SemeAd 2025
XXVIII Seminários em Administração

10, 11, 12, 13 e 14 de novembro de 2025
ISSN 2177-3866

Data Leadership Scale: Validation and Analysis of its Relationships with Project Success

MARIANA MOTTA FLECK
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS)

ANTONIO CARLOS GASTAUD MAÇADA
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS)

MAURICIUS MUNHOZ DE MEDEIROS
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL (UFRGS)

Agradecimento à orgão de fomento:
We thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the financial support in the development of this research.

DATA LEADERSHIP SCALE: VALIDATION AND ANALYSIS OF ITS RELATIONSHIPS WITH PROJECT SUCCESS

Introduction

In increasingly complex, data-driven organizational contexts, leadership has emerged as a central element for project success (Fleck & Maçada, 2025; Jansen, Vera, & Crossan, 2009; Mikalef Wetering, & Krogstie, 2021; Nisar *et al.*, 2020; NorenaChavez & Thalassinos, 2023; Shamim, Zeng, Shariq, & Khan, 2019; Zhao *et al.*, 2024). Moreover, decentralized data management structures have proven more effective than centralized models, fostering greater collaboration and organizational impact (Ahmed, Shaheen, & Philbin, 2022; Saif, 2020). Thus, how organizational resources are managed directly influences the effectiveness of data projects, and effective leaders play a central role in developing competencies and organizational practices (Nisar *et al.*, 2020; Shamim *et al.*, 2019).

However, the lack of analytical knowledge and capabilities among leaders remains a recurring challenge for organizations seeking data-driven decisions (Haude, Blohm, & Lagardère, 2024; Tabesh, Mousavidin, & Hasani, 2019). In response, many companies have invested in training programs focused on developing data leadership, as exemplified by the initiative at Lufthansa (Haude *et al.*, 2024).

A recent systematic review identified four main types of data leadership: visionary, decision-maker, driving, and educator (Fleck & Maçada, 2025; Haude *et al.*, 2024). Each type encompasses specific responsibilities with implications for data project success and organizational performance (Haude *et al.*, 2024; Schmidt, Dierendonck, & Weber, 2023). Although previous studies have addressed data leadership from a multidimensional perspective, the literature to date has not identified empirical efforts that quantitatively relate and measure these constructs in association with project success. This methodological gap hinders the advancement of quantitative empirical research on the topic and the formulation of evidence based organizational plans and policies. Furthermore, the absence of validated scales limits organizations' ability to diagnose leadership types, monitor their evolution, and understand their effective contribution to data-driven project outcomes.

Given this scenario, the objectives of this study are to develop and validate a scale for measuring data leadership types and to analyze its relationship with project success. It is expected that the instrument will contribute to theoretical advancement in the field of data management and leadership by providing an empirical basis for future research, and to organizational practice by offering a diagnostic tool applicable in leadership development, training, and strategic alignment processes in data-driven environments.

Scale validation is a crucial step to ensure the quality of measures used in academic and organizational research, with face and content validity being particularly important in the early development stages (Hinkin, 1998; Hudson, 2014). Content validity refers to the extent to which an instrument's items adequately represent the conceptual domain of the construct, while face validity concerns the clarity, comprehensibility, and perceived relevance of items by experts or respondents (Hinkin, 1998; Righi *et al.*, 2013). To ensure this adequacy, items should be evaluated by experts, allowing adjustments based on convergence and response analysis (Righi *et al.*, 2013). After expert validation, empirical data were collected for factor analysis to identify latent factors, explore the instrument's dimensionality, and verify whether items group according to the expected theoretical structure (Hair, 2019). Finally, structural equation modeling assessed the measurement model's quality and its relationships (Hair, Hult, Ringle, & Sarstedt, 2022).

This study is structured as follows: first, we present the main theoretical concepts; next, we detail the method; then, we describe the analyses performed; and finally, we discuss the results and present the conclusions, describing contributions, limitations, and suggestions for future research.

2 Theoretical Background

2.1 The relationship between Data Leadership and Project Success

The concept of project success has evolved significantly over the past decades. Initially tied to meeting scope, schedule, and budget – the so-called "Iron Triangle" – success came to embody a broader perspective, also considering strategic impact, sustainable value creation, and adaptation to organizational change (Serrador & Turner, 2015; Turner & Zolin, 2012). In highly complex initiatives, such as technology and data projects, factors like strategic alignment, qualified leadership, and adaptability have proven critical to achieving success (Müller & Turner, 2007; Shenhar & Holzmann, 2017).

The absence of adequate leadership is frequently cited as a cause of project failures, especially in innovative and dynamic contexts (Afzal, 2014; Shenhar *et al*., 2002). Conversely, leaders who can mobilize teams, make strategic decisions, and promote adaptive practices positively influence project outcomes, particularly when they tailor their management style to project characteristics such as innovation level, complexity, and urgency (Müller & Turner, 2007; Shenhar *et al*., 2020).

In this context, data leadership has emerged as a critical factor for project success. Traditionally associated with technical roles within information technology (IT), data leadership has assumed a strategic position, reflected in the growing prominence of the Chief Data Officer (CDO), whose responsibilities extend beyond technical management to encompass organizational transformation and governance (Lee *et al*., 2014). Effective data leaders are essential for articulating strategic visions, aligning data with organizational objectives, promoting data literacy, and fostering a culture oriented toward data-driven decision-making (Menukin, Mandungu, Shahgholian, & Mehandjiev, 2023; Schmidt *et al*., 2023). Moreover, data leaders play a central role in facilitating cross functional integration and driving dynamic organizational capabilities—such as opportunity sensing, resource mobilization, and continuous transformation—elements considered critical to the success of projects in complex, dynamic environments (Mikalef *et al*., 2021).

However, a significant knowledge gap regarding data persists among leaders, which can compromise project outcomes and organizational analytical maturity (Haude *et al*., 2024; Tabesh *et al*., 2019).

Below, we discuss the main characteristics of each leadership type and their potential relationships with organizational project success.

2.1.1 Visionary Data Leadership

Visionary data leadership is responsible for establishing a comprehensive strategic vision for data use, promoting its centrality within the organization, and inspiring analytical transformation (Shamim *et al*., 2019). This type of leadership acts as a driving force for innovation by articulating a data-based future vision capable of mobilizing resources and aligning organizational efforts (Fernandes *et al*., 2022; Jansen *et al*., 2009). Visionary leaders play a central role in formulating organizational strategy, particularly by anticipating trends and aligning data with the company's competitive positioning (Shao, 2019). Evidence suggests that organizations operating under a clear data vision can optimize resources, reduce risks, and increase return on investment (Brocchi *et al.*, 2018). Conversely, the absence of such a vision significantly reduces the team's effectiveness in delivering value (Gartner, 2021). Thus, visionary data leadership constitutes an essential condition for steering projects toward success.

2.1.2 Decision-Maker Data Leadership

Decision-maker data leadership is characterized by the ability to demonstrate the value of data for organizational decision-making, fostering the democratization of information access, responsible data use, and the establishment of robust governance (Zhang, Sun, & Zhang, 2022). This data leadership type contributes to legitimizing analytical initiatives by aligning stakeholders with strategic priorities and demonstrating tangible data use value, while also structuring governance mechanisms that ensure reliable access, ethical use, and cross functional coordination throughout the analytical transformation journey (Brocchi *et al.*, 2018; Ferreira, Merendino, & Meadows, 2023). Engaged data governance leadership has proven fundamental to avoiding project investment failures, ensuring regulatory compliance, and safeguarding decision integrity (Abraham, Schneider, & vom Brocke, 2019; Janssen *et al.*, 2020). Evidence indicates that the lack of effective governance jeopardizes up to 80% of data Investments (Gartner, 2024). Therefore, decision-maker data leadership plays a decisive role in legitimizing and sustaining data-driven projects.

2.1.3 Driving Data Leadership

Driving data leadership operationalizes the data strategy by connecting systems, repositories, and personnel, ensuring that applications integrate with organizational routines (Tabesh *et al.*, 2019; Zhang *et al.*, 2022). This leadership is essential to prevent siloed structures and guarantee that data generates real business value (Brocchi *et al.*, 2018; Saif, 2020). Effective performance of this leadership type depends on coordinating tools, data architecture, and human resources, enabling the implementation of use cases that have a direct impact on organizational performance (Tabesh *et al.*, 2019). Driving leaders also help create competitive advantage by aligning data with strategic objectives (Fernandes *et al.*, 2022; Norena-Chavez & Thalassinos, 2023). Thus, driving data leadership is central to project execution and value creation.

2.1.4 Educator Data Leadership

Educator data leadership is responsible for promoting analytical literacy, acting as a facilitator of data-driven organizational learning (Sharma, Mithas, & Kankanhalli, 2014). Its primary function is to increase employees' knowledge and confidence in data use, fostering knowledge sharing across units and continuous skill development (Kadarsah, Govindaraju, & Prihartono, 2023). To that end, effective data educators encourage hands-on learning and disseminate responsible data practices at all organizational levels (Tabesh *et al.*, 2019). Low confidence in data use remains a persistent challenge in many organizations (Qlik & Accenture, 2022), while evidence shows that only a small portion of analytical projects produce actionable *insights* (Deloitte, 2020). Conversely, investments in data literacy and analytical culture can increase an organization's market value by up to 5% (Qlik, 2018). Therefore, educator data leadership strengthens the data culture and raises organizational analytical maturity, which translates into greater project success.

2.2 Definition of Constructs and Research Model

Recent studies indicate that data project success is associated with the actions of different leadership types, which vary in focus, scope, and role within the data value chain (Mikalef *et al.*, 2021; Nisar *et al.*, 2020; Zhao *et al.*, 2024). This article adopts the data leadership types proposed by Haude *et al.* (2024) and identified in a systematic literature review (Fleck & Maçada, 2025), classifying data leadership into four primary types: Visionary, Decision-Maker, Driver, and Educator. **Table 1** summarizes the constructs under investigation.

Table 1. **Constructs and definitions.**

Construct	Definition	References
Visionary Data Leader	Ability to establish a strategic and inspiring vision for data use, promoting innovation and organizational transformation.	Fernandes *et al.* (2022); Haude *et al.* (2024); Jansen *et al.* (2009); Mikalef *et al.* (2021); Shamim *et al.* (2019); Shao (2019).
Decision Maker Data Leader	Ability to demonstrate the value of data in decision-making, fostering information democratization and implementing data governance.	Abraham *et al.* (2019); Brocchi *et al.* (2018); Ferreira *et al.* (2023); Haude *et al.* (2024); Janssen *et al.* (2020); Saif (2020); Schmidt *et al.* (2023); Zhang *et al.* (2022).
Driving Data Leader	Responsible for operationalizing the data strategy by integrating resources, tools, and people to generate value in projects.	Brocchi *et al.* (2018); Fernandes *et al.* (2022); Haude *et al.* (2024); Mikalef *et al.* (2021); Saif (2020); Tabesh *et al.* (2019); Zhang *et al.* (2022).
Educator Data Leader	Focused on developing an analytical culture and empowering employees through data literacy and organizational learning.	Ferraris *et al.* (2018); Haude *et al.* (2024); Kadarsah *et al.* (2023); Mikalef *et al.* (2021); Tabesh *et al.* (2019).
Project Success	Achievement of project objectives considering scope, time, and cost, as well as strategic impact, sustainable value, and adaptation to complexity.	Shenhar *et al.* (2020); Shenhar & Holzmann (2017); Shenhar (2007); Turner & Zolin (2012).

The theoretical proposition is that the four dimensions of data leadership are related to project success. **Figure 1** presents the research model with its dimensions and indicators.

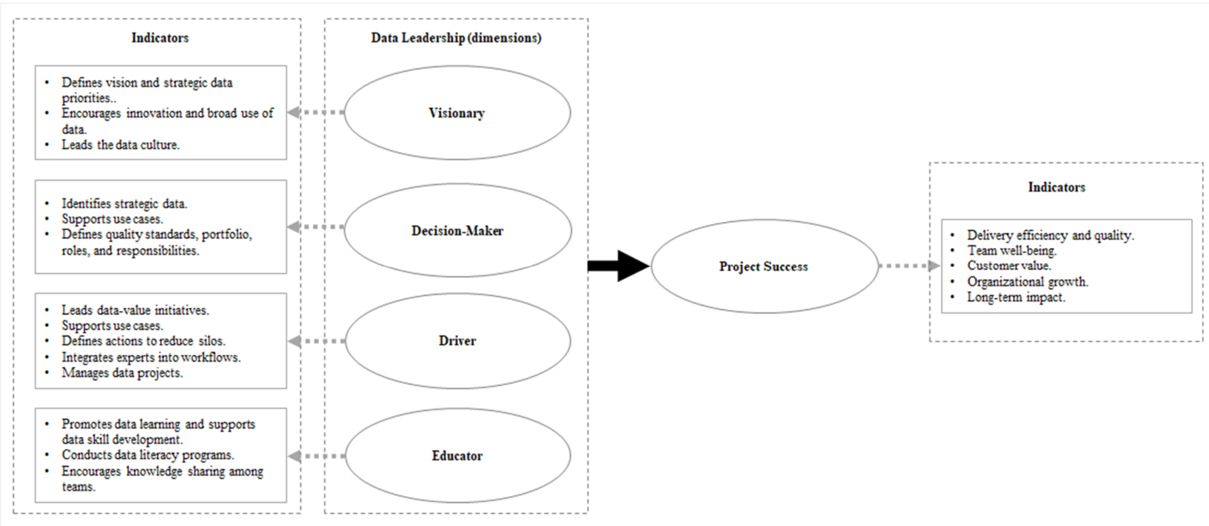


Figure 1. Research model.

Consequently, the conceptual discussion and the developed research model support the following hypotheses:

> ***H1.** Data leadership is a multifaceted factor reflected in four dimensions (visionary, decision-maker, driving, and educator).*
>
> ***H2.** Each type of data leadership - visionary (H2a), decision-maker (H2b), driving (H2c), and educator (H2d) - is positively related to the success of data and analytics projects.*

Although prior studies (Fleck & Maçada, 2025; Haude *et al*., 2024; Schmidt *et al*., 2023) have addressed data leadership from a multidimensional perspective, the literature reviewed to date has not identified empirical efforts that quantitatively relate and measure these constructs in association with project success, which justifies the innovative nature, originality, and theoretical relevance of this study.

3 Method

3.1 Instrument Development

The scale-construction process followed the guidelines proposed by MacKenzie *et al*. MacKenzie *et al*. (2011), and was conducted in multiple stages, emphasizing theoretical validity and methodological rigor. First, the constructs related to data leadership types were defined based on a recent systematic literature review (Fleck & Maçada, 2025). From this theoretical foundation, corresponding items were generated by adapting the roles and responsibilities described by Haude *et al*. (2024) and integrating empirical evidence on leadership's effect on project success (Müller & Turner, 2007; Shamim *et al*., 2019; Shenhar *et al*., 2020; Tabesh *et al*., 2019; Zhao *et al*., 2024).

The initial instrument version contained 20 items across four constructs: Visionary Data Leader, Decision-Maker Data Leader, Driving Data Leader, and Educator Data Leader. To ensure face and content validity, a closed-format card-sorting exercise was conducted with seven experts in information management, strategy, and data analysis, following Faria (2010), and Wood & Wood (2008). Experts were asked to group items according to the predefined constructs. The results were analyzed via a similarity matrix to assess convergence among experts.

Next, items were evaluated for clarity, relevance, and representativeness using a 4-point ordinal scale (1 = inadequate item to 4 = highly adequate item), as recommended by Polit & Beck (2006). The Content Validity Index (CVI) was calculated based on the proportion of experts rating each item ≥ 3. Following Adrian *et al*. (2019), items with a CVI ≥ 0.80 were deemed satisfactory. As a result of this validation stage, wording adjustments and item additions/removals reduced the total number of items to 19. As a result of the face-validation process, wording adjustments, additions, and deletions were made, and the total number of items for the four data-leadership constructs was reduced to 19.

3.2 Expert Evaluation

Content and face validity of the instrument were assessed by a panel of experts selected based on their academic and professional experience in information management, strategy, and data analysis. Expert selection followed the recommendation of Polit and Beck (2006), who suggest involving at least three evaluators. The authors also note that, as the number of experts increases, it becomes statistically more difficult to achieve high levels of agreement, thereby requiring more stringent criteria for item acceptance. In this study, seven experts participated in the card sorting and CVI evaluations; they consisted of university professors, researchers, and data team leaders from various sectors and market segments.

These specialists possessed relevant qualifications, including applied research experience, leadership of data teams, and academic publications in the field. Selection criteria also considered participants' familiarity with the constructs under analysis: data leadership and project success. **Table 2** summarizes the profiles of the experts who took part in the card sorting and CVI exercises.

Table 2. **Profile of experts** (n = 07).

#	Education	Position	Sector	Experience (years)
E1	PhD in Management	Strategy and Projects Manager	Public Sector	20
E2	PhD in Management	Governance and Management Advisor	Education	5
E3	Master's Degree	Procurement Analyst and Researcher	Agribusiness	2
E4	Postgraduate	Data Analyst and Researcher	Education	7
E5	Master's Degree	IT Technician	Healthcare	10
E6	Postgraduate	Data Director	Public Sector	8
E7	PhD in Management	Data Scientist	Public Sector	10

In addition to evaluating item content through the card sorting and CVI tests, the experts also provided face validity feedback via qualitative comments on the scale's overall structure, clarity, and relevance. These contributions were essential for refining the instrument and ensuring alignment between the items and the proposed theoretical constructs (Polit & Beck, 2006; Righi *et al*., 2013).

3.3 Analysis Procedures

After the research instrument was validated by the experts, the measurement model was tested by collecting empirical data that were subjected to factor analysis to assess the quality of the constructs.

This stage allows for the identification of latent factors, the exploration of the instrument's dimensionality, and the verification of whether the items group according to the expected theoretical structure (Hair, 2019). Polit and Beck (2006) emphasize that content validity, although essential in the initial stages of scale development, is insufficient to ensure an instrument's robustness without the empirical support provided by techniques such as exploratory factor analysis (EFA). Moreover, factor analyses offer objective evidence of construct validity by revealing how items relate to each other and to the latent factors, and are recommended in rigorous scale development and refinement protocols (Hinkin, 1998). Preliminary statistical tests for sample analysis and exploratory factor analysis of the data leadership dimensions were conducted using IBM SPSS v20.

The evaluation of the research model was carried out using SmartPLS v4 software, employing the Partial Least Squares Structural Equation Modeling (PLS-SEM) technique. The choice of PLS-SEM over covariance-based SEM (CB-SEM) was mainly driven by the characteristics of the available sample. PLS-SEM is recognized for being more robust with smaller sample sizes, as it focuses on prediction and the identification of relationships among constructs, which fully aligns with the objectives of this study. In summary, selecting PLS-SEM aligns with the need for a method that optimizes analysis given the available sample, ensuring the feasibility of the analysis and the extraction of insights from the data.

4 Data Analysis

To validate the research instrument, similarity analysis via card sorting and Content Validity Index (CVI) procedures were first conducted. Thereafter, empirical data were analyzed to validate the measurement model. The following subsections present these analytical procedures and their results.

4.1 Face and Content Validation

4.1.1 *Similarity Analysis with Card Sorting*

Among the content-validation procedures recommended in the literature, the card-sorting technique - particularly in its closed format, where experts categorize items according to predefined constructs - stands out (Wood & Wood, 2008). **Table 3** displays the similarity matrix results for this stage.

Table 3. **Card Sorting similarity matrix** (n = 7; items = 20).

Item	V	DM	D	E
Provide thought leadership for idea generation and innovation	**100**			
Monitor progress toward the vision	**86**	14		
Develop a global vision for data	**72**		28	
Define the strategic importance of data	**58**	28	14	
Inspire widespread use of data	**58**	14	14	14
Enable data-based decision-making	14	**72**		14
Democratize the use of data-driven decision tools	14	**72**	14	
Select data initiatives to implement	14	**72**	14	
Define data-quality standards and accountability	14	**44**	28	14
Break down data silos	14	**44**	14	28
Determine the value of specific data assets	14	14	**72**	
Show how to turn data into business value	14	14	**72**	
Integrate data experts into workflows	14		**72**	14
Design and implement data-based use cases	14	28	**58**	
Demonstrate that data drives the business	14	28	**44**	
Promote data learning				**100**
Support continuous development of data expertise				**100**
Empower people to develop new skills			14	**86**
Conduct programs to develop data literacy		14		**86**
Encourage knowledge sharing among cross functional teams			28	**72**

Legend: Visionary (V), Decision-Maker (DM), Driving (D), Educator (E).
Note: Percentage values; highest similarity percentages in each row are in bold.

This approach verifies the alignment between items and conceptual domains, supporting the analysis of representativeness and theoretical coherence of scale components. It also provides qualitative insights for instrument refinement, ensuring content validity (Hudson, 2014; Righi *et al*., 2013).

The test was administered online using the UXtweak platform (https://uxtweak.com/) in May 2024. The grouping of the 20 items into corresponding categories followed a predefined theoretical structure, as recommended for closed card sorting (Faria, 2010). Analysis via the similarity matrix - a recognized effective technique for *card-sorting* analysis (Righi *et al*., 2013); allowed identification of convergence patterns among experts and visualization of semantic relationships between items. Overall, items were classified as expected, indicating satisfactory evidence of content validity. However, the items “Define data-quality standards and accountability,” “Break down data silos,” and “Demonstrate that data drives the business” showed lower agreement among experts, indicating a need to revise their wording.

4.1.2 Content Validity Analysis

In this stage, seven experts with relevant experience participated. To analyze the data, the Content Validity Index (CVI) was used. **Table 4** shows the proportion matrix.

Table 4. **CVI proportion matrix** (n = 7; Items = 22).

Item	Experts							Total Agreement	CVI Item
	E1	E2	E3	E4	E5	E6	E7		
VD1.	3	3	3	4	4	4	4	7	**1,00**
VD2.	4	3	4	3	4	4	4	7	**1,00**
VD3.	4	4	4	3	4	2	4	6	**0,86**
VD4.	3	3	3	2	4	2	3	5	0,71
VD5.	3	4	3	4	4	2	4	6	**0,86**
VD6.	1	3	2	4	4	4	4	5	0,71
VD7.	4	4	3	2	3	3	4	6	**0,86**
TD1.	4	3	3	3	3	4	3	7	**1,00**
TD2.	4	4	4	4	3	3	4	7	**1,00**
TD3.	4	3	3	3	4	4	1	6	**0,86**
TD4.	3	4	3	4	1	1	3	5	0,71
TD5.	4	4	3	4	4	4	4	7	**1,00**
TD6.	4	4	4	4	2	4	4	6	**0,86**
ID1.	3	1	3	3	2	2	3	4	0,57
ID2.	1	4	4	4	2	3	4	5	0,71
ID3.	1	2	3	4	3	2	4	4	0,57
ID4.	4	4	2	4	1	3	4	5	0,71
ID5.	4	4	3	3	4	4	4	7	**1,00**
ED1.	3	4	3	4	1	4	4	6	**0,86**
ED2.	3	4	4	4	2	4	4	6	**0,86**
ED3.	3	3	3	4	3	4	4	7	**1,00**
ED4.	4	4	3	3	4	4	4	7	**1,00**
Total	19	20	20	20	15	16	21	**AVE-IVC[2]=**	**0,85**
								UA-IVC[3]=	**0,36**
%	**86,4**	**90,9**	**90,9**	**90,9**	68,2	72,7	**95,5**	**Média**	**0,86**

Note: Items with CVI ≥ 0.80 are highlighted in bold.

The CVI is calculated as the proportion of experts rating an item ≥ 3, indicating favorable judgment (Polit & Beck, 2006). Experts assessed each item for clarity, relevance, and representativeness on a 4-point ordinal scale, following Davis (1992) and Lynn (1986). Total CVI values ≥ 0.80 were considered satisfactory, consistent with prior studies (Adrian *et al.*, 2019). These results allowed verification of the level of agreement among experts regarding the scale’s content validity.

4.1.3 Synthesis of Adjustments in the Development of the Research Instrument

In addition to the quantitative evaluation, the experts also provided qualitative feedback on item wording and conceptual scope, in line with best practices recommended for assessing face validity (Davis, 1992; Hudson, 2014; Righi *et al.*, 2013).

The main suggestions included wording refinements, sentence simplification, and removal of redundant items. These changes not only improved clarity and comprehensibility but also reinforced that the constructs were well defined and adequately covered the proposed theoretical domain, a critical element for content validity (Lynn, 1986; Polit & Beck, 2006). Adjustments occurred between the card sorting and CVI stages, with additions, deletions, and item reformulations. While 20 items were used in the card sorting exercise, 22 items were evaluated in the CVI stage. No new conceptual elements emerged, which underscores the theoretical and practical consistency and representativeness of the original structure.

The adjustments made based on the experts' suggestions are presented in **Table 5**.

Table 5. **Adjustments to the research instrument.**

Variable Description	Adjustment
VD1. Leads the development of a corporate vision for data projects VD2. Leads the strategic prioritization of data projects VD3. Encourages widespread use of data in organizational projects VD4. Encourages the building of data-driven- innovation projects VD5. Helps demonstrate how data projects drive the business TD1. Indicates how to monetize data projects into commercial value TD3. Defines data quality standards TD4. Defines the data owner(s) ID1. Leads the process of data monetization ID2. Supports implementation of data-driven use cases ID3. Defines measures to reduce data silos ID5. Leads the management of data projects ED1. Promotes learning of new data skills	Content adjusted - *card sorting stage*
VD6. Encourages the building of data-driven innovation projects VD7. Leads the process of developing a data culture TD6. Encourages data-based decision-making	Content adjusted – *card sorting stage*
ED3. Empowers people to develop new skills VD5. Helps demonstrate how data projects drive the business	Content adjusted – *card sorting stage*
VD6. Encourages the building of data-driven innovation projects TD6. Encourages data-based decision-making TD1. Identifies which data have strategic value for the organization	Item removed – IVC stage
TD2. Democratizes the use of data tools TD4. Defines data professionals' roles and responsibilities TD5. Defines the data project portfolio ID1. Leads initiatives that turn data into tangible organizational results ID2. Supports implementation of data driven use cases ID5. Manages the development of data projects	Item removed – IVC stage

Note: other items remained unchanged.

4.3 Sample Analysis

4.3.1 Sample Size Determination and Data-Collection Procedures

The population comprises managers working with data analytics in organizations operating in Brazil across various economic sectors. The unit of analysis in this study is "managers"; thus, the target population included professionals, executives, and managers who work with data and employ data analytics in their organizations. When presenting the data-collection form, respondents were instructed to base their answers on their own perceptions and experiences within their organization. They were informed that this was scientific research and asked to authorize the publication of their responses in aggregated form, ensuring the confidentiality of individual information.

The minimum sample size was estimated using *G*Power V3.1.9.4 software* (Faul, Erdfelder, Buchner, & Lang, 2009). Following Hair, Hult, Ringle, and Sarstedt's (2022) recommendations, we set the statistical power to 0.80, a medium effect size ($f^2 = 0.15$), and, considering that the independent variable has four predictors (the four data leadership dimensions), calculated a minimum sample of 85 respondents.

Data were collected via an electronic survey administered through Google Forms during June 2025. No evidence of automated or random responding was detected. After this qualification screening, the final sample comprised 91 valid responses, exceeding the recommended minimum.

4.3.2 Test for Systematic Bias in the Sample

Because the data are primary, we needed to ensure that no systematic bias affected the collected information. We applied Harman's single-factor test (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), where a single extracted factor should explain less than 50% of the variance. The test indicated that one factor accounted for 42.7% of the variance - below the 50% threshold - demonstrating that common-method bias is not a significant concern. Additionally, a four-factor solution explained 65.9% of the variance in the phenomenon.

4.3.3 Descriptive Analysis of the Sample

The sample was predominantly male (71.6%). The most common academic background among participants was in business fields (management, economics, or finance), followed by technology or information systems, engineering, and statistics.

Tables 6 provide detailed respondent profiles.

Table 6. **Sample Profile** (n = 91).

G	N	(%)	Experience (years)	N	(%)	Education	N	(%)	Time allocation (%)	N	(%)
F	26	28.6	≤1	2	2.2	Business	33	36.2	≤20	2	2.2
M	65	71.6	1<x≤3	17	18.6	IS/IT	24	26.4	20<x≤40	16	17.6
			3<x≤8	27	29.6	Engineering	14	15.4	40<x≤60	19	20.9
			8<x≤15	26	28.6	Statistics	9	9.9	60<x≤80	29	31.9
			x>15	19	20.9	Sciences	4	4.4	x>80	25	27.4
						Law	2	2.2			
						Others	5	5.5			

Position	N	(%)	Economic Sector	N	(%)
Data Analyst	17	17.6%	Technology	26	28.6
Business Analyst	13	13.2%	Government	15	16.4
Data Coordinator/Supervisor	11	11.0%	Retail/Commerce	9	9.9
Data Engineer	9	9.9%	Consulting	9	9.9
Data Manager	8	8.8%	Financial Services	9	9.9
C-Level/Data Director	8	7.7%	Education	7	7.7
Project/Product/Innovation Manager	4	1.1%	Healthcare	6	6.6
Researcher	4	4.4%	Healthcare	3	3.3
Tech Lead	4	4.4%	Construction/Real Estate	2	2.2
Data Scientist	4	4.4%	Manufacturing	2	2.2
Data Product Manager	3	3.3%	Manufacturing	2	2.2
Design Analyst	2	2.2%	Fashion	1	1.1
Projects/Innovation Analyst	2	2.2%			
Business Manager	2	2.2%			

Regarding years of experience with data analytics, the vast majority have more than 3 years, with nearly half having over 8 years, suggesting a professionally mature sample. Most respondents dedicate more than half of their working time to data-related activities, reinforcing the technical specialization of the sample. The most common positions are data coordinator or supervisor, indicating a strong presence of mid-level leadership positions. The most represented sector is Technology, followed by Government, Financial Services, Retail, and Consulting, reflecting the importance of data-oriented professions.

Additionally, to ensure the sample's appropriateness, respondents were asked which types of analytical work processes they engage in within data projects. Process categories were inferred from the literature.

Figure 2 depicts the frequency analysis of these processes in the sample.

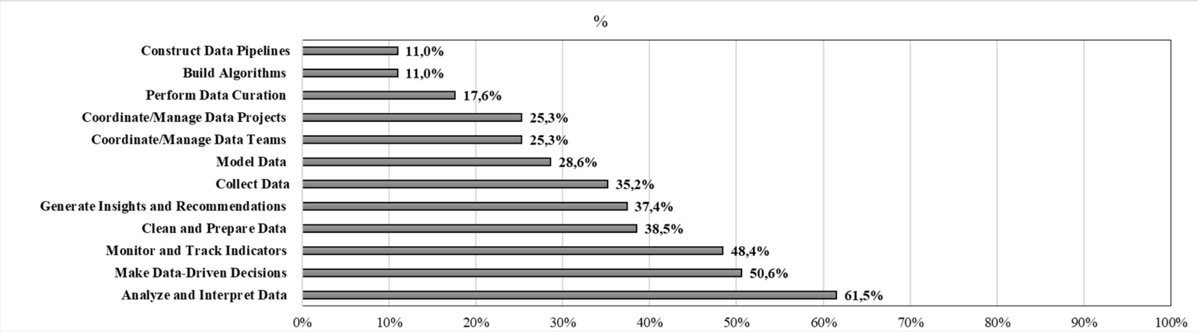


Figure 2. Most frequent analytical work processes in data projects (n = 91).

It can be observed that the most frequent processes (≈ 50% or more) in the sample are: analyzing and interpreting data, making data-driven decisions, and monitoring indicators. Between 35% and 38% of respondents engage in data collection, data cleaning, and generating insights. Processes such as data modeling, managing data teams and projects, and data curation appear in 15%–30% of cases. The least frequent processes (≤ 15%) are building data pipelines and algorithms. Therefore, the sample is deemed appropriate for the study's objectives.

4.4 Exploratory Factor Analysis of the Data Leadership dimensions

The items for the dependent variable "Project Success" scale had already been empirically validated in previous studies (Shenhar & Dvir 2007). However, although the four data-leadership dimensions are described in the literature, they have not been measured with empirical data until this study; thus, a rigorous examination of these four factors' dimensionality was required. We conducted an Exploratory Factor Analysis (EFA) following Hair *et al.* (2019).

Overall, normality issues were not detected based on skewness and kurtosis coefficients (Hair *et al.*, 2019). The dataset met adequacy conditions for factor analysis: Bartlett's test of sphericity was significant ($p < .05$), and the Kaiser–Meyer–Olkin (KMO) measure exceeded .60 (KMO = .863).

All variables had anti-image correlations above .50 (Hair *et al.*, 2019). However, variables TD2, TD3, and ID1 were removed for failing to reach communalities > .50 (Hair *et al.*, 2019), and VD1 was excluded because it cross-loaded (loading > .50) on both VD and TD factors (Hair *et al.*, 2019).

After these removals, a repeated dimension-reduction test in SPSS v20 using Varimax rotation (eigenvalue = 1) confirmed four factors. This solution explained a cumulative 71.62% of total variance—exceeding the recommended 60% threshold (Hair *et al.*, 2019).

Subsequently, we assessed scale reliability using Cronbach's alpha (α) and convergent validity via Average Variance Extracted (AVE). To achieve a satisfactory AVE for the VD construct, VD5 - whose loading conflicted with the TD construct - was excluded. Following this adjustment, both indicator reliability and convergent validity met recommended thresholds:

AVE values for all constructs exceeded .50, and internal consistency also reached acceptable levels (Hair *et al.*, 2019). **Table 7** presents the EFA results.

Table 7. **EFA – Data-Leadership dimensions:** reliability and convergent validity.

Item	Communality	Factor Loadings by Component				Adjustment
		Educator	Visionary	Driving	Decision-Maker	
VD2	.766	.091	.394	.053	**.747**	Dimension changed [VD>TD]
VD3	.664	.291	**.654**	.402	.061	
VD4	.692	.245	**.768**	.338	.071	
VD5	.695	.379	.629	.052	.437	Variable removed for AVE adjustment.
TD1	.538	.184	**.739**	.019	.135	Dimension changed [TD>VD]
TD4	.639	.388	.011	.184	**.656**	
TD5	.732	.052	.056	.239	**.837**	
ID2	.677	.067	.302	**.752**	.045	
ID3	.670	**.772**	.227	.075	.203	Dimension changed [ID>ED]
ID4	.707	.393	.120	**.738**	.163	
ID5	.710	.101	.045	**.774**	.328	
ED1	.730	**.721**	.322	.307	.101	
ED2	.781	**.760**	.288	.353	.101	
ED3	.731	**.850**	.063	-.042	.207	
ED4	.728	**.773**	.318	.246	.011	
Qtd. Itens		**5**	**3**	**3**	**3**	Dimensions with 3 or more items [OK]
AVE (%)		**.603**	**.521**	**.570**	**.563**	AVE > 0.5 [OK]
αC		**.898**	**.754**	**.771**	**.737**	αC > 0.7 [OK]

Next, we proceeded to analyze discriminant validity by the HTMT criterion (Hair *et al.*, 2019). **Table 8** demonstrates discriminant validity, assessed via HTMT < 0.85.

Table 8. **EFA – Data-Leadership dimensions:** discriminant validity.

Construct	Educator	Driving	Decision-Making	Visionary
Educator Data Leadership	-			
Driving Data Leadership	.590	-		
Decision-Maker Data Leadership	.533	.590	-	
Visionary Data Leadership	.722	.670	.538	-

Note: Discriminant validity assessed by the HTMT criterion.

Thus, the data-leadership measurement model fulfilled quality criteria, demonstrating internal consistency, convergent and discriminant validity across all constructs. Out of 19 initial items, only four were removed, leaving 15 items representing the four constructs, each with at least three indicators.

With EFA complete, we proceeded to evaluate the overall measurement model by examining the relationships between the four data-leadership dimensions and the dependent variable, Project Success. The results of this analysis are presented in the next section.

4.5 Evaluation of the Research Model

As justified in the Analysis Procedures section of this study, the PLS-SEM technique was chosen to evaluate both the measurement model of the constructs and the structural model for exploratory analysis of the hypothesized relationships.

In the measurement-model analysis, the following quality criteria were considered, as recommended by Hair *et al*. (2019, 2022): (i) Outer loadings of observable variables > 0.70; (ii) Cronbach's alpha (α) and (iii) Composite Reliability (CR) of constructs > 0.70; (v) Average Variance Extracted (AVE) for convergent validity of constructs > 0.50; (v) Discriminant validity, assessed via HTMT < 0.85; and (vi) Standardized Root Mean Residual (SRMR) < 0.08.

Table 9 presents the results for these quality and validation criteria.

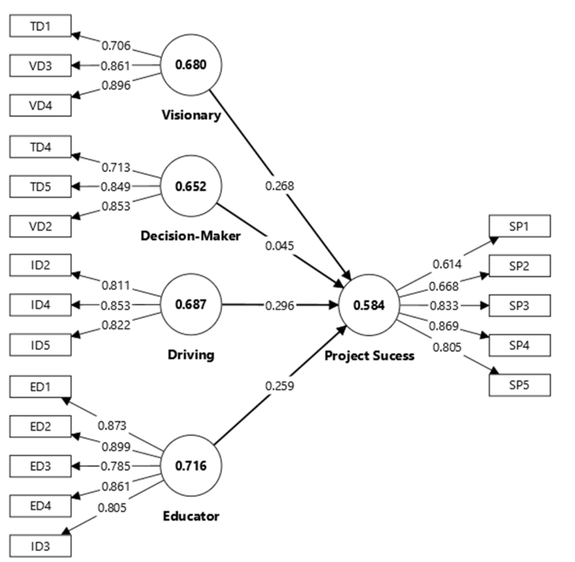
Table 9. **Measurement model evaluation.**

Construct	**Educator**	**Driving**	**Project Success**	**Decision-Making**	**Visionary**
Educator Data Leadership	**-**				
Driving Data Leadership	**.590**	**-**			
Project Success (PS)	**.666**	**.754**	**-**		
Decision-Maker Data Leadership	**.533**	**.590**	**.477**	**-**	
Visionary Data Leadership	**.722**	**.670**	**.738**	**.538**	-
αC	**.902**	**.773**	**.817**	**.736**	**.763**
CC	**.923**	**.777**	**.837**	**.770**	**.788**
AVE	**.716**	**.687**	**.584**	**.652**	**.680**

Legend: Cronbach's alpha (α); Composite Reliability (CR); Average Variance Extracted (AVE). **Note**: Discriminant validity assessed by the HTMT criterion.

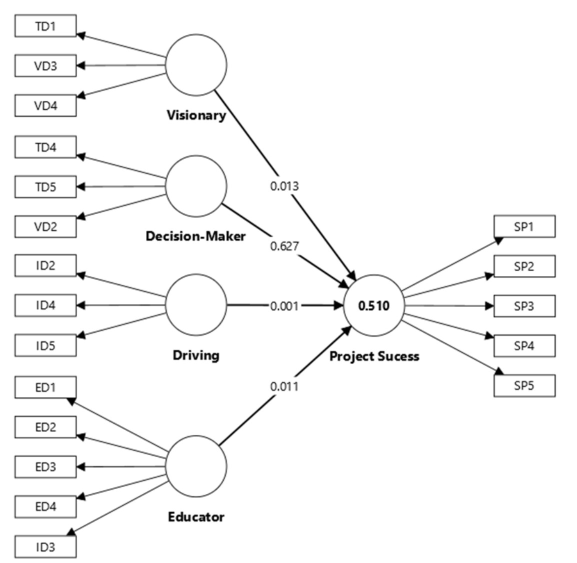
The model performed well across all of these quality criteria, indicating that the instrument is a useful tool for collecting data and testing research hypotheses. **Figure 3** presents the measurement and structural models.

Measurement Model



Note 1. Construct loadings for variables.
Note 2. Path coefficients between constructs.
Note 3. AVE in the constructs.

Structural Model



Note 1. p-Values on the arrows between constructs.
Note 2. R^2 on the construct (dependent variable).

Figure 3. Measurement and structural model via PLS-SEM.

The structural model was evaluated as recommended by Hair *et al.* (2019, 2022). The model exhibits a large Pearson's coefficient of determination ($R^2 = 0.510$) and strong predictive relevance ($Q^2 = 0.44$). Fit is satisfactory (SRMR = 0.078).

Path-analysis shows a statistically significant relationship for ID → SP ($\beta = 0.296$; $f^2 = 0.112$; $p = 0.001$) and for ED → SP ($\beta = 0.259$; $f^2 = 0.074$; $p < 0.05$) and VD → SP ($\beta = 0.268$; $f^2 = 0.079$; $p < 0.05$). All three effects are small (f^2). However, TD → SP is not significant ($\beta = 0.045$; $f^2 = 0.003$; $p > 0.05$).

These results will be discussed in detail in the following section.

5 Discussion of Results

This study aimed to fill an important methodological gap in the literature by proposing and validating a multidimensional scale to measure types of data leadership and their relationship with project success. The proposed instrument was developed on a theoretical foundation drawn from a systematic literature review (Fleck & Maçada, 2025) and recent studies on data leadership (Haude *et al*., 2024; Schmidt *et al*., 2023; Tabesh *et al*., 2019).

During item development, face and content validity were assessed. Establishing face and content validity underscores the importance of grounding instrument construction in clear, representative items aligned with the specific conceptual domain. Experts were invited to categorize and review items for clarity, relevance, and representativeness. Combining card sorting and the Content Validity Index (CVI) confirmed item consistency with the proposed theoretical constructs. Items achieving a CVI ≥ 0.80 were deemed valid, demonstrating adequate initial validity.

Empirical data were then collected to test the instrument. The exploratory factor analysis results indicated necessary adjustments to validate the measurement model, which will be discussed in detail in the subsequent subsections.

5.1 Discussion on the development and validation of the Data Leadership scale

First, it is worth noting that, unlike previous conceptual studies on data leadership (Haude *et al*., 2024; Schmidt *et al*., 2023), this research advances the empirical construction of a multidimensional scale. While Fleck & Maçada (2025), Haude *et al.* (2024) e Schmidt *et al.* (2023) had already proposed different types of data leadership, no validated quantitative instruments existed to measure these dimensions simultaneously and test their relationship with project success. By operationalizing the four leadership types, this study makes an original contribution by enabling empirical diagnostics and comparative analyses.

The empirical test revealed the need to remove five originally planned variables due to low communalities - TD2, TD3, and ID - or cross-loadings between different constructs, as in the cases of VD1 and VD5 (leads the development of a data culture). Their removal may be linked to the semantic overlap already noted during card sorting, especially between the TD and ID constructs. Similarity analysis had indicated some difficulty in clearly separating strategic responsibilities (TD) from operational ones (ID) when experts categorized items. This suggests that in organizational practice, these leadership roles may overlap, particularly in environments with less mature data governance. Furthermore, Content Validity Index (CVI) criteria showed that some items fell below the expected threshold, signaling representativeness issues.

For example, ID3 ("defines measures to reduce data silos") may overlap with TD2 ("democratizes the use of data tools"), since both address ensuring broad data accessibility—whether by breaking down silos or democratizing data (Awasthi & George, 2020). Similarly, TD3 ("defines data quality standards") could be conflated with TD4 ("defines data professionals' roles and responsibilities"), as in practice the distinction between setting standards and defining roles is subtle; both relate to data governance (Medeiros & Maçada, 2022). A comparable overlap may exist between ID1 ("leads initiatives that turn data into tangible organizational outcomes") and TD1 ("identifies which data have strategic value for the organization"), as both concern specific stages of converting data into value—nuances that practitioners may not easily differentiate in everyday settings.

VD1 ("leads the development of a corporate vision for data projects") may have been confused with TD5 ("defines the data project portfolio"), since both pertain to data project strategy management. Finally, VD5 ("leads the process of developing a data culture") showed little specificity for the broad, multidimensional phenomenon of a data-driven culture, and thus may not have stood apart from other items addressing cultural aspects—mindset, data literacy, accessibility, and governance (Javed & Akhlaq, 2024).

Both conceptual and empirical findings confirm the validity of the proposed multidimensional structure, directly addressing the theoretical gap identified. The final model, with 15 items across four constructs, demonstrates internal consistency and convergent validity in all dimensions, with AVE values exceeding 0.50. This reinforces that data leadership is not a unidimensional phenomenon but a multifaceted one, involving distinct leadership types with specific cognitive, behavioral, cultural competencies, and actions, as earlier qualitative studies suggested (Haude *et al.*, 2024; Mikalef *et al.*, 2021; Tabesh *et al.*, 2019). By operationalizing these dimensions in an integrated manner, this scale enables more sophisticated analyses of leadership's role in analytical projects.

Therefore, empirical evidence supports Hypothesis H1, confirming that data leadership is a multifaceted factor reflected in four dimensions (visionary, decision-maker, driving, and educator).

5.2 Exploring the relationship between Data Leadership types and Project Success

Empirical evidence supports hypotheses H2a (Visionary), H2c (Driver), and H2d (Educator), confirming that these dimensions of data leadership are positively related to the success of data and analytics projects. However, surprisingly, hypothesis H2b was not supported - there is no empirical confirmation that Data Decision Maker leadership affects project success. Each of these findings is discussed below.

(H2a) Visionary Data Leadership → Project Success [β = 0.268, *p-Value* > 0,05]
The statistically significant, high intensity relationship between visionary leadership and project success confirms the recognized role of strategic vision in organizational alignment (Fernandes *et al.*, 2022; Jansen *et al.*, 2009). Although its influence operates at a more abstract level, visionary leadership is essential for setting direction and inspiring long term transformation. Thus, when analysis focuses on specific project deliverables (such as scope, value generated, or operational alignment), the effects of visionary leadership become perceptible. However, it must be acknowledged that in many organizations a strategic data vision is not yet fully internalized or disseminated (Medeiros, Hoppen, & Maçada, 2020; Quach *et al.*, 2022), which limits its practical efficacy throughout the project lifecycle.

(H2b) Decision-Maker Data Leadership → Project Success [β = 0.045, p-Value > 0,05]
The most surprising finding is the weak relationship between decision maker leadership and project success, which lacks statistical significance. One hypothesis to explain this result lies in the potential confusion between control and trust: leaders overly focused on governance and standardization may, in certain contexts, stifle project fluidity and hinder innovation. Moreover, governance and decision-making practices can be perceived as detached from execution and sometimes disconnected from value delivery. It is also possible that the practical difficulty of

operationalizing data governance in organizations with low maturity generates frustration or resistance, negatively impacting project outcomes. These indications warrant deeper investigation in future studies.

(H2c) Driver Data Leadership → Project Success [β = 0.296, *p-Value* = 0.001]
The strong and significant relationship observed between driver leadership and project success suggests that the ability to operationalize the data strategy is a critical factor in generating tangible results (Baecker *et al.*, 2025; Medeiros, Maçada, & Freitas, 2020). This dimension involves actions such as implementing use cases, managing workflows, and integrating experts—elements directly connected to effective project execution. The literature already highlights that the absence of integration among technology, processes, and people is one of the main causes of failures in analytics initiatives (Brocchi *et al.*, 2018; Saif, 2020), which may justify the high weight of this leadership style on success. It can be inferred that, in complex organizational contexts, the pragmatic and coordinating role of the driver leader is critical for transforming analytical potential into real value.

(H2d) Educator Data Leadership → Project Success [β = 0.259, *p-Value* > 0,05]
The positive, high intensity relationship between educator leadership and project success corroborates the idea that developing analytical competencies and fostering a datadriven culture are important conditions, albeit with more diffuse and long term effects on shifting toward a data-driven mindset (Huynh, Veglio, & Gunkel, 2025; Javed & Akhlaq, 2024; Medeiros & Maçada, 2022). Investments in data literacy elevate an organization's analytical maturity (Cezar & Maçada, 2021; Medeiros *et al.*, 2020) and reduce the risk of project failure, but their effects may not be immediate. Nevertheless, although this leadership strengthens the cultural and cognitive foundation for data-oriented decisions, its direct contribution to the success of specific projects tends to be indirect and progressive.

In summary, the findings indicate relationships of varying strength between data leadership dimensions and project success. These insights shed light on the extent to which different leadership types can influence project success in data contexts.

6 Final Considerations

The study fully achieved its objective of developing and validating a scale to measure data leadership types and their relationship with data project success. Based on evidence gathered through face and content validity processes, as well as the execution of exploratory factor analysis, the research instrument was both developed and validated. It is expected that this instrument will contribute to theoretical advancement in the field of data management and leadership by providing an empirical foundation for future investigations, and to organizational practice by offering a diagnostic tool applicable to leadership development, training, and strategic alignment processes in data-driven environments.

Appendix A presents the final version of the scale.

6.1 Theoretical and Managerial Implications

This study stands out for its originality and innovation by empirically developing and validating the first multidimensional data leadership scale, establishing the relationship between different leadership types and data project success. Moreover, the standardization of data leadership types, combined with the initial validation of their indicators, provides a basis for empirical investigations and enables inter-organizational comparisons.

From a practical standpoint, the scale offers a useful tool for managers, consultants, and data professionals seeking to diagnose the leadership types present in their organizations. Identifying these types can inform the design of development initiatives, training programs, and strategies to align leadership types with the objectives of data oriented projects. Furthermore, the instrument supports the construction of a data-driven organizational culture.

6.2 Limitations and Future Research

Although the sample of 91 participants exceeded the minimum requirement of 85 responses, it is important to acknowledge that a larger sample would be preferable to achieve greater statistical power and generalizability of the results. This limitation may affect the extent to which the study's findings can be extrapolated to broader populations or different contexts, suggesting that future research with a larger number of participants could further deepen and validate the results presented here.

The face and content validity evidence and factor analysis reported here provide a foundation for subsequent steps, such as structural equation modeling to test hypotheses in research models investigating relationships between data leadership and other related constructs (Hair *et al.*, 2022). There are numerous possibilities, given that leadership interrelates with other constructs linked to data-driven governance and culture, such as data literacy, a data-driven mindset, data stewardship, data-driven decision-making, organizational agility, and performance, among others.

It is also recommended that future studies expand the respondent sample and conduct additional analyses, including examining variations in project performance according to organizational sector and management maturity level. These analyses could strengthen the robustness and applicability of the scale across different contexts. Continuing this validation process may consolidate the scale as a reliable instrument for measuring data leadership and its impact on data project success.

References

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*, 49, 424–438.

Adrian, C., Abdullah, R., Jusoh, Y., & Atan, R. (2019). A Content Validity Study for Big Data Analytics Implementation Model (p. 5).

Afzal, M. H. B. (2014). Large Scale IT Projects: Study and Analysis of Failures and Winning Factors. *IETE Technical Review*, 31(3), 214–219.

Ahmed, R., Shaheen, S., & Philbin, S. P. (2022). The role of big data analytics and decision-making in achieving project success. *Journal of Engineering and Technology Management*, 65, 101697.

Atkinson, R. (1999). Project management: Cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. *International Journal of Project Management*, 17(6), 337–342.

Awasthi, P., & George, J. J. (2020). A Case for Data Democratization, in Proceedings of the 26th *Americas Conference on Information Systems (AMCIS)* (Vol. 23), Virtual conference, August 10.

Baecker, J., Weking, J., Hein, A., & Krcmar, H. (2025). Organizational data strategy: Unveiling key elements and strategic types. *Journal of Information Technology*, 0(0).

Brocchi, C., Grande, D., Rowshankish, K., Saleh, T., & Weinberg, A. (2018). Designing a data transformation that delivers value right from the start. *McKinsey and Company*, October.

Cezar, B. G.d.S, & Maçada, A. C. G. (2021). Data literacy and the cognitive challenges of a data-rich business environment: an analysis of perceived data overload, technostress and their relationship to individual performance. *Aslib Journal of Information Management*, *73*(5), 618-638.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5(4), 194–197.

Deloitte. (2020). Data Literacy, success, analytics. Deloitte, Belgium. https://www2.deloitte.com/be/en/pages/technology/articles/data-literacy.html

Faria, M. M. de. (2010). Card Sorting: Noções sobre a técnica para teste e desenvolvimento de categorizações e vocabulários. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, 8(1), Artigo 1.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160.

Fernandes, C. I., Veiga, P. M., Ferreira, J. J., Rammal, H. G., & Pereira, V. (2022). Assessing strategic leadership in organizations: Using bibliometric data to develop a holistic model. *Journal of Business Research*, 141, 646–655.

Ferraris, A., Mazzoleni, A., Devalle, A., & Couturier, J. (2018). Big data analytics capabilities and knowledge management: Impact on firm performance. *Management Decision*, 57(8), 1923–1936.

Ferreira, C., Merendino, A., & Meadows, M. (2023). Disruption and Legitimacy: Big Data in Society. *Information Systems Frontiers*, 25(3), 1081–1100.

Fleck, M. M., & Maçada, A. C. G. (2025). Leadership in Management Information Systems: A Systematic Review. *AMCIS 2025 Proceedings*.

Gartner. (2021). 3 Ways to Be an Influential Chief Data Officer. Gartner. https://www.gartner.com/en/publications/3-ways-for-chief-data-officers-to-grow-their-influence-and-business-impact

Gartner. (2024). *Gartner Predicts 80% of D&A Governance Initiatives Will Fail by 2027, Due to a Lack of a Real or Manufactured Crisis*. Gartner.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2019). *Multivariate data analysis*. Cengage learning. Hampshire, United Kingdom, 633.

Hair, J., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*.

Haude, C., Blohm, I., & Lagardère, X. (2024). *How Lufthansa Shapes Data-Driven Transformation Leaders*. MIT Sloan Management Review.

Hinkin, T. R. (1998). A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. *Organizational Research Methods*, 1(1), 104–121.

Hudson, W. (2014). Card Sorting. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/card-sorting

Huynh, M. T., Veglio, V., & Gunkel, M. (2025). Conceptualizing the data-driven mindset: An application of the mindset theory of action phases. *Technovation*, 146, 103293.

Jansen, J. J. P., Vera, D., & Crossan, M. (2009). Strategic leadership for exploration and exploitation: The moderating role of environmental dynamism. *The Leadership Quarterly*, 20(1), 5–18.

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly*, 37(3), 101493.

Javed, B., & Akhlaq, A. (2024). A Systematic Review of Exploring the Multiple Dimensions of Data-Driven Culture. *International Journal of Trends and Innovations in Business & Social Sciences*, 2(4), 522–536.

Kadarsah, D. R. A. K., Govindaraju, R., & Prihartono, B. (2023). The Role of Knowledge-Oriented Leadership in Fostering Innovation Capabilities: The Mediating Role of Data Analytics Maturity. IEEE Access, 11, 129683–129702. *IEEE Access*.

Lee, Y., Madnick, S., Wang, R., Zhang, H., & Wang, F. (2014). A Cubic Framework for the Chief Data Officer: Succeeding in a World of Big Data. *MIS Quarterly Executive*, 13(1).

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385.

MacKenzie, S., Podsakoff, P., & Podsakoff, N. (2011). Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. *MIS Quarterly*, 35, 293–334.

Medeiros, M. M. D., Hoppen, N., & Maçada, A. C. G. (2020). Data science for business: benefits, challenges and opportunities. *The Bottom Line*, *33*(2), 149-163.

Medeiros, M.M.d., & Maçada, A.C.G. (2022). Competitive advantage of data-driven analytical capabilities: the role of big data visualization and of organizational agility, *Management Decision, 60*(4), 953-975.

Medeiros, M.M.d., Maçada, A.C.G., & Freitas Júnior, J.C.d.S. (2020). The Effect of Data Strategy on Competitive Advantage, *The Bottom Line, 33*(2), 201-216.

Menukin, O., Mandungu, C., Shahgholian, A., & Mehandjiev, N. (2023). Guiding the integration of analytics in business operations through a maturity framework. *Annals of Operations Research*.

Mikalef, P., van de Wetering, R., & Krogstie, J. (2021). Building dynamic capabilities by leveraging big data analytics: The role of organizational inertia. *Information & Management*, 58(6), 103412.

Müller, R., & Turner, J. R. (2007). Matching the project manager's leadership style to project type. *International Journal of Project Management*, 25(1), 21–32.

Nisar, Q. A., Nasir, N., Jamshed, S., Naz, S., Ali, M., & Ali, S. (2020). Big data management and environmental performance: Role of big data decision-making capabilities and decision-making quality. *Journal of Enterprise Information Management*, 34(4), 1061–1096.

Norena-Chavez, D., & Thalassinos, E. (2023). Impact of big data analytics in project success: Mediating role of intellectual capital and knowledge sharing. *Journal of Infrastructure, Policy and Development*, 7(3), Artigo 3.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, *88*(5), 879.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497.

Qlik, & Accenture. (2022). *The seven principles of data literacy*. The Data Literacy Project.

Quach, S., Thaichon, P., Martin, K. D., Weaven, S., & Palmatier, R. W. (2022). Digital technologies: tensions in privacy and data. *Journal of the Academy of Marketing Science*, 1-25.

Righi, C., James, J., Beasley, M., Day, D. L., Fox, J. E., Gieber, J., Howe, C., & Ruby, L. (2013, maio 7). Card Sort Analysis Best Practices—JUX. *JUX - The Journal of User Experience*.

Saif, T. H. D., Nitin Mittal, and Irfan. (2020, fevereiro 3). *What Separates Analytical Leaders From Laggards?* MIT Sloan Management Review.

Schmidt, D. H., van Dierendonck, D., & Weber, U. (2023). The data-driven leader: Developing a big data analytics leadership competency framework. *Journal of Management Development*, 42(4), 297–326.

Serrador, P., & Turner, R. (2015). The relationship between project success and project efficiency. *Project Management Journal*, 46(1), 30-39.

Shamim, S., Zeng, J., Shariq, S. M., & Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information & Management*, 56(6), 103135.

Shao, Z. (2019). Interaction effect of strategic leadership behaviors and organizational culture on IS-Business strategic alignment and Enterprise Systems assimilation. *International Journal of Information Management*, 44, 96–108.

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems, 23*(4), 433–441.

Shenhar, A. J., & Dvir, D. (2007). *Reinventing project management: the diamond approach to successful growth and innovation*. Harvard Business Review Press.

Shenhar, A. J., Tishler, A., Dvir, D., Lipovetsky, S., & Lechler, T. (2002). Refining the search for project success factors: A multivariate, typological approach. *R&D Management*, 32(2), 111–126.

Shenhar, A., & Holzmann, V. (2017). The Three Secrets of Megaproject Success: Clear Strategic Vision, Total Alignment, and Adapting to Complexity. *Project Management Journal*, 48(6), 29–46.

Shenhar, A., Holzmann, V., Dvir, D., Shabtai, M., Zonnenshain, A., & Orhof, O. (2020). If You Need Innovation Success, Make Sure You've Got the Right Project. *IEEE Engineering Management Review*, PP, 1–1.

Tabesh, P., Mousavidin, E., & Hasani, S. (2019). Implementing big data strategies: A managerial perspective. *Business Horizons*, 62(3), 347–358.

Turner, R., & Zolin, R. (2012). Forecasting Success on Large Projects: Developing Reliable Scales to Predict Multiple Perspectives by Multiple Stakeholders over Multiple Time Frames. *Project Management Journal*, 43(5), 87–99.

Wood, J., & Wood, L. (2008, novembro 7). Card Sorting: Current Practices and Beyond - JUX. *JUX - The Journal of User Experience*, Volume 4, 1–6.

Zhang, Q., Sun, X., & Zhang, M. (2022). Data Matters: A Strategic Action Framework for Data Governance. *Information & Management*, 59(4), 103642.

Zhao, G., Xie, X., Wang, Y., Liu, S., Jones, P., & Lopez, C. (2024). Barrier analysis to improve big data analytics capability of the maritime industry: A mixed-method approach. *Technological Forecasting and Social Change*, 203, 123345.

APPENDIX A – VALIDATED RESEARCH INSTRUMENT

Construct	Item
Data Visionary	VD3. Encourages broad use of data in organizational projects.
	VD4. Encourages the development of data-driven innovation projects.
	TD1. Identifies which data have strategic value for the organization.
Data Decision Maker	**VD2. Leads the strategic prioritization of data projects.**
	TD4. Defines roles and responsibilities for data professionals.
	TD5. Defines the data project portfolio.
Data Driver	ID2. Supports the implementation of data-based use cases.
	ID4. Integrates data experts into workflows.
	ID5. Manages the development of data projects.
Data Educator	**ID3. Defines measures to reduce data silos.**
	ED1. Promotes learning of new data skills.
	ED2. Supports the continuous development of data expertise.
	ED3. Conducts programs to develop data literacy.
	ED4. Encourages sharing of data related knowledge across cross functional teams.
Project Success	SP1. Ensures project efficiency by securing deliveries within established standards.
	SP2. Fosters team well-being in the work environment.
	SP3. Ensures projects meet client needs and add value to their operations.
	SP4. Ensures projects align with organizational growth, generating positive financial impact.
	SP5. Promotes the long-term effects of projects, preparing the organization for future challenges and opportunities.

Note 1. Items whose dimension was changed during the factor analysis are highlighted; all other items remained unchanged.

Note 2. Variables removed during the empirical phase: VD1. Leads the development of a corporate vision for data projects; VD5. Leads the process of developing a data culture; TD2. Democratizes the use of data tools; TD3. Defines data quality standards; and ID1. Leads initiatives that transform data into tangible results for the organization.