HOGESCHOOL ROTTERDAM

Towards Data Literacy in Higher Education

Alfons Looman, Mels de Kloet, Shengyun (Annie) Yang, and Leo Klienbannink

Rotterdam, januari 2022

Abstract

Data has been deemed fundamental to the digital transformation in society and businesses. It can influence all spheres of a company. Data-driven insights enable companies to make better decisions and thus to become more competitive in the market. The demand for data-literate employees and citizens needs to be supplied through the current generation of students. To help meet this demand, this research aims to better understand the status quo of data literacy in today's higher education and investigates how students' data literacy can be improved during their Bachelor learning trajectories. Four specific studies were carried out to define and distinguish the key concepts, explore the demand for data literate employees in the job market, examine students' data literacy levels, and experiment with a possible method to enhance students' data competences. The results drawn from our empirical research reveal that our current education does not yet adequately equip our students with the level of data literacy required by the real business world. Moreover, the findings demonstrate the effectiveness of practice-oriented, data-driven research to enhance students' overall data competences and lead to a win-win-win situation for the main stakeholders. The contributions and recommendations based on this research are also discussed.

Acknowledgements

This report would not have been possible without the many contributions we received from individuals and organisations. First, we would like to thank the Taskforce for Applied Research SIA for their help in financing this project through KIEM funding.

For this research endeavour we needed student respondents. For privacy reasons, we were not able to invite them directly, so our gratitude goes out to all the colleagues who helped us to invite students from several study programmes of the RUAS Business School (both of its current components, HRBS and RBS) to participate in this research. Special thanks to Tanya Dimitrova, Helen de Haan-Cao, Martijn de Ruiter, Lex Hoozemans, Abdel el Yalte, Maltie Koediep, Steph Lausberg, Arend Keemink, Ingrid Bogert, Marije Grift-Groot and Kiran-Ramroop for their support. We thank all study programmes, lecturers and students who participated in our survey and experiment, including its testing. Without your time and effort we would not have succeeded.

We thank all consortium partners that contributed to our studies. Three companies, namely Hyperion BV, Live Action Play and Added Venue BV, contributed to our field experiment in the 'Datawerkplaats' (data workplace) project within the minor programme Controlling at the RUAS Business School. They provided real current business problems with which students could develop their data skills and data literacy. Special thanks go to the company representatives for their exceptional help, coordination and guidance: Derek Göbel represented Monetor (Hyperion, BV), Ruud Tjepkema represented Games at Work (Live Action Play) and Schiecentrale (Added Venue, BV). Moreover, IT Campus Rotterdam organises the cooperation between education and businesses in the region. Its director, Marijn Roelof, introduced us to cooperating company Monetor. We would like to extend our gratitude to Marijn for his introduction to our business partner.

We further wish to thank our directors who govern and manage the RUAS business domain: Zakia Guernina, executive board member, Fred Feuerstake, Director of HRBS, Mariska de Wit, Director of RBS, Arjen van Klink, Director of Research Centre Business Innovation and all study programme managers, for facilitating all staff involved and for their general support and interest in this project.

During the last month leading up to the presentation of our report there were many deadlines and administrative burdens. Without our amazing support team we would not have overcome them. We therefore thank our support team: Adri-Janne Vaartjes, Wikke Heij-Popken, Jade Irawan, Nicole Pothof and Tamara Rademaker-Minkman. Additionally, we would like to thank Arjen van Klink for reviewing this report, and Deborah Sherwood for her edits.

Abbreviations and note about the study programmes in this research

RUAS	Rotterdam University of Applied Sciences (vocational university)
HR	Hogeschool Rotterdam (same as above, in Dutch)
HRBS	Hogeschool Rotterdam Business School (Business faculty at RUAS with courses offered in Dutch)
RBS	Rotterdam Business School (Business faculty at RUAS with courses offered in English)

Note: The RUAS Business School currently comprises two separate faculties: HRBS and RBS. These will be merged within the next few years and will share the name HRBS. For this research, we collected interview and observations from both faculties (though the majority were from HRBS). To indicate this, in this research we refer to all studies and participants as being part of the RUAS Business School.

Content

1. Introduction

The digital transformation that our society and businesses are going through is driven by data. The large-scale collection, processing and analysis of data has enabled the rise of new technologies, such as Big Data, the Internet of Things (IOT) and Artificial Intelligence (AI). Meanwhile, the increased adoption of these new technologies further promotes the collection, analysis and use of data. To function effectively in this dynamic business environment, companies need to work in tandem with data. Nowadays nearly all processes have a significant digital component: many processes require data to function and deliver data as output.

Data is important in all business processes, such as marketing, accounting, sales, customer service, production, services, human resource management (HRM) and research and development (R&D). Data-driven insights allow companies to make better decisions and thus to become more competitive in the market. For example, companies can use data to better predict the right time to launch a new product for promising sales, identify the bottlenecks in their business processes for improvement, and gain a full view of their customers for precision marketing. Companies that are aware of the value of data use it to make better decisions and develop an organisational data strategy. In any business sector, the successful use of data provides a competitive advantage.

Nevertheless, to create value from data, companies need adequate knowledge of collecting, managing and analysing high quality data. In addition, it is also important to be aware of how the data and the analysis thereof can add value in a company. In turn, we see an increasing demand for employees with data competences in the labour market to realise data-intensive management.

1.1 Research Motivation

The cutting-edge technologies in the digital transformation, most of which revolve around data, can only be used by those who have 'the ability to collect, manage, evaluate and apply data in a critical manner' (Ridsdale et al., 2015), i.e., those who are data literate. Data literacy allows employees to use these new data-driven technologies with success and deal with data in an acceptable and appropriate way in business.

Data literacy is an important competency for a successful data-driven societal future. Prior research has emphasised two major reasons for dedicating more attention to data literacy in education.

First, we all live in an information society in which people recognise the need to deal effectively with data and information to solve problems, achieve social integration, optimise personal

development and actively participate in society (Prado & Marzal, 2013; Stephenson & Caravello, 2007).

Second, every aspect of the lives of young people, digital natives, school children and students leaves a data trail. Thus, they need a thorough understanding of data to manage their personal data effectively and safely (Bowler et al., 2017).

The demand for data-literate employees and citizens needs to be supplied through the current generation of students. Therefore, we at the Research Centre Business Innovation, one of the major research forces within Rotterdam University of Applied Sciences (RUAS), feel the need to enhance data literacy in our education and cultivate data-literate students prepared for this business and societal demand.

It is important for us to examine the current level of data literacy of our students in relation to job market expectations. This means that we need to understand the current data-driven issues that human resource departments search for when filling vacancies, the data skills that are needed to tackle these issues, and whether Bachelor students at RUAS acquire these skills during their training as starting professionals. Therefore we pose the following central research question:

How can we define, measure and promote data literacy in higher education?

Based on the results drawn from this research, all higher education institutes can gain a clearer picture of the concept of data literacy. This can provide a handle to curriculum designers so that they can more easily match the learning goals (the needs of the labour market) to learning activities (the data skills offered in courses). We also aim to create a frame of reference for conversations between educational partners and government. In addition, we expect that this and future data literacy research will keep education stakeholders in dialogue with each other about new methods and forms of education to improve and promote data literacy and related skills.

1.2 Research Scope

From a business perspective, we focus on small and medium-sized enterprises (SMEs) in the Rotterdam-The Hague Metropolitan Region (MRDH) due to three major reasons.

First, SMEs are the regional labour engine, as they provide 60% of employment in the Rotterdam area. Sustainable, circular, digital and data-driven ways of working are essential elements for the future of SMEs – not only to meet regulations and new laws, but also to meet the ever increasing expectations of customers and buyers. SMEs should rise to the occasion to become more data-driven, because this innovation will deliver attractive opportunities and create structural competitive advantages. Futureproof SMEs will shape a vital economy, labour market and city.

Companies that invest in the data literacy of their employees will become more resilient in times of crisis or pandemics like Covid-19. But the challenges in making these investments are huge, especially for SMEs with a relatively small budget and staff, and which have a knowledge shortage in comparison to large corporations (MKB010>>Next, 2021).

Second, the municipality of Rotterdam wants to stimulate and support SMEs in their transformation into the new economy. In particular, with the SME programme Innovation 2019-2022 (MKB010>>Next, 2021), the municipality of Rotterdam wants to inspire, inform and activate 5,000 SMEs to become more digital and to become sustainable companies (Ondernemen010, 2019). Nevertheless, the majority of SMEs find it difficult to keep up with the fast pace of innovation and struggle with the question of how they can shape it. Since an enormous number of innovations in the digital transformation have been enabled based on a data-driven approach, it is crucial for us to examine the data literacy needs of SMEs.

Third, from various parties, both inside and outside the university, including the consortium partners, we see an increased demand to support SMEs in the transformation to a new digital data-driven economy. The biggest labour suppliers for SMEs are the regional vocational institutes, such as RUAS. These institutes exist primarily to fulfil the labour needs of SMEs, so we need to understand what these needs are. Are they simply to replace the accountant and marketing professional that are ready for retirement, or do SMEs now need data-literate young professionals who can help them innovate and be successful through this digital transition and beyond? If vocational institutes can meet the challenge of educating a new generation of data-literate young professionals in all domains, we can ease the labour market's high demand for data literacy and help the municipality of Rotterdam and the city's SMEs reach their goals.

From an education perspective, data literacy is essential in all domains. However, in this research we focus on the business studies for three reasons. First, the Research Centre Business Innovation is connected to the RUAS Business School, and as such our mission is to strengthen business education, business research and regional businesses themselves. Second, professions in the business domain are often administrative in nature, and therefore they are usually the first ones transformed by automation and digitisation. Third, via the business domain we can make a much bigger impact than any other domain, as financial, business, administration and operations managers are employed in all sectors.

1.3 Structure of the Report

In this research we carried out four studies within the RUAS Business School, based on multiple research methods. Section 2 presents the first study, which aims to better understand the concept of data literacy. Section 3 introduces the second study, which examines to what extent students' data literacy has been addressed in various courses and programmes. Section 4 focuses on the third study, which explores the level of data literacy among students and lecturers.

Section 5 explains the fourth study, which investigates whether our education meets the data literacy requirements of companies. Section 6 further discusses the research findings drawn from the four studies. Section 7 concludes the research with a summary of the research findings, a discussion of the limitations, and suggestions for future research.

2. Study 1: Data Literacy

2.1 Introduction

Advanced analytics, including descriptive, diagnostic, predictive and prescriptive capabilities, are a top business priority. Gartner (2014) describes these four types in its analytics continuum model, starting from descriptive information in hindsight, via insight through diagnostics and predications, to the final step of optimisation with prescriptive analytics. The final stage is about using data to automate decision-making in order to realise organisation goals. This continuum is comparable to data maturity levels that are used to describe companies' analytics capabilities.

For organisations to benefit from these analytics endeavours, data literate staff are essential. The first step is to train people in these skills, and education institutes are a major base for cultivating talents. Researchers and policymakers agree that there is both an increasing societal need and a growing business demand for data savvy and data literate young professionals. To cater to this need, higher education has begun to focus attention on data literacy.

However, the way in which this should take form is not clear. Defining the concept of data literacy would therefore be helpful in order to structure, improve and speed up this discussion, not only within higher education but also between higher education and the labour market. As data literacy is a new concept, the term is currently not widely used in higher education. It is often used in tandem with or confused with digital literacy. To be able to use this concept we need a working definition. Thus the following research question for this study is put forth:

What is data literacy?

We aim to propose an accurate and precise definition that can mitigate the ambiguity with other concepts and yet be able to evolve in a fast-changing, technology-rich environment. We also attempt to break up the concept of data literacy into underlying competencies so that education programmes can incorporate and apply it to learning goals within their curricula, and to come up with a way to measure it. If we can measure data literacy, we can examine the status quo and state the goals we intend to reach.

2.2 Research Method

We carried out desk research to explore what is already known about data literacy and the associated concepts in existing research and policy documents online. Desk research fits the purpose of this reconnaissance because of its low cost and quick results. We found several relevant frameworks around these concepts in the scientific literature, reports from international and national policy makers, and some private initiatives that are worth mentioning.

2.3 Research findings

2.3.1 Data

In our search to define 'data' for the purposes of our research, we considered that the word 'data' is used in Dutch as well as in English, yet this is new to the digital age. The literal translation of 'data' in the traditional sense of the word is 'gegevens'. We therefore used the Van Dale (the most prominent dictionary of the Dutch language) as a starting point. It provides one relevant definition for 'gegevens': known facts from which to draw inferences. It can be argued that if the term 'data' is used in Dutch instead of the word 'gegevens', it is more likely to mean digital data that is used in the digital domain. In support of this, the Dutch etymological dictionary (Philippa et al., 2004), describes 'data' as data, facts and units that can be processed with and within computers.

In current English definitions, Eurostat (2022) states in its glossary that 'data' refer to characteristics or information, usually numerical, that are collected through observation and that data are typically the results of measurements and can be visualised using graphs or images. The Merriam-Webster dictionary provides three definitions for data: 1. factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation; 2. information in digital form that can be transmitted or processed; 3. information output by a sensing device or organ that includes both useful and irrelevant or redundant information and must be processed to be meaningful.

In our current highly digitised society, most data is captured digitally, so for this research project we refer to data and data-driven in dealing with digital data and digital datasets. The advantage of using the English terminology in an otherwise Dutch context is that we are more likely to intrinsically and automatically connect 'data' to the digital domain, as we see in the definitions by Philippa and Merriam-Webster.

Commercial providers of data analysis services distinguish between information-driven and data-driven processes. They see information-driven as a way of working in which data is combined to form Key Performance Indicators, based on definitions and standards that the organisation sets itself. Data-driven working is described as being directly based on raw, original, real-time data – for example, the real-time display of advertisements based on search behaviour, or the start-up/control of maintenance based on sensor data that monitors wear and tear.

This differentiation points out that in an information-driven way of working, data has already been altered or manually created and structured, whereas in data-driven work raw data is used to make decisions. A data-driven business model, therefore, is realised when data is a source of great importance in the business process (Hartmann et al., 2014). In other words, being 'data-driven' indicates that organisations set up decision-making, control and processes based on

(mostly automated) digitally recorded observations and registrations of applications, systems and sensors.

Since this research is motivated by data-driven management in the business world, we define 'data' in this research on data literacy from this point of view: as the original raw digital registration of facts from events and processes.

2.3.2 Literacy

Evolutionary development is helpful in explaining the concept of literacy. Primary knowledge is what was adopted during the millions of years of evolution of our species; secondary knowledge includes language acquisition, which formed during the last 100,000 years among homo sapiens. Literacy traditionally refers to written language, having developed only about 6,000 years ago as a sort of cultural 'expertise' for which evolution does not provide a corresponding specialised neural substrate (Bulajic et al., 2019).

In recent history literacy can be divided into the period until 1950, when literacy was understood solely as alphabetic literacy (word and letter recognition), and the period from 1950 onwards, when literacy became used as a wider concept and included a more dynamic process which became known as functional literacy (Dijanošić, 2009). In 1978 at the UNESCO General Assembly a functionally literate person was defined as someone 'who can engage in all those activities in which literacy is required for the effective functioning of his group and community and also for enabling him to continue to use reading, writing and calculation for his own and the community's development'. We see in this definition (UNESCO, 1979) that calculation is also included in the definition of functional literacy. Handling and calculating with numbers are also referred to as numeracy.

Other sources also consider literacy to be an individual good as well as a social good for the development of communities. For an individual it builds up one's cognitive skills, and as a social good it builds a more informed citizenry, according to Prado and Marzal (2013). Their research hypothesised that each media or information skill has three purposes. The first is to create democracy, participation and active citizenship. The second goal is to create a knowledge-based economy, competitiveness and choice. The final goal is to promote lifelong learning, cultural expression and personal fulfilment. These principles can be used to create the social, economic and cultural benefits for both individuals and society as a whole.

Going forward, when referring to 'literacy' in this study, we will use the 1978 UNESCO definition for functionally literate people (see above). The advantage of this definition of literacy is that it is dynamic and open to adjustment based on changes in society.

We thus recognise literacy as a dynamic moving target shaped by local, cultural, economic and technological development. In our current data-rich environment it is obvious that the concept of literacy needs to adjust and include the ability to work with data to advance one's own and the community's development. The disadvantage of such a dynamic concept is that it is difficult to compare over time.

2.3.3 Digital literacy

During our research we were asked on several occasions how it was going with the research on digital literacy. We replied that this research is about data literacy, not digital literacy. Though digital and data are often used in tandem, they are not synonymous. Digital is used in many ways relating to digital signals and computer technology, for instance the Internet. But as discussed earlier, when we refer to data in this research, we do actually mean digital data. So how do these two forms of literacy relate to each other?

Digcomp 2.1 is the most recent European Union Digital Competence Framework for Citizens (Carretero et al., 2017); an update is expected to be released in 2022. This framework mentions five competence areas related to digital literacy: information and data literacy; communication and collaboration; digital content creation; safety; and problem solving. The fact that information and data literacy together form the first digital competence in the framework is not surprising, as all other digital competence areas in fact require some level of data literacy to develop, as they deal with sharing, creating, protecting and using data-driven technology. In this sense, data literacy can be interpreted as the fundamental competence that other areas of digital literacy build upon.

In the Digcomp framework, the competences of information and data literacy are: 'browsing, evaluating and managing data, information and digital content'. Analysing this framework, we can disambiguate data literacy and digital literacy in two ways. First, we can establish information and data literacy as one of the five competence areas, 'data literacy' thus being a part of the broader concept 'digital literacy'. Second, we establish that not only data but also data derivatives such as information and digital content are mentioned. Information comes into being when data is combined, synthesised, used, visualised, applied, etcetera. Digital content is in fact a form of processed or artistically created data. We expect someone competent in the field of data to also be competent in handling information, digital content, products, tools and services that were derived from, created or built upon data.

2.3.4 Data literacy

In the 2010s, researchers were increasingly engaged with the concept of data literacy. Wolff et al. (2016) refer to data literacy as:
the ability to use a research process to ask and answer questions from large and small datasets, taking into account the ethical use of data. It is based on practical and creative core skills with

the ability to expand the knowledge of specialised skills in the field of data processing in accordance with the objectives. These skills include the ability to select, clean, analyse, visualise, critique and interpret data, as well as communicate stories from data and use data as part of a design process.

Ridsdale et al. (2015) took an interdisciplinary approach that led to a more concise definition: 'Data literacy is the ability to collect, manage, evaluate and apply data in a critical manner'. All these researchers noted that the definition must be able to change and evolve with the input of the stakeholders; all the definitions share many common understandings about the competences of data literacy.

For our research we largely adopt the definition as posed by Mandinach and Gummer (2013), who identify eight broad competencies including data identification, acquisition, management, evaluation, processing, analysis, interpretation and ethical use. These can evolve over time, as the role of data in society also changes over time. New data technology will require novel tools, governance and ethical considerations, but the underlying core competencies will remain relevant. We agree with Wolff et al. (2016) that communicating data is crucial to data literacy.

In this research, we define data literacy as follows:

The ability of an individual to identify, acquire, manage, evaluate, interpret, apply, use and communicate data in a critical, ethical and responsible manner to reach a certain goal.

Yang (2020) operationalised the concept of data literacy in higher education based on the work of Prado and Marzal (2013) and of Maybee and Zilinski (2015). The researchers analysed and synthesised competences of data literacy from these existing frameworks and translated them into education goals (see Table 1). We emphasise six core competences: (1) understanding and awareness of data; (2) access to and acquirement of data; (3) engaging in data; (4) planning for and managing data; (5) synthesising, visualising and representing data; and (6) using data properly and ethically. Sub-competences associated with the core competences further define the important fundamentals of data.

Table 1. Scheme of data literacy instructions
(Adapted from and Prado and Marzal (2013) and Maybee and Zilinski (2015))

Core competences	Specific competences	Descriptions
Awareness	What is data?	Learners need to know what is meant by data and be aware of the various possible types of data.
	Data in society	Learners need to be aware of the role of data in society, how they are generated and by whom, and their possible applications, as well as the implications of their use.
Access	Data sources	Learners need to be aware of the possible data sources, be able to evaluate them and select the ones most relevant to an informational need or a given problem.
	Obtaining data	Learners need to be able to detect when a given problem or need cannot be (totally or partially) solved with the existing data and, as appropriate, undertake research to obtain new data.
Engage	Reading and interpreting data	Learners need to be aware of the various forms in which data can be presented (written, numerical or graphic), and their respective conventions, and be able to interpret them.
	Evaluating data	Learners need to be able to evaluate data critically based on data evaluation criteria (including authorship, method of obtaining and analysing data, comparability, inference and data summaries).
Manage	Data and metadata collection and management	Learners need to be aware of the need to save the data selected or generated and of descriptive or other data associated therewith, for due identification, management and subsequent reuse.
	Preserving data	Learners need to be aware of curation practices for long-term storage and use.
Communicate	Producing elements for data synthesis	Learners need to be able to synthesise, visualise and represent the results of data analysis in ways suited to the nature of the data, their purpose and the audience targeted in the inquiry.
Use	Data handling	Learners need to be able to prepare data for analysis, analyse them in keeping with the results sought and know how to use the necessary tools.
	Ethical use of data	Learners need to make ethical use of data, acknowledging the source when obtained or formulated by others, and making sure that used methods are deployed and results interpreted transparently and honestly.

We adopt this scheme of data literacy instructions to examine the data literacy of individuals, as it fits the purpose of this reconnaissance and our working definition of a data literate person.

2.3.5 Demand for data competences in the labour market

The possibilities of billions of people connected by mobile devices, with unprecedented processing power, storage capacity and access to knowledge are unlimited. These possibilities will be multiplied by emerging technology breakthroughs in fields such as artificial intelligence, robotics, the Internet of Things, autonomous vehicles, 3-D printing, nanotechnology, biotechnology, materials science, energy storage and quantum computing. All of these data-driven technologies are described by Schwab (2017) as marking the current (fourth) industrial revolution.

The World Economic Forum (2020) mapped out the impact of the fourth industrial revolution on the labour market. In addition to the potential scale of layoffs, the World Economic Forum identified strategies for a successful transition from shrinking to growing jobs, roles and sectors. Under the influence of the Covid-19 pandemic, the underlying trend in the technological augmentation of labour has accelerated even more.

The Future of Jobs Report 2020 (World Economic Forum, 2020) indicated that the pace of technology adoption (including the cloud, big data, and e-commerce) will continue to be high over the next five years. Employers expected that by 2025 50% of all work will be carried out by machines and 50% by humans. More than 97 million new jobs will be created, while 85 million redundant jobs will disappear due to the shift in the division of labour between people and machines. Lund et al. (2021) from McKinsey confirms this in their Future of Work after Covid-19 report; they see companies enlisting automation and AI to cope with Covid-19 disruptions, adding to these already surging trends.

Skills that will become more important in the next five years, according to World Economic Forum (2020), are critical thinking, analysis, problem solving and self-management skills, such as active learning, resilience, stress tolerance and flexibility. These skills are essential for employees of tomorrow in the green economy, data economy, AI economy, technology, cloud computing and product development. Over eighty percent of employers are on the verge of digitising work processes in the short term, including a significant shift to remote working, partly prompted by the Covid-19 pandemic.

It is expected by the World Economic Forum (2020) that 40% of the core skills for employees who remain in their roles will change and that 50% of employees will need new skills within five years. Employers offer new skills development options to 70% of employees, but only 42% make use of them. Companies need to invest more in social cost-benefit analyses, because upskilling or reskilling employees gives a broader and deeper result than can be represented by a company's net profit.

We expect to see these worldwide trends also in the regional labour market. To get an idea of the demand for data literacy in business vacancies in Rotterdam we did a local search on a national Dutch job website 'De Nationale Vacaturebank' on September 10th, 2021. This yielded a list of 308 jobs at the 'HBO' level (meaning those who have graduated from vocational universities) in which data is a part of the job as per the mentioning of 'data' in the job description. We used the following filters: Rotterdam, HBO, Commercial, Consultancy, Finance, HR, Marketing, and Procurement, with the search word 'data' in the location 'Rotterdam'. At the time, a total of 1684 relevant job openings at the HBO level were found, or 18% of the openings on this job website in the economic domain in Rotterdam had data as a component of the job. We expect that this finding can be generalised to the job market throughout the Netherlands; if that is true, a substantial proportion of vacancies require data literate young business professionals to fit the needs of the organisation hiring them. Therefore, we expect that jobseekers who have a higher level of data literacy will have substantially greater chances of being selected for these openings, and that they will then be able to help organisations in their digital transformation. Furthermore, not all businesses may yet recognise the need for data-savvy staff, and may therefore not mention related competences in their job openings yet.

2.4 Summary of Study 1

We depart from UNESCO's definition that people need functional literacy to function effectively and develop one's community. We thus recognise literacy as a dynamic moving target shaped by local, cultural, economic and technological developments. In Study 1 we recognise that society and businesses are becoming more and more data-rich and data-driven, and thus that data competences need to be considered in a concept of data literacy. Also, higher education needs to cater to this surging demand in society and business by developing data literacy among the current generation of students. By 'data' we mean raw digital data that can be processed with and within computers. The use of this data can lead to advantages for citizens, businesses and society. We distinguish 'information-driven', in the form of KPIs and spreadsheets, from 'data-driven', which is powered by raw, real-time data about events and processes.

Based on our findings, we define data literacy as follows: the ability to identify, acquire, manage, evaluate, interpret, apply, use and communicate data and data products in a critical, ethical and responsible manner to reach a certain goal. We also propose a competence scheme that aligns with our definition and which can be used in educational settings to assess students and lecturers on their data literacy competences.

3. Study 2: Current Promotion of Data Literacy in Courses

3.1 Introduction

We explored the data literacy concept in Study 1. We now have a clear definition of what data literacy is, and we have also illuminated the importance of data literacy competencies for society, business and education. Because of its importance, it needs to be given greater attention in education, and courses need to be focused on these competencies. In this second study we attempt to find out if data literacy is promoted in the study programmes offered specifically at the at the RUAS Business School. In turn, the following research question is presented:

To what extent is data promoted in current courses at the RUAS Business School?

In this study we explore which programmes offer courses that are expected to increase students' data literacy, and to what extent. To promote data literacy, one would expect to find courses on data-related topics such as databases, data analytics, data management, data governance, data ethics, data strategy and data science.

3.2 Research Method

This study adopted content analysis. The RUAS Business School does not have a structured indexed catalogue from which we could create a crosstab of courses and subjects for all programmes; therefore we looked for an alternative. All programmes need to register their courses in Osiris, which is the education management database used by RUAS. In Osiris this section of the database is referred to as the education catalogue. As mentioned in the introduction, we expected that courses which cover data-related topics would mention 'data' in their descriptions. We analysed the content of the education catalogue in search of the word 'data' in the names of the courses, the learning goals and course descriptions to see if, where and how often we literally found the word 'data' used either as a noun or as an adjective. We considered also including the Dutch word 'gegevens' and maybe other keywords, but we expected this would result in too many hits and manual filtering for this exploratory project, so we decided to only include 'data'.

We also considered analysing all curricula of all study programmes and interviewing their course managers, but because this is an explorative study into data literacy we only used this quantitative content analysis to see how many courses mentioned 'data' in the course database, as this was considered the most time-efficient approach. There were instances in the dataset where 'data' was used in Dutch to point to the plural of 'date' instead of the topic 'data' (for

instance the word 'inleverdata' – 'deadlines'); these were removed to prevent invalid results. In this study we tried to assess the status quo as of December 2020.

The catalogue was analysed by counting how often the word 'data' occurred in certain attributes within the course table. The attributes that lent themselves best to this analysis were:

- name of the course
- learning objectives
- content description

The RUAS Business School education catalogue in the digital Osiris education system has a limit on how many results can be displayed or downloaded, so we had to get our data via an Osiris superuser. We received a download of the catalogue for 2020 in an Excel sheet on December 10th, 2020 from one of the school's managers. This dataset was filled solely with courses that had been updated for the year 2020, to prevent using outdated data. Based on the course codes (unique codes that distinguish the courses uniquely from each other), we extracted the descriptions of each course. We looked at how often the word 'data' appeared in the names, learning objectives, and content descriptions of the courses in the catalogue.

Before we continue with the results, we need to acknowledge that the data quality is very low. After analyses of the data, it became clear the data was incomplete in two ways. First, the list of courses we received was incomplete. We learned that not all courses are updated each year, so courses that were not updated were not included in our dataset for 2020. Second, it became clear that while course codes in Osiris are needed to centrally register results, learning goals and course descriptions are not necessarily stored in Osiris. Programmes can use other systems to store this data and disseminate content to their staff and students. So there may well have been courses that offer data literacy competencies which did not show up in our search because the information was posted elsewhere.

These two reasons for incomplete information led to lots of missing values in the course catalogue, perhaps falsely indicating a lack of a data-intensive courses at several programmes. In future research the data quality should be improved upon or another research method should be employed. In this study we discuss the issues surrounding data quality and then discuss the research findings, bearing in mind these data quality issues.

Completeness: More than 30% of the learning objective fields were empty or filled with a generic text such as 'follows' or 'see manual on Cumlaude' (Cumlaude is the almost retired learning management system used to publish course materials for students) or simply 'see manual'. Maybe it would be possible to connect these systems and databases to get a more complete picture. This would require access to all these systems, which could prove difficult in practice, as it appears that each of the fifteen programmes govern their own curricula and the related data. This would also require technical assistance from systems experts who could create the needed

queries on this data. It seems there is no faculty-wide data policy which allows for the time efficient comparison of courses, their learning goals or the subjects they cover.

We know we are missing data because some programmes and minors use other systems. For instance, the part-time Business programmes run through the Career Academy register course descriptions via BrightSpace. When we asked one of the lecturers at the Career Academy about their curriculum on Oct 11th, 2021, we were provided with the following information:

> The Career Academy provides large semester modules (30 ECTS). At least 6 of the total of 25 modules are very much data-driven and infused with new technology. These data-driven modules are:
> - Data Driven Organiseren (Business, IT & Management),
> - HR Insight (Human Resource Management),
> - Business Intelligence & Control (Bedrijfskunde, Finance & Control),
> - Business Insight (Commerciële Economie),
> - Artificiële Intelligentie (an elective in Business, IT & Management),
> - Digitale Transformatie (an elective in the Commerciële Economie track).

The data intensity at the Career Academy was not shown in our content analysis in Study 2 because the Career Academy registers course content in a separate system called BrightSpace.

Another important programme with missing data is the full-time Business IT & Management (BIM). When we asked one of the lecturers at BIM about their curriculum on Jan 12th , 2021, we were provided with the following information:

> At BIM we now have a course guide per block of 15 ECTS, no longer per course. We have two new course blocks ready now for the first semester of the new academic year:
> - Getting Started
> - On the Move
>
> Both course blocks are full of data topics like data collection, databases, data modelling, data communication and awareness of the value of data.

The data intensity at BIM was not shown in our content analysis in Study 2 because at the time of data collection these course blocks were not yet updated or registered in the Osiris catalogue dataset we received.

The same is true for some minors within the RUAS Business School that are focused on the theme 'digital' which were not part of our dataset but which are all probably data-intensive, among which 'Marketing in an online World', 'Privacy, Security, Risk', 'Digital Marketing', and two new minors 'Ethics & Digitalisation' and 'Data-driven Solutions'. The programme Finance & Control also uses an alternative system, 'Acte', to register their course descriptions, so they were not

considered in this part of the analysis. Improving on this is beyond the scope of this research endeavour but will hopefully picked up on by faculty management.

Accuracy: The data that was available (not empty or referring to other systems) at first appeared to be accurate. We took a random sample of 5 from our dataset of 465 records and used manual inspection to see if the description and learning goals were logically aligned to the respective courses. For these five courses we also checked if the responsible course manager/lecturer registered in Osiris was indeed working for the respective programme, and we found no anomalies.

Uniqueness: On further detailed analysis we found multiple instances of (partial) duplicates; on several occasions there were multiple course codes in 2020 referring to the same course. We assume this is due to mistakes or errors found after registration in the course database, but after results are registered on a course, the content cannot be changed or deleted, so a new version of the course must be inserted into the database. The duplicate courses that had unique codes but identical content would have created invalid results, so these were deleted from the dataset.

Reliability: The reliability of the data is low. It seems that agreements were only made on the length and positions of the course codes, as in most of the records the course codes display some structured features. In 81% of records in our dataset the course code starts with three letters for the programme abbreviation, then three letters for the course name abbreviation, and finally four positions to indicate the place of the course within the four-year Bachelor programme.

There does not appear to be guidance, control or governance on the data regarding learning goals and course descriptions. Through manual visual inspection we discovered that name, learning objectives and content are so called 'open fields', which prevents using comparable pre-set values. This creates variation because hundreds of course managers have different interpretations of how to describe the name, learning objectives and content of a course. Staff probably fill in these fields only according to what they think is needed by their own colleagues and students, not according to what would be needed to make it possible to perform school-wide analyses of these courses. It seems necessary to calibrate the requirements for the input, for example in what detail and by which standards course goals and descriptions should be described.

3.3 Research Findings

We analysed the data at three levels, namely the course level based on name of course, the level of learning objectives based on entered description of learning objectives, and the level of the content based on the course description. Table 2 shows that a total of 465 courses were registered in our dataset (i.e., the courses in the RUAS Business School's 2020 education catalogue), with a total of 3370 ECTS (European Credit Transfer and Accumulation System). In the

dataset we found 433 courses with 3114 ECTS pertaining to regular programmes and 32 courses pertaining to minor programmes.

Table 2. Overview of the courses in the RUAS Business School's 2020 course dataset derived from the Osiris education catalogue

Name	Course count	ECTS	'data' in course name	'data' in learning objectives	'data' in course description	'data' in ECTS	% ECTS with 'data'
Programme							
Niet opleidingsspecifiek (Not programme specific)	41	362	0	0	0	0	0.0%
Accountancy	33	196	1	1	3	14	7.1%
Bedrijfskunde (Business Administration)	24	167	0	0	0	0	0.0%
Business, IT & Management	2	4	0	0	0	0	0.0%
Commerciële Economie (Commercial Economy)	26	160	0	0	3	24	15.0%
Creative Marketing & Sales	45	238	0	3	6	18	7.6%
Finance & Control	76	339	1	0	0	0	0.0%
Finance, Tax & Advice	12	41	0	0	0	0	0.0%
Financial Services Management	1	2	0	0	0	0	0.0%
Global Marketing & Sales	55	305	0	0	7	68	22.3%
Human Resource Management	22	200	0	0	0	0	0.0%
International Business Administration	16	46	0	0	0	0	0.0%
Management, Economie & Recht (Management, Economy & Law)	7	28	0	0	0	0	0.0%
Marketing of Social Business	37	174	0	0	2	10	5.7%
Entrepreneurship & Retail Management	11	102	0	0	0	0	0.0%
Career Academy	25	750	0	0	0	0	0.0%
Programme Total	***433***	***3114***	***2***	***4***	***21***	***134***	***4.3%***
Minor							
Niet opleidingsspecifiek (Not programme specific)	24	210	0	4	9	94	44.8%
Commerciële Economie (Commercial Economy)	8	46	0	0	0	0	0.0%
Minor Total	***32***	***256***	***0***	***4***	***9***	***94***	***36.7%***
Total	**465**	**3370**	**2**	**8**	**30**	**228**	**6.8%**

Again, it should be stressed that lots of data are missing from the used dataset, the data presented here indicates which programmes had updated courses in 2020 at the time of data collection. Also, this data may indicate unknown variables that were or were not selected in Osiris during data collection that affected which data was returned by the Osiris query. We aim to state the status quo as a starting point for future research.

Before we move on it's important to clarify the structure of programmes at the RUAS Business School. A full regular Bachelor programme consists of 240 ECTS, which is usually divided into 8 semesters, all accounting for 30 ECTS. A minor consists of 30 ECTS and is a regular part of the Bachelor in the last phase, but students can choose their minor by preference, and the programme provides room for this, traditionally in the 7th semester. Commerce programmes work with a community structure, where the first two years are small scale, subject-oriented communities and in the last two years students pursue a generic Commerce programme, which is the same for all Commerce communities.

3.3.1 Course level – Names of the courses

Among the 465 courses examined, only two mentioned data in their course name. Analysis of the course names shows that the study programme 'Accountancy' has one course that includes 'data analytics' as part of the course that has a total of 5 ECTS, and the study programme 'Finance & Control' shows one course 'Database Design' for a total of 2 ECTS (see Figure 1).

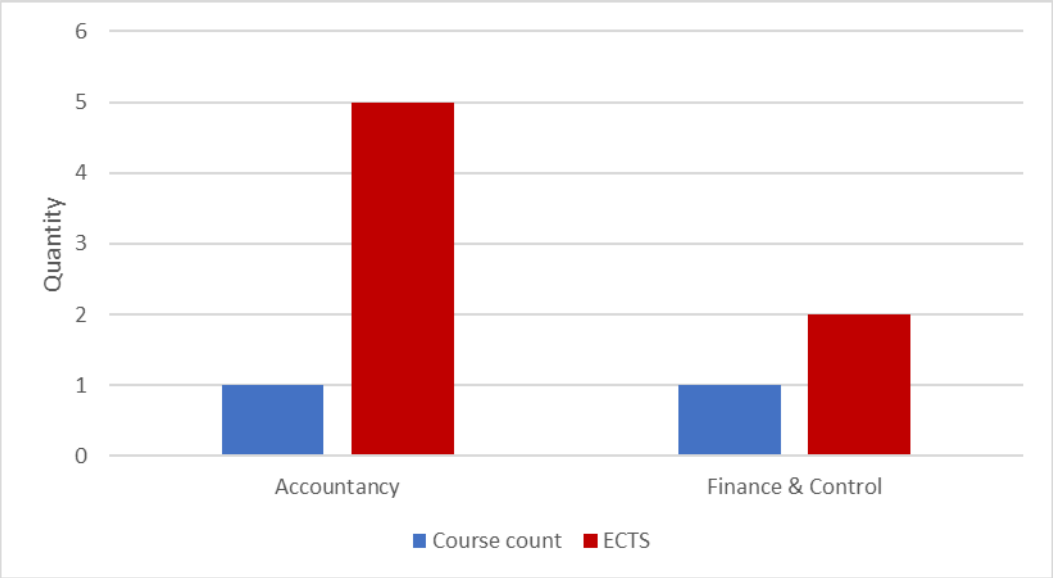

Figure 1. 'Data' in course names

Both courses were offered in the second year of the Bachelor programme. Perhaps from the former Institute for Financial Management (IFM)[1], it seems logical that Finance programmes are the focus for 'data' because of the importance of currency and data reporting within Finance, as compared to Commerce, Business Administration & Management (the current HRBS component of the RUAS Business School consists of the former institutes IFM, IBK & COM that were merged into HRBS in 2019). We conclude that two out of fifteen programmes show data in some of their course names.

[1] HRBS now consists of the former institutes Instituut voor Financieel Management (IFM), Instituut voor Bedrijfskunde (IBK) & Instituut voor Commercieel Management (COM), that were merged into HRBS in 2019. In future the English programmes from the internationally oriented Rotterdam Business School (RBS) will also be merged into HRBS.

3.3.2 Learning objectives - Entered descriptions of learning objectives

Among the 465 courses examined, eight mention data in their learning objectives, aggregated into two programmes and two minors. Four of these minor courses add up to two full minors, namely 'Business Model Innovator' and 'Fieldlab Digital Economy', each consisting of 30 ECTS. Minor programmes were tied to specific programmes in the past but are now in principle independent of any specific education programme.

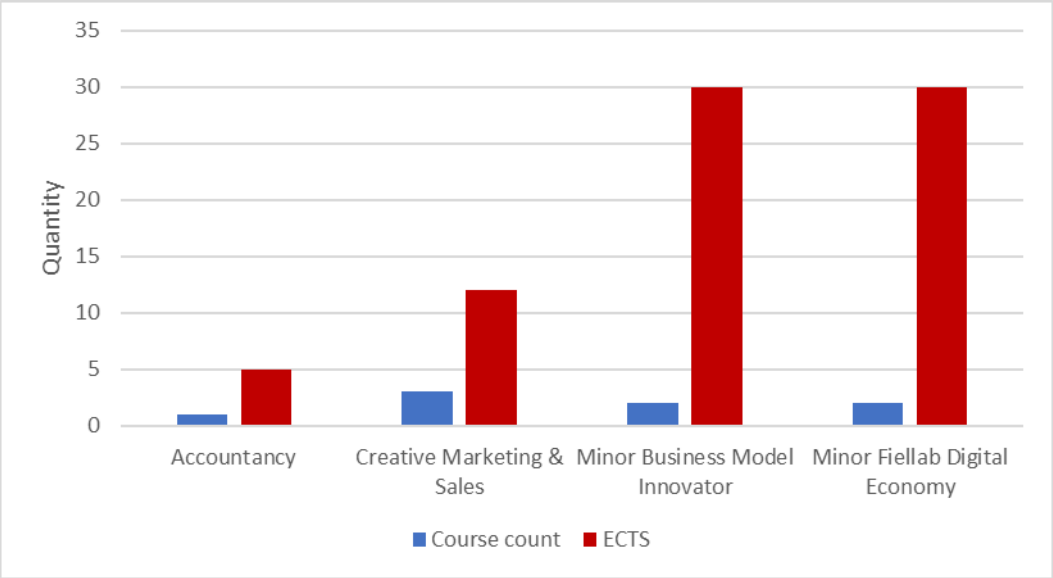

Figure 2. 'Data' in learning objectives

The fifth course is 'data analytics', mentioned in the former paragraph in the 'Accountancy' programme. The last three courses were connected to the Commerce community 'Creative Marketing & Sales'. The three courses have unique and attractive names: 'Draw your conclusion', 'Keep them', and 'Sensing'. According to the description, these courses revolve around research skills, such as formulating research questions, collecting data, (data) analysis and statistics, with a focus on digital commercial data and analytics for first- and second-year students, totalling 12 ECTS. None of the other courses in the dataset show data as part of their learning objectives.

It is striking that while 'Finance & Control' (the name of a programme) was found at the level of course name, 'data' did not appear in the learning objectives for any courses in that programme. One would expect that a term that is found in the course name would also be found in lower levels like learning goals and course description. The reason for this is that 'Finance & Control' is one of the programmes that register course goals and descriptions in a separate system called Acte. The reason that they chose another system was that it allows them to take a more data-driven approach within the programme's curriculum. As this is not a faculty-wide initiative, there are adverse effects for data analysis on a broader scale such as the RUAS Business School as a whole.

3.3.3 Content level - Course descriptions

Among the 465 courses examined, 31 courses from 5 programmes and 4 minors mentioned data in the course descriptions, for 279 ECTS in total. For the programme 'Accountancy' three courses were listed that deal with financial data, among which was one that also was mentioned in previous paragraphs, totalling 14 ECTS for years 2 and 3.

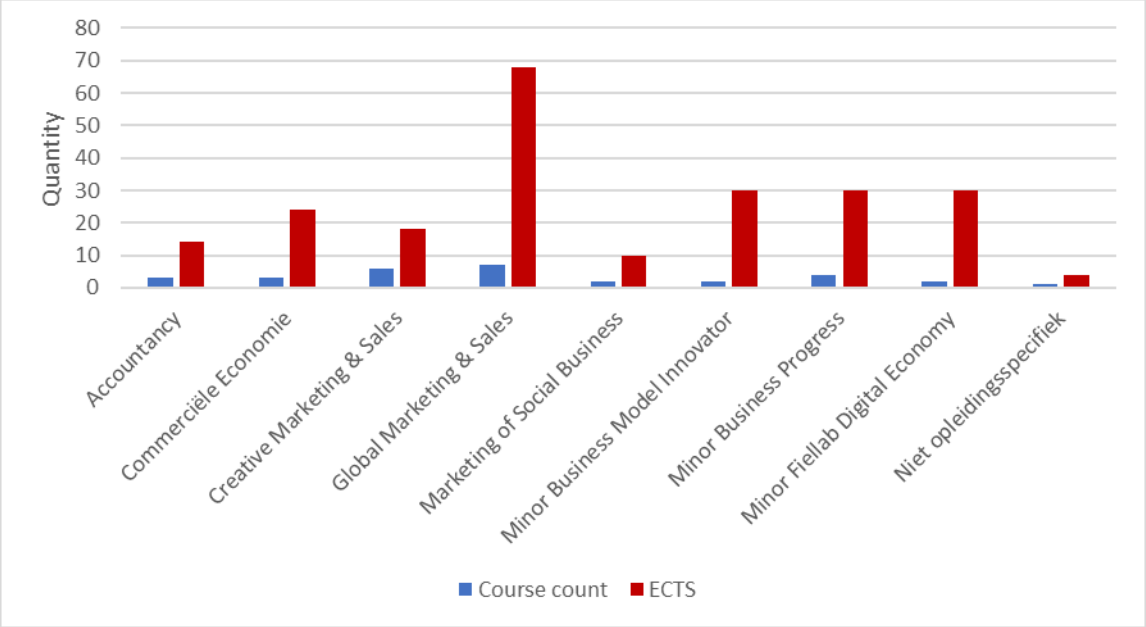

Figure 3. 'Data' in course descriptions (see Table 2 for translations)

Apart from the minors and generic courses, the rest of the courses that mentioned data in their course descriptions were all tied to commerce programmes. The programme with the most 'data' within its course descriptions was 'Global Marketing & Sales'. Eight courses totalling 98 ECTS were found during the analysis, which is approximately 40% of the total ECTS required for a full Bachelor programme. In these eight courses the following data-related subjects were touched upon:

- analysis of commercial data
- the use of 'big data'
- automated data driven CRM systems
- data analysis
- SEO/SEA data
- data management
- consumer behaviour data
- creating value from data

In our analysis of the course descriptions, the programme 'Marketing of Social Business' revealed two courses totalling 10 ECTS that included topics on (customer and supplier) databases and general research. The programme 'Creative Marketing & Sales' revealed two more courses in which data was used in general research competences, apart from the three already mentioned in the former paragraph, totalling 18 ECTS.

The general programme 'Commerce' showed three courses, totalling 47 ECTS, that included 'data' in their course descriptions. These are all third and fourth year courses that build on the community programmes, 'Marketing of Social Business', 'Creative Marketing & Sales', and 'Global Marketing & Sales'. These Commerce courses entail internships and thesis preparation, and general research competencies in the field of Commerce are central to each. The course descriptions specifically mentioned literature databases and data processing, analysis and building conclusions on the analysis. This makes the Global Marketing & Sales programme a very data intensive programme, combining the community programme and generic Commerce programme with 145 ECTS out of a total of 240 ECTS, so more than 60%, relating to data.

Apart from these programmes, four minors showed up as having 'data' in their course descriptions. The two not mentioned in the previous paragraphs are the 'WO Doorstroom' minor and 'Business Progress' minor. The WO Doorstroom minor had one course totalling 4 ECTS. According to the description this is a market research course heavy in data analytics and statistics. The Business Progress minor is the last in this list, having 4 courses totalling 30 ECTS that mention data, namely KPIs and data analytics.

3.4 Summary of Study 2

We must be cautious in drawing conclusions from this study because of the data quality issues mentioned earlier. Further research is needed with better quality data to assess which courses in which programmes could and do promote data literacy.

For the education programmes, we conclude that 5 out of the 12 programmes present in the Osiris dataset that we used showed 'data' in their registered courses. If we look at the ECTS involved, we see that 230 ECTS out of 3370 ECTS of all courses in our dataset contain the word 'data' in the sense of data literacy, which is 6.8% of the total ECTS volume. This is an indication of how much data is promoted in the respective programmes and of the differences between programmes. The Commerce programme 'Global Marketing & Sales' appears the most data-intensive programme based on our data, with 60% of registered courses containing the word 'data', where some other programmes show none.

The third and fourth year programmes grant students opportunities to integrate data-intensive programmes into their personal curricula by way of minors, internships and graduation projects. As we see in this study, there are several minors at the RUAS Business School that promote data literacy.

4. Study 3: Perceived Data Literacy of Students and Lecturers

4.1 Introduction

This study was exploratory and attempted to examine the status quo on students' and lecturers' data literacy, in the RUAS Business School.

We are interested in the general data literacy level of students, and the differences between study programmes and study years. As society and business become more and more data driven, the labour market will require increasing levels of data literacy in new hires. These insights can help to assess if students at the RUAS Business School will be able to satisfy this increasing demand. We can also use this benchmark in future studies to see if students are progressing and if investments in education to promote data literacy are paying off.

We are also interested in the general data literacy of lecturers, because without data literate staff it will be challenging to promote data literacy in our education. We hope to see if there are differences among programmes and also among lecturers in their core competences of data literacy. We furthermore want to know if lecturers think their data literacy is sufficient for their job and if they are open to improving their data literacy through professionalisation. These insights can help us to assess if and how we can use data literacy and assessment thereof in higher education. In turn, the following research question in this study is constructed:

What is the status quo on students' and lecturers' data literacy in the RUAS Business School?

4.2 Research Method

We carried out an online questionnaire survey via MS Forms to examine the perceived level of students' and lecturers' data literacy within the two faculties of the current RUAS Business School: HRBS and RBS. We used the scheme of data literacy instructions (see Table 1) to operationalise data literacy in 23 items divided over the 6 core competences on a five-point Likert scale ('very much agree' to 'very much disagree').

In addition, we asked what the participants' primary education programmes were. For students, we asked what year they started their study. For lecturers, we added two questions:

- 'Do you consider your personal data literacy sufficient to perform well in your job?'
- 'If you were compensated for your time, would you like to improve your data literacy?'

Because of privacy regulations we were not allowed to mass-mail the 12,141 students of the RUAS Business School, so we asked several staff members to invite their students for the survey. This introduced a bias, but was still the best way to get results. We also promoted the survey via the RUAS intranet. We set up a bilingual survey, as we have programmes in Dutch (HRBS) and English (RBS), as well as Dutch and international students whose mother tongue is not Dutch.

We performed a pre-test with students and colleagues. Some ambiguities in the survey questions were addressed, and some items were split up so we could get more valid results by addressing only one issue per question.

4.3 Research Findings

4.3.1 Students' perceived data literacy

We received 115 valid responses from 12 study programmes between June 14th and July 12th, 2021. Table 3 exhibits the overview of these responses.

Table 3. Overview of student responses (see Table 2 for translations)

Characteristics	**Frequency**	**Percentage**
First year at RUAS		
2015 - 2016 or earlier	7	6.1
2016 - 2017	7	6.1
2017 - 2018	28	24.3
2018 - 2019	41	35.7
2019 - 2020	12	10.4
2020 - 2021	20	17.4
Study programme		
Accountancy full-time	15 (2.6%)*	13.0
Bedrijfskunde part-time	5 (3.4%)	4.4
Business IT & Management part-time	9 (6.2%)	7.8
Business IT & Management full-time	12 (1.8%)	10.4
Commerciële Economie part-time	1 (0.9%)	0.9
Commerciële Economie full-time	1 (N.A.)	0.9
Creative Marketing & Sales full-time	1 (N.A.)	0.9
Finance & Control full-time	27 (2.8%)	23.5
Human Resource Management part-time	2 (3.2%)	1.7
International Business full-time	32 (1.0%)	27.8
Marketing of Social Business full-time	2 (N.A.)	1.7
Others	8 (0.2%)	7.0
Total	115	100

Note: * The percentage scores next to the number of students from a specific study programme indicates how the entire programme is represented within this sample. These percentages were

small compared to the total number of students within each study programme. Thus, the representativeness of each study programme may be low. As can be seen from Table 3, more than half of the participating respondents followed the study programmes 'International Business' or 'Finance & Control'. A significant percentage (60%) of the respondents started their study at RUAS between the years 2017-2019. This implies that in our sample, the majority were in the last year of their Bachelor programmes and finalising their studies.

An overview of the students' data literacy is exhibited in Table 4, presenting the values of mean, max, min, and standard deviation of each core and overall competency. Appendices I and II illustrate a comprehensive overview of the different core competences based on study programmes and starting years, respectively. The score of the overall competence is the average score of the six core competencies.

Table 4. Overview of students' data literacy

Competence	**Mean**	**Min**	**Max**	**SD**
Awareness	4.1	3.3	4.9	0.5
Access	3.9	3.0	4.5	0.5
Engage	3.8	2.9	4.5	0.5
Manage	3.9	2.8	4.7	0.6
Communicate	3.8	2.8	4.7	0.6
Use	3.9	2.8	4.8	0.6
Overall	3.9	3.0	4.3	0.5

A one-way ANOVA analysis was performed to examine the difference among different study programmes and starting years (see Table 5).

Table 5. Students' perceived level of data literacy

Competence	**df**	**Sum sq**	**Mean Sq**	**F-value**
Based on study programme				
Awareness	11	6.9	0.6	1.7
Access	11	7.5	0.7	1.5
Engage	11	7.8	0.7	1.3
Manage	11	9.6	0.9	1.4
Communicate	11	12.9	1.2	2.0*
Use	11	9.6	0.9	1.5
Overall	11	7.4	0.7	1.7
Based on starting year				
Awareness	5	1.8	0.4	0.9
Access	5	1.1	0.2	0.4
Engage	5	1.3	0.3	0.4
Manage	5	2.2	0.4	0.7
Communicate	5	2.4	0.5	0.8
Use	5	1.9	0.4	0.6
Overall	11	7.4	0.7	1.7

Note: * $p < 0.05$

The results indicate that students in our sample perceive similar, moderate to high levels of data literacy regardless of their study programme or study year. As shown in Table 5, 'communicate', which refers to communicating data, was the only core competence that revealed a significant difference among students in different study programmes ($F(11, 103) = 2.0$, $p < 0.05$). A post-hoc Tukey test further showed that study programmes 'International Business' and 'Finance & Control' differed significantly. These two programmes together represent half the sample, and are therefore the most likely to yield significant results in this analysis. Where finance is more focussed on analytics, a business degree will focus more on management, commerce & human resource skills, which all require interpersonal or 'soft' skills. This may explain why International Business students score significantly higher than Finance & Control students on the core competence 'communicate'.

Figure 4 illustrates the students' perceived level of the overall competences based on study programmes. Within the participating study programmes, students of 'Business, IT & Management' (part-time) formed the group with the highest overall data literacy (Mean = 4.3, SD = 0.4).

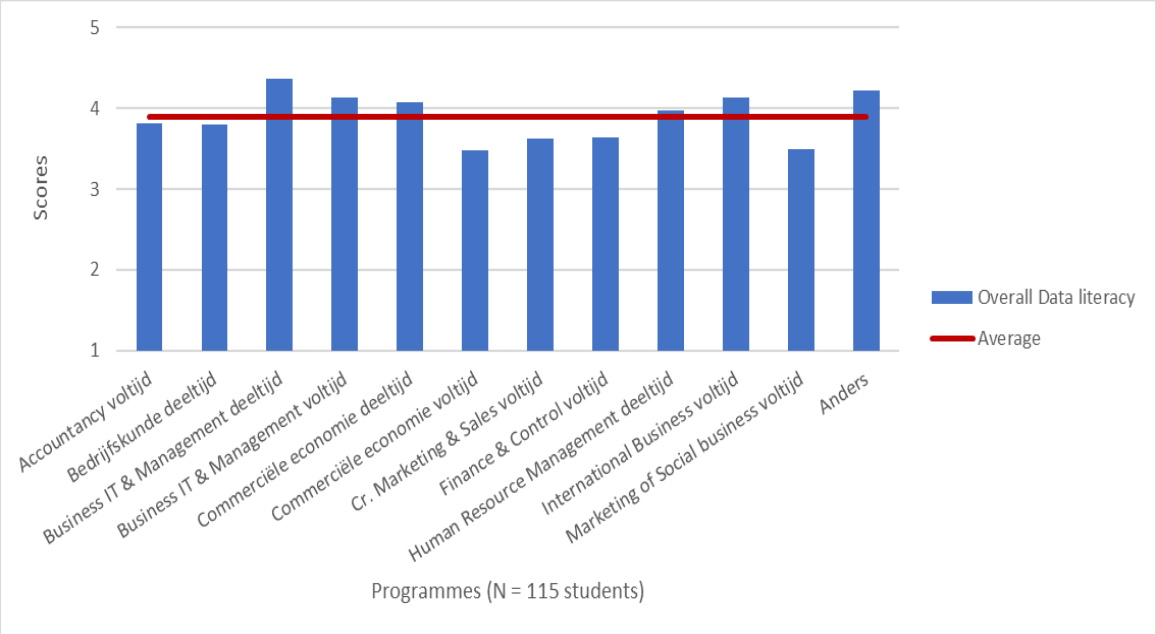

Figure 4. Students' perceived level of overall data literacy based on study programmes (see Table 2 for some translations; voltijd = full-time, deeltijd = part-time, anders = other)

Figure 5 exhibits the students' perceived level of their overall competences based on their start years. It shows that students who started their study in academic year 2018-2019, the 4th year students in this study, comprised the group with the highest overall data literacy. Students who started in the academic year 2016-2017 showed the lowest overall data literacy score. In general, the students who started their studies in years 2018-2019 and 2019-2020 perceived relatively higher levels of data literacy than the others. The group 2018-2019 is a group that potentially contains the largest proportion of students in the sample who studied nominally and who were, at

the time of data collection, ready for their thesis in semester 8. This group had the most and most recent programme experience to strengthen their data literacy.

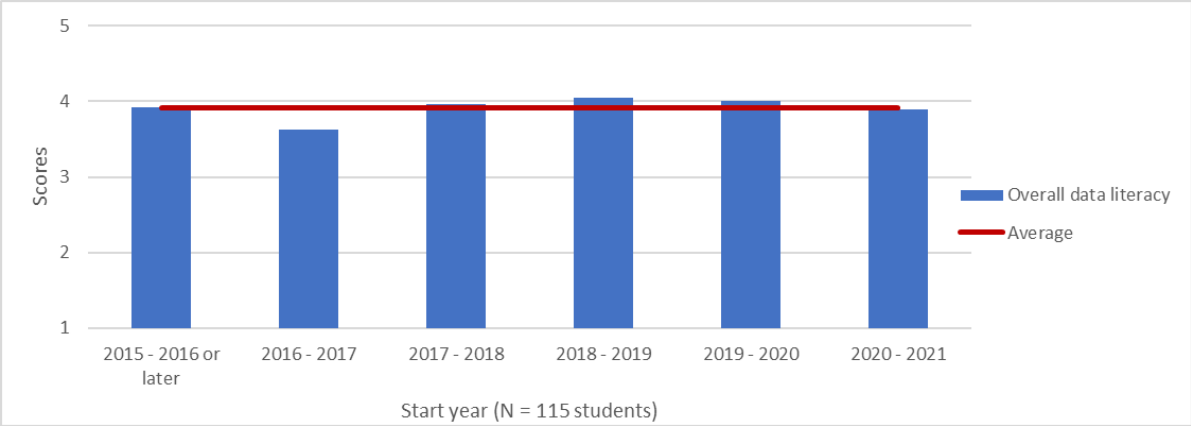

Figure 5. Students' perceived level of overall data literacy based on starting years

A few other factors may be at play here. Students who started their studies in 2017 and earlier were taking more time than the 4 years their study programme nominally provides. This may suggest that they represent a generally less competent group of students who also score lower on data competences. The decrease in recent years may reflect the influence of the Covid-19 crisis and its effects on study results and acquired competences in general.

Further, we examined the difference between full- and part-time study programmes. Table 6 lists the means of students' perceived level of their overall and specific competences between these two programme types. While competence of use was the same between the two programme types, all other competences were different, though not all significant.

Table 6. Means of students' perceived data literacy in full- and part-time programmes

Study programmes	**n**	**Over all**	**Aware-ness**	**Access**	**Engage**	**Manage**	**Com-municate**	**Use**
Full-time	90	3.8	4.1	3.9	4.0	3.8	3.7	3.9
Part-time	17	4.1	4.5	4.1	4.1	4.2	3.9	3.9
t-statistics		0.2	2.6*	1.2	0.5	2.1*	1.0	0.3

Note: * $p < 0.05$

An independent T-test was subsequently carried out to compare the data literacy of full- and part-time students. The analysis revealed that these two groups of students differ significantly for competences 'awareness' ($t(105) = 2.6$, $p < 0.05$) and 'manage' ($t(105) = 2.1$, $p < 0.05$), while there are no significant differences for overall data literacy or the core competences 'access', 'engage', 'communicate', and 'use'.

In addition, the results shown in Figure 6 indicates that part-time Bachelor students, in general, tend to perceive higher data literacy than full-time Bachelor students. This could be due to the

different backgrounds of students in these two programmes; the students who participate in part-time programmes are usually working in administrative or business roles either full-time or part-time during the day, while studying mainly evenings and weekends.

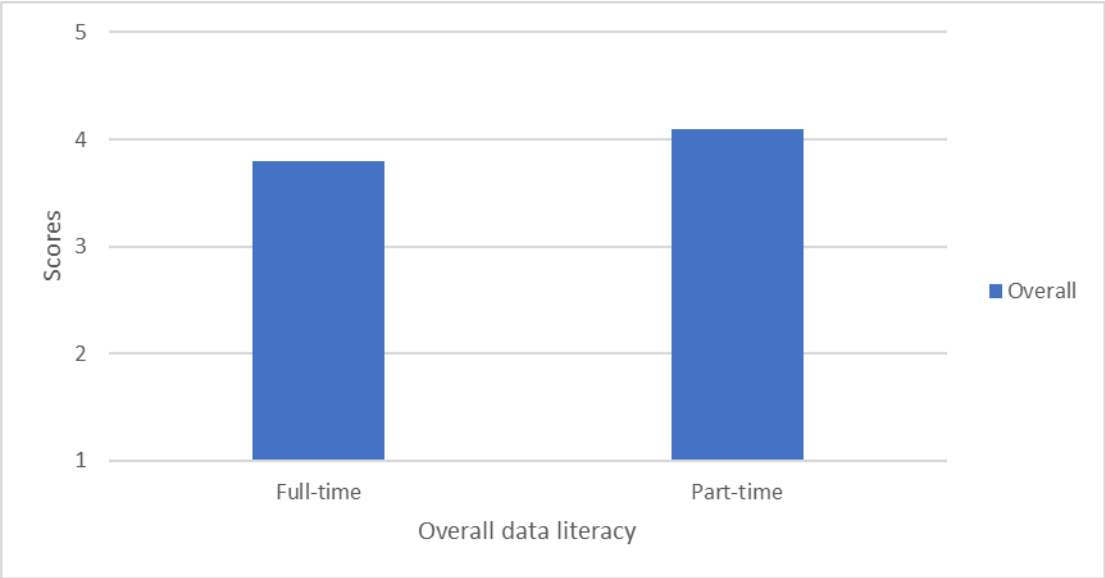

Figure 6. Students' data literacy scores

4.3.2 Data literacy of lecturers

We conducted another online questionnaire survey to examine the self-perceived level of data literacy among lecturers at the RUAS Business School between June 24th and July 4th, 2021. We received 64 valid responses from 15 education programmes. Table 7 presents the overview of these responses.

Table 7. Overview of lecturer responses (see Table 2 for translations)

Characteristics	**Frequency**	**Percentage**
Language		
Dutch	43	67
English	21	33
Educational programme		
International Business full-time	13	20
Ondernemerschap & Retail Management full-time	4	6
Finance & Control part-time	1	2
Finance & Control full-time	3	5
Business IT & Management full-time	6	9
Business IT & Management part-time	1	2
Finance, Tax and Advice full-time	5	8
Bedrijfskunde full-time	9	14
Commerciële Economie full-time	7	11
Commerciële Economie full-time	1	2
Accountancy full-time	4	6
Global Marketing & Sales full-time	1	2
Marketing of Social Business full-time	1	2
Human Resource Management full-time	1	2
Human Resource Management part-time	1	2
Others	5	8
Total	64	100

This overview shows that programmes 'International Business' (English programme, n = 13) and 'Bedrijfskunde' (Dutch programme: Business Administration, n = 9) were the largest respondent contributors. The 'language' overview shows the percentages of respondents that filled in their questionnaire in either Dutch or English. One-third of the respondents filled in the English version. This suggests that international lecturers also participated in this survey.

Table 8 presents the average perceived level of the lecturers' data literacy, including the overall and the six core competences. The score for overall competence is the average score of the six core competencies, ranging from 1 indicating low to 5 indicating high. The participating lecturers perceived their data literacy as ranging from 3.4 to 4.4, which can be described as moderate to high.

Table 8. Lecturers' perceived level of data literacy

Competence	Mean	Min	Max	SD
Awareness	4.4	4.0	4.7	0.3
Access	3.7	3.1	4.3	0.5
Engage	4.0	3.2	4.5	0.5
Manage	3.8	3.1	4.4	0.5
Communicate	3.4	2.5	4.2	0.7
Use	3.6	2.5	4.3	0.7
Overall	3.9	3.3	4.4	0.4

Figure 7 further illustrates that the core competence 'awareness' (M = 4.4, SD = 0.3) has the highest mean, followed by 'engage' (M = 4.0, SD = 0.5). On the contrary, the core competences 'communicate' (M = 3.4, SD = 0.7) and 'use' (M= 3.6, SD = 0.7) were given a relatively lower score. These results indicate that while lecturers are confident in recognising the importance and possibilities with data, they are challenged by using data and delivering the informative output from data analyses.

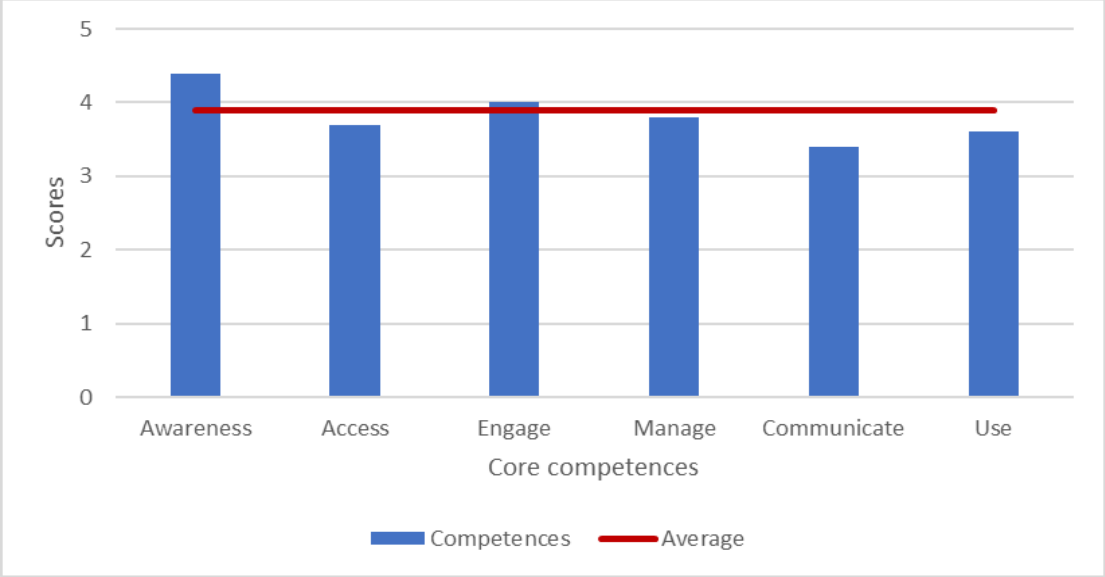

Figure 7. Lecturers' scores on core competences of data literacy

The highest is 'awareness', followed by 'engage', 'manage', 'access', 'use' and 'communicate'. The order is as expected, though we would expect 'access' to be just as high as 'engage', as one needs to access data before one can engage. As data-driven ways of working are all around and mentioned in literature and media, awareness was thus expected to be high. On the other hand, as lecturers are expected to communicate with their students, their 'communicate' competence was likely to be just as high.

An overview in Appendix III presents the lecturers' perceived level of each core and overall competency based on the study programmes. Due to the very limited number of respondents from each study programme, it is difficult to compare the difference between the programmes and draw conclusions.

Next, we reviewed the lecturers' perception towards the sufficiency of their current data literacy for their work and their willingness to improve their data literacy. Table 9 exhibits the questions and the answers in the survey, and the responses from the participating lecturers. The results show that the majority (almost 85%) of the lecturers considered their level of data literacy sufficient or somewhat sufficient for their current work. However, a small group of lecturers (approximately 6%) admitted that they did not have sufficient level of data literacy to fulfil their work needs. The other 9% are not sure if their data literacy is sufficient compared to their work needs.

Table 9. Lecturers' survey responses

Questions and answers	**Frequency**	**Percentage**
Do you consider your personal data literacy sufficient to perform well in your job?		
Sufficient	32	50.0
Somewhat sufficient	22	34.4
Neither sufficient neither insufficient	6	9.4
Somewhat insufficient.	2	3.1
Insufficient	2	3.1
If you were compensated for your time, would you like to improve your data literacy?		
Yes	50	78.1
No	14	21.9
Total	64	100

Subsequently, participants were asked to answer if they were willing to improve their current data literacy level. Table 9 indicates that 78.1% lecturers from our sample of participating lecturers would like to improve their data literacy if compensated for their time.

A simple linear regression test revealed that lecturers' perceived sufficiency of their data literacy for work significantly predicted their willingness to improve ($\beta = -0.59$, $p < 0.05$). These findings

suggest that, a large part of our lecturers are willing to improve their data literacy, and those who perceive a relatively lower level of sufficiency for their work are more likely to be willing to improve.

A multiple regression was performed to examine the influence of lecturers' perceived level of their core data competences on their perceived sufficiency of data literacy for their work and on their willingness to improve. Table 10 displays the test results. It can be seen that the core data competence 'communicate' ($\beta = 0.5$, $p < 0.05$) and 'overall' competence ($\beta = 0.9$, $p < 0.001$) significantly predicted lecturers' perceptions towards the sufficiency of their level of data literacy for their work. It shows that when a lecturer perceived a higher level of competence to 'communicate' data and 'overall' effectively deal with data, she perceived her data literacy more sufficient for her work.

Table 10. Effects of lecturers' perceived data competences on level of sufficiency & willingness to improve

	Level of Sufficiency		**Willingness to improve**	
Competence	***β***	***t-value***	***β***	***t-value***
Awareness	0.2	0.7	-0.3	-1.6
Access	-0.1	-0.2	0.2	1.7
Engage	-0.1	-0.5	-0.1	-0.7
Manage	0.0	0.1	-0.0	-0.2
Communicate	0.5*	2.3	0.1	1.2
Use	0.3	1.6	-0.1	-0.9
Overall	0.9***	7.2	0.0	0.6

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.4 Summary of Study 3

The average perceived data literacy among students in our sample was 3.9 – a moderate to high score – which is more than researchers expected to find based on interactions with students. Because of low sample size, not many significant results were found between programmes, though International Business students perceive significantly higher communication competence than Finance & Control students. Students in their last year of the Bachelor study showed higher perceived data literacy than students who had been studying longer or shorter, though this was not a significant result. Part-time students showed higher data literacy than full-time students across the board, significantly for 'awareness' and 'manage'.

Lecturers' average perceived data literacy was 3.9, the same as for students. Researchers expected higher data literacy for lecturers than for students based on the assumption that students need to learn from lecturers, but the data did not result in a overall difference. There are, however, differences between the mean spread among core competences: the core

competences for students range from 3.8 (engage and communicate) to 4.1 (awareness), whereas lecturers' scores range from on average 3.4 (communicate) to 4.4 (awareness). The fact that awareness scored higher than the other competences was an expected result.

The percentage of lecturers who considered their data literacy insufficient for their job was 6.2%, though these four respondents were all willing to improve on their data literacy. In fact, 78.1% of our total sample population would like to improve their data literacy levels. This probably indicates that lecturers recognise the importance of improving their data literacy skills in an ever-digitising work environment. Therefore, this study encourages the RUAS Business School to offer opportunities to its lecturing staff for improving their level of data literacy.

5. Study 4: How to Prepare Students for SMEs' Demand for Data Literacy

5.1 Introduction

Vocational institutes educate students to prepare them for their future professions. Most of them will come to work for SMEs in the region. In this study we aim to find out how can we best prepare students for their future place in a data-driven labour market. How can we prepare them to be an inspiring data-driven force in an organisation that has yet to achieve data maturity? How do we know if, how and to what degree data literacy is important for SME companies. In this study we explore just this. Therefore, the following research question for this study is put forth:

How can we effectively improve students' data literacy to meet SMEs' needs?

5.2 Research Methods

This study was done by conducting a field experiment in collaboration with the minor programme 'Controlling'. This minor programme contained a project-based practical assignment executed between September 1st, 2020 and February 1st, 2021, called the 'Datawerkplaats' (data workplace) project. In this assignment, students were required to tackle real business questions by taking a data-driven approach. These business questions were provided by this programme's cooperating consortium partners; most of them were SMEs. Students were all in their fourth year and from different disciplines. Thus, Datawerkplaats allowed us to investigate whether the data competences that students had acquired from their studies fulfil the needs of SMEs.

Moreover, we propose that practice-oriented research is an effective way to improve our students' data literacy. The business world is changing rapidly. The ability of formal education to stay up to the minute is under strain. Companies can provide real-world learning experiences to supplement students' formal education and allow them to receive valuable and relevant insights as part of their studies. In addition, based on Study 3, we examined the correlation among the core competences. Results drawn from both students and lecturers demonstrated an imperative correlation among all the competences (see Appendix IV). This indicates that to improve an individual's data literacy, we must take a comprehensive view and approach that allows the individual to experience all the competences simultaneously. Therefore, we argue that it is crucial for students, particularly, before they graduate, to gain hands-on experience of an entire data-driven project.

In this experiment, students participated in a real data project, or part thereof, in a relatively complex context of a running organisation with mostly real data, but with limited time pressure

and less need for cost management to make it fit for an educational purpose. For example, the research project started with a sponsor organisation's idea for innovation of its business process. A group consisted of students from different study programmes, forming a team with different domain knowledge; and the students needed to go through the entire cycle of a data analytics project, from identification of a business problem to insights, and to take into consideration the technical use of data.

The procedure for the Datawerkplaats project is as follows. Students are placed in groups of approximately 5 students. Student groups start with a course on data journalism, in which they get an assignment to replicate a data-driven piece of journalism. In this course they become acquainted with the flow of a data-driven project within a current societal subject like climate, corona, gender, etcetera: the question, the needed data, the collection and cleaning process, evaluation of data quality, analysis and interpretation and communication of findings. After this course they are assigned to an organisation with a real business case about gaining control of a business issue with the help of data, so students needed to employ a data-driven approach. Guest lectures on data topics and finance/business courses are organised on the side to provide input for their assignment.

In this experiment, four companies and organisations provided nine projects. These nine projects were completed by 43 students in groups of four to five students, and two lecturers supervised and assessed the projects to evaluate if and at what level students reached the learning goals.

The five learning goals were:

1) Students can define and evaluate a KPI;
2) Students can collect, transform and clean data;
3) Students can evaluate the quality of data;
4) Students can analyse and interpret data; and
5) Students can communicate findings based on data.

5.3 Research Findings

5.3.1 The change of students' data literacy

During the experiment, we measured the students' data literacy levels before and after they participated in the research project, based on an online questionnaire. Thirteen students participated in the survey. Table 11 exhibits the average scores of the core and overall competences given by the students at the two different points in time.

Table 11. Students' data literacy perceived by students themselves and organisation

Core competence	**Students**			**Org1**		**Org2**	
	Before	*After*	*t-statistics*	*Before*	*After*	*Before*	*After*
Awareness	3.1	4.4	5.54***	3.2	4.7	2.7	4.0
Access	3.0	4.1	5.42***	3.0	4.0	4.0	4.0
Engage	3.4	4.3	2.69*	3.0	4.0	4.0	4.0
Manage	3.2	4.2	3.08**	3.0	4.0	3.0	4.0
Communicate	3.2	4.1	4.11**	3.0	4.0	3.7	4.0
Use	3.2	4.2	3.09**	2.8	4.0	3.3	4.0
Overall	3.2	4.2	4.97***	3.0	4.1	3.1	4.0

Note: * p < 0.05, ** p < 0.01, *** p < 0.001

The dependent t-test demonstrated a remarkable improvement in all of the competences. Figure 8 further illustrates the improvement in each area of competence.

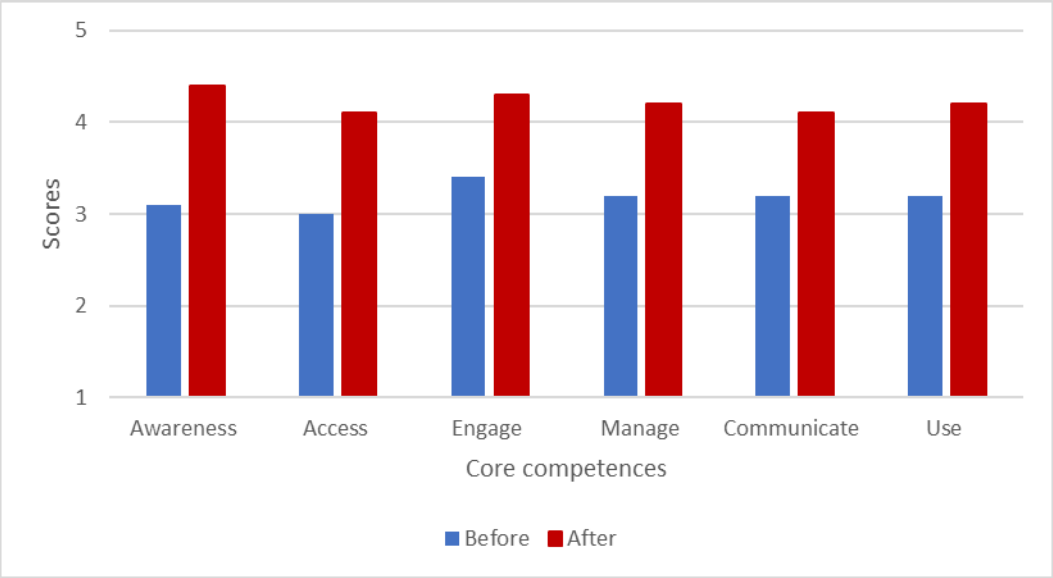

Figure 8. Change in student's perceived data literacy levels

We also invited the project partners to evaluate the students' data literacy before and after their research project, based on a similar online questionnaire (see Table 11). Although only two organisations participated in the survey, their similar evaluations indicated the common opinions of the organisations – both organisations gave a much higher scores to the students' competences after the projects.

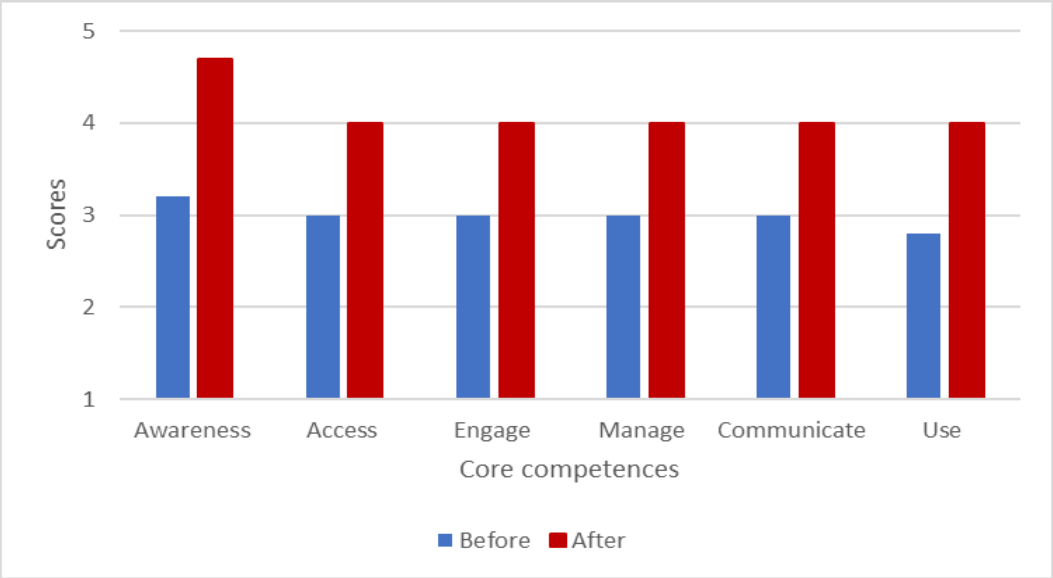

Figure 9a. Change in students' data literacy as perceived by Org1

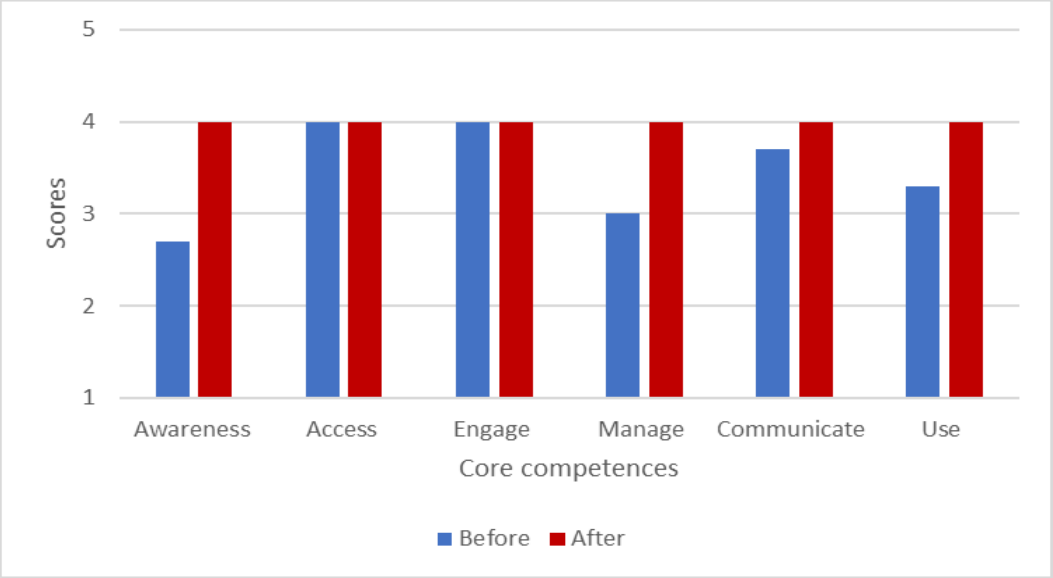

Figure 9b. Change in students' data literacy as perceived by Org2

Figures 9a and 9b illustrate the students' data competences as perceived by these two organisations, respectively. To clarify, the participating companies and students in this questionnaire were not working on the same projects. The analysis here aims to reflect the general perceptions of students and companies.

Some interesting differences were captured. For example, according to Org1, the lack of data awareness was striking at the start of the project. In the first stages of the project student's awareness of data was perceived to be very low, and because of this the awareness of value to be captured with data was also low. Org1 & Org 2, while very different in size, sector and data maturity, both saw similar considerable growth in awareness during the project. Org2 saw even lower data awareness at the start of the project than Org1. This can be explained by the fact that

Org2 was a large public organisation. So on top of the lack of awareness of data there was also a lack of awareness regarding how such an organisation operates and the efforts that are required for large organisations to take the data-driven approach. In the latter case, domain knowledge and awareness on different levels had to be built up by students before data awareness could really develop. This could also show that large data-mature organisations have higher expectations, and this high expectation creates lower assessments of students with objectively similar levels of data literacy.

The competence 'communication' shows a difference before the project, as Org1 saw students at level three on 'communicate' before the project, while Org2 saw students almost on level four before the project. Organisation size may be an influence here; smaller organisations like Org1 are typically more informal and require more initiative and creativity from employees, which may not be the focus of business school programmes. Org2 was the bigger organisation, where more formal business etiquette and communication are standard, which also may be what students do get taught in the business school in their first three years.

The assignments for students ranged from creating a data-driven game based on cocktail flavours to creating a tool for deciding whether data sources should be categorised as open data. This underlines the idea that a data-driven approach can be taken in nearly any business, process or development. In totally different assignments in different sectors, organisations see students' overall data literacy as moderate at the start, increasing on average 1 point on our data literacy scale, which translates to a roughly 25% increase after their Datawerkplaats experience. This suggests that across the board organisations have similar ideas about what makes someone data literate. It also suggests that the general idea of a data-driven approach appeals to managers in different assignments, different sectors, private and public, and that data literacy competences generalise reasonably well between domains and sizes of organisations.

5.3.2 Companies' needs for data competences

In the experiment, one group of students conducted semi-structured interviews with two representatives. One of them (R1) represented two companies, namely Org1 & Org4, and the other (R2) represented one company, namely Org3. The purpose of the interviews was to understand their perceptions towards the students' level of data literacy and explore the organisations' specific needs for data competences. In this section, we present the results drawn from these two interviews.

Interview results – Org1 & Org4

R1 works for two companies. The first, Org1, runs an escape room with which people are trained on various topics, such as AVG, Cybersecurity or Office 365. In addition, R1 also represented Org4, which is a multifunctional event location where, among other things, theatre and music performances are organised. In addition to the performances, they also have a bar with an

extensive drinks menu. R1 himself is not yet deeply familiar with the matter of collecting data and translating data into decision making. R1 does not expect students to have a lot of experience with working with data; he thinks it is more important that people are aware of data and the fact that data can be used to steer an organisation. Understanding the value of data is therefore important. Especially in the events sector and the training industry, almost nothing is done with data, while the potential is enormous. If data is used at all, it is still in its infancy and not at all used to its full potential.

R1 does not expect new employees to have a lot of knowledge of programmes such as Excel or Python. They must be experienced but also show signs of being eager to learn new skills. Org1 thinks it is more important that people have insight and see things. 'As a student who has graduated with a higher professional education, it is important that you know what you can do with data. You could teach everyone skills such as Excel.' Knowing when data are contaminated (low quality) but can still be used is an example of a skill that is more difficult to learn. In addition, it is important that with the available data, a context is outlined around and correct conclusions are drawn.

Within Org1 there are no specifically data-related functions. R1 does not really expect to hire a pure data analyst but instead expects that people who work for him have some knowledge of data. This allows employees to better combine the different aspects of the work and give better feedback after a run of the escape room. Clients see that data is becoming increasingly important, also within the events sector. An organisation that specialises in advising on how events can function more efficiently, for example, would surely attract clients. R1 finds it important that employees can think of ways to use data to organise events more efficiently by conveying information in the right way.

When hiring new employees, particular attention is paid to whether the candidates have experience working with data. It may even be decisive for his choice. This also applies to working with freelancers. Working with data is not yet an issue in the sector in which Org1 operates, yet people quickly underestimate the value of data. If you have information in advance about the group you are going to train, you can steer much better, according to R1.

Finally, R1 indicated that his experiences with working with students so far had been quite positive. He also indicated that in the future, if he were to participate in a similar project again, he would give the students a little more guidance. In addition, students spent too long cleaning up data, while much of what the Org1 does is with a wink and not everything has to be so very serious.

Interview results – Org3

Org3 is a tech scale-up organisation where a lot of data are used, mainly focused on exchange rates. Employees are therefore expected to have experience working with data. R2 entered the project without too many expectations, since data skills are very dependent on the person. Yet

according to R2, the students were on the low side when it comes to data skills. As an example, R2 indicated that the students only had a low knowledge of Excel, which according to R2 is a kind of entry-level programme. To work really well with data, students need experience with more than just Excel, for example Python and R. With these programmes, the students had zero experience. In addition, working with data is also very much about the experience that a person has, and that was also very low among the students. In addition to a deep knowledge of Excel and Visual Basic for Applications, Org3 also expects new employees to have knowledge of Python or R, databases, open-source analytics and BI tools.

In addition to knowledge of various programmes, Org3 expects skills like insight and problem-solving. Because business economists usually have less experience with these programmes, and the knowledge they do have is usually not deep enough, Org3 mainly hires IT people who have more experience working with data. Org3 has worked with about 60 students in the past, of which about 6 met the expectations that Org3 has. Currently Org3 has no vacancies in the field of data, but they are always looking for self-driving data engineers. Self-management is very important here, the management at fast growing companies has little time to guide juniors. They have to come up with a solution independently for an assignment and implement it. This independence was still lacking among the students, and R2 needed to give too much guidance.

R2 also indicated that the last few years have seen a turnaround in which more and more data scientists are being hired and are needed. These data scientists are very international and often offer their services via the Internet. In addition, people from abroad are often better with data skills than Dutch students, and these people are often cheaper. For example, Org3 works a lot with people from Eastern Europe.

Finally, R2 indicated that data skills are not integrated enough in the schools. Students do learn to work with Excel, but that's often where it ends. In addition, Org3 did not find the minor programme deep enough. He indicated that to make true progress the students would need to fiddle and mess around with data much more than happens within the current minor. Students need to enjoy working with data if they wish to progress; they need to find joy in dedicating hours to mastering their dataset. You don't have to be a trained data specialist, as long as you know where data comes from and how you can get started with it.

5.3.3 Main stakeholders' perspective on students' data literacy

At the end the experiment, we facilitated a focus group, which consisted of students, lecturers and companies, to further investigate these main stakeholders' perception towards students' data literacy and to discuss the extent to which our students' data literacy fulfils the organisations' needs.

The discussion in the focus group was very interesting and confirmed the results from previous analyses. In addition to these insights provided about data literacy, the discussion delivered more information about how participants experienced participating in these practice-oriented, data-driven research projects.

During the focus group discussion, data literacy was described as 'the effective and successful use of data for answering questions, solving problems and supporting decision-making'. The results of this discussion were used to indicate the different perspectives on data literacy from the three participants: students, companies and lecturers.

Students' perspective on their data literacy

Students indicated their data literacy knowledge before participating in the experiment as low and neglected throughout their Bachelor study. Participating students in the focus group described the progress they made as 'going from nothing to something' and were motivated to continue developing their data literacy. Accordingly, this practice-oriented research is seen as a useful and recognised learning activity that complements the current Bachelor curriculum.

Organisations' perspective on students' data literacy

Both companies that participated in the focus group agreed on the low data literacy level of the students at the beginning of the programme. However, there was a noticeable learning curve during the minor. This learning curve resulted in the creation of new tools, instruments and data analysis strategies, which both companies indicated as substantial additions to their current methods for gathering, analysing and interpreting data. The participating companies indicated the following points as vital elements to closing the gap between what students are capable of and what could benefit the organisations:

- Improve proactivity from the student side during the collaboration.
- Increase data awareness about the quality and gathering process of data.
- Improve acknowledgment of the value for businesses through data.

Lecturer perspective on students' data literacy

Lecturers who participated in the focus group discussed that the students showed development of a good understanding of what to do with data throughout the research project. Nevertheless, there is still a gap between what students are capable of and what companies need. Students showed interest in the data analyses processes and delivered useful tools to improve the current data handling procedures of participating companies. These companies indicated the above points as substantial barriers yet to overcome before students meet the data literacy requirements of companies. Therefore, lecturers need to focus on improving these three points to decrease the gap between what companies want and what students are capable of.

5.4 Summary of Study 4

Based on the results yielded from this study, we can conclude three things. First, participant companies and organisations assessed the data literacy of students as moderate across the board and were expecting a higher level to start with. The project began at the start of year four of the Bachelor; at this point in time students had had three years of business school training. Second, a project-based practical assignment like the Datawerkplaats seems to be effective in improving data literacy in students, increasing competence across the board by about 25% percent, though complex organisations and assignments can slow down or hinder the development of data literacy for starting professionals. Third, more focus on the awareness of data and its value would serve companies' needs; advice was put forward by the participating organisations to increase the data-intensity of projects at the RUAS Business School and thus to increase the effectiveness of the institute's education.

6. Discussions

6.1 Measurement of Data Literacy

In this explorative research, we measured the perceived data literacy of individuals based on the scheme of data literacy (see Table 1). In the existing literature we have not found a well-established measurement of individual data literacy, particularly for higher education. Therefore we have adopted this scheme, which was developed based on the discussion on data literacy in undergraduate education (Calzada Prado & Marzal, 2013; Maybee & Zilinski, 2015). The major advantage of this approach is that this scheme clearly identifies the core competences and sub-competences of data literacy and their corresponding learning goals.

To evaluate the measurement model, Cronbach's alpha reliability test was performed. This analysis investigates if the consistency level of each scale aligns with the measurements used within the datasets (Odom & Morrow, 2006). Table 12 shows that Cronbach's alpha values of all the scale items, based on the dataset of students and lecturers, reached 0.8 or 0.9. The literature defines a Cronbach's alpha score of higher than 0.6 as sufficient for exploratory research (Hair et al, 2019). The results in Table 12 demonstrate that all of the Cronbach's alpha scores for this study exceed the standard of 0.6, indicating that it is acceptable to use these items as measurements for the core competences (Pallents, 2013). Therefore, in this study, we did not drop any measurement item for analysis.

Table 12. Assessment of internal consistency of the measurement

Core competence	Cronbach's Alpha 'Students'	Cronbach's Alpha 'Lecturers'
Awareness	0.9	0.8
Access	0.9	0.9
Engage	0.8	0.9
Manage	0.8	0.9
Communicate	0.9	0.9
Use	0.9	0.9

Nevertheless, during the data analysis, we noticed that the adoption of the 5-point Likert scale sometimes limits the interpretation of the findings. In the future, other types of measurements can also be adopted, such as rating scale and ranking. The use of multiple types of measurements will allow us to extract more insights and discern the differences. Another interesting question may be how well this subjective measurement reflects the objective data literacy of the subjects: are lecturers more critical or more modest than students? This would however require an instrument for the objective measurement of data literacy.

6.2 Students' Perceived Data Literacy

Study 3 reveals that students generally perceive themselves as having similar levels of data literacy despite their divergent study experiences. In other words, students in the first and fourth years perceive the same level of data literacy competences. This raises a question: In their learning trajectories, do our students improve their data literacy?

Slight differences are captured among the students. For example, as delineated in Study 3, students from 2018-2019 had the highest level of perceived overall data literacy, while students from 2016-2017 showed the lowest level. Has there been any major change to the curricula since the academic year 2018-2019? What changes have contributed to this improvement?

Due to the limited sample size and the limited information, the answers to these questions cannot be firmly concluded. However, to enhance the data literacy in higher education, it is crucial for us to design our curricula by taking into consideration a clear learning objective of data literacy. In this way we can help our students to systemically build up and improve their data literacy throughout their Bachelor studies.

The results drawn from Study 3 shows that the part-time students' data literacy, in general, is perceived to be higher than that of the full-time students. One possible reason is that part-time students are working in organisations or companies; they have more opportunities to deal with data in a real business environment. Thus, they possess higher levels of data competencies, such as awareness, access, engage, manage, communicate and use.

Specifically, the most important differences between part-time and full-time students are age, years of work experience and field of work. Full-time students predominantly work in typical student jobs like distribution and delivery services, whereas older part-time students work predominantly in an administrative business setting more likely to demand data competences. In turn, part-time students may indeed have a higher level of data literacy than full-time students.

Another possible reason is that part-time students seem to have more opportunities to pursue a data-intensive curriculum at the RUAS Career Academy than students in full-time programmes. The Career Academy was able to build a new curriculum in the last few years. They were not hindered by running programmes and could start with a data-intensive curriculum right away.

These findings of the difference between part-time and full-time programmes emphasise the importance of practice in the real business environment and support the findings of Study 4, where students demonstrated the improvement of their data literacy after participation in the minor programme 'Datawerkplaats'.

Nevertheless, the same perceived level of the core competence 'use' among full-time and part-time students implies that these part-time students may have not trained in their work about how to systematically handle and ethically use the data. This also highlights the weakness of data-driven management in companies – the use of data should be further enhanced.

The field experiment based on Datawerkplaats revealed that students' awareness of data in the real business environment is limited. Consultancy agency Gartner (2014) states that advanced analytics is a top business priority. But the first requirement of this is data awareness to get the analytics started. This is a fundamental competence of a data literate staff, which organisations need in order to start benefiting from analytics endeavours.

6.3 Lecturers' Perceived Data Literacy

Only 64 lecturers from the Business School participated in the survey. As such, we are not able to identify differences between different study programmes. However, overall, we can see that lecturers perceive themselves to have moderate to high levels of data competences. We see higher awareness among lecturers than among students, but overall lecturers score equally high as the students do. This raises the question of how the students' data literacy was developed: was it by way of courses taught at university, or maybe also via other means outside of school, for instance work or online resources?

Moreover, most of the participating lecturers consider their data literacy sufficient for their work and show a high willingness to further improve their data literacy. Nevertheless, there could be a bias of this sample: the participating lecturers are likely to be those who are interested in data, willing to learn and adept at most core competences. Thus, the findings may not be well generalised to the full population of lecturers.

Study 3 shows that almost 10% of participating lecturers showed neutral attitudes towards the sufficiency of their data literacy in relation to their work needs (see Table 8). This may imply that it is not clear for all lecturers what data competences and what level of data literacy they should possess to perform their work in education. One possible reason is changes in the curricula design and the competences requested of lecturers. All these changes make lecturers uncertain about whether their current level of data literacy is sufficient. Another possible reason is the lack of specific information and requirements for lecturers' data competences to fulfil their teaching objectives.

Figure 10 shows that among the lecturers who already have confidence in the sufficiency of their data literacy, a high percentage, more than 75%, are still willing to improve, which is wonderful but also actually a necessity in this space. Developments in technology require lecturers to continually educate themselves to keep their skills up to date.

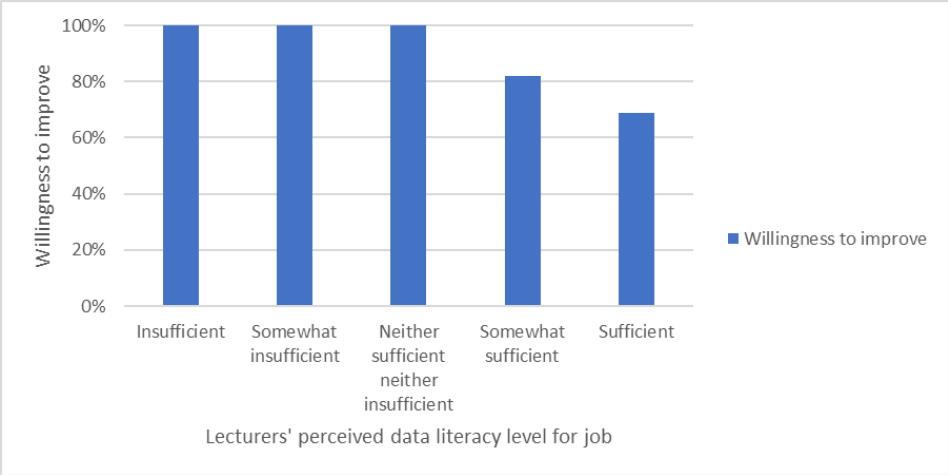

Figure 10. The trend of lecturers' willingness to improve their data literacy

According to Figure 10, as confidence in personal data literacy decreases, we see an increase up to 100% in willingness to improve among lecturers with moderate to low confidence about the sufficiency of their data literacy for their jobs. When asked if they want to improve, respondents may wonder if this can be done on the boss's time or if they need to invest their private time. As we were only interested in the willingness to improve during worktime, the survey asked specifically if they would like to improve, provided their time was compensated. So these numbers reflect the assumption that the employer will facilitate professionalisation in data literacy.

6.4 Organisation's Perceived Data Literacy of Students

Organisations showed low assessments of graduates' data literacy. They pointed out the inadequate level of different data competences and the lack of experience in using current data tooling among students. But they confirmed the student's learning capabilities throughout the research project. Within just one semester, students improved remarkably in all the data competences; they can quickly grasp new knowledge about the use of data and learn to use a new software or tool.

Organisations, including SMEs, emphasised their needs for data literate employees so that they will be able to take a data-driven approach and be more competitive in the market. Meanwhile, there seems to be a shortage of sufficiently data literate employees in the Netherlands, and in turn, recruitment is often done internationally. This impresses upon us the urgency to improve our students' data literacy in Dutch higher education.

A striking finding drawn from the research is that practice-oriented, data-driven research projects do not only improve the students' data literacy, but also the lecturers' data literacy by supervising the projects, and the company's organisational level of data literacy by working on a data analytics project. Also, we see organisations' focus on their data-driven endeavours strengthen while participating in a project, leading to a win-win-win situation.

7. Conclusions

7.1 Summary of the Findings

Data literacy is essential for an individual to better integrate and interact with others in this information society. It is also of paramount importance for companies to take a data-driven approach in their business and management. Data is driving a broad digital transformation in our current society, which in turn drives the demand for data literate young professionals. That is why education needs to cater to this surging demand in society and business. Therefore, in this research, we attempt to answer the central research question: How can we define, measure and promote data literacy in higher education? Four research sub-questions were constructed accordingly.

The first research sub-question is: What is data literacy? Having disambiguated different kinds of literacy and recognised the core competences of data literacy as fundamental to the more general and broader concept of digital literacy, we have defined data literacy as the ability of an individual to identify, acquire, manage, evaluate, interpret, apply, use and communicate data in a critical, ethical and responsible manner to reach a certain goal. To assess an individual's data literacy, we used six core competences, namely Awareness, Access, Engage, Manage, Communicate and Use.

The second research sub-question is: To what extent is data promoted in current courses at the RUAS Business School? We found high variance between programmes and between different kinds of curricula. In one programme more than half of the ECTS could be attributed to courses dealing with data in one way or another. Minor programmes showed higher data intensity across the board than regular programmes did, and students who don't see much data in their regular curriculum do have ample opportunity to pursue a minor that promotes data literacy. Evaluating courses and minor programmes together, we found that overall only a small percentage of the ECTS for courses deal with data. We are cautious about drawing strong conclusions, as we encountered severe data quality issues. Development of a data strategy and data management infrastructure at the institute could inspire programmes to work in a more data-driven manner, which in turn could inspire the development of more data-intensive curricula in programmes as they see fit. We will see increasing levels of data intensity in future research rounds, as several programmes are working on a new data-rich curriculum as this paper goes to press.

The third research sub-question is: What is the current perceived data literacy of students and lecturers? We found moderate to high levels of perceived data literacy, with no difference between students and lecturers. Part-time students pursuing a degree at the RUAS Business School Career Academy were found to be more data literate than their full-time counterparts. This can be caused by the fact that part-time students are required to deal with data in their

professional work environments, but also because of a data-intensive curriculum at the Career Academy. Most lecturers view their data literacy as sufficient for their job, but even so, most of them are willing to improve their data literacy if their employer would facilitate this.

The fourth research sub-question is: How can we effectively improve students' data literacy to meet SMEs' needs? The study confirmed that companies need professionals who are aware of data and the value thereof. We found that employing a project-based practical assignment seems to be effective in improving data literacy in students. This is effective because it is real, but complexity should be contained or limited for students totally new to data. More of these initiatives with even higher data intensity could improve student's data literacy to meet SMEs' needs.

In short, this research has proposed a working definition of data literacy for education. A measurement to evaluate a student's and lecturer's data literacy has been operationalised based on the core competences identified in this definition. To promote data literacy in higher education, we shall create a learning environment that synthesises education, research and business. This all-round and highly interactive learning environment will not only help students more efficiently improve their data competences but will also bring benefits to lecturers, researchers and business practitioners. This win-win-win situation can further sustain the synthesis of education, research and business in our higher education.

7.2 Contributions to Education

This research has made several major contributions to education. First, conducting the empirical study has enhanced the awareness of data literacy among students, lecturers and study programme managers. This awareness may encourage these main stakeholders to further acknowledge the significance and rationale of data literacy in higher education. Thus, they are more likely to build up intrinsic motivation to address data literacy in their studies and work.

Second, the involvement of companies and organisations in this study has demonstrated the gap between what the real world requires and what we have equipped our students with. The fact that business partners judged students' data literacy as low at the beginning of the experiment implies that they had high expectations. This can be explained by the highly data-driven nature of the participating organisations, which may be considered a general quality among organisations doing business in the twenties.

Third, the examination of the students' and lecturers' perceived data literacy illustrates the individual challenges of data competences. The results thus provide a snapshot that can help us understand the general levels of students' and lecturers' data competences. This information can help us to better design curricula and support programmes for lecturers' continuing professionalisation.

Last, the field experiment exemplified how practice-oriented, data-driven research projects can improve students' data literacy and create a win-win-win situation for students, lecturers and companies. The mutual benefits derived from this type of research projects can foster a substantiable synthesis among research, education and business practices.

7.3 Contributions to Research

There is no universal definition of data literacy; different definitions are constructed by different industries and business sectors. To create a common understanding of data literacy in higher education, this study has developed a clear-cut definition that focuses on individual competences in dealing with data.

To evaluate a student's or a lecturer's level of data literacy, this research has operationalised the concept and synthesised competences of data literacy in higher education based on the work of Maybee and Zilinski (2015) and of Prado and Marzal (2013). The resulting framework has made it possible to examine an individual's level of data literacy against education goals.

Furthermore, this is one of the first studies to empirically examine students' and lecturers' data literacy levels, identify the gap between the needs of businesses and student competences in data literacy, and experimentally determine how practice-oriented research with a business partner can help improve the data literacy of the main stakeholders (students and businesses). This study has demonstrated the importance of bringing research and business partners into formal education programmes.

7.4 Recommendations for Education

Based on our research findings, several recommendations can be made. First, study programmes that attempt to enhance students' data literacy need to systematically incorporate the training of data competences throughout the four-year Bachelor learning trajectory. In other words, the training ought to start in Year 1 and build further in Years 2, 3, and 4, gradually building structure and adding complexity. Students will then have the opportunity to reach their full data literacy potential.

University management needs to be aware of some possible influential factors of the data intensity of curricula. For example, lecturers' propensity to change and innovate could be a factor when it comes to creating data-intensive curricula. Data literacy is much more important now than it was 30 years ago; and this, in combination with lecturers' personal backgrounds, may also influence how data is incorporated in curricula. The different content focus among the study programmes could also lead to differences. It is expected that Commerce has seen a more pervasive digital transformation than, for instance, Human Resources Management. Whether the school organisation and programme management are organised in a data-driven fashion could

also lead to differences in study programmes. This may also lead to different hiring decisions; a data literate manager may be more inclined to ask candidates about their data literacy competences.

Next, due to the limited availability of Dutch study materials and resources, and the challenges in translation, study programmes may consider the adoption of bilingual (Dutch and English) learning materials to address the teaching of data competences. For example, Business, IT & Management will adopt some English literature in their new tracks in the 3rd and 4th years. Managers of courses offered in Dutch could leverage this advantage of language, for example by adopting bilingual teaching materials.

Furthermore, we would like to refer to constructive alignment in education, which is to say that learning goals, learning environment and assessment should be aligned for education to be successful. In the final stages of a vocational programme, it is important to embed the actual business environment into the learning environment. Assuming data literacy is an important learning goal, the real business environment desiring a data-driven approach and assessment on the data-driven solution to a real business problem creates optimal alignment.

In addition to a practice-oriented data-driven research project with a company, we may also add the elements of the real world to other teaching methods. For example, with regard to the practices in controlled learning environment, instead of just dummy data, we could use a training dataset that developed from real company data, reflecting the actual business phenomenon. Our Research Centre Business Innovation has recently developed a training dataset for process mining in healthcare to reflect the actual business process in this domain and a teaching case, with ready-to-use teaching note, of data monetisation to illustrate how companies can extract real value from data. Both can be directly used by the study programmes in our university.

Despite the possible bias of the sample of lecturers, we were happy to see that there is a group of lecturers who have adequate level of data literacy for work and great interest in data competences with the Business School of RUAS. We shall empower these driving forces to create data-driven culture and promote data literacy in their own programmes with their great insights into the domain. For example, guest lectures can be organised and given by our own lecturers who are adept at certain data competences. University management should facilitate the peer learning and knowledge sharing among the lecturers.

It should be pointed out that while study programmes may consider the corresponding data literacy of its lecturers, it does not mean that all the lecturers should have the same level of data literacy or the proficiency in each data competence. Different courses or teaching activities require different data competences. It is more important for the study programme managers and curricula committees to carefully craft out the association between their learning objectives and

the corresponding requirement for the lecturers in terms of data literacy. The availability of this information will also help the university and study programmes to recruit the right lecturers.

Last but not least, for successful research endeavours it is advised that a high level data strategy for course data is developed and implemented. This does not align well with the trend of decentralisation in RUAS, but for some core processes it would benefit the whole university and make transparent where collaboration could be realised. This could make the organisation more transparent and create much needed insights that can be generated on the spot, releasing much time from staff and lecturers. An extra benefit to this is that lecturers will be inspired by the benefits of a data driven approach, and adopt it into their teaching.

7.5 Limitations and Future Research

One of the noticeable limitations of this study is its small sample size. We were only able to obtain 115 student and 64 lecturer respondents, respectively. In turn, it was not possible for us to form a representative sample of each study programme. However, at the aggregate level, based on the diversity of the components of the respondents, to a certain extent our results do reflect the overall situation.

Another limitation is the inaccurate and incomplete information of course descriptions extracted from Osiris. The quality of the raw data was low, mainly due to the decentralisation of data management within our organisation.

In future research, we shall draw more representative samples of students and lectures with an effective incentive scheme for a higher response rate and an efficient distribution method to reach the populations. Different data management systems and decentralised documents need to be incorporated into the research to increase the accuracy, completeness and timeliness of the data.

Finally, future research can be further extended to explore demand trends for different data competences in different major business domains. To achieve this we propose that a permanent longitudinal research project be put in place to get a clear view on this evolving demand. As a result, study programmes will be able to better design their curricula by incorporating the specific, developing needs of the corresponding real-world domains.

References

Bowler, L., Acker, A., Jeng, W., & Chi, Y. (2017). "It lives all around us": Aspects of data literacy in teen's lives. *Proceedings of the association for information science and technology*, *54*(1), 27-35.

Bulajic, A., Despotovic, M., & Lachmann, T. (2019). Understanding functional illiteracy from a policy, adult education, and cognition point of view: Towards a joint referent framework. *Zeitschrift für Neuropsychologie*, *30*, 109-122. https://doi.org/10.1024/1016-264X/a000255

Carretero, S., Vuorikari, R., & Punie, Y. (2017). The digital competence framework for citizens. *Publications Office of the European Union*.

Dijanošić, B. (2009). Prilozi definiranju pojma funkcionalne nepismenosti. *Andragoški Glasnik*, *13(1)*, 1–64.

Eurostat. (2022). *Eurostat Statistics Explained Glossary*. Retrieved 8-1-2022 from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Category:Glossary

Gartner. (2014). *Gartner Says Advanced Analytics Is a Top Business Priority*.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). *Big Data for Big Business?*

A Taxonomy of Data-driven Business Models used by Start-up Firms. https://cambridgeservicealliance.eng.cam.ac.uk/resources/Downloads/Monthly%20Papers/2014_March_DataDrivenBusinessModels.pdf

Lund, S., Madgavkar, A., Manyika, J., Smit, S., Ellingrud, K., Meaney, M., & Robinson, O. (2021). The future of work after COVID-19. *McKinsey Global Institute*, *18*.

Mandinach, E. B., & Gummer, E. S. (2013). A Systemic View of Implementing Data Literacy in Educator Preparation. *Educational Researcher*, *42*(1), 30-37. https://doi.org/10.3102/0013189x12459803

Maybee, C., & Zilinski, L. (2015). Data informed learning: A next phase data literacy framework for higher education.

MKB010>>Next. (2021). *Digitaal & duurzaam doorpakken*. https://www.ondernemen010.nl/apps/ondernemen010.nl/netwerken/mkb010/over-mkb010-next/index.xml

Ondernemen010. (2019). *Koersen richting de nieuwe economie*

Het peloton in versnelling

Programma Innovatie (brede) MKB 2019-2022.

Philippa, M., Debrabandere, F., & Quak, A. (2004). Etymologisch Woordenboek van het Nederlands, deel 1: A t/m E.

Prado, J. C., & Marzal, M. Á. (2013). Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. *Libri*, *63*(2), 123-134. https://doi.org/doi:10.1515/libri-2013-0010

Ridsdale, C., Rothwell, J., Smit, M., Bliemel, M., Irvine, D., Kelley, D., Matwin, S., Wuetherick, B., & Ali-Hassan, H. (2015). *Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report*. https://doi.org/10.13140/RG.2.1.1922.5044

Schwab, K. (2017). *The fourth industrial revolution*.

Stephenson, E., & Caravello, P. S. (2007). Incorporating data literacy into undergraduate information literacy programs in the social sciences: A pilot project. *Reference services review*.

UNESCO. (1979). *Records of the General Conference Twentieth*

Session: Resolutions. . Paris: UNESCO

Wolff, A., Gooch, D., Cavero Montaner, J. J., Rashid, U., & Kortuem, G. (2016). Creating an Understanding of Data Literacy for a Data-driven Society. *The Journal of Community Informatics*(12(3)), 9-26. https://openjournals.uwaterloo.ca/index.php/JoCI/article/view/3275/4298

World Economic Forum. (2020). *The Future of Jobs Report 2020*. http://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf

Yang, S. (2020). Turning data into strategic assets. https://www.hogeschoolrotterdam.nl/contentassets/2c5a652aff6e4a78a557b4170b7ca697/yang_turning-data-into-strategic-assets.pdf

Appendix

Appendix I. Descriptives of students' perceived data literacy per study program

Mean	Awareness	Access	Engage	Manage	Communicate	Use	Overall
Accountancy voltijd	4.2	3.8	4.0	3.7	3.6	3.6	3.8
Anders	4.2	4.2	4.3	4.0	4.2	4.4	3.8
Bedrijfskunde deeltijd	4.3	4.0	3.5	3.9	3.7	3.4	4.1
Business IT & Management deeltijd	4.7	4.3	4.4	4.5	4.0	4.3	3.5
Business IT & Management voltijd	4.3	4.1	4.1	4.2	4.1	4.1	3.6
Commerciële Economie deeltijd	4.7	4.0	4.0	4.0	4.0	3.8	4.1
Commerciële Economie voltijd	3.5	3.0	3.0	4.0	3.7	3.8	4.2
Creative Marketing & Sales voltijd	3.5	4.5	3.3	3.3	3.3	3.8	4.4
Finance & Control voltijd	3.9	3.7	3.8	3.5	3.4	3.5	4.1
Human Resource Management deeltijd	4.3	3.9	4.0	3.8	4.0	3.9	3.6
International Business voltijd	4.3	4.2	4.2	3.9	4.0	4.2	4.0
Marketing of Social Business voltijd	3.7	3.4	3.5	3.5	3.0	3.8	3.5
Max	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
Accountancy voltijd	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Anders	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Bedrijfskunde deeltijd	4.8	5.0	5.0	5.0	5.0	5.0	5.0
Business IT & Management deeltijd	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Business IT & Management voltijd	5.0	5.0	5.0	5.0	5.0	5.0	4.9
Commerciële Economie deeltijd	4.7	4.0	4.0	4.0	4.0	3.8	4.1
Commerciële Economie voltijd	3.5	3.0	3.0	4.0	3.7	3.8	3.5
Creative Marketing & Sales voltijd	3.5	4.5	3.3	3.3	3.3	3.8	3.6
Finance & Control voltijd	5.0	5.0	5.0	5.0	5.0	5.0	3.6
Human Resource Management deeltijd	5.0	4.0	4.0	4.3	4.0	4.0	4.0
International Business voltijd	5.0	5.0	5.0	5.0	5.0	5.0	4.1

Marketing of Social Business voltijd	4.0	4.0	4.0	4.0	4.0	4.0	3.5
Min	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
Accountacy voltijd	2.5	1.5	1.3	1.3	1.0	1.0	1.4
Anders	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Bedrijfskunde deeltijd	3.8	2.8	2.0	2.0	2.3	1.0	2.3
Business IT & Management deeltijd	3.8	3.8	3.7	3.3	3.0	2.8	3.4
Business IT & Management voltijd	3.2	3.3	3.3	3.7	3.3	3.0	3.4
Commerciële Economie deeltijd	4.7	4.0	4.0	4.0	4.0	3.8	4.1
Commerciële Economie voltijd	3.5	3.0	3.0	4.0	3.7	3.8	3.5
Creative Marketing & Sales voltijd	3.5	4.5	3.3	3.3	3.3	3.8	3.6
Finance & Control voltijd	3.0	2.0	2.0	1.7	2.0	2.0	2.0
Human Resource Management deeltijd	3.5	3.8	4.0	3.3	4.0	3.8	3.8
International Business voltijd	3.3	2.8	3.0	2.3	2.7	3.0	3.4
Marketing of Social Business voltijd	3.3	2.8	3.0	3.0	2.3	3.5	3.0
Standard deviation	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
Accountacy voltijd	0.6	0.7	0.9	0.8	0.9	1.0	0.8
Anders	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Bedrijfskunde deeltijd	0.4	0.8	1.3	1.1	1.0	1.5	1.0
Business IT & Management deeltijd	0.4	0.4	0.5	0.6	0.6	0.7	0.4
Business IT & Management voltijd	0.5	0.6	0.5	0.5	0.5	0.6	0.4
Commerciële Economie deeltijd	NA	NA	NA	NA	NA	NA	NA
Commerciële Economie voltijd	NA	NA	NA	NA	NA	NA	NA
Creative Marketing & Sales voltijd	NA	NA	NA	NA	NA	NA	NA
Finance & Control voltijd	0.7	0.7	0.8	0.8	0.8	0.8	0.6
Human Resource Management deeltijd	1.1	0.2	0.0	0.7	0.0	0.2	0.3
International Business voltijd	0.5	0.5	0.5	0.7	0.7	0.5	0.4
Marketing of Social Business voltijd	0.5	0.9	0.7	0.7	1.2	0.4	0.7

Appendix II. Descriptives of students' perceived data literacy per start year

Mean	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
2015 - 2016 or earlier	4.1	4.0	4.1	4.0	3.5	3.8	3.9
2016 - 2017	3.7	3.7	3.7	3.5	3.5	3.7	3.6
2017 - 2018	4.3	3.9	4.0	3.8	3.8	3.9	4.0
2018 - 2019	4.3	4.0	4.1	4.0	3.9	4.1	4.0
2019 - 2020	4.2	4.1	4.1	4.0	3.7	3.9	4.0
2020 - 2021	4.2	4.0	4.0	3.7	3.7	3.7	3.9
Max	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
2015 - 2016 or earlier	5.0	5.0	5.0	5.0	5.0	5.0	4.9
2016 - 2017	4.3	4.5	4.3	4.7	4.7	4.8	4.3
2017 - 2018	5.0	5.0	5.0	5.0	5.0	5.0	5.0
2018 - 2019	5.0	5.0	5.0	5.0	5.0	5.0	5.0
2019 - 2020	5.0	5.0	5.0	5.0	5.0	5.0	5.0
2020 - 2021	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Min	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
2015 - 2016 or earlier	3.0	2.8	3.0	2.3	2.0	2.0	2.6
2016 - 2017	3.0	3.0	3.0	2.0	2.7	2.8	2.8
2017 - 2018	3.0	2.0	2.0	2.0	2.0	2.0	2.2
2018 - 2019	2.5	1.5	1.3	1.3	1.0	1.0	1.4
2019 - 2020	3.0	3.5	3.3	1.7	2.3	2.5	3.1
2020 - 2021	2.0	2.0	2.0	2.0	2.0	1.0	2.0
Standard deviation	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
2015 - 2016 or earlier	0.7	0.7	0.7	0.8	0.9	1.0	0.7
2016 - 2017	0.5	0.5	0.6	0.9	0.7	0.7	0.6
2017 - 2018	0.6	0.7	0.9	0.8	0.8	0.8	0.7
2018 - 2019	0.6	0.7	0.7	0.7	0.8	0.7	0.6
2019 - 2020	0.6	0.4	0.4	1	0.7	0.7	0.5
2020 - 2021	0.8	0.8	0.9	0.8	0.9	1.0	0.8

Appendix III. Descriptives of lecturers' perceived data literacy per study program

Mean	Awareness	Access	Engage	Manage	Communicate	Use	Overall
Anders	4.3	4.0	4.2	4.0	3.7	3.8	4.0
Accountancy	4.7	4.1	4.2	4.0	3.8	3.8	4.2
Bedrijfskunde	4.4	3.9	4.2	3.4	3.4	3.6	3.8
Business IT & Management	4.9	4.8	4.8	4.9	4.6	4.6	4.8
Commerciële Economie	4.5	4.1	4.3	3.5	3.9	3.6	4.0
Creative Marketing & Sales	4.7	4.0	4.3	5.0	3.0	4.3	4.3
Finance & Control	4.5	4.3	4.3	4.3	4.0	3.7	4.3
Finance, Tax and Advice	4.4	4.1	4.3	4.1	3.7	3.8	4.1
Global Marketing & Sales	4.3	4.0	4.0	5.0	4.0	4.0	4.3
Human Resource Management	4.0	2.4	3.5	2.8	2.7	2.9	3.1
International Business	4.3	3.7	3.7	3.6	3.2	3.4	3.7
Marketing of Social Business	3.3	1.0	1.3	1.3	1.0	1.8	1.5
Ondernemerschap & Retail Management	4.6	3.9	4.3	3.8	3.8	3.6	4.0
Max	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
Anders	4.7	5.0	5.0	4.3	5.0	4.5	4.8
Accountancy	5.0	4.8	5.0	5.0	5.0	4.5	4.8
Bedrijfskunde	5.0	5.0	5.0	4.0	4.7	5.0	4.7
Business IT & Management	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Commerciële Economie	4.8	4.5	4.7	4.3	4.7	4.3	4.5
Creative Marketing & Sales	4.7	4.0	4.3	5.0	3.0	4.3	4.3
Finance & Control	4.8	4.8	4.7	4.7	4.7	4.5	4.7
Finance, Tax and Advice	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Global Marketing & Sales	4.3	4.0	4.0	5.0	4.0	4.0	4.3
Human Resource Management	4.2	3.0	4.0	3.3	3.3	3.0	3.3
International Business	5.0	5.0	5.0	5.0	4.7	5.0	4.9
Marketing of Social Business	3.3	1.0	1.3	1.3	1.0	1.8	1.5
Ondernemerschap & Retail management	5.0	5.0	5.0	5.0	5.0	4.8	4.9
Min	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
Anders	3.8	3.3	3.0	3.7	2.3	2.0	3.2
Accountancy	4.2	3.5	3.0	3.3	3.0	2.0	3.1
Bedrijfskunde	4.0	2.8	3.3	2.3	1.3	1.8	2.8

Business IT & Management	4.7	4.3	4.0	4.3	4.0	4.0	4.4
Commerciële Economie	4.2	3.0	3.3	2.0	3.0	2.3	3.5
Creative Marketing & Sales	4.7	4.0	4.3	5.0	3.0	4.3	4.3
Finance & Control	3.7	4.0	4.0	4.0	2.0	2.0	3.4
Finance, Tax and Advice	3.5	3.5	3.3	3.3	2.7	2.0	3.2
Global Marketing & Sales	4.3	4.0	4.0	5.0	4.0	4.0	4.3
Human Resource Management	3.8	1.8	3.0	2.3	2.0	2.8	2.9
International Business	3.5	1.8	1.7	1.0	1.0	1.0	2.1
Marketing of Social Business	3.3	1.0	1.3	1.3	1.0	1.8	1.5
Ondernemerschap & Retail management	4.0	3.0	4.0	3.0	3.0	2.5	3.6

Standard deviation	**Awareness**	**Access**	**Engage**	**Manage**	**Communicate**	**Use**	**Overall**
Anders	0.4	0.7	0.8	0.3	1.0	1.0	0.6
Accountancy	0.4	0.5	0.9	0.7	0.9	1.2	0.7
Bedrijfskunde	0.4	0.8	0.5	0.6	1.0	1.0	0.6
Business IT & Management	0.1	0.3	0.4	0.3	0.4	0.5	0.2
Commerciële Economie	0.2	0.5	0.5	0.9	0.6	0.7	0.3
Creative Marketing & Sales	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Finance & Control	0.5	0.3	0.3	0.3	1.3	1.1	0.6
Finance, Tax and Advice	0.6	0.6	0.6	0.7	0.9	1.2	0.7
Global Marketing & Sales	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Human Resource Management	0.2	0.9	0.7	0.7	0.9	0.2	0.3
International Business	0.5	1.0	1.0	1.0	1.3	1.2	0.8
Marketing of Social Business	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Ondernemerschap & Retail management	0.4	0.8	0.5	0.9	0.9	1.0	0.6

Appendix IV. Correlation matrix of core competences

A. Correlation matrix of core competences based on students

Core competence	***Awareness***	***Access***	***Engage***	***Manage***	***Communicate***	***Use***
Awareness	**1.00**	0.76***	0.71***	0.71***	0.73***	0.73***
Access	0.76***	**1.00**	0.74***	0.64***	0.73***	0.73***
Engage	0.71***	0.74***	**1.00**	0.68***	0.71***	0.76***
Manage	0.71***	0.64***	0.68***	**1.00**	0.67***	0.74***
Communicate	0.73***	0.73***	0.71***	0.67***	**1.00**	0.82***
Use	0.73***	0.73***	0.76***	0.74***	0.82***	**1.00**

Note: * p < 0.05, ** p < 0.01, *** p < 0.001

B. Correlation matrix of core competences based on lecturers

Core competence	***Awareness***	***Access***	***Engage***	***Manage***	***Communicate***	***Use***	***Sufficiency***	***Improvement***
Awareness	**1.00**	0.69***	0.67***	0.62***	0.74***	0.69***	0.57***	-0.05
Access	0.69***	**1.00**	0.81***	0.75***	0.76***	0.69***	0.53***	0.16
Engage	0.67***	0.81***	**1.00**	0.66***	0.81***	0.67***	0.54***	0.08
Manage	0.62***	0.75***	0.66***	**1.00**	0.64***	0.71***	0.50***	0.05
Communicate	0.74***	0.76***	0.81***	0.64***	**1.00**	0.82***	0.70***	0.10
Use	0.69***	0.69***	0.67***	0.71***	0.82***	**1.00**	0.67***	0.01
Sufficiency	0.57***	0.53***	0.54***	0.50***	0.70***	0.67***	**1.00**	-0.25*
Improvement	-0.05	0.16	0.08	0.05	0.10	0.01	-0.25*	**1.00**

Note: * p < 0.05, ** p < 0.01, *** p < 0.001