for publication of authentic reliable statistics analysis on the study variables.

3.1. Operational Measures of Variables

The key dependent variable is Nigeria's economic growth proxied with gross domestic product. Financial development proxies employed derived from standard proxies established by Word Bank (2018) for financial debt, access, stability and efficiency, which indictors service as operational variables of measure.

- **Gross Domestic Product:** Consist of market value finished goods and services produced in Nigeria annually. This is measured in billions of Naira.
- **Financial Depth:** This is ratio of credit disbursed to the private sector in Nigeria to gross domestic product. It's measured in percentage.
- **Financial Access:** This is measured as the ratio of number of accounts in all the operating bank branches as stated by Apex Bank to 100, 000 adults in Nigeria. It is measured in percentage.
- **Financial Efficiency:** This is captured as the difference between the lending rate and the average deposit rate. It is measured in percentage/or ratio per annum.
- **Financial Stability:** This is ratio of performing bank loans and advances to all loans and advances granted by banks in Nigeria. It is measured in percentage.

3.2. Model Specification

The main objective of this research is to evaluate the contribution of financial development on economic growth in Nigeria. Investigation therefore will employ following model.

$$GDPt = f(PSC, BAPA, LDS, AQR) \quad (1)$$

Where:
GDP = Gross domestic product
PSC = Credits to the private sector as ratio of gross domestic product.
BAPA = Ratio of number of bank accounts per 100,000 adults
LDS = Lending-deposit spread
AQR = Asset quality ratio

Because GDP is in level form, while PSC, BAPA, LDS, AQR are in ratios/percentage, the growth rate of gross domestic product was employed to compare likes. As such, equation 1 such, is re-written as follows.

$$GDPR = f(PSC, BAPA, LDS, AQR) \quad (2)$$

Where:
GDPR = Rate of GDP growth (in percentage) and other variables remains as earlier defined.
For estimation purposes, equation 2 is re-written as.

$$GDPR = \beta_0 + \beta_1 PSC + \beta_2 BAPA + \beta_3 LDS + \beta_4 AQR + ei \quad (3)$$

Where:
β_0 = Constant Parameters
β_1, β_2, β_3, β_4, = Estimation parameters
μ_1 =Error term

3.3. Method of Data Analysis

3.3.1. Graphical Presentations and Descriptive Statistics

This includes the presentation of employed figures in summarized tables and charts. Descriptive statistics are essential as the study just gave us the primary data, it will be difficult to visualize what the data were showing, especially if there are lots of it. Descriptive statistics, therefore, enable us to present the data in a more meaningful way, which allows simpler interpretation of data.

3.3.2. Stationarity (Unit Root) Test

Considering the specific purpose of this study, unit root tests were performed on all study time-series variables to find out the magnitude or contrary to, the data associated with unit root properties so to avoid spurious estimates

3.3.3. Autoregressive Distributive Lag (ARDL)

ARDL a statistical method utilized to investigate long-run connection between two or more factors in a time series context. Usually used when variables might be integrated of different orders, or where time scope is less than 30 periods. The ARDL test authorize approximation of parameters in a dynamic model that includes both short-run and long-run effects. It is particularly useful when analyzing cointegration relationships, where variables are linked in the long run but may exhibit short-term dynamics. The general steps involved in conducting an ARDL test are as follows:

Determining the appropriate lag length: Lag length determination is necessary for error correction and ARDL estimation. This is because investments in some material variables in some periods in the past may be very significant in explaining current output as observed in agriculture, banking and other business sectors in the economy. As such in this study, branches established in the past may significantly contribute to the deposits, lending and investment outbursts in the current year.

4. RESULTS AND DISCUSSIONS

4.1. Line Graph

Fig. 1. Trend of Gross Domestic Product (GDP): 1994 - 2023. *Source: E-views output.*

Figure 1 plots the annual gross domestic product from 1994 to 2023. As the figure shows gross domestic product fluctuated over period under review. Highest gross domestic product observed over that period was 14.6% in 2002, indicating a significant surge in economic growth. The lowest gross domestic product occurred in 2002, with a contraction of -1.92%. This decline can be attributed to the global economic downturn triggered by COVID-19 and associated lockdown measures.

Fig 2. Trend of Private Sector Credit (PSC): 1994 - 2023. *Source: E-views output*

Figure 2 shows the visual trend and movement of private sector credit which shows the aggregate levels of credit disbursed by operating banks to the private sector over the period of study. From this pictorial representation, it could be observed that private sector credit fluctuated. This trend is most noticeable since 2007 after which, a momentous decrease was witnessed, until 2012. However, private sector credit again rose after 2012, but economy continued to fluctuate between 2013 to 2023.

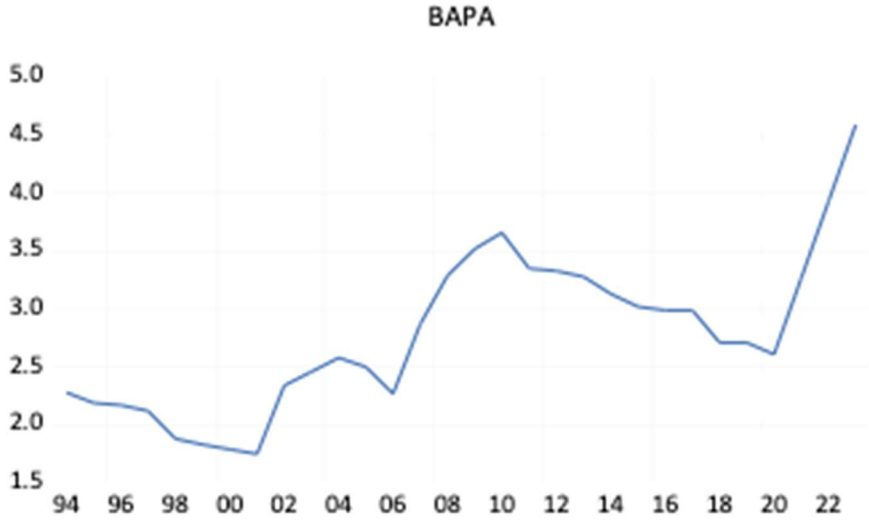

Figure 3. Trend of Number of Bank Accounts Per 100,000 Adults (BAPA): 1994 - 2023.*Source: E-views output*

Figure 3 shows the visual trend and movement of bank accounts per 100,000 adults from 1994 to 2023. It could be observed that the number of bank accounts per 100,000 adults has been fluctuating. This trend is most noticeable since 2011 after which, a momentous decrease was witnessed until 2020, after which it rose again.

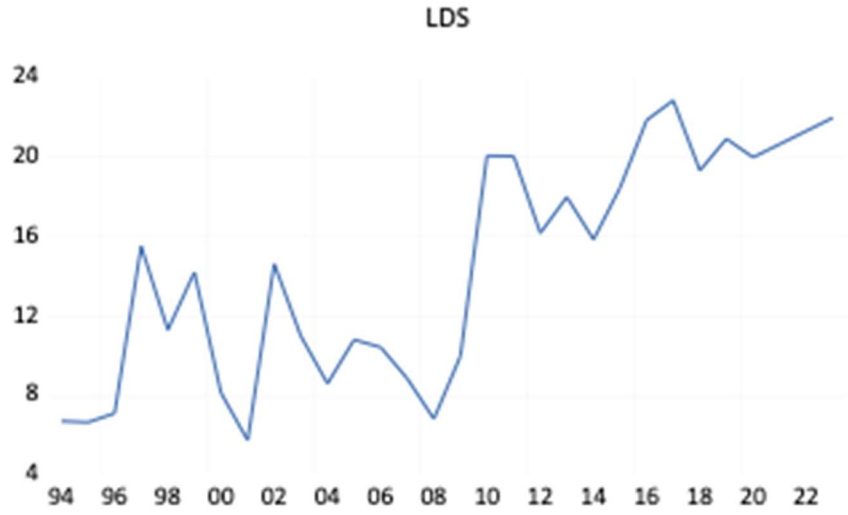

Figure 4. Trend of Lending-Deposit Spread (LDS): 1994 – 2023, *Source: E-views output*

Graphical trend above displays a fluctuating trend in Lending-deposit spread. The lending-deposit spread showed an increase in the beginning. The highest lending-deposit spread value recorded was 22.82 in 2017. The lowest lending-deposit spread occurred in 2006 and 2007, with values of 10.44 and 8.87, respectively, thus suggesting a narrower spread between lending and deposit rates during those years.

Figure 5. **Trend of Asset Quality Ratios (AQR): 1994 - 2023.** *Source: E-views output*

Figure 5 above shows the trend of asset quality ratios. The highest asset quality ratio value observed was 90.00% in 2001, indicating a relatively high level of asset quality ratio in financial institutions. The lowest asset quality ratio occurred in 2005, with a value of 60.43%. Thus, indicating a precarious asset quality ratio in financial institutions in that year.

4.2. Unit Root Test (Augmented Dickey Fuller Test)

Stationarity tests were carried out on data shown in table 1. The aim was to ascertain stationarity properties of available data to determine whether they are reliable for employment in this research work.

Table 1. Results of Augmented Dickey- Fuller (ADF) Unit Root Test at First Difference. Source: E-views Output

Variable	ADF test Statistic	Critical Value 5%			Order of Integration	
		1%	5%	10%		Prob.
GDP	-5.007282	-3.689194	-2.971853	-2.625121	I(1)	0.0004
PSC	-5.012783	-3.699871	-2.976263	-2.627420	I(1)	0.0004
BAPA	-2.795576	-3.689194	-2.971853	-2.625121	I(0)	0.0717
LDS	-5.350714	-3.699871	-2.976263	-2.627420	I(1)	0.0002
AQR	-6.255726	-3.689194	-2.971853	-2.625121	I(1)	0.0000

From table 4.1 above, gross domestic product, private sector credit, lending-deposit spread, and asset quality ratios were found stationary at first difference. However, number of bank accounts per 100,000 adult population failed the stationary at first difference. The prevalence of fractional/mixed integration even at first difference among the study variables, therefore, creates room for employment of auto regressive distributive bound test technique as necessary for analysis of the study data.

4.3. ARDL Model Estimation (Short Run)

Table 2. ARDL Model Estimation (Short Run).

Dependent Variable: GDP				
Method: ARDL				
Date: 07/24/24 Time: 13:32				
Sample (adjusted): 1996 2023				
Included observations: 28 after adjustments				
Maximum dependent lags: 2 (Automatic selection)				
Model selection method: Akaike info criterion (AIC)				
Dynamic regressors (2 lags, automatic): PSC BAPA LDS AQR				
Fixed regressors: C				
Number of models evaluated: 162				
Selected Model: ARDL(1, 1, 1, 0, 2)				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.***
GDP(-1)	**0.435939**	**0.136408**	**3.195847**	**0.0050**
PSC	**-0.799692**	**0.190817**	**-4.190885**	**0.0005**
PSC(-1)	**0.784282**	**0.239894**	**3.269279**	**0.0043**
BAPA	**10.19326**	**1.394670**	**7.308720**	**0.0000**
BAPA(-1)	**-8.553290**	**1.924095**	**-4.445358**	**0.0003**
LDS	**-0.366534**	**0.139326**	**-2.630764**	**0.0170**
AQR	**0.061402**	**0.060169**	**1.020491**	**0.3210**
AQR(-1)	**-0.126443**	**0.066797**	**-1.892949**	**0.0746**
AQR(-2)	**0.082855**	**0.061245**	**1.352839**	**0.1929**
C	**2.620163**	**7.350730**	**0.356449**	**0.7256**
R-squared	**0.865971**	**Mean dependent var**		**5.518571**
Adjusted R-squared	**0.798956**	**S.D. dependent var**		**4.079791**
S.E. of regression	**1.829295**	**Akaike info criterion**		**4.318191**
Sum squared resid	**60.23373**	**Schwarz criterion**		**4.793978**
Log likelihood	**-50.45467**	**Hannan-Quinn criter.**		**4.463644**
F-statistic	**12.92209**	**Durbin-Watson stat**		**1.792852**
Prob(F-statistic)	**0.000004**			

The coefficient of determination (R-squared) is 0.865971. It indicates approximately 86.59% variations in gross domestic product described discrepancies in employed explanatory variables in the model. The adjusted R-squared is 0.798956. Adjusted R^2 considers, any other adjustments effected for the number of predictors in the model. The F-statistic is 12.92209 and it involves probability of 0.000004, indicating the whole sample is substantially significant. The Durbin-Watson statistic is 1.792852, which offers no significant autocorrelation present across all model's residues. With respect to explanatory variables in the model, it is observed as follows:

Private Sector Credits: The coefficient of private sector credits variable is -0.547589. It is statistically significant (t-Statistic = 0.7996920, Prob.* = 0.0043), indicating that private sector credits have positive effects on gross domestic product. An increase in private sector credits relates to growth in gross domestic product in the short run.

Number of Bank Accounts Per 100,000 Adults (BAPA): The coefficient of the BAPA variable is 10.190326. It is statistically significant at 0.05 level (t-Statistic = -4.445358, Prob.* = 0.0010), indicating that number of bank accounts per 100,000 adults has a positive effect on gross domestic product. An increase in bank account per 100,000 adults is associated with an increment in gross domestic product.

One Year Lagged Value of Number of Bank Accounts Per 100, 000 Adults (BAPA)

This is statistically significance at 0.05 level (t-Statistic = -3.767332, Prob.* = 0.0003). It indicates that the lagged value of bank account per 100, 000 adults have a negative effect on the current gross domestic product. This suggests that past values of bank accounts per 100,000 adults have negative influence on current gross domestic product. This is seen as normal, because it is only the current values of bank deposits that can influence transactions, not past ones.

Lending-deposit Spread: The coefficient of the lending-deposit spread variable is -0.3665340. It is statistically significant at 0.05 level (t-Statistic = -2.195095, Prob.* = 0.0170). It indicates that lending-deposit spread has a negative implication on gross domestic product. An increase in lending-deposit spread relates to decrease in gross domestic product.

Asset Quality Ratio: The coefficient of asset quality ratio variable is 0.061402. It is not statistically significance at 0.05 level (t-Statistic = 1.020491, Prob.* = 0.3210). It indicates that asset quality ratio does not have a significant effect on GDP and as such, not a significant determinant of GDP growth at 0.05 level in the study's model.

These findings show that, in the short run, lagged gross domestic product values has a positive and statistically significance effects on gross domestic product. This proves that short term, past gross domestic product values positively influence current gross domestic product. While current private sector credits relate negatively with Nigeria's gross domestic product, it could be observed that lagged private sector credits has positive and statistically significant impacts on gross domestic product. This could partly attributed to the gestation or maturity periods which could be required for investments to be productive. This suggests that past values of private sector credits positively influence current gross domestic product in the short run. Basically, short run (ARDL) tests are accepted as litmus test.

4.4. Discussion of Findings

Concluded from the observations, the following were discussed:

Depth of Financial Institutions: The depth of financial institutions proxied by ratio of private sector credits to Nigeria's gross domestic product showed a profound impact on the growth of the economy. Although the coefficient is negatively signed in the long run, ARDL analysis (table 4.7) as -0.529885, this could be attributed to excessive cost of bank credits in Nigeria. The fact remains that the multiplier effect of bank credits is obvious on the economy. Most of these credits are significantly short-term commercial bank lending which attracts serious costs. However, it is expected that once the lending-deposit spread is addressed by the Central bank, this sign will change in the interest of the economy.

Access to Financial Services: Level of access to financial services proxied by number of bank accounts per 100, 000 adults have a strongly positive and significant effect on Nigeria's gross domestic product. Banking and other financial services grow through network (branch) expansion. Branch expansion ensures greater deposit mobilization, which generates associated businesses including credits and service incomes and commissions. In this direction, wider branch networks implicate expanded businesses which in turn, generates growth in the economy by facilitating financial product services to the economy.

Efficiency of Financial Institutions: The efficiency of financial institutions is proxied by lending-deposit spread. In table 4.7, lending-deposit spread has no significant statistical long run relationship with Nigeria's gross domestic product. The coefficient of -0.20631 (negatively signed) implicates the obvious fact that the spread is usually high in Nigeria, which at some point in time overall cost of credit is considered, attains up to 27% p.a. The result is in appropriate pricing of bank deposits (very low interest on savings and time deposits) which often, act as a serious disincentive to save among savers. The result shows significant unwillingness on the part of the public to deposit their funds into banks thereby, hindering deposit mobilization, investments and economic growth.

Stability of Financial Institutions: The stability of financial institutions proxied by asset quality ratio although positively signed is insignificantly related to Nigeria's gross domestic product. The result provides strong support for 2004/05 banking sector restructuring/consolidation in Nigeria. Asset quality ratios of Nigerian banks deteriorated very sharply from the year 2000, as elaborated by Soludo (2004). Weak loan repayment and debt recovery efforts have remained the key banes of Nigeria's banking sector. String legislation to ensure adherence to credit contracts are critical to performing/quality risk assets. Deteriorating risk assets weaken a nation's investment base and deaccelerate the economy. These trends agree with the assertions of Freixas and Rochet (2008) which observed the stable financial institutions inspire investors' confidence and ensure stable flow of credits in the economy.

5. CONCLUSION AND RECOMMENDATIONS

5.1. Conclusion

Given variables involved in this research, it is concluded that; credits to Nigeria's private sector and number of bank accounts are statistically material in predicting Nigeria's gross domestic product, while asset quality ratio and lending-deposit spread are not. Further, explanatory variables of this study – credit to private sector, asset quality ratio, lending-deposit spread and bank accounts per 100,000 supports or promotes Nigeria's economic growth and vice-versa.

5.2. Recommendation

Based on the findings from the analyses, the following recommendations were made:

- Central Bank of Nigeria should increase the prescribed loan deposit ratio for operating banks. This will enable operating banks to increase funding of investment projects in the private sector, Through the implied multiplier effects, disbursed credits/investments will accelerate Nigeria's economic growth.

- Nigerian banks should intensify their loan recovery efforts. This will ensure that customers do not divert borrowed funds into unproductive ventures, thereby reducing productive investments in the economy.

- Nigerian banks should be more aggressive in deposit mobilization through enhanced product development to develop more strategic sector specific credit/deposit product for better financial intermediation.

- Central bank of Nigeria should lower the requirements for network/branch expansion for existing financial institutions. This will even facilitate services and expansion of financial service providers like Point of Sale (POS) operators as they often require the services of nearest offices of operating banks for their own enhanced service delivery.

References

[1] Adeyemi, B., & Yusuf, T. (2023). Financial Development and Sectoral Growth: Evidence from Nigeria. *Journal of Development Economics and Policy*, 16(3), 72-88.
[2] Agbo, O., & Okeke, E. (2022). External Debt, Financial Development, and Economic Growth: Evidence from Nigeria. *Journal of International Economics and Finance*, 11(3), 89-107.
[3] Amadi, A., & Edeh, P. (2021). Financial Development and Economic Growth: Evidence from Nigeria. *Journal of Macroeconomic Studies*, 14(3), 22-40.
[4] Arestis, P., & Demetriades, P. O. (1991). Financial development and economic growth: assessing the evidence. *Economic Journal,* 107 (442): 783-799.
[5] Balogun, T., & Nwankwo, U. (2023). Institutional Quality, Financial Development, and Economic Growth Nexus in Nigeria: An ARDL Approach. *Journal of Development Economics and Policy*, 18(2), 51-69.
[6] Bilir, L. K., Chor, D., & Manova, K. (2019). Host-country financial development and multinational activity. *European Economic Review*, *115*, 192-220.
[7] Fashina, O. A., Asaleye, A. J., Ogunjobi, J. O., & Lawal, A. I. (2018). Foreign aid, human capital and economic growth nexus, Evidence from Nigeria. *Journal of International Studies,* 11(2): 104-17.
[8] Freixas, X. & Rochet, J. (2023). Microeconomics of banking, MIT Press,
[9] Fry, M. (1988). *Money, interest and banking in economic development.* Baltimore. The John Hopkins University Press.
[10] Ghirmay, T. (2004). Financial development and economic growth in Sub-Saharan African countries: Evidence from time series analysis. *African Development Review*, *16*(3), 415-432.
[11] Greenwood, J. & Jovanovic, B. (1990). Financial development, growth, & the distribution of income, *Journal of Political Economy*, *98*(5), 1076–1107.
[12] Hasan, I., Koetter, M., & Wedow, M. (2009). Regional growth and finance in Europe: Is there a quality effect of bank efficiency? *Journal of Banking & Finance*, *33*(8), 1446-1453.
[13] Ibrahim, M., & Alagidede, P. (2018). Effect of financial development on economic growth in sub-Saharan Africa. *Journal of Policy Modeling*, *40*(6), 1104-1125.
[14] Lawal, A., & Ibrahim, S. (2023). Financial Development, Inflation, and Economic Growth: Evidence from Nigeria. *Nigerian Journal of Economic Research*, 20(4), 93-110.
[15] Levine, R (2005). Finance and growth: Theory and evidence in Aghion, P and Durlauf. S. N. (eds.), *Handbook of Economic Growth*, London: Elsevier 1A, 865-934.
[16] Levine, R. (1997). Financial development and economic growth: Views and agenda. *Journal of Economic Literature*, *35*(2), 688–726.

[17] Levine, R. (1997). Financial Development and Economic Growth: Views and Agenda. *Journal of Economic Literature*, 35(2), 688-726.
[18] Levine, R. (2004). Finance and growth: Theory and evidence. Carlson School of Management. *University of Minnesota and the NBER*, *321*.
[19] McKinnon, R. I. (1973). *Money and capital in economic development*, Washington: Brookings Institute. Washington, DC.
[20] Ojo, F., & Ahmed, M. (2023). Financial Sector Competition and Economic Growth in Nigeria: A DSGE Model Approach. *Journal of Banking and Economic Development*, 12(2), 71-90.
[21] Pagano, M., & Jappelli, T. (1993). Information sharing in credit markets. *The Journal of Finance*, *48*(5), 1693-1718.
[22] Patrick, H. (1966). Financial development and economic growth in underdeveloped countries. *Economic Development and Cultural Change*, *14*(2), 174–189.
[23] Rewilak, J. (2017). The role of financial development in poverty reduction. *Review of development finance*, *7*(2), 169-176.
[24] Robinson, J. (1952). The generalization of the general theory, In *The Rate of Interest and Other Essays,* London: Macmillan. 69–142.
[25] Shaw, E.S. (1973). *Financial deepening in economic development:* New York, Oxford University Press York.
[26]
[27] Soludo, C. C. (2004). Consolidating the Nigerian banking industry to meet the development challenges of the 21st century, Being an address delivered at the special meeting of the bankers' committee, Abuja (July).
[28] World Bank (2018). *World development annual report*: The World Bank Group.

Appendix A. Data on Gross Domestic Product (GDP), Financial Institutions Depth (private sector credit) (PSC), Financial Institutions Access (Bank accounts per 100,000 adults) (BPA), Financial Institutions Efficiency (lending-deposit spread) (LDS), and Financial Institutions Stability (asset quality ratios) (AQR) in Nigeria over the period of 1994 to 2023.

Year	GDP (#'b) (A)	Credit to Private Sector (B) (#'b)	Bank Branches (C)	Nigeria Population (D) (Millions)	Bank Population by 100,000 (E) Bank account 100,000 people	Lending Rate (F) (%)	Deposit Rate (G) (%)	Non-performing Loan (H) (#'b)	Performing Loan (I) (#'b)
1994	21897.47	143.42	2403	105,355,783	379.28	21	14.27	3.392483	140.03
1995	21881.56	180	2368	108,011,465	388.84	20.79	14.12	4.195518	175.81
1996	22799.69	238.6	2407	110,732,904	398.64	20.86	13.74	6.422407	232.17
1997	23469.34	316.21	2407	113,522,705	408.68	23.32	7.81	16.95778	299.25
1998	24075.15	351.96	2185	116,385,750	418.99	21.34	10.03	16.05719	335.9
1999	24215.78	431.17	2185	119,327,073	429.58	27.19	12.99	32.4876	398.68
2000	25430.42	530.37	2193	122,352,009	440.47	21.55	13.4	36.78106	493.59
2001	26935.32	764.96	2193	125,463,434	451.67	21.34	15.58	0.0006	764.96
2002	31064.27	930.49	3010	128,666,710	463.2	30.19	15.57	0.4135	930.08
2003	33346.62	1096.54	3247	131,972,533	475.1	22.88	11.88	50.5015	1046.03
2004	36431.37	1421.66	3492	135,393,616	487.42	20.82	12.21	212.2915	1209.37
2005	38777.01	1838.39	3467	138,939,478	500.18	19.49	8.68	543.6776	1294.71
2006	41126.68	2290.62	3233	142,614,094	513.41	18.7	8.26	17.41565	2273.2
2007	43837.39	3668.66	4200	146,417,024	527.1	18.36	9.49	240.8424	3427.82
2008	46802.76	7899.14	4952	150,347,390	541.25	18.7	11.86	428.7275	7470.41
2009	50564.26	9889.58	5436	154,402,181	555.85	22.62	12.63	515.3447	9374.24
2010	55469.35	10518.17	5809	158,578,261	570.88	22.51	2.46	561.1546	9957.02
2011	58180.35	9600.02	5454	162,877,076	586.36	22.42	2.4	1339.918	8260.11
2012	60670.05	13293.64	5564	167,297,284	602.27	23.79	7.62	2073.347	11220.29
2013	63942.85	14461.41	5639	171,829,303	618.59	24.69	6.71	3409.524	11051.89
2014	67977.46	16753	5526	176,460,502	635.26	25.74	9.89	4042.916	12710.09
2015	69780.69	18688.42	5470	181,181,744	652.25	26.71	8.26	3900.771	14787.65
2016	68652.43	21025.24	5570	185,989,640	669.56	27.29	5.46	3668.855	17356.39
2017	69205.69	22459.18	5714	190,886,311	687.19	30.6	7.78	4326.574	18132.61
2018	70536.35	22646.33	5301	195,874,683	705.15	28.16	8.85	4850.307	17796.02
2019	72094.09	25676.87	5437	200,963,599	723.47	30.57	9.67	6237.968	19438.9
2020	70800.54	29030.01	5385	206,052,515	741.79	28.64	8.67	7858.071	21171.94
2021	73382.77	32868.49	6918	211,441,431	761.19	28.12	7.49	10944.55	21923.94
2022	75965	36706.97	8451	216,830,347	780.59	27.6	6.31	14031.04	22675.94
2023	78547.23	40545.45	9984	222,219,263	799.99	27.08	5.13	17117.52	23427.94

Source(s): Central Bank of Nigeria Statistical Bulletin (2018, 2023). World Bank: World Bank (2023). Global Financial Development Data Base-Measures of Financial Development

The Business of Trust: Towards a Sociological Orientation on Social Enterprise Projects with Disadvantaged Citizens

James MOIR [a*]

[a] *Abertay University, Faculty of Design, Informatics and Business, Old College, Bell Street Dundee, DD1 1HG j.moir@abertay.ac.uk*

Abstract

This conceptual paper is concerned with the issue of trust in the context of working with disadvantaged citizens on social enterprise projects. Trust is a major aspect of these kinds of project given that they often seek to empower people who face the sharp end of inequalities such as those living in deprived communities. The paper begins by considering the nature of social enterprises and how they seek to collaborate with citizens from disadvantaged communities. Research on collaborating with these communities is examined before turning to how the trust process has been modelled and applied in organizational investigations. A critical appraisal of this conventional approach is offered from the point of view of considering how mistrust from citizens in such communities is related to how they may have been disidentified in the past. From there the paper goes on to consider the application of the sociologist Harold Garfinkel's notion of trust conditions. The implications for considering these as pre-requisite for external agencies initiating social enterprise projects are considered.

Keywords: social enterprise, disadvantsge, trust conditions, sociology

1. INTRODUCTION

This paper is concerned with the issue of trust in the context of collaborating with disadvantaged citizens on social enterprise projects. Trust is a major aspect of these kinds of project given that they often seek to empower people who face the sharp end of inequalities such as those living in deprived communities. Gaining trust in such circumstances can prove difficult where citizens feel that they have little stake in society, or in seeking to improve their lives through working in a business context. Societal inequalities erode trust at the most basic level of identity and sense of agency. Such erosion has an effect upon citizens' ability and willingness to interact with others immediately beyond their community. The centrality of trust within the operation of social enterprises raises questions beyond conventional psychological trust-building models and invokes sociological questions around identity and interaction. The paper begins with an overview of the nature of social enterprise projects in terms of their transformational aims with respect to tackling issues that accompany social inequalities, as well as the centrality of trust for these projects. The conventional psychological treatment of trust is then critically considered based upon a tri-partite model conceptualized around calculus, knowledge, identification as sub-types. I then move on to consider the issue of trust from a sociological point of view in terms of its accomplishment in interaction and how this can be derailed through inequalities. In advancing this view I draw upon the work of Garfinkel (2019 [1962]) in which he points to the fragility of identity and interaction as well as his work paper (Garfinkel, 1963) on trust as a condition for interaction. Finally, this perspective is applied to considering the nature of social enterprise projects as an inherently trust-bound interactive

* Corresponding author.

process. All regions face similar challenges in delivering sustainable transport solutions to meet their current and future mobility requirements. Transport authorities are aware of the real needs specific to their region but often find it difficult to identify detailed information on targeted solutions that would deliver direct and tangible positive outcomes.

2. DISADVANTAGED CITIZENS, SOCIAL ENTERPRISES AND TRUST

Portales (2019: 10) considers social enterprises in the terms of being private sector companies "who see social challenges as business opportunities [...] consolidating business models with social purpose." These social business models therefore address social problems, such as those associated with deprived communities, in order to generate income and potentially scale up their operation. This often meets private company corporate social responsibility initiatives and can involve an interface between the private market and the third sector charity organizations Portales, 2019: 22). This can involve consulting with one another on matters such as donations, sponsorship, community investment initiatives, and payments for services or products. In other words, there are now established models where companies and third sector organizations work together in producing goods or services that meet the needs marginalized groups in, particularly those in deprived areas. SE projects are, therefore, much more embedded and part of local communities than traditional commercial businesses, as they seek to address social problems and issues within these communities.

Social enterprise projects often involve collaborating with local communities in ways that seek to exploit a sense of entrepreneurship. Scholars in the area present entrepreneurship as a means of addressing societal ills such as disadvantage, and a way of developing self-empowerment and wealth creation (Bruton et al., 2013). However, entrepreneurial action is set within social and cultural parameters (Bruton et al., 2010). In other words, entrepreneurship must be considered within a range of overlapping and intersecting historical, cultural, social, political, and institutional context that limit the scope for engaging in social enterprise projects (Aidis et al., 2008; Roundy, 2019; Urbano et al., 2019; Welter, 2011). In the context of social enterprise projects in deprived communities the question arises as to how these various boundaries often exclude the engagement of citizens and how they can be overcome, in order to maximize the involvement of actors and institutions alike. In this regard, social enterprises are part of an entrepreneurial ecosystem (Spanuth & Urbano, 2024) that involve different formal and informal actors and institutions (Audretsch et al., 2021; Auschra et al., 2019; Carayannis et al., 2019). Enabling disadvantaged citizens to engage with, and across, these different elements is a considerable hurdle and one that should not be underestimated.

In thinking about citizen engagement in social enterprises, it is tempting to adopt a neoliberal position on entrepreneurship in terms of overcoming embedded structural inequalities through notions of motivation, agency, and effort (Martinez Dy, 2020; Mole and Mole, 2010; Ahl and Marlow, 2012). However, considering entrepreneurship from a wider from a sociological perspective, disadvantage can be considered as being associated with associated with a range of factors such as social class (Anderson and Miller, 2003), ethnicity (Ram et al., 2017; Vershinina et al., 2019); gender (Ahl and Marlow, 2012), disability (Kasperova and Kitching, 2014), age (Mallett and Wapshott, 2015), migrant status (Sepulveda et al., 2011), and refugee status (Al-Dajani and Marlow, 2013). These distinct factors often co-occur in terms of being intersectional (Anthias, 2013), and therefore disadvantaged communities may not be homogenous in composition. This is important when considering the issue of trust as a feature of social enterprise community engagement. For example, Putnam (2007) found lower levels of trust in ethnically diverse communities, even when controlling for associated variables such as social class and income.

Trust is therefore a vital aspect of networks upon which social enterprises rely upon. Building and maintaining trust can be challenging, particularly in disadvantaged communities where formal institutions may be treated with some suspicion or where their influence is weak in comparison to the stronger bonds of informal inter-personal ties. However, if trust can be achieved, then social enterprises can use it effectively to support their aims and impacts. In order to build trust in collaborating with disadvantaged communities, social enterprises need to publicly demonstrate their commitment to improving the lives of those in these communities (Mason, 2012) as well as meeting the expectations of institutional actors (Czinkota et al., 2020). Trust is therefore a matter of social practice in terms of respecting and listening to communities in ways that demonstrate respect and the inclusion of their interests in a manner where partnership equality is central.

The above raises the issue of disadvantage and social capital Bertotti et al., 2012; Sacchetti and Campbell, 2014; Hidalgo et al., 2024). Trust between social enterprises and the communities they serve is crucial in terms of supporting the networks that enable their functioning. However, the social capital associated with a business perspective, even if it is in pursuit of social ends, may prove problematic. This issue has been raised in earlier work by Fenton et al. (1999) who argue that the adoption of a "business-like" culture in contrast to a more altruistic perspective can damage trust

between citizens and organizations. Smaller, local not-for-profit organizations were considered as being more trustworthy given the community-based beliefs and values they hold. They found that blurring of boundaries between sectors between the business world and the third sector eroded trust (Fenton et al., 1999, p. 25). Having said this, Hussain et al. (2025) point out that disadvantaged groups can and do engage in entrepreneurial activity in ways that may not follow explanations rooted in conventional psychological models such as the Theory of Planned Behaviour (Ajzen 1991). Their qualitative research focused on unemployed women residing in social housing in a deprived urban area of the United Kingdom. Participants' biographical narratives were associated with multiple layers of disadvantage, and accessibility to fewer resources. However, their positionalities (Anthias, 2002; 2013), revealed a motivation towards engaging in entrepreneurship and, in some cases, enabling them to convert their entrepreneurial intentions into action.

While Seanor and Meaton (2008:26) argue that social enterprise involves "more than the promotion of business ethos and practices" it is worth noting that they ways in which that business ethos is discursively constructed is important. For example, Levander's (2010) study of Swedish social enterprises, found that success stories are constructed around the transformation of 'problem identities' into 'resourceful identities'. This reformulation of disadvantaged identities strengthens the view of community engagement as agents of change that benefit society and addresses the challenges of disadvantage and deprivation. However, there remains the issue of trust and how citizens in such communities engage with external agencies that seek to promote social enterprises as a means of tackling disadvantage. The case study conducted by Finlayson and Roy (2018) in Scotland is informative in this respect. Their study draws attention to misunderstanding arising in communicating the purpose of the project, and how this led to confusion and lack of trust. In such cases, discursive legitimacy building becomes essential, emphasizing the need to empower communities to co-create solutions and actively involve them in the development and implementation of projects. Finally, Granados and Rosli (2020) found that stakeholder engagement is a crucial factor for social enterprises, particularly in the local community. Cooperation with local stakeholders lead to maximizing their impact and building reputation and credibility.

Social enterprise initiatives often need to enter into partnerships with local community leaders in their capacity as gatekeepers for acceptance into a community. It is therefore essential that these local leaders perceive the social enterprise and the relationships in entails as being beneficial to all rather than seeming threatening to their status (Raghubanshi et al., 2021). Once entry is gained, then a social enterprise can seek to become embedded in a community by effectively communicating its mission and activities. This is important given that, unlike commercial enterprises, which seek to market themselves and build social networks for the purpose of increasing market share, social enterprises utilise networks for the purpose of advocacy in relation to raising awareness and seeking policy changes on social issues of importance to local communities. However, all of this rests on the issue of trust. Building closeness and community support and trust is crucial when serving vulnerable communities (Seelos et al., 2011). Collaboration is therefore a key approach for building closeness and establishing trust and legitimacy.

Therefore, the success of these kinds of projects crucially rests upon citizen involvement in co-production. Co-production enables citizens from disadvantaged communities to be involved in the development and delivery of social enterprises, which make them ideal vehicles for corresponding social innovation. (Eriksson, 2022). What is therefore vital is that people from these communities can assume various aspects of the co-production process within a project (Cepiku et al., 2020: 3-4). Co-production is valuable in enabling the facilitation and development of skills in communities through the application of experience and a process of knowledge transfer (Wu et al., 2015: 2251). While social enterprises are not risk averse, they can be operated in such a way as to foster accountability (Defourny and Nyssens, 2021; Evers and Ewert, 2021). This can operate at the individual or collective level in which individual acts of co-production involve spontaneous or informal acts as a part of the delivery of a service, while collective acts depend upon "involve formally organized and institutionalized activities done together with others" (Pestoff, 2015: 4). Typically, these kinds of projects involve some combination of the two, particularly with respect to extended co-production efforts (Pestoff, 2015). Social enterprises therefore work best when different actors provide a mix of volunteering and professional-based activities (Defourny and Nyssens, 2021; Evers and Ewert, 2021).

Communication and interaction are vital for successful projects with institutions being connected in such a way as to support the development of social enterprises (Terrestrial et al., 2020: 887–888). There is scope here for the public sector to support such social enterprises through partnership working, innovations in policy, and the provision of learning resources. This approach enables a more integrated ecosystem where collaborative work can enhance social innovations in achieving service co-production (Lindsay et al., 2018). Utilizing digital technologies can also support this kind of ecosystems for enabling social enterprises as a means to social innovation (Mazzucato, 2019). Digital approaches and services that are related to social innovations can aid in achieving scalability and sustainability. Digital

co-production can encourage a shift from separate services towards a service ecosystem, and it therefore offers the possibility of enhancing citizen inclusion (Paananen et al., 2021; Perikangas and Tuurnas, 2023).

However, gaining citizen trust and engagement cannot be simply taken for granted as a process issue. Thus, while social enterprises often seek address systemic problems associated with disadvantaged communities, there is a need to ensure that all participants gain from there being equality in terms of input from different actors Windrum et al. (2016: 153–154) point out that while a service innovation is useful in seeking to improve the lives of those in disadvantaged communities, citizens must be treated as active participants rather than passive users of services. They must be included in co-creating and delivering innovations through their active participation and user insight. In other words, there has to be, from the outset, an understanding that such projects are not simply foist upon communities for the sake of some external rationale, but rather a sense of agency by citizens in playing a part in solving their problems. External organizations such as those in the third sector, private sector businesses, or knowledge-based institutions such as universities, can obviously play a major part in driving forward such innovations, but often the central goal to be kept in mind is that such organizational inputs are based on an intermediary capacity that supports citizens.

3. MODELLING THE TRUST-BUILDING PROCESS

The trust-building process has been the subject of investigation with general models proposed for how interpersonal trust develops, strengthens or declines, and can be disrupted. One of the main influential models focuses upon three bases that can lead to the development of trust: a rational base (calculus), then on to sharing experiences (knowledge), and finally achieving a high-relational base (identification) (Lewicki & Bunker, 1995, 1996). However, this kind of linear approach assumes a pathway to trust and may not capture the trust process in all its complexities. Some scholars have sought to take a non-linear approach by focusing on 'small wins' (Bryson, 1988). This perspective considers trust development in terms of a more iterative and cyclical process whereby repeated interactions over time (Huxham & Vangen, 2013) are cumulative, as well as offering a non-hierarchical path in which people treat each other on an initial trust basis and commit to a higher level of risk (Huxham & Vangen, 2013; Jones & George, 1998; Lewicki et al., 2006).

However, the tripartite model Shapiro et al. (1992) remains the main basis for considering the bases of trust in terms of calculus, knowledge, and identification (Lewicki & Bunker, 1995; Sheppard & Tuchinsky, 1996). Calculus-based trust involves deciding to engage in an exchange relationship on a beneficial in some way or other. This level of trust is taken as being foundational and is the basis on which subsequent levels are then built. Although this foundational level assumes trust, it is considered as being rooted in opportunism and as such may also be based on little or no trust or distrust (Bijlsma-Frankema & Costa, 2005; Lewicki et al., 2006). The next level, knowledge-based trust, relies upon knowing how to predict the behaviour of another in the exchange relationship. This derives from knowledge built up through multiple interactions over a period of time in which an individual can then gain an understanding of the other party as the basis for their decisions. The third and final level is identification-based trust which relies upon perceptions of similarity between the parties in terms of shared beliefs, preferences, or values that are internalized through interactions. It is argued that at this level of trust, power differences are not salient due to the development and acquisition of a sense of common identity (Kasten, 2018). This common identification is therefore rooted in an understanding of common aims that override the possibility of opportunistic intent being raised. These bases are assumed to be organized in a hierarchal manner with identification being considered as the most advanced level of trust (Sheppard & Tuchinsky, 1996). However, it is also thought of as being the level that is most difficult to attain given the higher level of interdependence. It is also considered as posing a greater level of risk should trust be breached and a state of mistrust occurs.

Although the model provides a broad-brush approach to trust development it does not deal with the dynamics of the relationship between the bases in terms of how interactions are patterned. Other factors that can impinge upon the trust development process have been examined, for example individuals who tend to exhibit high or low levels of trust, or the cognitive processes involved in evaluating a trustee's intention (McKnight et al., 1998). However, the main issue here, as with the model in general, is that the level of analysis is rooted in psychological processes and states devoid of any sociological input. There is little in the way of a consideration of the antecedents of trust and how certain groups in society may find themselves put in a disadvantageous position when it comes to operating within dominant forms of relations where their voices are less heard or not heard at all.

Farnese et al. (2022) used the model as the tripartite trust model as the basis for their study that focused on organizational culture in two social enterprises in Italy. Their study examined the micro-processes that underpin the model's bases 'down' at the level of institutional routines where trust becomes embedded. This more dynamic approach to model found a non-linear path of trust-building in operation including a norm of reciprocity. It was also

found that there was back-and-forth between cognitive and affective processes across all the bases. This more complex overlapping understanding overcomes the hierarchal basis of the tripartite model and instead suggests that a growth in trust can be nurtured by several bases at the same time. Möllering (2006) also adopted a more dynamic approach focusing on the reflexive process that can be part of the trust process as well as alongside the conventional bases. This involves individuals in a process continuously reflecting on their exchanges with others, potentially reevaluating them, and amending their expectations about the future exchanges. This reflexive process can be considered as a means through which the trust-building process is enhanced in a transitional manner (Koole, 2020).

While modelling the trust process has no doubt been useful from the point of view of understanding organizational relations and dynamics it has not been applied to an understanding of the beneficiaries of social enterprises and their relationships with the initiators of these projects, particularly where mistrust may rooted in the systematic disidentification of citizens from disadvantaged communities. Moreover, much of the work in this area is based upon psychological understandings of trust in terms of conventional conceptualizations of mental processes and states. This view tends to focus on trust at the level of cognitive and emotional processes within individuals rather than considering the actual interactional logics themselves and how these derive from normative methods of sense making that are prevalent in society. A fundamental question is how such normative methods align with power and privilege in terms of the ways in which they are related to matters of practical sociological understandings of how to proceed in interactions. In other words, how can trust be viewed in terms of the cultural resources drawn upon social actors as they engage with one another? A related question is what is the impact of a misalignment in the deployment of these resources in local situated practices? These kinds of questions take the issue of trust beyond that of mental processes and states and instead place it within the logics of the rules and expectations that social actors orient towards. This also raises the question of trust being something that is forged through interaction or something where it is a precondition that enables exchanges to proceed in a meaningful way. This leads to the following discussion on trust conditions and how these are crucial for those setting up social enterprises with disadvantaged citizens.

4. TRUST CONDITIONS

It has been argued that when considering social enterprises, trust is a crucial issue for those citizens who are involved in the co-creation process. There must be sufficient trust for citizens to engage with representatives of the organizations involved in these kinds of projects. However, this is not easy given that they may see themselves as marginal agents in the process in terms of the very inequalities they face in their everyday lives. What is required is a sociological approach in which trust is framed as the central issue where trust is a precondition in how we view social enterprises. In other words, we need to treat more seriously the 'social' in social enterprise by considering how this is instantiated within the interactions between project organizers and citizens.

One approach where trust has been explored and as a matter of crucial importance is ethnomethodology. This sociological perspective is concerned with the ethno (or 'folk') methods (the methods of constructing intelligibility) in the process of ongoing interaction. It was, from the outset, concerned with matters of social justice and the ways in which power and privilege affect those whose identities are normatively considered as marginal. The founder of this approach, Harold Garfinkel (1967), argued that inequalities bring about a breakdown in what he referred to as "trust conditions" Garfinkel (2019 [1962]). This breakdown results in unintelligible interaction for the social actors involved. Moreover, his investigations into how people deal with incongruous information about others revealed that those who resist such information also resist challenging their views on inequalities. Therefore, the importance of this work lies in understanding the way in which disadvantage and inequality can lead to troubling issues for the conduct of interactions between those who are part of such communities and others who are not. Social inequalities are often intersectional in nature, for example, those in ethnic minorities and the intersecting issue of poverty. When citizens from these disadvantaged minorities engage in interactions with those from majority identities problematic interactional issues can occur leading to the reinforcement of prevailing social expectations thereby making them resistant to change.

Therefore, the (pre)conditions for mutually intelligible interaction, rest upon a reciprocal understanding of, and commitment to, common practices or methods of sense assemblage. In cases where these conditions are absent then interaction becomes problematic for participants. When this occurs on a repeated basis, those who can be considered as possessing disadvantaged identities are at risk of being put in more marginal and vulnerable positions and also are more at risk of being blamed by those in a more dominant position for any resulting interactional trouble (Rawls and David, 2005). These commonly involve an infringement of 'morality of appearances' (Goffman, 1959: 251) and the failure to legitimate another's claim to their identity as performed in interaction. What this amounts to is that in

interacting with dominant or majority others, citizens from disadvantaged backgrounds may be treated in such a way that nullifies or violates their identities.

Goffman (1959:13) argued that for the interaction order to proceed smoothly the production of identities extends a moral obligation to be treated in a manner in keeping with of the definition of the situation. In other words, that an alignment of social actors' normative expectations about how to proceed should be present. However, those from disadvantaged communities who face inequalities are often faced with situations where their self-presentations are not considered legitimate, or where on the basis of precious interactions they anticipate this and therefore where trust conditions are not met. This can lead to interactions where those inequalities can lead to prejudiced expectations being confirmed within the details of the exchanges by those in the dominant majority. It therefore little wonder that repeated interactional troubles lead those in marginal communities to either avoid exchanges with those in majority positions or to where they attempt to adopt defensive manoeuvres. These kinds of reactions are rooted in the tacit shared understanding of those communities. Such defence mechanisms can make it appear to those in dominant majority positions that they are in keeping with 'type' by being 'difficult' or non-cooperative, thereby maintaining conventional stereotypes.

Identity is a fragile social accomplishment (Garfinkel, 2006 [1948]; Goffman, 1983; Rawls, 1987) that can be impeded or problematized unless recognized by others during the course of interaction. For those from disadvantaged communities with marginal identities, it is common for their identities to be called into question and made problematic. Berger (1963: 103) made the point that 'human dignity is a matter of social permission' and those who face inequalities can often find that such permission is denied when interacting with those in the dominant majority. This can result in citizens from minorities staying within the confines of their own communities where interaction orders are more conducive to the recognition of their identities (Rawls, 2000). In so doing, the prevailing divide between majority and minority interaction orders are maintained as almost separate from each other. Trust conditions that span these orders become weaker and interactions become more problematic. The net result is the segregation of such communities who become deemed as 'problem areas.'

Garfinkel's early work focused on the role that inequality plays in the resulting disparities in the treatment of citizens in the United States. This work was concerned with how such inequality led to damage to identity and the process of sense-making in interaction. His studies ranged over a number of such damaged identities and their interactional problems, including, Jewish pre-medical candidates, and transsexual people (see Garfinkel, 1967: 116 – 185), the difficulties faced by the sight-impaired, mentally ill, and the intellectually and physically disabled (Turowetz & Rawls; 2021). This work demonstrated the 'practical methodologies' (Garfinkel, 1967: 183) used by those persons to pass as 'normals'. Such attempts to pass involved seeking to understand the tacit interactional rules that the majority control, thereby forcing them into guessing what others such as white men, medical school admissions officers, the police and those who gender normative, make of their actions in with contexts where they face unfamiliar expectations.

His later breaching studies grew out his work on the systematic troubles faced by minorities, that Garfinkel was able to reveal the interactional work of majority social actors put into accomplishing the seeming 'naturalness' of everyday social world. Such breaches expose the constitutive expectancies that are used to navigate interaction and make apparent the commitment to equality and reciprocal treatment in everyday life. In other words, interaction requires and implicit level of trust in order to accomplish meaningful interaction and the production of social order. In contexts where there are inequalities and interactional expectations troubled in some way or other, the result is a breakdown in the accomplishment of constitutive ordering practices. This is a crucial point in considering the nature of trust. In the conventional modelling of trust, it is something that is built up (or challenged) through the course of interactions. In other words, interactional participants can gain or lose trust through their exchanges without the intelligibility of the interaction being at stake. In this view trust is akin to being something like persuasion. Through interactions participants are persuaded to trust one another. Trust can be won or lost through the exercise influence co-participants have over one another. However, Garfinkel's notion of trust conditions does not adhere to this position and begins with the analytical prior question of trust as being a constitutive condition on which identities are dependent. Trust is therefore contingent upon a tacit understanding of the constitutive rules of the situation and the mutual commitment by participants to them.

Garfinkel's position is bound up with that of Durkheim (1893) in terms and the view that a social contract is a pre-requisite for society, and that this in turn rests upon a morality of social justice. This insight was also taken up by Parsons (1937), who took the view that 'rational' citizens exist only through social relations and that an implicit social contract is a prerequisite for shared meaning. Garfinkel's position is rooted in these views in terms of social justice being crucial for the accomplishment of social order through interaction. By drawing upon Durkheim's notion of 'constitutive practices' as well as the Durkheim/Parsons focus on implicit social contract Garfinkel was able to

examine the nature of trust conditions (Garfinkel, 2019 [1962]: Chapter 4). In doing so, his concern was with the ways in which inequalities disrupt the reciprocal requirements of this implicit social contract.

Later, his work on the tacit knowledge that underpins social practices and the conditions of trust they require developed beyond inequalities and became part of a broader focus on the constitutive nature of reciprocity in interaction across a range of settings. However as, Turowetz and Rawls (2021: 8) write the "persistent misinterpretation of Garfinkel's position is a loss to both sociology and the general aim of achieving a just society." Nevertheless, others have built upon his approach and examined various disadvantaged and marginal identities such as Black Americans (Rawls, 2000; Rawls and Duck, 2017; Rawls et al., 2018), autistic children (Turowetz, 2015), and person deemed to be police 'suspects' (David et al., 2018). Cutting across these studies is a focus on how such marginal identities are 'troubled' in interaction where trust conditions break down. Citizens from marginal positions are more likely to face attributions of blame, be sanctioned in some way or other, or face violence. In showing that minorities are more vulnerable to interactional breakdowns than their majority counterparts, Garfinkel was able to demonstrate the social practices through which social inequalities can create interactional misunderstandings that in turn perpetuate and reinforce those inequalities. Citizens from disadvantaged minorities have their own experience and heightened consciousness of such interactional troubles.

Given the above concern with trust as a vital element in interaction and sense-making, it is therefore important to its role in social enterprises. Citizens from disadvantaged communities are likely to be wary of organizational actors who seek to establish communication with them about setting up such projects. Their experience is likely to lead them to be sceptical about the intentions of such actors as well as making for make for difficult interactions. It is one thing to model trust as a (mental) process that is built up through the course of interactions through various organizational mechanisms. However, it is another thing to consider the antecedents of trust as something that lies in tacit and shared understandings of the rules of interaction in particular contexts. Mistrust for Garfinkel is rooted in the latter, where people from minorities do not share these understandings or where they must try and second guess what they are. Repeated interactional troubles for these group leads to mistrust from the outset for those who are official initiators of social enterprises. These typically tend to be representative from universities, charities and other third sector bodies. While the mission of social enterprises is laudable, officials who come into contact with disadvantaged minorities are therefore, at least initially, liable to me met with mistrust.

5. IMPLICATIONS FOR SOCIAL ENTERPRISES

There are two sets of inter-related implications arising from the argument above: conceptual and practical. In taking the conceptual implications first it is apparent that the ethnomethodological approach to trust conditions is at odds with the conventional tripartite trust model and the other variations of it. The conventional model takes trust as something that can be developed in a linear manner over time moving from calculus-based thinking through to being rooted in knowledge and finally to being identification-based. The attempts to make the model more dynamic through investigating the non-linear nature of the trust bases in social enterprises by Farnese et al. (2022) is helpful but nevertheless still situates trust within the broad outline of the model. An ethnomethodological approach also frees up conceptualizing trust as a linear process and permits flexibility in the form of constitutive trust practices. However, for Garfinkel these practices require as a pre-condition equality and reciprocity. He also treated disadvantaged groups as cases of interest precisely because they reveal how some categories of marginalized citizens encounter difficulties in achieving "normal" meanings and identities in interaction with those who are in majority groups. Trust for these minority groups becomes problematized in two ways: first, they are aware of moral order that consigns them to be counted a s being citizens who are not party to the majority rules of interaction; and second, although they may try to interact with persons of a majority category, such attempts means that they on guard for contingencies that could lead to misunderstandings thereby resulting in mistrust.

The arguments above set a case for adopting a more sociological perspective in examining the issue of trust by local citizens in disadvantaged communities in those who initiate social enterprise projects. This is set within the broader context of social enterprise involving more than the promotion of business ethos and practices. Social enterprise is fundamentally concerned with sociological themes of social justice and community cohesion. These underlying social values drive social enterprise projects that work with disadvantaged citizens (Pearce, 2003). However, while this admirable mission means that such projects involve reaching out to local disadvantaged communities, there is the potential for mistrust at the outset given the embedded ethnocentric nature of the process of interaction. Garfinkel's argument that what he called "trust conditions" rest upon reciprocity and cooperation sit well with the ethos of social enterprise. However, while trust may be treated as a matter of persuasion by those responsible for initiating social enterprise projects in terms of "winning over" local disadvantaged communities, their "buy-in" is

treated as a matter of changing attitudes towards seeing the benefits of setting up micro-businesses for the good of those communities. Taking Garfinkel's "trust conditions" as a starting point means putting in place the necessary set of prerequisites prior to setting up such projects. In practice this means a long lead-in time is required to establish these trust conditions where project initiators and other organizational actors become familiar to local disadvantaged communities. Moreover, this requires a deep understanding of how those local communities operate, their ways of seeing the world and interacting in it. This is not just a practical matter of enabling better communication between project organizers and local participants but rather is key to sensitizing organizers to the differences between themselves as majority members of society to the ways in which inequity in society leads to disadvantaged communities relying on their own networks and how trust operates at a local level.

What all of the means is that those setting up social enterprise projects need to understand how the category 'disadvantage' becomes reified in a way that is seen as requiring some form of intervention as in order to attempt to alleviate the 'problem.' However, the issue with such problematization is that it can lead to a managerialist type of thinking in which solutions are sought in terms of a business ethic that may be out of kilter with the ways in which local communities operate. Again, this requires a much more reflexive stance to be adopted by social enterprise organizers so that they collaborate with communities and understand them from the 'inside.

6. CONCLUSION

Social enterprises are needed now more than at any point due to declining public investment, citizen disenfranchisement, and increasing social divisions. However, many social enterprise projects are initiated without considering how (mis)trust features in the initiation of these projects. The systematic ways in which trust is problematized in disadvantaged or deprived communities through interactions with those with majority recognized identities need to be recognized. These minority communities commonly have a history of troubled interactions with officials from the majority. Such repeated instances of failure of reciprocity of legitimate identities in interaction leads to various reactions such as mistrust, eschewing such interactions, or resistance.

However, it is also the case that social enterprises can open the door to the establishment of trust through interactions and processes of co-creation and co-production rooted in norms of inclusivity, equality, and a sense of social justice. Garfinkel's work demonstrates the social practices involved in the assemblage of sense-making and what happens when those in minority disadvantaged positions are not 'in the know' of how to engage with those practices which the majority seem to perform so effortlessly. For this reason, it has been argued that social enterprise organisers must take the business of trust seriously in setting up such projects. Trust is not an abstract conceptual matter but rather something that is inherent in the ways in which interaction is conducted. Gaining funding for social enterprise projects is often difficult but gaining and building trust with citizens who may benefit from such project may potentially be a much more difficult prospect.

References

[1] Ahl, H. & Marlow, S. (2012), Exploring the dynamics of gender, feminism and entrepreneurship: advancing debate to escape a dead end?" Organization, 19 (5), 543-562.
[2] Al-Dajani, H. & Marlow, S. (2013), Empowerment and entrepreneurship: a theoretical framework. International Journal of Entrepreneurial Behavior and Research, 19 (5), 503-524.
[3] Aidis, R., Estrin, S. & Mickiewicz, T. (2008) Institutions and entrepreneurship development in Russia: a comparative perspective. Journal of Business Venturing, 23, 656–67.
[4] Ajzen, I. (1991). The theory of planned behavior. Organizational behavior and human decision processes, 50(2), 179-211.57–183.
[5] Anderson, A.R. & Miller, C.J. (2003), Class matters': human and social capital in the entrepreneurial process. The Journal of Socioeconomics, 32 (1),17-36.
[6] Anthias, F. (2002), Where do I belong? Narrating collective identity and translocational positionality, Ethnicities, 2 (4), 491-514.
[7] Anthias, F. (2013), Hierarchies of social location, class and intersectionality: towards a translocational frame. International Sociology, 28 (1), 121-138.
[8] Audretsch, D. B., Belitski, M. & Cherkas, N. (2021) Entrepreneurial ecosystems in cities: the role of institutions. PLoS ONE, 16, e0247609.
[9] Auschra, C., Schmidt, T. & Sydow, J. (2019) Entrepreneurial ecosystems as fields: integrating meso-level institutional theory. Zeitschrift für Wirtschaftsgeographie, 63, 64–78.
[10] Berger, P. L. (1963). Sociology as a form of consciousness: Invitation to Sociology. New York: Doubleday.
[11] Bertotti, M., Harden, A., Renton, A., & Sheridan, K. (2012). The contribution of a social enterprise to the building of social capital in a disadvantaged urban area of London. Community Development Journal, 47(2), 168-183.
[12] Bijlsma-Frankema, K., & Costa, A. C. (2005). Understanding the trust-control nexus. International Sociology, 20(3), 259-282.

[13] Bruton, G. D., Ahlstrom, D. & Li, H. L. (2010) Institutional theory and entrepreneurship: where are we now and where do we need to move in the future? Entrepreneurship Theory and Practice, 34, 421–440
[14] Bruton, G.D., Ketchen, D.J., Jr & Ireland, R.D. (2013) Entrepreneurship as a solution to poverty. Journal of Business Venturing, 28 (6), 683-689.
[15] Bryson, J. (1988). Strategic planning: Big wins and small wins. Public Money & Management, 8(3), 11–15.
[16] Carayannis, E. G., Grigoroudis, E., Stamati, D., & Valvi, T. (2019) Social business model innovation: A quadruple/quintuple helix-based social innovation ecosystem. IEEE Transactions on Engineering Management, 68, 235–24
[17] Cepiku, D., Marsilio, M., Sicilia, M. & Vainieri, M. (2020), Introduction. Cepiku, D., Marsilio, M., Sicilia, M. and Vainieri, M. (Eds), The co-production of public services, Palgrave Macmillan, Cham, pp. 1-10.
[18] Czinkota, M., Kaufmann, H. R., Basile, G. & Ferri, M. A. (2020) For Benefit company (fBComp): An innovative social-business model. The Italian case. Journal of Business Research, 119, 377–387.
[19] David G., Rawls A.W. & Trainum J (2018) Playing the interrogation game: Rapport, coercion and confessions in police interrogations. Symbolic Interaction 41(1), 3–24.
[20] Defourny, J. & Nyssens, M. (2021), Social Enterprise in Western Europe: Theory, Models and Practice, Routledge: New York.
[21] Durkheim, E. ([1893]1933). Division of Labor in Society, trans. George Simpson. New York Free Press.
[22] Eriksson, E. (2022), Coproduction and inclusion: a public administrator perspective. International Public Management Journal, Vol. 25 No. 2, pp. 217-240.
[23] Evers, A. & Ewert, B. (2021), Understanding Co-production as a social innovation. Loeffler, E. and Bovaird, T. (Eds), The Palgrave Handbook of Co-production of Public Services and Outcomes, Palgrave Macmillan: Cham, 133-153.
[24] Farnese, M. L., Benevene, P., & Barbieri, B. (2022). Learning to trust in social enterprises: The contribution of organisational culture to trust dynamics. Journal of Trust Research, 12(2), 153-178.
[25] Fenton, N., Passey, A. & Hems, L. (1999), Trust, the voluntary sector and civil society. International Journal of Sociology & Social Policy, 19 (7/8), 21-42.
[26] Finlayson, E. & Roy, M. J. (2018) Empowering communities? Exploring roles in facilitated social enterprise. Social Enterprise Journal, 15, 76–93.
[27] Garfinkel H (2019 [1962]) Parsons' primer. Stuttgart: Springer.
[28] Garfinkel H. (1963) A conception of an experiments with 'trust' as a condition of concerted actions. In: Harvey OJ (ed.) Motivation and social interaction: cognitive approaches. New York: Ronald Press, 187–238.
[29] Garfinkel, H. 1967 Studies in Ethnomethodology. Englewood Cliffs, N.J.: Prentice- Hall.
[30] Garfinkel H (2006 [1948]) Seeing Sociologically. Boulder, CO: Paradigm.
[31] Goffman E (1959) Presentation of Self in Everyday Life. New York: Double Day.
[32] Goffman, E. (1983). The interaction order: American Sociological Association, 1982 presidential address. American sociological review, 48(1), 1-17.
[33] Granados, M. L. & Rosli, A. (2020) Fitting In' vs. Standing Out': how social enterprises engage with stakeholders to legitimize their hybrid position. Journal of Social Entrepreneurship, 11, 155–176.
[34] Hidalgo, G., Monticelli, J. M., & Vargas Bortolaso, I. (2024). Social capital as a driver of social entrepreneurship. Journal of Social Entrepreneurship, 15(1), 182-205.
[35] Hussain, S., Vershinina, N., & Carey, C. (2025). Defying the odds? Multiple disadvantage as a source of entrepreneurial action. International Journal of Entrepreneurial Behavior & Research, 31(1), 219-245.
[36] Huxham, C., & Vangen, S. (2013). Coping with trust. In C. Huxham & S. Vangen (Eds.), Managing to collaborate: The theory and practice of collaborative advantage (pp. 138–156). London and New York: Routledge.
[37] Jones, G. R., & George, J. M. (1998). The experience and evolution of trust: Implications for cooperation and teamwork. Academy of Management Review, 23(3), 531–546.
[38] Kasperova, E. & Kitching, J. (2014), Embodying entrepreneurial identity. International Journal of Entrepreneurial Behavior and Research, 20 (5), 438-452.
[39] Koole, B. (2020). Trusting to learn and learning to trust: A framework for analyzing the interactions of trust and learning in arrangements dedicated to instigating social change. Technological Forecasting and Social Change, 161, 120260.
[40] Levander, U. (2010). Social enterprise: Implications of emerging institutionalized constructions. Journal of social entrepreneurship, 1(2), 213-230.
[41] Lewicki, R. J., & Bunker, B. B. (1995). Trust in relationships: A model of development and decline. In B. B. Bunker, J. Z. Rubin, & Associates (Eds.), Conflict, cooperation, and justice: Essays inspired by the work of Morton Deutsch (pp. 133–173). San Francisco, CA: Jossey-Bass.
[42] Lewicki, R. J., & Bunker, B. B. (1996). Developing and maintaining trust in work relationships. In R. M. Kramer & T. R. Tyler (Eds.), Trust in organizations: Frontiers of theory and research (pp. 114–139). Thousand Oaks, CA: Sage.
[43] Lewicki, R. J., Tomlinson, E. C., & Gillespie, N. (2006). Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. Journal of Management, 32 (6), 991–1022.
[44] Lin, N. (1999). Social networks and status attainment. Annual review of sociology, 25(1), 467-487.
[45] Lindsay, C., Pearson, S., Batty, E., Cullen, A.M. & Eadson, W. (2018), Co-production and social innovation in street-level employability services: lessons from services with lone parents in Scotland. International Social Security Review English Edition, 71 (4) 33-50.
[46] Mallett, O. & Wapshott, R. (2015), Making sense of self-employment in late career understanding the identity work of olderpreneurs. Work, Employment and Society, 29 (2), 250-266.
[47] Martinez Dy, A.M. (2020), Not all entrepreneurship is created equal: theorising entrepreneurial disadvantage through social positionality, European Management Review, 17 (3) 687-699.
[48] Mason, C. (2012) Isomorphism, social enterprise and the pressure to maximise social benefit. Journal of Social Entrepreneurship, 3, 74–95.
[49] Mole, K.F. & Mole, M. (2010), Entrepreneurship as the structuration of individual and opportunity: a response using a critical realist perspective: comment on Sarason, Dean and Dillard, Journal of Business Venturing, 25 (2), 230-237.

[50] Mazzucato, M. (2019), Governing missions in the European union. European commission. Directorate General for Research and Innovation, available at: https://ec.europa.eu/info/sites/info/files/research_and_innovation/contact/documents/ec_rtd_mazzucato-report-issue2_072019.pdf
[51] McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. Academy of Management Review, 23(3), 473–490
[52] Möllering, G. (2006). Trust: Reason, routine, reflexivity. Bingley: Emerald Group Publishing.
[53] Paananen, R., Alanko, R. and Kainulainen, S. (2021), "Digitaalinen Zekki-palvelu yhdistaa Nuorten tarpeet ja tarjolla olevan tuen", Kasvun Tuki Aikakauslehti, 1, 48-59,
[54] Parsons, T. (1937). The structure of social action. New York: The Free Press.
[55] Pearce, J. (2003), Social Enterprise in Anytown, Calouste Gulbenkian Foundation: London
[56] Perikangas, S., & Tuurnas, S. (2024). Design for inclusive digital co-production. Public Management Review, 26(6), 1731-1751.
[57] Pestoff, V. (2015), Co-production as social and governance innovation in public services. Polityka Społeczna, 11 (1) 2-8.
[58] Portales, L. (2019). Social innovation and social entrepreneurship. Fundamentals, Concepts, and Tools. Cham, Switzerland: Palgrave Macmillan.
[59] Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century. The 2006 Johan Skytte Prize Lecture. Scandinavian Political Studies, 30, 137–174.
[60] Putnam, R. D., Nanetti, R. Y., & Leonardi, R. (1994). Making democracy work: Civic traditions in modern Italy. Princeton: Princeton University Press.
[61] Raghubanshi, G., Venugopal, S. & Saini, G. K. (2021) Fostering inclusive social innovation in subsistence marketplaces through community-level alliances: an institutional work perspective. Industrial Marketing Management, 97, 21–34.
[62] Ram, M., Jones, T. & Villares-Varela, M. (2017), Migrant entrepreneurship: reflections on research and practice, International Small Business Journal, 35 (1), 3-18.
[63] Rawls, A. W. (1987). The interaction order sui generis: Goffman's contribution to social theory. Sociological theory, 136-149.
[64] Rawls A. W. (2000) 'Race' as an interaction order phenomenon: WEB Du Bois's 'double consciousness' thesis revisited. Sociological Theory 18(2): 241–274.
[65] Rawls A. W. & David G (2005) Accountably other: Trust, reciprocity and exclusion in a context of situated practice. Human Studies 28(4): 469–497.
[66] Rawls AW and Duck W (2017) Fractured reflections of high status black male presentations of self: Non-recognition of identity as a barrier to exercising legitimate authority. Sociological Focus 50(1): 36–51.
[67] Rawls A. W., Duck W. & Turowetz, J. (2018) Problems establishing identity/residency in a city neighborhood during a black/white police-citizen encounter: Reprising Du Bois' conception of submission as 'submissive civility'. City & Community 17(4): 1015–1050.
[68] Roundy, P. T. (2019) "It takes a village" to support entrepreneurship: intersecting economic and community dynamics in small town entrepreneurial ecosystems. International Entrepreneurship and Management Journal, 15, 1443–1475
[69] Sacchetti, S., & Campbell, C. (2014). Creating space for communities: social enterprise and the bright side of social capital. Journal of Entrepreneurial and Organizational Diversity, 3(2), 32-48.
[70] Seelos, C., Mair, J., Battilana, J. & Dacin, M. T. (2011) The embeddedness of social entrepreneurship: understanding variation across local communities. In Marquis, C., M. Lounsbury, & R. Greenwood, (Eds), Communities and organizations. Bingley: Emerald Group, pp. 333–363.
[71] Seanor, P., & Meaton, J. (2008). Learning from failure, ambiguity and trust in social enterprise. Social Enterprise Journal, 4(1), 24-40.
[72] Sepulveda, L., Syrett, S. & Lyon, F. (2011), Population superdiversity and new migrant enterprise: the case of London, Entrepreneurship and Regional Development, 23 (7-8), 469-497.
[73] Shapiro, D. L., Sheppard, B. H., & Cheraskin, L. (1992). Business on a handshake. Negotiation Journal, 8 (4), 365–377.
[74] Sheppard, B. H., & Tuchinsky, M. (1996). Micro-OB and the network organisation. In R. M. Kramer & T. R. Tyler (Eds.), Trust in organisations Frontiers of theory and research (pp. 140–165). London: Sage.
[75] Spanuth, A., & Urbano, D. (2024). Exploring social enterprise legitimacy within ecosystems from an institutional approach: A systematic literature review and research agenda. International Journal of Management Reviews, 26(2), 211-231.
[76] Terstriep, J., Rehfeld, D. and Kleverbeck, M. (2020), "Favourable social innovation ecosystem(s)? – an explorative approach", European Planning Studies, Vol. 28 No. 5, pp. 881-905.
[77] Turowetz J (2015) The interactional production of a clinical fact in a case of autism. Qualitative Sociology 38(1): 57–78.
[78] Turowetz, J., & Rawls, A. W. (2021). The development of Garfinkel's 'Trust' argument from 1947 to 1967: Demonstrating how inequality disrupts sense and self-making. Journal of Classical Sociology, 21(1), 3-37.
[79] Urbano, D., Aparicio, S. & Audretsch, D. (2019) Twenty-five years of research on institutions, entrepreneurship, and economic growth: what has been learned? Small Business Economics, 53, 21–49.
[80] Vershinina, N., Rodgers, P., McAdam, M. and Clinton, E. (2019), Transnational migrant entrepreneurship, gender and family business", Global Networks, 19 (2),238-260.
[81] Windrum, P., Schartinger, D., Rubalcaba, L., Gallouj, F., & Toivonen, M. (2016). The co-creation of multi-agent social innovations: A bridge between service and social innovation research. European Journal of Innovation Management, 19(2), 150-166.
[82] Welter, F. (2011) Contextualizing entrepreneurship - conceptual challenges and ways forward. Entrepreneurship Theory and Practice, 35, 165–184
[83] Wu, L. W., Lii, Y. S., & Wang, C. Y. (2015). Managing innovation through co-production in interfirm partnering. Journal of Business Research, 68(11), 2248-2253.

The Origins of the Firm: Theoretical Insights and Empirical Evidence from Italian SMEs - A Preliminary Study

Paolo ROFFIA [a*]

[a] *University of Verona, Department of Management, Via Cantarane 24, Verona, 37124, Italy,*
paolo.roffia@univr.it[b]

Abstract

This paper explores the fundamental question of why firms exist by examining the theoretical and psychological drivers behind their emergence. Building on the foundational insights from behavioral and entrepreneurial perspectives, particularly the Theory of Planned Behavior and the Theory of Entrepreneurial Orientation, the paper also reviews alternative theoretical explanations from Ronald Coase's Theory of Transaction Costs to Property Rights Theory, Resource-based View, and Evolutionary Economics. Using a quantitative methodology based on a sample of 111 Italian small and medium-sized enterprises (SMEs), this study empirically explores the underlying motivations behind firm creation. Partial Least Squares Structural Equation Modeling (PLS-SEM) was employed to test the theoretical relationships. The findings indicate that attitude and subjective norms significantly influence the intention to establish a firm, whereas resource-based capabilities and perceived behavioral control little or no substantial effect. Through historical examples, theoretical frameworks, and empirical testing, the article concludes that firms emerge not only as efficient solutions to coordination problems but also as institutional responses to social, psychological, and entrepreneurial stimuli.

Keywords: new business birth, theory of Planned behaviour, SMEs, entrepreunerial orientation

1. INTRODUCTION

The question of why firms exist lies at the heart of organizational economics and business history. While markets are efficient mechanisms for coordinating economic activity through prices, they are not without friction. Firms, as organizational entities that internalize transactions, represent an alternative coordination mechanism (Coase, 1937). But what are the precise conditions under which firms arise? What inefficiencies or constraints give rise to these structured organizations? This paper seeks to address these questions by exploring the rationale for the firm's existence from theoretical, behavioral, historical, and empirical perspectives.

Understanding the origins of firms is not merely a theoretical pursuit but a practical concern, especially in contexts where entrepreneurship plays a key role in economic development and employment generation. The question gains even more relevance in the case of small and medium-sized enterprises (SMEs), which often emerge not from corporate planning but from individual initiative, social context, and perceived opportunity (OECD, 2023).

* Corresponding author.

Investigating how and why these firms come into being offers insights not only into entrepreneurial behavior but also into broader institutional and policy dynamics.

The emergence of firms can be viewed through multiple lenses—transaction cost economics, behavioral theory, institutional theory, and the psychology of entrepreneurship. Each provides partial explanations but rarely captures the full complexity of entrepreneurial action. By integrating macroeconomic theories of coordination and micro-level models of individual intention, this paper aims to develop a more holistic framework for understanding firm birth. The study particularly emphasizes the intersection between behavioral intention—shaped by attitudes, social norms, and perceived control—and the availability of resources and competencies, as theorized in Ajzen's (1991) Theory of Planned Behavior.

This integrated approach is crucial in contemporary economies where innovation and entrepreneurial dynamism often stem from individual decision-making rather than large-scale industrial strategy. Moreover, it helps bridge the divide between abstract economic theory and the lived experiences of entrepreneurs, especially in regions with strong traditions of family business and artisanal enterprise, such as northeastern Italy. The Italian case offers unique insight into how traditional values, personal motivations, and resource constraints converge to influence firm formation (Beck & Demirgüç-Kunt, 2006).

The focus on Italian SMEs provides a valuable empirical setting where historical legacies, institutional characteristics, and entrepreneurial culture interact in unique ways. Italy's economic structure, heavily reliant on SMEs, offers fertile ground for testing theoretical models of firm emergence. This context allows for the examination of how individual-level intentions, combined with contextual enablers or constraints, translate into firm creation.

Methodologically, this study adopts a quantitative design and employs Partial Least Squares Structural Equation Modeling (PLS-SEM) to analyze data from 111 SMEs. This approach allows for the examination of complex theoretical relationships involving latent constructs and variables derived from the Theory of Planned Behaviour (Ajzen, 1991) and the Resource-Based View (Barney, 1991). By integrating these frameworks, the model captures both motivational and capability-based dimensions of entrepreneurship, thereby enriching the analytical perspective.

By synthesizing insights across disciplines and validating them through empirical testing, this paper contributes to a deeper understanding of why and how firms are born—an inquiry that remains central to both economic theory and entrepreneurial practice (Shane & Venkataraman, 2000).

The rest of the paper is organized as follows: Section 2 reviews the literature on macro- and micro-level theories of firm emergence, outlining key conceptual frameworks and defining the research hypotheses. Section 3 describes the research methodology, including data collection, sample characteristics, and the rationale for using PLS-SEM. Section 4 presents the empirical results, assessing both the measurement and structural models. Section 5 discusses the implications of the findings in light of existing theory and contextual factors. Finally, Section 6 concludes with a summary of contributions, limitations, and avenues for future research.

2. LITERATURE REVIEW AND RESEARCH HYPOTHESIS

Firms have existed in various forms across centuries, from medieval merchant guilds and artisanal family workshops to early joint-stock companies and contemporary multinational corporations. Their evolution has been closely intertwined with technological, legal, institutional, and socio-economic transformations. The modern firm—characterized by bureaucratic hierarchy, capital intensity, and managerial specialization—emerged during the Industrial Revolution, when the coordination of large-scale operations outgrew the efficiency limits of market-based contracts (Chandler, 1977). Railroads, textile mills, and steel plants exemplified enterprises where centralized planning and vertical integration became essential.

The legal infrastructure that enabled the modern corporation, including limited liability, perpetual succession, and corporate personhood, significantly lowered investment risks and catalyzed the rise of capital-intensive business models (Hansmann & Kraakman, 2000). These developments also coincided with broader institutional shifts, such as urbanization, the formation of labor markets, and the growing separation between ownership and control (Berle & Means, 1932; Zingales, 2000). Collectively, these factors positioned the firm as a stable and adaptive organizational unit for coordinating resources in increasingly complex environments.

Notably, small and medium-sized enterprises (SMEs) have consistently been at the forefront of this evolution. According to the OECD (2023), SMEs represent over 99% of all enterprises and account for 60–70% of employment in many economies. They play a vital role in innovation, regional development, and societal resilience (Beck & Demirgüç-Kunt, 2006), underscoring their importance in any theory of firm emergence.

The academic discourse surrounding firm formation has produced two major theoretical clusters: macro-level theories, which explore institutional, structural, and resource-based drivers, and micro-level theories, which investigate the behavioral, psychological, and entrepreneurial underpinnings of individual firm founders.

2.1 Macro-Level Theories of Firm Emergence

The foundational contribution of Ronald Coase in The Nature of the Firm (1937) remains essential to macro-level theorizing. Coase argued that firms arise to minimize transaction costs—those incurred in using the price mechanism, such as searching, contracting, monitoring, and enforcing agreements. When market coordination becomes inefficient relative to internal governance, firms emerge as cost-saving alternatives.

Building on Coase, the incomplete contracts theory (Grossman & Hart, 1986; Hart & Moore, 1990) highlights the inability to specify all future contingencies in contracts. In such cases, assigning residual control rights through firm ownership becomes a more efficient governance mechanism. This aligns with the property rights theory (Alchian & Demsetz, 1972), which posits that firms organize ownership and incentive structures to reduce agency costs and safeguard asset specificity.

The nexus of contracts perspective (Jensen & Meckling, 1976) further conceptualizes the firm as a web of contractual relationships, rather than a monolithic entity. This view emphasizes governance, incentive alignment, and risk-sharing among stakeholders.

Moreover, Holmström and Roberts (1998) provided critical insights into the boundaries of the firm, suggesting that firms arise not only due to transactional considerations, but also due to the need for control, coordination, and information processing within complex tasks.

From a strategic perspective, the resource-based view (RBV) (Barney, 1991) explains firm existence through their capacity to bundle and exploit valuable, rare, inimitable, and non-substitutable (VRIN) resources. Complementing this, the knowledge-based view (KBV) (Kogut & Zander, 1992) emphasizes the firm's role as a repository of tacit knowledge, routines, and organizational learning—elements that markets struggle to replicate.

Foss (1996) and Foss & Mahnke (2000) argue that firms are superior vehicles for integrating and coordinating dispersed knowledge, enabling dynamic capabilities that drive innovation and adaptation.

Lastly, evolutionary economics (Nelson & Winter, 1982) introduces a dynamic lens. Firms are seen as evolving entities shaped by routines, technological trajectories, and path dependencies. Rather than optimizing at a given point in time, firms adapt incrementally through organizational learning, variation, and selection processes—situating firm birth within broader historical and institutional ecosystems (Hodgson, 2002).

In sum, macro-level theories portray firms as institutional responses to coordination inefficiencies, uncertainty, and the need for sustained competitive advantage in complex environments.

2.2 Micro-Level Theories and Psychological Motivations

While macro theories underscore structural and institutional rationales, micro-level approaches highlight the individual-level intentions, traits, and motivations that underlie firm creation. Among these, Ajzen's (1991) Theory of Planned Behavior (TPB) stands out as a leading framework for predicting entrepreneurial intention. TPB asserts that intention is shaped by three components: attitude toward the behavior, subjective norms, and perceived behavioral control (PBC). When individuals perceive firm creation as desirable, socially supported, and personally feasible, the likelihood of entrepreneurial action increases.

Krueger et al. (2000) found that intention-based models like TPB outperform demographic and personality-based predictors of entrepreneurial behavior. Furthermore, perceived behavioral control has been shown to correlate positively with both new venture creation and performance (Kautonen et al., 2015).

Complementing TPB is the Entrepreneurial Orientation (EO) framework (Lumpkin & Dess, 1996), which captures behavioral dimensions such as risk-taking, proactiveness, autonomy, innovativeness, and competitive aggressiveness. These dimensions not only characterize the mindset of firm founders but also influence firm-level strategies and outcomes. A meta-analysis by Rauch et al. (2009) confirmed EO's consistent association with firm growth and success across contexts.

Emerging research has integrated these theories with environmental enablers. For instance, access to capital, mentorship, digital infrastructure, and supportive institutions significantly mediate the translation of intention into action (Saeed et al., 2014; Liñán & Fayolle, 2015). Shane and Venkataraman's (2000) framework further

emphasizes the opportunity-individual nexus, suggesting that entrepreneurship emerges from the interaction between opportunity recognition and individual characteristics.

Finally, psychological constructs such as self-efficacy (Bandura, 1997), need for autonomy, and locus of control (Rotter, 1966) have also been implicated in entrepreneurial intention formation. These factors underscore the role of agency, motivation, and belief systems in firm birth.

Collectively, micro-level theories reveal that firm formation is not merely an economic calculation but a behavioral and motivational process, shaped by personal goals, cognitive evaluations, and contextual conditions.

Based on the above, we formulate the following research hypothesis:

Hp1: Attitude positively influences the intention to start a new business;

Hp2: Subjective norms positively influence the intention to start a new business;

Hp3: Behavioural control (competencies and external support) positively influences the intention to start a new business;

Hp4: Behavioural control (competencies) positively influences attitude;

Hp5: Behavioural control (competencies) positively influences new product or technology development;

Hp6: New product or technology development positively influences the intention to start a new business;

3. METHODOLOGY

To complement the theoretical discussion, this study adopted a quantitative research design using a sample of 111 small and medium-sized enterprises (SMEs) located in the north-eastern Italian provinces of Verona and Vicenza. These regions were selected to ensure a relatively homogeneous sample in terms of industrial environment and accessibility to local business associations (e.g., Confindustria and API).

A structured questionnaire was distributed to SMEs in the manufacturing, distribution, and construction sectors, which together represent the majority of economic activity in the sampled area. The questionnaire covered a wide range of topics, including motivations for starting the business, perceived opportunities, resource availability, and behavioral drivers. Appendix A reports the list of questions related to firm born considered in our study. A total of 137 responses were collected, and after removing duplicates and non-eligible entries, 111 valid responses remained. SMEs included in the final sample employed between 10 and 249 workers, in line with EU definitions, and excluded entities under liquidation, with special legal forms, or lacking public financial information.

To mitigate potential biases, several procedures were implemented. Non-response bias was tested by comparing early and late respondents and found to be non-significant. Common method bias was evaluated using Harman's single-factor test, confirming that no single factor dominated the variance structure.

The data were analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM), a technique well-suited for modeling latent variables with relatively small sample sizes. For this purpose, we employed SmartPLS Software, version 4.1. PLS-SEM enabled the simultaneous evaluation of measurement and structural models, allowing for testing of complex relationships among psychological, strategic, and economic variables. All model constructs were reflective, and their reliability and validity were confirmed through standard metrics: Cronbach's alpha, composite reliability, Average Variance Extracted (AVE), and discriminant validity via HTMT ratios and the Fornell-Larcker criterion. The global model fit was also within acceptable thresholds (SRMR < 0.10, NFI > 0.7).

4. RESULTS

4.1. Results

Descriptive statistics—comprising minimum, maximum, mean, and standard deviation—for the responses provided by the SMEs in our sample are available upon reasonable request. Likewise, the Pearson correlation matrix for the study variables can be provided on request. In general, correlations among the variables were below 0.50, with only a few instances approaching 0.70.

4.2. Measurement model assessment

In our model, we assume that all latent variables are antecedents of their indicators and, therefore, our constructs are reflective, reason why we use the calculation procedure based on the consistent PLS-SEM algorithm. For the

initial assessment of the model, we examined construct reliability, indicator reliability, convergent validity, and discriminant validity. Detailed results are available upon request.

As a first step, we checked for collinearity by inspecting the variance inflation factors (VIF) values that remained under 3 for all variables (the max value was 1.949).

The reliability of the constructs has been evaluated using Chronbach's alpha, composite reliability rho_a, and composite reliability rho_c. All values were greater than 0.7, being 0.822 for Cronbach's alpha, 0.842 for composite reliability Rho_a, and 0.829 for composite reliability Rho_c, thus confirming construct reliability.

Regarding indicator reliability, all factor loadings exceed the 0.728 value, overpassing or being very close to the suggested threshold of 0.7.

Convergent validity was evaluated using Average Variance Extracted (AVE), which should have been greater than 0.5. Attitude constructs obtained a value of 0.710, confirming the validity of the proposed reflective scales' internal consistency.

To check the discriminant validity, we inspected the heterotrait-monotrait (HTMT) ratios of the correlations and their 95% one-sided bootstrap confidence intervals. All HTMT values were below 0.90 (the maximum value was 0.602). In addition, we also used the Fornell-Larcker criterion to support discriminant validity, checking that the correlation between each pair of constructs did not exceed the square root of the AVE of each of the constructs. For this reason, discriminant validity can be confirmed. As a last step, we evaluated the global quality of the model using the standardized root mean square of residuals (SRMR), which was 0.034 for the saturated model and 0.060 for the estimated model, therefore an acceptable fit considering the recommended threshold of 0.10. As additional criterion for the validity of the model we evaluated NFI that was 0.945 and 0.915 for saturated and estimated model respectively.

4.3. Structural model assessment and path analysis

As a first step, we checked for collinearity in the structural model, inspecting variance inflation factors (VIF) values for all variables and constructs. None of them exceeded the threshold of 5 also being lower to the more conservative threshold of 3 (values ranged from 1 to 1.949) (Samueli et al. 2019); therefore, we conclude that collinearity is not at critical levels.

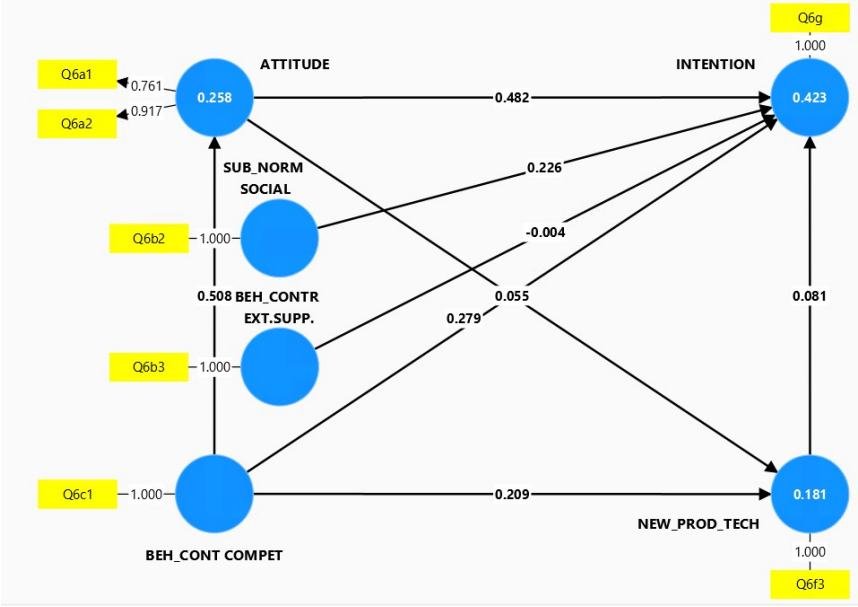

Fig. 1. Model path diagram

Second, we assessed the significance and relevance of the structural model analyzing direct effects and path coefficients. The structural model is presented in Figure 1. A bootstrapping procedure with 20,000 resamples was then used to calculate p-values and f-squared effects. We run bootstrapping using one-tailed p-values at 5%, as we assumed in our hypotheses that coefficients had a positive or negative sign (Kock, 2015). The results with path coefficients and p values are available upon request.

The findings do confirm ($p<0.05$) that the intention to set-up a new firms is positively influenced ($p<0.001$) by ATTITUDE, and by SUBJECTIVE NORMS ($p<0.05$), whereas BEHAVIOURAL CONTROL (both via external support and via competencies) is statistically NOT significant. In our model Behavioural control (competencies) slightly ($p<0.1$) influences positively NEW_PROD_TECH_DEV, but this last does NOT influence INTENTION. Conversely Behavioural control (competencies) positively influences ATTITUDE. Therefore Hp1, Hp2, and Hp4 are confirmed, Hp5 only slightly whereas Hp3 and Hp6 are rejected.

5. DISCUSSION

The analysis of the results obtained from the selected sample allows us to partially confirm the validity of the Theory of Planned Behavior (TPB) in predicting the intention to establish a new firm. While Attitude and Subjective Norms are both statistically significant and align with theoretical expectations, Perceived Behavioral Control—as conceptualized through external support and access to competencies—did not show consistent statistical significance. This suggests that while cognitive evaluations of desirability and social expectations may serve as strong precursors to entrepreneurial intention, the feasibility dimension, typically captured through behavioral control, is less predictive in this specific SME context.

This outcome echoes previous research indicating that entrepreneurial intentions often stem more from intrinsic motivations and social context than from the availability of external enablers (Krueger et al., 2000; Liñán & Fayolle, 2015). In the context of SMEs, the intention to establish a new firm appears to be driven primarily by the individual's aspiration to become an entrepreneur, reinforced by perceived societal approval and normative pressure. This supports the idea that firm birth is not merely a function of environmental or institutional readiness, but of personal agency and belief systems, consistent with Ajzen's (1991) model.

Conversely, the lack of significance for perceived behavioral control (external support) suggests a limited role of institutional and financial scaffolding—such as banks or public bodies—in shaping entrepreneurial intention within the studied sample. This may be particularly true in cases where the firm's origin is rooted in earlier periods or in more traditional contexts, in which businesses were established through self-reliance, family support, or informal networks rather than through structured ecosystems of support. This finding may also reflect the historical embeddedness of many SMEs in the Italian context, where entrepreneurial ventures have often emerged from individual resilience and contextual necessity, rather than from well-developed startup infrastructures.

Nonetheless, competencies—interpreted through the lens of the resource-based view (RBV)—show a meaningful, though indirect, influence on firm formation. Specifically, competencies are positively linked to both the development of new products and technologies, as well as to entrepreneurial attitude. This suggests that internal capabilities function as antecedents to the strategic mindset required for launching a business, even if their direct impact on intention is not strongly supported in the structural model. The results support the notion that competence-based advantages, particularly those that are tacit and path-dependent, enhance the capacity for innovation, which in turn reinforces the founder's confidence and readiness to start a new venture (Barney, 1991; Kogut & Zander, 1992).

This relationship also mirrors arguments from evolutionary economics, particularly the work of Nelson and Winter (1982), who assert that firms emerge and evolve based on routines and competencies that shape behavior over time. The finding that competencies influence product and technological development but not directly intention may indicate that innovation capacity is a necessary but not sufficient condition for entrepreneurship—highlighting a potential gap between capability and action. This reinforces the need for more integrative models that link micro-level intention with macro-level innovation processes.

Moreover, the rejection of some hypotheses (Hp3 and Hp6) indicates that not all theoretical assumptions hold uniformly across contexts. For example, the weak or absent impact of external behavioral control (support) may reflect structural or cultural specificities in the Italian SME environment, where entrepreneurial ecosystems may be underdeveloped or perceived as ineffective. This nuance emphasizes the importance of context-sensitive entrepreneurship research, as advocated by recent institutional and behavioral studies (Saeed et al., 2014; Shane & Venkataraman, 2000).

In any case, the results reveal both strengths and limitations of the study, highlighting the need for further investigation. Indeed, we aim to refine our analysis by identifying and assessing the potential impact of outliers that may have skewed the results. Additionally, expanding the sample size to include a broader range of firms—across different geographic areas or industry sectors—could provide a more robust dataset and improve external validity. Future research should also extend the data collection period, which may uncover temporal dynamics or longitudinal patterns not visible in a cross-sectional design.

Overall, the findings contribute to a richer understanding of firm emergence by integrating psychological, strategic, and institutional perspectives. They underscore that firms are not merely efficiency-driven structures, as posited by transaction cost economics, but are also behavioral expressions of entrepreneurial intention and resource mobilization, deeply embedded in individual motivations and contextual realities.

6. CONCLUSION, CONTRIBUTIONS, LIMITATIONS AND FUTURE DIRECTIONS

This study set out to explore the underlying motivations behind firm formation, drawing on a combination of theoretical frameworks including the Theory of Planned Behavior (TPB), the Resource-Based View (RBV), and institutional economic perspectives. By applying Partial Least Squares Structural Equation Modeling (PLS-SEM) to a sample of 111 Italian SMEs, we examined how psychological drivers, strategic competencies, and external support structures influence the intention to start a business.

Our findings make several contributions to existing theory. First, they provide partial empirical support for the TPB by confirming that attitude and subjective norms are significant predictors of entrepreneurial intention. This reinforces the importance of individual-level cognitive and social factors in the entrepreneurial decision-making process. Second, the study extends the RBV by highlighting the indirect role of competencies—though not always statistically significant in influencing intention, competencies appear to foster innovation and attitude, which are key antecedents to firm creation. These findings suggest the need to move beyond linear interpretations of theory and toward more nuanced, interaction-based models that consider both internal and external enablers. Lastly, by revealing the limited influence of perceived behavioral control (external support), the study invites a reevaluation of the assumed centrality of institutional infrastructure in early-stage entrepreneurship—at least within certain historical or cultural contexts.

From a practical standpoint, the results suggest that policymakers and support organizations should focus not only on financial aid or institutional programs but also on shaping the perceived desirability and social legitimacy of entrepreneurship. Entrepreneurial education, role models, and peer networks may be more effective in stimulating firm creation than formal support structures alone. In particular, fostering entrepreneurial attitudes and innovation orientation in potential founders could have a meaningful impact on start-up dynamics, especially among SMEs. Furthermore, while competencies may not always directly drive intention, they are critical in enabling the execution of innovative strategies once the firm is established, underlining the value of early investment in capability development.

This study is not without its limitations. The relatively small sample size, while appropriate for PLS-SEM, constrains the generalizability of the findings. The data were collected from a specific geographical and cultural context (northeastern Italy), which may limit applicability to other regions or countries. The cross-sectional design also prevents us from capturing how intentions evolve over time or how they translate into actual firm creation and survival. Additionally, while constructs were partially measured with validated scales, the complexity of entrepreneurial motivation may require the inclusion of other psychological, contextual, or temporal variables not covered in this model.

Future studies could adopt a longitudinal approach to track the transition from intention to action, thereby bridging the gap between entrepreneurial cognition and behavior. Expanding the sample to include different regions, industries, and cultural contexts would also enhance external validity and allow for cross-country comparisons. Further investigation could also focus on moderating variables such as gender, education, prior experience, or access to digital technologies, which may influence the strength and direction of the relationships identified here. Finally, integrating qualitative methods—such as interviews or case studies—could provide deeper insight into the lived experiences and narratives that underlie entrepreneurial intention and firm birth.

Ultimately, the firm is born not only as an efficient response to market failures or coordination problems, but as a product of human intention, embedded knowledge, and social context. Understanding the genesis of the firm therefore requires a multidisciplinary lens—one that appreciates the interplay between institutional structures, strategic resources, and the personal aspirations of entrepreneurs. As the entrepreneurial landscape continues to evolve amid digitalization and socio-economic shifts, revisiting the question of why firms are born remains as relevant as ever.

Funding/Acknowledgements

This paper has received no external funding.

Conflict of Interest

The authors declare NO CONFLICT of INTEREST.

References

[1] Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2), 179–211.
[2] Alchian, A. A., & Demsetz, H. (1972). Production, information costs, and economic organization. American Economic Review, 62(5), 777–795.
[3] Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.
[4] Barney, J. B. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99–120.
[5] Chandler, A. D. (1977). The visible hand: The managerial revolution in American business. Harvard University Press.
[6] Coase, R. H. (1937). The nature of the firm. Economica, 4(16), 386–405.
[7] Foss, N. J., & Mahnke, V. (2000). Competence and governance perspectives on the boundaries of the firm. In N. J. Foss & V. Mahnke (Eds.), Competence, governance, and entrepreneurship (pp. 25–48). Oxford University Press.
[8] Grossman, S. J., & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. Journal of Political Economy, 94(4), 691–719.
[9] Hart, O., & Moore, J. (1990). Property rights and the nature of the firm. Journal of Political Economy, 98(6), 1119–1158.
[10] Hodgson, G. M. (2002). The legal nature of the firm and the myth of the firm-market hybrid. International Journal of the Economics of Business, 9(1), 37–60.
[11] Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(4), 305–360.
[12] Kautonen, T., van Gelderen, M., & Fink, M. (2015). Robustness of the Theory of Planned Behavior in predicting entrepreneurial intentions and actions. Entrepreneurship Theory and Practice, 39(3), 655–674.
[13] Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. Organization Science, 3(3), 383–397.
[14] Kock, N. (2015). One-tailed or two-tailed P values in PLS-SEM? International Journal of e-Collaboration (IJeC), 11(2), 1–7.
[15] Liñán, F., & Fayolle, A. (2015). A systematic literature review on entrepreneurial intentions: Citation, thematic analyses, and research agenda. International Entrepreneurship and Management Journal, 11(4), 907–933.
[16] Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the entrepreneurial orientation construct and linking it to performance. Academy of Management Review, 21(1), 135–172.
[17] Nelson, R. R., & Winter, S. G. (1982). An evolutionary theory of economic change. Harvard University Press.
[18] OECD (2023). OECD SME and Entrepreneurship Outlook 2023. OECD Publishing, Paris.
[19] https://doi.org/10.1787/33d62f15-en
[20] Rauch, A., Wiklund, J., Lumpkin, G. T., & Frese, M. (2009). Entrepreneurial orientation and business performance: An assessment of past research and suggestions for the future. Entrepreneurship Theory and Practice, 33(3), 761–787.
[21] Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs: General and Applied, 80(1), 1–28.
[22] Saeed, S., Yousafzai, S. Y., & Engelen, A. (2014). On cultural and macroeconomic contingencies of the entrepreneurial orientation–performance relationship. Entrepreneurship Theory and Practice, 38(2), 255–290.
[23] Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. Academy of Management Review, 25(1), 217–226.

Appendix A. Questionnaire submitted to companies. List of questions

How much do you think the following elements may have influenced the BIRTH of the COMPANY (FOUNDATION)? (Please reply on a Likert scale from 1 to 7. where: 1=not at all; 7=extremely)

(Q6.a1) Desire to become an entrepreneur
(Q6.a2) Desire for "self-employment"
(Q6.b2) Strong social value of entrepreneurship
(Q6.b3) Support from banks, the state, and institutions for entrepreneurship
(Q6.c1) Presence of product/service-specific skills
(Q6.f3) The company introduced NEW PRODUCTS or TECHNOLOGIES
(Q6.g) The INTENTION to START the company was HIGH

The Evolution of the OECD Corporate Governance Principles - A Comparative Analysis (1999–2023).

Paolo ROFFIA [a*]

[a] *University of Verona, Department of Management, Via Cantarane 24, Verona, 37124, Italy,*
paolo.roffia@univr.it[b]

Abstract

This paper explores the evolution of the OECD Corporate Governance Principles from 1999 to 2023, assessing their development and influence on global corporate governance systems. As a globally recognized benchmark, the OECD Principles have shaped governance practices in response to major financial crises, technological progress, and shifting economic landscapes. Over time, they have been widely adopted and adapted across jurisdictions, informing national governance codes and corporate strategies in diverse legal and economic contexts. Using a content analysis approach, this study identifies significant revisions, emerging themes, and their broader implications. The findings indicate a gradual shift from a shareholder-centric model toward a more inclusive framework that emphasizes sustainability, stakeholder engagement, and digital transformation. Across editions, there is a consistent strengthening of board accountability, transparency, and risk management practices. The 2023 revision, in particular, reflects contemporary priorities by incorporating ESG factors, corporate resilience, and digital governance. A comparative review of the four editions highlights the growing emphasis on corporate responsibility and adaptability in the face of economic and technological change. The study underscores the importance of ongoing refinement of governance principles to meet evolving global challenges and to maintain corporate trust and stability. By offering a detailed examination of the OECD Principles' progression, this research enhances our understanding of governance dynamics and provides meaningful insights for policymakers, corporate leaders, and investors operating in today's rapidly changing environment.

Keywords: OECD, corporate Governance, OECD Principles of CG, Institutional theory

1. INTRODUCTION

The evolution of corporate governance principles is deeply intertwined with broader societal changes, economic transformations, and technological advancements. Indeed, Corporate governance frameworks are not static; they evolve in response to regulatory shifts and societal expectations (Tricker, 2015). Over the past few decades, globalization, financial crises, and the digital revolution have significantly reshaped corporate governance structures. These developments have necessitated regulatory adaptations to ensure market stability, enhance transparency, and promote sustainable business practices.

In the late 20th century, corporate governance was primarily focused on maximizing shareholder value, with limited attention to broader societal concerns. However, the 21st century has seen a paradigm shift toward a more inclusive governance approach, emphasizing environmental, social, and governance (ESG) considerations, stakeholder rights,

* Corresponding author.

and technological governance. The rise of digital economies, increasing environmental awareness, and the growing demand for corporate accountability have further propelled governance reforms.

The OECD Corporate Governance Principles, first introduced in 1999, have played a pivotal role in shaping global governance frameworks. These principles were revised in 2004, 2015, and most recently in 2023, reflecting ongoing economic, technological, and regulatory developments. Each revision aimed to address emerging challenges, including financial crises, corporate fraud, sustainability concerns, and digital transformation.

This paper aims to provide a comparative analysis of the OECD Corporate Governance Principles across their four editions, highlighting key changes and their implications. By assessing the evolution of these principles, the study contributes to corporate governance literature by identifying major governance trends and evaluating their impact on businesses, investors, and policymakers.

The article is structured as follows: Section 2 provides a comprehensive literature review on the OECD Corporate Governance Principles, their purpose, and their impact. Section 3 outlines the research methodology used for analyzing the evolution of the principles. Section 4 presents key findings, comparing the different editions of the principles and discussing their implications. Finally, Section 5 concludes with insights into future corporate governance trends and policy recommendations.

2. THEORETICAL BACKGROUND AND LITERATURE REVIEW

The OECD Corporate Governance Principles were first introduced in 1999 by the Organization for Economic Co-operation and Development (OECD) to establish internationally recognized guidelines on corporate governance. These principles serve as a foundational framework for policymakers, investors, and corporations, aiming to enhance transparency, accountability, and efficiency in corporate governance structures. Over time, the principles have undergone revisions to address emerging economic, technological, and regulatory challenges, leading to updates in 2004, 2015, and 2023.

Institutional theory provides a useful lens for understanding how corporate governance evolves in response to external pressures. According to DiMaggio and Powell (1983), institutional isomorphism explains why governance standards converge across countries due to regulatory, normative, and cognitive pressures. Institutional pressures often lead organizations to adopt similar structures and practices, regardless of their efficiency (Scott, 2014). Similarly, Meyer and Rowan (1977) argue that governance frameworks often reflect symbolic conformity to legitimacy expectations rather than purely efficiency-driven changes. These theoretical perspectives help contextualize the evolving role of corporate governance principles in a globalized and interconnected economy.

The OECD Corporate Governance Principles provide a framework that enhances market confidence, protects investor rights, and promotes fair and efficient financial markets. The principles focus on core governance issues, including board responsibilities, shareholder and stakeholder rights, transparency, and ethical business conduct. Their global influence is evident in corporate governance codes across multiple jurisdictions, including the European Union, the United States, and Asia.

The principles have been widely used as a benchmark for assessing governance frameworks across countries. They have guided corporate governance reforms globally, leading to improved standards in board accountability, shareholder rights, and corporate transparency. Their role in shaping financial market regulations became particularly prominent following the 2008 financial crisis, where governance weaknesses contributed to economic instability.

Several scholars have analyzed the impact and effectiveness of the OECD Corporate Governance Principles.

Indeed, in 2004 Clarke examined how the 1999 principles influenced corporate governance frameworks in Asia, highlighting their role in stabilizing financial markets following the 1997 Asian financial crisis.

Aguilera & Jackson (2010) explored cross-country variations in the implementation of the OECD principles, noting that differences in legal and institutional frameworks affect their enforcement and effectiveness.

In 2013 Mallin discussed the relationship between OECD principles and firm performance, demonstrating that companies adhering to strong governance frameworks tend to exhibit higher financial stability and investor confidence.

García-Castro et al. (2017) studied the impact of sustainability integration within corporate governance reforms, emphasizing the increasing role of ESG factors in recent updates to the OECD principles.

Despite their widespread adoption, the OECD principles have faced criticism. Some researchers argue that the principles are too broad, allowing varying interpretations that limit enforceability (Bebchuk & Weisbach, 2010). Others contend that the principles overemphasize shareholder rights, while inadequately addressing the role of stakeholders (Freeman et al., 2020). Additionally, their applicability in emerging economies remains debated, as weaker regulatory environments pose challenges to enforcement (La Porta et al., 1999).

Considering these discussions, this study explores the following hypothesis:

Hp1: Corporate governance codes are dynamic and have evolved in response to economic, social, and technological changes, reflecting the shifting expectations of stakeholders.

3. METHODOLOGY, RESEARCH APPROACH, DATA COLLECTION

This study employs a qualitative content analysis methodology to compare the four editions of the OECD Corporate Governance Principles (1999, 2004, 2015, and 2023). Content analysis is an established research approach used to systematically analyze textual data, enabling the identification of patterns, shifts in emphasis, and thematic changes across different editions of the principles (Krippendorff, 2018).

3.1 Research Approach

The study follows a systematic document analysis, examining OECD publications, related policy documents, and academic literature to identify key modifications, thematic trends, and governance priorities. Each edition of the OECD principles is reviewed against major economic, financial, and regulatory events influencing their development (Bowen, 2009).

3.2 Data Collection

The primary data sources for this research include:

- Official OECD reports and publications on corporate governance (OECD, 1999; OECD, 2004; OECD, 2015; OECD, 2023).
- Policy documents from international regulatory bodies, such as the G20, World Bank, and Financial Stability Board, which have referenced or adopted the OECD principles (FSB, 2016).
- Academic studies and empirical research on the effectiveness and impact of the principles (Aguilera & Jackson, 2010; Clarke, 2004; La Porta et al., 1999).
- Industry reports and corporate governance codes from various jurisdictions that have implemented OECD guidelines (Mallin, 2013).

3.3 Analytical Framework

The content analysis follows a structured framework to examine changes across the four editions of the OECD principles. The analysis is categorized based on the following themes:

- Board Responsibilities: Evolution in governance expectations for board independence, composition, and risk oversight (Bebchuk & Weisbach, 2010).
- Shareholder Rights and Protection: Changes in policies affecting shareholder engagement, voting mechanisms, and legal protections (Freeman et al., 2020).
- Transparency and Disclosure: Developments in financial reporting, audit requirements, and information dissemination (García-Castro et al., 2017).
- Sustainability and ESG Factors: Integration of environmental, social, and governance (ESG) considerations over time (OECD, 2023).
- Digitalization and Technological Governance: Adaptation of corporate governance standards to emerging technologies, cybersecurity, and data governance (FSB, 2017).
- Stakeholder Engagement: Expansion of governance frameworks to include non-shareholder stakeholders such as employees, customers, and communities (Mallin, 2013).

3.5 Comparative Analysis Method

A comparative analysis is conducted by systematically coding and categorizing governance principles across the four editions. Each principle is analyzed in terms of:

- Consistency and Continuity – Identifying principles that have remained largely unchanged across editions (OECD, 1999; 2004; 2015; 2023).
- Modifications and Expansions – Tracking principles that have been modified or expanded to address emerging governance challenges (Aguilera & Jackson, 2010).
- New Additions – Highlighting governance elements introduced in response to contemporary issues such as sustainability and digitalization (García-Castro et al., 2017).
- Regulatory and Economic Context – Assessing the influence of major financial crises, corporate scandals, and global governance trends on the revisions (La Porta et al., 1999).

3.6 Validity and Reliability

To ensure the reliability of the content analysis, multiple sources are cross-referenced, and findings are validated against secondary research, such as peer-reviewed articles and corporate governance reports. Expert opinions from governance practitioners and policymakers are also considered to support the interpretation of findings (Krippendorff, 2018).

The methodology allows for a structured comparison of the evolution of OECD Corporate Governance Principles, highlighting key governance trends and their implications for corporate governance globally.

4. FINDINGS AND DISCUSSION

4.1. The OECD Corporate governance principes and their evolution

The OECD Corporate Governance Principles provide a framework for corporate governance aimed at improving market confidence, financial stability, and sustainable economic development.

The first edition, published in 1999, consisted of approximately 46 pages structured into five key principles, with a total of around 50 paragraphs detailing best practices and recommendations. These principles were developed through a collaborative process involving OECD member states, including the United States, the United Kingdom, France, Germany, and Japan, alongside representatives from international financial institutions, corporate governance experts, and business organizations.

Initially, the principles were widely adopted by OECD countries, serving as a benchmark for national corporate governance codes in economies such as Canada, Australia, and European Union member states. Over time, their influence extended beyond OECD members, with emerging economies like Brazil, India, and South Africa incorporating aspects of the framework into their regulatory reforms. This broad application underscores the OECD Principles' role as a global standard for corporate governance practices.

The principles included in the first edition emphasized six core areas:

- Ensuring the Basis for an Effective Corporate Governance Framework – Establishing clear regulatory and legal structures that promote good governance practices.
- The Rights and Equitable Treatment of Shareholders and Key Ownership Functions – Protecting investor rights and ensuring fair treatment of all shareholders.
- Institutional Investors, Stock Markets, and Other Intermediaries – Promoting transparency and accountability in market transactions and corporate oversight.
- The Role of Stakeholders in Corporate Governance – Recognizing stakeholder rights and encouraging cooperation between businesses and broader societal interests.

- Disclosure and Transparency – Strengthening reporting obligations, financial disclosures, and corporate accountability.

- The Responsibilities of the Board – Enhancing board independence, risk management, and ethical leadership.

These principles have evolved over time in response to global economic shifts, financial crises, and emerging challenges. Over two decades, the events and challenges that good corporate governance has faced have been diverse: tackling fraud and corruption (Bebchuk & Weisbach, 2010), the increasing importance of all stakeholders and not just shareholders (Freeman et al., 2020), the protection of minority investors (La Porta et al., 1999), resilience as the ability to withstand and react to crises (Aguilera & Jackson, 2010), digitalization (FSB, 2017), business sustainability (García-Castro et al., 2017), and environmental protection (OECD, 2023). Corporate scandals, while detrimental in the short term, often act as valuable catalysts for governance reform by revealing systemic weaknesses and prompting the recalibration of oversight mechanisms (Coffee, 2007). Each of these aspects has influenced the continuous development and adaptation of corporate governance principles to meet the evolving expectations of market participants and regulators alike.

The analysis of the OECD Corporate Governance Principles from 1999 to 2023 reveals several significant trends and changes in corporate governance frameworks. These findings demonstrate the evolving priorities of global corporate governance and highlight the increasing complexity of regulatory expectations. The key findings are as follows:

- Shift from Shareholder to Stakeholder Governance: Early editions of the principles focused primarily on shareholder rights and market efficiency. However, the 2023 edition places a strong emphasis on stakeholder engagement, incorporating environmental, social, and governance (ESG) factors as central to corporate governance policies. Stakeholders are identified by their interests in the corporation, regardless of whether the corporation has any corresponding functional interest in them (Donaldson & Preston, 1995).

- Strengthened Board Responsibilities and Accountability: The principles have progressively reinforced the role of boards in risk management, ethics, and sustainability, requiring greater transparency and independence. Boards play a critical role in aligning managerial incentives with shareholder interests and monitoring firm performance (Adams et al, 2010).

- Integration of Digitalization and Technological Governance: The 2023 edition explicitly addresses the risks and opportunities posed by digitalization, including cybersecurity, AI governance, and data protection. Corporate governance must now reckon with the implications of datafication, algorithmic control, and surveillance (Zuboff, 2019).

- Greater Emphasis on ESG Considerations: Sustainability and long-term value creation have become fundamental principles in the latest edition, aligning governance frameworks with global climate and social responsibility goals. Firms that embed sustainability into their core strategy demonstrate superior performance in long-term value creation (Eccles et al., 2014).

- Improved Regulatory Frameworks and Compliance Standards: Over the years, corporate governance principles have evolved to include stricter compliance requirements, disclosures, and enforcement mechanisms to prevent financial fraud and misconduct. In response to financial misconduct and governance failures, regulatory frameworks have increasingly emphasized transparency, accountability, and robust compliance mechanisms (Solomon, 2017).

Table 1: OECD Corporate Governance Principes – Comparative analysis

Edition	Key Focus Areas	Key Revisions & Updates	Pages
1999	Shareholder rights, market efficiency	Established foundation for corporate governance, emphasizing shareholder rights, equitable treatment, and transparency. Focused on investor protection and fraud prevention	46

2004	Board responsibilities, disclosure requirements	Strengthened board responsibilities, disclosure requirements, and responses to corporate scandals (e.g., Enron, WorldCom). Introduced director independence and improved risk management	69
2015	Risk management, executive compensation	Addressed post-financial crisis issues, enhancing risk management, executive compensation regulations, and long-term value creation. Expanded focus on institutional investors and their role.	60
2023	ESG integration, digitalization, stakeholder engagement	Introduced provisions for sustainability, digitalization, and an expanded role of stakeholders. Integrated ESG factors into governance structures, addressing climate change and digital transformation.	53

Source: Our elaboration

4.2. Discussion

Beyond these core changes, the discussion also evaluates how regulatory frameworks and market structures have influenced the implementation of the OECD principles. Countries with well-developed financial markets and strong institutional frameworks have adopted these principles more effectively, while developing economies often face implementation challenges due to weaker enforcement mechanisms and regulatory capacity. Indeed, institutional diversity continues to shape national governance systems, challenging convergence theories (Jackson & Deeg, 2008). Weak legal and regulatory institutions often necessitate alternative governance structures in emerging markets (Khanna & Palepu, 2000).

Additionally, the increasing complexity of corporate structures has necessitated stronger internal governance mechanisms, prompting firms to adopt more rigorous board oversight and compliance measures. The role of institutional investors has also evolved, with a growing emphasis on active ownership and responsible investment strategies aligning with ESG considerations.

In light of the results obtained and the previous discussion, Hp1 can be confirmed, meaning that corporate governance codes are dynamic and have evolved in response to economic, social, and technological changes, reflecting the shifting expectations of stakeholders.

Despite the advancements made, challenges remain in ensuring uniform application of the principles across jurisdictions. The lack of a standardized enforcement mechanism limits their impact in some regions, and disparities in regulatory interpretation can lead to inconsistencies in governance practices. Moving forward, greater international collaboration and harmonization of governance standards may be necessary to enhance the effectiveness of these principles in fostering global corporate stability and resilience.

5. CONCLUSION

The OECD Corporate Governance Principles have undergone significant changes to remain relevant in an evolving business environment. This study's content analysis reveals key trends, highlighting the growing emphasis on sustainability, digitalization, and broader stakeholder considerations. The increasing integration of ESG principles reflects the recognition that corporate governance must adapt to address global economic and social challenges. As digital technologies continue to transform corporate decision-making processes, boards must prioritize cybersecurity, data governance, and AI ethics in their governance frameworks.

While the OECD principles provide a strong foundation for corporate governance, their effectiveness depends on national implementation and enforcement. Future research could further explore the effectiveness of these principles in practice and their impact across different regions. Additionally, comparative studies could examine how different jurisdictions interpret and enforce these governance principles, shedding light on best practices and areas for improvement.

Despite some limitations due to the methodology and the lack of empirical data, our study has several implications for policymakers, professionals and investors.

Policymakers must ensure that governance regulations remain adaptable to emerging challenges, particularly in digital security and sustainability. Businesses must integrate ESG considerations into corporate strategy to align with evolving stakeholder expectations. Investors must leverage governance assessments to guide responsible investment decisions and risk management.

We plan to expand this work with an analysis of corporate governance recommendations issued by other organizations, as well as to integrate the analysis with the corporate governance codes adopted in individual countries.

This discussion underscores the dynamic nature of corporate governance and the need for ongoing refinement of governance frameworks to address emerging challenges and opportunities.

Funding/Acknowledgements

This paper has received no external funding.

Conflict of Interest

The authors declare NO CONFLICT of INTEREST.

References

[1] Adams, R. B., Hermalin, B. E., & Weisbach, M. S. (2010). The role of boards of directors in corporate governance: A conceptual framework and survey. Journal of Economic Literature, 48(1), 58–107.
[2] Aguilera, R. V., & Jackson, G. (2010). Comparative and international corporate governance. Academy of Management Annals, 4(1), 485-556.
[3] Beattie, V., & Smith, S. J. (2012). Value creation and business models: Refocusing the intellectual capital debate. The British Accounting Review, 45(4), 243-254.
[4] Bebchuk, L. A., & Weisbach, M. S. (2010). The state of corporate governance research. The Review of Financial Studies, 23(3), 939-961.
[5] Clarke, T. (2004). Theories of corporate governance: The philosophical foundations of corporate governance. Routledge.
[6] Coffee, J. C. (2007). Gatekeepers: The professions and corporate governance. Oxford University Press.
[7] DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147-160.
[8] Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. Academy of Management Review, 20(1), 65–91.
[9] Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. Management Science, 60(11), 2835–2857.
[10] Financial Stability Board (FSB). (2017). Summary Report on Financial Sector Cybersecurity Regulations, Guidance and Supervisory Practices. Available at: https://www.fsb.org/2017/10/summary-report-on-financial-sector-cybersecurity-regulations-guidance-and-supervisory-practices/. Accessed 15/03/2025.
[11] Freeman, R. E., Harrison, J. S., & Zyglidopoulos, S. (2020). Stakeholder theory: Concepts and strategies. Cambridge University Press.
[12] García-Castro, R., Ariño, M. A., & Canela, M. A. (2017). Corporate social responsibility, governance, and shareholders' value: An analysis of ESG practices. Journal of Business Ethics, 144(4), 707-721.
[13] Jackson, G., & Deeg, R. (2008). Comparing capitalisms: Understanding institutional diversity and its implications for international business. Journal of International Business Studies, 39(4), 540-561.
[14] Khanna, T., & Palepu, K. (2000). The future of business groups in emerging markets: Long-run evidence from Chile. Academy of Management Journal, 43(3), 268–285.
[15] Krippendorff, K. (2018). Content analysis: An introduction to its methodology. Sage publications
[16] La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1999). Corporate ownership around the world. The Journal of Finance, 54(2), 471-517.
[17] Mallin, C. A. (2013). Corporate governance. Oxford University Press.
[18] Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. American Journal of Sociology, 83(2), 340-363.
[19] OECD. (1999). OECD Principles of Corporate Governance. Organisation for Economic Co-operation and Development.
[20] OECD. (2004). OECD Principles of Corporate Governance. Organisation for Economic Co-operation and Development.
[21] OECD. (2015). G20/OECD Principles of Corporate Governance. Organisation for Economic Co-operation and Development.
[22] OECD. (2023). OECD Principles of Corporate Governance. Organisation for Economic Co-operation and Development.
[23] Scott, W. R. (2014). Institutions and organizations: Ideas, interests, and identities. Sage publications.
[24] Solomon, J. (2017). Corporate Governance and Accountability. Wiley.
[25] Tricker, B. (2015). Corporate Governance: Principles, Policies, and Practices. Oxford University Press.
[26] Zuboff, S. (2019). The age of surveillance capitalism. Profile Books.

Two Sides of Corporate Performance: Profitability Financial Ratios versus Sector-Specific KPIs

Paolo ROFFIA [a*]

[a] *University of Verona, Department of Management, Via Cantarane 24, Verona, 37124, Italy,*
paolo.roffia@univr.it[b]

Abstract

Financial ratios such as Return on Assets (ROA), Return on Sales (ROS), and Return on Equity (ROE) are widely employed to assess corporate financial performance, offering standardized metrics for cross-sector benchmarking and supporting strategic decision-making. However, these ratios, being broad and heavily reliant on financial accounting data, may not adequately capture performance at the operational, process, or activity level. Key Performance Indicators (KPIs) serve as a valuable complement to financial ratios by providing more detailed and context-specific insights. Unlike standardized financial ratios, KPIs are flexible and can be tailored to monitor business-critical areas and align with strategic objectives and Critical Success Factors (CSFs). As a result, they are commonly used as tools for performance monitoring and managerial control. In several industries—such as railways, insurance, airlines, retail, and telecommunications—sector-specific KPIs (e.g., cost per kilometer, load factor, or sales per square meter) have long played a pivotal role in driving operational efficiency and strategic alignment. This study investigates whether financial ratios and sector KPIs operate as complementary indicators of performance by analyzing a sample of large Italian companies in industries with established KPI traditions. Through an empirical analysis of ROA, ROS, and sector-specific KPIs for selected firms over the 2012–2013 period, the study identifies a strong correlation between KPI rankings and financial performance. Firms performing well in sector-specific KPIs also tend to exhibit superior financial ratios, suggesting that KPIs may serve as leading indicators of financial success in the medium term.

Keywords: KPI, financial performance, performance management, sector KPI

1. INTRODUCTION

The role of financial ratios (FRs) and Key Performance Indicators (KPIs) in performance measurement has been extensively studied in corporate finance and strategic management. Over the decades, financial ratios have established themselves as essential tools for assessing corporate financial health, efficiency, and profitability. Financial ratios such as Return on Assets (ROA), Return on Sales (ROS), and Return on Equity (ROE) have long been fundamental tools for evaluating a company's financial health and operational efficiency (Bauwhede, 2009). These indicators provide quantitative insights into profitability, liquidity, activity, and leverage, enabling comparisons across firms and industries (Horrigan, 1968). However, financial ratios alone are not always sufficient to capture the full complexity of a firm's performance, as they are heavily reliant on financial accounting processes and may not reflect real-time operational effectiveness (Kaplan & Norton, 1992). Since financial ratios are backward-looking and shaped by

* Corresponding author.

accounting policies, they may not fully account for the operational and strategic factors influencing a firm's performance.

To address these limitations, Key Performance Indicators (KPIs) have gained prominence as complementary tools, offering a more dynamic and sector-specific view of performance (Parmenter, 2015). Unlike standardized financial ratios, KPIs can be tailored to industry needs, reflecting critical success factors (CSFs) and strategic objectives (Rockart, 1976). KPIs provide forward-looking insights and offer a real-time perspective on operational performance, making them valuable for both short-term tactical adjustments and long-term strategic planning (Mtau & Rahul, 2024). For example, the load factor (%) is a crucial KPI for the airline industry, sales revenue per square meter is a key performance measure in retail, and ARPU (Average Revenue Per User) is widely used in telecommunications (Cento, 2009; Grewal et al., 2017; Gerpott, Rams, & Schindler, 2001).

Sector-specific KPIs have been extensively analyzed in industry studies, showing their importance in benchmarking and strategic planning. Eccles (1991) emphasized that KPIs provide operational insights that financial ratios may overlook, particularly in industries where real-time decision-making is critical. Marr (2012) further highlighted that KPIs are more aligned with process efficiency and managerial control, whereas financial ratios primarily measure outcomes rather than drivers of performance. Despite their widespread application in corporate decision-making, the interplay between financial ratios and sector KPIs remains underexplored in academic literature.

This study aims to bridge this research gap by investigating whether financial ratios and sector KPIs are two sides of the same coin, analyzing their correlation in large Italian companies across four key industries: airlines, retail, telecommunications, and insurance. By analyzing financial ratios and sector KPIs for Italian firms during the 2011–2013 period, this study aims to determine whether sector KPIs can serve as predictive indicators of financial performance. The findings have implications for both practitioners and scholars, offering insights into how different industries balance financial metrics and operational KPIs to achieve long-term sustainability and competitive advantage.

The study has limitations, particularly regarding the sample size and the focus on a limited time period. Additionally, the research is exploratory in nature, meaning that its conclusions should be interpreted with caution. Nonetheless, we believe that this study has several practical insights for professionals and contributes to scholarly debates on performance measurement.

The remainder of the paper is structured as follows: Section 2 presents the theoretical background and literature review, tracing the evolution of financial ratios and sector-specific KPIs and discussing their respective strengths and limitations. Section 3 outlines the research methodology, including data selection and analysis approach. Section 4 presents the findings and discussion, comparing sector KPIs with financial ratios across industries. Finally, Section 5 concludes the study, highlighting key implications, limitations, and potential directions for future research.

2. THEORETICAL BACKGROUND AND LITERATURE REVIEW

2.1 Financial Ratios and Key Performance Indicators

The use of financial ratios (FRs) and Key Performance Indicators (KPIs) in corporate performance measurement has evolved over time. While FRs date back to the late 19th century (Horrigan, 1968), KPIs emerged in the 1970s as part of a broader shift toward strategic performance management (Rockart, 1976).

Both FRs and KPIs serve the fundamental purpose of enabling organizations to establish effective control mechanisms, yet they differ in approach. Financial ratios have traditionally been used by both internal and external stakeholders to monitor financial performance across periods and between different companies, particularly for benchmarking processes. These ratios can be classified into four main categories:

- Profitability ratios – Measure the company's ability to generate returns on assets or equity. Key indicators include ROCE, ROA, ROI, ROS, and ROE.
- Liquidity ratios – Assess a firm’s ability to meet short-term obligations.
- Activity ratios – Evaluate operational efficiency.
- Leverage ratios – Indicate financial risk and debt dependency.

Financial ratios are inherently linked to financial statements, which are the outcome of accounting processes and adjusting entries. They serve as a historical snapshot rather than a forward-looking performance measure. As a result,

financial ratios may be insufficient for assessing real-time operational efficiency and strategy execution (Kaplan & Norton, 1992).

There are several issues to be aware of in the interpretation of financial ratios, such as the comparability among periods – i.e. due to changes in the accounting process of the same company – or among companies, when they aggregate financial information differently. In any case, financial analysis reports to us a company's financial information and not operational information: from this point of view FRs are like an old picture, that tell us about the present and the past but enough about the future.

Despite the limits, financial ratios are a widely used tool, both in internal and external analysis, by managers, executives, board of directors, creditors, banks, investors and regulatory authorities in their respective activities (Beaver, 1966; Altman, 1968; Penman, 2007).

In contrast, KPIs provide a more dynamic and strategic perspective, focusing on key success factors specific to each industry (Neely, 1999; Ittner & Larcker, 2003). KPIs are widely used by internal stakeholders, such as managers and executives, as well as external analysts and investors, who employ them for performance measurement and strategic decision-making (Neely, 1999). Unlike financial ratios, KPIs:

- Are not standardized, allowing for sector- or company-specific customization.
- Offer real-time insights into operational effectiveness.
- Align more closely with business strategies and process efficiency (Marr, 2012).

KPI can be divided into two broad categories, such as:

- Financial: measure financial health and performance of an organization. Some examples of financial KPIs are revenue growth, cash flow, return on investment (ROI), inventory turnover, Debt-to-Equity ratio, etc. The target of financial measures from the BSC approach is associated with the main goals of a company in that perspective: to survive, to succeed and to prosper (Kaplan and Norton, 1992).
- Non-financial: relate to different aspects of a business other than the financial perspective. For instance, some examples of non-financial measures are customer satisfaction, product or service quality, employee motivation, brand reputation, innovation rate, etc.

The control process of the company at all levels can take advantage of monitoring those quantities, identifying variations and possible reasons. From this perspective, KPIs can be considered "monitoring elements" , provided that four associated values are defined for each: a target value (to be achieved), a threshold (distinguishing poor performance from good performance), a current value, and the worst value, which serves as a critical alert for managers to take action (Maté, Trujillo, and Mylopoulos, 2012).

Internal users, both at top and middle level management, tend to identify firm specific KPIs, like the number of new customers or the number of new products launched. External users focus on KPIs with a wider use, preferably for the entire sector (Domínguez et al., 2019).

A significant advancement in the use of KPIs linked to business strategy was made in 1992 by Robert Kaplan and David Norton with the introduction of the Balanced Scorecard (BSC). In this framework, organizational performance is evaluated from four interconnected perspectives: financial, customer, internal process, and learning and growth. For each perspective, companies must establish goals and define their respective measures, providing a comprehensive overview of the company's results. The authors also stress that managers should focus on monitoring a few critical measures rather than frequently adding new ones. Today, the BSC remains one of the primary frameworks from which organizations derive their KPIs.

2.2 Sector-Specific KPIs

Some industries have developed long-standing KPIs that provide unique performance insights beyond traditional financial ratios. Examples include:

- Airline sector: Load factor (%) – Measures capacity utilization and operational efficiency.
- Retail sector: Sales revenue per square meter – Evaluates store profitability.
- Telecom sector: Average Revenue Per User (ARPU) – Assesses profitability per customer.

Sector KPIs partially address the limitations of financial ratios, offering a more focused and timely performance measure. However, they remain industry-specific and may not fully capture a company's overall financial health (Eccles, 1991).

Sector-specific KPIs have been extensively studied across industries, demonstrating their role in performance management. In the airline industry, load factor (%) has been established as a key efficiency measure, directly linked to profitability (Cento, 2009). Similarly, in the retail sector, sales per square meter has been widely adopted to evaluate store productivity (Grewal et al., 2017). Telecommunications companies rely on ARPU to track revenue streams from customers, which is a strong predictor of long-term profitability (Gerpott, Rams, & Schindler, 2001).

Nonetheless, this study explores whether profit financial ratios and sector KPIs are fundamentally different measures or interrelated components of performance evaluation and management. While they rely on distinct data sources, we hypothesize that firms that perform well in sector KPIs also exhibit strong financial ratios, forming a predictive relationship between the two. In summary, we posit the following research hypothesis:

Hp1: Sector KPI and Profit financial ratios are strongly correlated, so that a company high positioning in the sector KPI correspond to a similar positioning for financial performance ratio list.

3. METHODOLOGY, RESEARCH APPROACH, SAMPLE DATA

To investigate the relationship between financial ratios and sector KPI we decided to adopt the selected case-study methodology, identifying some sectors and some companies to analyze in a limited time period (Yin, 2009; Eisenhardt, 1989). We are aware that because of this approach, the results of our analysis have limited validity. Nonetheless the exploratory nature of this research and the difficulty in obtaining information on internal indicators for a broader set of companies, we believe can justify this choice, especially considering the interesting empirical insights we expect to derive from this study.

The first step of our research was the selection of business sectors for which to analyze KPIs. We requested they have: 1. consensus and use among companies, and 2. availability for external users. According to our literature review, sector KPIs are different from financial ratios and are sector specific, so they are comparable only within the same sector.

We identified the following sectors: insurance, gross distribution, mobile phone operators, railways companies, fast foods chains, on-line travel agencies. For those sectors, we prepared a list of primary companies operating in Italy (in terms of financial turnover greater than 10 million euros and number of employees that had to be a minimum of 50), most of which had headquarters in Italy. Financial data were extracted mainly from the AIDA databank (Bureau Van Dick), whereas sector KPIs were collected from the company website or from journal articles. Sector KPIs were not available for all the companies identified so some of them had to be excluded. This was the main difficulty we identified to enlarge our study to get a solid statistical approach. The years considered were 2011, 2012 and 2013 for the availability of sector KPI data.

In the second step, the number of sectors was then reduced considering only those with: 1. at least a two-year period availability (we considered years 2012 and 2013), both for sector KPIs and financial ratios and 2. at least three competitors to analyze.

The sector KPIs were selected considering the operating side of the business, going deep into the key drivers of its profitability. To be homogeneous with sector KPIs, the economic-financial ratios were selected to reflect the operating profitability of the entire business (excluding extraordinary financial elements) and not only sales or margins. We considered one ratio only for financial performance and one sector KPI for all four industries. In most cases we used ROI ratio (EBIT over Total Assets), except for insurance companies for which we used the ROA ratio (Net Income/Total Assets) due to the specificity of the income statement.

Definitively, the four sectors identified were: 1. Airline companies; 2. Large-scale retail channel companies; 3. Mobile phone operators, 4. Insurance companies.

The list of KPIs and financial ratios identified is reported below (Table 1):

Table 1: Sectors and financial ratios considered in this study

N.	Sector/Industry	KPI	Financial Ratio
1	Airline companies	Load factor %	ROI
2	Large scale retail channel	Sales revenues/square meters	ROI

3	Mobile Phone operators	Average Revenue Per User	ROI
4	Insurance companies	Combined ratio	ROA

Source: Our elaboration

In the airline sector the sector KPI identified for this analysis was the load factor (in %). The load factor KPI has a wide use in airline companies and in all the other transport services. It is an easy and effective measure of economies of scale that are one of the most important factors. Despite of production companies, in these service companies, the excess of capacity cannot be stored, so a right control of the capacity used is fundamental. According to the level of service and pricing policy, airline companies can be divided into two groups: traditional and no frill companies' business model. The load factor (in percentage) was the KPI selected for the airline sector. In table 2 are presented the load factors for the seven main companies operating in Italy: Air Dolomiti, Air France KLM, Alitalia, Deutsche Lufthansa, EasyJet, Meridiana, Ryanair.

The sales revenues/square meters (SRSM) KPI has a long tradition in the large-scale retail channel sector. It is quite simple to calculate, but effective as a measure of general performance. It measures the sales revenues (euro) that each square meter of the point of sale can generate (an average) in one year. The business model adopted by the company "traditional" or "discount" influences this KPI in a benchmarking process. Six companies have been selected for the analysis during the period 2011-2013: Auchan, Conad, Coop Italia, Despar, Essselunga, Eurospin Italia, Gruppo PAM, Selex. The SRSM values are reported in Table 4.

The ARPU (average revenue per unit) was the KPI identified for the analysis in the mobile phone sector. It reflects the value that each active SIM can generate, both for phone traffic and data traffic, in one month or in a year. Revenues for data traffic are more and more relevant, due to the strong competition in voice traffic. Four companies were considered: H3G, Telecom Italia, Vodafone Italia, and Wind. The ARPU values for the selected companies are included in Table 6.

For the insurance industry the sector KPI identified for this analysis was the combined ratio. The combined ratio is calculated by taking the sum of losses and expenses incurred divided by earned premiums. For an insurance company it is a general expression of profitability in the daily operating activity, without considering profits from investments. Normal values for a profitable insurance company are a few below 100%. Values for on-line insurance companies tend to be better than others, as they have a lower cost for expenses to the network. Four companies of this Industry were analysed: Cattolica Assicurazioni, Generali Assicurazioni, Unipol, Vittoria Assicurazioni. Table 8 reports the combined ratios of our selected companies.

After collecting the data for each company relevant to our analysis, we prepared two tables for each sector (one for financial performance values and the other sector KPI). In each table we included one column to highlight the relative position of each company in the list. Our analysis therefore focused on comparing the relative ranking of each company within the industry list.

4. FINDINGS AND DISCUSSION

In table 2 are presented the load factors (%) of the seven companies analysed for our first industry (Airline). All the companies in all the years considered (2011-2013) maintain a growing trend with values over 60%. For 2012 and 2013 we include also the relative positioning of each company in the list. (Pos. Column). EasyJet is at top in all years.

Table 2: : Load factor % - Sector KPI for airline companies

N.	Company	2011	2012	Pos.	2013	Pos.
1	Air Dolomiti	60,0%	65,0%	6	70,0%	6
2	Air France KLM*	82,0%	83,2%	2	83,8%	2
3	Alitalia	72,8%	74,6%	5	74,8%	5
4	Deutsche Lufthansa*	77,6%	78,8%	4	79,8%	4
5	EasyJet	87,3%	88,7%	1	89,3%	1
6	Meridiana	61,0%	63,0%	7	-	-
7	Ryanair	82,0%	81,2%	3	83,0%	3

Source: Our elaboration

EasyJet and Ryanair, which have a "NO FRILL" business model, had the highest load factor (near 90% both in 2012 and 2013). AirFrance-KLM and Deutshe Lufthansa were the best companies with the traditional business model (load factor near 85-80%), while Alitalia had the worst value, only 75% (in our sample).

Table 3: ROI ratio (%) for airline companies

N.	Company	2011	2012	Pos..	2013	Pos.
1	Air Dolomiti	60,0%	65,0%	6	70,0%	6
2	Air France KLM*	82,0%	83,2%	2	83,8%	2
3	Alitalia	72,8%	74,6%	5	74,8%	5
4	Deutsche Lufthansa*	77,6%	78,8%	4	79,8%	4
5	EasyJet	87,3%	88,7%	1	89,3%	1
6	Meridiana	61,0%	63,0%	7	-	-
7	Ryanair	82,0%	81,2%	3	83,0%	3

Source: Our elaboration

If we consider the economic financial ratio ROI (Table 3) the best performance was realized by Easyjet and Ryanair that are ahead of the companies with the "traditional" business model.

Combining the position of companies in the two tables, Easyjet and Ryanair are at the top, while Alitalia and Meridiana are at bottom.

The sector KPI considered for the large scale retail channel was the ratio Sales/square meters of point of sale (Table 4). Eight companies were considered. If we exclude Esselunga, that has one of the highest in Europe, other retail channels have values near 5-6,000 euros per square meter, both for "traditional" and "discount" business model.

Table 4: Sales / Square meters - Sector KPI for large-scale retail channels

N.	Company	2011	2012	Pos.	2013	Pos.
1	Auchan	5,480	5,410	4	4,995	6
2	Conad	6,470	6,310	3	5,962	4
3	Coop Italia	7,320	7,230	2	7,030	2
4	Despar	4,750	4,610	6	4,530	8
5	Essselunga	18,090	18,050	1	17,734	1
6	Eurospin Italia	n.a.	n.a.	-	6,706	3
7	Gruppo PAM	n.a.	n.a.	-	5,218	5
8	Selex	4,680	4,670	5	4,702	7

Source: Our elaboration on data from http://www.gdonews.it/2014/05/18/le-redditivita-per-metro-quadrato-dei-discount-sono-eccellenti/

The second position for sector KPI was for Coop Italia both in 2012 and 2013, and Eurospin (discount business model) had the next, whereas Despar has the last.

Table 5: ROI ratio (%) for large-scale retail channels

N.	Company	2011	2012	Pos.	2013	Pos.
1	Auchan	-4.63	-3.79	8	-5.03	7
2	Conad	0.27	0.42	5	1.37	5
3	Coop Italia	1.33	0.26	7	0.60	6
4	Despar	4.77	4.65	3	n.a.	-
5	Essselunga	8.95	6.64	2	7.10	2
6	Eurospin Italia	7.92	11.06	1	12.90	1
7	Gruppo PAM	1.97	2.09	4	1.73	4

8	Selex	3.61	0.32	6	2.64	3

Source: Our elaboration

If we consider the ROI ratio (table 5), Eurospin had the top position both in 2012 and 2013, followed by Esselunga, Selex and Gruppo PAM. Eurospin and Esselunga had top position both in KPI and in ROI tables.

Combining the position of each company in table 4 and table 5, Esselunga and Eurospin had the highest values in both tables, while Auchan had the lowest.

In Table 6 are presented the ARPU (Average Revenues Per User) of the main competitors in the Italian mobile phone arena. The leadership in sector KPI is for Vodafone Italia, both in 2012 and 2013, while Telecom Italia and Wind are the next. H3G is the last.

Table 6: A.R.P.U. - Sector KPI for mobile phone operators

N.	Company	2011	2012	Pos.	2013	Pos.
1	H3G	16.02	13.03	4	12.30	4
2	Telecom Italia	17.80	15.91	2	14.30	2
3	Vodafone Italia	19.82	18.12	1	15.66	1
4	Wind	15.65	14.36	3	13.17	3

Source: http://www.cmcrp.org/wp-content/uploads/2013/11/ARPU-by-Firm-and-Country.xlsx

ROI financial ratio for main Italian mobile phone operators are presented in table 7. The position of each competitor is aligned to the position in the sector KPI table in both years 2012 and 2013.

Table 7: ROI ratio (%) for mobile phone operators

N.	Company	2011	2012	Pos.	2013	Pos.
1	H3G	-2.20	0.30	4	-0.30	4
2	Telecom Italia	12.10	12.70	2	11.90	2
3	Vodafone Italia	50.80	35.40	1	17.50	1
4	Wind	9.60	8.70	3	6.70	3

Source: Our elaboration

The sector KPI considered for the Insurance industry was the combined ratio (Table 8). During the period 2011-2013 all the four companies selected (Cattolica, Generali Italia, Unipol and Vittoria Assicurazioni) maintained values below 100%. The trend during the period 2011-2013 was positive, as the ratio expenses plus cost for payments divided by payments received dropped down approximately to 92% for “traditional” companies (and 90% for on-line ones).

Table 8: A.R.P.U. - Combined ratio (%) - Sector KPI for insurance companies

N.	Company	2011	2012	Pos.	2013	Pos.
1	Cattolica Assicurazioni	96.2%	94.9%	2	94.5%	4
2	Generali Assicurazioni	96.8%	97.6%	3	91.5%	2
3	Unipol Assicurazioni	95.5%	99.3%	4	92.2%	3
4	Vittoria Assicurazioni	97.3%	92.9%	1	90.7%	1

Source: our elaboration from www.fisac-cgil.it/3831/lab-la-novestrale-dei-gruppi-assicurativi

If we look at the relative position of each company (column “position - pos.”), in 2012 combined ratio revealed the best performance for Vittoria Assicurazioni (on-line business model), followed by the other three “traditional” business models: Cattolica Assicurazioni, Generali Assicurazioni and Unipol. In 2013 the best performance was again realized by Vittoria Assicurazioni, followed by Generali Assicurazioni, then Unipol and Cattolica.

Table 9: ROA ratio (%) for insurance companies

N.	Company	2011	2012	Pos.	2013	Pos.
1	Cattolica Assicurazioni	0.12	0.86	2	0.89	2
2	Generali Assicurazioni	0.43	0.37	4	0.52	4
3	Unipol Assicurazioni	-0.68	0.86	3	0.62	3
4	Vittoria Assicurazioni	3.30	2.62	1	2.18	1

Source: Our elaboration

Considering ROA financial ratio (Table 9), the best performance was realized by Vittoria Assicurazioni both in 2012 and 2013, followed by Cattolica Assicurazioni, then Unipol and Generali Assicurazioni.

Considering the combined position in the two tables, Vittoria Assicurazioni had the best position both in sector KPI and in ROA financial ratio in 2012 and 2013. Cattolica Assicurazioni had the second position for both ratios in 2012, but not in 2013 where they had the second place in ROA but the last in sector KPI.

At this point, we can summarize our findings as follows.

In the airline sector, companies with higher load factor percentages (e.g., EasyJet and Ryanair) consistently ranked at the top for ROI, reinforcing previous findings that capacity utilization is linked to financial performance in aviation (Cento, 2009).

Similarly, in the retail sector we found a positive relationship between sales revenue per square meter and ROI ratio, with Esselunga and Eurospin leading in both measures (Kaplan & Norton, 1996; Grewal et al., 2017). This supports the idea that retail KPIs serve as strong predictors of financial success, aligning with Grewal et al. (2017).

In the mobile phone sector the results indicate that ARPU (Average Revenue Per User) is a reliable indicator of financial strength, with Vodafone Italia consistently outperforming competitors in both ARPU and ROI (Kim et al., 2004; Gerpott et al., 2001). Similar findings have been reported in prior telecom studies (Gerpott, Rams, & Schindler, 2001).

Lastly, in the Insurance industry the combined ratio was found to be a valid operational efficiency metric, with Vittoria Assicurazioni leading in both combined ratio and ROA. This supports previous research highlighting the importance of cost-efficiency in insurance profitability (Cummins & Weiss, 2013).

Therefore, in all four industries considered, the relative rankings of companies based on sector KPIs were aligned with those obtained for economic and financial performance.

According to our results, Hp1 can be confirmed, i.e. sector KPI and Profitability financial ratios are strongly correlated, so that a company high positioning in the sector KPI correspond to a similar positioning for financial performance ratio list.

5. CONCLUSION

This study aimed to examine the relationship between profitability financial ratios and sector-specific KPIs in four key industries: airlines, retail, telecommunications, and insurance. Using data from Italian large companies between 2011 and 2013, the findings reveal a strong correlation between sector KPIs and financial performance metrics, supporting the hypothesis that companies ranking high in sector KPIs also demonstrate strong financial performance ratios. Different industries can balance financial metrics and operational sector KPIs to achieve long-term sustainability and competitive advantage.

The study has limitations, such as the sample size, time period considered, industry scope, and countries considered. We also did not account for external factors that can moderate this relationship that we have not studied in terms of causation. Another limit is that we did not incorporate real-time operational KPIs, which could provide additional insight into firm performance.

Nonetheless we believe that this study has several practical insights for professionals because it highlights the usefulness of sector KPI and financial ratios jointly to obtain a comprehensive view of performance and make better strategic decisions (Ittner & Larcker, 2003; Marr & Schiuma, 2003). It also contributes to investment decisions, because investors can use sector KPIs as early indicators of financial success, improving their ability to identify high-performing firms before traditional financial reports are published. In addition, policymakers could develop industry-specific reporting guidelines that integrate both financial and operational performance measures to enhance corporate transparency.

We also contribute academic research on performance measurement and corporate finance bridging the gap between financial and non-financial measures, highlighting the predictive power of KPIs, which can be used as leading indicators of financial success, complementing traditional financial ratios, offering a foundation for more in-depth research on the causal relationship between KPIs and financial performance.

In conclusion, financial ratios and sector KPIs serve as complementary performance measures. Therefore, we can affirm that 'financial performance ratios and sector KPIs are two faces of the same coin'.

Funding/Acknowledgements

This paper has received no external funding.

Conflict of Interest

The authors declare NO CONFLICT of INTEREST.

References

[1] Adams, S. (2013). Performance measurement and management: A strategic approach to financial ratios and KPIs. Routledge.
[2] Altman, E. I. (1968). "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." The Journal of Finance, 23(4), 589-609.
[3] Bauwhede, H. V. (2009). "Financial Reporting Quality and Corporate Governance: The Influence of Board Independence." Corporate Governance: An International Review, 17(5), 447-469.
[4] Beaver, W. H. (1966). "Financial Ratios as Predictors of Failure." Journal of Accounting Research, 4, 71-111.
[5] Bikker, J. A., & Bos, J. W. B. (2008). Bank Performance: A Theoretical and Empirical Framework for the Analysis of Profitability, Competition, and Efficiency. Routledge.
[6] Bryde, D. (2005). Project Management Concepts, Methods, and Applications. Palgrave Macmillan.
[7] Cento, A. (2009). The Airline Industry: Challenges in the 21st Century. Physica-Verlag.
[8] Cummins, J. D., & Weiss, M. A. (2013). "Analyzing Firm Performance in the Insurance Industry Using Frontier Efficiency Methods." Geneva Papers on Risk and Insurance - Issues and Practice, 38(4), 520-557.
[9] Domínguez, E., Pérez, B., Rubio, A. L., & Zapata, M. A. (2019). A taxonomy for key performance indicators management. Computer Standards & Interfaces, 64, 24-40.
[10] Eccles, R. G. (1991). "The Performance Measurement Manifesto." Harvard Business Review, 69(1), 131-137.
[11] Eisenhardt, K. M. (1989). "Building Theories from Case Study Research." Academy of Management Review, 14(4), 532-550.
[12] Fama, E. F., & French, K. R. (1993). "Common Risk Factors in the Returns on Stocks and Bonds." Journal of Financial Economics, 33(1), 3-56.
[13] Gerpott, T. J., Rams, W., & Schindler, A. (2001). "Customer Retention, Loyalty, and Satisfaction in the German Mobile Cellular Telecommunications Market." Telecommunications Policy, 25(4), 249-269.
[14] Grewal, D., Roggeveen, A. L., & Nordfält, J. (2017). "The Future of Retailing." Journal of Retailing, 93(1), 1-6.
[15] Horrigan, J. O. (1968). "A Short History of Financial Ratio Analysis." Accounting Review, 43(2), 284-294.
[16] Hübner, A., Kuhn, H., & Sternbeck, M. G. (2013). "Retail Category Management: State-of-the-Art Review of Quantitative Research and Future Directions." Omega, 40(2), 119-128.
[17] Kaplan, R. S., & Norton, D. P. (1992). "The Balanced Scorecard: Measures that Drive Performance." Harvard Business Review, 70(1), 71-79.
[18] Kim, M. K., Park, M. C., & Jeong, D. H. (2004). "The Effects of Customer Satisfaction and Switching Barriers on Customer Loyalty in Korean Mobile Telecommunications Services." Telecommunications Policy, 28(2), 145-159.
[19] Ittner, C. D., & Larcker, D. F. (1998). "Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction." Journal of Accounting Research, 36, 1-35.
[20] Ittner, C. D., & Larcker, D. F. (2003). "Coming Up Short on Nonfinancial Performance Measurement." Harvard Business Review, 81(11), 88-95.
[21] Kaplan, R. S., & Norton, D. P. (1996). The Balanced Scorecard: Translating Strategy into Action. Harvard Business Press.
[22] Kim, M. K., Park, M. C., & Jeong, D. H. (2004). "The Effects of Customer Satisfaction and Switching Barriers on Customer Loyalty in Korean Mobile Telecommunications Services." Telecommunications Policy, 28(2), 145-159.
[23] Larcker, D. F., Richardson, S. A., & Tuna, I. (2005). "How Important is Corporate Governance?" Journal of Accounting and Economics, 39(1), 3-43.
[24] Marr, B. (2012). Key Performance Indicators: The 75+ Measures Every Manager Needs to Know. Pearson.
[25] Marr, B., & Schiuma, G. (2003). "Business Performance Measurement – Past, Present and Future." Management Decision, 41(8), 680-687.
[26] Maté, A., Trujillo, J., & Mylopoulos, J. (2012). Conceptualizing and specifying key performance indicators in business strategy models. In Conceptual Modeling: 31st International Conference ER 2012, Florence, Italy, October 15-18, 2012. Proceedings 31 (pp. 282-291). Springer Berlin Heidelberg.
[27] Mtau, T. and Rahul, N. Optimizing Business Performance through KPI Alignment: A Comprehensive Analysis of Key Performance Indicators and Strategic Objectives. American Journal of Industrial and Business Management, 14, (2024), p. 66-82 . doi: 10.4236/ajibm.2024.141003.

[28] Neely, A. (1999). "The Performance Measurement Revolution: Why Now and What Next?" International Journal of Operations & Production Management, 19(2), 205-228.
[29] Parmenter, D. Key performance indicators: developing, implementing, and using winning KPIs, John Wiley & Sons, Inc., New York, NY, USA, 2015.
[30] Penman, S. H. (2007). Financial Statement Analysis and Security Valuation. McGraw-Hill.
[31] Rockart, J. F. (1976). "Chief Executives Define Their Own Data Needs." Harvard Business Review, 57(2), 81-93.
[32] Shleifer, A., & Vishny, R. W. (1997). "A Survey of Corporate Governance." Journal of Finance, 52(2), 737-783.
[33] Tsai, W. H., Hsu, W., Chen, C., & Lin, W. (2011). "An Integrated Approach for Airline Passenger Value Analysis Using KPIs and Data Mining." Expert Systems with Applications, 38(6), 7110-7119.
[34] Walsh, P. (1996). Performance Measurement and Management: Accounting, Financial Ratios, and KPIs. McGraw-Hill.
[35] Yin, R. K. (2009). Case Study Research: Design and Methods. SAGE Publications.

Tracing the Evolution of Corporate Governance Codes in Europe

Paolo ROFFIA [a*]

[a] *University of Verona, Department of Management, Via Cantarane 24, Verona, 37124, Italy, paolo.roffia@univr.it*[b]

Abstract

Corporate governance across Europe has experienced substantial evolution over the past thirty years, influenced by recurring financial crises, regulatory innovations, and a growing emphasis on sustainability and stakeholder engagement. In response to these shifts, five major European economies—France, Germany, Italy, Spain, and the United Kingdom—have collectively amended their corporate governance codes 56 times, demonstrating an ongoing commitment to adapting governance practices to changing economic conditions and regulatory demands. These developments highlight the dynamic and responsive nature of governance systems, aiming to align corporate practices with the expectations of increasingly complex financial markets and broader societal concerns. This paper traces the historical progression of corporate governance codes in Europe, focusing on major revisions, emerging themes, and their implications for corporate governance structures and investor relations. By conducting a comparative analysis of these five countries—selected due to their economic significance and leadership within the European context—the study examines the transition from a shareholder-centric paradigm to a more inclusive, stakeholder-oriented governance model. The analysis reveals a consistent trend toward refining governance frameworks to incorporate digital transformation, environmental, social, and governance (ESG) considerations, and stronger board accountability. Furthermore, the paper investigates how variations in national governance models shape corporate conduct and influence the broader movement toward regulatory harmonization in the European Union. The study concludes with recommendations for enhancing EU corporate governance practices to meet future challenges, underscoring the need for ongoing reform, improved transparency, and the integration of sustainability-focused policies to support resilient and accountable corporate behavior.

Keywords: corporate governance, corporate governance code, Institutional theory, listed companies, Europe

1. INTRODUCTION

Corporate governance plays a foundational role in defining the relationship between firms, stakeholders, and regulatory systems. In Europe, its evolution has been shaped by economic integration, capital market liberalization, and the increasing influence of global standards such as those developed by the OECD and the European Commission.

Corporate governance serves as a fundamental pillar in ensuring transparency, accountability, and sustainability in business practices. Within the EU, corporate governance codes have evolved in response to economic shifts, regulatory reforms, and global trends. The development of these codes has been particularly reactive to financial turbulence, corporate scandals, and the growing demand for sustainable and socially responsible business conduct. "Corporate governance codes provide a vital mechanism for reinforcing investor confidence and enhancing economic stability"

* Corresponding author.

(OECD, 2015). The introduction of governance frameworks such as the European Commission's recommendations, national governance codes, and the EU Shareholder Rights Directives has played a pivotal role in shaping corporate governance across member states. Moreover, the shift from a shareholder-centric to a stakeholder-oriented perspective reflects broader institutional and cultural changes across the continent.

This paper aims to explore the evolution of corporate governance codes within Europe, analyzing key regulatory changes, their impact on corporate structures, and emerging governance trends. By tracing historical developments and comparing governance practices across Europe, this study provides insights into the future trajectory of European corporate governance.

This paper is structured as follows: Section 2 explores the historical development of EU corporate governance codes, focusing on early frameworks, the impact of financial crises, and the shift towards stakeholder governance. Section 3 presents a comparative analysis of key governance codes across selected EU member states. Section 4 discusses emerging trends in corporate governance, including ESG integration, digitalization, and shareholder rights. Finally, Section 5 provides conclusions and insights into the future trajectory of corporate governance within the EU.

2. THE DEVELOPMENT OF CORPORATE GOVERNANCE CODES IN EUROPE

More than 30 years have passed since the first leading European corporate governance code was issued in 1992. The 1992 Report of the Committee on the Financial Aspects of Corporate Governance – also known as the Cadbury Code, named after the committee's chairperson – was a relatively simple code but established fundamental principles and practices for good governance in listed companies (Haxhi & Aguilera, 2012). Before that, other smaller corporate governance codes had emerged in the USA (1978) through the Business Roundtable's report "The Role and Composition of the Board of Directors of the Large Publicly Owned Corporation," followed by regulatory frameworks developed by the U.S. Security Exchange Commission and the New York Stock Exchange. Similar initiatives were undertaken in Hong Kong (1989) with the Code of Best Practice and in Ireland (1991) with the Statement of Best Practice on the Role and Responsibility of Directors of Publicly Listed Companies (Aguilera & Cuervo-Cazurra, 2004).

Since then, more than 90 countries have formulated corporate governance codes, with notable acceleration in recent years (Cuomo et al., 2016; Haxhi & Aguilera, 2012). European nations, including France, Spain, Germany, and Italy, have adapted and revised their governance codes over time (EU, 2002). Even smaller economies and listed companies in various jurisdictions have adopted corporate governance principles due to increasing demands from international financial markets (Yoshikawa & Rasheed, 2009).

2.1 Early Governance Frameworks

The foundation of European corporate governance principles can be traced back to the 1990s when the European Commission issued its first recommendations on corporate governance. These early frameworks focused on strengthening shareholder rights, increasing transparency, and enhancing board responsibilities (European Commission, 1999). However, the first comprehensive European corporate governance code was the 1992 UK Cadbury Report, which introduced the concept of board independence, internal controls, and financial transparency. It laid the groundwork for subsequent codes across the continent. This was later developed into the Combined Code on Corporate Governance in 1998, which integrated principles from the Cadbury, Greenbury, and Hampel reports. The Combined Code became a benchmark for best practices, influencing governance structures throughout Europe (Tricker, 2015).

2.2 The Impact of Financial Crises

The early 2000s and the aftermath of the 2008 financial crisis prompted significant reforms in corporate governance. "The financial crisis of 2008 exposed fundamental weaknesses in corporate governance structures, leading to the revision of key regulatory frameworks" (Financial Stability Board, 2010). The crisis highlighted deficiencies in risk management, executive remuneration, and board oversight, leading to the adoption of the 2010 Green Paper on Corporate Governance and the revision of the Shareholder Rights Directive in 2017 (European Commission, 2017). National codes began to incorporate more stringent risk controls, clearer rules on remuneration disclosure, and enhanced shareholder engagement mechanisms. The French AFEP-MEDEF Code, for instance, introduced binding say-on-pay provisions, while the German DCGK added rules for more rigorous supervisory board evaluations (Bebchuk & Roe, 2004).

2.3 The Shift Towards Stakeholder Governance

Corporate governance codes have generally been introduced as part of a soft law approach, meaning they lack the binding force of statutory regulations (hard law) but are typically issued by multi-actor committees (Cuomo et al., 2016). However, in response to major financial scandals, such as Enron and Parmalat, certain measures have been implemented to enhance transparency and accountability (Haxhi & Aguilera, 2012). One of the fundamental principles underpinning these codes is the "comply or explain" approach, allowing issuers the flexibility to either adhere to governance guidelines or provide justifications for deviations—an application of "freedom with accountability" (Aguilera & Cuervo-Cazurra, 2004, 2009; Mallin, 2013).

Over the last two decades, corporate governance codes in Western Europe have evolved in alignment with academic recommendations and changing economic and social conditions (Gregory & Simmelkjaer, 2002; Haxhi & Aguilera, 2014). These codes have played a crucial role in shaping governance best practices, particularly regarding board composition, director independence, performance evaluations, executive remuneration, internal controls, and risk management processes (Chams et al., 2019).

Recent years have witnessed a shift from a shareholder-centric approach to a more inclusive governance model that integrates stakeholder interests, ESG considerations, and corporate sustainability. "Modern corporate governance should balance the interests of various stakeholders, including employees, communities, and investors, to ensure long-term value creation" (Freeman et al., 2020). The EU Green Deal and the Sustainable Finance Disclosure Regulation (SFDR) have reinforced the role of corporate governance in achieving long-term environmental and social objectives (European Parliament, 2020). Although the United Kingdom officially left the European Union in 2020, it has maintained alignment with EU corporate governance trends. The UK Corporate Governance Code continues to emphasize principles such as board accountability, stakeholder engagement, and ESG integration. Many of the post-Brexit reforms in the UK have preserved EU-compatible language and objectives, suggesting a deliberate effort to remain attractive to international investors who prioritize strong governance standards (Light, 2021). As such, the UK's governance approach continues to influence the broader European landscape.

Considering these discussions, this study explores the following research hypothesis:

Hp1: Corporate governance codes are dynamic and have evolved in response to economic, social, and technological changes, reflecting the shifting expectations of stakeholders.

3. METHODOLOGY, RESEARCH APPROACH, DATA COLLECTION

3.1 Research Approach

This study adopts content analysis as a methodological foundation, specifically through a systematic document analysis approach. Content analysis enables the structured examination of textual data, while systematic document analysis provides the procedural rigor necessary for analyzing official reports and policy documents (Krippendorff, 2018; Bowen, 2009). Together, these methods facilitate the identification of patterns, thematic shifts, and regulatory priorities across governance frameworks.

The study follows systematic document analysis methodology, reviewing European Commission directives, national corporate governance codes, and related academic literature to identify key modifications, governance trends, and regulatory priorities. Each governance code is analyzed in the context of major economic, financial, and regulatory events that have influenced their development (Bowen, 2009).

We considered the codes of corporate governance proposed by national commissions of the five most important Western European countries in terms of annual GDP in 2022 (Statista, 2024), namely Germany, the United Kingdom, France, Italy, and Spain. These countries were chosen because of their economic relevance and because they (excluding the Netherlands) have the most important financial markets, accounting for a major part of the European financial markets. This choice is also interesting because major differences exist among the selected countries' law systems (Soltani & Maupetit, 2015).

3.2 Data collection

The primary data sources for this research include:

- Official reports and publications from the European Commission and national governance institutions (European Commission, 1999, 2017; EU, 2002).

- Policy documents from international regulatory bodies, such as the Financial Stability Board, the OECD, and the International Sustainability Standards Board (FSB, 2016; OECD, 2015; IFRS, 2023).

- Empirical research and academic studies on the effectiveness and impact of governance codes (Aguilera & Jackson, 2010; Cuomo et al., 2016; Haxhi & Aguilera, 2012).

- Industry reports and governance frameworks implemented across different EU member states (Mallin, 2013).

3.3 Analytical Framework

The content analysis is structured around several key themes to assess governance evolution, specifically within the context of EU corporate governance codes:

- Board Responsibilities: The evolution of governance expectations for board independence, composition, and risk oversight in EU countries, with particular reference to the Shareholder Rights Directives and national adaptations (European Commission, 2017; OECD, 2015).

- Shareholder Rights and Protection: Strengthening shareholder participation, voting rights, and protection mechanisms in governance frameworks, as seen in the UK Corporate Governance Code and the German dual-board system (Financial Reporting Council, 2018; Schmidt & Spindler, 2007).

- Transparency and Disclosure: Developments in financial reporting, audit requirements, and corporate disclosure obligations under EU Directives and national laws, including the EU Non-Financial Reporting Directive (European Parliament, 2020).

- Sustainability and ESG Factors: The integration of environmental, social, and governance (ESG) considerations into corporate governance codes, particularly influenced by the EU Green Deal and the Sustainable Finance Disclosure Regulation (García-Castro et al., 2017; European Commission, 2020).

- Digitalization and Technological Governance: The adaptation of corporate governance principles to emerging technologies, cybersecurity risks, and AI governance, as reflected in the European Commission's Digital Finance Strategy and national codes (FSB, 2016; IFRS, 2023).

- Stakeholder Engagement: Expanding governance frameworks to incorporate the interests of non-shareholder stakeholders, including employees, communities, and broader society, aligning with the principles outlined in the AFEP-MEDEF Code in France and Scandinavian governance models (Mallin, 2013; Thomsen, 2016).

3.4 Comparative Analysis Method

A comparative analysis is conducted by systematically coding and categorizing governance principles across selected states. Each governance code is assessed based on:

- Consistency and Continuity: Identifying principles that have remained largely unchanged over time (European Commission, 1999, 2017; OECD, 2015).

- Modifications and Expansions: Tracking governance elements that have been updated or revised in response to new regulatory challenges (Aguilera & Jackson, 2010).

- New Additions: Highlighting emerging governance topics such as digitalization, sustainability, and corporate social responsibility (García-Castro et al., 2017).

- Regulatory and Economic Context: Examining the role of financial crises, corporate scandals, and global governance trends in shaping governance codes (La Porta et al., 1999).

3.5 Comparative Analysis Method

To ensure the reliability of this content analysis, multiple data sources are cross-referenced, and findings are validated against secondary research, such as peer-reviewed academic articles and corporate governance reports. Additionally, expert opinions from governance practitioners and policymakers are considered to contextualize the study's findings (Krippendorff, 2018).

The methodology enables a structured assessment of the evolution of EU corporate governance codes, providing insights into governance trends and their implications for businesses, investors, and regulators.

4. FINDINGS AND DISCUSSION

This section examines corporate governance codes across selected EU member states, detailing the historical development, key revisions, and unique features of each framework. European corporate governance models have been shaped by a mix of legal traditions, financial market structures, and political dynamics (Aguilera & Cuervo-Cazurra, 2004). Governance frameworks across EU countries have experienced notable adaptations, often influenced by corporate scandals, economic crises, and international regulatory trends (Mallin, 2013).

The German Corporate Governance Code (DCGK) was first introduced in 1998 and has undergone multiple revisions to enhance transparency and shareholder protection. Germany follows a dual-board system consisting of a management board and a supervisory board, ensuring a clear separation of executive and oversight roles. The 2009 revision strengthened risk management frameworks, while the 2020 update emphasized ESG considerations and executive compensation transparency (Schmidt & Spindler, 2007).

In France the AFEP-MEDEF Code, established in 1995, has played a crucial role in shaping corporate governance in France. French corporate governance emphasizes a balanced approach between shareholder rights and stakeholder engagement. The 2018 revision introduced stricter guidelines on executive remuneration and board diversity, while the 2022 update reinforced climate-related reporting requirements (Bebchuk & Roe, 2004).

The UK Corporate Governance Code, originally introduced in 1992 as the Cadbury Code, has been a benchmark for corporate governance best practices. The UK follows a comply-or-explain approach, requiring listed companies to either adhere to governance recommendations or justify deviations. The 2018 revision placed a greater focus on stakeholder engagement and board accountability, aligning with the increasing emphasis on ESG principles (Financial Reporting Council, 2018).

Italy adopted its first corporate governance code in 1999, aligning with EU regulatory frameworks. The Italian governance framework emphasizes the role of independent directors, shareholder rights, and board transparency. The 2020 revision introduced new sustainability disclosure requirements, reflecting the growing importance of ESG factors in corporate governance (Mallin, 2013).

Spain introduced its first corporate governance code in 1996, with subsequent updates focusing on enhancing board independence and shareholder rights. The 2015 revision aligned Spanish governance practices with EU directives, reinforcing transparency and risk management mechanisms. The 2020 revision further strengthened ESG reporting and board diversity requirements (European Commission, 2017).

The emerging trends we found in European Corporate Governance Codes are:

- ESG Integration: Increasing emphasis on environmental and social responsibilities in corporate governance (García-Castro et al., 2017).
- Digitalization and AI Governance: Adapting corporate governance to technological advancements and cybersecurity challenges (Brynjolfsson & McAfee, 2014).
- Strengthening Shareholder Rights: Enhanced transparency and engagement mechanisms in corporate decision-making (La Porta et al., 1999).
- Board Diversity and Inclusion: Greater emphasis on gender diversity and broader representation in corporate boards (Terjesen et al., 2009).

The evolution of corporate governance codes across Europe has been driven by a combination of economic crises, regulatory initiatives, and shifts in corporate strategy. Well-developed financial markets and strong institutional frameworks have facilitated the adoption and enforcement of governance codes in some countries, while others have struggled with compliance due to weaker regulatory mechanisms (Aguilera & Cuervo-Cazurra, 2004).

Corporate governance frameworks in Germany and France have been shaped by stakeholder-oriented models, balancing shareholder interests with those of employees and the broader economy (Thomsen, 2016). By contrast, the UK and Spain have emphasized shareholder primacy, relying on market-driven governance mechanisms to enhance corporate transparency and accountability (Mallin, 2013).

The increasing complexity of corporate structures has necessitated stronger internal governance mechanisms, prompting firms to adopt more rigorous board oversight and compliance measures. Institutional investors have also played a growing role in corporate governance, emphasizing active ownership and responsible investment strategies that align with ESG considerations (García-Castro et al., 2017).

Based on the findings obtained and the previous analysis, it is evident that corporate governance codes are dynamic and have evolved in response to economic, social, and technological changes, reflecting shifting stakeholder expectations and thus confirming Hp1. However, despite advancements, challenges remain in ensuring uniform application of governance principles across jurisdictions. The lack of standardized enforcement mechanisms limits their impact in some regions, while disparities in regulatory interpretation can lead to inconsistencies in governance practices (Cuomo et al., 2016).

Looking ahead, greater international collaboration and harmonization of governance standards may be necessary to enhance the effectiveness of corporate governance codes in fostering market stability and corporate resilience (OECD, 2015).

5. CONCLUSIONS, IMPLICATIONS AND LIMITATIONS

The evolution of European corporate governance codes reflects a dynamic interplay between regulatory developments, economic changes, and societal expectations. "The adaptability of governance frameworks is crucial in addressing emerging global challenges such as digital transformation and climate change" (Aguilera & Jackson, 2010). While substantial progress has been made in enhancing governance practices, ongoing challenges such as regulatory harmonization and sustainability require continuous adaptation of governance frameworks.

Future research should focus on assessing the effectiveness of recent governance reforms and exploring innovative governance solutions to address emerging global challenges. Additionally, an in-depth analysis of corporate governance practices in listed corporations should be conducted to examine how these codes are implemented in practice. This approach will help bridge the gap between theoretical governance frameworks and their real-world application, providing valuable insights into compliance levels, boardroom behaviors, and governance effectiveness across different industries and markets.

5.1 Implications for Scholars and Practitioners

For scholars, this study provides a foundation for further research into the impact of evolving governance codes on corporate performance, regulatory compliance, and stakeholder engagement. Future academic work should examine the interplay between governance regulations and firm behavior, exploring how different industries adapt to new governance requirements and whether corporate governance codes achieve their intended objectives.

For practitioners, including policymakers, corporate executives, and investors, the findings highlight the importance of aligning governance frameworks with contemporary business challenges. Strengthening board accountability, fostering sustainable corporate practices, and integrating digital governance mechanisms should be a priority. Investors can leverage governance assessments to guide responsible investment decisions, ensuring compliance with ESG and other regulatory expectations. Policymakers, in turn, must continue refining governance codes to enhance market confidence, economic stability, and corporate responsibility.

5.2 Limitations of the Study

This study has several limitations. First, the research primarily focuses on European countries, meaning that the findings may not be fully generalizable to corporate governance systems outside Europe. Second, the study employs a qualitative content analysis methodology, which, while effective for identifying patterns and thematic changes, does not include quantitative assessments of governance effectiveness. Future studies could complement this research by incorporating empirical data and statistical analysis to measure the real-world impact of governance reforms. Finally, while this study explores governance codes at a policy level, further research should investigate their practical implementation within individual corporations, assessing compliance and deviations in corporate governance practices across different industries and economic environments.

Funding/Acknowledgements

This paper has received no external funding.

Conflict of Interest

The authors declare NO CONFLICT of INTEREST.

References

[1] Aguilera, R. V., & Jackson, G. (2010). Comparative and international corporate governance. Academy of Management Annals, 4(1), 485-556.
[2] Bebchuk, L. A., & Roe, M. J. (2004). A theory of path dependence in corporate governance. Stanford Law Review, 52(1), 127-170.
[3] Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. W. W. Norton & Company.
[4] European Commission. (1999). EU corporate governance framework. Brussels: European Commission.
[5] European Commission. (2017). Revised Shareholder Rights Directive. Brussels: European Commission.
[6] European Parliament. (2020). The European Green Deal and corporate governance. Strasbourg: European Parliament.
[7] Financial Reporting Council. (2018). The UK Corporate Governance Code. London: FRC.
[8] Financial Stability Board. (2010). Principles for sound compensation practices. Basel: FSB.
[9] Freeman, R. E., Harrison, J. S., & Zyglidopoulos, S. (2020). Stakeholder theory: Concepts and strategies. Cambridge University Press.
[10] García-Castro, R., Ariño, M. A., & Canela, M. A. (2017). Corporate social responsibility, governance, and shareholders' value: An analysis of ESG practices. Journal of Business Ethics, 144(4), 707-721.
[11] Light, G. (2021). Corporate Governance After Brexit: Continuity or Change? European Business Law Review, 32(1), 89–107.
[12] La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1999). Corporate ownership around the world. The Journal of Finance, 54(2), 471-517.
[13] Krippendorff, K. (2018). Content analysis: An introduction to its methodology (4th ed.). Sage Publications.
[14] Schmidt, R. H., & Spindler, G. (2007). Path dependence, corporate governance, and complementarity. International Review of Law and Economics, 27(4), 386-409.
[15] Terjesen, S., Sealy, R., & Singh, V. (2009). Women directors on corporate boards: A review and research agenda. Corporate Governance: An International Review, 17(3), 320-337.
[16] Tricker, B. (2015). Corporate Governance: Principles, Policies, and Practices. Oxford University Press.
[17] Thomsen, S. (2016). The Nordic corporate governance model. Management and Organization Review, 12(1), 189-204.

Teaching, Learning and E-learning (IAC-TLEl)

AI Tools in Online Tests

Murat POLAT [a*]

[a] *Anadolu University, Education Faculty, Assessment & Evaluation, Eskisehir, Türkiye, mpolat@anadolu.edu.tr*

Abstract

The use of artificial intelligence, which has become a very popular research topic in the context of assessment and evaluation methods in education, can meet very important needs. Therefore, a review technique was chosen in this research to provide an in-depth discussion. First, the tools and methods that can be used by artificial intelligence in language education, their possible advantages and disadvantages, future applications and the role of teachers in future applications are discussed. Second, the extent to which assessment and evaluation tools using artificial intelligence can provide feedback on the assessment of learning will be discussed based on research findings. Third, the extent to which the use of artificial intelligence in assessment and evaluation can affect the accuracy, speed and usability of potential assessment applications will be investigated. Finally, the ethical aspect of using artificial intelligence for assessment and evaluation will be discussed, emphasizing what can be done with traditional assessment and evaluation but cannot be done with artificial intelligence.

Keywords: artificial intelligence, language education, assessment applications, traditional assessment

1. INTRODUCTION

In foreign language education, technology is now used in many areas to make individuals' lives easier, and this convenience has also begun to be felt especially in some branches of science. One of these areas is certainly the education. Considering the impact of technology on education, significant progress has been made in child education, especially in the last 20 years (Tuncer & Şen, 2018). One of these innovations is undoubtedly artificial intelligence, which we encounter in almost everywhere in our daily lives. Artificial intelligence has become a transformative force in education thanks to the many conveniences it promises. This new era, which can be described as revolutionary, includes technologies such as natural language processing (NLP), raw data analytics, and machine learning, which have the potential to take teaching and learning to a new level. By taking advantage of the power of artificial intelligence, educators can quickly prepare classroom materials, test content, create multi-dimensional feedback, and gain deeper insights into individual learning needs. The ability to use and process data allows educators to personalize their instruction, provide targeted interventions, and support each student as effectively as possible on their unique learning journey (Pirim, 2006). Stiggins argues that measurement and assessment, which can take longer than expected, can lead to a significant workload that teachers spend almost a third of their professional time on (2014). Despite their critical importance, research has shown that teachers generally need more support and training than currently provided, as well as new generation assessment and evaluation tools/methods that focus on AI-based assessment and evaluation (Mede and Atay, 2017; Tsagari, 2011).

Due to the coronavirus outbreak that started worldwide in 2019, there was a sudden and mandatory transition to distance education in many parts of the world, and most of the courses started to be provided through online platforms. This situation brought with it the need for artificial intelligence-based digital assessment and evaluation, and it was a

* Corresponding author.

challenging experience for teachers who were not familiar with these applications. Although teachers have sufficient knowledge about the concepts and applications related to assessment and evaluation and know how to use tests effectively to improve teaching and learning, this does not mean that teachers are effective evaluators in the digital environment. Therefore, it has become mandatory for teachers to have skills related to assessment and evaluation applications carried out in the digital environment. As a result, it is very important to determine the assessment and evaluation applications carried out in the education and training environment using artificial intelligence-supported tools and methods, that is, to keep up with the times, to save time in terms of assessment, to increase interaction and to be able to make individual decisions. In the meantime, when we consider the development of artificial intelligence in numbers, we see a tremendous increase in investments and research projects. To reiterate, technology is developing faster than ever with developments such as fifth-generation mobile internet connection, augmented reality, and blockchain. It can be said that artificial intelligence (AI) technology, which increasingly affects the stages of obtaining, processing, reporting, and using information worldwide, is one of the fundamental technological developments contributing to this change. For this reason, investments in AI technology research in universities and the private sector have gained significant momentum. In 2019, private AI investments in the global arena exceeded $70 billion (AI Index Annual Report, 2019). By 2020, the Massachusetts Institute of Technology (MIT) alone had invested $1 billion in new AI programs, while Microsoft had hired more than 8,000 AI researchers (Fabian, 2019). Although the demand for AI technologies is rapidly increasing on a global scale, it can be argued that the education system, including higher education institutions, is not sufficiently funded and students do not receive sufficient training to equip them with digital skills and AI technologies. Therefore, it is thought that students are not sufficiently exposed to innovative technological changes in education (Heintz, 2021) and artificial intelligence technology. Therefore, in the new knowledge-based society, it becomes difficult for higher education graduates to meet the increasing and changing demands of the labor market.

Considering the significant growth of AI and its applications in education, a number of questions such as:

1. What are the most common AI-generated assessment techniques?

2. What are the most important advantages of using AI-generated assessment techniques?

3. What is the function of teacher using AI-generated assessment techniques?

need to be asked and answered to provide a better view of the role of this technology in pedagogical issues. The literature included in this review study was considered within the framework of these research questions and the answers to these questions were carefully sought in studies focusing on the subject.

2. METHODOLOGY

A document review was carried out in order to access the latest research data, to bring together the relevant discussions on the subject and to explore the details more comprehensively. Document review or analysis, also known as documentary scanning, is the process of examining existing records and documents to obtain data. Document analysis involves reading and reporting the information and documents dealing with a specific problem or question (Karasar, 2005). This process sometimes involves compiling and examining written works such as books, journals and newspapers, and sometimes compiling and reading documents published in electronic media, which is a meticulous process that requires extensive work. Ekiz describes this whole process as the collection and processing of information and documents that might be used the for the research (2015). This preparation can sometimes be done by collecting documents that are accessible to everyone, while sometimes it can be done by a researcher who is allowed to access very specific information and documents. According to O'Leary (2017), document analysis is the process of gathering, analysing, weighing, and assessing data and documents that will help the researcher to avoid wasting research resources. In short, this type of research involves collecting and synthesizing information, documents and studies that have been written, conducted or presented to the public by other people or institutions in relation to the research problem they are focusing on. Similarly, the stages of document analysis have been approached in different ways by researchers. For example, O'Leary (2017) lists these stages as:

(1) planning for all eventualities;
(2) collecting documents,

(3) checking the reliability of the documents and questioning their intentional and unintentional evidence,
(4) analyzing the data.

On the basis of the different stages mentioned here, one can conclude that document analysis is carried out according to a specific plan, and appropriate methods should be employed to obtain, examine and analyse documents suitable for the purpose.

3. RESULTS

The literature was reviewed according to the research questions identified in the study and each research question was discussed in detail under separate headings and supported by examples from the literature.

3.1. What are the measurement and evaluation tools and methods used in education with artificial intelligence?

Artificial intelligence is changing the perception of education, making teachers and students ask the same question: "How much should we know?". It is crystal clear that the use of AI tools in various educational settings will help teachers improve their professional quality and increase the pace of learning, if properly implemented. To illustrate, Halaweh believes that the growing importance of AI has attracted the attention of researchers, who are actively investigating alternative methods to integrate a range of technological tools into the classroom in order to promote sustainable learning (2023). Thus, with the growing amount of educational data and the increasing pace of machine learning codes, the intensive use of AI has the potential to transform the way one learns, uses and checks his/her skills. As such, the benefits of using AI tools in testing are numerous. For example, Huang et al. (2023) argue that the use of AI for administrators, teachers, and students can be advantageous in developing critical thinking, gaining new insights, and fostering imagination. It has also been reported that artificial intelligence can play a crucial role in motivating students (Xia et al., 2022), increasing their participation in class (Lin et al., 2021), improving learner interaction reducing test anxiety (Ren, 2020), testing learning outcomes (Karsenti, 2019), monitoring real-time academic performance (Khan et al. 2021), and stimulating greater interest in learning (Haves & Arya, 2023).

If AI-based assessment kits are designed in accordance with the curriculum, they save time and opportunity in determining much more economical measurement and evaluation tools to measure goals and objectives, and in designing new tools (Khan et al., 2021). Where, when and how AI-based assessment kits can be used in measurement and evaluation has been discussed in detail by many researchers in the literature (Çukurova & Luckin, 2018; Karsenti, 2019):

- Computer-based testing platforms
- Formative assessment tools
- Digital assessment tools
- Learning management systems
- Plagiarism detection
- Predictive analytics tools
- Digital portfolios
- Gamification tools
- Automated grading tools
- Data visualization tools
- Learning analytics tools

Considering potential usefulness of AI in scoring and providing feedback on classical exams, we can say that the use of AI has the potential to become an indispensable assistant that can save time for language teachers, especially by allowing them to speed up and automate the marking process and providing students with fast and accurate feedback on their work (Öz, 2014). Artificial intelligence can analyze drafts, exam papers and even project reports to provide feedback regarding organization, vocabulary richness, grammar and mechanics in many languages. By using more AI systems in assessment, teachers can focus more on some basic educational needs, such as lesson planning, material development, and student counselling, and can save significant amount of time (Adıgüzel et al. 2023). To illustrate, the AI that powers Turnitin, a very popular homework submission program works through the NLP algorithms to analyze learners' written output and deliver feedback on vocabulary, grammar and spelling (İşler &

Kılıç, 2021). In addition, the same software can detect plagiarism by comparing different texts, allowing graders to grade assignments more quickly and accurately.

AI can also provide predictive analysis of learner achievement by incorporating student participation in classes, mid-term or end-of-term test data, and summative assessments. In addition to predicting achievement, these analytics can be used to identify students who might require extra help or guidance, enabling teachers to intervene in a timely and well-informed manner (Owan et al., 2023). Learning Management Systems (LMS), where AI is most commonly used in this sense, are platforms that allow experts to create, present and develop supporting materials, assignments, quizzes and feedback tools for students. With a little training, all teachers can choose the right LMS system, track their students' progress, and use these platforms to receive feedback on both student and teacher performance at the end of the process. Some popular examples include Moodle, Canvas and Blackboard (Suh & Ahn, 2022).

3.2. Benefits and Challenges of Assessment and Evaluation Using Artificial Intelligence in Education

Developments especially in the last two decades reveal that artificial intelligence technology has assumed important roles in areas that everyone is interested in such as architecture, economy, communication, sports and politics, and has become indispensable for people who are familiar with this convenience. Realizing this transformation, the superpowers such as the United States, Russia, China, the United Kingdom and the European Union countries are integrating artificial intelligence-based and even artificial intelligence education into their K-12 curricula and consider it as a component of their basic education systems (Steinbauer et al., 2021). Discussions regarding where artificial intelligence can be used most effectively suggest that there is an interdisciplinary connection due to the fact that the layered structure of this technology interconnects many different fields and thus, when used well, two targets can be hit by one arrow. For this reason, artificial intelligence education gives the opportunity to challenge what the human brain loves the most and can provide interested young minds with the opportunity to develop metacognitive skills such as reasoning, critical thinking, mathematical modelling and hypothetical problem solving.

On the other hand, compared to traditional educational methods, tools and training based on artificial intelligence offer the possibility of customizing learning environments, increasing the efficiency of education and extending the duration of learning. The use of artificial intelligence makes it possible to monitor learning speed in real time, and in this way it is possible to prepare personalized learning content for students whose needs we determine by analyzing their development, what they do well and what they do poorly. To do this, artificial intelligence can individually examine data generated at many different times, such as quizzes, homework and projects, to determine the individual's real-time learning path and provide feedback for improvement (Adıgüzel et al. 2023). Thanks to AI, people can learn at their own pace, speeding up or slowing down as they wish, and if they need help in other areas, this can be done quickly and economically. Today, AI-enabled personalized learning software such as Knewton or DreamBox, which have been created to meet such educational needs, are being used to create learning plans and content for their customers' educational needs at different speeds and with different content (Owan et al. 2023). In this way, students can experience adaptive learning through content and tests designed specifically for them, and progress without stress.

AI-based tutoring systems (ITS) are ideal for extracurricular educational tasks that normally take up a lot of a teacher's time, such as aiding or feedback. These systems not only provide educational services in the background, but can also identify an individual's perception and learning curve and determine the appropriate goal, time and level of achievement. AI-powered learning modules are also not indifferent to the motivation needs of students (Mena-Guacas et al., 2023). The fact that success takes time and patience is clearly encoded, and can be conveyed to the young people who will spend the most time with them. To illustrate, Carnegie Learning's AI-powered math modules, which are designed to develop an individual's problem-solving skills, not only provide personalized feedback according to the individual's strengths and weaknesses, but also calibrate the content, pace and level of difficulty for the next lesson. Such software adapts to each customer's learning speed and prepares training and exercise variations rich in content and suitable for their speed. Since this would be very difficult and laborious to do in school, teachers can use software such as Knewton, Carnegie Learning and ALEKS to track their students' learning pace by looking at ITS data outside of class, and use AI to determine what they should focus on in their lessons (Seufert et al., 2021).

3.3. Teacher's function in the use of AI in assessment

AI-based assessment has been rapidly integrated into education in order to offload teacher work onto machines and free up time for out-of-class tasks such as lesson planning, content preparation, activity enrichment, change-led development, and greater attention to student needs. Meanwhile, it cannot be assumed that teachers have no role or responsibility in AI-based assessments. Therefore, the fact that technology has replaced classical assessment methods

and that AI tools are used for diagnosis, measurement and feedback does not mean that teachers are obsolete, have lost their function and that machine learning has taken over the whole stage (Dillenbourg, 2016). Instead of the pointless debate about whether artificial intelligence will change the role or even replace the teacher in the classroom, which was fashionable once, we should think much more about how this technology can be used optimally and how it can make teachers' lives easier so that they can focus more on providing quality education, which is their main focus (Hrastinski et al., 2019). Thus, there are some critical tasks that teachers will undertake during this transformation (Owan et al., 2023):

1. Defining the objectives
2. Designing the content
3. Informing the student
4. Control the process
5. Thinking about how it could be better
6. Individualize teaching
7. Monitoring progress
8. Small touches of critical thinking
9. Identifying students who are struggling
10. Calibrating test content and level
11. Ensure accuracy
12. Provide feedback

In fact, educational programmers are responsible for designing educational content together with material developers and for announcing learning outcomes. In this way, teachers clearly understand what is expected of them and can use various testing tools and ideas to measure the existence of the identified learning outcomes. Involving students in such decisions not only increases the validity, acceptability and reliability of the assessments, but also motivates students to a higher level as their own ideas are considered. While AI provides speed and accuracy in the application, evaluation and feedback of the prepared tests, the clear understanding of this evaluation and feedback by the students, the transformation of this evaluation and feedback into an action plan for the future, in other words, the use of exams to guide education is a complex and multidimensional process that machine learning cannot overcome, but teachers can. (Yolcu, 2024). In fact, teachers can benefit from AI at this stage, they can benefit from the machine learning database to evaluate the performance of individuals and provide them with special feedback by identifying their strengths and weaknesses (Akyel & Tur, 2024). As a result, it is clear that the use of AI will change the role of the teacher both inside and outside the classroom, and the teacher will take on the role of planner, diagnostician and evaluator in addition to being an instructor. (Çam et al., 2021). These increased responsibilities for teachers may seem like more work, but when one considers the time to be saved from tasks such as preparing exams, evaluating and giving feedback using AI, he/she will understand that this exchange is not a bad idea at all.

The most critical role in the use of AI in education lies with school leaders and administrators. It is essential to ensure that the artificial intelligence tools used to determine student performance are unbiased, fair, valid and consistent, that the data collected from students is protected with a high level of security, and that these files are never used for purposes other than educational issues. Another issue that institutions moving to the use of AI and teachers adopting this technology should be aware of is that the software, service providers and websites they are served by do not use the data they collect without the knowledge of the users and do not share it with third parties or organizations. Care should be taken to ensure data security, and consideration should be given to why free service providers or websites offer this service for free, and possible misuse should be considered in advance (Lin et al. 2021). Teachers can benefit from the output of AI-created tests to re-shape or tailor education content according to students' learning needs (Yolcu, 2024). If, for example, there are short quizzes, true/false tests or fill-in-the-blank questions on an AI-based page for measurement and assessment purposes, and if the advertisements of the pages that students can see or even accidentally click on contain objectionable content or redirect to websites, these providers should not be used and should not be held responsible for any problems that may arise (Yolcu, 2024).

4. DISCUSSION & CONCLUSION

This review study investigates different uses of AI-based assessment tools in educational environments, the opportunities and problems they present, the changing facets of learning and teaching, and the ethical dimension of the whole process with its future implications for our children. The findings of the research have led to the conclusion

that AI tools can respond more effectively to students' expectations, thanks to the rapidly increasing technological possibilities, and represent an important opportunity in terms of providing faster, cheaper and more varied assessment options.

It has also been shown that the need for feedback, which we can see as an essential part of the assessment process, can be met both more quickly and in more detail, and with less manpower and time, thanks to this technology. If we look at the quality of work of AI-based tools in measurement and assessment in the field of education, the first thing we can say is that, thanks to AI, much more valid assessment of success can be made compared to classical methods in terms of validity, and considering the fact that more tests are made, taking into account the number of tests used in the process, their variety and predictability, AI-based tools can be more advantageous, provided that they include other components such as reliability and usability.

In addition to test validity, another important criterion in the evaluation and comparison of assessments is the need for reliability. It can be assumed that achievement tests created by AI tools using their own algorithms can give very good results in terms of reliability, due to the fact that the test content and item difficulty are written in such a way as to give close values to each other. Assuming that achievement tests administered at different times are administered with a fixed ability level (theta) and a personalized criterion, the hypothesis that the measured traits can be tested more reliably is an assumption that is also accepted by Item Response Theory and can be considered highly rational. In addition, the economic, simple and accessible products that AI tools can offer to increase the accuracy, appropriacy and efficiency of the assessment process have shown that this type of measurement is superior to traditional methods in terms of usability.

It has also been concluded that teachers with busy schedules and workloads can reduce this burden by using tests prepared with the help of artificial intelligence, and in addition to making more varied and faster assessments, they can also benefit from the advantages of this technology in terms of providing feedback. This advantage was also pointed out underlying the fact that AI tools can be used not only to give effective feedback but also to support needy learners. Meanwhile, the teacher is expected to know and plan how much work each pupil needs, how intensively and how often they will do it, like a football team coach. The results of the research show that this new and powerful weapon, AI, when used correctly and in moderation, can make a very significant contribution and offer great solutions to the changing and ever diversifying and renewing educational needs of students.

This transformation has also proven the need for educators with sufficient digital measurement and evaluation knowledge and skills to be able to select, use and develop AI tools (Suh & Ahn, 2022). As Vasconcelos and dos Santos (2023) stated in their study, artificial intelligence tools should be provided to students to demonstrate upper-level cognitive skills like reasoning, reflective and critical thinking, to reveal their creativity potential, to solve problems they may encounter in daily life on their own without getting help, and to identify their own abilities and use them effectively. Huang et al. (2018) also mentioned similar needs in his study and said that the integration of artificial intelligence into assessment is an opportunity that can simultaneously offer versatility and diversity that enables skill-focused learning designed for the individual.

On the other hand, significant shortcomings and problems have been identified in the use of AI tools in assessment and evaluation. The most prominent of these can be listed as technological competence, prejudice against change, ethical problems, lack of transparency, software and system costs, integration with classical methods used, infrastructure problems and user motivation. Among these problems, the one that needs the fastest solution can be considered ethical, because using an assignment, a paper or an article produced by an AI without declaring that it is not original, passing an exam, documenting a competency or fulfilling a responsibility is ethically extremely wrong and constitutes a crime (Gregori et al., 2018). This is because those who can do this with their own labour, but are weaker than machine learning in this respect, will be overtaken and considered less successful, which is both unfair and a violation of rights.

The results of the study suggest that in order to have an effective, fair and authentic assessment in education and training, the relevant institutions should regulate the use of such software by law, monitor it and clearly explain to the public what will be done in case of violation. Moreover, it should not be forgotten that the problems encountered by both students and teachers in using AI tools, and the studies on how to solve them, are actually seen as steps in development, and that all these difficulties and efforts will pave the way for the emergence of much more efficient evaluation platforms. To support these efforts, relevant ministries, universities and private initiatives should provide material and moral support to the process, and organizing motivational activities for teacher training and content development will accelerate the process and facilitate efficiency.

Looking at the possibilities of AI-based tools from the perspective of an assessment specialist rather than an educator, it is possible to see how seriously these applications can contribute to the field and how quickly this branch of science can develop and become a much more interdisciplinary field thanks to machine learning. Thanks to this

technology, it will be possible to develop more robust test items, calculate their internal reliability coefficients more quickly and effectively, and make ability estimates much more realistic and evidence-based.

Another advantage is that feedback, which is an indispensable link in the effectiveness of assessment, is provided by these tools in a more targeted, rapid and self-checking way (Nazaretsky et al., 2022). The evolution of these developments is highly relevant to the perspectives of assessment professionals on AI-assisted assessment and evaluation tools and methods. If there is a positive attitude and everyone understands the importance of innovation, it will be easier to see how many areas AI evaluations can be useful in the near future. Eliminating human evaluation errors and biases, and creating designs and algorithms that allow for absolute objectivity in assessment and evaluation, is an opportunity to give back the time and energy that people spend making decisions, so they can do other things better.

References

[1] Adıgüzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring thetransformative potential of ChatGPT. *Contemporary Educational Technology*, *15*(3), 1-13.

[2] AI Index Annual Report (2019). Artificial Intelligence Index Report 2019. https://hai.stanford.edu/sites/g/files/sbiybj10986/f/ai_index_2019_report.pdf

[3] Akyel, Y. & Tur, E. (2024). Eğitim bilimlerinde yapay zekânın potansiyeli ve beklentiler, zorluklar ve gelecek yönelimleri. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 25*(1),645-711.

[4] Çam, M. B., Çelik, N., Turan-Güntepe, E., & Durukan, Ü. G. (2021). Öğretmen adaylarının yapa zeka teknolojileri ile ilgili farkındalıklarının belirlenmesi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 18*(48), 263-285.

[5] Çukurova, M., & Luckin, R. (2018). Measuring the impact of emerging technologies in education: Apragmatic approach. In J. Voogt, G. Knezek, R. Christensen, and K. Lai (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education* (s. 1-19). Springer

[6] Dillenbourg, P. (2016). The evolution of research on digital education. *International Journal of Artificial Intelligence in Education, 26*(2), 544-560.

[7] Ekiz, D. (2015). *Bilimsel araştırma yöntemleri.* Anı Yayıncılık.

[8] Fabian, W. (2019). The Artificial Industry and Global Changes. https://www.forbes.com/sites/cognitiveworld/2019/11/27/the-artificial-intelligenceindustry-and-global-challenges/#34ce06603deb

[9] Gregori, P., Martínez, V., & Moyano-Fernández, J. J. (2018). Basic actions to reduce dropout rates in distance learning. Evaluation and Program Planning, 66, 48– 52. https://doi.org/10.1016/j.evalprogplan.2017.10.004

[10] Halaweh, M. (2023). ChatGPT in education: strategies. *Contemporary Educational Technology*, 15(2).

[11] Heintz, F. (2021). Three interviews about K-12 AI education in America and Singapore. *KI—Künstliche Intelligenz, 35*(2), 233-237.

[12] Hrastinski, S. O., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., Jaldemark, J., & Utterberg, M. (2019). Critical imaginaries and reflections on artificial intelligence and robots in post-digital K-12 education. *Post-Digital Science and Education, 1*(2), 427-445. https://doi.org/10.1007/s42438-019-00046-x

[13] Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society, 26*(1), 112-131.

[14] Karasar, N. (2005). *Bilimsel araştırma yöntemi.* Nobel Yayın Dağıtım.

[15] Karsenti, T. (2019). Artificial intelligence in education: The urgent need to prepare teachers for tomorrow's. *Formation et Profession [Education and Profession]*, 27(1), 112-116.

[16] Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An artificial intelligence approach to monitor student performance and devise preventive measures. *Smart Learning Environments, 8*(1), 1-18.

[17] Lin, P. Y., Chai, C. S., Jong, M. S., Dai, Y., Guo, Y., & Qin, J. (2021). Modelling the structural relationship among primary students' motivation to learn artificial intelligence. *Computers and Education: Artificial Intelligence,* 2(100006). doi: https://doi.org/10.1016/j.caeai.2020.100006

[18] Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 43-60.

[19] Mena-Guacas, A. F., Urueña Rodríguez, J. A., Santana-Trujillo, D. M., Gómez-Galán, J., & López-Meneses, E. (2023). Collaborative learning and skill development for educational growth of artificial intelligence: A systematic review. *Contemporary Educational Technology*, 15(3).

[20] Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology, 53*(4), 914-931. doi: https://doi.org/10.1111/bjet.13232

[21] O'Leary, Z. (2017). The essential guide to doing your research project. SAGE Publications Inc.

[22] Owan, V., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *EURASIA Journal of Mathematics, Science and Technology Education, 19*(8).

[23] Öz, H. (2014). Turkish teachers' practices of assessment for learning in English as a foreign language. *Journal of Language Teaching*, 5, 775-785.

[24] Pirim, H. (2006). Yapay Zekâ. Journal of Yasar University. 1(1), 81-93

[25] Ren, X. (2020). Artificial intelligence and depression: How AI powered chatbots in virtual reality games may reduce anxiety and depression levels. *Journal of Artificial Intelligence Practice, 3*(1), 48-58.

[26] Seufert, S., Guggemos, J., & Sailer, M. (2021). Technology-related knowledge, skills, and situation and emerging trends. *Computers in Human Behavior, 115*(106552). doi: https://doi.org/10.1016/j.chb.2020.106552

[27] Steinbauer, G., Kandlhofer, M., Chklovski, T., Heintz, F., & Koenig, S. (2021). A differentiated discussion about AI education K-12. *Kunstliche Intelligenz, 35*(2), 131-137.
[28] Stiggins, R. (2014). Improve assessment literacy outside of schools. *Phi Delta Kappan*, 96(2), 67-72.
[29] Suh, W., & Ahn, S. (2022). Development and validation of a scale measuring student attitudes. doi:https://doi.org/10.1177/21582440221100463
[30] Tsagari, D. (2011). Investigating the 'assessment literacy' of EFL state school teachers in Greece. In D. Tsagari, & I. Csépes (eds.), *Classroom-based language assessment*, (169-190). Frankfurt am Main: Peter Lang,
[31] Tuncer, M. & Gezer-Şen, B. (2018). Okul öncesi eğitimde toplumsal cinsiyet eşitliğine yönelik Türkiye'de yapılan çalışmalar. *Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi (ASEAD)*, 5(12), 870-882.
[32] Vasconcelos, M. A., & dos Santos, R. P. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *EURASIA Journal of Mathematics, Science and Technology Education, 19*(7). doi: https://doi.org/10.29333/ejmste/13313
[33] Xia, Q., Chiu, T. K., Lee, M., I., T., Dai, Y., & Chai, C. S. (2022). A self-determination theory design approach for inclusive and diverse artificial intelligence (AI) K-12 education. *Computers& Education*,*189*(104582). doi: https://doi.org/10.1016/j.compedu.2022.104582
[34] Yolcu, H. H. (2024). Yapay genel zekâ çağında öğretmen rolünün yeniden tanımlanması: öngörüler. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi, 10*(1), 155-167.

English Coursebook Evaluation

Murat POLAT [a*]

[a] *Anadolu University, Education Faculty, Assessment & Evaluation, Eskisehir, Türkiye, mpolat@anadolu.edu.tr*

Abstract

The purpose of this study is to analyse the data to be collected by using a rubric to be developed to determine the functionality of foreign language activities in the 10th grade English textbook used by the Ministry of National Education and their effect on language teaching by using the Multifacet Rasch Model. The study will focus on the extent to which the skills framework and the holistic educational approach, which are emphasized in the "Turkish Century Education Model", are reflected in the content of the book. The 10 units of the book subjected to the study will be evaluated by an evaluation board consisting of a total of nine people, five experts from foreign language education, two experts from curriculum development and two experts from measurement and evaluation disciplines. These experts will examine the content of the book with the "Coursebook Evaluation Rubric" to be developed for this study. The main purpose of the study is to determine how effective the foreign language activities in the textbook will be in terms of their usefulness and difficulty levels, and how effective they will be in terms of holistic educational goals that will enable the acquisition of skills and the utilization of different skills together. In addition, the functionality of the category levels of the language activity criteria in the rubric to be developed and whether they are at a statistically acceptable level will be tested. Finally, in the light of the results compiled from the study, it will be analysed how effective and explanatory the MFRM can be in empirical studies with multiple surfaces and where each surface is to be examined in detail.

Keywords: lanuage teaching, coursebook evaluation, MFRM, evaluation Rubric

1. INTRODUCTION

Most of the English books used in Türkiye have remained in mind as course materials that were brought from abroad, marketed with significant amounts of money and serious difficulties, used in schools and then thrown away after a certain period of time without properly measuring their effectiveness. There were no significant initiatives in this regard until almost the beginning of the 2000s, and instead of writing its own foreign language education books, Turkey preferred to buy foreign language resources from other countries, prepared course programs and contents based on their content, and paid the price for this, especially with the failure it experienced in English education. In the last 20 years, especially as a result of the efforts of the Ministry of National Education (MEB), foreign language books of the quality and content that can be taught in public schools have been written, and the country has gained its own resources and a pool of materials that it can choose in this sense. The purpose of this research is to examine the data to be collected with the scale that will be developed in order to determine the functionality of the foreign language activities in the 10th grade English course book used by the MEB and its effect on language learning, using the Multi-Surface Rasch Model (MSRM). The research particularly focused on how much the skills framework and holistic education approach, which are emphasized in the "Turkish Century Education Model", are reflected in the content of the book. The 10 units of the book in question will be evaluated by a nine-person evaluation board consisting of three experts from each discipline of foreign language education, program development and measurement and evaluation. The experts in question will examine the content of the book using the "Textbook Evaluation Scale" to be developed

* Corresponding author.

for this study and will determine to what extent the different foreign language activities in the book will contribute to the use and development of different language skills.

Considering the impact of technology on education, significant progress has been made in children's education, especially in the last 20 years (Fırat, 2020; Tuncer & Şen, 2018). One of these innovations is undoubtedly artificial intelligence, which we encounter in almost every field. Artificial intelligence has become a transformative force in education thanks to the many conveniences it promises. This new era, which can be described as revolutionary, includes technologies such as natural language processing (NLP), raw data analytics and machine learning, which have the potential to take teaching and learning to a new level. By harnessing the power of artificial intelligence, educators can rapidly prepare classroom material, test content, create multi-dimensional feedback, and gain deeper insights into individual learning needs. This ability to use and process data allows educators to customize their teaching, provide targeted interventions and support each student as effectively as they can on their unique learning journey (Pirim, 2006).

It is quite rational and practical to use these innovative technologies in assessment and evaluation processes in education, which are sometimes perceived as the last ring and often as a process that takes place after education rather than as part of it. According to Stiggins, assessment and evaluation, which might take longer time than it is expected, could lead to an important workload on which teachers spend almost a third of their professional time (2014). This situation can be presented as an evidence of the necessity to integrate technology into educational practices in particular, and therefore, the critical role of employing artificial intelligence in assessment and evaluation, together with other technological developments in teaching and learning. Research has shown that despite their critical importance, teachers generally need more support and training than they are currently being provided as well as a new generation of assessment and evaluation tools/methods that focus on artificial intelligence-based assessment and evaluation (Kafadar, 2022; Mede & Atay, 2017; Tsagari, 2011).

Due to the worldwide coronavirus pandemic in the early 2020s, there was a sudden and obligatory shift to distance learning in most parts of the world, and most courses were provided via online platforms. This situation brought with it the need for artificial intelligence-based digital assessment and evaluation, which was a challenging experience for teachers who were not familiar with these practices. Although teachers have sufficient knowledge of the concepts and applications related to assessment and evaluation and know how to use tests effectively to improve teaching and learning, this does not mean that teachers are effective evaluators in the digital environment. Therefore, it has become essential for teachers to have skills related to measurement and evaluation applications carried out in the digital environment. All in all, it is very important to determine the measurement and evaluation applications carried out in the education and training environment by using tools and methods supported by artificial intelligence, that is, to keep up with the times, save time in terms of evaluation, increase interaction, and be able to make individual decisions. Meanwhile, when we consider the development of AI in numbers, we can see a tremendous increase in investments and research projects. To reiterate, technology is developing faster than ever, with developments such as fifth-generation mobile internet connectivity, augmented reality and block chain (Arslan, 2020; Marr, 2019). It can be said that artificial intelligence (AI) technology, which is increasingly impacting the stages of obtaining, processing, reporting and using information worldwide, is one of the key technological developments contributing to this change. For this reason, investment in AI technology research in universities and the private sector has gained considerable momentum. In 2019, private AI investments in the global arena exceeded $70 billion. By 2020, the Massachusetts Institute of Technology (MIT) alone had invested $1 billion in new AI programs, while Microsoft had hired over 8,000 AI researchers (Fabian, 2019). Although the demand for AI technologies is skyrocketing day by day on a global scale, it can be argued that the education system, including higher education institutions, is not adequately funded and learners are not adequately trained to equip them with digital skills and AI technologies. Therefore, it is believed that learners are not sufficiently exposed to innovative technological changes in education (Heintz, 2021), including AI technology. Therefore, in the new knowledge-based society, it is becoming harder for higher education graduates to meet the increasing and changing demands of the labor market.

Given the significant growth of AI and its application in education, a number of questions should be asked and answered in order to provide a better view of the role of this technology in pedagogical issues.

1. What are the most common AI-generated assessment and evaluation techniques?
2. What are the most important advantages and disadvantages of using AI-generated assessment and evaluation techniques?
3. What is the role of the teacher using AI-produced assessment and evaluation techniques?

The literature included in this review study was addressed within the framework of these research questions, and answers to these questions were carefully sought in the studies focusing on the issue.

2. METHODOLOGY

A document review was carried out in order to access the latest research data, to bring together the relevant discussions on the subject and to explore the details more comprehensively. Document review or analysis, also known as documentary scanning, is the process of examining existing records and documents to obtain data. Document analysis involves reading and reporting the information and documents dealing with a specific problem or question (Karasar, 2005). This process sometimes involves compiling and examining written works such as books, journals and newspapers, and sometimes compiling and reading documents published in electronic media, which is a meticulous process that requires extensive work. Ekiz describes this whole process as the collection and processing of information and documents that might be used the for the research (2015). This preparation can sometimes be done by collecting documents that are accessible to everyone, while sometimes it can be done by a researcher who is allowed to access very specific information and documents. According to O'Leary (2017), document analysis is the process of gathering, analyzing, weighing, and assessing data and documents that will help the researcher to avoid wasting research resources. In short, this type of research involves collecting and synthesizing information, documents and studies that have been written, conducted or presented to the public by other people or institutions in relation to the research problem they are focusing on. Similarly, the stages of document analysis have been approached in different ways by researchers. For example, O'Leary (2017) lists these stages as:

(1) planning for all eventualities;

(2) collecting documents,

(3) checking the reliability of the documents and questioning their accuracy,

(4) analyzing the data.

On the basis of the different stages mentioned here, one can conclude that document analysis is carried out according to a specific plan, and appropriate methods should be employed to obtain, examine and analyze documents suitable for the purpose (Altunışık et al., 2010).

3. RESULTS

The literature was reviewed according to the research questions identified in the study and each research question was discussed in detail under separate headings and supported by examples from the literature.

3.1. What are the measurement and evaluation tools and methods used in education with artificial intelligence?

Artificial intelligence is changing the perception of education, making teachers and students ask the same question: “How much should we know?” (Papapicco, 2020). It is crystal clear that the use of AI tools in various educational settings will help teachers improve their professional quality and increase the pace of learning, if properly implemented. To illustrate, Halaweh believes that the growing importance of AI has attracted the attention of researchers, who are actively investigating alternative methods to integrate a range of technological tools into the classroom in order to promote sustainable learning (2023). Thus, with the growing amount of educational data and the increasing pace of machine learning codes, the intensive use of AI has the potential to transform the way one learns, uses and checks his/her skills. As such, the benefits of using AI tools in testing are numerous. For example, Huang et al. (2023) argue that the use of AI for administrators, teachers, and students can be advantageous in developing critical thinking, gaining new insights, and fostering imagination. It has also been reported that artificial intelligence can play a crucial role in motivating students (Xia et al., 2022), increasing their participation in class (Lin et al., 2021), improving learner interaction (Kumar, 2019), reducing test anxiety (Hou et al, 2022; Ren, 2020), testing learning outcomes (Karsenti, 2019; Luo et al., 2022), monitoring real-time academic performance (Khan et al. 2021), and stimulating greater interest in learning (Haves & Arya, 2023).

If AI-based assessment kits are designed in accordance with the curriculum, they save time and opportunity in determining much more economical measurement and evaluation tools to measure goals and objectives, and in designing new tools (Khan et al., 2021). Where, when and how AI-based assessment kits can be used in measurement

and evaluation has been discussed in detail by many researchers in the literature (Çukurova & Luckin, 2018; Holmes et al., 2019; Karsenti, 2019):

- Computer-based testing platforms
- Formative assessment tools
- Digital assessment tools
- Learning management systems
- Plagiarism detection
- Predictive analytics tools
- Digital portfolios
- Gamification tools
- Automated grading tools
- Data visualization tools
- Learning analytics tools

Considering potential usefulness of AI in scoring and providing feedback on classical exams, we can say that the use of AI has the potential to become an indispensable assistant that can save time for language teachers, especially by allowing them to speed up and automate the marking process and providing students with fast and accurate feedback on their work (Öz, 2014). Artificial intelligence can analyze drafts, exam papers and even project reports to provide feedback regarding organization, vocabulary richness, grammar and mechanics in many languages. By using more AI systems in assessment, teachers can focus more on some basic educational needs, such as lesson planning, material development, and student counselling, and can save significant amount of time (Adıgüzel et al. 2023). To illustrate, the AI that powers Turnitin, a very popular homework submission program works through the NLP algorithms to analyze learners' written output and deliver feedback on vocabulary, grammar and spelling (İşler & Kılıç, 2021). In addition, the same software can detect plagiarism by comparing different texts, allowing graders to grade assignments more quickly and accurately.

AI can also provide predictive analysis of learner achievement by incorporating student participation in classes, mid-term or end-of-term test data, and summative assessments. In addition to predicting achievement, these analytics can be used to identify students who might require extra help or guidance, enabling teachers to intervene in a timely and well-informed manner (Owan et al., 2023). Learning Management Systems (LMS), where AI is most commonly used in this sense, are platforms that allow experts to create, present and develop supporting materials, assignments, quizzes and feedback tools for students. With a little training, all teachers can choose the right LMS system, track their students' progress, and use these platforms to receive feedback on both student and teacher performance at the end of the process. Some popular examples include Moodle, Canvas and Blackboard (Suh & Ahn, 2022).

Another medium where AI can be used effectively is in learning analytics tools. Owing to this technology, students can analyze their own learning speed and have the ability to self-assess their performances. The system uses data mining and artificial intelligence algorithms to give members insights into their participation, retention, interaction and own learning performance. Experts may prefer learning analytics tools to track the progress of their students, identify slow learners or those experiencing some problems, and use existing data to improve the level of performance tested periodically (Önder et al., 2023). Computer-based testing (CBT) software applications allow educators to benefit from many different assessment methods, such as true/false questions, matching, multiple choice, fill-in-the-blank items online or face-to-face. CBT platforms can be recommended to assess learner competence, save money and time in assessment, and provide quick and accurate feedback to learners, especially in summative assessment applications (Bassey et al., 2020). JAMB CBT, ExamSoft, ProProf, UNICAL Postgraduate e-exams and Questionmark can be mentioned among such educational platforms.

Unlike summative testing, formative assessment tools allow teachers to determine student performance with reliable data, provide feedback based on the results obtained, and enable timely development or modification of the program or materials thanks to real-time measurement. The most common formative testing applications include Socrative, Nearpod and Mentimeter. In addition to these testing platforms, digital portfolios that students can use allow them to store and view their writing at any time, self-critique their own learning pace and effectiveness, and receive feedback from their teachers and peers in their classes. Examples of these digital portfolios include WordPress and Google Sites. Another AI opportunity, data visualization tools, allow experts to analyze and report on data from testing educational outcomes and create visual tools such as graphs, figures, and charts needed to present these learning outcomes (Owan et al., 2023). Examples include Tableau, Infogram and Google Data Studio.

Another way we can use AI to measure student performance is through game-like tools. These tools use attractive elements to stimulate children by engaging them in learning activities more, making learning fun and giving them a

sense of competition. Educators, especially those working with younger age groups, can use these fun tools to increase children's participation in activities, make them more active learners, and give them instant feedback on their strengths and weaknesses (Türker, 2023). Examples include Quizlet, Classcraft and Kahoot. Finally, to prevent ethical breaches, AI-based plagiarism detection systems have become modern methods in many academic institutions, where NLP algorithms are widely used to check assignments and detect possible plagiarism. Among such systems are Ithenthicate, Grammarly, Turnitin and Copyscape.

Teachers can then use AI systems to customize their exams and assignments, quizzes or interactive course content, saving a lot of time. While traditional assessment and measurement are affected by human error and bias, AI systems generally do not have such limitations. As a result, AI tools are quite advanced in what they can do in a limited amount of time and are completely unbiased in terms of reliable grading and personalized feedback.

3.2. Benefits and Challenges of Assessment and Evaluation Using Artificial Intelligence in Education

Developments especially in the last two decades reveal that artificial intelligence technology has assumed important roles in areas that everyone is interested in such as architecture, economy, communication, sports and politics, and has become indispensable for people who are familiar with this convenience. Realizing this transformation, the superpowers such as the United States, Russia, China, the United Kingdom and the European Union countries are integrating artificial intelligence-based and even artificial intelligence education into their K-12 curricula and consider it as a component of their basic education systems (Chiu, 2021; Steinbauer et al., 2021). Discussions regarding where artificial intelligence can be used most effectively suggest that there is an interdisciplinary connection due to the fact that the layered structure of this technology interconnects many different fields and thus, when used well, two targets can be hit by one arrow. For this reason, artificial intelligence education gives the opportunity to challenge what the human brain loves the most and can provide interested young minds with the opportunity to develop metacognitive skills such as reasoning, critical thinking, mathematical modelling and hypothetical problem solving.

On the other hand, compared to traditional educational methods, tools and training based on artificial intelligence offer the possibility of customizing learning environments, increasing the efficiency of education and extending the duration of learning. The use of artificial intelligence makes it possible to monitor learning speed in real time, and in this way it is possible to prepare personalized learning content for students whose needs we determine by analyzing their development, what they do well and what they do poorly. To do this, artificial intelligence can individually examine data generated at many different times, such as quizzes, homework and projects, to determine the individual's real-time learning path and provide feedback for improvement (Adıgüzel et al. 2023). Thanks to AI, people can learn at their own pace, speeding up or slowing down as they wish, and if they need help in other areas, this can be done quickly and economically. Today, AI-enabled personalized learning software such as Knewton or DreamBox, which have been created to meet such educational needs, are being used to create learning plans and content for their customers' educational needs at different speeds and with different content (Owan et al. 2023). In this way, students can experience adaptive learning through content and tests designed specifically for them, and progress without stress.

AI-based tutoring systems (ITS) are ideal for extracurricular educational tasks that normally take up a lot of a teacher's time, such as aiding or feedback. These systems not only provide educational services in the background, but can also identify an individual's perception and learning curve and determine the appropriate goal, time and level of achievement. AI-powered learning modules are also not indifferent to the motivation needs of students (Mena-Guacas et al., 2023). The fact that success takes time and patience is clearly encoded, and can be conveyed to the young people who will spend the most time with them. To illustrate, Carnegie Learning's AI-powered math modules, which are designed to develop an individual's problem-solving skills, not only provide personalized feedback according to the individual's strengths and weaknesses, but also calibrate the content, pace and level of difficulty for the next lesson. Such software adapts to each customer's learning speed and prepares training and exercise variations rich in content and suitable for their speed. Since this would be very difficult and laborious to do in school, teachers can use software such as Knewton, Carnegie Learning and ALEKS to track their students' learning pace by looking at ITS data outside of class, and use AI to determine what they should focus on in their lessons (Seufert et al., 2021).

To prevent mechanical and spelling errors, NLP tools can provide feedback on correct spelling, grammar and punctuation rules, helping students to make fewer mistakes when preparing homework and to both see and learn from their mistakes in situations that require writing skills. Tools that use NLP reinforce students use more cognitive skills while studying and they also provide feedback where necessary, allowing students use their critical thinking skills better. This technology contributes to students' intellectual activities such as production and writing by analyzing and interpreting not only homework, but also extracts from everyday life, such as students' various social media posts

(Huang et al., 2023). For example, Grammarly software provides its users with very important writing support, including suggestions on sentence formation, grammar, spelling and punctuation thanks to NLP. This allows students to write with less pressure and more support.

Artificial intelligence can determine students' social preferences by using data collected from their past activities on the Internet, and thus create more engaging and user-specific educational content (Türker, 2023). For example, software called Smart Sparrow uses artificial intelligence to create learning modules that are tailored to students' individual needs and can be adapted to different contents. For example, a chemistry teacher can use AI to create different scenarios that can be done in the lab according to students' hobbies and interests.

Machine learning and AI can also be used in virtual assistant services to help students maintain discipline and schedule. Thanks to AI, students can stay more motivated, organized and focused during their busy study schedule, use their time more efficiently and thus study more effectively (Kim & Park, 2019). AI tools such as Brainly, where machine learning is used at the highest level, can be very useful in helping students organize and discipline their studies.

Another issue is where teachers will get the feedback they need. AI-based measurement and evaluation tools can help teachers determine how effective their teaching is, what they should spend more time on, where they should repeat, and what they should pay attention to. When it comes to measurement and evaluation, they can take a lot of burden off teachers by providing practical, fast and reliable measurement tools. AI-assisted diagnostic tools, with hundreds of examples in education, can help teachers evaluate students' academic success, guide teachers in terms of what students know and how much they know in a more general sense starting from basic and moving to the more complex (Huang et al., 2023).

Artificial intelligence-based educational software can also be beneficial when students fall behind in attendance and absenteeism (Delgado et al. 2020). These tools can track school records, exam results, pass and fail status of absent students, and can be used as an early warning system in cases where intervention is needed. With all these opportunities being used so widely, and with AI providing personalized educational opportunities, control and monitoring opportunities, it may be possible in the near future to move the educational environment from schools to other places, creating educators produced by machine learning rather than teachers, and providing optimal and effective educational opportunities where the true potential of the human mind can be revealed (Türker, 2023).

On the other hand, there are also a number of backwash effects of artificial intelligence in education. For example, it should be mentioned that students copy and paste text from some sources without making any changes or adding anything of their own, neglecting to refer to it with appropriate methods when quoting, which inevitably leads to academic plagiarism (Halaweh, 2023). The following is a list of back-wash effects observed when AI tools are used in assessment;

- Cost
- Inadequate training
- Prejudice
- Student motivation and participation
- Technical difficulties
- Ethical concerns
- Habits
- Transparency
- Feedback and support
- Stakeholder input
- Resistance to change
- Human role
- Standardization
- Limited scope
- Data management
- Integration with existing systems

In the design and development of AI tools to be used in measurement and evaluation activities in education, the lack of participation or insufficient participation of other stakeholders, such as educators and programmers, parents, students or employers, can be a significant problem. Developing AI tools without this feedback can lead to problems such as non-adoption, lack of interest and lack of trust. Think about a case what a teacher using one of these tools in the classroom would do if he or she received criticism from students that the software is inadequate, too easy or too difficult during his/her preparation for the class. Even some AI software developers often do not take into

consideration the opinions and expectations of students, who are the most important and real users of artificial intelligence (Çukurova & Luckin, 2018; Seufert et al., 2021). Therefore, if we want to give more space to the AI tools in assessment and evaluation at schools, more feedback should be obtained from all stakeholders, starting with students (Holmes et al., 2019).

One of the problems with the use of AI in educational measurement is the lack of accountability and transparency of the process. In the preparation and scoring of a test, it may not be easy to approve the data used by artificial intelligence algorithms to create test items and say that the results are reliable, and this paradox can undermine stakeholder confidence in the accuracy and impartiality of AI decisions. Just like human intelligence, machine learning can be biased, which can lead to hasty, biased and even incorrect assessments. The reason for this problem is that machine learning is only as good as what it is taught. If there is a bias or distortion in the data it is fed, the software will naturally do the same (Owan et al. 2023).

Moreover, another issue that arises with the use of AI is that the removal of the human factor from the measurement and evaluation part of the educational process can have an impact on the student. When obtaining test scores and measurement results, some critical data, clues or details that only human intelligence can notice may be missed when AI performs scoring. Given the type of items used and the variety of questions, tests scored by AI tend to include questions that measure low-level cognitive skills, which limits what can be measured and scored. However, this is a rare situation when a human rater is used, as imagination, creativity and many of the knowledge and skills we can use to solve problems in everyday life can be measured using classical measurement methods such as interviews, debates and problem solving (Türker, 2023).

Another problem encountered in the use of AI in assessment is ethical issues such as data storage, sharing and rights violations, as well as comparing what AI tools can do with what human raters can do, which can even create an ethical problem. Prospective and technical capacity can also be seen as a problem because teachers who will agree or disagree on having their exams done by AI may lack different skills such as understanding the functioning of AI algorithms and observing their reflections on the assessment of learning (Aşık et al. 2023).

Teachers who adopt the use of AI for assessment should receive the necessary training, just as they would do to drive a car. You can arrive at your destination much faster, more comfortably and safely with a car if you know how to use it and follow many rules when using it. The benefits of measuring with artificial intelligence may not be endless (Nazari et al., 2021), but expert opinion should be sought on how to integrate them into the system in which they will be used and how to use them in the most harmless but adaptable way.

In addition, the evaluation systems and infrastructures created by taking into account the existing measurement tools should be reviewed, and the necessary technological investments and transfers should be made to avoid problems, and this transformation should be carried out smoothly (İşler & Kılıç, 2021). This transformation may be more expensive and cumbersome than expected, and many public schools may not have the financial and physical facilities to cope with it. This problem may also lead to inequality between schools and students, which is a serious issue that needs to be considered. Another problem needs to be mentioned is the attitude of parents because parents who have a negative attitude towards these new technologies may be resistant to their children being educated using such technologies. They may prefer classical methods and not allow change and transformation, which may slow down the application of artificial intelligence measurement and disrupt the process (Şad & Göktaş, 2013).

The tools to be used in AI-assisted measurement may also reveal some issues related to aptitude for this technology. Students who are familiar with such technologies may appear to have an advantage over students who are less familiar with them during the assessment phase. AI-based assessment tools used for measurement are mostly designed to provide the measurement process through formative tests that are standardized using different methods. However, relying on standardization to minimize bias leads to a failure to consider participants' prior experience, the diversity of opportunities available to them, and other influential stakeholders, which may jeopardize the validity of the measurement (Owan et al. 2023).

Finally, there is the infrastructure and technical facilities required to use this technology. In cities or institutions where AI-assisted assessment tools are used, there should be reliable, stable energy sources and data transmission facilities. Power outages, internet connection disruptions or other technical problems that may occur during AI-based evaluation and assessment can jeopardize the process and cause data loss, which is a situation that can never be tolerated in this evaluation. When AI is used for diagnostic rather than evaluative purposes, the amount of data can double or even quadruple, so it is extremely important that this data set is obtained and delivered securely and smoothly. Those using this technology must provide assurances to all stakeholders regarding this data security (Dillenbourg, 2016).

3.3. Teacher's function in the use of AI in assessment

AI-based assessment has been rapidly integrated into education in order to offload teacher work onto machines and free up time for out-of-class tasks such as lesson planning, content preparation, activity enrichment, change-led development, and greater attention to student needs. Meanwhile, it cannot be assumed that teachers have no role or responsibility in AI-based assessments. Therefore, the fact that technology has replaced classical assessment methods and that AI tools are used for diagnosis, measurement and feedback does not mean that teachers are obsolete, have lost their function and that machine learning has taken over the whole stage (Dillenbourg, 2016). Instead of the pointless debate about whether artificial intelligence will change the role or even replace the teacher in the classroom, which was fashionable once, we should think much more about how this technology can be used optimally and how it can make teachers' lives easier so that they can focus more on providing quality education, which is their main focus (Hrastinski et al., 2019). Thus, there are some critical tasks that teachers will undertake during this transformation (Owan et al., 2023):

1. Defining the objectives
2. Designing the content
3. Informing the student
4. Control the process
5. Thinking about how it could be better
6. Individualize teaching
7. Monitoring progress
8. Small touches of critical thinking
9. Identifying students who are struggling
10. Calibrating test content and level
11. Ensure accuracy
12. Provide feedback

In fact, educational programmers are responsible for designing educational content together with material developers and for announcing learning outcomes. In this way, teachers clearly understand what is expected of them and can use various testing tools and ideas to measure the existence of the identified learning outcomes. Involving students in such decisions not only increases the validity, acceptability and reliability of the assessments, but also motivates students to a higher level as their own ideas are taken into account. While AI provides speed and accuracy in the application, evaluation and feedback of the prepared tests, the clear understanding of this evaluation and feedback by the students, the transformation of this evaluation and feedback into an action plan for the future, in other words, the use of exams to guide education is a complex and multidimensional process that machine learning cannot overcome, but teachers can. (Yolcu, 2024). In fact, teachers can benefit from AI at this stage, they can benefit from the machine learning database to evaluate the performance of individuals and provide them with special feedback by identifying their strengths and weaknesses (Akyel & Tur, 2024). As a result, it is clear that the use of AI will change the role of the teacher both inside and outside the classroom, and the teacher will take on the role of planner, diagnostician and evaluator in addition to being an instructor. (Çam et al. 2021). These increased responsibilities for teachers may seem like more work, but when one considers the time to be saved from tasks such as preparing exams, evaluating and giving feedback using AI, he/she will understand that this exchange is not a bad idea at all.

The most critical role in the use of AI in education lies with school leaders and administrators. It is essential to ensure that the artificial intelligence tools used to determine student performance are unbiased, fair, valid and consistent, that the data collected from students is protected with a high level of security, and that these files are never used for purposes other than educational issues (Köse et al., 2023). Another issue that institutions moving to the use of AI and teachers adopting this technology should be aware of is that the software, service providers and websites they are served by do not use the data they collect without the knowledge of the users and do not share it with third parties or organizations. Care should be taken to ensure data security, and consideration should be given to why free service providers or websites offer this service for free, and possible misuse should be considered in advance (Lin et al. 2021). Teachers can benefit from the output of AI-created tests to re-shape or tailor education content according to students' learning needs (Yolcu, 2024; Seyrek et al., 2024). If, for example, there are short quizzes, true/false tests or fill-in-the-blank questions on an AI-based page for measurement and assessment purposes, and if the advertisements of the pages that students can see or even accidentally click on contain objectionable content or redirect to websites, these providers should not be used and should not be held responsible for any problems that may arise (Yolcu, 2024).

Another point that should not be forgotten is that the main purpose of AI assessment tools is to evaluate students' knowledge and that the data evaluated by these tools belongs to real people, i.e. students; if it does not, it is fake and constitutes a crime (Lin et al. 2021). At this point, teachers need to ensure the accuracy of AI-created tests, providing evidence that they measure the skills they are ought to measure, and confirm that ethical rules are followed, just as with traditional exams. Decisions should not be made solely on the basis of the results of assessment using AI tools; some other verified methods should also be conducted. (Owan et al. 2023). Therefore, when using AI tools/methods, teachers should include real life problems for in-class performance assessment as well as student participation (Çam et al. 2021; İşler & Kılıç, 2021).

Last but not least, the fact that the purpose of the algorithms of the tools used for evaluation and measurement of AI tools in education is hidden may be a sign that important ethical problems may arise. There are important criticisms about the inclusion of artificial intelligence software in all kinds of evaluation systems without much calculation of their results and the lack of transparency and accountability in data processing processes (Ivanov, 2016; Rospigliosi, 2023; Rusmiyanto et al., 2023). In deciding which artificial intelligence tools to use and whether the algorithms they have will be used in the same way or whether they will undergo a series of development processes, expert help should definitely be sought and information should be obtained about their tested versions. In order to ensure auditability and transparency, data usage, coding and network formations should be monitored and reports should be requested. Because the unfair use of this data also includes the responsibility of information security, validity and reliability of the exam. It should be considered that the potential jeopardization of the data obtained from students and teachers without knowing who will use it for what purpose will lead to unreal exam results and inequalities, as well as the emergence of data obtained from students' different preferences and behaviors for different purposes.

4. DISCUSSION AND CONCLUSION

This review study investigates different uses of AI-based assessment tools in educational environments, the opportunities and problems they present, the changing facets of learning and teaching, and the ethical dimension of the whole process with its future implications for our children. The findings of the research have led to the conclusion that AI tools can respond more effectively to students' expectations, thanks to the rapidly increasing technological possibilities, and represent an important opportunity in terms of providing faster, cheaper and more varied assessment options.

It has also been shown that the need for feedback, which we can see as an essential part of the assessment process, can be met both more quickly and in more detail, and with less manpower and time, thanks to this technology. If we look at the quality of work of AI-based tools in measurement and assessment in the field of education, the first thing we can say is that, thanks to AI, as Sarıoğlu (2023) also stated in his study, much more valid assessment of success can be made compared to classical methods in terms of validity, and considering the fact that more tests are made, taking into account the number of tests used in the process, their variety and predictability, AI-based tools can be more advantageous, provided that they include other components such as reliability and usability.

In addition to test validity, another important criterion in the evaluation and comparison of assessments is the need for reliability. It can be assumed that achievement tests created by AI tools using their own algorithms can give very good results in terms of reliability, due to the fact that the test content and item difficulty are written in such a way as to give close values to each other. Assuming that achievement tests administered at different times are administered with a fixed ability level (theta) and a personalized criterion, the hypothesis that the measured traits can be tested more reliably is an assumption that is also accepted by Item Response Theory and can be considered highly rational. In addition, the economic, simple and accessible products that AI tools can offer to increase the accuracy, appropriacy and efficiency of the assessment process have shown that this type of measurement is superior to traditional methods in terms of usability.

It has also been concluded that teachers with busy schedules and workloads can reduce this burden by using tests prepared with the help of artificial intelligence, and in addition to making more varied and faster assessments, they can also benefit from the advantages of this technology in terms of providing feedback. This advantage was also pointed out by Yılmaz (2024) underlying the fact that AI tools can be used not only to give effective feedback but also to support needy learners. Meanwhile, the teacher is expected to know and plan how much work each pupil needs, how intensively and how often they will do it, like a football team coach. The results of the research show that this new and powerful weapon, AI, when used correctly and in moderation, can make a very significant contribution and offer great solutions to the changing and ever diversifying and renewing educational needs of students.

This transformation has also proven the need for educators with sufficient digital measurement and evaluation knowledge and skills to be able to select, use and develop AI tools (Suh & Ahn, 2022). As Vasconcelos and dos Santos

(2023) stated in their study, artificial intelligence tools should be provided to students to demonstrate upper-level cognitive skills like reasoning, reflective and critical thinking, to reveal their creativity potential, to solve problems they may encounter in daily life on their own without getting help, and to identify their own abilities and use them effectively. Huang et al. (2018) also mentioned similar needs in his study and said that the integration of artificial intelligence into assessment is an opportunity that can simultaneously offer versatility and diversity that enables skill-focused learning designed for the individual.

On the other hand, significant shortcomings and problems have been identified in the use of AI tools in assessment and evaluation. The most prominent of these can be listed as technological competence, prejudice against change, ethical problems, lack of transparency, software and system costs, integration with classical methods used, infrastructure problems and user motivation. Among these problems, the one that needs the fastest solution can be considered ethical, because using an assignment, a paper or an article produced by an AI without declaring that it is not original, passing an exam, documenting a competency or fulfilling a responsibility is ethically ~~extremely~~ wrong and constitutes a crime (Gregori et al., 2018). This is because those who can do this with their own labor, but are weaker than machine learning in this respect, will be overtaken and considered less successful, which is both unfair and a violation of rights.

The results of the study suggest that in order to have an effective, fair and authentic assessment in education and training, the relevant institutions should regulate the use of such software by law, monitor it and clearly explain to the public what will be done in case of violation. Moreover, it should not be forgotten that the problems encountered by both students and teachers in using AI tools, and the studies on how to solve them, are actually seen as steps in development, and that all these difficulties and efforts will pave the way for the emergence of much more efficient evaluation platforms. To support these efforts, relevant ministries, universities and private initiatives should provide material and moral support to the process, and organizing motivational activities for teacher training and content development will accelerate the process and facilitate efficiency.

Looking at the possibilities of AI-based tools from the perspective of an assessment specialist rather than an educator, it is possible to see how seriously these applications can contribute to the field and how quickly this branch of science can develop and become a much more interdisciplinary field thanks to machine learning. Thanks to this technology, it will be possible to develop more robust test items, calculate their internal reliability coefficients more quickly and effectively, and make ability estimates much more realistic and evidence-based.

Another advantage is that feedback, which is an indispensable link in the effectiveness of assessment, is provided by these tools in a more targeted, rapid and self-checking way (Nazaretsky et al., 2022). The evolution of these developments is highly relevant to the perspectives of assessment professionals on AI-assisted assessment and evaluation tools and methods. If there is a positive attitude and everyone understands the importance of innovation, it will be easier to see how many areas AI evaluations can be useful in the near future. Eliminating human evaluation errors and biases, and creating designs and algorithms that allow ~~for~~ absolute objectivity in assessment and evaluation, is an opportunity to give back the time and energy that people spend making decisions, so they can do other things better.

References

[1] Adıgüzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring thetransformative potential of ChatGPT. *Contemporary Educational Technology*, *15*(3), 1-13.

[2] Akyel, Y. & Tur, E. (2024). Eğitim bilimlerinde yapay zekânın potansiyeli ve beklentiler, zorluklar ve gelecek yönelimleri. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 25*(1),645-711.

[3] Altunışık, R., Coşkun, R., Bayraktaroğlu, & S., Yıldırım, E. (2010). Sosyal bilimlerde araştırma yöntemleri. Sakarya. Sakarya Kitabevi.

[4] Arslan, K. (2020). Eğitimde yapay zekâ ve uygulamaları. *Batı Anadolu Eğitim Bilimleri Dergisi, 11*(1), 71-88.

[5] Aşık, F., Yıldız, A., Kılınç, S., Aytekin, N., Adalı, R., & Kurnaz, K. (2023). Yapay zekanın eğitime etkileri. *International Journal of Social and Humanities Sciences Research, 10*(98), 2100-2107.

[6] Bassey, B. A., Owan, V. J., Ikwen, E. U., & Amanso, E. O. I. (2020). Teachers' attitudes towards learners with disability scale (TALDS): construction and psychometric analysis. *Psychometric Analysis, 6*(4), 518-530.

[7] Chiu, T. K. (2021). A holistic approach to the design of artificial intelligence (AI) education for K-12 schools. *TechTrends, 65*(5), 796-807.

[8] Çam, M. B., Çelik, N., Turan-Güntepe, E., & Durukan, Ü. G. (2021). Öğretmen adaylarının yapa zeka teknolojileri ile ilgili farkındalıklarının belirlenmesi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 18*(48), 263-285.

[9] Çukurova, M., &Luckin, R. (2018). Measuring the impact of emerging technologies in education: Apragmatic approach. In J. Voogt, G. Knezek, R. Christensen, and K. Lai (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education* (s. 1-19). Springer

[10] Delgado, H. O., de Azevedo Fay, A., Sebastiany, M. J., & Silva, A. D. (2020). Artificial intelligence adaptive learning tools. *BELT-Brazilian English Language Teaching Journal, 11*(2).

[11] Dillenbourg, P. (2016). The evolution of research on digital education. *International Journal of Artificial Intelligence in Education, 26*(2), 544-560.
[12] Ekiz, D. (2015). *Bilimsel araştırma yöntemleri.* Anı Yayıncılık.
[13] Fabian, W. (2019). The Artificial Industry and Global Changes. https://www.forbes.com/sites/cognitiveworld/2019/11/27/the-artificial-intelligenceindustry-and-global-challenges/#34ce06603deb
[14] Fırat, M. (2020). Öğrenci destek servislerinde doğal dil işleme: GPT-3 örneği. 12th ICoSReSSE International Conference of Strategic Research in Social Science and Education.
[15] Gregori, P., Martínez, V., & Moyano-Fernández, J. J. (2018). Basic actions to reduce dropout rates in distance learning. Evaluation and Program Planning, 66, 48– 52. https://doi.org/10.1016/j.evalprogplan.2017.10.004
[16] Halaweh, M. (2023). ChatGPT in education: strategies. *Contemporary Educational Technology*, 15(2).
[17] Hawes, D., & Arya, A. (2023). Technology solutions to reduce anxiety and increase cognitive availability in students. *IEEE Transactions on Learning Technologies, 16*(2), 278-291.
[18] Heintz, F. (2021). Three interviews about K-12 AI education in America and Singapore. *KI—Künstliche Intelligenz, 35*(2), 233-237.
[19]
[20] Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning.* Boston: Center for Curriculum Redesign. https://curriculumredesign.org/wp-content/uploads/AIED-Book-Excerpt-CCR.pdf
[21] Hou, G., Anicetus, U., & He, J. (2022). How to design font size for older adults: A systematic literature review with a mobile device. *Frontiers in Psychology, 13*(931646), 1-19
[22] Hrastinski, S. O., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., Jaldemark, J., & Utterberg, M. (2019). Critical imaginaries and reflections on artificial intelligence and robots in post-digital K-12 education. *Post-Digital Science and Education, 1*(2), 427-445. https://doi.org/10.1007/s42438-019-00046-x
[23] Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society, 26*(1), 112-131.
[24] Ivanov, S. (2016). Will robots substitute teachers? Yearbook of Varna University of Management, *9,* 42-47.
[25] İşler, B., & Kılıç, M. Y. (2021). Eğitimde yapay zeka kullanımı ve gelişimi. *e-Journal of New Media / Yeni Medya Elektronik Dergi – eJNM, 5*(1), 1-11.
[26] Kafadar, T. (2022). Oyunlaştırmanın eğitimdeki yeri. In T. Kafadar, ve C. A. Asena (Eds.), *Eğitimde oyunlaştırma, (1-16)*. Nobel Akademik Yayıncılık.
[27] Karasar, N. (2005). *Bilimsel araştırma yöntemi.* Nobel Yayın Dağıtım.
[28] Karsenti, T. (2019). Artificial intelligence in education: The urgent need to prepare teachers for tomorrow's. *Formation et Profession [Education and Profession]*, 27(1), 112-116.
[29] Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An artificial intelligence approach to monitor student performance and devise preventive measures. *Smart Learning Environments, 8*(1), 1-18.
[30] Kim, J., & Park, N. (2019). Development of a board game-based gamification learning model for training on the principles of artificial intelligence learning in elementary courses. *Journal of the Korean Association of Information Education, 23*(3), 229-235.
[31] Köse, B., Radıf, H., Baysal, İ., & Demirci, N. (2023). Öğretmen görüşlerine göre eğitimde yapay zekanın önemi, *Journal of Social, Humanities and Administrative Sciences, 9*(71):4203-4209.
[32] Kumar, N. S. (2019). Implementation of artificial intelligence in imparting education and evaluating student performance. *Journal of Artificial Intelligence, 1*(1), 1-9.
[33] Lin, P. Y., Chai, C. S., Jong, M. S., Dai, Y., Guo, Y., & Qin, J. (2021). Modelling the structural relationship among primary students' motivation to learn artificial intelligence. *Computers and Education: Artificial Intelligence,* 2(100006). doi: https://doi.org/10.1016/j.caeai.2020.100006
[34] Luo, Y., Han, X., & Zhang, C. (2022). Prediction of learning outcomes with a machine learning in blended courses. *Asia Pacific Education Rev*. doi: https://doi.org/10.1007/s12564-022-09-6
[35] Marr, B. (2019). The seven biggest technology trends in 2020 everyone must get ready for now. *https://www.forbes.com/sites/bernardmarr/2019/09/30/the-7-biggest-technologytrends-in-2020-everyone-must-get-ready-for-now/#54c7e5612261*
[36] Mede, E., & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 43-60.
[37] Mena-Guacas, A. F., Urueña Rodríguez, J. A., Santana-Trujillo, D. M., Gómez-Galán, J., & López-Meneses, E. (2023). Collaborative learning and skill development for educational growth of artificial intelligence: A systematic review. *Contemporary Educational Technology*, 15(3).
[38] Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology, 53*(4), 914-931. doi: https://doi.org/10.1111/bjet.13232
[39] Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of artificial intelligence powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon, 7*(5).
[40] O'Leary, Z. (2017). The essential guide to doing your research project. SAGE Publications Inc.
[41] Owan, V., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *EURASIA Journal of Mathematics, Science and Technology Education, 19*(8).
[42] Önder, A., Öztaş, G. S. & Akçapınar, G. (2023). Öğrenme analitiği sürecine yönelik modellere genel bir bakış: Kavramsal bir çerçeve önerisi. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi, 9*(1), 92-117.
[43] Öz, H. (2014). Turkish teachers' practices of assessment for learning in English as a foreign language. *Journal of Language Teaching*, 5, 775-785.
[44] Papapicco, C. (2020). Google mini: Italian example of artificial pro-sociality. *Online Journal of Communication and Media Technologies*, 10(3).
[45] Pirim, H. (2006). Yapay Zekâ. Journal of Yasar University. 1(1), 81-93
[46] Ren, X. (2020). Artificial intelligence and depression: How AI powered chatbots in virtual reality games may reduce anxiety and depression levels. *Journal of Artificial Intelligence Practice, 3*(1), 48-58.

[47] Rospigliosi, P. A. (2023). Artificial intelligence in teaching and learning: what questions should we ask of ChatGPT?. Interactive Learning Environments, 31(1), 1-3.
[48] Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A., & Sari, M. N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. Journal on Education, 6(1), 750-757.
[49] Sarıoğlu, S. (2023). Bilimsel süreç becerilerinin yapay zeka ile yordanması, öğrenciler ve üstün yetenekli öğrencilerdeki etkililiği. Doktora Tezi, Uludağ Üniversitesi, Bursa.
[50] Seufert, S., Guggemos, J., & Sailer, M. (2021). Technology-related knowledge, skills, and situation and emerging trends. *Computers in Human Behavior, 115*(106552). doi: https://doi.org/10.1016/j.chb.2020.106552
[51] Seyrek, M., Yıldız, S., Emeksiz, H., Şahin, A., & Türkmen, M. T. (2024). Öğretmenlerin eğitimde yapay zeka kullanıma yönelik algıları. *International Journal of Social and Humanities Sciences Research (JSHSR), 11*(106), 845-856. (Seyrek et al., 2024)
[52] Steinbauer, G., Kandlhofer, M., Chklovski, T., Heintz, F., & Koenig, S. (2021). A differentiated discussion about AI education K-12. *Kunstliche Intelligenz, 35*(2), 131-137.
[53] Stiggins, R. (2014). Improve assessment literacy outside of schools. *Phi Delta Kappan*, 96(2), 67-72.
[54] Suh, W., & Ahn, S. (2022). Development and validation of a scale measuring student attitudes toward artificial intelligence. *Sage Open, 12*(2). doi:https://doi.org/10.1177/21582440221100463
[55] Şad, S. N., & Göktaş, Ö. (2013). Öğretim elemanlarının geleneksel ve çağdaş ölçme değerlendirme yaklaşımlarının incelenmesi. *Ege Eğitim Dergisi, 14*(2), 79-105.
[56] Tsagari, D. (2011). Investigating the 'assessment literacy' of EFL state school teachers in Greece. In D. Tsagari, & I. Csépes (eds.), *Classroom-based language assessment*, (169-190). Frankfurt am Main: Peter Lang,
[57] Tuncer, M. & Gezer-Şen, B. (2018). Okul öncesi eğitimde toplumsal cinsiyet eşitliğine yönelik Türkiye'de yapılan çalışmalar. *Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi (ASEAD)*, 5(12), 870-882.
[58] Türker, O. (2023). Eğitim bağlamında yapay zeka yazılımları. E. Bay ve O. Zahal (eds.), *Eğitim Bilimleri Alanında Akademik Çalışmalar*. Ankara: Gece Kitaplığı, 189-205.
[59] Vasconcelos, M. A., & dos Santos, R. P. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *EURASIA Journal of Mathematics, Science and Technology Education, 19*(7). doi: https://doi.org/10.29333/ejmste/13313
[60] Xia, Q., Chiu, T. K., Lee, M., I., T., Dai, Y., & Chai, C. S. (2022). A self-determination theory design approach for inclusive and diverse artificial intelligence (AI) K-12 education. *Computers& Education*,*189*(104582). doi: https://doi.org/10.1016/j.compedu.2022.104582
[61] Yılmaz, G. (2024). Disleksi Öğrenciler İçin en etkili yapay zeka uygulamaları ve chrome uzantıları. :https://tr.linkedin.com/pulse/disleksi-%C3%B6%C4%9Frenciler-i%C3%A7in-en-etkili-_____yapay-zeka-ve-chrome-gi%CC%87zem-yilmaz-jscsf?trk=public_post_feed-article-content
[62] Yolcu, H. H. (2024). Yapay genel zekâ çağında öğretmen rolünün yeniden tanımlanması: öngörüler. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi, 10*(1), 155-167.

The Role of Psychologists in Primary Education

Aneta SOŁTYSIAK-JEDNASZEWSKA [a*]

[a] *St Mary`s Institute Limited, Dublin, Ireland, soltysiak.ane@gmail.com*

Abstract

Psychologists play a crucial role in primary education by supporting students' cognitive, emotional, and social development. They assess learning difficulties, provide early intervention strategies, and collaborate with teachers and parents to create inclusive and supportive learning environments. Additionally, psychologists address behavioural challenges, promote mental well-being, and implement programs to enhance students' resilience and self-regulation skills. Their contributions help ensure that children receive the necessary support to thrive academically and emotionally, fostering a positive foundation for lifelong learning. Psychologists also help in designing school-wide programs that reduce stress, improve peer relationships, and prevent bullying, ultimately shaping a healthier school culture. Through their expertise, psychologists empower educators to better understand individual student needs, enabling more tailored and effective teaching approaches.

Keywords: psychology, primary education, role of psychologist in primary education, mental disorders, resilience, self-regulation-skills, behavioural challenges

1. INTRODUCTION

The acquisition of knowledge is not merely an intellectual endeavor, and neither is education itself. Indeed, the faculty of learning can unhinge a child's curiosity, compromising affective engagement in and motivational responses to developmental encounters. We can observe that learners acquire knowledge differently at the beginning of schooling than at the end [1]. This is an ongoing process being under variety of intrinsic and extrinsic changes. It is here that the psychologist working in the primary school can be most useful. This represents a beginning orientation to a gradual awareness with regard to the vital instrumental role of psychology in primary education [2].

In what follows, this work will attempt to demonstrate the necessity of the psychological perspective, embarking ultimately upon the following aspects: the role of the psychologist in the schools, the functions they fulfill, the approaches they adopt, and the guidelines they utilize in doing so. There shall be recommended a more comprehensive reconceptualization of the needs of children in the schools that are psychological in nature. It is also argued that psychological practice, namely, psychological services, are part and parcel of the totality and direction of future trends in psychological outlook in primary schools and education generally.

What do psychologists do in the schools? They are involved in multiple tasks, including theoretical and practical duties. They are involved with education, with the children for whom education is meant, and the environment, or, most accurately, the environments in which children are educated. They view the problem and its related solutions from different orientations, and they function in the school setting in terms of these orientations. They may also direct the child, the teacher, the school or educational system bureaucracy, or any of the multiple participants in the schooling process. They may, therefore, become directly involved in such functions or take part indirectly in them.

* Corresponding author.

2. HISTORICAL DEVELOPMENT OF PSYCHOLOGISTS IN PRIMARY EDUCATION

The present-day pupil support group in primary schools in Poland consists of a counselor-teacher, and a school psychologist. The role of the school psychologist is relatively new. Therefore the historical background is recounted here. The theoretical basis of this paper is the ecological system theory introduced by American ecological psychologist Urie Bronfenbrenner in the 1970s. This theory emphasizes the interaction of the child with its environment. This theory emphasized that *the environmental systems affecting individual development includes both individuals` own biological factors, and the proximal environmental factors and the distal environmental factors* [3].

Granville Stanley Hall was the precursor of child psychology. *He used the questionnaire, originated by Galton, in studying children`s display of anger, fear, play, cry, etc.* [4]. Hall`s work was a cornerstone in studying children`s problems at the end of the XIX and beginning of the XX century.

The aim of the psychologist in education was to understand how the knowledge was acquired by the learners. They applied these principles through notes in studies of children, adults, and animals. Their study has applications when it comes to mainstreaming and diversity. The first American research to study the child in the field of psychology had a strong impact on education in the United States and in Europe, in parallel with the application of psychological services to education. From the second half of XIX century, the laboratories of psychology developed these principles to apply them to topics such as the fundamental processes of thought, memory, and movement. The first experimental psychology laboratory was created by doctor Wilhelm Wundt at the University of Leipzig, Germany in 1979 [5]. Wiliam James established psychology department and opened a psychology laboratory at Harvard University in 1875. James offered a programme entitled “The Relation between physiology and psychology” for the undergraduate students [4].

The study of the child was considered until 1950 by psychologists. It is then that some psychologists began to help children in disarray entering school. It is then that the school became the subject of study by the child. It is observed that in the 1960s, psychometrics, medical services, and experimental schools developed, and a center was opened to help underprivileged children and worked to teach educators. This marks the shift from helping the child in human groups to helping schools receive children as individuals.

Current psychological treatment include varied theoretically and treatable approaches. Today psychologist work with children focuses on such aspects of the individuum like: child-centered, family focused, community based practices in planning, implementing and evaluating the child`s treatment [6].

3. THEORETICAL FRAMEWORKS FOR UNDERSTANDING THE ROLE OF PSYCHOLOGISTS

The work of psychologists in primary education is guided by many theoretical frameworks and historical advances in this discipline. The resulting developmental theories offer basic principles about how humans learn, grow, and change, offering guidance on assessment, intervention, and collaboration strategies. Theories of learning from these perspectives inform us about how to assess, intervene, and collaborate for improved student functioning in all aspects of schooling, for 'academic' outcomes and 'mental health' outcomes [7].

A full discussion of frameworks is beyond the scope of this paper. That said, a “theoretical framework” is a broad range of cognitive beliefs that inform practice or other functions within an organization. For schools, our “beliefs” about children and the nature of disability inform many district and school practices, including response to intervention, programming, evaluation, and the practice of school psychologist broadly. The everyday practice of school psychologists draws upon a foundation of theoretical beliefs whether we are using evidence-based practice, promising practices, or best practices. Understanding why we do “what we do”, particularly when it comes to our assessment of children and exclusionary practices, is the necessary foundation for making any future change that will improve child, family, and school outcomes. However, the growing psychological support needs of children and youth have outstretched the possibilities and capacities of psychological care system [8]. Despite the workload, difficult student cases and low earnings, school psychologists can be proud of the excellent results of their work [9].

4. PSYCHOLOGICAL ASSESSMENT AND INTERVENTION STRATEGIES IN PRIMARY EDUCATION

Psychological assessment offers a series of methods and techniques that can be very useful as tools of analysis for educators in order to determine the educational needs of their students. In the first place, performance or achievement tests can be used to analyze the levels of knowledge and mastery of cognitive competencies in children and adolescents. In this way, it is possible to identify early those students who may present educational problems such as learning disorders, deficits in the development of specific competencies, and lag in learning the study content, among

others. On the other hand, instruments like emotional coding seek to assess the emotional status of students. Given that academic learning, to a large extent, is emotionally influenced, it is important to identify all those students who are emotionally low and work on the causes of this mood [10].

Intervention time of psychologist work with a child is crucial. In general, there is a conviction that the sooner the psychologist start working with the child who needs help, the better results can be achieved. Moreover, in an educational context where the idea is to prevent the emergence of problems as soon as possible, if left untreated, they can accompany students throughout their training. When different needs are detected, intervention techniques and strategies can be implemented. The development of a personalized educational intervention plan for children with special educational needs and high capacities, as well as for children requiring learning support, is considered basic [11]. The starting point for each student must always be individualization. Each student requires specific objectives and must work on the necessary contents or competencies to develop those identified objectives. Evidence-based knowledge and experience on effective teaching strategies or intervention actions that lead to good student performance and well-being contribute to implementing a quality educational intervention.

The personal educational intervention plan is revised at least once a year, based on the average result of measures of the overall developmental curriculum assessments of the students' peers. If it is not possible, improvements, needs, proposals for changes, and/or recommendations are taken into account in the plan in the student's educational intervention report as appropriate. To evaluate and carry out the above-mentioned interventions, psychopedagogical follow-up that involves working simultaneously with the school context, teachers, and the family is considered essential. The figure of an educational psychologist who continuously assesses, diagnoses, and treats the evolution of each of the students could ensure an optimization of the strategies used, in addition to optimizing the expected results. It is important to involve teachers in the elaboration of the plan. Regular meetings must be held between teachers and the educational psychologist to exchange strategies, assess results, and adapt the educational action plan to the results obtained, always based on a common criterion in the evaluation of teaching strategies [12]. In short, a close relationship between the counseling team, teachers, and the student's family would ensure that the strategies proposed in the different school contexts are the same. Likewise, this relationship would favor the implementation of a parental action plan that reinforces some of the actions or strategies performed in class.

4.1. Collaboration with Teachers and Parents

In order to provide pupils with psychological support (cognitive, emotional, volitional education, prevention of conflicts between children, bullying, etc.), both conceptual and practical cooperation between child development psychologists and educators (teachers) and parents is essential. Child development psychologists collaborate with teachers, hold methodological meetings where they can present behavioral and developmental case studies. These reports help teachers to evaluate their actions in working with children, to understand the children's motives, and so on.

Psychologists can also provide training sessions for teachers in order to help them recognize when a child is in need of psychological help. Various ways of involving parents in the activities of the society were suggested by nearly half of respondents.

Additional comments in this respect suggest that psychologists may come to the kindergarten to meet parents, invite parents to participate in workshops at school, and hold parent meetings at school. Parents becoming actively involved have the opportunity to share their experience with other parents and reflect on their children`s day to day school challenges and constrains.

Parents are an integral and valuable part of any work with children, and their involvement can lead to positive outcomes for children. They can support the work done in school and work on the same projects outside school. This not only helps to link school and home but adds relevance to the work being carried out. It is generally felt that when working with children, a team approach is vastly effective, as a multitude of ideas and solutions to support the child provide a richer pool of interventions required. A team approach can include teachers, parents, ancillary staff, educational psychologists, and child psychologists. In primary school, it has been found that many popular initiatives have been led by a psychologist but required the cooperation of other staff.

4.2. Addressing Diversity and Inclusion in Primary Education

As educators and child advocates, we must start by acknowledging that children come to us with the wealth of varied experiences that different family backgrounds bring. One of our most challenging and rewarding tasks is to find out about these backgrounds and treat them as a gift to be shared. Delivering at least part of that education in another language, or in another culture, brings even more opportunities—chances to collaboratively develop

curriculum, values, traditions, and a code of ethics responsive to a changing society. Few things are more uplifting than a multiracial, multiethnic, multilingual group of children or parents laughing or smiling together.

Carefully examining the components of cultural difference, psychology, generational change, and socio-economic barriers, we seek to debunk the purported benefits of an affluent homogeneous private education in the context of today's world. Inclusivity must assure that all students, regardless of home background, have access to the instructional opportunities all their peers enjoy. *Inclusion is understood as a dynamic process of transforming the cultures, policies and practices of educational communities to provide a quality education response to all learners* [13]. Psychologists in primary education classrooms address diversity and therefore inclusivity. They champion equal opportunities and practices, thereby providing quality education to children from varied backgrounds. They are active in providing supportive services to prevent learning problems, mental health issues, and emotional difficulties [14].

Excellent school psychologists are proactive and advocate to impact policies and practices so that they are culturally responsive and equitable. This led to inclusive programs and practices that contribute to more effective school and classroom management, and in the end enhance the academic and social achievements of students. Regular professional development in diversity competence is a priority for every school and mental health practitioner in school and community settings.

4.3. Promoting Positive Mental Health and Well-being

There is a proactive aspect of the psychologist's role: the promotion of positive mental health and well-being. Early mental health education and awareness can contribute to young people developing positive emotional strength and problem-solving strategies, and can help create an environment of openness and responsiveness to emotional and mental health issues. A supportive school environment promotes emotional health and helps foster resilience. Strategies to encourage the development of a range and balance of coping styles include anti-bullying programs, alternative dispute resolution, teaching problem-solving, and supportive school environments. Such an approach is discussed in programs providing a variety of strategies to build students' skills in dealing with suicidal comments and confidences, and those designed to assist a student following a perceived suicide risk incident [15].

The promotion of social and emotional learning programs in primary schools should be an ongoing process. Social and emotional learning is both a primary prevention strategy in preventing the development of depressive and anxiety disorders, and an early intervention for young people with the beginning signs of emotional distress. Such a program can be the basis for a school's philosophy and can have positive impacts on an entire school community. Planning for preventive mental health programs in primary schools should be carried out in collaboration with the students' educators and the students' home community. The latter are important for intervention, advice, and support, and can help to make programs more feasible for school staff and effective for school students. Evaluation is an important next goal of mental health promotion and mental illness prevention efforts in communities. It helps to determine a program's and a program's elements' effectiveness, their sustainability, and any necessary modification of a program both before and after its implementation. Providing the best opportunity for promoting mental well-being within the primary years is an essential component of academic success.

Whereas the work of school psychologists is focused on treatment of students` mental disorders and mental problems, detection and prevention can be equally important for children`s further educational progress and development. When the child emotional, mental, social, etc. problems are not noticed and neglected, the chances of self-harm, becoming aggressive, attempting suicide or dropping the school are higher.

4.4. Prevention and Intervention of Behavioral Issues

Many behavioral issues manifest in the general classroom at the primary education level. Though public school teachers are trained to differentiate instruction to meet the diverse needs of all students, they may not be prepared to manage various challenging behaviors. Disruptive, defiant, aggressive, and avoidant behaviors may be dominant; however, instances of impulsivity, immaturity, attention and persistent inattention, and off-task behavior can impinge upon a classroom teacher's ability to deliver instruction to any held standard. It is a short-sighted approach to wait for students to emerge into a manifestation of a learning disorder prior to addressing their presenting behaviors. Beyond addressing disruptive behavior, early intervening psychological strategies on forbidding behavioral issues and addressing behavior plans are important for the educational and psychological sustainability of students, their families, and educational systems.

In attempting to stop minor behavioral issues that often precede disruptive behaviors, psychologists should enhance prosocial skills among students and assist students and teachers alike in cooperative developmental strategies usable in the classroom in real time. As an entry-level strategy, this is important to block the development of severe

behavior issues. Once minor behavioral issues increase in frequency or intensity or persist, treatment of core symptoms with psychological intervention is advisable.

School psychologists address these lower-level behavioral functioning issues systematically by:
1) addressing the underlying cause of attention issues,
2) providing conflict resolution strategies and resources for students to problem-solve,
3) modifying the environment to reduce or prevent access to desired activities.
School psychologists have a variety of treatment methods to replace less desirable behavior with behavior that is more socially acceptable [16].

4.5. Supporting Students with Special Educational Needs

Children with special educational needs (SEN) could have categorical disabilities related to those known as communication disorders, specific learning disabilities, or cognitive and intellectual impairment due to conditions such as Down syndrome, autism spectrum disorder, and multiple disabilities, as well as other factors. Non-categorical disabilities could affect academic learning. These disorders include attention, behavior, memory, perception, and the processing of information. They are collectively called hidden disabilities and include attention deficit hyperactivity disorder (ADHD), behavior disorders, and learning difficulties in areas such as reading, mathematics, and spelling. In addition to hidden disabilities that could affect academic achievement, children might exhibit a range of challenges such as chronic illness, environmental, and cultural disadvantages [17].

An SEN (Special Educational Needs) assessment is conducted by educational psychologists who base it on the child's needs to plan interventions that can help facilitate the child's inclusive education and improved academic success. Typically, the role of educational psychologists involves some or all the following interventions:

1) Assessment and diagnosis – practitioners acknowledge and choose a model of SEN assessment that fosters a good educational fit that is adjusted to the range of special needs and can provide recommendations for curriculum differentiation to inform teaching and learning strategies that allow pupils access to the class material.

2) Intervention plan – the focus of assessment should also be on the identification of strategies that will foster the child's inclusion in regular education and help them academically. Overall, emphasis should be placed on information that is related to educational decision-making. The strategies or interventions outlined in the assessment plan must be rooted in or demonstrate a clear understanding of research and evidence-based practices.

3) Advocacy – any cognitive profile that is developed through academic assessment should help identify areas of deficit and provide information that could be used by the parent and psychologist to advocate for certain resources or types of educational placements that the child needs. Parents need training in advocacy skills in order to benefit the child with special needs and access the appropriate resources at home and with public institutions.

4) Collaboration – the most powerful instrument in addressing the educational development of children with special needs lies in collaboration among service providers. This means that the educational psychologist needs to have contact with the classroom teacher for continuous feedback and to gain a more comprehensive perspective of a student's learning in a broader context. Similarly, every part of the assessment must be checked with special educators and the teachers and those available in the field of learning support in order to provide suitable and comprehensive support.

5) Consultancy and advising should always be provided when changes need to be made in the school environment concerning instruction, teachers' expertise, teacher-student ratios, curriculum, and resources.

Conclusively, the SEN assessment and diagnosis should be a functional instrument pointing the way to how children with special needs will be included and their strengths used in the educational context.

5. CHALLENGES AND FUTURE DIRECTIONS

Challenges and Limitations. There are several challenges that impact the way psychologists work in primary education. Funding and staffing shortages make it difficult for schools to offer psychological services. Although results for literacy are high, many children struggle to read, and many children's mental health is compromised. There

are no sufficiently available data that might indicate effective strategies the governments could roll out, and although educational reforms are underway in many countries. The reality is also that many schools lack basic security. Furthermore, parents and caregivers seek services that are funded and delivered by service providers who have been trained about mental disorders. They might avoid psychological services at schools if the service is branded as a joint problem-solving and educational service.

Technological advances have been making innovations in psychological practice for the delivery of services via the internet, telephone, and directly to mobile devices. Finally, psychologists and schools are invited to radically rethink approaches that are currently used. Best teaching practice is known to occur when teachers are supported to adapt their approaches to meet the individual differences within their classrooms; psychological services might best engage highly diverse students by adapting service approaches as well. Interventions that primarily target the child with psychological treatments while largely ignoring the child's context have shown to be less effective than interventions that are implemented at various levels. Furthermore, long-term mental health treatment services are overwhelmingly indicated to be most effective when the external context of the individual facilitates its effectiveness. Current health settings have determined that schools are the primary mental health service context for children. Psychologists may have a range of views about whether this is true. If effectively implemented, over time, there may be less pressure for schools to provide support for children with disabilities. There are relatively few data on the impact of low-intensity interventions on long-term outcomes of children.

References

[1] Mirolli, M. & Baldassarre, G. (2013). *Intrinsically Motivated Learning in Natural and Artificial Systems.* Springer Berlin Heidelberg, 58.
[2] Pintrich, P. R. (2000). *Educational psychology at the millennium: A look back and a look forward. Educational Psychologist, 35.* Routledge, 221-226.
[3] Guanghui, C. (2025). *Ecological Systems Theory. The ECPH Encyclopedia of Psychology.* Springer Nature, 441.
[4] Singh, A. K. (1991). *The Comprehensive History of Psychology*. Motilal Banarsidass, 138-149.
[5] Freedheim, D. (2003). *Handbook of Psychology, History of Psychology.* Wiley, 518.
[6] Craighead, W. & Nemeroff, C. (2004). *The Concise Corsini Encyclopedia of Psychology and Behavioral Science.* Wiley, 170-171.
[7] Palombo, J., Bendicsen, H. & Koch, B. (2009). *Guide to Psychoanalytic Development* Theories. Springer New York, 21-39.
[8] Conoley, J. C., La Greca, A. M. & Hughes, J. M. (2001). *Handbook of Psychological Services for Children and Adolescents.* Oxford University Press, 74.
[9] Rathavon, N., Rilley-Tilman, T. C. & Burns, M. K. (2017). *Effective School Interventions. Evidence-Based Strategies for Improving Students Outcomes.* Guildford Publications, 3.
[10] Wong, R. (2020). *When no one can go to school: does online learning meet students` basic learning needs.* Interactive Learning Environments, Volume 31, 2023, Issue: 1.Taylor and Francis.
[11] Levin, J. R. & Kratochwill, T. R. (2015). *Single-Case Research Design and Analysis. New Directions for Psychology and Education.* Taylor & Francis, 167.
[12] Bray, M. A. & Kehle, T.J. (2011). *The Oxford Handbook of School Psychology.* Oxford University Press, 21 -23.
[13] Kairiene, D., Bahdanovich Hanssen, N. & Alisauskiene, S. (2023). *Interprofessional and Family-Professional Collaboration for Inclusive Early Childhood Education and Care*. Springer International Publishing, 140.
[14] Merrell, K. W. Ervin, R. A. & Peackock, G.G. (2011). *School Psychology for the 21st Century. Second Edition.* Guilford Publishing 215.
[15] Matsumoto, K., Hamatani, S. & Singh, T. (2024). *Psychological Intervention for Suicidal Ideation, Behaviour, and Attempted Suicide.* Frontiers Media SA, 11.
[16] Peacoock, G. G., Merrell, K. W. & Ervin, R.A. (2012). *Practical Handbook of Psychology. Effective Practices for 21st Century.* Guilford Publications, 503.
[17] McNamara, B. E. (2017). *Bullying and Students With Disabilities. Strategies and Techniques to Create a Safe Learning Environment for All.* SAGE Publications, 17.

Implementation of Integrated Quality Management in Educational Institutions

Fethi KAYALAR [a*]

[a] *Erzincan B. Y. University), Faculty of Education,TURKEY fkayalar@erzincan.edu.tr*

Abstract

Integrated quality management of education has begun to be implemented in the world of education along with the rapid development of science and technology. The purpose of this study is to explore educational quality management and the process approach, steps for implementing the process approach, the benefits of the process approach, and the application of the process approach in educational institutions to improve quality. Quality management is used for continuous improvement of institutions. In quality management, there are eight principles used, but there are still educational institutions that have not implemented the principles of quality management, especially the principle of the process approach. The method used is a qualitative approach with literature studies. The principle of the process approach views that the results of this study can be achieved more effectively and efficiently, when the managed activities are combined into an interrelated process and function as a continuous system. A quality management system based on this principle is very important because it is based on the relationship of interrelated processes that enable the achievement of organizational goals with the satisfaction of interested parties. The conclusion of this study is the process approach includes input, process, and output, the principle of the process approach is in line with Deming's principle, namely planning, doing, checking and acting. The process approach can be a recommendation in determining the expected goals for educational institutions in improving quality.

Keywords: Integrated Quality Management, Educational Management, Teacher Training, Educational Quality Management

1. INTRODUCTION

Management is a unique process consisting of functions such as planning, organizing, motivating and controlling the actions taken to determine and achieve predetermined goals with the use of human resources and other resources. Educational management is considered as a conscious and planned effort to create a learning environment and learning process for students to effectively develop their spiritual religious power, self-control, personality, intelligence, high morality and the potential to acquire the skills that they, the society, the nation and the state need.

Integrated quality management essentially refers to all activities of the general management function. It is a process that starts from planning, organizing and controlling, and extends to the leadership that determines quality policies, goals and responsibilities and their implementation with management tools such as planning, control, quality assurance and improvement.

In absolute terms, educational quality refers to a feature that defines the degree of "goodness" of the service and graduates produced by a particular educational institution according to agreed national and international standards, meets the needs of stakeholders and avoids unsuccessful products as unqualified graduates.

* Corresponding author.

Currently, various efforts to improve the quality of education continue to be carried out by many parties to respond to the many quality problems faced in the world of education such as the quality of graduates, the quality of teaching and the quality of professionalism of teacher performance. These efforts are based on an awareness of how important the role of education is in developing human resources and developing national character for the progress of society and the nation, because the dignity of a nation is very much determined by the quality of its education.

Effective educational institutions in improving the quality of education require good quality management to face the competitive atmosphere and future orientation. In quality management, management attention cannot be separated from the desires, needs and expectations of customers, both internal and external. Because the assessment of the good and bad quality of an educational institution is very dependent on the assessment of customers. Therefore, to understand quality management in education so that an education can provide maximum service and can satisfy customer expectations, in this study we will discuss quality management in educational institutions.

2. CONCEPT OF EDUCATIONAL QUALITY

The meaning of quality can be different for each person, because quality has many criteria and is very dependent on the context. Lexically in the dictionaries, quality is a measure of the good or bad of an object, condition, level, or degree of skill, intelligence, and so on. Quality is defined as the suitability of product use (fitness for use) to meet customer needs and satisfaction or quality as conformity to specifications. Meanwhile, according to Deming (1982), quality is conformity to market needs or whatever is the need and desire of consumers. The quality of education is the level of intelligence of the nation's life that can be achieved from the implementation of the National Education System.

Although there is no universally accepted definition of quality, there are several similarities from the existing definitions. This means that defining quality requires a comprehensive view. In this case, there are several elements that can make something said to be of quality. First, quality includes efforts to meet or exceed customer expectations. Second, quality includes products, services, people, processes, and the environment. Third, quality is a condition that is always changing, meaning that what is considered quality today may be considered less quality at another time. Fourth, quality is a dynamic condition related to products, services, people, processes, and environments that meet or exceed expectations (Umiarso and Gojali, 2011; Díez, Iraurgi and Sanchez, 2018).

So the quality of education is the degree of excellence in managing education effectively and efficiently to produce academic and extracurricular excellence in students who are declared to have graduated for one level of education or completed a certain learning program. It must also be known what is included in the quality dimension. Garvin defines eight dimensions that can be used to analyse the characteristics of product quality, namely (a) Performance; (b) Features, characteristics or specialties and complementary characteristics; (c) Reliability; (d) Conformance; (e) Durability; (f) Service ability; (g) Aesthetics; and (h) Perceived quality.

The indicators or criteria that can be used as benchmarks for educational quality are the final results of education, direct results of education which are used as benchmarks for measuring the quality of education in an educational institution such as written tests, checklists, anecdotes, rating scales, and attitude scales, educational processes, input instruments that interact with raw input, namely students, and raw input and the environment.

This question is very important: *"Who should decide whether a school has succeeded in providing a quality service?"*. Customers are the arbiters of quality and the institution itself will not be able to survive without them. Quality can be defined as something that satisfies and exceeds the desires and needs of customers. This definition is also known as quality in perception. This quality can be called quality that only exists in the eyes of the beholder. This is a very important definition. This is because there is one risk that is often overlooked in this definition, namely the fact that customers are the ones who make decisions about quality. And they make this assessment by referring to the best products that can survive the competition (Sailis, 2011).

According to Sailis (2011) there are two groups in quality standards, namely:

1. Product and Service Standards, consisting of: (a) Conformity to specifications; (b) Conformity to objectives and benefits; (c) Zero effects; and (d) Always good from the start.
2. Customer Standards, consisting of: (a) Customer satisfaction; (b) Meeting customer needs; and (c) Delighting customers.

3. QUALITY MANAGEMENT IN EDUCATION

Quality Management (QM) in education is a philosophy of continuous improvement, which can provide a set of practical tools to every educational institution in meeting the needs, desires, and expectations of customers, now and for the future.

According to Umiarso and Gojali (2011), QM is a functional management with an approach that is continuously focused on improving quality so that its products are in accordance with quality standards of the community served in the implementation of public service tasks and community development. Meanwhile, according to Sudiyono (2004) Integrated Quality Management (IQM) is an activity that seeks to optimize the competitiveness of an organization through continuous improvement of its products, services, people, processes and environment.

IQM is a fundamental and comprehensive value system in managing an organization with the aim of improving performance sustainably in the long term by paying special attention to achieving customer satisfaction while still paying adequate attention to meeting the needs of all stakeholders of the organization concerned. Quality issues in IQM require the involvement and responsibility of all parties in the organization. Therefore, the integrated quality management approach is not only partial, but comprehensive by involving all parties interested in the products produced. Quality issues are also no longer interpreted and viewed as technical problems, but are more oriented towards realizing consumer or customer satisfaction. IQM also involves physical and non-physical factors, such as organizational culture, leadership style and followers. The integration of these factors will result in increased and meaningful service quality.

IQM in education is an effort to optimize educational institutions and organizations in order to satisfy customers. Thus, integrated quality management in education is related to:

- Customers, both internal and external,
- Quality, which means the quality of individual and institutional service
- Commitment from all components in the organization,
- Teamwork,
- Improvements made continuously,
- Freedom carried out in a controlled manner,
- Unity of purpose,
- Overall personnel involvement.

Quality education services are those that meet or exceed customer needs and are able to provide educational services appropriate to the needs of their customers, whether they are students, businesses or the community (Sohel-Uz-Zaman, 2016; Sun and Guozhen, 2024).

In integrated quality management, schools are acclaimed as service units, namely learning servants. As a service unit, the school serves customers, both internal and external customers. Several sources of quality that can support the implementation of IQM optimally in education include:

- Commitment of top management (principal) to quality;
- Management information system;
- Potential human resources;
- Involvement of all functions; and
- Philosophy of continuous quality improvement.

IQM is a method that can help education professionals cope with a constantly changing environment. It can be used as a tool to form bonds between schools, the business world, and government. The association will enable educational professionals in schools or districts to be equipped with the resources needed to develop quality programs. IQM is a major aspect of total management. Total Quality Management (TQM) is a methodology that makes it easier to manage change, form a focus for change, form a more flexible infrastructure, respond quickly to the demands of changing society, and help education overcome cost and time constraints.

Changes to TQM begin with adopting a division of tasks regarding the implementation of quality at the school board level, administrators, teachers, administrative staff, students, parents, and the community. Activities begin with formulating the vision and mission of the school, department or study program, and sections of school education. The vision of TQM is focused on finding the needs of graduate users (customers), preparing to involve the community as a whole in quality improvement programs, developing a system to measure the added value of education, a support system that allows teachers, administrative staff to manage change, and making continuous improvements with the aim that the school product is heading in a better direction (Ejionueme and Oyoyo, 2015; Magd and Karyamsetty, 2020; Mamabolo, Malatji ve Mphahlele, 2022).

For educational organizations, the adaptation of IQM can be said to be successful if it shows these symptoms:

a. The level of product consistency in providing public services and implementing development for the benefit of continuously improving the quality of human resources.

b. Mistakes in work that have an impact on causing dissatisfaction and complaints from the community being served are decreasing.

c. Time discipline and work discipline are increasing.
d. The inventory of organizational assets increasingly perfect, controlled, and not decreasing or disappearing without knowing the cause.
e. Waste of funds and time in working which can be prevented.
f. Improvement of work skills and expertise continuously carried out, so that the method or way of working is always able to adapt to changes and developments in science and technology as the most effective, efficient and productive way of working. Therefore, the quality of products and services continues to increase.

3.1. Obstacles in Implementation of Integrated Quality Management

IQM is hard work and takes time to develop a quality culture. Hard work and time are two important things that must be considered, as the journey of the quality work mechanism will be hampered if these two things cannot run well. While IQM requires a champion mindset that can face challenges and changes in education, quality improvement is a process that requires attention and care. Because staying in place while competitors continue to develop is a sign of failure.

The biggest obstacle is the manager's, director's or principal's anxiety about adopting new methods and approaches. In IQM, they must trust that their employees will work together to carry the vision of their organization forward. Some are sometimes hesitant to share their vision with their subordinates because they are worried about losing their status and this is considered a decrease in their status.

It is not just managers who are obstacles to quality development. Some employees who are overly concerned about the consequences of empowerment can also be obstacles to quality. Sometimes they tend to like things that are static. They need to brainstorm about the importance and benefits of change. Therefore, IQM should not be just jargon and advertising. Because these things can lead to loss of enthusiasm, skepticism, cynicism and distrust of change. Most IQM obstacles involve elements of anxiety and uncertainty. Fear of the unknown or fear of doing something different, trusting others and making mistakes is a very powerful resistance mechanism. Employees will not be able to do their best work unless they feel trusted and their opinions are heard. Deming argued that 'eliminating fear and anxiety' is essential in the effort to bring about a quality revolution

3.2. Reasons for Failure in Educational Quality

According to Deming, if managers really pay serious attention to quality, then they need to understand the reasons for quality failure. In order to solve a problem well, it is necessary to understand its reasons well. And the analysis of quality failures is one of the most important results of Deming's research. He divides the reasons for failure into two categories: 'general' and 'special'. Common reasons are reasons resulting from system failure. This system problem is an internal problem of the institutional process. These problems can only be overcome by changing the institution's systems, processes and procedures. Other reasons, which he calls special reasons, are external reasons that lead to changes within the system.

Common reasons for low educational quality include poor curriculum design, substandard buildings, poor working environment, inappropriate systems and procedures, irregular work schedules, insufficient resources and insufficient personnel development. If it is determined that the error or failure is due to a system, policy or resource problem, this is a 'common cause' failure. From an administrative perspective, these reasons need to be eliminated and systems and procedures need to be redesigned, created and developed. This may require policy changes or new training. In order to identify the root and extent of a problem, it is necessary to make an effort to investigate failure data and perform regular checks. A common mistake in the education world is the lack of research and analysis on the reasons for the low level of achievement of goals, and also the failure to carry out this research and analysis as a subject of administrative actions.

On the other hand, specific causes of failure often result from procedures and rules that are not followed. Such failures can also result from individual staff members lacking the skills, knowledge, and qualifications required to be teachers or educational administrators. Specific causes of quality problems may include members' lack of knowledge and skills, lack of motivation, communication errors, or equipment problems.

If a problem is due to specific causes, then it can be resolved without the need for policy changes or system redesign. Changing the system may not be appropriate and may lead to more fatal failures. The source of the failure must be identified and resolved. Management of specific objectives is also a management responsibility. Of course, other staff members can take up the issue and resolve it, but sometimes they may not have the authority.

The implications of the distinction between common and special causes are important for managers. Do quality failures stem from special or general causes? If the problems employees encounter cannot be solved only by

motivation, it is no use providing them with motivational training. Problems and errors are often attributed to individuals, whereas the real causes of these errors are policy and system errors. Only those who can redesign the system can solve the problems created by the system. Most of these problems are caused by weak or inadequate management (Papanthymou and Darra 2018; Purnomo and Haryati, 2023; Queen-Mary and Mtapuri 2014).

Identifying the causes of quality failures and correcting them is a basic duty of a manager. Often the wrong solution or the wrong person is assigned to solve a problem. Also, often individuals are blamed for people who are not their own fault. In such cases, they are disappointed when their efforts fail. It takes effort to produce intelligent insights to overcome quality failures. Deming (1982) stated that in most cases, when something goes wrong, the culprit is not necessarily the staff. However, in reality, teachers become the scapegoats for the failures in the education system. In IQM, it is stated that successful quality development requires permanent commitment from management. IQM also emphasizes that commitment is not just about encouraging the efforts of others. In practical terms, commitment is management's awareness that it is responsible for finding a solution to a mistake.

3.3. Integrated quality Management Leadership in Education

In his book Edward Sallis, Spanbauer (1989)offers a leadership model for empowering teachers as follows:

- Engaging teachers and all staff in problem-solving activities using basic scientific methods, statistical quality principles, and process control.
- Choosing to ask their opinions about things and about how they run projects and not just telling them how they should behave.
- Conveying as much management information as possible to help develop and improve their commitment. Asking staff for their opinions about which systems and procedures are preventing them from delivering quality to customers: students, parents, and work partners.
- Understanding that the desire to improve the quality of teachers is not in accordance with a top-down management approach. Transferring responsibility and control for professional development directly to teachers and technical workers.
- Implementing systematic and continuous communication between everyone involved in the school.
- Developing problem-solving and negotiation skills in order to resolve conflicts.
- Having a helpful attitude without having to know all the answers to every problem and without feeling inferior.
- Providing learning materials for quality concepts such as team building, process management, customer service, communication and leadership. Setting a good example, by displaying the desired characteristics and taking the time to look around the institution and listen to the wishes of other teachers and customers.
- Learning to act as a coach and not a boss, Provide autonomy and dare to take risks.
- Providing balanced attention in providing quality for external customers such as students, parents and others, and to internal customers such as teachers, board of teachers, and other workers (Hairiah, 2015).

4. CONCLUSION

Educational quality is the degree of excellence in the effective and efficient management of education to produce academic and extracurricular excellence in students who are declared to have graduated from a certain level of education or completed a certain learning program.

IQM in Education is a philosophy of continuous improvement that can provide any educational institution with a set of practical tools to meet the needs, desires and expectations of its customers now and in the future.

The concerns experienced by principals in adopting new methods and approaches, their fear of the unknown or doing something different, their reliance on others and making mistakes are obstacles to the implementation of IQM. The reasons for the failure in educational quality include poor curriculum, substandard buildings, poor working environment, inappropriate systems and procedures, irregular work schedules, inadequate resources, inadequate staff development, lack of knowledge and skills of members, lack of motivation and poor communication.

References

[1] Deming, W. E. (1982). Improvement of quality and productivity through action by management. *National Productivity Review*, 1(1): 12-22.

[2] Díez, F.; Iraurgi, I. & Sanchez, A. V. (2018). Quality Management in Schools: Analysis of Mediating Factors. *South African Journal of Education,* v38 n2

[3] Ejionueme, L. K. & Oyoyo, A. O. (2015). Application of Total Quality Management (TQM) in Secondary School Administration in Umuahia Education Zone. *Journal of Education and Practice*, v6 n27
[4] Hairiyah (2015). The Concept of Integrated Quality Management in Education (Konsep Manajemen Mutu Terpadu Dalam Pendidikan). *LITERASI,* Volume VI, No. 1
[5] Magd, H., & Karyamsetty, H. (2020). Organizational Performance and Sustainability in Manufacturing and Service through TQM Implementation. *Open Journal of Business and Management,* 8(06)
[6] Mamabolo, M. M.; Malatji, K. S. & Mphahlele, L. K. (2022). The Role of Supervisors in the Implementation of the Integrated Quality Management System in Schools. *South African Journal of Education,* v42 n1
[7] Nasution, M. N. (2005). *Total Quality Management* (Manajemen Mutu Terpadu) (Jakarta: Ghalia Indonesia.
[8] Papanthymou, A.& Darra, M. (2018). The Implementation of Total Quality Management in Greek Higher Education: The Case of Electronic Administrative Services. *International Education Studies*, v11 n7
[9] Purnomo, S. & Haryati, T. (2023). Quality Management in Educational Institutions (Manajemen Mutu dalam Lembaga Pendidikan). *Jurnal Pendidikan Terintegrasi JPTR Falkutas Keguruan dan Ilmu Pendidikan* Volume 4 No 1
[10] Queen-Mary, T. N. & Mtapuri, O. (2014). Teachers' Perceptions of the Integrated Quality Management System: Lessons from Mpumalanga, South Africa. *South African Journal of Education,* volume 34, no1
[11] Sailis, E. (2011). *Total Quality Management in Education* (Ed: Ahmad Ali Riyadi). Yogyakarta: IRCiSoD.
[12] Sohel-Uz-Zaman, A. S. M. (2016). Implementing Total Quality Management in Education: Compatibility and challenges. *Open Journal of Social Sciences,* 4(11).
[13] Spanbauer, S. J. (1989). Measuring and Costing Quality in Education: Using Quality and Productivity Methods to Improve Schools. Fox Valley Technical Coll., Appleton, WI.
[14] Sun W. & Guozhen Y. (2024). Optimizing Online Teaching: Total Quality Management in Action for Quality Assurance Measures. *Journal of Education and Learning,* Vol. 13, No. 4
[15] Umiarso & Gojali, I. (2011). *Manajemen Mutu Sekolah di Era Otonomi Pendidikan.* Yogyakarta: IRCiSoD.

Enhancing Spanish Language Learning through Virtual Visits to the Prado Museum: A Theoretical - Pedagogical Approach

Kamela MARKU [a*], Megi PLAKU [b]

[a] *Phd Candidate in Translation, 00355 69 52 24 639 (University of Tirana, Faculty of Foreign Languages, Spanish Department, Elbasani Street, Tirana, Albania,kamela_marku@hotmail.coml)*
[b] *PhD Candidate in Lingustics, 00355 69 64 30 049 (Polytchnic University of Tirana, Dëshmorët e Kombit Boulvard, Tirana, Albania, megplaku@gmail.com)*

Abstract

This article delves into the effectiveness and the enrichment of learning Spanish for students of Spanish as a Foreign Language. Teachers can make use of the museum's wide collection of Spanish and European art in order to integrate the cultural aspect in the teaching process, by providing students with a vivid connection to the Spanish-speaking world and culture. Virtual museums tours can serve as a very dynamic and interactive tool where students get familiar, understand and acquire better the right vocabulary; they acquire the right linguistic and cultural competence. This approach of teaching creates a friendly learning environment for students where, fun, knowledge and interaction are all integrated. This study employs a qualitative approach, examining case studies of SFL programs that have implemented virtual museum tours. Data is gathered from educator feedback and analysis of student outcomes to assess how virtual visits contribute to linguistic and cultural learning objectives. It discusses best practices for incorporating virtual visits into the SFL curriculum, highlighting both the benefits and challenges of using digital resources in language education.

Keywords: Spanish language, teaching, virtual visit, effectiveness

1. INTRODUCTION

The Prado Museum, is a very famous museum which is located in Madrid, Spain. It is one of the world's most visited tourist attraction as a marvelous art institution, where one can encounter him/herself with masterpieces by artists such as Velazquez, Goya or El Greco, Van der Weyden, Hieronymus Bosch, Titian and Rubens which form only part of the Prado Museum's collection, considered by many to be one of the richest in the world for the quality and variety of its paintings.[1] From the didactic point of view, for students learning Spanish, art can serve as a gateway to a deeper comprehension of culture and language. However, it should be emphasized that not all learners have the opportunity to travel to Spain and experience its artistic heritage themselves, in person. Virtual museum tours help bridge this gap by providing an interesting and beneficial way that integrates cultural exploration into language instruction. This paper will examine, from the theoretical point of view, the benefits of using the Prado

* Corresponding author.

Museum's virtual resources in SFL classrooms and how they can enhance students' language skills while fostering cultural appreciation.

2. LITERATURE REVIEW

2.1 Cultural Richness in Language Learning

Culture plays a fundamental role in language acquisition. As Kramsch (1993) points out, language is more than a communication tool; it is a means of transmitting cultural values and traditions. That being said, when it comes to traditional language teaching, there are some classic components, such as grammar and vocabulary prioritization. Today's approach towards language teaching and learning is complex and dynamic. It is focused on competencies such as communication, critical thinking, and cultural competence. Culture can be a very interesting and fruitful tool in teaching Spanish to SFL students, because there is an integration of both curiosity for the culture and goal of learning the language. By merging these two aspects, the objectives are easily achieved.

2.2 Virtual Museums as Educational Tools

Technology has become the subject of the last century. Advancements in technology have made virtual museums widely accessible, allowing students to interact with art and culture in innovative ways. Research by Falk and Dierking (2012) highlights how digital platforms stimulate curiosity and encourage active learning. Virtual museum visits remove geographical and financial barriers, offering students an opportunity to explore authentic cultural content from anywhere in the world.

2.3 Integrating Technology in SFL Education

Studies (Stockwell, 2012) show that technology enhances student engagement and motivation in language learning. Digital tools such as virtual tours, podcasts, and interactive activities provide immersive experiences that cater to diverse learning styles. Virtual museum visits align well with this trend, offering both synchronous and asynchronous opportunities for language practice.

3. METHODOLOGY

This study employs a qualitative approach, examining case studies of SFL programs that have implemented virtual museum tours. Data is gathered from educator feedback and analysis of student outcomes to assess how virtual visits contribute to linguistic and cultural learning objectives.

4. THE BENEFITS OF VIRTUAL VISITS TO THE PRADO MUSEUM FOR EDUCATION

4.1 Improving Language Skills

Acquisition of Vocabulary: Students can be taught the specialized terminology related to art (for example, la paleta [palette], el marco [frame]). They can describe Goya's El 3 de mayo de 1808 with a particular emphasis on the colors, emotions, and composition by making use of the right vocabulary. While being presented to different pieces of art and their explanation, there is also a decrease of cognitive load, which increases retention and comprehension because students will have a visual representation of what they are listening to.

Grammar: Virtual visits allow students to practice grammar in a great context. For example, students can narrate scenes from Las Meninas by Velazquez’s using the pretérito imperfecto.

Students' listening skills: Some art pieces, for instance, "La Anunciacion” by Fra Angelico, can be accompanied by audio guides in Spanish, and students can be provided with comprehension questions to answer. This increases students’ attention and strengthens their listening skills.

4.2 Broadening Understanding Of Different Cultures

Behind the Art: Art is a window to Spain’s history. Students discussing “Las Meninas” can analyze the Spanish Habsburg dynasty and its influence on art.

Art and National Identity: Students can use Spanish art terminology when comparing Goya's El Coloso to Neoclassicism as an example of Romanticism.

General Symbols: Cultural symbols in Spanish history can be discussed on the basis of Goya's Saturno "devorando a su hijo"

4.3 Enhancing Learners' Participation and Interest

Interactive Exploration: Learners craft their own meanings from the visuals of El Jardín de las Delicias by Bosch because they are able to zoom into the details. It draws their attention and tames them to dwell on their own perceptions.

Group Discussions: Artworks such as El 3 de mayo de 1808 create discussions that compel learners to think critically while using opinionated expressions in Spanish. It also encourages collaboration and communication among students where they exchange thoughts and opinions.

Scavenger Hunts: Tasks like "Locate a Goya painting and narrate its emotional tone" encourage self-directed learning and self-explanatory activity.

4.4 Promoting Active Learning and Critical Thinking

Comparing Artistic Styles: When analyzing the works of Velázquez and Goya, students may notice differences in the use of light and shadow.

Personal Interpretation: Students can examine power and surrender after watching La rendición de Breda, and use phrases like "desde mi perspectiva" with their commentary.

5. PRACTICAL CONSIDERATIONS FOR IMPLEMENTING VIRTUAL MUSEUM VISITS

Here, we make some suggestions on how to implement Virtual Museum Visits as a teaching tool in a Spanish as a Foreign Language teaching class.

5.1 Pre-Visit Preparation

After selecting the students, teachers must provide the appropriate vocabulary and collections to ensure students understand what the key Prado Museum covers before the virtual tour. The provision of short videos or readings before a visit may be very useful.

5.2 Activities During the Virtual Visit

Learners can do active exercises like responding to discussion questions or participating in a set discussion based on a piece of art.

5.3 Post-Visit Reflection

By inviting learners to prepare a written or oral presentation in Spanish from their observations, they are able to integrate different facets of knowledge they have acquired.

Art topics can be eloquently presented as progress indicators for masters level students' language proficiency.

6. CONCLUSION

The combination of language and culture can be seamlessly achieved through virtual tours of the Prado Museum. By integrating language from Spain's art perspective, teachers can foster a more effective and stimulating educational atmosphere. As technology advances, so do the resources available, and virtual tours of museums will most likely form a vital part of teaching English as a foreign language in Spain, offering students valuable interactive learning opportunities to improve their language and culture skills.

References

[1] Falk, J. H., & Dierking, L. D. (2012). *The Museum Experience Revisited*. Left Coast Press.
[2] Kramsch, C. (1993). *Context and Culture in Language Teaching*. Oxford University Press.
[3] Stockwell, G. (2012). Technology and Language Learning: A Guide for Teachers and Learners. Routledge.
[4] Stern, H. H. (1992).*Issues and Options in Language Teaching*. Oxford University Press
[5] Lamy, M.-N., & Hampel, R. (2007).*Online Communication in Language Learning and Teaching*. Palgrave Macmillan.
[6] Leipert, M., & Smidt, S. (2020). "Virtual Field Trips: A Means of Enhancing Cultural Understanding in Language Learning." *International Journal of Educational Technology in Higher Education, 17*(1), 1-15.
[7] Miller, J., & Murphy, T. (2018).Digital Storytelling in the Language Classroom: A New Tool for Spanish Education. Modern Language Journal, 102(2), 219-235.
[8] Moncada, E. (2015). "The Integration of Museum Resources into Foreign Language Learning." *Foreign Language Annals, 48*(2), 258-276
[9] García-Sánchez, I. M., & García-Sánchez, L. A. (2018). "The Use of Online Museum Tours in Spanish Language Learning." *Journal of Language Teaching and Learning, 12*(3), 45-61.
[10] Bennett, M. J. (2011).Developing Intercultural Competence Through Education. Peter Lang Publishing.

The Influences of Grammarly on EFL Students' Writing Performance: A Case Study at a University in Vietnam

Thi Thanh Ha NGUYEN [a*]

[a]*University of Economics – Technology for Industries, Faculty of Foreign Languages,*
No. 456 Minh Khai, Hai Ba Trung, Hanoi, Vietnam, Corresponding email: nttha.khoann@uneti.edu.vn

Abstract

This study investigates the impact of Grammarly, an automated writing assistant, on the writing performance of EFL students at a Vietnamese university. Using a quasi-experimental design, 44 undergraduate students were divided into experimental (n=22) and control (n=22) groups. The experimental group received traditional writing instruction supplemented with Grammarly, while the control group received only traditional instruction over eight weeks. Pre-test and post-test writing samples were evaluated using an analytical scoring rubric. Results revealed that while both groups demonstrated significant improvement in writing performance, the experimental group showed substantially more significant gains than the control group. These findings suggest that Grammarly is an effective supplementary tool for enhancing EFL writing skills, particularly in addressing grammatical accuracy and structural issues. The study contributes to understanding how automated writing assistants function as scaffolding mechanisms within Vygotsky's (1978) Sociocultural Theory while reducing cognitive load, as Sweller's (1988) Cognitive Load Theory proposed. Recommendations include integrating Grammarly as a complementary tool rather than a replacement for traditional instruction and conducting further research on its long-term effects.

Keywords: Grammarly, EFL writing performance, Vietnamese EFL students

1. INTRODUCTION

Technology has emerged as a revolutionary transformation throughout the fast-changing English language education setting, which modifies established educational methods and student learning paths. Automated writing assistants now receive prominent attention within technological innovations because they show promise to boost the writing skills of English as a Foreign Language (EFL) students. One of the most popular automated writing assistants, Grammarly, provides users with immediate feedback concerning grammar, spelling, punctuation, and style elements, which could help EFL learners resolve their long-standing problems with written communication. Writing proficiency is demanding for EFL students when their access to genuine English usage remains restricted. The writing abilities of Vietnamese students learning English as a foreign language share the typical difficulties that arise when non-English speakers attempt writing according to English standards (Nguyen & Nguyen, 2017). Students in Vietnamese EFL classrooms face additional problems due to overcrowded classes, insufficient individual input, and insufficient practice time, which is common in Vietnamese EFL classes (Le, 2019). Educators and researchers persist in developing innovative strategies for enhancing traditional writing instruction because they want to provide students with additional support systems.

* Corresponding author.

Manual writing support tools such as Grammarly offer an appealing solution to overcome these issues. Researchers have established that automated writing assistants deliver advantages in different learning environments. According to O'Neill and Russell (2019), university students decreased their writing surface errors through Grammarly, yet Cavaleri and Dianati (2016) demonstrated higher grammatical accuracy levels in EFL learners who used the platform. The scarcity of scholarly research investigating how Grammarly affects Vietnamese EFL students' writing performance generates a fundamental knowledge gap regarding this technology's operations in that educational setting. The current studies that evaluated different Grammarly features produced inconsistent outcomes because of limited research regarding its effectiveness with Vietnamese EFL students. The research findings about Grammarly diverge between positive outcomes for detecting and correcting errors (Qassemzadeh & Soleimani, 2016) and failure to address advanced writing quality components, including coherence and argumentation (Ghufron & Rosyida, 2018). Research on the theoretical bases that link automated writing assistants to academic outcomes still requires more development, specifically regarding established theories of second language acquisition and cognitive growth.

The research targets these knowledge gaps by examining how much Grammarly influences the writing abilities of English as a Foreign Language learners at a Vietnamese educational institution. This research measures writing improvement between students who use Grammarly and those following traditional writing instructions to verify the tool's effectiveness at the Vietnamese university. The research applies Vygotsky's Sociocultural Theory and Sweller's Cognitive Load Theory to develop theoretical knowledge about automated writing assistant effects on language learning processes. The investigation abides by this research question: *How significantly does Grammarly affect the writing skills of EFL students?*

2. LITERATURE REVIEW

2.1. The role of technology in EFL writing

Using technology in English as a Foreign Language (EFL) writing instruction contributes to significant teaching practice changes, yet it presents benefits and difficulties. Research validates that technological resources boost students' writing output, actively participate in assignments, and teach them self-guided learning methods. According to Al-Wasy (2020), who analyzed 18 empirical studies containing information from 1,281 EFL learners, technology integration resulted in substantial writing skill improvement. The study showed that high school and university students obtained the most benefits, particularly when drafting and editing their papers. The effect measures of beginner learners indicated that technology-based interventions deliver more substantial results to advanced English language learners. Chang and Szanajda (2016) performed a quasi-experimental research that merged blogs into EFL writing courses following the process-genre approach. Student writing skills showed improvement through blog writing and more positive attitudes toward blog writing tasks. Interactive blogs enabled students to receive higher peer language exposure and more possibilities for written interaction, thus supporting their writing skills. Yabukoshi and Mizumoto (2024) investigated the technology preferences and academic writing problems of EFL learners at a university through their study. The survey demonstrated that university students primarily relied on smartphones for academic purposes, including Weblio, Google Translate, and online dictionary platforms. Students with access to various technological resources still face problems organizing their work, generating texts, and correcting errors, which shows that accessing technology does not automatically fix writing problems. According to Angraini et al. (2024), a systematic research investigation examined articles from the SIELE Journal from 2019 through 2024 that studied technology usage in EFL writing education. Through their review, researchers discovered that educational technology raises student involvement, improves writing skills, and generates better instructional feedback. Successful deployment of digital tools in education faces impediments because students and instructors exhibit different digital proficiencies, schools have inconsistent technical resources, and teachers need complete training programs.

2.2. Grammarly as an AI-based writing assistant

The AI-based assistant tool Grammarly offers users live feedback, which helps them improve their grammar, style, and vocabulary. The key functions of Grammarly assist writers with grammar correction, locating mistakes and fixing them, and style enhancement through clarity and conciseness recommendations and vocabulary suggestions that present alternatives for language variety. The program's component features help writers create more precise texts while attaining a professional finish. Research supports that Grammarly demonstrates the beneficial potential for EFL students who write in English. Research shows that this technology produces significant declines in grammatical errors, thus improving the

quality of students' compositions. A study by Dizon and Gold (2023) found that EFL students who used Grammarly decreased language anxiety while enhancing their learning self-governance because the tool proved to have benefits beyond error detection. The real-time feedback feature of Grammarly enables students to learn independently by showing them their mistakes instantly to correct them immediately. The quick feedback from Grammarly helps students improve their writing abilities because they better recognize standard writing mistakes and improve their skills in subsequent assignments (Bailey & Lee, 2020; Dizon & Gold, 2023; Huang, 2020). Particularly, users appreciate that Grammarly provides not only an easy-to-use interface but also convenient accessibility. That results in enabling learners to conduct revisions faster and with less effort (Thi & Nikolov, 2021).

However, several drawbacks exist to using Grammarly in EFL writing practice. The main issue lies in students leaning too much on the tool because this learning behavior restricts their ability to build self-editing competence and master language rules. The continuous use of Grammarly hinders student learning when users accept solutions instead of studying grammatical rules (Dizon & Gayed, 2021). Fan (2023) argued that Grammarly's fixing-errors highlights makes students lose interest in advanced writing aspects such as coherence and cohesion or lexical resources. Grammarly's primary feedback focuses on grammatical precision, but it provides weak support for developing advanced writing abilities, including content structure and persuasive arguments. Although Grammarly successfully identifies errors at the surface level, the tool fails to improve students' writing coherence and cohesion (Ghufron & Rosyida, 2018). Students suffer from restricted Grammarly functionality due to technical problems and premium-blocking limitations that affect their writing benefits, specifically in regions with resource constraints (Dewi, 2023).

2.3. Theoretical framework

Through sociocultural and cognitive load theoretical perspectives, we can study how Grammarly supports language learning while improving writing skills in EFL instruction.

According to Vygotsky's Sociocultural Theory, cognitive development occurs through social context-mediated activities that involve cultural artifacts (Vygotsky, 1978). Within this theoretical structure, tools and signs act as bridging mechanisms that help students perform their cognitive activities. Grammarly is an educational tool that promptly evaluates writing content to provide personalized feedback that helps learners advance their writing abilities. The real-time feedback from Grammarly helps students detect their grammatical mistakes, which leads to better writing accuracy (Dizon & Gold, 2023). The educational approach known as Zone of Proximal Development (ZPD) demonstrates that appropriate support allows students to accomplish tasks that surpass their independent abilities, as described by Vygotsky. The suggestions provided by Grammarly function as training steps to help students advance inside their ZPD as they learn language norms. The tool assists language learning by reducing foreign language anxiety and promoting learner autonomy, thus reinforcing its position as a sociocultural aid for language education development (Dizon & Gold, 2023).

Swetler's Cognitive Load Theory (1988) explains that learning can reach its full potential through educational strategies which reduce superfluous cognitive workload so learners can focus their mental resources on core activities. Learned writings in foreign languages require mental effort because mastering their grammar structures and vocabulary and syntactical elements creates an inherent cognitive burden. Grammarly relieves such cognitive strain by performing automatic error identification and correction thus freeing students to concentrate on advanced writing elements like content generation and structural organization. Automation in Grammarly's feedback system decreases learner editing work to enhance written content quality (Bailey & Lee, 2020). Learning students can optimize their ideas and argument development through Grammarly since it takes care of standard error management which enables better composition quality. The use of Grammarly as an educational tool results in better cognitive processing according to Sweller's theory because it diminishes unnecessary mental loads when composing in EFL writing environments.

2.4. Previous studies

Various scholarly research proves that Grammarly positively affects the performance of EFL writers. Mohammadi et al. (2023) revealed that Iranian students with Grammarly support performed better in writing tasks than learners who received standard teaching methods. According to Sistani and Tabatabaei (2023), automated feedback generated by Grammarly improved writing abilities in Iranian EFL learners while matching teacher-generated feedback under specific situations. Research has evaluated how Grammarly affects writing tasks and investigates its effects on affective aspects. Dizon and Gold (2023) analyzed how Grammarly affected Japanese EFL students' foreign language anxiety (FLA) and learner autonomy traits. Through their research, they found that Grammarly usage decreased foreign language anxiety while enabling autonomy, concluding that automated writing assessment tools positively impact students' emotional

control and self-regulation. Nhi and Phuong (2023) researched EFL postgraduate students who judged the effectiveness of Grammarly and peer feedback for enhancing academic writing abilities. All test subjects recognized Grammarly's quick detection of grammatical mistakes, yet they chose peer feedback since it provided detailed, constructive information. The research demonstrates that automated tools function best with human feedback for teachers to provide complete writing improvement techniques. Abu Guba et al. (2023) examined how Grammarly worked for beginners in learning English as a foreign language. According to the research results, Grammarly enhanced complete writing ability, although specific word formation and usage mistakes remained unchanged because more trained-based instructional methods surpass computer-based support. Dewi (2023) investigated EFL students' perspectives on using Grammarly for automated writing evaluation. Most students had favorable experiences with Grammarly, which provided quick feedback about grammar, punctuation, and vocabulary functions. The software presents valuable assistance to writers, but its free version has limited capabilities while occasional technical errors emerge, thus demonstrating that Grammarly requires enhancement to become a complete writing instruction solution. However, lower-proficiency EFL students demonstrated limited critical thinking about Grammarly feedback (Anastasia et al., 2024). Students might miss the complete benefit of automated feedback tools without proper guidance.

3. METHODOLOGY

3.1. Setting

The current research took place within the framework of a public Vietnamese university where researchers studied 44 English-major sophomore students. Thirty-four females represented 77.3% of the participants, and 10 males comprised 22.7% of the group. All participants were between 19 and 22 years old. The researcher categorized all 44 participants into separate groups where the experimental section included 23 students, who comprised 52.3% of the sample, and the control group consisted of 21 students, who made up 47.7%. The participants studied English for 3–5 years with 34.1%, whereas 22.7% studied for 6–8 years, and 43.2% studied English for more than 9 years. The instructional material for this course was "Great Writing 2," published by National Geographic Learning. The textbook devotes its content to paragraph development while teaching students to properly design paragraphs along with description techniques, rationale methods, agency approaches, and sentence unity. The program provides exercises that develop vocabulary skills and grammar instruction to establish fundamental abilities in academic writing.

The teaching environment matches the broader English education system of universities across Vietnam. English education is a mandatory subject and independent study area within Vietnamese university programs, but their credit load ranges differently. The number of English courses in non-English major programs generally remains limited, which affects how much prominence English proficiency receives (Le, 2019). The designed curriculum for English-major students builds complete language skills through an educational path leading to vocational requirements involving advanced English proficiency. The structured curriculum at Vietnamese educational institutions does not solve all the problems that affect English education across Vietnam. The obstacles that hinder English teaching progress include adopting conventional teaching models, excessive class-student limits, and funding shortages to support sophisticated teaching approaches (Le, 2019). Using Grammarly could improve writing education by providing students with real-time customized feedback, which assists in overcoming specific instructional issues present in this learning environment.

3.2. Research design and data collection instrument

A quasi-experimental research design was applied to assess how Grammarly affected the writing ability of EFL students. When researchers cannot implement random participant assignment in experimental and control groups, they use quasi-experimental design methods, which evaluate causations through similarities between non-randomized groups (Scribbr, 2020). Educational settings often utilize this research protocol since the existing classrooms determine the structure of group assignments. Grammarly's effectiveness was studied through writing tests before and after testing both experimental and control group participants. The assessment of participants occurs before and after an intervention, enabling researchers to measure changes in their skills and knowledge through pre- and post-testing methods. This research used tested materials from a question bank that received institution-level administrator approval as writing assessment tools. The institutionally approved instruments serve to increase measurement accuracy because they were explicitly designed to evaluate student performance within educational settings.

3.3. Data analysis

The evaluation of writing performance changes for English as a Foreign Language students because Grammarly involved using IBM SPSS Statistics version 27.0 and both independent samples t-tests and paired samples t-tests. The researcher evaluated mean differences between the experimental group, which utilized Grammarly, and the control group, which did not use independent samples t-test across pre-test and post-test measures. The independent samples t-test evaluates mean differences between two unassociated groups provided that the continuous dependent variable shows standard distribution patterns and both groups share equal variance levels. A t-test analysis followed the execution of Levene's Test for Equality of Variances, which determined the interpretation of t-test results according to its results. Students' writing performance was assessed through the paired samples t-test because researchers compared students' writing scores before and after the intervention by viewing participants as matched pairs. The paired observations test determines if the mean difference between corresponding values (previous scores vs. subsequent scores) is statistically different from zero by assuming a normal distribution of these pair differences. The research examined student writing performance changes by evaluating matching scores using Grammarly to measure individual writing improvement.

4. FINDINGS AND DISCUSSION

4.1. The comparison of the experimental and controlled learners' writing performance

Statistical data from the independent samples t-test showed no evidence of significant difference when comparing pre-test and post-test results between control and experimental groups.

Table 1. Independent sample t-test on the experimental and controlled learners' writing performance

Independent Samples Test

		Equality of Variances		t-test for Equality of Means						
									of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Pre	Equal variances assumed	1,499	0,228	-0,076	42	0,940	-0,0166	0,2187	-0,4578	0,4247
	Equal variances not assumed			-0,076	41,442	0,939	-0,0166	0,2166	-0,4538	0,4206
Post	Equal variances assumed	5,849	0,020	1,373	42	0,177	0,3236	0,2356	-0,1519	0,7991
	Equal variances not assumed			1,398	38,533	0,170	0,3236	0,2314	-0,1447	0,7919

An independent sample t-test with t = -0.076, df = 42, and p = 0.940 demonstrated that both groups possessed similar writing abilities before starting the study. The assumption of equal variances was checked out (F = 1.499 and p = 0.228). After the intervention, the independent t-test analysis (t = 1.373, df = 42, p = 0.177) revealed no statistically significant difference in writing performance between the control group and experimental group, although a violation of the Levene's test for equality of variances (F = 5.849, p = 0.020) did occur. The research indicates Grammarly offers writing feedback, yet its impact on performance enhancement between groups remains insignificant.

The findings matched what scholars had discussed about AI-based writing aids through Grammarly. The benefits students achieve from Grammarly depend on how much they interact with its features, including grammar correction, style enhancement, and vocabulary suggestions. Research evidence indicates Grammarly helps lower errors and improve grammar but does not help improve advanced writing competencies like coherence or argumentation. The research evidence from this study shows that Grammarly does not substantially improve writing abilities until students implement feedback for learning instead of accepting automated responses. The study results support established theoretical models that explain writing enhancement through Grammarly. As per Vygotsky's Sociocultural Theory (1978), Grammarly exists as an educational tool that delivers scaffolding support to students. System-dependent success exists for students to benefit from this tool, although how they apply and retain feedback remains important. The process-based writing assistance from Grammarly in managing linguistic errors reduces students' cognitive workload yet produces minimal observable consequences on total writing quality regarding the Cognitive Load Theory (Sweller, 1988). The study confirms that Grammarly enables writing support, but increased performance in writing requires students to actively learn from and reflect upon the feedback the tool offers. Future research must investigate students' use of Grammarly feedback and establish any persistent writing advancement resulting from its usage.

4.2. The differences of writing performance within the experimental and controlled groups before and after the intervention

The paired samples test demonstrates substantial writing performance improvement across both groups, yet they differ substantially in magnitude.

Table 2. Paired sample t-test to illustrate the differences of writing performance within the experimental and controlled groups before and after the intervention

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	of the Difference Lower	of the Difference Upper			
Pair 1	Post_exp. - Pre_exp.	0,6783	0,3288	0,0686	0,5361	0,8205	9,892	22	0,000
Pair 2	Post_con. - Pre_con.	0,3381	0,4006	0,0874	0,1557	0,5204	3,868	20	0,001

The experimental group's writing performance with Grammarly support, substantially exceeded that of the control group. Cross-group analysis showed that experimental participants experienced a mean score elevation of 0.6783 (SD = 0.3288) to the degree that surpassed statistical significance with a t-value of 9.892 alongside a pdf of 22 and a p-value below 0.001. The control group participants also demonstrated statistically important progress through their 0.3381 mean score rise (SD = 0.4006), as confirmed by the t = 3.868 test statistic with df = 20 and p = 0.001. The data illustrates that experimental subjects achieved approximately twice the growth rate of control participants during the study period. The experimental group demonstrated a more impactful intervention because its t-value was higher than the control group. The conclusion supports the impact of Grammarly on student writing based on an analysis of difference confidence intervals, which shows that experimental students scored between 0.5361 and 0.8205 points better than control students, who scored between 0.1557 and 0.5204 points.

The research findings match what previously emerged regarding AI-based writing assistants that benefit EFL learners. The study by Mohammadi et al. (2023) showed that students who used Grammarly achieved better results than their peers receiving traditional teaching. According to Sistani and Tabatabaei (2023), student writing skills improved when using Grammarly automated feedback at similar levels to teacher-provided feedback across specific writing scenarios.

Vygotsky's (1978) Sociocultural Theory explains the experimental group's noteworthy improvement because Grammarly operates as an educational tool that provides supporting scaffolds for students' learning process. The immediate feedback from Grammarly assists learners in recognizing their grammatical mistakes, which, in turn, results in better writing accuracy (Dizon & Gold, 2023). Grammarly offers students training steps derived from written suggestions, enabling them to develop their language proficiency inside their ZPD. The technology analyzes writing errors automatically and fixes them to enable students to prioritize their attention on enhancing content development alongside structural planning according to Cognitive Load Theory (Sweller, 1988). By reducing the participants' cognitive and mental workload, the experimental students demonstrated superior progress to those in the control group.

Although the experimental group demonstrated significant progress, the study still shows evidence of the limitations that researchers have previously found with Grammarly. The identification system of Grammarly recognizes surface-level issues yet remains insufficient for advancing student writing coherence and cohesion (Ghufron & Rosyida, 2018). The engineering approach Grammarly uses for error identification (Fan, 2023) results in students neglecting more complex aspects of writing, such as coherence and lexical resources. The experimental group achieved substantially more significant improvement, due to which Grammarly demonstrates potential as a supportive system to enhance EFL writing skills when teachers integrate it into their teaching strategies. Dizon and Gold (2023) agreed with findings showing that automated writing assessment technology improves student emotional control and self-regulation, leading to better writing results. While Grammarly does not replace conventional instruction, it proves to be an indispensable additional resource for EFL writing education.

5. CONCLUSION

The research examined how Grammarly affected writing abilities among EFL students at a Vietnamese university. Writing performance considerably improved in both experimental and control groups, but the experimental participants using Grammarly achieved significantly higher gains than the control group. The study evidence indicates that Grammarly functions successfully as an additional tool for developing EFL writing proficiency due to its ability to detect

grammatical mistakes and structural problems. The substantial educational gains in the experimental group match both Vygotsky's Sociocultural Theory and Sweller's Cognitive Load Theory. Grammarly functions as an automated scaffolding system that detects mistakes directly in students' development zones to limit their mental workload so they can concentrate on complex aspects of writing. Writing skill development thrives in EFL learners because they receive two advantages from this method.

Nevertheless, the study presents promising results but requires attention to various specific research limitations. The non-randomized assignment method used in the quasi-experimental design introduces selection bias, which damages the internal validity of research findings. The findings from the research have limited external validity because the study worked with only 44 participants. The study limited its analysis to numerical writing assessment data since it failed to assess students' perspective on their writing process and their particular writing enhancement or engagement dynamics. The research duration was potentially insufficient for studying how well students maintained their writing skills and whether Grammarly-assisted learning lasted post-usage. The research omitted controlling factors such as students' experience with technology, expressive characteristics, and their outside English-language learning activities that might affect writing development.

Following the analysis and restrictions discovered, various suggestions arise for educational practice and research. Educators of EFL must incorporate Grammarly as an aid but not as a substitute for their standard teaching procedures. Students should learn to evaluate Grammarly feedback independently from the help of teachers by analyzing the key principles of grammar following a review of the suggestions. Institutions should obtain premium Grammarly accounts due to the free version restrictions described by Dewi (2023). Further research must use mixed research approaches to study the quantitative performance benefits and the qualitative aspects experienced with Grammarly use. Extended research over time will demonstrate if students maintain their writing progress while evaluating their dependency on such tools. Researchers need to examine how diverse linguistic abilities react to Grammarly's automatic responses to text and assess which particular writing elements show the best results from this tool-based feedback. Research that compares Grammarly to other AI writing assistants would offer important knowledge about different technological solutions for EFL writing education. This research would benefit the development of extensive models for technological tools in language education.

References

[1] Abu Guba, M. N., Awad, A., & Abu Qub'a, A. (2023). Grammarly in teaching writing to EFL learners at low levels: How useful is it? *World Journal of English Language, 14*(3), 1–12. https://doi.org/10.5430/wjel.v14n3p1

[2] Al-wasy, B. Q. (2020). The effectiveness of integrating technology in EFL/ ESL writing: A meta-analysis. *Interactive Technology and Smart Education, 17*(4), 435 454. https://doi.org/10.1108/ITSE-03-2020-0033

[3] Angraini, D., Sulistiyo, U., Haryanto, E., & Riady, Y. (2024). Integration of Technology in EFL Writing Instruction: A Systematic Review of Insights from SIELE Journal Articles. *PPSDP International Journal of Education*, *3*(2), 302–320. https://doi.org/10.59175/pijed.v3i2.317

[4] Anastasia, W., Murtisari, E., & Isharyanti, N. (2024). Lower-proficiency EFL students' use of Grammarly in writing: Behavior, cognition, and affect. *The JALT CALL Journal, 20*(1), 1–25. https://doi.org/10.29140/jaltcall.v20n1.1089

[5] Bailey, D. R., & Lee, A. R. (2020). An exploratory study of Grammarly in the language learning context: An analysis of learner perceptions. *Journal of Educational Computing Research*, 58(5), 1051–1064.

[6] Cavaleri, M. R., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning, 10*(1), A223-A236.

[7] Chang, W.-Y., & Szanajda, A. (2016). How computer technology transforms writing performance: An integration of the process/ genre approach and blogs in EFL writing courses. *International Journal for 21st Century Education*, 3(Special), 169-185. https://doi.org/10.21071/ij21ce.v3iSpecial.5715

[8] Dewi, R. S. (2023). Challenges in integrating Grammarly into EFL writing instruction: Student perspectives. *Journal of Language Teaching and Research*, 14(3), 456–463.

[9] Dizon, G., & Gold, J. (2023). Exploring the effects of Grammarly on EFL students' foreign language anxiety and learner autonomy. *The JALT CALL Journal*, 19(3), 299–316.

[10] Dizon, G., & Gayed, J. M. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *The JALT CALL Journal*, 17(2), 103–120

[11] Fan, N. (2023). Exploring the effects of automated written corrective feedback on EFL students' writing quality: A mixed-methods study. *SAGE Open*, 13(2), 23-43.

[12] Ghufron, M. A., & Rosyida, F. (2018). The role of Grammarly in assessing English as a foreign language (EFL) writing. *Indonesian Journal of Applied Linguistics*, 8(2), 372–381.

[13] Han, Y., Zhao, S. & Ng, L-L. (2021). How technology tools impact writing performance, lexical complexity, and perceived self-regulated learning strategies in EFL academic writing: A comparative study. *Front. Psychol. 12:752793*. doi: 10.3389/fpsyg.2021.752793

[14] Huang, H.-T. D. (2020). Learner perceptions of using a mobile application for English pronunciation learning: A case study in Taiwan. *Journal of Educational Computing Research*, 58(5), 978–1007.
[15] Le, V. C. (2019). Current English education in Vietnam: Policy, practices, and challenges. In A. Burns & J. C. Richards (Eds.), *International perspectives on English language teaching* (pp. 167–180). Palgrave Macmillan.
[16] Mohammadi, R., Borna, P., & Zoghi, R. (2023). The effect of employing Grammarly on Iranian EFL learners' writing performance. *International Journal of Research in English Education, 8*(5), 187–197. http://ijreeonline.com/article-1-869-en.html
[17] Nguyen, T. H., & Nguyen, T. N. (2017). Challenges of learning to write in English: A case study of Vietnamese EFL learners. *European Journal of English Language Teaching, 3*(1), 33-48. https://doi.org/10.5281/zenodo.1286168
[18] Nhi, H. Y., & Phuong, H. N. (2023). EFL postgraduate students' perceptions on the use of Grammarly and peer feedback to improve their academic writing skills. *International Journal of TESOL & Education, 3*(2), 1–14. https://doi.org/10.54855/ijte.25512
[19] O'Neill, R., & Russell, A. (2019). Grammarly: Help or hindrance? Academic Learning Advisors' perceptions of an online grammar checker. *Journal of Academic Language and Learning, 13*(1), A88-A107.
[20] Qassemzadeh, A., & Soleimani, H. (2016). The impact of feedback provision by Grammarly software and teachers on learning passive structures by Iranian EFL learners. *Theory and Practice in Language Studies, 6*(9), 1884-1894. https://doi.org/10.17507/tpls.0609.23
[21] Scribbr. (2020). Quasi-experimental design: Definition, types & examples. Retrieved from https://www.scribbr.com/methodology/quasi-experimental-design/
[22] Sistani, H., & Tabatabaei, O. (2023). Effects of teacher vs. Grammarly feedback on Iranian EFL learners' writing skill. *International Journal of Foreign Language Teaching and Research, 12*(45), 85–97. https://journals.iau.ir/article_703271.html
[23] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
[24] Thi, N. P., & Nikolov, M. (2021). Evaluating digital tools for the EFL classroom: Grammarly as a case study. *International House Journal of Education and Development*, 52.
[25] Tran, T. T. T. (2025). Enhancing EFL writing revision practices: The impact of AI- and teacher-generated feedback and their sequences. *Education Sciences, 15*(2), 232. https://doi.org/10.3390/educsci15020232
[26] Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
[27] Yabukoshi, T. & Mizumoto, A. (2024). University EFL learners' use of technology and their perceived difficulties in academic writing. *LET Kansai Chapter Collected Papers*, 22, 117-129. https://doi.org/10.50924/letkansai.22.0_117

Transport, Logistics and Tourism (IAC-TLT)

Food Packaging and Health: Analysis of the Suitability and Harmfulness

Nijolė Janina VASILIAUSKIENĖ [a*], Aldona VOSYLIŪTĖ[b]

[a] *Kauno kolegija Higher Education Institution, Faculty of Medicine, Pramonės pr. 20, Kaunas, Lithuania, nijole.vasiliauskien@go.kauko.lt*
[b] *Kauno kolegija Higher Education Institution, Faculty of Business, Pramonės pr. 20, Kaunas, Lithuania, aldona.vosyliute@go.kauko.lt*

Abstract

Food packaging plays an important role in protecting products from external influences and ensuring their safety. However, due to the migration of chemicals into food, certain packaging materials may pose a health risk. This article discusses the most common harmful substances found in food packaging, their effects on human health, and safer packaging alternatives. Research shows that some materials used in packaging can present risks to human health. Migration of chemicals from packaging to food is one of the main problems. Another problem is the use of recycled materials in food packaging, which creates a risk of introducing residues of chemicals that were previously present in the packaging. Safer alternatives, such as biodegradable and natural packaging, need to be used to address these issues. Research shows that stricter control of packaging materials, clearer legislation and ongoing consumer education are essential to ensure food safety and health protection.

Keywords: food packaging, migration, plastics, bisphenol A (BPA), phthalates

1. INTRODUCTION

Food packaging is an integral part of the modern food industry, with important implications for both product protection and consumer health. Food packaging performs several important functions: it protects food from the environment, extends its shelf life, facilitates the transportation of food and informs consumers about the composition, nutritional value, proper storage and consumption conditions. However, the benefits come with potential risks, as some of the materials used in food packaging can have adverse health effects, especially if their constituents migrate into food.

The safety and environmental sustainability of food packaging has become increasingly important in recent decades. Research shows that certain chemicals used in plastic, metal, glass or paper packaging can have endocrine disrupting, carcinogenic or allergenic properties. For example, bisphenol A (BPA), phthalates and some perfluorinated substances have been linked with endocrine disruption and other health risks. This has led to a growing number of consumers and regulators to question the suitability and potential harm of food packaging.

This study aims to analyse the sources of scientific information on the suitability of food packaging materials and their potential harmful effects on human health, and propose solutions to reduce adverse health effects of food packaging.

* Corresponding author.

The objectives to achieve this are to:

1. Discuss the composition of commonly used food packaging materials and their interaction with food;
2. Identify which chemicals migrate into food and how they may affect human health;
3. Analyse the use of recycled and natural materials in packaging and their impact on food safety;
4. Make recommendations on the regulation of the production of packaging, the use of safer alternatives and the raising of consumer awareness.

Food packaging plays an important role in ensuring the safety and quality of food, but its potential impact on health and the environment must also be considered. To reduce the negative impact of food packaging, it is necessary to promote safer and more sustainable packaging materials, strengthen regulation and raise consumer awareness. Only a joint effort can help ensure a healthier and cleaner future.

It is hoped that this analysis will contribute to a broader understanding of the safety of packaging and will encourage the search for more sustainable and healthier alternatives, both in the food industry and in the choices made by consumers.

2. METHODOLOGY OF THE STUDY

The methodology of the study includes a number of important steps to provide a comprehensive assessment of the suitability of food packaging and its potential risks to health.

Previous studies on the health effects of food packaging are analysed in the literature review. Information sources have been collected from reputable scientific databases such as PubMed, ScienceDirect and Google Scholar.

The literature analysis covers:

- Properties of food packaging materials (plastic, paper, glass, metal) and their interaction with food.
- The effects of chemicals such as bisphenol A (BPA), phthalates and perfluorinated compounds on the human body.
- Research into biodegradable and innovative packaging alternatives.

In order to assess the safety of packaging materials, laboratory tests are carried out.

Experimental research includes:

- Compositional analysis: spectroscopic and chromatographic techniques are used for the identification of potentially harmful compounds in food packaging.
- Chemical migration studies: assessment of the extent to which chemicals in packaging can be released into food.
- Research on microplastics: impact of plastic packaging on the formation and migration of microplastic particles into food.

Through analysis of the scientific data collected, the aim is to establish the relationship between the type of packaging, its composition and potential health effects. Based on the results of the analysis of scientific sources, recommendations are made on the use of the safest food packaging to minimise adverse effects on consumer health. This methodology, based on the latest research and data, will allow a thorough assessment of the suitability and potential harmfulness of food packaging.

3. COMPOSITION OF FOOD PACKAGING AND ITS INTERACTION WITH FOOD

The composition of a food packaging is a key factor in determining its suitability for use and its potential impact on human health. As well as protecting products from external factors (moisture, light, microbes), food packaging can interact with the food. Packaging made of different materials has different physical and chemical properties that can have an impact on the safety, quality and shelf life of food products. The growing focus on human health and the environment is bringing new challenges to the use of food packaging. Research shows that some materials used in

packaging can have adverse health effects on consumers, particularly through the migration of chemical compounds into food. The use of synthetic substances such as bisphenol A (BPA) and phthalates in plastic packaging is one of the most pressing issues. These substances can interact with food, especially at high temperatures, according to Lim et al. (2019). In addition, these scientists note that the use of recycled plastics in food packaging further increases the concentration of harmful compounds, as the manufacturing processes of such materials do not always ensure their cleanliness.

Rodríguez-Ramos et al. (2024) point out that plastics are an integral part of modern life, finding applications in various industries, including packaging, agriculture, and food processing. Plastics (PET, HDPE, PVC, LDPE, PP, PS) are synthetic polymers that are commonly used in the packaging of food products because of their versatility, light weight and resistance. The composition and properties of each type of plastic are different. Plastics (polythene (PE), polypropylene (PP), polycarbonate (PC), including recycled materials) are analysed for their wide use and possible migration of BPA and phthalate into food (Wang et al., 2022; Vaferi et al., 2025; Urbelis and Cooper, 2021).

Polyethylene terephthalate (PET) is used to make bottles for drinks and sauce jars. Antimony trioxide (a catalyst used in production) may be present in these products. Microplastic particles and acetaldehyde can get into water when PET plastic bottles are reused or stored in the sun or at high temperatures.

High-density polyethylene (HDPE) is used in the manufacture of milk bottles, cream jars, juice bottles and oil containers. It is chemically inert but can absorb fat.

Polyvinyl chloride (PVC) is used in the manufacture of food films and bottles. It contains stabilizers that can release phthalates, which can migrate into fatty foods such as cheese and meat.

Low-density polyethylene (LDPE) is used for bread bags and plastic shopping bags. It is chemically inert but not resistant to high temperatures.

Polypropylene (PP), used in the manufacture of microwave boxes, snack packs and yoghurt containers, is resistant to high temperatures but sensitive to UV rays.

Polystyrene (PS) is used to make disposable cups and containers for food products. Hot fatty foods (e.g. kebabs, soups) can release styrene, which is potentially carcinogenic.

Microplastics are small plastic particles that can enter the human body through food and water. Although the long-term health effects of microplastics are not fully understood, there are concerns about potential toxicity and migration of chemicals into the body.

Nanomaterials such as nanoclay, titanium nitride or titanium oxide can improve the properties of packaging. However, there are concerns about their potential migration into food and health effects (ECHA, 2023).

Paper and board are natural materials made from wood fibres, easily recyclable and rapidly biodegradable in the environment, but they are not resistant to moisture and fat, so they are often coated with a layer of plastic, wax or aluminum. These coatings may release chemicals. The production of food packaging such as paper bags, food cartons, plastic-lined coffee cups and egg cartons is energy and water intensive. Most paper packaging (e.g. potato sacks) contains PFAS (perfluorinated compounds). These are resistant to grease and water, but their compounds can migrate into food and accumulate in the human body. Disposable cups often have an inner plastic (PE) coating that can release hazardous compounds when microwaved. Baking paper is often impregnated with fluorinated compounds (PFAS), which have the potential to migrate into food products. The packaging of butter may contain residues of paraffin, which can migrate into the fat over time. Bamboo boxes may be coated with synthetic lacquers which can release chemicals. Formaldehyde impregnation is sometimes used to increase the durability of wooden forks and knives.

Glass is made from natural materials (silica sand, limestone, soda ash). It is inert, does not interact with food, is 100% recyclable without loss of quality, but is heavy, increases transport costs and carbon emissions, is fragile and requires careful handling. Glass is the most environmentally friendly choice for long-lasting products (Hischer & Weidema, 2010). Glass packaging is the safest. However, the inner coating of the lid of glass containers may contain bisphenol A (BPA), which interacts with acidic products such as tomato sauce and jams.

Made from aluminium or steel, metal food packaging is durable and reusable. The metal is highly recyclable, resistant to moisture and weathering, but has a high energy footprint during production and high production and recycling costs. Metal used in canning may be coated with Bisphenol A to protect against corrosion. Canning jars often have an inner epoxy coating containing BPA. This can migrate into acidic produce such as pineapple or tomatoes. Aluminium foil can release aluminium ions when it comes into contact with acidic products (e.g. slices of lemon or tomato sauce).

Food packaging materials that have the potential to interact with food, especially when exposed to high temperatures, fats or acids, include the following:

- Bisphenol A (BPA), which is used in plastics and can be found in plastic bottles and food cans.
- Phthalates are used to soften plastics and are found in plastic packaging for fatty products such as cheese and butter.

- Perfluorinated compounds (PFAS), found in fast food packaging (such as hamburger boxes) and paper cups.

Chemical migration into food occurs by diffusion when certain plastic components (e.g. additives, monomers) migrate into food that contacts the packaging. Factors influencing this process include:

- Migration is accelerated by high temperatures (e.g. microwave or heating). It was found, that antimony concentrations in water stored in PET bottles increase at high temperatures (>40 °C) (Westerhoff et al., 2008).

- Plastics are more likely to interact with fatty, acidic or alcoholic foods. For example, phthalates, which are fat soluble, migrate more easily into oil than into water. Studies have shown that the concentration of phthalates in stored foods increases within 30 to 40 days (Tsumura et al., 2003). Chemical migration increases with prolonged storage of food in plastic packaging.

In a study by EPA (2015), the effects of phthalates in fatty foods were investigated. The results showed that phthalate migration increases significantly when food is stored in vacuum-packed plastic containers. The authors emphasized that these substances have carcinogenic properties and can cause fertility problems.

Another major concern is the use of lesser-known chemical compounds, such as perfluoroalkyl substances (PFAS), in packaging that is resistant to grease and water. Kwiatkowski et al. (2020) found that PFAS migrating from food packaging may be associated with thyroid dysfunction and immunodeficiency. Biedermann-Brem and Grob (2009) found that BPA migration in polycarbonate packaging increases when food is heated (>60 °C). Lower quality plastics, recycled materials, or packaging that has not been properly treated during the manufacturing process may contain higher levels of migrating compounds.

To reduce the risk of chemical migration and environmental contamination, next-generation food packaging is being developed (Kishore et al., 2022; Ruiz-de-Maya and Ferrer-Berna, 2025; Cruz and Varzakas, 2023; Maurizzi et al., 2022). Sustainable packaging is an important part of the circular economy for the reduction of environmental pollution, energy consumption and health impacts. These alternatives focus on using renewable materials, reducing resource consumption, and allowing packaging to be recycled or composted.

- Bioplastics are made from plant-based materials (such as PLA), but their durability and degradability still need to be studied. Renewable resources such as corn starch, sugar cane fibre or potato starch are used for production of disposable cutlery, cups, bags and disposable straws. Studies have shown (Álvarez-Chávez et al., 2012) that PLA is an effective alternative to single-use plastics, but that it is only biodegradable under certain conditions.

- Active packaging can extend the shelf life of foods such as meat and fish by releasing antibacterial compounds or controlling moisture.

- Edible packaging is made from seaweed, starch, or other natural ingredients and can be consumed together with food (e.g., edible chocolate cups). However, they have a short shelf life and are more expensive than traditional packaging.

Environmental concerns are also inextricably linked to food packaging issues. Traditional plastic packaging is not only harmful to the environment because it takes a long time to break down, but can also indirectly affect human health when microplastics enter the food chain (Torres et al., 2014). The development of alternatives, such as biodegradable materials, is a promising solution. However, these technologies have not yet been widely adopted due to economic and technological limitations.

4. MIGRATION OF CHEMICALS INTO FOOD AND THE IMPACT ON HEALTH

Several factors, including the type of packaging material, the type of food, storage conditions and temperature, affect the migration of chemicals from food packaging to food.

Bisphenol A (BPA) can migrate into food and act as an endocrine disruptor. It can cause hormone disruption and the following effects:

- Problems with the reproductive system, such as a decrease in fertility or a decrease in sperm count.

- Hormonal imbalance that results in a malfunction of estrogen and testosterone.

- Metabolic disorders, such as increased risk of obesity, diabetes, and heart disease.

Scientists have consistently shown that certain chemicals used in packaging can migrate into food and be a source of health problems. They (Lim et al., 2009; Muzeza et al., 2023) link bisphenol A (BPA), commonly used in plastics, to endocrine disruption and increased cancer risk. Researchers report that the migration of BPA into food increases by more than 60% when plastic packaging is stored at temperatures above 40 °C. According to researchers, this chemical acts as an endocrine disruptor that has been linked to type 2 diabetes and prostate cancer. In addition, studies in animals have shown that BPA can cause changes in behaviour and other health problems (Vilniaus Medicinos Portalas, 2025). Due to its potential health and environmental effects, the European Union (EU) has placed certain restrictions on bisphenol A. As of June 2011, BPA has been banned in the manufacture of baby bottles throughout the EU. Its use in thermal paper has also been restricted since January 2020. To protect the health of consumers, the EU has also set limits on the migration of BPA from food packaging into food (ECHA, 2023). The USA Food and Drug Administration (2018) is restricting the use of BPA in certain products. Packaging material migration limits (e.g. BPA $\leq$ 0.05 mg/kg food) are in place in the EU and USA.

Phthalates can migrate into food and cause endocrine disruption and liver and kidney disease. Their negative effects are:

- Decreased intellectual potential and increased risk of autism spectrum disorders in children.
- Reproductive disorders, such as reduced levels of testosterone in men and an increased risk of infertility.
- Metabolic disorders that are associated with an increased risk of developing type 2 diabetes and heart disease.

Lim et al. (2009) confirm the direct effects of these compounds on fertility and the endocrine system. Researchers state that fatty foods that are stored in plastic packaging contain phthalates at a level of 2 to 5 mg/kg. The fact that phthalates can cause reproductive disorders and have long-term effects on children's development is confirmed by many scientists (Basso et al. 2022; Minatoya and Kishi, 2021; Vilmand et al., 2023; Sharpe, 2024; The World Health Organization (2010) and the United Nations are promoting the reduction of phthalates in food packaging and the search for alternatives. In Lithuania, the use of phthalates is also subject to regulation for the protection of human health and the environment.

Perfluorinated compounds can also migrate into food and accumulate in the body, causing the following health problems:

- Excessive levels of PFAS can be the cause of liver inflammation and even cancerous lesions.
- PFAS are associated with an increased risk of thyroid and kidney cancer.
- PFAS weaken the immune system and the body's resistance to infection.

Kwiatkowski et al. (2020) note that PFAS compounds used as grease-resistant coatings migrate into foods even at low temperatures, and their health effects have been linked to thyroid dysfunction and elevated cholesterol levels.

In the EU, materials that are in contact with food are subject to regulations to ensure that chemicals from the packaging do not migrate into the food in dangerous quantities. However, some substances, such as bisphenol A and DEHP (bis(2-ethylhexyl) phthalate), are of concern because of their potential health effects. Their use is therefore restricted.

Certain types of food packaging, especially those made of paper or recycled plastic, may contain heavy metals such as lead, cadmium or arsenic. The accumulation of heavy metals in the body can cause:

- Neurological disorders associated with memory loss and changes in behavior.
- Impaired kidney function, increased risk of stone formation.
- Developmental disorders in children such as stunted growth and impaired cognitive function.

The migration of chemicals from food packaging into food can pose a risk to human health. This is particularly the case when substances with endocrine disrupting properties are used. To minimise the migration of chemicals into food and protect consumer health, it is important to comply with regulations and use safe packaging materials.

Considering the health risks posed by chemicals migrating to food, it is recommended to:

- Avoid plastic containers, which contain BPA and phthalates, and choose glass or stainless steel instead.
- Do not heat food in plastic containers as high temperatures can increase the migration of harmful chemicals into the food.
- Avoid fast food packaging, especially if it is coated with a perfluorinated compound (PFAS).
- Choose recycled and biodegradable packaging as it is less polluting and safer for health.
- Pay attention to packaging labels and choose products with 'BPA-free' and 'PFAS-free' claims.

In order to ensure food safety and consumer health, it is essential to carefully assess the packaging materials in use, promote the introduction of safer alternatives and raise public awareness of potential risks.

5. USE OF RECYCLED AND NATURAL MATERIALS IN PACKAGING AND FOOD SAFETY IMPLICATIONS

In recent years, there has been a growing interest in the use of recycled and natural materials in packaging as a means of reducing environmental impact and ensuring food safety.

Recycled materials include recyclable plastics, paper, cardboard and metals. Their use reduces the amount of waste, but can also be a source of hazardous chemicals that can migrate into food products and pose a risk to health.

- Recycled plastics may contain residual chemicals from previous uses and are less stable (e.g. plastic bottles (PET), food cartons).
- Recycled paper may contain mineral oil residues (MOAHs, MOSHs) from ink (e.g. pizza cartons, fast food wrappers, egg cartons).
- Recycled metal is safer, but may have protective coatings containing BPA (e.g. cans, drinks cans).
- Recycled glass is chemically inert, but lids may contain BPA or other resins (e.g. jars with metal lids).

The main risk is the difficulty in ensuring full purity of the recycled materials. This can lead to the migration of harmful chemical compounds into the food. Studies show that recycled plastics often have higher concentrations of pollutants because hazardous compounds are not always removed during the manufacturing process (Kwiatkowski et al., 2020). This increases the risk of chemical release, especially when food is heated or stored in such containers. Therefore, recycled packaging must be rigorously tested to ensure it meets safety standards.

The production of environmentally friendly packaging often requires less energy and water than the production of conventional plastic packaging. This is particularly true in the food industry, where large volumes of packaging consume a lot of resources. Natural materials such as starch, cellulose, bioplastics, wax, bamboo or glass are becoming increasingly popular as greener alternatives. Natural materials often have a lower potential for migration of chemicals into food, which can make them safer for consumers, and they degrade naturally, reducing waste. However, they also have some disadvantages. Some natural materials may be less resistant to moisture, grease or mechanical damage, requiring additional coatings or additives that may pose a risk to food safety.

- Bioplastics (PLA, PHA) are biodegradable and contribute to the preservation of food products (e.g. disposable plates, straws).
- Waxed paper is safer than plastic, but may contain paraffin (e.g. baking paper, butter wrapping paper).
- Bamboo and wood can be impregnated with chemicals (e.g. boxes, tea packages).
- Glass is chemically inert and does not release chemicals (e.g. canned foods, drinks).

Packaging made from bio-based and natural materials can reduce the risk of chemical exposure, but has not yet been widely adopted due to high production costs (Torres et al., 2014). The main risk is that some natural materials may be less durable and require additional impregnation, which may have an impact on food safety. The EU and

Lithuania impose strict requirements on food packaging to ensure that it is safe to use. However, excessive regulation can be a barrier to business and innovation.

6. RECOMMENDATIONS

The packaging of food products plays a key role in ensuring the safety, quality and shelf life of food products. However, certain chemicals present in food packaging can migrate into food and pose health risks. Based on the analysis of the scientific research, the following recommendations are made to reduce the potential adverse health effects of food packaging:

- Avoid food in packaging that may contain harmful substances such as bisphenol A (BPA) or phthalates.
- Follow the instructions on the package, especially when heating food in the microwave. To prevent chemicals from migrating into the food, use only packaging that is designed for this purpose.
- Use safe packaging materials that comply with safety standards and are approved for use in contact with food.
- Regularly test the composition of packaging materials and their interaction with food to ensure that there is no migration of harmful substances.
- Strengthen regulation by reviewing and updating food packaging safety legislation in light of the latest research and technological developments.
- Promote research to support the identification of new safe packaging materials and technologies that reduce the migration of chemicals into food.
- Educate the public by informing consumers of the potential risks associated with certain packaging materials and encouraging them to choose safer alternatives.
- Promote sustainability by encouraging the use of sustainable, biodegradable and recyclable packaging that reduces environmental impact and ensures food safety.

EU regulations, such as Regulation 1935/2004, require that all materials intended to come into contact with food must be safe and stable and must not have an adverse effect on the composition, odour or taste of the food. Manufacturers are required to conduct tests to make sure that the packaging meets all of the safety requirements. These include chemical migration tests, which determine whether packaging releases harmful substances into food.

7. CONCLUSIONS

The most dangerous chemicals in food packaging that can migrate into food and cause long-term health effects are bisphenol A (BPA), phthalates, and perfluorinated compounds (PFAS). Bioplastics (PLA, PHA) are biodegradable and contribute to the preservation of food. Natural materials often have a lower potential for migration of chemicals into food, which can make them safer for consumers, and they degrade naturally, reducing waste.

The migration of chemicals from food packaging into food is a serious public health concern. Studies show that bisphenol A, phthalates, perfluorinated compounds, and heavy metals can have long-term health effects, particularly on the endocrine system, the liver, reproductive function, and the immune system. In order to reduce these impacts, there is a need for stricter regulation of food packaging and the use of safer alternatives.

The use of recycled and natural materials in food packaging can have both advantages and disadvantages in terms of food safety. Recycled materials often help reduce environmental impact. However, they may contain residual chemical compounds from previous uses. Natural materials (e.g., bioplastics, bamboo, glass) are safer, but they are less resistant to moisture and grease. When choosing the right packaging, the key is its contact with the food and the health implications that can result from it.

Ensuring that food packaging is safe for consumers and the environment requires stricter regulation of its production, promoting the use of sustainable materials, and increasing consumer awareness of the composition and health effects of food packaging. By following these recommendations, it is possible to reduce the potential adverse health effects of food packaging and ensure a high level of food safety.

References

[1] Álvarez-Chávez, C. R., Edwards, S., Moure-Eraso, R. & Geiser, K. (2012). Sustainability of bio-based plastics: General comparative analysis and recommendations for improvement. *Journal of Cleaner Production*, 23(1), 47-56.

[2] Basso, C. G., de Araújo-Ramos, A. T., Martino-Andrad, A. J. (2022). Exposure to phthalates and female reproductive health: A literature review. *Reproductive Toxicology*, Volume 109, 61-79. https://doi.org/10.1016/j.reprotox.2022.02.006

[3] Biedermann-Brem, S. & Grob, K. (2009). Release of bisphenol A from polycarbonate baby bottles: Water hardness as the most relevant factor. *European Food Research and Technology*, 228, 679-684.

[4] Cruz, R. M. S. and Varzakas, T. (2023). Frontiers on Sustainable Food Packaging. *Foods.* 12, 349. https://doi.org/10.3390/

[5] European Chemicals Agency (2023). Bisphenols. https://www.echa.europa.eu/hot-topics/bisphenols

[6] Europos Parlamento ir Tarybos reglamentas (EB) Nr 1935/2004, 2004 m. spalio 27 d. dėl žaliavų ir gaminių, skirtų liestis su maistu. https://eur-lex.europa.eu/eli/reg/2004/1935/oj

[7] Hayder, M., van Wezel, A. P., Gruter, G-J.M., Astefanei, A. (2024). What if you eat nanoplastics? Simulating nanoplastics fate during gastrointestinal digestion. *Chemosphere. Volume 365.* https://www.sciencedirect.com/science/article/pii/S0045653524021751

[8] Hischier, R., & Weidema, B. (2010). Life cycle assessment of packaging systems for beverages. *Environmental Impact Assessment Review*, 30(5), 380-391.

[9] Lim, D. S., Kwack, S. J., Kim, K. B., Kim, H. S., Lee, B. M. (2009). Potential risk of bisphenol A migration from polycarbonate containers after heating, boiling, and microwaving. *Journal of Toxicology and Environmental Health, Part A, Volume 72*. https://www.ncbi.nlm.nih.gov/pubmed/20077198

[10] Kishore, A., Mithul Aravind, S. , Singh, A. (2022). Bionanocomposites for active and smart food packaging: A review on its application, safety, and health aspects. *Journal of Food Processing Engineering.* DOI: 10.1111/jfpe.14320

[11] Kwiatkowski, C. F., Andrews, D. Q., Birnbaum, L. S., Bruton, T. A., DeWitt, J. C., Knappe, D. R. U., Maffini, M. V., Miller, M. F., Pelch, K. E., Reade, A., Soehl, A., Trier, X., Venier, M., Wagner, C. C., Wang, Z. and Blum, A. (2020). Scientific Basis for Managing PFAS as a Chemical Class. *Environmental Science & Technology Letters.* 532-543DOI: 10.1021/acs.estlett.0c00255.

[12] Maurizzi, E., Bigi, F., Quartieri, A., De Leo, R., Volpelli, L. A., Pulvirenti, A. (2022). The Green Era of Food Packaging: General Considerations and New Trends. *Polymers.* 14, 4257. https://doi.org/10.3390/polym14204257

[13] Minatoya, M. and Kishi, R. (2021). A Review of Recent Studies on Bisphenol A and Phthalate Exposures and Child Neurodevelopment. *International Journal of Environmental research and Public Health.* 18, 3585. https://doi.org/10.3390/ijerph18073585

[14] Muzeza, C., Ngole-Jeme, V. and Msagati, T. A. M. (2023). The Mechanisms of Plastic Food-Packaging Monomers Migration into Food Matrix and the Implications on Human Health. *Foods.* 12, 3364. https://doi.org/ 10.3390/foods12183364

[15] Rodríguez-Ramos, R., Santana-Mayor, A. Herrera-Herrera, A.V., Socas-Rodríguez, B., Rodríguez-Delgado, M. A. (2024). Recent advances in the analysis of plastic migrants in food. *Trends in Analytical Chemistry.* www.elsevier.com/locate/trac

[16] Ruiz-de-Maya, S. and Ferrer-Bernal, E. (2025). Sustainable food packaging: engagement through the public discourse on social networks. *Journal of Product & Brand Management.* 34/2 (2025) 251–264

[17] Sharpe, R. M. (2024). Endocrine disruption and male reproductive disorders: unanswered questions. *Human Reproduction*, 2024, 39(9), 1879–1888. https://doi.org/10.1093/humrep/deae143

[18] Torres, I. F., Bastida, F., Hernández, T., Bombach, P., Richnow, H. H., García, C. (2014). The role of lignin and cellulose in the carbon-cycling of degraded soils under semiarid climate and their relation to microbial biomass. *Soil Biology and Biochemistry.* 10.1016/j.soilbio.2014.04.007.

[19] Tsumura, Y., Ishimitsu, S., Kaihara, A., Yoshii, K., Nakamura, Y., & Tonogai, Y. (2003). Phthalates leaching from polyvinyl chloride food packaging. *Food Additives and Contaminants*, 20(5), 470-478.

[20] United Nations Environment Programme (UNEP). Plastics and their additives in ecosystems. https://www.unep.org/inc-plastic-pollution/session-5.2

[21] Urbelis, J. H and Cooper, J. R. (2021). Migration of food contact substances into dry foods: A review. *Food Additives & Contaminants: Part A, Vol. 38, No. 6, 1044–1073* https://doi.org/10.1080/19440049.2021.1905188

[22] U.S. Food and Drug Administration. (2018). Questions and Answers on Bisphenol A Use in Food Contact Application. https://www.fda.gov/food/food-packaging-other-substances-come-contact-food-information-consumers/questions-answers-bisphenol-bpa-use-food-contact-applications

[23] US Environmental Protection Agency/EPA (2015). Phthalates in consumer products. https://www.epa.gov/saferchoice/alternatives-certain-phthalates-partnership /

[24] Vaferi, B., Dehbashi, M., Yousefzadeh, R. & Alibak, A. H. (2025). Prediction of the packaging chemical migration into food anwater by cutting-edge machine learning techniques. Scientific Reports. doi.org/10.1038/s41598-025-92459

[25] Vilmand, M., Beck, I. H., Bilenberg, N., Andersson, A. M., Juul, A., Schoeters, G., Boye, H., Frederiksen, H., Jensen, T. K. (2023). Prenatal and current phthalate exposure and cognitive development in 7-year-old children from the Odense child cohort. *Neurotoxicology and Teratology,* Volume 96, 107161. https://doi.org/10.1016/j.ntt.2023.107161

[26] Vilniaus Medicinos Portalas (2025). Mokslininkai įspėja: buityje naudojami plastiko gaminiai gali kenkti sveikatai https://www.vlmedicina.lt/lt/mokslininkai-ispeja-buityje-naudojami-plastiko-gaminiai-gali-kenkti-sveikatai

[27] Wang, M., Liu, Y., Liang, G., Ding, H., Zhou, X., Qin, S., Zhao, P., Han, L. (2022). Migration analysis and health impact assessment of phthalates in takeaway food packaging materials. *Journal of Food Safety.* DOI: 10.1111/jfs.13021

[28] Westerhoff, P., Prapaipong, P., Shock, E., & Hillaireau, A. (2008). Antimony leaching from polyethylene terephthalate (PET) plastic used for bottled drinking water. *Water Research*, 42(3), 551-556.

[29] World Health Organization. (2010). Toxicological and Health Aspects of Bisphenol A. https://iris.who.int/bitstream/handle/10665/44624/97892141564274_eng.pdf

Rural Tourism as Alternative Solution to Overtourism

Andriy POPOVYCH [a*]

[a] *Independent Researcher, apopovych@kse.org.ua*

Abstract

The problems associated with tourism development have become an important topic for discussion in the world press and academia due to importance of the sector to the world economy. High demand for some tourist destination leads to overtourism, which compromises tourism resource base with negative consequences to the nature and the industry. Hence, there is a need of defining alternative forms of tourism to solve this problem from the standpoint of sustainable development. The study uses an aggregated model of innovation diffusion presented by logistic function and the data of UNWTO, European Travel Commission, etc., for some EU countries and regions. The quantitative and graphical results obtained allows concluding that ecological/rural/agri-/green tourism could be regarded as an alternative to the traditional mainstream tourism.

Keywords: sustainable tourism, innovation diffusion model, ecotourism, rural tourism, alternative tourism

1. INTRODUCTION

The world tourism industry grows fast. The number of international trips was increasing rapidly before and continues to grow after the pandemic, reaching 1.3 billion people worldwide in 2023, almost doubling over the past fifteen years. The total outbound tourism spending has increased and reached US$1.5 trillion, underscoring the importance of this sector for the global economy [13].

The increasing demand for tourist services causes excessive consumption of resources and leads to negative consequences for the local environment and economy. According to the World Travel and Tourism Council, up to 40% of international tourist trips in 2018 were to one of the 300 world's most popular sites [12]. As a result, many destinations are overwhelmed with number of tourists leading to the spread of anti-tourism sentiment.

The main socio-economic trends that contribute to excessive tourism are increasing physical and economic accessibility, the growth of well-being of an increasing mass of the planet's population, the development of social networks, the formation of a mentality of success and the need for impressions. Among the measures taken to combat excessive tourism, only various prohibitions and restrictions, tourist taxes stand out, which have not yet achieved the desired effect.

The negative trends in the world tourism has determined the necessity of finding the different scenario for the industry within the framework of sustainable development. The alternative approach to the tourism management recognizes the role of all stakeholders and the right of both current and future generations to use tourist resources in the long term. It contributes to the growth of social responsibility of the tourism business and reflects the

* Corresponding author.

evolutionary shift in behavioral patterns of consumers of tourist services. This study aims to support this view with empirical findings using the simple economic model.

2. THE PROBLEM OF OVERTOURISM AND ITS ALTERNATIVE

Overtourism is a phenomenon that major tourist destinations around the world are facing. The excessive tourism or overtourism occurs when the number of visitors exceeds the number of accommodation places and other important facilities, when the infrastructure of the tourist destination, the level of service and the value of authenticity suffer [14].

The World Tourism Organization (UNWTO) defines overtourism as “the impact of tourism on a destination or part of it that excessively affects the perceived quality of life of its citizens and/or the quality of visitors' experiences in a negative way”. The development of new technologies, transport infrastructure, population growth, fashion trends and the emergence of low-cost air carriers have increased the availability of travel and contributed to the growth of the tourism industry, including in countries with developing economies. However, the uncontrolled increasing concentration of tourists in certain destinations has a negative impact on both territories and local communities.

One of the main functions of the tourism industry is to provide tourists with a quality experience that aligns with their values and expectations, and is crucial for the long-term sustainable development of the territory. The quality of the tourist's consumer experience largely determines their satisfaction, which, in turn, seems possible if the emotional, social or epistemic value of the tourism product (service) is created [8].

Recently, ethics, moral values, concern for the environment and its ecosystems, as well as the desire to have a positive impact on local communities have become the determining factors when choosing a vacation spot for a tourist [4]. Tourists are increasingly striving to make conscious and responsible travel decisions, are less susceptible to the demonstration effect, and are more likely to visit little-known places rather than well-known tourist centers - "baits" that suffer from overtourism [10].

The concept of sustainability in tourism has become a widely used term and has gone beyond the narrowly defined concept of ecotourism to encompass all aspects of tourism [6]. This notion is used as an alternative to industrialized mass tourism, the main goal of which is to maximize profits. The sustainable tourism recognizes that the future of the tourism sector depends on the protection of life in all its diversity. At the same time, since cultural and natural resources are at the core of its activities, tourism is interested in combining efforts to overcome economic and environmental problems, for example, by preserving natural areas, which ensures an increase in tourism revenues [11].

Alternative forms of tourism are developing in many parts of the world but especially in the most developed countries and among the residents of large cities. In this case, they are united by a responsible attitude to environment, in particular, a willingness to travel to less visited and known places in order to contribute to solving the problem of overtourism in popular destinations and to provide socio-economic assistance to local communities [9].

Although the terms "ecological tourism", "green tourism", "agritourism", "rural tourism" as variants of sustainable tourism each have its own specifics they all constitute an alternative to the mainstream tourism by nature of management that does not lead to resource degradation, since part of the benefits from its development are used to restore and protect the latter [1]. Among them, the rural tourism is the most often used form of alternative tourism and is the prime subject of this study.

3. METHODOLOGY

Modeling of sustainable tourism processes allows for a scientific basis for empirical tools for effective management of tourism processes in a much more cost-effective way, compared to a natural experiment. Empirical models can be actively used in the process of marketing planning to improve the quality of management of the international and regional tourism economy [3].

The best formal representation of the aggregated scheme of interactions and mutual influences is the model of diffusion of innovations in the form of logistic function and its sigmoidal (S-shaped) curve, which models the dynamics of the growth of the probability of a certain event, as the control parameters (risk factors) change. The simplest logistic function can be written as:

$$P(t) = 1/(1+e-t) \quad (1)$$

where P is a variable depending on time t.

The initial stage of growth of the logistic curve approximately corresponds to the exponent, and as it becomes saturated, the growth slows down, goes through a linear phase and, finally, in the mature period, it practically stops [2].

In its turn, the logistic function is a solution to a simple nonlinear differential equation of the first order:

$$dP/dt = P(1 – P) \quad (2)$$

where P is a variable depending on time t and with the initial condition P(0)=P0.

The logistic differential equation is a continuous version of the logistic mapping, which after integration takes one of the two most common forms of recording the logistic dependence. The first is the model of an unlimited population, also called Malthusian. The second is the Malthusian model generalized for the case of a limited population, where its size grows over time to equilibrium value P = K:

$$dP/dt = rP(1 – P/K) \quad (3)$$

where P is the population size, t is time, r is the growth rate, called the Malthusian parameter, K is the carrying capacity of the environment or the equilibrium population density [15].

A complex of economic growth, social stability, environmental well-being, cultural development, management excellence and other positive factors creates the positive impact on the growth of tourist flows. The negative impact is created through factors of decreasing tourist flows, such as negative social changes at the destination (such as anti-tourism protests), environmental degradation, bad publicity and some types of cultural transformation and other similar negative factors. In addition, there is a two-way interaction resulting from interpersonal contacts between actual and potential tourists, related to the attractiveness of the location.

In the case of modeling tourist flows and obtaining an aggregated representation of the sustainable tourism development, it is possible to consider two differential equations as a simplification of natural processes. The first equation models the process of involving potential consumers in tourism per unit of time, and the second models the outflow of actual and potential tourists due to negative experience of visiting the location. These processes depend on many interacting factors, both positive and negative, which are expressed in growth coefficients. The coefficient of attractiveness is positive and reflects the intensity and quality of activity to draw tourists, and the coefficient of unattractiveness is the likelihood of “overlooking” a tourism destination due to various negative factors, conditions and shortcomings. Therefore:

$$Y(t) = rY(1 – Y/K) \quad (4)$$

where Y represents the time-varying number of travelers, r is coefficient of attractiveness, K is are the maximum contingent of tourists a tourist destination can accept,

$$X(t) = sX(1 – X/G) \quad (5)$$

while X reflects the time-varying number of those who wanted but refused to travel for various other than economic reasons, G is potential maximum of “refusers”. In this case, the number of those who refused also includes those who changed their travel direction and the potential of those who postponed or refused to travel.

Together, these two processes provide the final assessment of actual consumers of tourism products:

$$Z(t) = Y(t) – X(t) \quad (6)$$

where Z is the resulting number of travelers.

To apply this model, it is necessary to calculate the potential capacity for sustainable tourism. According to The World Tourism Organization tourism carrying capacity (TCC) as "the maximum number of people who can visit a tourist destination at the same time without causing destruction of the physical, economic, socio-cultural environment and an unacceptable decrease in the quality of visitor satisfaction."

There are different types of tourism carrying capacity. The first, physical carrying capacity is the maximum number of visits, determined by the ratio between the available space and the time required to visit the site. The second, economic carrying capacity is the maximum number of visits that a destination can accommodate tourism

functions at which the increased income generated by tourism development is offset by tourism-induced inflation. The ecological carrying capacity of tourism depends on the ecological and physical parameters, the carrying capacity of resources, ecosystems and infrastructure [7]. The social carrying capacity of tourism is expressed in the decrease of local tolerance for tourism to the level of protests. All these indicators are imperfect [16], due to the huge variety of tourist sites that cannot be covered by a single methodology, but a useful approach to visitor management in vulnerable areas [5].

The described model can be applied to simulate and analyze situations related to overtourism in various destinations. For example, by imagining hypothetical scenario: the destination currently visited by 1 million tourists becomes popular with a projected annual growth rate of 10% (y1), a maximum capacity of 10 million, and the number of cancellations grows by 5% per year (x1) could reach 5 million, it would take 150 years to reach a steady state (z1) (Fig. 1).

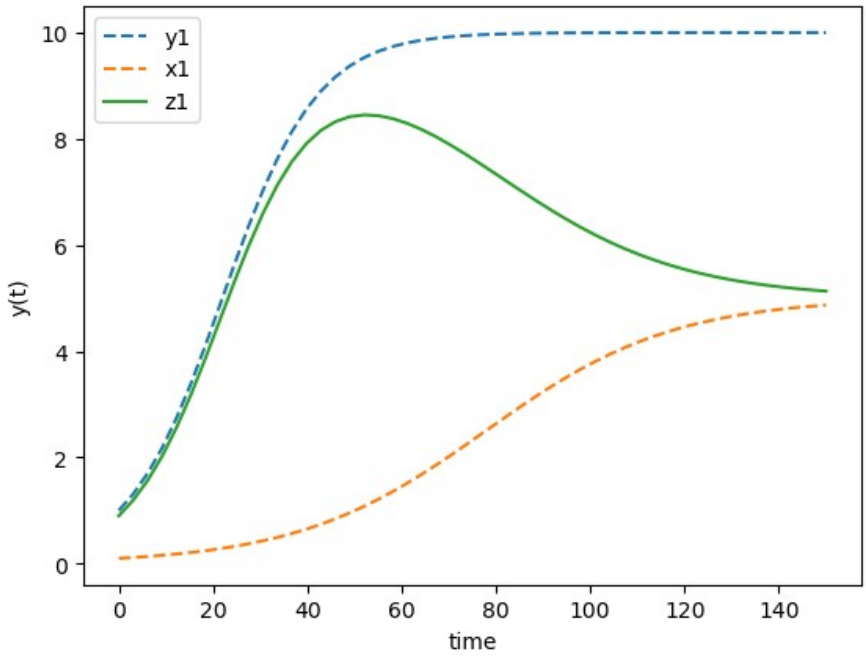

Fig.1. A hypothesized scenario of a number of tourist to a destination

4. MODEL EVALUATION AND DISCUSSION

The empirical application of the above-described model is conducted on the data collected from the following sources: UN Tourism (UNWTO), The European Travel Commission (ETC), Polska Organizacja Turystyczna (POT). In this case, the data up to 2019 is examined, as the last pre-pandemic period. According to the first source, the number of inbound tourism from 2010 to 2019 increased from 970 million to 1.465 billion, which implies an average cumulative annual growth of 5%. Putting this information into the first part of the model leads to the conclusion that such development of tourist flows is not sustainable because it corresponds to the maximum capacity of 5 billion or 60% of the world's population (Fig.2).

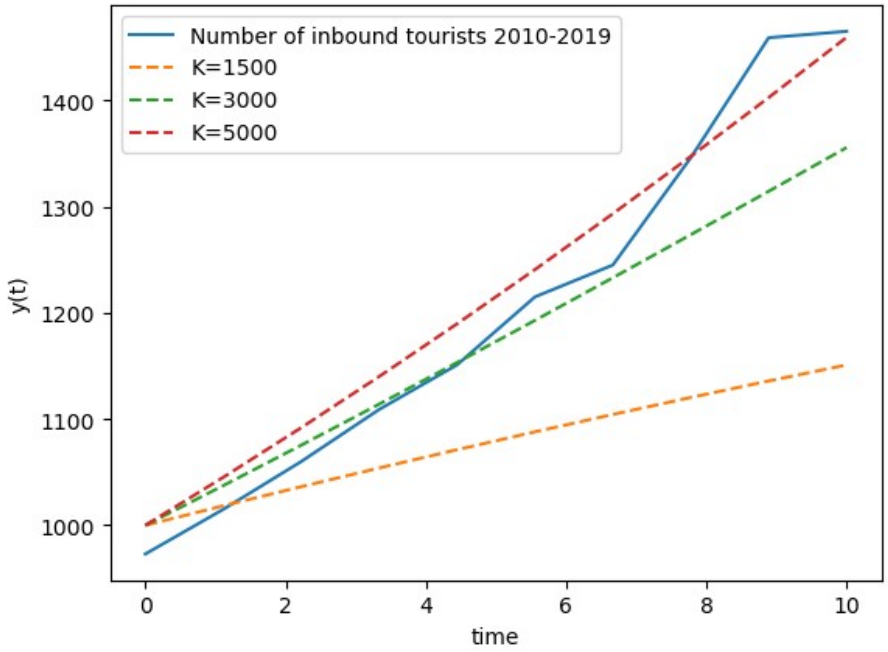

Fig.2. A simulation of development of the number of international tourists

The problem of overtourism manifests itself primarily in European countries, which are the main tourist destinations, and it becomes especially noticeable in capitals. At the same time, according to such indicators as the occupancy rate, number of available hotel rooms, the ratio of foreign to local tourists, number of tourists per capita in capitals significantly exceeds similar indicators on a national scale (Table 1). This indicates uneven tourist flows not only between countries, but also within them, which creates a problem of overtourism, but also shows the potential of non-capital regions.

Table 1. Key features of selected tourist destinations

Country	Occupancy rate, %	Beds available per 1000 population	Ratio Domestic vs International tourists	Number of tourists per 1 resident
Italy	55	37	1.0	1.1
Rome	75	51	0.5	6.8
Germany	62	20	0.9	4.7
Berlin	78	33	2.8	3.8
Poland	51	10	4.0	4.4
Warsaw	73	16	3.3	5.7

Source: UNWTO, ETC, PTO

To simulate tourist flows using the above-described model regions known for their tourist centers in rural areas, as an alternative to the above-mentioned metropolitan tourist destinations, were selected in the respective countries. Regions such as Italian Tuscany, German Saxony, and Polish Mazowsze can serve as examples of alternative tourism destinations. To estimate the first equation of the model, the average cumulative growth rates of the number of tourists in these destinations from 2010 to 2019 are used, and the throughput is taken at the level of 2019. The second equation of the model is estimated hypothetically on a proportional basis (Table 2).

Table 2. Key development indicators of selected tourist destinations

City/Region	Y0	r	K	X0	m	G
Rome	19.9	9%	35.0	1.4	7%	3.0
Tuscany	4.7	8%	6.1	0.4	6%	0.7
Berlin	14.9	7%	25.0	1.2	5%	2.2
Saxony	3.1	8%	4.0	0.2	6%	0.5
Warsaw	13.5	6%	20.0	1.2	4%	2.0
Mazowsze	4.4	5%	5.7	0.4	3%	0.7

Source: UNWTO, ETC, PTO

The results of simulation of the development of tourist flows using the innovation diffusion model is graphically represented in the following figure (Fig. 3), where the more famous tourist centers are on the left and less renowned are on the right.

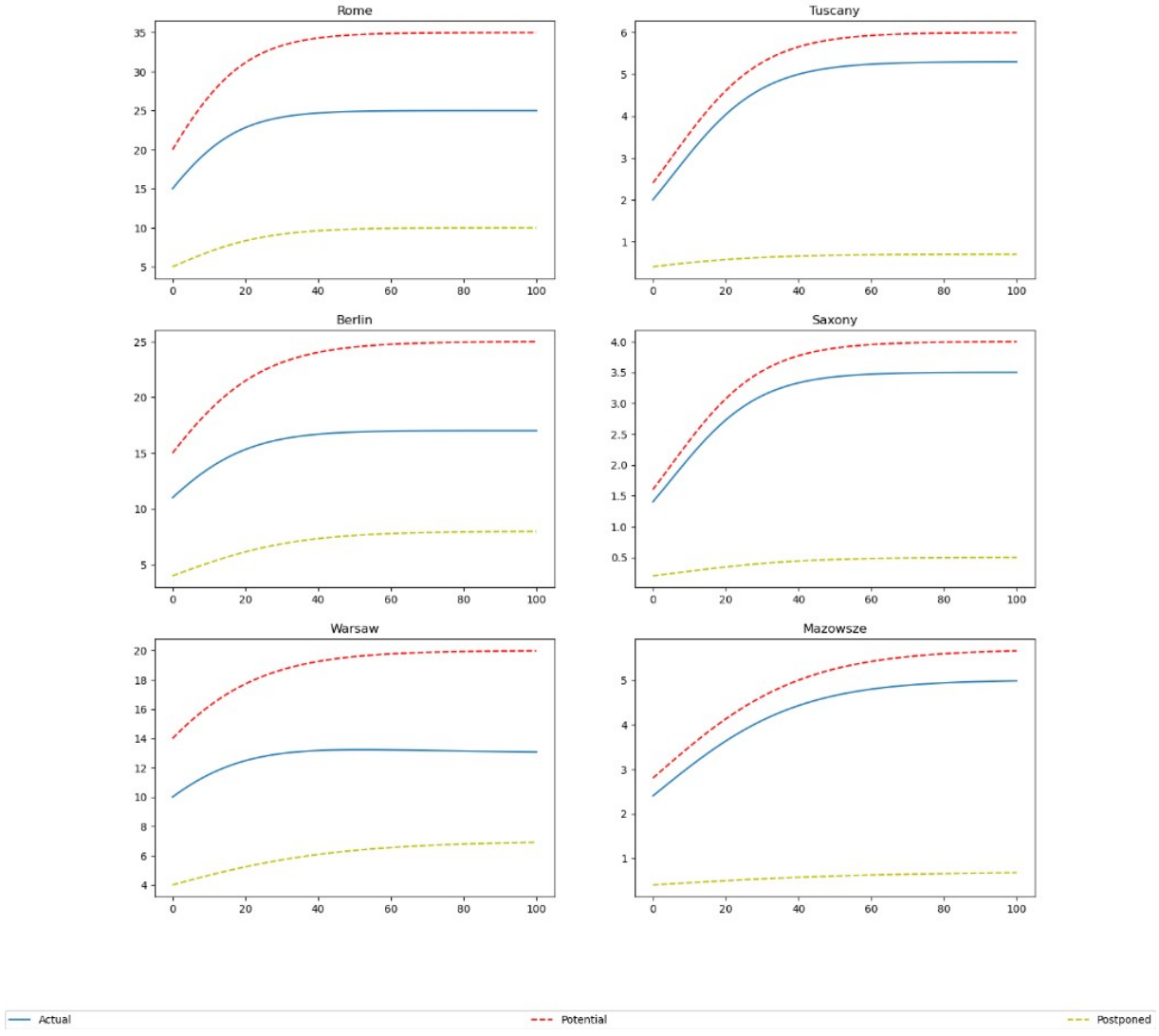

Fig.3. Simulation of the development of the number of tourists in selected locations

The analysis of graphs obtained using the innovation diffusion model to simulate tourist flows in different locations makes it possible to compare the dynamics of changes in travelers in the most popular and less popular tourist places and draw the following conclusions. In the developed tourist centers, the tourists' number reaches the carrying capacity faster than in places focused on rural tourism. In addition, rural tourism is less susceptible to the influence of negative reactions from travelers, which leads to a significantly smaller impact of the number of "refusers" on the final number of visitors, which explains that at initial stages it has higher dynamics than mainstream tourism. It is according to these indicators that rural tourism surpasses traditional tourism in terms of sustainable development.

References

[1] An, Wookhyun, and Silverio Alarcón. "How can rural tourism be sustainable? A systematic review." Sustainability 12, no. 18 (2020): 7758.
[2] Dorini, Fábio Antonio, M. S. Cecconello, and Leyza Baldo Dorini. "On the logistic equation subject to uncertainties in the environmental carrying capacity and initial population density." Communications in Nonlinear Science and Numerical Simulation 33 (2016): 160-173.
[3] Ganglmair-Wooliscroft, Alexandra, and Ben Wooliscroft. "Diffusion of innovation: The case of ethical tourism behavior." Journal of Business Research 69, no. 8 (2016): 2711-2720.
[4] León-Gómez, Ana, Daniel Ruiz-Palomo, Manuel A. Fernández-Gámez, and Mercedes Raquel García-Revilla. "Sustainable tourism development and economic growth: Bibliometric review and analysis." Sustainability 13, no. 4 (2021): 2270.
[5] Long, Cheng, Song Lu, Jie Chang, Jiaheng Zhu, and Luqiao Chen. "Tourism environmental carrying capacity review, hotspot, issue, and prospect." International Journal of Environmental Research and Public Health 19, no. 24 (2022): 16663.
[6] Lu, Jiaying, and Sanjay K. Nepal. "Sustainable tourism research: An analysis of papers published in the Journal of Sustainable Tourism." Journal of sustainable Tourism 17, no. 1 (2009): 5-16.
[7] Mexa, Alexandra, and Harry Coccossis. "Tourism carrying capacity: a theoretical overview." The challenge of tourism carrying capacity assessment (2017): 53-70.
[8] Rasoolimanesh, S. Mostafa, Sundari Ramakrishna, C. Michael Hall, Kourosh Esfandiar, and Siamak Seyfi. "A systematic scoping review of sustainable tourism indicators in relation to the sustainable development goals." Journal of Sustainable Tourism 31, no. 7 (2023): 1497-1517.
[9] Rosalina, Putu Devi, Karine Dupre, and Ying Wang. "Rural tourism: A systematic literature review on definitions and challenges." Journal of Hospitality and Tourism Management 47 (2021): 134-149.
[10] Tahiri, Alberta, and Idriz Kovaci. "The theory of sustainable tourism development." Academic Journal of Business 3, no. 1 (2017): 343-349.
[11] Torres-Delgado, Anna, and Jarkko Saarinen. "Using indicators to assess sustainable tourism development: A review." New research paradigms in tourism geography (2017): 31-47.
[12] UN World Tourism Organization, International Tourism Highlights, 2019 Edition, All Regions, UN Tourism, Madrid, https://www.e-nwto.org/doi/pdf/10.18111/9789284421152
[13] UN World Tourism Organization, World Tourism Barometer, volume 22, issue 2, May 2024, UN Tourism, Madrid, https://www.e-unwto.org/toc/wtobarometereng/22/2
[14] Wehrli, R., Egli, H., Lutzenberger, M., Pfister, D., Schwarz, J., & Stettler, J. (2011). „Is there De-mand for Sustainable Tourism?–Study for the World Tourism Forum Lucerne 2011—. ITW Working Paper Series Tourism 001/2011, Lucerne University of Applied Sciences and Arts.
[15] Yoshioka, Hidekazu. "A simplified stochastic optimization model for logistic dynamics with control-dependent carrying capacity." Journal of biological dynamics 13, no. 1 (2019): 148-176.
[16] Zekan, Bozana, Christian Weismayer, Ulrich Gunter, Bernd Schuh, and Sabine Sedlacek. "Regional sustainability and tourism carrying capacities." Journal of Cleaner Production 339 (2022): 130624.

AI in Hospitality: Navigating Innovation, Ethics, and Cultural Representation in Albania

Matilda NACO [a*], Denisa MILLO [b], Kozeta SEVRANI [c]

[a] *University of Tirana, Faculty of Economy, Department of Marketing & Tourism, Rruga Arben Brozit, Tirana, Albania, matilda.naco@unitir.edu.al*
[b] *University of Tirana, Faculty of Economy, Department of Statistics and Applied Informatics, Rruga Arben Brozit, Tirana, Albania, denisa.millo@unitir.edu.al*
[c] *University of Tirana, Faculty of Economy, Department of Statistics and Applied Informatics, Rruga Arben Brozit, Tirana, Albania, kozeta.sevrani@unitir.edu.al*

Abstract

The integration of AI and smart technologies is transforming the hospitality industry, enhancing efficiency and guest experiences. However, as AI automation expands, ensuring human-centered and culturally inclusive implementation is essential. This study examines AI adoption in Albania's hospitality sector, focusing on its impact on hotel management while preserving human interaction and addressing ethical challenges. Using a mixed-method approach, including qualitative interviews with hotel professionals and quantitative AI adoption analysis, findings highlight AI's benefits in service personalization, booking management, and revenue optimization. However, concerns persist regarding depersonalized services, algorithmic biases, and workforce displacement. Language barriers and cultural representation remain challenges for AI-powered customer interactions.
AI also presents opportunities for inclusivity through translation technologies, interactive guest engagement, and digital storytelling. The study underscores the need for ethical AI frameworks to mitigate biases and ensure fair representation. Workforce adaptability and upskilling are crucial for smooth AI integration without harming employment.
Despite AI's potential, adoption in Albania faces challenges such as high implementation costs, limited technical expertise, and resistance to change. The paper concludes with recommendations for balancing automation with human connection, ensuring AI policies align with Albania's cultural and economic landscape.

Keywords: Artificial Intelligence in Hospitality, Smart Hotel Technologies, Human-Centered AI, Cultural Inclusivity in AI, AI Adoption in Developing Markets

1. INTRODUCTION

The hospitality industry is undergoing a profound transformation, driven by the integration of Artificial Intelligence (AI) and smart technologies. These digital innovations offer vast potential for operational efficiency, personalized guest services, and competitive advantage in global markets. From chatbots and dynamic pricing systems to AI-powered concierge services and smart room technologies, the hospitality sector is rapidly embracing

* Corresponding author.

automation [1]. However, the success of AI integration is not solely dependent on its technological capabilities, it also hinges on its ethical implications, cultural sensitivity, and human-centered design [2].

In Albania, a country with a rich cultural heritage and growing tourism economy, the adoption of AI must be carefully aligned with local values and socio-economic realities. The Albanian hospitality sector is characterized by small- and medium-sized businesses, many of which emphasize warm, personalized service and operate in informal economies with limited access to high-tech infrastructure. As such, the challenge is not only to introduce AI technologies but to do so in a way that respects and enhances the unique character of Albanian hospitality. This paper explores the integration of AI in Albania's hospitality industry through the lenses of innovation, ethics, and cultural representation. Drawing from a mixed-method study, it highlights both the opportunities, and the challenges associated with AI adoption, and it offers recommendations for a culturally grounded and ethically responsible approach to AI implementation.

2. LITERATURE REVIEW

The global hospitality sector has experienced substantial innovation through AI-driven technologies, including virtual assistants, smart rooms, and customer data analytics [3]. These tools enhance operational efficiency and elevate guest experience by enabling personalization at scale. In highly digitized markets such as the United States, Japan, and Western Europe, AI is now integral to hospitality operations [4]. However, studies show that the implementation of AI in hospitality also introduces complex ethical issues. [5] and [6] caution against depersonalization, data misuse, and algorithmic bias particularly when AI systems lack cultural contextualization. These challenges are especially relevant for countries like Albania, where hospitality is rooted in personal service and local traditions. Tourist perceptions also shape the success of AI in hospitality. [7] found that while tourists appreciate the convenience of AI, they also value the human touch and express concerns about surveillance and cultural misrepresentation. This underscores the importance of designing transparent, culturally sensitive AI tools.

In developing countries, infrastructure limitations and workforce skill gaps hinder AI adoption. [8] emphasize the role of smart technologies in streamlining hotel operations yet highlight the need for ethical implementation that respects labor markets and traditional service models. Similarly, [9] argue that ethical AI must be grounded in local epistemologies and belief systems, advocating for “actionable ethics” rather than abstract principles. The concept of value-sensitive design, proposed by [10], offers a relevant framework for Albania. This approach integrates human values and stakeholder input into the design of AI systems, ensuring alignment with social justice and cultural preservation. [11] further advocate for embedding local narratives and aesthetics into tourism technologies, which is particularly important for destinations like Albania that rely heavily on cultural storytelling and authenticity. Literature also points to the gap between global AI tools and the needs of non-Western users. [12] highlights how AI algorithms often prioritize upscale accommodations and overlook local businesses, reinforcing economic disparities. [13] identifies language barriers as a persistent issue, with few AI tools supporting Albanian or its dialects limiting accessibility for operators and guests.

The literature reveals both the promise and the pitfalls of AI in hospitality. It stresses the importance of ethical governance, local customization, and stakeholder inclusion elements that are crucial for effective and sustainable AI adoption in the Albanian context.

3. METHODOLOGY

This research employs a mixed-method approach that combines qualitative interviews with hotel professionals and quantitative analysis of AI adoption [14], [15]. This methodology is particularly well-suited to exploring the nuanced challenges of AI-powered customer interactions in Albania’s hospitality sector. It allows for a comprehensive and multidimensional understanding of complex issues such as depersonalized service, algorithmic bias, workforce displacement, language barriers, and cultural representation. A total of 120 hotel staff and managers were approached to complete structured questionnaires, with 108 valid responses collected from seven cities across Albania. Each of these 108 participants also took part in in-depth qualitative interviews, ensuring consistency between the two strands of data. This integration of quantitative and qualitative data enhances the depth and breadth of the research by aligning statistical trends with individual perspectives and reflections.

The quantitative component of the study involved a structured survey administered across a stratified sample of accommodation establishments. The sample was stratified by hotel size, level of technology investment in the past five years, workforce indicators such as AI-related training, and customer feedback on personalization and cultural sensitivity [16]. The survey data was cleaned, integrated, and transformed into a common dataset. Dummy variables were used to convert qualitative elements into numerical values suitable for statistical analysis. Descriptive statistics

were produced to highlight the distribution of AI adoption across different types of services and technologies. Histograms illustrated the frequency of adoption across six AI technologies, and bar charts visualized perceived benefits, innovation, and centralized management. A Spearman correlation test was conducted to investigate relationships between AI adoption levels and reported implementation challenges, though no strong correlations were found.

The qualitative component relied on purposive sampling and included interviews with 108 hotel professionals across various roles and accommodation types. These participants represented a diversity of hotel categories, from luxury international chains to family-owned guesthouses and spanned key tourism destinations such as Tirana, Saranda, Berat, Gjirokastër, Shkoder, Pogradec as well as rural and heritage sites. Interviewees included front-desk staff, IT specialists, marketing professionals, and senior managers, many of whom were multilingual and had direct experience with AI implementation. The interviews explored perceptions of ethical implications, cultural misalignment, service transformation, and workforce changes.

By integrating methods, this research moves beyond binary success/failure metrics and instead reveals how AI impacts people, processes, and perceptions within Albania's evolving tourism sector. Triangulation across data sources increases the validity of the findings by corroborating themes and patterns. For example, workforce anxieties identified in interviews were compared with HR trends in staff training and retraining. Similarly, qualitative concerns about algorithmic bias were examined in relation to discrepancies in guest feedback across cultural groups.

Quantitative data provided insight into macro-level patterns of AI integration, while qualitative findings revealed the micro-level realities of ethical tension, staff adaptation, and guest experience. This dual-level inquiry, combining thematic coding and statistical analysis, ensures a holistic understanding of AI's transformative role in the Albanian hospitality industry.

4. FINDINGS AND DISCUSSIONS

This research used qualitative coding and thematic analysis to extract patterns, sentiments, and insights from questionnaire responses and in-depth interviews. Through open and axial coding [15], [17], both free-text and multiple-choice responses were synthesized into high-level themes reflecting the attitudes, experiences, and aspirations of hospitality professionals across Albania.

Participants acknowledged that AI technologies are gradually becoming embedded in core service areas. The most commonly offered services among the surveyed hotels were restaurants and bars, parking, laundry, room service, and event spaces functions with high human involvement but also high potential for automation and smart enhancement. Many establishments already use AI in booking and revenue management through dynamic pricing and forecasting tools, and in guest personalization via chatbots and CRM systems capable of tailored, multilingual communication. Operational efficiency has also improved, as several respondents noted benefits from AI-assisted maintenance and housekeeping systems. Quantitative analysis confirmed these impressions, with operational efficiency cited by 69% of participants as the top benefit, followed closely by enhanced customer experience (68%) and cost reduction (60%). Respondents identified a range of challenges that limit the full potential of AI adoption. High implementation costs remain the most significant barrier, mentioned by 65% of participants. Other notable concerns include lack of internal technical expertise (45%), insufficient training (40%), tight budgets (38%), and complexity of integration with existing systems (36%). In response to these challenges, many stressed the need for modular, cost-effective, and well-supported technologies.

Sustainability also emerged as a relevant theme. Many participants recognized the potential of smart technologies to contribute to environmental goals, especially in areas like water and energy conservation. Among the 108 respondents, 78% prioritized technologies that reduce water and energy use, 72% emphasized consumption monitoring, and 60% favored smart lighting and temperature systems as efficient and sustainable options.

In addition to these operational and environmental dimensions, ethical and human-centered concerns were repeatedly raised. While AI offers clear advantages in terms of efficiency, many respondents warned of the risks of depersonalization. Participants described fully automated services as lacking the spontaneity and emotional warmth traditionally associated with Albanian hospitality. Algorithmic bias was another commonly cited concern, particularly in relation to AI systems that do not accurately interpret regional dialects or culturally specific behavior. Hotel staff also voiced fears of workforce displacement, as automation increasingly takes over functions previously handled by human employees.

Cultural and linguistic representation stood out as a key area for improvement. Around 82% of participants stated that current AI platforms do not adequately support the Albanian language or reflect cultural nuances. Respondents noted that chatbots trained on international datasets often misinterpret idioms, customs, or etiquette,

resulting in awkward or inappropriate guest interactions. Nonetheless, many saw promises in using AI to support cultural translation, enhance storytelling through AR/VR applications, and promote accessibility for guests with disabilities.

When asked about future aspirations, respondents expressed strong interest in more advanced AI tools (64%), autonomous maintenance systems (64%), and improved data protection (50%). Cybersecurity and privacy were recognized as prerequisites for long-term AI adoption, with suggestions including modern security technologies (75%), system upgrades (50%), and targeted training in digital safety (45%).

Beyond the technological dimension, the study revealed the importance of cultural responsiveness, ethical design, and participatory implementation strategies. Many professionals stressed the necessity of aligning innovation with core values of trust, guest safety, and cultural integrity. They advocated national guidelines to address issues of data transparency, fairness, and localization, ensuring that AI is not only effective but also respectful and contextually appropriate.

The following table summarizes the findings through a comparative analysis of three major challenge areas ethical considerations, cultural sensitivity, and stakeholder collaboration drawing on concrete survey insights and real-world examples, along with corresponding recommendations.

Challenge	*Survey Insight*	*Illustrative Example*	*Recommended Response*
Innovation–Ethics Balance	*63% of hotel managers express concern over data misuse related to AI systems. Only 27% have adopted AI for predictive pricing due to privacy and compliance concerns.*	*A Tirana-based hotel installs an AI booking engine but disables guest profiling due to GDPR-related uncertainty and fear of reputational harm.*	*Develop a national "AI Ethics Framework for Tourism"; create GDPR-aligned training resources and case studies tailored for hoteliers.*
Cultural Sensitivity	*48% of respondents reported that AI tools fail to reflect authentic Albanian experiences. 82% avoid Albanian-language chatbots due to low accuracy and poor localization.*	*An AI concierge in Gjirokastër suggests mass-produced souvenirs over handmade local crafts, reflecting data bias toward generic commercial offerings.*	*Fund Albanian language corpora and regional cultural datasets; require AI systems to prioritize local vendor representation in recommendations.*
Stakeholder Collaboration	*70% of small hotel and guesthouse operators report being excluded from AI-related decision-making. Only 22% of AI implementations involved local consultation.*	*A state-backed tourism app disproportionately promotes urban hotels, marginalizing rural guesthouses in Theth and Berat.*	*Establish co-design workshops and inclusive AI review councils that include rural stakeholders, cultural institutions, and tourism boards.*

Fig. 1. (a) AI Adaptation Challenges in Albanian Hospitality: Survey-Backed Analysis

The cumulative evidence indicates that AI implementation in the Albanian hospitality sector cannot rely solely on technological advancement. Instead, it must be framed within a broader, multidimensional strategy that incorporates ethical standards, cultural sensitivity, and inclusive collaboration. AI should not be seen as a replacement for human labor or traditional hospitality values, but rather as a supportive tool to enrich guest experiences and empower the workforce. To ensure successful integration, the development of national guidelines tailored specifically to tourism and hospitality is essential. These should include ethical protocols for data privacy, algorithmic transparency, and cultural alignment. Establishing such a framework would foster trust among stakeholders while offering a reliable foundation for both operators and technology providers.

Equally important is the upskilling of the workforce. As AI systems become more prevalent, staff must acquire the digital competencies needed to effectively manage and complement these technologies. This calls for coordinated training programs in partnership with vocational schools, universities, and industry actors. Locally adapted AI tools must be prioritized, particularly those that support the Albanian language and local cultural contexts. Current international systems often fall short in delivering seamless communication or culturally relevant interactions. Promoting the development of home-grown or regionally adapted AI solutions can address this gap and

better serve diverse hospitality environments. Cross-sector collaboration is also crucial. Many smaller and family-run businesses have been left out of digital transformation initiatives. Inclusive engagement among hotel operators, tourism boards, technology developers, and cultural institutions will ensure that resulting innovations reflect the collective needs of the sector. A human-centered approach must guide every stage of AI development and implementation. AI should be intuitive, respectful of users, and empathetic in design, serving as a tool that enhances not replaces the uniquely human touch at the heart of Albanian hospitality.

5. CONCLUSION

To ensure ethical, inclusive, and effective AI integration in Albania's hospitality sector, several strategic actions emerge from the findings. First, the development of national AI guidelines tailored for the hospitality industry is vital. These should incorporate standards for data ethics, cultural sensitivity, and linguistic accessibility to foster public trust and regulatory clarity [18]. Second, there is a clear need to invest in workforce training and AI literacy. Equipping current employees with the knowledge and skills to interact with and manage AI tools will help bridge the gap between technological capability and practical application [19].

Promoting localized AI development is another key priority. Encouraging domestic startups and academic institutions to design culturally attuned AI systems would ensure better alignment with the needs of Albanian hospitality. These efforts must be supported through public-private partnerships that can offset implementation costs and facilitate pilot projects in both urban and rural contexts. Underpinning all these recommendations is the necessity of human-centered design systems must not only be functional but empathetic, intuitive, and responsive to user needs [20]. AI offers transformative potential for Albania's tourism and hospitality industries. It can enhance operational efficiency, deliver personalized guest experiences, and position the country competitively on a global scale. Yet the success of such transformation hinges on more than technological advancement. It depends on ethical foresight, inclusive policy, cultural respect, and meaningful collaboration among all stakeholders.

As Albania navigates this technological frontier, a carefully designed AI strategy grounded in its local values and social dynamics can make it a model for culturally sustainable innovation. With the right investments in infrastructure, workforce development, and cross-sector partnerships, Albania can modernize its hospitality industry while preserving the authenticity that defines its tourism identity.

The integration of AI into Albania's hospitality sector presents a unique opportunity to modernize services and elevate the country's global tourism standing. Yet, this transition must be approached with care. The essence of Albanian hospitality its warmth, cultural richness, and human connection must remain at the heart of this evolution.

If AI technologies are to succeed in this sector, they must be implemented in a way that reinforces cultural identity, protects employment, and promotes social inclusion. The path forward involves more than simply acquiring tools; it requires building capacities, strengthening collaboration, and creating a shared vision for the future of hospitality in Albania.

By embracing a locally driven, ethically grounded, and human-focused approach, Albania can not only overcome the challenges of digital transformation but can also become a model for responsible innovation in tourism and beyond.

Acknowledgements: This research was supported by the National Agency for Scientific Research and Innovation (AKKSHI).

References

[1] Khlusevich, A., Inversini, A., Schegg, R. (2024). Artificial Intelligence and Hospitality: A Challenging Relationship. In: Berezina, K., Nixon, L., Tuomi, A. (eds) Information and Communication Technologies in Tourism 2024. ENTER 2024. Springer Proceedings in Business and Economics. Springer, Cham. https://doi.org/10.1007/978-3-031-58839-6_27.

[2] Milwood, Pauline & Hartman-Caverly, Sarah & Roehl, Wesley. (2023). A Scoping Study of Ethics in Artificial Intelligence Research in Tourism and Hospitality. 10.1007/978-3-031-25752-0_26.

[3] Ivanov, S., Webster, C., & Berezina, K. (2017). Adoption of Robots and Service Automation by Tourism and Hospitality Companies. Revista Turismo & Desenvolvimento, 27/28, 1501–1517.

[4] Ivanov, S., & Umbrello, S. (2021). The Ethics of Artificial Intelligence and Robotization in Tourism and Hospitality - A Conceptual Framework and Research Agenda. Journal of Smart Tourism, 1(4), 9-18. https://doi.org/10.52255/smarttourism.2021.1.4.3 (Original work published 2021).

[5] Omar, Shida & Aliah, Nur & Mooi Kwong, Anthony. (2024). The Ethical Implications of AI-Powered Personalization in Hospitality and Tourism. 10.4018/979-8-3693-6755-1.ch006.

[6] Kim, Hyunsu & So, Kevin Kam Fung & Shin, Seunghun & Li, Jing. (2024). Artificial Intelligence in Hospitality and Tourism: Insights From Industry Practices, Research Literature, and Expert Opinions. Journal of Hospitality & Tourism Research. 49. 10.1177/10963480241229235.

[7] Binesh, Fatemeh & Baloglu, Seyhmus. (2023). Are we ready for hotel robots after the pandemic? A profile analysis. Computers in Human Behavior. 147. 107854. 10.1016/j.chb.2023.107854.

[8] Tom Dieck, M. Claudia & Jung, Timothy. (2018). A Theoretical Model of Mobile Augmented Reality Acceptance in Urban Heritage Tourism. Current Issues in Tourism. 21. 154-174. 10.1080/13683500.2015.1070801.

[9] Chi, Oscar Hengxuan & Denton, Greg & Gursoy, Dogan. (2020). Artificially intelligent device use in service delivery: a systematic review, synthesis, and research agenda. Journal of Hospitality Marketing & Management. 29. 1-30. 10.1080/19368623.2020.1721394.

[10] El Hajal, Georges. (2024). The future of AI: hospitality and tourism perspectives. Research in Hospitality Management. 14. 1-3. 10.1080/22243534.2024.2395708.

[11] El Hajal, Georges & Yeoman, Ian. (2024). AI and the future of talent management in tourism and hospitality. Current Issues In Tourism. 1-18. 10.1080/13683500.2024.2439395.

[12] Rancati, Elisa & Abashidze, Irakli & D'Agata, Alessandro. (2025). AI-based Machine Translations in Tourism and Hospitality: An Exploratory Study of Travelers' Perceptions and Future Challenges. European Scientific Journal, ESJ. 21. 1. 10.19044/esj.2025.v21n7p1.

[13] Inhyouk Koo, Umer Zaman, Hojung Ha, Shahid Nawaz, Assessing the interplay of trust dynamics, personalization, ethical AI practices, and tourist behavior in the adoption of AI-driven smart tourism technologies, Journal of Open Innovation: Technology, Market, and Complexity, Volume 11, Issue 1, 2025, 100455, ISSN 2199-8531, https://doi.org/10.1016/j.joitmc.2024.100455.

[14] Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving Integration in Mixed Methods Designs: Principles and Practices. Health Services Research, 48(6pt2), 2134–2156. https://doi.org/10.1111/1475-6773.12117

[15] Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using Mixed-Methods Sequential Explanatory Design. Field Methods, 18(1), 3–20. https://doi.org/10.1177/1525822X05282260.

[16] Umbrello, S. (2019). Beneficial Artificial Intelligence Coordination by Means of a Value Sensitive Design Approach. Big Data and Cognitive Computing, 3(1), 5. https://doi.org/10.3390/bdcc3010005

[17] Manzoor, Sheikh & Ullah, Rezwan & Khattak, Afraseyab & Ullah, Munsif. (2024). Exploring tourist perceptions of artificial intelligence devices in the hotel industry: impact of industry 4.0. Journal of Travel & Tourism Marketing. 41. 272-291. 10.1080/10548408.2024.2310169.

[18] Gretzel, Ulrike & Sigala, Marianna & Xiang, Zheng & Koo, Chulmo. (2015). Smart tourism: foundations and developments. Electronic Markets. 25. 10.1007/s12525-015-0196-8.

[19] Morch, A. (2025). AI & Humans: A Perfect Harmony for Elevating the Hotel Guest Experience. Hospitality Net. https://www.hospitalitynet.org/opinion/4125634.html

CIAE
Czech Institute of Academic Education z.s.

Title:

"Proceedings of IAC 2025 in Prague", ISBN 978-80-88203-42-1

April 2025 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 978-80-88203-42-1

Czech Institute of Academic Education z.s.