Guiding the Algorithm: Harnessing artificial intelligence to nurture SMEs management control systems

Laura Broccardo, Elisa Ballesio, Daniele Giordino, Elisa Giacosa[*]

Received: 22 August 2024
Accepted: 28 May 2025

Abstract

The diffusion of artificial intelligence (AI) amongst organizations is profoundly revolutionizing business processes, such management control (MC) practices. However, the existing literature reveals a significant gap in understanding the contribution of AI-empowered tools to management control systems (MCSs), particularly within small and medium enterprises (SMEs), where these systems are often rudimental, informal, subjective, and short-period-oriented. To address this gap, the authors employed a mixed-methods approach, featuring an open-coding and thematic analysis of AI-generated answers to MC questions related the analysis of liquidity and profitability ratios from 37 Italian SMEs, as well as a qualitative investigation, with in-depth interviews with 15 respondents, to highlight their perceptions. Our investigation is informed by the social construction of technology theory, to explore the relevant social groups, interpretative flexibility, technological frames and patterns of closure related to AI-empowered tools and their contribution to SMEs' MCSs. We connect our findings to two frames, which highlight, on the one hand the positive contribution of such tools to SMEs' MCSs through accurate and comprehensive responses; on the other hand, a negative frame highlights concerns related to inconsistency and verbosity of the answers, which may affect trust and usability. This manuscript aims to provide an overview of the possible contribution of AI-empowered tools to SMEs' MCSs. Both scholars and practitioners can benefit from the theoretical and practical implications provided, which complements the limited number of studies in the field. Future research avenues are discussed.

Keywords: Management Control, Artificial Intelligence, SMEs, Small and Medium Enterprises, Social Construction of Technology theory

---

[*] University of Torino, Department of Management "Valter Cantino". Corresponding author: laura.broccardo@unito.it.

Management Control (ISSN 2239-0391, ISSNe 2239-4397), 2025, 1 – Special Issue
Doi: 10.3280/MACO2025-001-S1002

1. Introduction

Digital technologies (DTs) have brought significant transformations in nearly every aspect of life, with Artificial Intelligence (AI) being one of the most significant advancements (Valentinetti and Rea, 2023). Nowadays, AI is employed in numerous fields, and for different purposes, including speech and image recognition, medical diagnostics, and the automation of routine tasks (Grassi and Lanfranchi, 2022; Oppioli *et al.*, 2023). These technological advancements are part of a broader process, which can be understood through three interrelated concepts: digitization, digitalization, and digital transformation (DT). Digitization entails converting from analog to digital form (O'Leary, 2023). Digitalization, on the other hand, involves the use of DTs to transform a business model, creating new chances for income and value creation (Valentinetti and Rea, 2025). On top of that, DT represents the most profound shift, as it entails leveraging DTs to create new business models and build new processes or restructure the existing ones (O'Leary, 2023).

When it comes to organizations, digitization, digitalization and DT are profoundly revolutionizing business strategies, models, and processes (Culasso *et al.,* 2024; Skare *et al*., 2023; Verhoef *et al*., 2021), mainly through the integration of DTs into various departments (Bankins *et al*., 2024), including firms' accounting, finance, and management control (MC) functions (Chowdhury, 2023; Sundström, 2024). Scholars have demonstrated that DTs such as robotics (Cooper *et al*., 2019), blockchain (Kostić and Sedej, 2022), cloud computing (Maelah *et al*., 2020), and big data (Vasarhelyi *et al*., 2015) are profoundly influencing management accounting (MA) and will continue to do so (Lombardi *et al*., 2021). Among all DTs, AI is thought to have the most impact on MC because, thanks to its ability to learn, connect, and adapt (Sundström, 2024), it makes it possible to find patterns in vast volumes of accounting data that may help businesses make decisions (Kureljusic and Karger, 2024).

Nonetheless, despite the usefulness of AI integration within management control systems (MCSs) and its continuous diffusion amongst businesses, there is limited research on AI and MA (Losbichler and Lehner, 2021), especially in small and medium enterprises (SMEs) (Merjane et al., 2024). While DTs and AI adoption is growing among larger firms, in SMEs this phenomenon is notably scant, despite its benefits and opportunities (Wei and Pardo, 2022). Furthermore, in SMEs, MCSs are informal, subjective, and short-period-oriented SME (Broccardo *et al*., 2017), forcing managers to make fixed decisions (Chowdhury, 2023). Secondly, in SMEs, MCSs tend to be rudimental, as they prioritize analyses involving ratios and financial statements figures, driven by external factors such as taxation or bank requirements (Garengo and Biazzo, 2012). With AI assistance in MC, these issues could be addressed by including a more significant amount of data in the forecast, thus giving the possibility to make long-term and dynamic decisions and eliminating subjectivity (Losbichler and Lehner, 2021).

However, despite this field of research having gained significant attention from scholars in the last period, the empirical link between AI implementation and MCS in SMEs remains unclear (Merjane *et al*., 2024). This lack of research underlines the need for additional empirical evidence on how AI-empowered tools (AIETs) could support and nurture SMEs MCSs (Losbichler and Lehner, 2021). Research on SMEs is particularly relevant, as they constitute a significant portion of the global economy (Cardoni *et al*., 2023; Giordino *et al.*, 2024). Moreover, there is a growing interest among policymakers, like the European commission, in enhancing SMEs' digitalization (Broccardo *et al*., 2024; Giordino *et al.*, 2024) In fact, only 22% of European SMEs employ AI (Omrani *et al*., 2024).

Furthermore, despite the growing interest in AIETs for MA (Culasso *et al.*, 2024; Secinaro *et al*., 2024), computer science disciplines dominate this field's research (Kureljusic and Karger, 2024). Indeed, there are significant research gaps on how AIETs may assess business performance, thus impacting MA (Rikhardsson and Yigitbasioglu, 2018). Therefore, the present body of research seeks to provide empirical evidence on the contribution of AI to MA, thus broadening our understanding of how such technology could be leveraged to enhance MCSs, with a particular focus on SMEs.

To this end, the present manuscript seeks to investigate how AIETs, namely ChatGPT assistants, could nurture SMEs' MC practices and systems, through the assessment of liquidity (fixed assets coverage, interests coverage, quick ratio, debt equity ratio) and profitability (ROI, ROE, ROS, turnover) ratios.

The research question formulated is as follows:

- RQ1: How could AIETs, specifically AI assistants, contribute to SME MCSs?

In order to answer the aforementioned question, eight ratios were calculated for 37 Italian SMEs and ChatGPT assistants were asked to comment on them. Thanks to this process, we were able to collect a total of 296 answers. Secondly, to assess the overall quality of the answer, we employed a mixed method approach, with an open-coding procedure and semi-structured interviews with 15 respondents.

The context and sample used in this study seek to provide empirical evidence to address both theoretical and contextual gaps. Indeed, SMEs represent the majority of the Italian business context (Perrini, 2006), and are under increasing pressure to enhance their DT (Broccardo *et al*., 2024).

The present manuscript draws insights from the Social Construction of Technology (SCOT) theory, to explore the relevant social groups, interpretive flexibility, technological frame, and patterns of closure associated with AIETs and their contribution to SMEs' MCSs.

This manuscript's empirical findings report the rise of two frames regarding AIETs in SME MCSs. On the one hand, a positive frame emphasizes how such tools could nurture SMEs' MC practices (Losbichler and Lehner, 2021; Nóbrega *et al*., 2023) through accurate and comprehensive responses (Aryal *et al*., 2024). On the other hand, a negative frame highlights concerns related to inconsistency and verbosity of the answers (Kabir *et al*., 2024; Jang and Lukasiewicz, 2023; Saito *et al*., 2023), which may affect trust and usability. These conflicting frames suggest that full closure and stabilization are yet to be achieved (Pinch and Bijker, 1987). However, recent technological improvements, such as the release of GPT-4o, indicate that stabilization may be progressing (Feng *et al*., 2024), as the concerns highlighted in the second frame are gradually being assessed, paving the way for a more effective contribution of AIETs to SME MCSs

The present body of work seeks to contribute to both theory and practice. Research findings gather novel insights, allowing the formulation of theories and notions to better comprehend the role of AI within today's SMEs' MCSs. Furthermore, the empirical evidence clarifies and underscores both the opportunities and limitations of AIETs in supporting SMEs' MCSs. By capturing social groups' interpretive flexibility concerning this technology, the present study provides a nuanced perspective that broadens our understanding of AI's evolving role within the MC field. However, it is important to acknowledge the manuscript's theoretical, contextual, and geographical limitations, highlighting the need for scholars to investigate the role of AI within firms' MCSs continuously.

This article is organized as follows. Section 2 reviews the literature and theoretical foundations. Section 3 details the methodology. Section 4 presents the study's findings, while section 5 discusses them. Conclusions, implications, limitations, and directions for future research are highlighted in Section 6.

2. Literature review

2.1 Management Accounting and its use in SMEs

Simons (1991) defines MA as:

"The formalized routines and procedures that use information to maintain or alter patterns in organizational activity".

In other words, the goal of MA is to assist managers in planning, controlling, and making strategic decisions (Burns and Scapens, 2000; Kaplan, 1984) by providing both financial and non-financial information for use in steering organizations toward the future (Carnegie *et al.*, 2021).

MA's primary objectives remain constant; however, the implementation of its methodologies varies across firms, particularly concerning their respective sizes (Broccardo *et al.*, 2025). In particular, in SMEs, MCSs tend to be rudimental, informal, subjective, and short-period-oriented (Broccardo, Ballesio, *et al.*, 2024). Indeed, the foregoing systems mainly rely on assumptions rather than on data flexibility, which forces them to make fixed decisions (Chowdhury, 2023). Furthermore, SMEs tend to prioritize operational and financial performance measurement, frequently driven by external factors such as taxation or bank requirements (Garengo and Biazzo, 2012). An investigation of the MA tools more frequently adopted by SMEs reveals a clear connection to analyses involving ratios and items that comprise the financial statement (Broccardo *et al.*, 2017). Therefore, it is evident that, while ratio analysis represents a rudimentary tool within MA, it remains widely utilized by SMEs.

On the other hand, there is empirical evidence showing that, especially in smaller entities, MA tends to be externalized to chartered accountants, due to entrepreneurs' lack of time and financial knowledge (Broccardo, Ballesio, *et al*., 2024).

2.2 Using artificial intelligence in management accounting

The concept of AI was first introduced by Nilsson (1983), who described it as the capacity of machines to perceive their environment and make decisions that enhance their chances of achieving their goals. On the other hand, Russell and Norvig (2010) define AI as systems that replicate cognitive processes typically associated with human traits such as language, learning, and problem-solving. Despite some semantic differences, current definitions all emphasize the growing capacity of AI systems to perform tasks and functions that are traditionally performed by humans (Dwivedi *et al*., 2023), such as reasoning, profound adaptation, and decision-making (Oppioli *et al.*, 2023). AI can change businesses' ecosystems and models (Sundström, 2024), improve decision-making processes (Agbon, 2024), and employ data to create more accurate predictions at a reduced cost (Kureljusic and Karger, 2024).

Researchers have shown that AI significantly influences MA more than other DTs, because it allows for the identification of patterns in large sets of accounting data, aiding businesses in their decision-making (Kureljusic and Karger, 2024). Indeed, MA figures are ideally suited for automated examination employing AI models, as they are typically rule-based and well-structured (Nik Abdullah *et al*., 2022; Secinaro *et al*., 2024). More specifically, due to its interrelations, MA data is beneficial for identifying patterns (Carnegie *et al*., 2021). Furthermore, due to the vast number of figures included in the balance sheet, income statement, and cash-flow statement, it is doubtful to identify all the correlations between data without using machine-based evaluations (Vlad and Vlad, 2021). This finding is in harmony with those of Baldwin-Morgan (1995), which state that AI solutions, like expert systems, help organizations capture and apply specialized knowledge in financial analysis and reporting, while neural networks, with their sophisticated pattern recognition abilities, improve the analysis of complex financial datasets and uncover relationships that traditional methods might overlook. Furthermore, recent advancements in deep learning algorithms allow for the identification of common characteristics among firms, thanks to the elaboration of specific balance sheet indicators (Cristiano, 2020). Recent studies have also uncovered the potential of data-mining techniques in cost management, asset management and budgeting (Amani and Fadlalla, 2017). Overall, there is empirical evidence showing that employing AI in MA enhances rapidity and efficiency in organizations, as it reduces manual data entries and improves data speed, quality and accuracy (Secinaro *et al*., 2024).

However, it is important to acknowledge that while AI can have a positive impact on MA, it is not without challenges. For instance, scholars have demonstrated how textual synthesis tools sometimes fall short in capturing the complexity of descriptive information in annual reports (Naidoo and Dulek, 2022). Furthermore, AI solutions might also produce "hallucinations," by generating fictitious, inaccurate, or unverified information (Roberts *et al*., 2024), potentially misleading decision-makers.

2.3 Social Construction of Technology Theory

The authors of this manuscript draw from the SCOT, as it is a suitable theoretical lens for empirical studies investigating the relationship between MA and AI (Gupta *et al*., 2024). In fact, unlike deterministic frameworks, such as the Technology Acceptance Model, the Unified Theory of Acceptance and Use of Technology, and the Technology, Organization and Environment framework, that focus on predefined factors or rigid models of adoption, SCOT does not assume that AI adoption in MA occurs through a predictable, rational process based on perceived usefulness or readiness in relation to technology. In fact, there is empirical evidence showing how AI does not replace traditional MCSs (Secinaro *et al*., 2024); instead, its adoption is very much dependent on the perception of different actors towards its reliability, trustworthiness and integration into existing workflows.

SCOT is defined as an approach that examines the relationship between society and technology (Pinch and Bijker, 1987). More specifically, according to SCOT, technology and society are collectively constructed. Central concepts of SCOT are relevant social groups, interpretive flexibility, technological frame, and closure or stabilization (Sovacool *et al*., 2023).

Relevant social groups are different actors who refer to a technology in the same basic way. Social groups are crucial in defining the issues that emerge when new technologies are developed (Pinch and Bijker, 1987). SCOT theorists identify four different social groups, namely producers, advocates, users, and bystanders. More specifically, *producers* are those who create an artifact with an economic and organizational interest; *advocates* are those devoted to lobbying, policy development and scholarly research on artifacts; *users* are those who utilize the artifact to improve their lives; *bystanders* are those whose opinions indirectly influence the perceptions of what is beneficial and ethical for themselves and others, carrying a social or moral interest (Humphreys, 2005).

Social groups are linked to the notion of interpretative flexibility, as they have different ideas and opinions about how to use a specific technology. In other words, by interacting with an artifact, relevant social groups influence its design and interpretation (Humphreys, 2005). In fact, as each social group has a different perception of the issues related to technology, it contributes to its development (Elle *et al*., 2010).

On top of that, there is the technological frame, which depicts the relationships between the various social groups and their negotiated interpretative flexibility regarding a technology (Pinch and Bijker, 1987).

As interpretive flexibility wanes over time, certain artifacts assume greater dominance, ultimately culminating in the creation of a singular entity, thereby bringing the process of social construction of technology to a close (Pinch and Bijker, 1987). This is the final concept of SCOT, namely closure or stabilization.

Regarding the context analyzed in this study, we will analyze the perceptions and interpretative flexibility of different social groups, who share the same basic view of the technology, on ChatGPT-4 assistants' answers on MA, thus assessing whether AI could contribute to SMEs' MCSs. Following Sovacool *et al*. (2023) study, a more expansive and flexible understanding of relevant social groups is adopted, as the original SCOT categories are found to be inapplicable to the present case. The present study comprised two groups of actors. The first group comprised one programmer and two MA experts, who were responsible for the creation and evaluation of the artifact's reliability. The second group comprised individuals situated within the broader societal context, including entrepreneurs and academics, who analysed a predetermined set of outputs generated by the tool, without having directly contributed to its creation. However, given the exploratory nature of this research, it is important to note that the actual artifact producers of the first group did not have an economic interest in the technology itself. Instead, they may be regarded as advocates or even bystanders. Moreover, the artifact was conceived with the intention of eliciting perceptions from potential users, rather than analyzing actual real-world implementations. Accordingly, the second social group included academics and entrepreneurs with varying degrees of financial literacy and familiarity with digital technologies.

3. Research method

To address our RQ, a mixed-method approach was used. First, we used open coding and thematic analysis to assess the quality and correctness of ChatGPT-4 answers (Kabir *et al*., 2024) when evaluating SMEs' performance ratios. Second, we performed semi-structured interviews with 15 respondents to assess their opinions on the artifact. A thorough explanation of every technique is given in the ensuing sections.

3.1. Data collection

3.1.1 Sample Definition

A sample of 37 Italian SMEs, randomly chosen from Bureau Van Dijk's

AIDA database, was selected for the present study. AIDA is a well-reputed database known for incorporating companies' financial statements details, as well as other information, such as geographical location, sector, year of incorporation, ownership, and equity participation in other firms, over a 10-year timespan.

The selection process was guided by specific criteria, which were in line with deliberate sampling techniques. First, businesses had to fit the European Commission's definition of SMEs. In other words, we exclusively searched for firms employing from 10 to 249 people and with a turnover higher than € 2 million but lower than € 50 million or a balance sheet total higher than € 2 million but lower than € 43 million (European Commission, 2003). As previously stated, this decision was made since SMEs account for over 99% of all businesses in the EU and Italy (European Commission, 2021), significantly impacting the country's economy and serving as a good representation of the competitive environment.

Beyond the aforementioned criteria, the final sample selection was guided by the concept of data saturation. Data saturation can be defined as the point beyond which no new insights can be obtained and where additional data collection will, at best, make only slight improvements in empirical contributions (Guest *et al*., 2006). Indeed, after carefully assessing the financial data retrieved from the AIDA database, we believed that data saturation was reached at 37 companies. Indeed, we judged that additional firms would not have been able fundamentally to change our findings or reveal newer patterns. An in-depth description of the selected companies' main details can be found in Appendix 1.

After defining our final sample, we utilized the companies' financial figures to calculate eight ratios, namely fixed assets coverage, interest coverage, quick ratio, debt-equity ratio, ROI, ROE, ROS, and turnover. We thus obtained a total of 296 different ratio figures.

3.1.2 ChatGPT-4 Assistants Creation

In order to better tune ChatGPT answers and to avoid hallucinations (Aryal *et al*., 2024), two MA experts, who are the authors of this study, and a programmer created 8 different GPT assistants for all the ratios analyzed. Each assistant was built on the OpenAI Platform and features a brief definition of each ratio, its calculation method, a description of all the variables used to calculate it, ratio interpretation, and threshold values specific for each industry, below which performance is considered negative. In order to create said assistants, we employed the ChatGPT-4 version.

3.1.3 ChatGPT-4 Answers Collection

Two MA experts and a programmer created a single question prompt for each of the eight ratios, which they then fed to the ChatGPT-4 version on the OpenAI Platform. A detail of the question prompt can be found in Appendix 2.

The GPT-4 version was chosen as it is the only one that has the possibility of creating AI-empowered assistants.

All ChatGPT-assistant answers were stored in CSV files. As the OpenAI Platform keeps track of past inputs and outputs, a new chat session was initiated, before feeding each question prompt to ChatGPT-assistants. A total of 296 answers were obtained.

3.2 Open coding and thematic analysis

The present section will illustrate the process employed to assess the 296 ChatGPT-assistants' answers, to gather the interpretative flexibility of the actors who created the artifact.

ChatGPT-assistant answers were labeled at the phrase level to evaluate their accuracy and quality (Kabir *et al.*, 2024). Two authors convened several times to develop, refine, and finalize the criteria for evaluating ChatGPT replies. First, the authors got familiar with the answers dataset. To ensure rigor and eliminate possible bias, two researchers independently classified twenty ChatGPT answers and provided comments about their observations. Following Kabir *et al.* (2024) work, the two scholars met to agree on their labeling. Then, they used thematic analysis to group the labels into four categories, namely correctness, consistency, conciseness, and comprehensiveness.

Correctness was assessed by comparing ChatGPT answers with relevant and reputed literature on MA. We searched for three categories of errors: factual, conceptual, and terminological. If an answer did not contain any of these errors, it was deemed to be totally right.

Consistency was measured by comparing the answers between them (Jang and Lukasiewicz, 2023). More specifically, for companies with similar ratio values, within the same industry or with comparable financial structures or size, we compared the ChatGPT comments to check whether they were coherent between them. It is important to note that inconsistency does not necessarily mean incorrectness: two answers can be both correct, but inconsistent between them.

Conciseness issues were related to the presence of redundant, irrelevant, and excess information (Kabir *et al.*, 2024). Furthermore, we considered as not concise all those answers which were wordy, verbose or vague.

Comprehensiveness was related to an overall assessment of the answer. When all aspects of the question were addressed in an answer, it was deemed comprehensive (Cohen *et al.,* 2023).

The answers labels were then compared. After agreeing on the labeling, all the authors evaluated all the 296 answers until a final consensus was reached.

It is important to note that labels are not mutually exclusive, as an answer may have many quality concerns, thus resulting in multi-labeling (Kabir *et al*., 2024).

3.3 Qualitative study

With the aim of exploring the interpretative flexibility of the other social groups, namely entrepreneurs, and academics, we performed a study involving 15 participants.

3.3.1 Participants selection

15 participants were recruited for our qualitative investigation. We deliberately decided to incorporate a variety of viewpoints from the entrepreneurial and academic context. In fact, 3 participants were business and management PhD students, 4 participants were MA research fellows, 5 participants were MA professors, and 3 participants were SME entrepreneurs.

Purposive sampling and chain referral sampling were the techniques used for participant selection (Patton, 2015). Purposive sampling allowed us to choose interviewees based on predetermined criteria pertinent to our research (Fernandez-Vidal *et al*., 2022). On the other hand, through chain referral sampling we asked selected participants to recommend other experts who might contribute to the present research (Mack *et al*., 2005). To guarantee a certain level of cultural integrity or to prevent cultural biases (Pelzang and Hutchinson, 2017), we selected Italian participants, as it is the same country as the chosen companies.

Participants were asked to rate their financial literacy, level of competence and experience with MA and MCS, and familiarity with digital technologies and generative AI tools using a Likert scale ranging from 1 to 5. Appendix 3 provides an in-depth analysis of the 15 interviewees.

3.3.2 Data collection and analysis

Data was collected through 15 semi-structured interviews with business and management PhD students, MA research fellows, MA professors, and SME entrepreneurs. Appendix 4 provides an in-depth analysis of the 15 interviews.

Interviews were conducted in person or via video, from March to July 2024. The duration of each interview was from 27 to 75 minutes.

Each interview was conducted by two authors together, ensuring consistency in the questioning process. Semi-structured guidelines were employed during all the interviews. Interviewees were provided with a CSV file containing the ChatGPT responses and were asked to comment on them. Participants were asked open-ended questions to rate the overall quality and correctness of each response, and provide their opinion on the practical usefulness of the artifact in MC practices.

Due to respondents' privacy concerns, the authors could not obtain permission to record every interview due to privacy concerns, but they meticulously took notes of the salient points and provided a concise summary of the key takeaways at the conclusion of each interview. This approach ensured accurate documentation of the interviews (Fernandez-Vidal *et al.*, 2022). Lastly, based on the theoretical saturation model of Glaser and Strauss (1968), the interview process ended when the authors felt that they could not obtain any more relevant insights for the present study (Saunders *et al.*, 2018).

To ensure consistency and reliability of the results, the authors conducted a cross-comparison between the obtained findings. More specifically, two researchers independently performed coding and thematic analysis. Their results were then compared, and no significant differences were found. Finally, the two remaining authors evaluated and discussed the coding and thematic analysis until a final consensus was reached.

To ensure the scientific rigor of our analysis, the Grounded Theory approach (Glaser and Strauss, 1968; Lincoln and Guba, 1985), combined with the methodology of Gioia *et al*. (2013) was employed. The data analysis process consisted of three steps. First, first-order concepts were noted through an open coding method (Strauss and Corbin, 1998). We searched for common threads about our interviewees' perceptions of ChatGPT answers. In line with Strauss and Corbin's (1998) axial coding method, we limited the emerging themes to a block of 25 first-order concepts by comparing and contrasting them. Second, we searched for second-order themes to aid in explaining first-order concepts (Gioia *et al*., 2013). Third, after identifying 25 first-order concepts, and 8 second-order themes, we reduced these latter into aggregate dimensions (Gioia *et al*., 2013), following Glaser and Strauss's (1968) theoretical saturation model.

4. Findings

The present section will illustrate the findings obtained from the analysis of the ChatGPT-assistants' answers and the interviews.

4.1 Open coding and thematic analysis results

This section showcases our results, organized inductively around the aforementioned categories, namely correctness, consistency, conciseness and comprehensiveness (see Appendix 5). These dimensions reflect the interpretative flexibility of the social group of the actors who created the artifact. Their design choices, guided by the goal of ensuring the quality of the answers have influenced the artifact output characteristics. For example, one key design choice consisted of employing custom-built assistants featuring notions from relevant and reputed literature, to better tune in ChatGPT answers, thus increasing their overall quality. That explains the low percentage of incorrect answers (3.4%) and the high number of correct responses (96.6%). The vast majority of errors were related to incorrect terminology and some hallucinations were found in answers related to ratios with multiple components, such as ROI.

However, utilizing custom-built assistants did not prevent some answers (23%) from being inconsistent between them. More specifically, in some cases, even if two companies, with similar characteristics such as sector and number of employees, had close ratio values, ChatGPT evaluations differ a lot between them. Appendix 6 provides an example of two inconsistent answers.

The main issue encountered while evaluating answers was related to conciseness. More specifically, most ChatGPT comments (85.1%) were not concise, meaning they contained redundant, irrelevant, and excess information or were wordy, verbose, or vague. Concerning redundant content, we noticed a number of ChatGPT answers repeating information already mentioned in other sections of the reply. Furthermore, we saw a few answers providing irrelevant information, not pertaining to the context analyzed. The main issues were related, though, to the presence of excess information, even if that was not mentioned in the assistants. More specifically, we noted that ChatGPT commented on the ratio trends, not only based on the ratio components but also by adding not-required insights or providing not-required solutions in case of positive and negative situations. Furthermore, the presence of excess information made many answers wordy, resulting in a prompt that was not clear. Lastly, we noticed an excessive use of conditional statements, such as "*It might be due to…*" and "*This could have been caused by…*" that made the answers vague. Appendix 6 provides some examples of not-concise answers. However, it is important to note that question recency has a significant impact on the conciseness of the answers. More specifically, answers generated after May 13th, 2024 (the release date of GPT-4o) are more concise than those obtained before said date. Furthermore, we also noted that answers related to ratios with many components, such as ROI and debt-to-equity ratio, are less concise. We observed this trend even after the GPT-4o release despite seeing a notable improvement in conciseness.

Finally, our findings show how the majority of ChatGPT answers (92%) are comprehensive, as in most cases all aspects of the question were addressed in the answer.

4.2 Qualitative study results

Thanks to the thematic analysis of the interviews, we explored how the other social groups – entrepreneurs and academics – interpret and evaluate ChatGPT answers. Appendix 7 illustrates the resulting framework, synthesizing their perceptions into three interpretative dimensions, namely "*Lack of conciseness*", "*Correctness*", and "*Applicability in SMEs*". Each dimension tackles the tool's perceived advantages, limitations and possible applications in the SME context. Follows a more thorough explanation of how these groups make sense of this technology.

According to our respondents, there are both strengths and challenges associated with the use of our artifact in SMEs. The main issue encountered is related to the lack of conciseness of the answers. First, entrepreneurs, who often work under time constraints and lack financial literacy, felt overwhelmed by the amount of information included in a single answer. Similarly, academics described many answers as redundant, due to the reiteration of the same concepts in multiple ways, or to the presence of unnecessary details. Additionally, some answers were found to be vague and wordy, presenting hypotheses and over-complicated sentences to explain simple concepts. In a rapidly moving business environment, our respondents stressed the need for short and straightforward answers to help in the decision-making process. This is particularly relevant, especially for those who lack financial literacy. Notably, they found the latest responses, generated after the ChatGPT-4o release, satisfying in terms of conciseness.

However, despite the conciseness issues, all our respondents recognized the answers' general correctness. The artifact was found to provide accurate information, with only minor terminology-related issues. Furthermore, the tool demonstrates a clear comprehension of the topic.

Academics identified some concerns regarding consistency. Specifically, they highlighted how the tool provided inconsistent evaluations for companies with similar characteristics – such as sector and number of employees – which had close ratio values. This raised concerns about the reliability of AI-generated assessments, particularly in professional and educational contexts where consistency is essential.

A crucial aspect of interpretative flexibility emerged when the interviewees were asked about the applicability of the artifact within SMEs. On the one hand, in SMEs concurrent control and short-term analysis are prioritized over ratio-analysis, which is seen as an additional practice entrepreneurs do once a year. On the other hand, our respondents reported how such a tool would support entrepreneurs in ratio analysis, making this practice more utilized in smaller firms for decision-making. Therefore, this could lead to developing more objective, formal and long-period oriented MCSs in smaller businesses. Additionally, our interviewees noted that, with an improvement in conciseness, the artifact would also support people who lack financial literacy in better understanding and implementing ratio analysis in their firms.

In summary, the interpretative flexibility of the artifact varies across social groups, reflecting their different needs and evaluation criteria. Entrepreneurs prioritize efficiency and ease of comprehension, while academics precision and conceptual rigor. All groups acknowledge how such an artifact would support MA in SMEs, mainly thanks to its ability to generate generally correct answers. However, for this to happen, the technology needs some refinements. Improving conciseness and reducing redundancy and vagueness would make the artifact an effective tool for nurturing SMEs MCSs.

5. Discussion

The present manuscript seeks to enrich the existing body of literature by providing new insights into the contribution of AIETs to MCSs in SMEs (Nóbrega *et al.*, 2023). To answer our RQ, we employed the SCOT lens (Pinch and Bijker, 1987) to explore the relevant social groups, interpretive flexibility, technological frame, and patterns of closure associated with AIETs and their contribution to MCSs. The interpretative flexibility of the different social groups taken into consideration in the present study led to the construction of two frames – one positive and one negative – shaping the perception of AIETs in MA in the SMEs context. The first positive frame – *AI as a MA-support tool for SMEs* – sees ChatGPT assistants as valuable tools that could provide significant support and enhance SMEs' MCSs. On the other hand, the second negative frame – issues of AIETs – focuses on some concerns of the artifact that could impact its reliability.

5.1 AI as a MA-support tool for SMEs

This frame focuses on the positive characteristics of the artifact, such as correctness and comprehensiveness of the answers, thus interpreting it as a tool that could enhance MA practices in SMEs, particularly by formalizing ratio analysis practice and providing support to entrepreneurs with limited financial literacy. In fact, through an open coding and thematic analysis, the artifact creators found that 96.6% of the answers were correct, and 92% comprehensive. The qualitative investigation confirmed these results, as both entrepreneurs and academics were satisfied with the overall correctness of GPT answers, demonstrating that the tool has a clear comprehension of the topic. Prior studies support our findings (Aryal *et al.*, 2024; Kabir *et al.*, 2024). The overall correctness of the responses is mainly attributed to the use of custom-built assistants featuring reputed literature (Aryal *et al.*, 2024). This design choice ensured that the only errors spotted were related to terminology (Albuquerque and Gomes dos Santos, 2024).

Furthermore, our empirical analysis shows how the artifact can address all aspects of the question in the vast majority of the answers (92%). This is in harmony with previous research, which stresses how Chat GPT provides comprehensive answers (Cohen *et al.*, 2023).

Beyond correctness and comprehensiveness, our respondents suggest that the artifact would prove helpful in supporting entrepreneurs in ratio analysis, making this practice more utilized in smaller firms. Indeed, SMEs MCSs are informal, subjective, short-period-oriented (Broccardo *et al.*, 2017), and assumptions-relying (Chowdhury, 2023), thus making ratio analysis often overlooked. To mitigate these issues, AI assistance would prove helpful by including a more significant amount of data in the forecast, thus giving the possibility to make long-term and dynamic decisions and eliminating subjectivity (Losbichler and Lehner, 2021).

5.2 Issues of AI-empowered tools

This frame challenges some of the artifact issues, namely the consistency and conciseness of the responses. These negative aspects raised some concerns among the different social groups about the artifact reliability.

Firstly, our findings highlight that some inconsistencies (23%) between answers were present, especially when evaluating similar companies with close financial ratios. Our results are consistent with prior literature, which documents that Chat GPT often provides different outputs between paraphrased input questions (Jang and Lukasiewicz, 2023). This could represent a problem for decision makers, who may struggle to trust the artifacts, if it provides different answers for similar scenarios (Krügel *et al*., 2023).

Secondly, conciseness emerged as the primary issue, with 85.1% of responses being redundant or containing excessive information. The qualitative investigation confirmed these results, as both entrepreneurs and academics expressed concerns over the tool's verbosity, describing answers as wordy, vague, and sometimes overwhelming. The foregoing result is in harmony with earlier academic literature, which documents how ChatGPT has a verbosity bias (Saito *et al*., 2023), tending to favor longer and wordy answers (Kabir *et al*., 2024). Furthermore, our results shed light on how more significant levels of excess information and conditional statements are likely to translate into not concise answers (Kabir *et al*., 2024), resulting in an output that is unclear and hard to understand (Xiong, 2024). This issue is particularly relevant for SME entrepreneurs with limited financial literacy, who need clear and concise insights rather than lengthy and complex explanations. However, we observed a notable improvement after the GPT-4o release (Feng *et al*., 2024), suggesting that ongoing technological advancements are addressing this limitation.

5.3 Towards closure and stabilization

The final theme identified concerns closure and stabilization or the absence thereof. Regarding the context analyzed in the present study, the interpretative flexibility of the different social groups indicates that full closure and stabilization have not yet been achieved. While a degree of rhetorical closure is emerging – highlighted by the widespread recognition of the artifact's ability to answer correctly and comprehensively, and thus its positive role in nurturing SMEs MCSs – some concerns remain about the responses' consistency and conciseness. The different social groups perceive the tool as a valuable support for SMEs MCSs, particularly in the form of ratio analysis, reinforcing the notion that AIETs can formalize MA and decision-making in smaller companies (Losbichler and Lehner, 2021;

Nóbrega *et al.*, 2023). However, responses inconsistencies and verbosity (Kabir *et al.*, 2024; Jang and Lukasiewicz, 2023) challenge this stabilization. According to Pinch and Bijker (1987), an artifact reaches its stabilization when these heterogeneous concerns – in this case wordiness and inconsistency – are adequately addressed. The recent improvements seen after the release of GPT-4o suggest that stabilization may be progressing (Feng *et al.*, 2024), but the process is still incomplete. This emphasizes the actual lack of closure related to the artifact, but, at the same time, the crucial role AI training methodologies will play in determining whether AIETs can achieve a true closure within SMEs, ultimately becoming fully integrated in MCSs, making them less informal, subjective and short-period-oriented (Losbichler and Lehner, 2021), and helping entrepreneurs assess their companies trends and performance (Nóbrega *et al.*, 2023), through ratio analysis, without delegating this process to chartered accountants.

6. Conclusion

The present study explores how AIETs, namely custom-built AI assistants featuring ChatGPT-4 technology, could contribute to SME MCSs. Through the lens of the SCOT, we assessed the relevant social groups, interpretive flexibility, technological frame, and patterns of closure associated with the artifact we built, in an attempt to identify its possible contribution to MA in smaller companies. Therefore, this paper bridges several research gaps and represents a valuable starting point for developing this subject.

6.1 Practical and Theoretical Implications

Our study makes several contributions to both academics and practitioners. First, it contributes to the AI and MC literature strand (Nik Abdullah *et al.*, 2022; Secinaro et al., 2024; Valentinetti and Rea, 2023). Second, we further our understanding of the SCOT, demonstrating that it can be effectively employed to analyze AIETs on SME MCSs (Gupta *et al.*, 2024). This highlights how such a technology is not merely a standalone artifact, but rather embedded within broader sociotechnical systems, shaped by the interactions between different social groups, such as entrepreneurs, academics, and technology designers. Third, our findings could prove helpful to SMEs seeking to utilize AI in MA. For example, entrepreneurs, with the aid of academics and programmers, could build customized GPT assistants to nurture and enhance their companies' MCSs, thus optimizing their decision-making processes. Fourth, our findings may offer valuable insights to both entrepreneurs and academics by highlighting the positive and negative aspects of GPT answers, thus providing helpful guidance on how to effectively utilize AI assistants for MC practices.

6.2 Limitations and Directions for Future Research

The present manuscript has various limitations which can serve as a basis for further research. First, while our investigation provides insights into the perceived usefulness and limitations of AIETs for SMEs' MCSs it does not constitute a real-world deployment of the tool within organizational contexts. While this approach is consistent with the interpretive and exploratory nature of early-stage SCOT-informed research, it limits our ability to assess actual adoption dynamics or behavioral changes in management control practices. Future research could build upon these exploratory findings by conducting longitudinal case studies or field experiments that observe the integration of AIETs into SMEs' MCSs over time. Second, the generalizability of our results is limited by our sampling approach and the small sample size. Further works should thus carry out empirical analyses in different contexts and locations. For instance, comparative studies between countries with different reporting or MA practices could prove particularly useful in highlighting the contribution of GPT assistants on MCSs. Furthermore, future research could explore the perception of large or micro-enterprises to get a more nuanced understanding of the impact of GPT assistants on MA. Third, the present manuscript analyzes only eight ratios. Future studies should analyze other ratios, as well as the economic, property, and financial balances. Lastly, this empirical body of work is informed and limited by the SCOT. Consequently, in order to gain a deeper understanding of the contribution of AI to MCSs, future further studies could employ other theoretical lenses like technology appropriation and contingency theory.

References

Agbon G. (2024), Who speaks through the machine? Generative AI as discourse and implications for management, *Critical Perspectives on Accounting*, *100*, 102761. DOI: 10.1016/j.cpa.2024.102761.

Albuquerque F., and Gomes dos Santos P. (2024), Can ChatGPT be a certified accountant? Assessing the responses of ChatGPT for the professional access exam in Portugal, *Administrative Sciences*, *14*(7), 152. DOI: 10.3390/admsci14070152.

Amani F. A., and Fadlalla A. M. (2017), Data mining applications in accounting: A review of the literature and organizing framework, *International Journal of Accounting Information Systems*, *24*, pp. 32-58. DOI: 10.1016/j.accinf.2016.12.004.

Aryal S., Do T., Heyojoo B., Chataut S., Gurung B. D. S., Gadhamshetty V., and Gnimpieba E. (2024), Leveraging multi-AI agents for cross-domain knowledge discovery, In *arXiv [cs.AI]*. DOI: 10.48550/ARXIV.2404.08511.

Baldwin-Morgan A. A. (1995), Integrating artificial intelligence into the accounting curriculum, *Accounting Education*, *4*(3), pp. 217-229. DOI: 10.1080/09639289500000026.

Bankins S., Ocampo A. C., Marrone M., Restubog S. L. D., and Woo S. E. (2024), A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice, *Journal of Organizational Behavior*, *45*(2), pp. 159-182. DOI: 10.1002/job.2735.

Broccardo L., Ballesio E., Crocco E., and Giordino D. (2024), Performance management systems in micro, small, and medium enterprises: Lessons learned from times of uncertainty, *Management Control*, *3*, pp. 15-38. DOI: 10.3280/maco2024-003002.

Broccardo L., Ballesio E., Yaqub M. Z., & Mohapatra A. K. (2025), A bridge to success: the role of management accountants' intellectual capital in driving organizational decision-making through knowledge management, *Journal of Knowledge Management*, *29*(5), pp. 1365-1411. DOI: 10.1108/jkm-04-2024-0452.

Broccardo L., Giacosa E., Culasso F., and Ferraris A. (2017), Management control in Italian SMEs, *Global Business and Economics Review*, *19*(5), 632. DOI: 10.1504/gber.2017.086781.

Broccardo L., Tenucci A., Agarwal R., and Alshibani S. M. (2024), Steering digitalization and management control maturity in small and medium enterprises (SMEs), *Technological Forecasting and Social Change*, *204*, 123446. DOI: 10.1016/j.techfore.2024.123446.

Burns J., and Scapens R. W. (2000), Conceptualizing management accounting change: an institutional framework, *Management Accounting Research*, *11*(1), pp. 3-25. DOI: 10.1006/mare.1999.0119.

Cardoni A., Hiebl M., and Paradisi A. (2023), Management accounting implementation in SMEs: A Structured Literature Review, *Management Control*, *2*, pp. 189-214.

Carnegie G., Parker L., and Tsahuridu E. (2021), It's 2020: What is accounting today?, *Australian Accounting Review*, *31*(1), pp. 65-73. DOI: 10.1111/auar.12325.

Chowdhury E. K. (2023), Integration of artificial intelligence technology in management accounting information system: An empirical study, in *International Series in Operations Research and Management Science* (pp. 35-46). Springer International Publishing.

Cohen S., Manes Rossi F., and Brusca I. (2023), Debate: Public sector accounting education and artificial intelligence, *Public Money and Management*, *43*(7), pp. 725-726. DOI: 10.1080/09540962.2023.2209290.

Cooper L. A., Holderness D. K., Sorensen T. L., and Wood D. A. (2019), Robotic process automation in public accounting, *Accounting Horizons*, *33*(4), pp. 15-35.

Cristiano E. (2020), Intelligenza artificiale, impatto sul sistema aziendale e performance, in Lombardi R., Chiucchi M.S., Mancini D. (Ed.), *Smart Technologies, Digitalizzazione e Capitale intellettuale: sinergie e opportunità* (pp. 157-175). Milano: FrancoAngeli.

Culasso F., Broccardo L., Giordino D., Crocco E. (2024), Performance management systems and digital technologies: an exploratory analysis of practitioners and academics' perspectives, *Industrial Management & Data Systems*, *124*(10), pp. 2870-2898. DOI: 10.1108/IMDS-11-2023-0834.

Dwivedi Y. K., Kshetri N., Hughes L., Slade E. L., Jeyaraj A., Kar A. K., Baabdullah A. M., Koohang A., Raghavan V., Ahuja M., Albanna H., Albashrawi M. A., Al-Busaidi A. S., Balakrishnan J., Barlette Y., Basu S., Bose I., Brooks L., Buhalis D., … Wright R. (2023), Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy, *International Journal of Information Management*, *71*, 102642. DOI: 10.1016/j.ijinfomgt.2023.102642.

Elle M., Dammann S., Lentsch J., and Hansen K. (2010), Learning from the social construction of environmental indicators: From the retrospective to the pro-active use of SCOT in technology development, *Building and Environment*, *45*(1), pp. 135-142. DOI: 10.1016/j.buildenv.2009.05.011.

European Commission (2003), *Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises*.

European Commission (2021), *The digitalisation of small and medium-sized enterprises in Italy*.

Feng Z., Li B., and Liu F. (2024), *A first look at financial data analysis using ChatGPT-4o,* DOI: 10.2139/ssrn.4849578.

Fernandez-Vidal J., Antonio Perotti F., Gonzalez R., and Gasco J. (2022), Managing digital transformation: The view from the top, *Journal of Business Research*, *152*, pp. 29-41. DOI: 10.1016/j.jbusres.2022.07.020.

Garengo P., and Biazzo S. (2012), Unveiling strategy in SMEs through balanced scorecard implementation: a circular methodology, *Total Quality Management and Business Excellence*, *23*, pp. 79-102.

Gioia D. A., Corley K. G., and Hamilton A. L. (2013), Seeking qualitative rigor in inductive research: Notes on the Gioia methodology, *Organizational Research Methods*, *16*(1), pp. 15-31. DOI: 10.1177/1094428112452151.

Giordino D, Troise C., Vanhaverbeke W., Culasso F. (2024), Stakeholder engagement processes for the made in Italy small- and medium-sized enterprises: Value co-creation in the stakeholder network, *Business Ethics, the Environment & Responsibility,* pp. 1-18. DOI: 10.1111/beer.12696.

Glaser B. G., and Strauss A. L. (1968), The discovery of grounded theory: Strategies for qualitative research, *Social Forces; a Scientific Medium of Social Study and Interpretation*, *46*(4), 555. DOI: 10.2307/2575405.

Grassi L., and Lanfranchi D. (2022), RegTech in public and private sectors: the nexus between data, technology and regulation, *Economia e Politica Industriale*, *49*(3), pp. 441-479. DOI: 10.1007/s40812-022-00226-0.

Guest G., Bunce A., and Johnson L. (2006), How many interviews are enough?: An experiment with data saturation and variability, *Field Methods*, *18*(1), pp. 59-82. DOI: 10.1177/1525822x05279903.

Gupta P., Mahajan R., Badhera U., and Kushwaha P. S. (2024), Integrating generative AI in management education: A mixed-methods study using social construction of technology theory, *The International Journal of Management Education*, *22*(3), 101017. DOI: 10.1016/j.ijme.2024.101017.

Humphreys L. (2005), Reframing social groups, closure, and stabilization in the social construction of technology, *Social Epistemology*, *19*(2-3), pp. 231-253. DOI: 10.1080/02691720500145449.

Jang M., and Lukasiewicz T. (2023). Consistency Analysis of ChatGPT. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing.*

Kabir S., Udo-Imeh D. N., Kou B., and Zhang T. (2024), Is stack overflow obsolete? An empirical study of the characteristics of ChatGPT answers to stack overflow questions, *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1-17.
Kaplan R. S. (1984), *The evolution of management accounting. Readings in accounting for management control*, pp. 586-621.
Kostić N., and Sedej T. (2022), Blockchain technology, inter-organizational relationships, and management accounting: A synthesis and a research agenda, *Accounting Horizons*, *36*(2), pp. 123-141. DOI: 10.2308/horizons-19-147.
Krügel S., Ostermaier A., and Uhl M. (2023), ChatGPT's inconsistent moral advice influences users' judgment, *Scientific Reports*, *13*(1), 4569. DOI: 10.1038/s41598-023-31341-0.
Kureljusic M., and Karger E. (2024), Forecasting in financial accounting with artificial intelligence – A systematic literature review and future research agenda, *Journal of Applied Accounting Research*, *25*(1), pp. 81-104. DOI: 10.1108/jaar-06-2022-0146.
Lincoln Y. S., and Guba E. G. (1985), *Naturalistic Inquiry*, SAGE Publications.
Lombardi R., Trequattrini R., Schimperna F., and Cano-Rubio M. (2021), The impact of smart technologies on the Management and strategic control: A structured literature review, *Management Control*, *1*, pp. 11-30. DOI: 10.3280/maco2021-001-s1002.
Losbichler H., and Lehner O. M. (2021), Limits of artificial intelligence in controlling and the ways forward: a call for future accounting research, *Journal of Applied Accounting Research*, *22*(2), pp. 365-382. DOI: 10.1108/jaar-10-2020-0207.
Mack N., Woodsong C., and Macqueenm K. M. (2005), *Qualitative research methods: A data collector's field guide Family Health International*, Family Health International.
Maelah R., Al Lami M. F. F., and Ghassan G. (2020), Management accounting information usefulness and cloud computing qualities among small medium enterprises, *International Journal of Management Studies*, *26*. DOI: 10.32890/ijms.26.1.2019.10511.
Merjane S., Touili C., Khalil M., Touili K., and Fikri M. (2024), Management control tools for Moroccan industrial companies: Application of Target Costing and Artificial Intelligence, *Journal of Autonomous Intelligence*, *7*(5), 1524. DOI: 10.32629/jai.v7i5.1524.
Naidoo J., and Dulek R. E. (2022), Artificial intelligence in business communication: A snapshot, *International Journal of Business Communication*, *59*(1), pp. 126-147. DOI: 10.1177/2329488418819139.
Nik Abdullah N. H., Krishnan S., Mohd Zakaria A. A., and Morris G. (2022), Strategic management accounting practices in business: A systematic review of the literature and future research directions, *Cogent Business and Management*, *9*(1). DOI: 10.1080/23311975.2022.2093488.
Nóbrega V., Lopes R., Costa A., Gonçalves R., Dias Á., Pereira L., and Dorner K. (2023), The impact of artificial intelligence in accounting: application in SMEs, *International Journal of Electronic Finance*, *12*(2), 192. DOI: 10.1504/ijef.2023.129923.
O'Leary D. E. (2023), Digitization, digitalization, and digital transformation in accounting, electronic commerce, and supply chains, *Intelligent Systems in Accounting Finance and Management*, *30*(2), pp. 101-110. DOI: 10.1002/isaf.1524.
Omrani N., Rejeb N., Maalaoui A., Dabić M., and Kraus S. (2024), Drivers of Digital Transformation in SMEs, *IEEE Transactions on Engineering Management*, *71*, pp. 5030-5043. DOI: 10.1109/tem.2022.3215727.
Oppioli M., Sousa M. J., Sousa M., and de Nuccio E. (2023), The role of artificial intelligence for management decision: a structured literature review, *Management Decision*. DOI: 10.1108/md-08-2023-1331.

Patton M. Q. (2014), *Qualitative research and evaluation methods: Integrating theory and practice*, Sage publications.

Pelzang R., and Hutchinson A. M. (2017), Establishing cultural integrity in qualitative research: Reflections from a cross-cultural study, *International Journal of Qualitative Methods*, *17*(1).

Perrini F. (2006), SMEs and CSR theory: Evidence and implications from an Italian perspective, *Journal of Business Ethics*, *67*(3), pp. 305-316.

Pinch T., and Bijker W. E. (1987), *The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other*, MIT Press.

Rikhardsson P., and Yigitbasioglu O. (2018), Business intelligence and analytics in management accounting research: Status and future focus, *International Journal of Accounting Information Systems*, *29*, pp. 37-58. DOI: 10.1016/j.accinf.2018.03.001.

Roberts J., Baker M., and Andrew J. (2024), Artificial intelligence and qualitative research: The promise and perils of large language model (LLM) 'assistance', *Critical Perspectives on Accounting*, *99*, 102722. DOI: 10.1016/j.cpa.2024.102722.

Russell S. J., and Norvig P. (2010), *Artificial intelligence: A modern approach*, Prentice Hall.

Saito K., Wachi A., Wataoka K., and Akimoto Y. (2023), Verbosity bias in preference labeling by Large Language Models, in *arXiv [cs.CL]*. DOI: 10.48550/ARXIV.2310.10076.

Saunders B., Sim J., Kingstone T., Baker S., Waterfield J., Bartlam B., Burroughs H., and Jinks C. (2018), Saturation in qualitative research: exploring its conceptualization and operationalization, *Quality and Quantity*, *52*(4), pp. 1893-1907. DOI: 10.1007/s11135-017-0574-8.

Secinaro S., Calandra D., Lanzalonga F., and Biancone P. (2024), The role of artificial intelligence in management accounting: An exploratory case study, in *Digital Transformation in Accounting and Auditing* (pp. 207-236), Springer International Publishing.

Simons R. (1991), Strategic orientation and top management attention to control systems, *Strategic Management Journal*, *12*(1), pp. 49-62.

Skare M., de las Mercedes de Obesso M., and Ribeiro-Navarrete S. (2023), Digital transformation and European small and medium enterprises (SMEs): A comparative study using digital economy and society index data, *International Journal of Information Management*, *68*, 102594. DOI: 10.1016/j.ijinfomgt.2022.102594.

Sovacool B. K., Iskandarova M., and Geels F. W. (2023), "Bigger than government": Exploring the social construction and contestation of net-zero industrial megaprojects in England, *Technological Forecasting and Social Change*, *188*, 122332. DOI: 10.1016/j.techfore.2023.122332.

Strauss A. L., and Corbin J. M. (1998), *Basics of qualitative research: Techniques and procedures for developing grounded theory*, SAGE Publications.

Sundström A. (2024), AI in management control: Emergent forms, practices, and infrastructures, *Critical Perspectives on Accounting*, *99*, 102701. DOI: 10.1016/j.cpa.2023.102701.

Valentinetti D., and Rea M. (2023), Intelligenza artificiale e accounting: le possibili relazioni, *Management Control*, *2*, pp. 93-116. DOI: 10.3280/maco2023-002005.

Valentinetti D., and Rea M. A. (2025), Factors influencing the digitalization of sustainability accounting, reporting and disclosure: a systematic literature review, *Meditari Accountancy Research*, *33*(2), pp. 633-680. DOI: 10.1108/medar-02-2024-2385.

Vasarhelyi M. A., Kogan A., and Tuttle B. M. (2015), Big data in accounting: an overview, *Accounting Horizons*, *29*(2), pp. 381-396.

Verhoef P. C., Broekhuizen T., Bart Y., Bhattacharya A., Qi Dong J., Fabian N., and Haenlein M. (2021), Digital transformation: A multidisciplinary reflection and research agenda, *Journal of Business Research*, *122*, pp. 889-901. DOI: 10.1016/j.jbusres.2019.09.022.

Vlad M., and Vlad S. (2021), The use of machine learning techniques in accounting. A short survey, *Journal of Social Sciences*, *4*(3), pp. 139-143.

Wei R., and Pardo C. (2022), Artificial intelligence and SMEs: How can B2B SMEs leverage AI platforms to integrate AI technologies?, *Industrial Marketing Management*, *107*, pp. 466-483. DOI: 10.1016/j.indmarman.2022.10.008.

Xiong H. (2024), Research on confusing responses based on ChatGPT, *Applied and Computational Engineering*, *57*(1), pp. 90-97. DOI. 10.54254/2755-2721/57/20241315.