INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research

International Journal of Innovation and Learning

ISSN online: 1741-8089 - ISSN print: 1471-8197
https://www.inderscience.com/ijil

Garbage in garbage out: Likert scale in management research

Borut Kodrič, Anita Trnavčevič, Tine Bertoncel, Nada Trunk Širca

DOI: 10.1504/IJIL.2025.10070455

Article History:

Received:	22 October 2024
Last revised:	15 December 2024
Accepted:	16 December 2024
Published online:	14 April 2025

Garbage in garbage out: Likert scale in management research

Borut Kodrič*, Anita Trnavčevič and Tine Bertoncel

Faculty of Management,
University of Primorska,
Koper, Slovenia
Email: borut.kodric@fm-kp.si
Email: anita.trnavcevic@fm-kp.si
Email: tine.bertoncel@fm-kp.si
*Corresponding author

Nada Trunk Širca

Faculty of Management,
University of Primorska,
Koper, Slovenia
and
International School for Social and Business Studies,
Celje, Slovenia
Email: nada.trunk@fm-kp.si

Abstract: Master's and PhD programs worldwide vary in the quality of their programs, learning opportunities, and outcomes, including those in management. Master's and PhD theses can be published in various forms, such as traditional theses, articles, or monographic publications. These academic works often incorporate data gathered through surveys, especially in the case of social sciences research, whether in quantitative studies or mixed methods research designs. Psychometric measurements and Likert scales are often employed in surveys. However, the proper use of Likert and other measurement scales is sometimes questionable due to inappropriate application. Management research holds dual values: one for researchers and another for practitioners, policymakers, consultants, and companies. Therefore, master's and PhD theses must contribute high-quality results, which are achieved through high-quality data collection instruments. To gain initial insight into the quality of questionnaires used in Master's and PhD theses in the field of management, this research was conducted.

Keywords: research; business education; learning; PhD thesis; quality of data; web-scraping; questionnaire; measurement scales; Likert scale; content validity.

Reference to this paper should be made as follows: Kodrič, B., Trnavčevič, A., Bertoncel, T. and Širca, N.T. (2025) 'Garbage in garbage out: Likert scale in management research', *Int. J. Innovation and Learning*, Vol. 37, No. 6, pp.1–21.

Biographical notes: Borut Kodrič received his PhD in the field of Statistics where he analysed the methodology of the total factor productivity index estimation. He is an Assistant Professor at the Faculty of Management University of Primorska, lecturing on business mathematics and statistics, research methodology in management and statistics in economics and finance. He has authored or co-authored various scientific papers and is currently involved in different research projects.

Anita Trnavčevič is graduated from the University of Toronto, Ontario Institute for Studies in Education. She has been working at the University of Primorska as a researcher and a Full Professor. Her research interests are research methodology, education management, and business models. She particularly values interdisciplinary research.

Tine Bertoncel received his PhD in Management, and studying Industry 4.0. His research primarily focuses on the impact of technological advancements, particularly artificial intelligence, on industry, examining how these innovations affect daily work practices and shape the future job market. He is currently a Professor at the Faculty of Management University of Primorska, where he is lecturing on business informatics and e-business.

Nada Trunk Širca earned her PhD in Social Sciences from Manchester Metropolitan University, UK. She currently works as a researcher and higher education teacher at MFDPŠ in Celje and at the University of Primorska. She also actively collaborates with EMUNI University in Slovenia. Her areas of expertise primarily include educational management, with a focus on various aspects of higher education, such as management, quality assurance, recognition of formal and informal learning, lifelong learning, and the internationalisation and modernisation of higher education systems both within and outside the EHEA. Additionally, she specialises in research methodology in social sciences, particularly qualitative methods.

This paper is a revised and expanded version of a paper entitled 'Garbage in, garbage out – the appropriate use of scales?' presented at Management International Conference, Trento, Italy, June 2024.

1 Introduction

Powell and Snellman (2004) in their definition of the knowledge economy, emphasise the scientific component of this concept, encompassing the entire process from R&D labs to companies, production lines, and final customers. This highlights the close relationship between R&D and the intellectual capabilities central to the knowledge economy. Similarly, other authors (e.g., Choong and Leung, 2022; Wirba, 2022) emphasise the systematic application of knowledge to drive economic growth and

competitiveness in the knowledge-based economy. These perspectives thus reflect the important role of intellectual resources and systematic knowledge application in promoting innovation and remaining competitive in the global market.

Building on this understanding, from a knowledge economy perspective, one of the indicators and predictors of a country's economic and social development is the number of employees in the R&D sector. This assumption aligns with the European Commission's statement: "Investment in research, development, education, and skills is a key policy area for the EU as it is essential for economic growth and the development of a knowledge-based economy" (Eurostat, 2022). Master's and PhD programs are critical components of this strategy since the R&D sector requires employees with high-level expertise and specialised knowledge, which are often the outcomes of master's and PhD programs. These programs are designed to develop advanced research skills, with students engaging in original research and contributing new knowledge and innovations to their fields. Hence, education systems are expected to support R&D and scientific production by implementing quality PhD and master's programs and producing quality graduates.

Among the 18.8 million tertiary education students in the EU in 2022, 29.4% (5,509 thousand) were studying for master's degrees, and 3.6% (670 thousand) were studying for doctoral degrees (Eurostat, 2024). Germany had the highest total number of tertiary students among EU countries, with the largest number of students pursuing master's degrees (1,133 thousand) and, by far, the highest number of students pursuing doctoral degrees (200 thousand). This was more than twice the number of doctoral students in any other EU country. The highest share of tertiary students studying for doctoral degrees in 2022 was recorded in Luxembourg (13.3%), followed by Czechia (6.4%). The lowest shares of doctoral students among the total number of tertiary education students were observed in Poland (2.0%) and Italy (1.8%).

Eurostat (2024) considering the views of many commentators, predicts an increased demand for highly skilled people in the coming years. Driven by digital technology, jobs are becoming more flexible and complex, leading employers to seek staff capable of managing complex information, thinking autonomously, being creative, using resources efficiently, and communicating effectively.

Meanwhile, the number of full-time equivalent researchers in the EU increased by more than one-third (40%) between 2010 and 2020, rising from 1.34 to 1.89 million (Eurostat, 2022). This significant increase indicates the growing emphasis on the knowledge economy within EU societies, which in turn necessitates steady enrolment in master's and PhD programs. As of 2024, there are over 21,300 master's programs and 4,030 PhD programs available in Europe, according to Top Universities (n.d.). While these programs vary by discipline, field, and specific aims, they share common goals. One of the aims of both master's and doctoral programs worldwide is to equip students with the most up-to-date knowledge, enabling them to conduct quality research and contribute to the global treasury of knowledge. The question of 'what constitutes research excellence?' often revolves around indicators such as research funding, productivity (number of publications), and peer review evaluations (Nowatzky and Underwood, 1995).

Master's and PhD theses are forms of research and can be judged according to published indicators for quality research. These indicators are usually summarised under metrics, peer review evaluation, impact, and outcomes, or what is described and

elaborated under RQ+ (Ofir et al., n.d.). Additionally, the European Research Area Council (ERAC) has developed a composite indicator.

"The Adjusted Research Excellence Index is a composite indicator selected by the European Research Area Council (ERAC) as the headline measure to monitor country performance with respect to ERA Roadmap Priority 1, 'effective national research systems'" [European Commission, (2020), p.22], it is obtained through the aggregation of four indicators that characterise countries' effectiveness of research systems, in terms of scientific and technological research excellence (the 'production' of high-impact publications and valued patents), and the ability of institutes to attract outstanding research grants and participate in researcher exchanges to pave the way for future excellence and develop efficient research capacity. The four indicators are defined as [Caperna, (2020), p.2]:

- The share of top 10% most highly cited publications per total publications [HICIT].
- PCT patent applications per population [PCT].
- ERC grants per public R&D [ERC].
- Participation in Marie Skłodowska-Curie actions: the number of MSCA researchers hosted per national MSCA researcher) [MSCA] [Caperna, (2020), p.3].

Indicators of excellence in research are similar, so it would be assumed that students who complete high-quality Master's and PhD programs would also produce high-quality theses. However, Starkey et al. (2023) argue that such a relationship is not linear and is partially due to the dynamics and complexity between publishing and research. In their view, it is possible to complete a high-quality PhD program, produce a low-quality article, improve it through peer review, and publish it in a low-quality journal (regarding the impact factor). For various reasons, it can become 'popular' or known among the wider public, consequently building readership and citation numbers.

Also, the theses are not an aim per se. They are published in the form of articles, scientific monographs, or made publicly available via different repositories. As such, they become sources of inspiration for other researchers or useful tools for the adaptation or elaboration of data collection instruments.

Across the EU in 2022, more than a fifth (22.1%) of all tertiary education students were studying business, administration, or law, while the second most common field was engineering, manufacturing, and construction, accounting for 15.5% of all students (Eurostat, 2024). The former, particularly business and administration, present unique challenges in measurement. In the social sciences, many phenomena are latent by nature, meaning they are not directly observable and must be inferred from indirect measures. This contrasts with the natural sciences, where phenomena often involve physical measurements that can be directly observed and quantified.

Dealing with abstract and subjective constructs like attitudes, beliefs, and behaviours requires indirect measurement through surveys, questionnaires, and observational methods, which can introduce various sources of error and bias. Measurement tools in social sciences, such as Likert scales and other psychometric instruments, must therefore be carefully designed to ensure validity and reliability, addressing issues like aligning response formats with the constructs being measured and mitigating potential response biases.

Considering the often poorly designed questionnaires, we have observed excellent research articles and master's and PhD theses, particularly focusing on the use of questionnaires in surveys. Similar to several authors (e.g., Carifio and Perla, 2007; Dolničar, 2013; Rossiter, 2011) who have problematised the use of the Likert scale, we identified specific weaknesses related to the design of questionnaires and the use of scales, specifically the Likert scale and its alignment with constructs being measured.

In our article, we therefore focus on the design of data collection instruments (questionnaires), especially addressing the alignment of response formats with the constructs being measured. We developed the following research questions:

RQ1 What measurement scales and response formats are used in questionnaires in selected Master's and PhD theses?

RQ2 How frequently are Likert scales and Likert-type scales aligned with the constructs being measured?

2 'From theory to practice': enhancing questionnaire design

Surveys and questionnaires are widely used in research and everyday professional life for practical reasons, such as customer satisfaction measurement, political such as different opinion polls on political preferences, and as part of official national surveys.

Literature is rich on research methodology/methods (e.g., Aguinis, 2024; Cooper, 2018; Myers, 2025), 'how to create' a good survey (e.g., Wolf et al., 2016), questionnaires survey (e.g., Lietz, 2010; Aithal and Aithal, 2020; Fife-Schaw, 2020; Taherdoost, 2016; Prendergast and OMeara, 2022; Panda and Mohapatra, 2024), on measurement scales and variables – nominal, ordinal, interval, and ratio (e.g., Schwab, 2004; Tharenou et al., 2007; Cohen et al., 2017), on ensuring validity and reliability (e.g., Taherdoost, 2016; Roebianto et al., 2023; Asher et al., 2023), and other questionnaire-related topics. For example, Bartol et al. (2023) analysed survey scales from the point of construct validity, criterion validity, and measurement invariance. Draugalis et al. (2008) for example, provide best practices for survey research reports. However, there is little empirical research on the quality and alignment of constructs, questionnaire items, and response format. Hill et al. (2022) state that researchers must ask themselves if a survey is the best tool to measure variables of the researcher's interest. And their response is 'no' in many cases.

Despite available literature and variety in research methods and question options, many questionnaires use or claim to use the Likert scale. If PhD and master's students decide on quantitative research, they often apply survey methodology and design questionnaires based on or including the Likert responding format even if there are better options for data collection. Hill et al. (2022) specifically mention the quality of items and aligned response format and provide examples of different options for asking questions and developing items and scales with aligned response format.

It seems that PhD and master's students often struggle to avoid the misuse of the Likert response format, despite the inclusion of research methods and methodology courses in management, business administration, and marketing programs across Europe and beyond. Additionally, online platforms like Coursera offer relevant courses. Nevertheless, the Likert response format continues to be a popular choice among many students in these fields, often being applied incorrectly.

Table 1 Sample questions and recommended phrasing

Best practice	*Problematic example*	*Recommended improvement*
Write positively worded questions	How often are you unable to start class on time?	How often do you start class on time?
Use questions and item-specific response format	I enjoyed the lecture (response format: strongly disagree to strongly agree)	How much did you enjoy the lecture? (response format: not at all to a great amount)
Avoid double-barreled items	How effective was the lecture and hands-on instruction?	How effective was the lecture instruction? How effective was the hands-on instruction? Or, the item could be written at a higher level of abstraction: How effective was the residency instruction?
Choose an appropriate number of response items (accuracy)	Did you like the activity? Yes No	How much did you like the activity? Not at all A little A moderate amount Quite a bit A lot
Attend to formatting and layout	How satisfied were you with your residency training?	How satisfied were you with your residency training? Not at all satisfied Somewhat satisfied Moderately satisfied Quite satisfied Extremely satisfied
Organise the survey items intentionally	First question on the survey: how often do you take illicit drugs?	First question on the survey: What is your favourite extracurricular activity?

Source: Hill et al. (2022, p.3)

3 Pitfalls of the use of the Likert scale

The Likert scale, developed by Rensis Likert, has unique features: it produces composite scores composed of individual responses to multiple items [Warmbrod, (2014), p.31]. It measures attitudes, values, opinions, prejudices, judgments, and other psychological constructs. Typically, the scale is used when respondents need to express their level of agreement or disagreement with a statement or item.

Likert (1932) developed a scale consisting of batteries of statements that define and describe the content and meaning of the measured construct. These statements express beliefs, preferences, judgments, or opinions. His work addressed the challenge of measuring character and personality traits by providing a procedure to quantify these qualities for data analysis.

The original Likert (1932) scale used a series of questions (items) with five response alternatives: strongly approve (1), approve (2), undecided (3), disapprove (4), and strongly disapprove (5). The construct's quantification is based on a summated

(composite) score for each respondent, calculated by summing their responses to each item on the scale. While Likert used a five-point responding format, other variations are appropriate, including the omission of the neutral response (Clason and Dormody, 1994). However, as other authors (Edwards, 1957; Oppenheim, 1992; Spector, 1992) have warned, generating and wording individual statements is crucial in designing a true Likert scale to ensure the instrument yields valid and reliable summated scores.

A series of verbal statements are supposed to express a range of positive expressions, views, sentiments, claims, or opinions about the attitude object or underlying construct, ranging from mildly positive to strongly positive, and similarly for negative statements (e.g., 1 = strongly disagree, 2 = disagree, 3 = undecided or neutral, 4 = agree, 5 = strongly agree). A respondent who is positive about the attitude object should agree with the positive statements and disagree with the negative ones. If the underlying construct is believed to be multi-dimensional, 6 to 8 items balanced by positive and negative statements need to be defined for each sub-dimension (Likert, 1932).

The use of the Likert scale has been 'simplified'. It is often designed and analysed as a single item per single item. Carifio and Perla (2007) provide extensive arguments for why the Likert scale is interval and how item-per-item analysis is incorrect. They also point to common misunderstandings about the terms, such as 'scale' versus 'response format' and the sloppy use of language. Their argumentation against persistent myths and urban legends about the Likert scale is a response to Jamieson (2005) and her view on (mis)use of the Likert scale. In their view, the Likert scale's purpose and use are defined. This interesting argumentation from both articles has somehow continued up to recent times. Jamieson (2005) have identified significant advancements in Likert scale development over 25 years, covering topics like construct validity, readability tests, and alternative measures of precision. They inform psychological researchers about the progress in Likert scale creation.

Hence, it can be argued that the Likert scale cannot measure other social phenomena and items that have become increasingly popular in research and applied use, such as a company's customer satisfaction. Instead of using the Likert scale to measure constructs like satisfaction, perception of effectiveness, and perception of importance, other scales need to be used. In such cases, Likert-type questions are more appropriate due to the alignment of the response format: strongly dissatisfied – dissatisfied – neutral – satisfied – strongly satisfied. Clason and Dormody (1994) identified Likert-type items as single questions (rather than a series of related statements) that use some aspect of the original Likert response alternatives. While multiple questions may be included in a research instrument, researchers do not attempt to combine the responses from these items into a composite scale. While the traditional Likert scale uses an agree-disagree continuum, Likert-type questions can employ various wordings to measure different dimensions, like frequency, importance, satisfaction, and likelihood. These variations allow researchers to tailor the questions to the specific context and type of information, such as employee efficacy, job or customer satisfaction, and self-perceptions of knowledge and competence.

Even in such cases, it is important to consider what the item is about and whether this type of question is appropriate. Other ordinal scales, asking about quantity, extent, and similar aspects, are available to researchers, PhD candidates, and master's students.

An extensive debate on psychometrics in marketing can be found in current literature. Rossiter (2011) provides a critique of the overuse of the Likert scale and psychometric measures. He proposes the C-OAR-SE method, which he considers a

'revolution' in marketing measurement. Rossiter (2011) states: "A Churchill-inspired researcher would likely borrow or invent loose multiple items representing a vague and usually undefined 'domain', put them in a questionnaire with faulty Likert answer scales, show that after deleting some items the scores on the remaining items correlate and produce a 'high alpha', and then claim to have captured the essence of 'coolness'! This is exactly what Churchill's approach would tell the researcher to do – and the researcher is much more likely to have the work published by following it" (1585). Rossiter has initiated an extensive debate, in which several authors participated, including Rigdon et al. (2011) and Lee and Cadogan (2016). His contribution is worth considering because it aligns with practical observations, and Dolničar (2013) conducted a similar study to the one presented in this article. She raised questions about the quality of questions and measures in surveys. She assessed 78 survey studies published in respected journals in the field of tourism, analysing the survey questions and response formats. She concluded that high-quality measures require a clear definition of what is being measured.

Likert scales and Likert-type scales are widely used. Warmbrod (2014, p.1) reports that "forty-nine percent of the 706 articles published in the *Journal of Agricultural Education* from 1995 to 2012 reported quantitative research with at least one variable measured by a Likert-type scale."

Asking the right, proper, and meaningful questions and using appropriate scales and response formats leads to reliable results and consequently meaningful recommendations for policymakers, companies, stakeholders, and researchers in management. An initial step in any PhD or master's study is learning about asking correct questions and using appropriate and correct scales and corresponding response formats.

This study of PhD and Master theses' survey questions has a limited scope - the use of poorly or wrongly formulated queries and associated inappropriate return options, in Dolničar (2013) words, or the use of 'faulty Likert answer scales', in Rossiter (2011) words.

4 Methodology

Within the introductory section, we pointed out the problematic nature of research within social science, particularly in business and administration, which present unique challenges in measurement. Given the latent nature of most phenomena being investigated, the use of indirect measures through surveying is usually the case.

Given these challenges, we found it crucial to focus specifically on PhD and master's theses within the management discipline. Management research is inherently multidisciplinary and transdisciplinary, encompassing various aspects of business administration, such as organisational behaviour, strategy, marketing, and finance. This focus allows for a deeper examination of how advanced students, who are expected to contribute original research, address the use of measurement scales and corresponding response formats, which are critical for accurately capturing these latent phenomena.

In addition, management research holds dual value: one for researchers and another for practitioners, policymakers, consultants, and companies. Therefore, master's and PhD theses within the management discipline must contribute high-quality results, which are achieved through high-quality data collection instruments.

4.1 Data collection and retrieval

The empirical part of the article is based on the analysis of secondary sources, specifically PhD and master's theses. The Open Access Theses and Dissertations (OATD) database (available at: https://oatd.org/) was utilised as the primary source for obtaining relevant dissertations. A comprehensive search was performed on the OATD website using the following keywords to retrieve dissertations written in English regardless the origin:

1 business administration
2 business management
3 corporate management
4 education management
5 hospitality management
6 human resource management
7 innovation management
8 knowledge management
9 marketing management
10 middle management
11 operations management
12 performance management
13 project management
14 public management
15 quality management
16 talent management.

These keywords were chosen to encompass the various facets of management previously mentioned, ensuring a diverse collection of dissertations for analysis.

Due to high traffic on the OATD website at the time of data collection, which hindered the efficiency of web scraping tools, an alternative approach was adopted. The number of search results displayed per page was maximised, and each HTML file containing the search results was manually downloaded and processed using Python to extract essential information. The following Python libraries were utilised:

- `os` for interacting with the operating system.
- `re` for regular expression operations.
- `bs4` (BeautifulSoup) and `lxml` for parsing HTML content.

The extracted information included:

- *Title of dissertation:* Identifying the research focus.

- *Abstract:* Summarising the study's content.
- *Link to PDF source:* Directing to the hosting website containing the dissertation PDF.
- *Date of publication:* Recording the time of the research.
- *Keywords:* Providing thematic context.
- *Unique identifier:* Assigning a distinct reference for each dissertation for subsequent retrieval and analysis.

To efficiently download the dissertation PDFs from the extracted links, a function was developed using the following Python libraries:

- `selenium` for browser automation.
- `pandas` for data manipulation.
- `urllib.parse` for URL handling.
- `concurrent.futures` for asynchronous execution.
- `time` for handling timing operations.

The function operates as follows:

1 *Initialisation:* Sets up a web browser using Selenium with Chrome options configured for automated downloading.
2 *URL handling:* Reads a list of dissertation URLs from a CSV file.
3 *Concurrent downloading:* Utilises `concurrent.futures` to download multiple PDFs simultaneously.
4 *Validation:* Checks each downloaded file to ensure it is a PDF.
5 *Storage:* Saves valid PDFs to a designated folder for analysis.

The downloaded PDFs were converted into text format using PyPDF and pandas to facilitate content analysis. Using pandas and regular expressions (re), the table of contents of each dissertation was parsed to identify the page numbers where questionnaires were located, then retrieved the questionnaire. The steps included:

1 *Parsing table of contents:* Extracting sections related to questionnaires or appendices.
2 *Page retrieval:* Isolating the pages containing the questionnaires.
3 *Automatic retrieval of questionnaires:* Each questionnaire was extracted automatically.
4 *Quality assessment:* Evaluating the retrieved questionnaires based on predefined criteria.

Information regarding the academic level of each dissertation (doctoral or master's thesis) was also extracted using pandas for classification purposes.

4.2 Sample

Based on the described procedure of data collection, 118 master's and PhD theses were retrieved. After an initial check of their appropriateness in terms of available questionnaires with closed-ended questions and measurement scales, 118 theses were included in the sample, with 60 being master's theses and 58 doctoral theses.

Table 2 Sample presentation

Thesis type	*No. of thesis*
Master thesis	60
Doctoral thesis	58
Total	*118*

To align with the research objectives, the survey questionnaires within sample theses were examined in greater detail, with a particular emphasis on the use of measurement scales, the correct formulation of corresponding questions and the use of response formats. The analysis was conducted through the following steps:

- Identifying the sections within each questionnaire, as questions are typically organised into sections with specific focuses.
- Identifying measurement scales and the response formats used within each section.
- Evaluating measurement scales and the response formats, especially addressing the alignment of response formats with the constructs being measured.

Results of the first step of analysis are presented in Table 3. Based on the specific topics or themes of the questions, we categorised them into the following five categories:

- Sections with *demographic questions:* These include questions about age, gender, education level, and other similar personal or organisational characteristics.
- Sections with questions aimed at assessing *the respondent's knowledge or awareness of certain facts:* These questions evaluate how much respondents know about specific topics or their awareness of certain issues.
- Sections focused on *the behaviour of individuals, groups, or practices within an organisation:* These questions explore actions, habits, or practices within a workplace or group setting.
- Sections centred on *people's attitudes, such as opinions or viewpoints:* These questions gather respondents' opinions, beliefs, or feelings about various subjects.
- Sections with other *focuses:* These include questions that measure satisfaction levels, preferences, or other miscellaneous topics.

A total of 560 sections were identified across the 118 evaluated questionnaires. Demographic questions were absent in 4 of the questionnaires used in master's theses and in 10 of those used in doctoral theses. Excluding demographic sections, the master's theses questionnaires had an average of 3.3 sections each, while the doctoral theses questionnaires had an average of 4.5 sections each.

5 Results and discussion

The analysis was conducted with the underlying assumption that there is a significant difference in the appropriate use of measurement scales between master's and PhD theses. This assumption stems from the belief that PhD theses are generally more scientific and rigorous than master's theses. Hence, the instruments of data collection, including questionnaires, questions, scales, and response formats, should be better aligned in PhD theses.

Table 3 Distribution of sections by questions' focus

Thesis type	*Section topic*					*Total*
	Demographics	*Knowledge*	*Practice*	*Attitudes*	*Other*	
Master thesis	56 (22.1%)	28 (11.1%)	68 (26.9%)	90 (35.6%)	11 (4.4%)	*253*
Doctoral thesis	48 (15.6%)	30 (9.8%)	116 (37.8%)	65 (21.2%)	48 (15.6%)	*307*
Total	*104 (18.6%)*	*58 (10.4%)*	*184 (32.9%)*	*155 (27.7%)*	*59 (10.5%)*	*560*

The results presented in Tables 3, 4, 5, 6, and 7 are thus broken down by both types of theses, where the differences in the use of measurement scales across both types of theses are analysed using contingency analysis and the corresponding chi-square test.

The distribution in Table 3 suggests that while both types of theses mostly value understanding behaviour, organisational practices and attitudes, doctoral theses place a somehow stronger emphasis on practices and behaviours, whereas master's theses are more inclined towards exploring attitudes (chi-square significance value < 0.001, contingency coefficient = 0,25).

The next step involved identifying the measurement scales and response formats used within each section. Considering how the questions were formulated, and which response formats were used, we categorised them into one of the following types:

- *Dichotomous questions:* These questions offer two possible responses, such as 'yes' or 'no'.
- *Multiple choice questions:* These questions provide several answer options, from which respondents select one or more.
- *Rating questions:* These questions ask respondents to rate an item on a scale (e.g., 1 to 5).
- *Likert scale questions:* These questions measure attitudes, character, and personality traits on a symmetric agree-disagree continuum, typically with five or seven points.
- *Likert-type questions:* Likert-type questions are not limited to agree-disagree formats but can employ various wordings to measure different dimensions.
- *Rank order questions:* These questions ask respondents to rank items in order of preference or importance.
- *Open-ended questions:* These questions allow respondents to answer in their own words, providing more detailed and qualitative data.

Table 4 Distribution of question types within sections valuing attitudes

Thesis type	*Question type*							*Total*
	Dicho-tomous	*Multiple choice*	*Rating*	*Likert scale*	*Likert-type scale*	*Rank order*	*Open-ended*	
Master thesis	54 (9.8%)	20 (3.5%)	2 (0.4%)	244 (44.2%)	227 (41.1%)	1 (0.2%)	4 (0.7%)	*552 (100%)*
Doctoral thesis	32 (6.0%)	32 (6.0%)	16 (3.0%)	308 (58.2%)	135 (25.5%)	1 (0.2%)	5 (0.9%)	*529 (100%)*
Total	*86 (8.0%)*	*52 (4.8%)*	*18 (1.7%)*	*552 (51.1%)*	*362 (33.5%)*	*2 (0.2%)*	*9 (0.8%)*	*1.081 (100%)*

Notes: χ^2 = 49.732 (sig. = 0.000), contingency coefficient = 0.209.

Table 5 Distribution of question types within sections valuing knowledge and fact awareness

Thesis type	*Question type*							*Total*
	Dicho-tomous	*Multiple choice*	*Rating*	*Likert scale*	*Likert-type scale*	*Rank order*	*Open-ended*	
Master thesis	33 (32.7%)	11 (10.9%)	4 (4.0%)	37 (36.6%)	0 (0.0%)	2 (2.0%)	14 (13.9%)	*101 (100%)*
Doctoral thesis	37 (34.9%)	9 (8.5%)	3 (2.8%)	42 (39.6%)	8 (7.5%)	0 (0.0%)	7 (6.6%)	*106 (100%)*
Total	*70 (33.8%)*	*20 (9.7%)*	*7 (3.4%)*	*79 (38.2%)*	*8 (3.9%)*	*2 (1.0%)*	*21 (10.1%)*	*207 (100%)*

Notes: χ^2 = 13.108 (sig. = 0.000), contingency coefficient = 0.244.

Table 6 Distribution of question types within sections valuing practices and behaviours

Thesis type	*Question type*							*Total*
	Dicho-tomous	*Multiple choice*	*Rating*	*Likert scale*	*Likert-type scale*	*Rank order*	*Open-ended*	
Master thesis	89 (15.1%)	81 (13.7%)	8 (1.4%)	292 (49.5%)	88 (14.9%)	3 (0.5%)	29 (4.9%)	*590 (100.0%)*
Doctoral thesis	34 (4.2%)	29 (3.5%)	34 (4.2%)	389 (47.5%)	287 (35.0%)	11 (1.3%)	35 (4.3%)	*819 (100.0%)*
Total	*123 (8.7%)*	*110 (7.8%)*	*42 (3.0%)*	*681 (48.3%)*	*375 (26.6%)*	*14 (1.0%)*	*64 (4.5%)*	*1409 (100.0%)*

Notes: χ^2 = 156.745 (sig. = 0.000), contingency coefficient = 0.316.

Table 7 Distribution of question types within sections valuing miscellaneous topics

Thesis type	*Question type*							*Total*
	Dicho-tomous	*Multiple choice*	*Rating*	*Likert scale*	*Likert-type scale*	*Rank order*	*Open-ended*	
Master thesis	6 (9.2%)	12 (18.5%)	3 (4.6%)	25 (48.5%)	4 (6.2%)	0 (0.0%)	15 (23.1%)	*65 (100.0%)*
Doctoral thesis	2 (0.4%)	12 (2.3%)	229 (43.4%)	50 (9.5%)	39 (7.4%)	0 (0.0%)	196 (37.1%)	*528 (100.0%)*
Total	*8 (1.3%)*	*24 (4.0%)*	*232 (39.1%)*	*75 (12.6%)*	*43 (7.3%)*	*0 (0.0%)*	*211 (35.6%)*	*593 (100.0%)*

Notes: χ^2 = 135.104 (sig. = 0.000), contingency coefficient = 0.431.

Table 8 Examples of inappropriate use of question types and response formats

Questionnaire theme	*Section topic*	*Question/item*	*Response format*
Organisational culture (doctoral thesis)	Operational internal audit and internal control	The planning and procedures of the operational internal audit function conform with the professional standards	Strongly agree to strongly disagree (7-point Likert-type scale)
		The risks the business faces, and operational internal audit procedures are reviewed regularly to ensure that the organisation does not face unidentified risks	Strongly agree to strongly disagree (7-point Likert-type scale)
Internal innovation, (product and process innovation) and firm performance (Master thesis)	Product innovation	Trial and error procedures are supported by our company	Five-point response format without any description of response items
Business ethics (Doctoral thesis)	Ethical interventions	Please indicate the extent to which your organisation has adopted the code of ethics	'Five-point': • Unknown • Known but considered unimportant • Known but considered important • Implementation is planned • Implemented

Results of the question categorisation are presented in Tables 4 to 7, with each table showing the distribution of question types within each section type (demographic sections were excluded from our analysis). When comparing the distributions of question types between master's and doctoral theses, significant differences were found in all sections except those valuing knowledge of facts awareness. The contingency coefficient ranged from 0.209 for sections valuing attitudes to 0.431 for sections valuing miscellaneous topics.

The distribution in Table 4 suggests that the Likert scale is the most widely used question type for valuing attitudes in both master's and doctoral theses, what is not surprising. Interestingly, the Likert-type question format is also quite often used within master's theses, when attitudes are the focus.

The two most widely used question types for valuing knowledge and facts awareness (Table 5) are the Likert scale (36.6% in master's theses and 39.6% in doctoral theses) and dichotomous questions (32.7% in master's theses and 34.9% in doctoral theses).

The distribution in Table 6 suggests that the Likert scale is the most widely used question type for valuing practices and behaviours in both master's and doctoral theses, as well, followed by the Likert-type question format within doctoral theses.

In sections valuing miscellaneous topics, the Likert scale is again the most widely used question type in master's theses, whereas the rating question type is prevalent when valuing miscellaneous topics in doctoral theses (Table 7).

Before continuing with the evaluation of how well the measurement scales and identified response formats align with the construct being measured, it is evident that the Likert scale is the prevalent question type. In the following paragraphs, however, we aim to answer the question of whether such frequent use of the Likert scale, regardless of the section focus, is appropriate or not.

As noted in Section 3, several authors have problematised the use of the Likert scale, with the most notable inaccuracies summarised by Carifio and Perla (2007). They highlighted "a problem of extremely poor and careless scale (test, questionnaire, interview, protocol, and so on) construction today, and even poorer and more careless data analysis than it is a problem of any inherent, conceptual, untestable, or uncorrectable problem with the Likert response format itself" [Carifio and Perla, (2007), p.115]. Addressing several persistent misconceptions, misunderstandings, and factual and empirical errors, myths, and untruths about Likert scales and their characteristics and properties, they stressed the importance of considering complexities and sophisticated nuances whenever any given measurement scale (or response format) is used, developed, or analysed.

Considering the conceptual, theoretical and empirical baseline already outlined, we evaluated the use of measurement scales and identified response formats presented in Tables 4 to 7, paying special attention to the following parameters:

1 Whether a question is organised as a single item question (each question within a section stands alone and is answered independently), matrix question (multiple questions are presented in a grid format sharing the same response format) or batteries of items (a series of related questions grouped together to measure a specific construct or topic).

2 Number of levels/points in case of scale response format (e.g., five-point, six-point, seven-point measurement scale).

3 The type of measurement scale (nominal, ordinal, interval or ratio).

4 Whether the question format is aligned with the construct or concept being measured within the section.

While the use of dichotomous, multiple choice, rating, rank order, and open-ended question types has been mostly without issues, diverse problems have been found with the use of Likert scale and Likert-type response formats. The numbers in Tables 9 and 10 primarily reflect cases of inappropriate use of these latter question types. However, before outlining the results, some examples of the most evident issues related to these two question formats are presented in Table 8.

In the first example given in Table 8 (question item within the section operational internal audit and internal control), ambiguity arises because the respondent's attitude towards the actual practice (i.e., whether procedures conform to standards) can be interpreted in two ways: it may reflect their opinion on the practice itself or indicate the extent to which the practice is present in the selected organisation. In the latter case, which the authors probably had in mind, the issue is the inappropriate use of a

scale-type response format. Procedures can either conform or not conform; there is no middle ground (e.g., 'somewhat agree').

Table 9 Evaluation of the appropriateness of questionnaire items used in master theses

Section topic	*Total no. of questions*	*Total no. of items*	*No. of inappropriate items*	*% of inappropriate items*
Knowledge/facts	101	123	54	43.9%
Practice/behaviour	590	680	403	59.3%
Attitudes	552	793	113	14.2%
Other	65	96	45	46.9%
Total	*1,308*	*1,692*	*615*	*36.3%*

Table 10 Evaluation of the appropriateness of questionnaire items used in doctoral theses

Section topic	*Total no. of questions*	*Total no. of items*	*No. of inappropriate items*	*% of inappropriate items*
Knowledge/facts	106	185	43	23.2%
Practice/behaviour	819	1309	692	52.9%
Attitudes	529	784	59	7.5%
Other	528	594	266	44.8%
Total	*1,982*	*2,872*	*1,060*	*36.9%*

Our proposal is to list each procedure within the business area (operational internal audit and internal control) that standards refer to and use the 'yes/no' conformity checkbox for each procedure (dichotomous question type). This approach would unambiguously reflect not only how many procedures conform or not, but also which ones do.

Similar issues can be outlined in relation to the second example, where ambiguity again arises because the respondents' attitude towards the actual practice (i.e., regular review to prevent risk) can be interpreted in two ways: it may reflect their support for the practice itself or indicate the extent to which the practice is present in the selected organisation. In the latter case, which the authors probably had in mind, the use of a scale-type response format is inappropriate. When respondents strongly agree with the statement, it can be intuitively interpreted as an activity regularly in place, which is conditionally acceptable. However, how should non-agreement with the statement be interpreted? It could reflect the activity being rarely performed or not performed at all. This ambiguity persists.

Our recommendation is to reformulate the question and use a response format that reflects the frequency of practicing a selected activity (never/rarely/sometimes/often/always), which would unambiguously reflect organisational practices.

We are outlining the third example (product innovation procedures) as an incorrect implementation of what is most likely a Likert-type scale, where the author provided a five-point scale with digits from 1 to 5 as the only response item descriptions. Providing clear explanations for each response item is important to ensure that respondents understand what each option represents, reducing ambiguity and improving the accuracy of the data collected.

In the last example outlined in Table 8, the issue again lies in the inappropriate question response format. For a question like "Please indicate the extent to which your

organisation has adopted the following ethical interventions", a Likert-type scale (not at all/to a small extent/to a moderate extent/to a great extent/to a very great extent) would be appropriate. This scale allows respondents to clearly indicate the degree to which their organisation has adopted the interventions, helps to reduce ambiguity, and provides a nuanced understanding of their responses. In our opinion, this response format should be used almost always when questions, correctly formulated, refer to the presence of certain organisational practices or individuals' behaviours.

Tables 9 and 10 summarise the results of our evaluation of question formulations and the use of response formats in both types of theses. Overall, we consider more than a third of the question item formulations and/or corresponding response formats to be inappropriate in the theses evaluated. Questionnaire items aimed at measuring the behaviour of individuals, groups, or practices within an organisation were the most problematic, with 59.3% of items in corresponding topics in master's thesis surveys and 52.9% of items in corresponding topics in doctoral theses being inappropriately designed. The second most problematic sections were those aimed at measuring satisfaction levels, preferences, or other miscellaneous topics (labelled as 'others' in Tables 9 and 10), where 46.9% of items in the corresponding sections of master's theses and 44.8% of items in the corresponding sections of doctoral theses were inappropriately designed.

5.1 Conclusions

While Carifio and Perla (2007) mainly focused on the incorrect analysis of data collected using Likert scales, our article addresses the misuse of Likert scales for evaluating topics or phenomena they were not designed for. According to the definition of content validity in C-OAR-SE theory [Rossiter, (2011), p.1568], we are actually dealing with answer-scale validity. This means ensuring that the measure is free from distortions caused by semantic confusion when respondents answer the questions. The issues we highlight indicate poor content validity in many measures used in management research, similar to the problems (Rossiter, 2011) pointed out in marketing research.

By using automated data retrieval and processing techniques, we efficiently gathered a large number of theses from various management disciplines. This method allowed for a systematic assessment of questionnaire quality in subsequent steps, providing valuable insights into academic research methodology.

We found that many questionnaires were poorly designed, particularly in their use of Likert scales and other measurement tools. While the management field emphasises psychometric measurement and frequently uses Likert scales, there are significant issues. As Carifio and Perla (2007) pointed out, the Likert response format should be tied to a single question item that fits the logical requirements and characteristics of a purposefully constructed Likert scale. Unfortunately, more than half of the theses we analysed lacked this critical understanding and proper use of Likert scales. Focusing on data collection instruments is just as important as choosing statistical methods and designing the entire survey, including sampling procedures and sample selection. Many other data collection methods and instruments are undervalued, even though they might offer better options for generating valid questions. If we assume that companies and policymakers use research results, our findings indicate a serious problem. The academic perspective is equally concerning, as our results call into question much of the knowledge in the social sciences. This situation underscores the urgency of Rossiter

(2011) call to rethink our entire approach to measurement. Without addressing these issues, both practical applications and academic integrity are at risk.

The results have some limitations, including:

a the sample size and the variety of included institutions, theses, and curricula underlying learning opportunities

b the narrow focus, as we examined only the alignment of response format with the construct being measured

c the lack of generalisability.

This narrow focus might inspire other researchers to expand future research on content validity and the statistical methods used to analyse survey data in PhD and master's theses.

Despite these limitations, there are some implications for master's and PhD supervisors and university policymakers. This limited research calls for a discussion about the emphasis and content of research methodology courses. It also signals to editors of scientific journals the importance of requesting the provision of instruments of data collection for submitted research articles.

References

Aguinis, H. (2024) *Research Methodology: Best Practices for Rigorous, Credible, and Impactful Research*, Sage, George Washington University.

Aithal, A. and Aithal, P.S. (2020) 'Development and validation of survey questionnaire & experimental data – a systematical review-based statistical approach', *International Journal of Management, Technology, and Social Sciences*, Vol. 5, No. 2, pp.233–251.

Asher, A., Briney, K. and Goben, A. (2023) 'Valid questions: the development and evaluation of a new library learning analytics survey', *Performance Measurement and Metrics*, Vol. 24, No. 2, pp.101–119, DOI: 10.1108/PMM-04-2023-0009.

Bampasidou, M. and McKinzie, W. (2020) 'Decision-making: how involved are our producers and hired managers?', *Journal of ASFMRA*, pp.81–87, https://www.jstor.org/stable/27033926.

Bartol, J., Vehovar, V. and Petrovčič, A. (2023) 'Systematic review of survey scales measuring information privacy concerns on social network sites', *Telematics and Informatics*, Vol. 85, November, DOI: 10.1016/j.tele.2023.102063.

Caperna, G. (2020) *The Adjusted Research Excellence Index 2020, Methodology Report*, Publication Office of the EU, Luxembourg.

Carifio, J. and Perla, R.J. (2007) 'Ten common misunderstandings, misconceptions, persistent myths and urban legends about likert scales and likert response formats and their antidotes', *Journal of Social Sciences*, Vol. 3, pp.106–116, DOI: 10.3844/jssp.2007.106.116.

Choong, K.K. and Leung, P.W. (2022) 'A critical review of the precursors of the knowledge economy and their contemporary research: implications for the computerized new economy', *Journal of Knowledge Economy*, Vol. 13, pp.1573–1610, DOI: 10.1007/s13132-021-00734-9.

Clason, D.L. and Dormody, T.J. (1994) 'Analyzing data measured by individual Likert-type items', *Journal of Agricultural Education*, Vol. 35, No. 4, pp.31–35.

Cohen, L., Manion, L. and Morrison, K. (2017) *Research Methods in Education*, 8th ed., Routledge, ISBN: 9781138209886.

Cooper, D.R. (2018) *Business Research: A Guide to Planning, Conducting, and Reporting Your Study*, 1st ed., Sage Publications Ltd.

Dolničar, S. (2013) 'Asking good survey questions', *Journal of Travel Research*, Vol. 52, No. 5, pp.551–574, DOI: 10.1177/0047287513479842.

Draugalis, J.R., Coons, S.J. and Plaza, C.M. (2008) 'Best practices for survey research reports: a synopsis for authors and reviewers', *American Journal of Pharmaceutical Education*, Vol. 72, No. 1, p.11.

Edwards, A.L. (1957) *The Social Desirability Variable in Personality Assessment and Research*, Dryden Press, New York.

European Commission/EACEA/Eurydice (2024) *The European Higher Education Area in 2024: Bologna Process Implementation Report*, Publications Office of the European Union, Luxembourg [online] https://eurydice.eacea.ec.europa.eu/sites/default/files/2024-05/Chapter_1_European_Higher_Education_Area_key_data_0.pdf.

European Commission (2020). *Adjusted Research Excellence Index* [online] https://composite-indicators.jrc.ec.europa.eu/explorer/explorer/indices/arei/adjusted-research-excellence-index (accessed 28 September 2024).

Eurostat (2022) *Human Resources in Science and Technology* [online] https://ec.europa.eu/eurostat/statistics-explained/index.php?q=Human_resources_in_science_and_technology (accessed 11 August 2024).

Eurostat (2024) *Tertiary Education Statistics* [online] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics#Graduates (accessed 28 September 2024).

Farrukh, A. and Sajjad, A. (2023) 'A critical review of literature review methodologies', in Rana, S., Singh, J. and Kathuria, S. (Eds.): *Advancing Methodologies of Conducting Literature Review in Management Domain, Review of Management Literature*, Emerald Publishing Limited, Leeds, Vol. 2, pp.103–123, DOI: 10.1108/S2754-586520230000002006.

Fife-Schaw, C. (2020) 'Questionnaire design', *Research Methods in Psychology*, pp.343–374.

Gengler, A.M. (2020) 'Emotions and medical decision-making', *Social Psychology Quarterly*, Vol. 83, No. 2, pp.174–194, https://www.jstor.org/stable/48588971.

Hill, J., Ogle, K., Santen, S.A., Gottlieb, M. and Artino, A.R. (2022) 'Educator's blueprint: a how-to guide for survey design', *AEM Education and Training*, Vol. 6, No. 4, p.e10796, DOI: 10.1002/aet2.10796.

Jamieson, S. (2005) 'Likert scales: how to (ab) use them', *Medical Education*, Vol. 38, No. 12, pp.1217–1218, DOI: 10.1111/j.1365-2929.2004.02012.x.

Jebb, A.T., Ng, V. and Tay, L. (2021) 'A review of key likert scale development advances: 1995–2019', *Frontiers in Psychology*, Vol. 12, Article 637547, DOI: 10.3389/fpsyg.2021.637547.

Lee, N. and Cadogan, J.W. (2016) 'Welcome to the desert of the real: reality, realism, measurement, and C-OAR-SE', *European Journal of Marketing*, Vol. 50, No. 11, pp.1959–1968, DOI: 10.1108/EJM-10-2016-0549.

Lietz, P. (2010) 'Research into questionnaire design: a summary of the literature', *International Journal of Market Research*, Vol. 52, No. 2, pp.249–272.

Likert, R. (1932) 'A technique for the measurement of attitudes', *Archives of Psychology*, Vol. 22, No. 140, p.55.

Myers, M.D. (2025) *Qualitative Research in Business and Management*, 4th ed., Sage Publications Ltd., University of Auckland, New Zealand.

Nowatzky, R. and Underwood, D.G. (1995) 'Indicators of research quality', *Education Policy Analysis Archives*, Vol. 3, No. 20 [online] https://www.researchgate.net/publication/49609878_Indicators_of_Research_Quality (accessed 29 September 2024).

Ofir, Z., Schwandt, T., Duggan, C. and McLean, R. (n.d.) *Research Quality Plus: A Holistic Approach to Evaluating Research* [online] https://idl-bnc-idrc.dspacedirect.org/server/api/core/bitstreams/26d17914-e289-467a-b45e-2e139085156c/content.

Oppenheim, A.N. (1992) *Questionnaire Design, Interviewing, and Attitude Measurement*, Printer Publishers, New York, NY.

Panda, A. and Mohapatra, S. (2024) 'Research design', in *The Online Healthcare Community*, pp.53–62, Emerald Publishing Limited, Leeds, DOI: 10.1108/978-1-83549-140-920241004.

Paquette, E.T., Palac, H., Bair, E., Schultz, B., Stenquist, N., Joffe, S. and Shukla, A. (2020) 'The importance of engaging children in research decision-making: a preliminary mixed-methods study', *Ethics & Human Research*, Vol. 42, No. 3, pp.12–20, https://www.jstor.org/stable/26991473.

Powell, W.W. and Snellman, K. (2004) 'The knowledge economy', *Annual Review of Sociology*, Vol. 30, pp.199–220, https://www.jstor.org/stable/29737691.

Prendergast, M. and OḾeara, N. (2022) 'Questionnaire design and implementation', in *Perspectives in Contemporary STEM Education Research*, pp.120–128, Routledge, New York.

Rana, S., Singh, J. and Kathuria, S. (2023) *Advancing Methodologies of Conducting Literature Review in Management Domain, Review of Management Literature*, Emerald Publishing Limited, Leeds, Vol. 2, DOI: 10.1108/S2754-586520230000002006.

Rigdon, E.E., Preacher, K.J., Lee, N., Howell, R.D., Franke, G.R. and Borsboom, D. (2011) 'Avoiding measurement dogma: a response to Rossiter', *European Journal of Marketing*, Vol. 45, Nos. 11/12, DOI: 10.1108/03090561111167306.

Roebianto, A., Savitri, S.I., Aulia, I., Suciyana, A. and Mubarokah, L. (2023) 'Content validity: definition and procedure of content validation in psychological research', *TPM*, Vol. 30, No. 1, pp.5–18.

Rossiter, J.R. (2011) 'Marketing measurement revolution: the C-OAR-SE method and why it must replace psychometrics', *European Journal of Marketing*, Vol. 45, Nos. 11/12, pp.1561–1588, DOI: 10.1108/03090561111167298.

Saltos-Echeverría, T.K. et al. (2023) 'Market management change process from public to private: a case study in Cotacachi, Ecuador', *Theoretical and Empirical Researches in Urban Management*, Vol. 18, No. 3, pp.66–87, https://www.jstor.org/stable/27232315 (accessed 7 October 2024).

Schwab, D.P. (2004) *Research Methods for Organizational Studies*, 2nd ed., Lawrence Erlbaum, Mahwah, NJ.

Spector, P.E. (1992) *Summated Rating Scale Construction: An Introduction*, Sage, Newbury Park, CA.

Starkey, J.C. et al. (2023) 'Research quality: understanding definitions of and challenges to quality in the knowledge production process', in Koerber, A., Starkey, J.C., Ardon-Dryer, K., Cummins, R.G., Eko, L. and Kee, K.F. (Eds.): *The Predatory Paradox: Ethics, Politics, and Practices in Contemporary Scholarly Publishing*, Open Book Publishers, Cambridge, UK, DOI: 10.11647/OBP.0364.

Taherdoost, H. (2016) 'Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research', *International Journal of Academic Research in Management*, Vol. 5, No. 3, pp.28–36.

Taherdoost, H. (2022) 'Designing a questionnaire for a research paper: a comprehensive guide to design and develop an effective questionnaire', *Asian Journal of Managerial Science*, Vol. 11, No. 1, pp.8–16.

Tharenou, P., Donohue, R. and Cooper, B. (2007) *Management Research Methods*, Cambridge University Press, Cambridghe.

Top Universities (n.d.) *List of PhD Programs in Europe* [online] https://www.topuniversities.com/programs/europe/phd?region=[4010]&study_level=[5]&globalsearch=[doctoral] (accessed 28 September 2024).

Warmbrod, J.R. (2014) 'Reporting and interpreting scores derived from Likert-type scales', *Journal of Agricultural Education*, Vol. 55, No. 5, pp.30–47, DOI: 10.5032/jae.2014.05030.

Wirba, A.V. (2022) 'Transforming Cameroon into knowledge-based economy (KBE): the role of education, especially higher education', *Journal of the Knowledge Economy*, Vol. 13, No. 2, pp.1542–1572, DOI: 10.1007/s13132-021-00776-z.

Wolf, C., Joye, D., Smith, T.W. and Fu, Y. (2016) *The SAGE Handbook of Survey Methodology*, SAGE Publications Ltd., DOI: 10.4135/9781473957893.