

A peek into the working day

Combining multiple data collection techniques to record employee behaviour and discover work patterns within an organisational context

Tea Šinik
2915960

Supervisors:

Utrecht University

Prof. dr. ir. H.A. (Hajo) Reijers

dr. I.M. (Iris) Beerepoot

Ernst & Young

Tijmen Baas

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Science in the Faculty of Science Department of Information and Computing Sciences

29th January 2023



**Utrecht
University**



A peek into the working day

Combining multiple data collection techniques to record employee behaviour and discover work patterns within an organisational context

Author: Tea Šinik
2915960
t.sinik@students.uu.nl
Master Business Informatics
Department of Information and Computing Science
Utrecht University

1st supervisor: Prof. dr. ir. H.A. (Hajo) Reijers
h.a.reijers@uu.nl
(Department/Research group) Business Process Management and Analytics
Utrecht University

2nd supervisor: dr. I.M. (Iris) Beerepoot
i.m.beerepoot@uu.nl
(Department/Research group) Business Process Management and Analytics
Utrecht University

External supervisor: Tijmen Baas
tijmen.baas@nl.ey.com
Senior Manager and Team Lead Helix Team EY Netherlands
Ernst & Young (EY)

Word count: 46.535
Page count: 48 (10 - 58)
Total page count: 133

29th January 2023, 't Harde

Abstract

A peek into the working day: Combining multiple data collection techniques to record employee behaviour and discover work patterns within an organisational context.

A detailed recording of the behaviour of the employees is crucial because this is directly related to the success of an organisation. These records have many purposes such as monitoring the key performance indicators (KPIs), measuring the performance of employees, or detecting unwanted or repetitive tasks that can be (partially) automated in the future. Nowadays, business processes are increasingly supported by information systems and involve various actors from multiple departments. The assumption is that these records are an excellent representation of the actual employee behaviour to discover work patterns. However, certain activities are not recorded (properly and/or accurately) or occur outside of the information system(s). Thus, it is not always evident to what extent the records reflect the actual behaviour of the employee(s). This gap has not been investigated properly because existing literature almost exclusively focused on one stand-alone technique.

The study aimed to fill the gap by combining multiple data collection techniques to record employee behaviour and discover work patterns within an organisational context. The goal was to present the (dis-)advantages, and the commonalities and differences between the techniques to examine the level of confidence that should be placed in the analysis of this type of data. The objective was to identify the shortcomings of each technique and the impact it can have on the quality of the records.

We applied two methods in this study. First, we conducted a systematic literature review to gain a deeper understanding of work patterns, and to find techniques suitable to record employee behaviour in an organisational setting. We proceeded with the search until theoretical saturation was reached. The literature findings were used to design the research method of the study, and ultimately to conclude to what extent our study has confirmed or refuted these findings. Second, we conducted a multiple-case study to record the behaviour of six employees within the Helix Team during the course of 6 working days. We combined three data collection techniques: non-participant observations, screen recording, and archival analysis (timesheet).

The techniques produced 136:04 hours of data and 849 records of activities. We analysed the accuracy, quality, and particularities and shortcomings. We identified 58 differences related to the timestamp, online versus offline activities, short-lived activities, overtime activities, and uncategorised activities. Our study showed that each technique yields unique results, and has its advantages and disadvantages. Due to the shortcomings of the techniques, we have seen that the use of only one technique **will not** produce a complete and accurate record of the activities that occur on the screen (online), in the hallway (offline), and in the wee hours (overtime). Therefore, it is necessary to combine multiple techniques with different characteristics to be able to (completely and accurately) record employee behaviour and discover work patterns within an organisational context.

Keywords: Data Collection Techniques · Employee Behaviour · Work Patterns · Non-Participant Observation · Screen Recording · Archival Analysis · Timesheet

Acknowledgements

This report concludes my time as a Master's student at Utrecht University. This journey has been one of the most challenging, yet rewarding experiences of my life. I am thankful for the support I received throughout this study.

To begin, I would like to express my gratitude to my first supervisor Prof. dr. ir. H.A. (Hajo) Reijers. Thank you for your guidance, detailed feedback, and support throughout this research. I appreciate you taking the time to assist me whenever I had questions despite your hectic schedule. I found it quite pleasant to work with you and even discuss the possibilities of a PhD project. Secondly, I would like to thank my second supervisor dr. I.M. (Iris) Beerepoot for introducing me to this exciting topic. I am deeply grateful for your tremendous support. You provided frequent, constructive, and detailed feedback while preparing for your PhD defence and during the absence of Hajo (sabbatical).

Combining a full-time study with a full-time job has not been easy. This would not have been possible without the patience, support, and flexibility of my colleagues at Ernst & Young. Many thanks to Tijmen Baas; my mentor and daily supervisor. Thank you for being so involved throughout these years, and for always taking the time out of your busy schedule to assist me.

Last but certainly not least, my sincere gratitude goes out to my parents Miroslav Šinik and Fikreta Relja-Šinik. Thank you for your endless love, support, encouragement, and patience whilst listening to my doubts and rants. You gave me the needed confidence and strength to combine a full-time job with a full-time study. This has not been an easy journey. There were a lot of sleepless nights (sorry mom!) and early mornings during an exam period (thank you dad for being my chauffeur!). However, we made it – always together – therefore I dedicate this thesis to you.

Volim vas najviše na svijetu!

Tea Šinik

29th January 2023, 't Harde

Contents

1	Introduction	10
1.1	Problem Statement	10
1.2	Research Aim and Objectives	11
1.3	Research Question	11
1.4	Contributions	12
1.5	Thesis Structure	12
2	Literature Review	13
2.1	Content of the Literature Review	13
2.2	Work Patterns	13
2.3	Categorisation of Data Collection Techniques	14
2.3.1	Direct Technique	14
2.3.2	Indirect Technique	17
2.3.3	Independent Technique	18
2.4	Synthesis	20
3	Research Method	22
3.1	Triangulation	22
3.2	Literature Review	23
3.3	Case Study	24
3.4	Knowledge Gathering	25
3.5	Participant Selection	26
3.6	Data Collection	27
3.7	Data Analysis	29
3.8	Threats to Validity	30
4	Analysis	32
4.1	Data Preparation	32
4.2	Comparison	32
4.3	Accuracy	33
4.4	Quality	33
4.4.1	(Offline) Activity only recorded with observation technique	33
4.4.2	(Short-lived) Activity only recorded with the screen recording technique	34
4.4.3	(Overtime) Activity only recorded with the timesheet technique	34
4.5	Particularities and shortcomings	35
5	Results	36
5.1	Timestamp	36
5.2	Online versus Offline Activities	37
5.2.1	Meeting	38
5.2.2	Personal	39
5.2.3	Tours	40
5.3	Short-lived Activities	41
5.4	Overtime Activities	43
5.5	Uncategorised Activities	44
6	Implications	47
6.1	Timestamp	47

6.2	Online versus Offline Activities	48
6.3	Short-lived Activities	49
6.4	Overtime Activities	50
6.5	Uncategorised Activities	50
7	Translation of Insights Into Practice	52
7.1	Tick Tock, What Time Is It	52
7.2	Back to the Office	53
7.3	Double the Work, and Half the Rating	54
8	Conclusion	55
8.1	Answers to Sub-Questions	55
8.2	Answer to Main Research Question	56
8.3	Limitations	57
8.4	Challenges in Research Context	57
8.5	Future Work	58
	Bibliography	67
	A Literature Review	68
	B Systematic Literature Review Protocol	77
	C Case Study Protocol	80
C.1	Preamble	80
C.2	Research Instruments	81
C.2.1	Non-participant Observation	81
C.2.2	Screen Recording	82
C.2.3	Archival Analysis	83
C.3	Data Collection & Analysis	85
C.4	Ethical Considerations & Guidelines	86
C.5	Informed Consent Form	86
D	Classification Scheme	92
D.1	Standardised Form	92
D.2	Time	95
D.3	Category	95
D.4	Purpose of the Activities	96
D.5	Sub-Activity	97
D.6	Medium	98
D.7	Participants	99
D.8	Initiation of the Activities	100
D.9	Duration	100
D.10	Field Notes	100
D.11	Client	101
D.12	Participant	101
E	Data Analysis	102
E.1	Alteryx	102
E.1.1	Check Empty Fields	102
E.1.2	Union Data Streams	102
E.1.3	Comparison Results Observation & Screen Recording	102
E.1.4	Richness of the Field Notes	103
E.2	Microsoft Excel	110
E.2.1	Total per Classification	110
E.2.2	Total per Classification per Participant	110
F	Results from Multiple-Case Study Extended	112
F.1	Category	113
F.2	Purpose of the Activities	115
F.3	Sub-Activity	117

F.4	Medium	121
F.5	Participant	125
F.6	Initiation of the Activities	128
F.7	Field Notes	129
F.8	Clients	131
F.9	Participant	133

List of Figures

2.1	The three types of process mining: <i>discovery, conformance, and enhancement</i> , adapted from [1, 102]	18
3.1	Research approach phases	22
5.1	The Activity IDs in the Mercury timesheet	39
B.1	Excel file containing the information documented for the Systematic Literature Review documentation	78
B.2	Word file containing the overview of data collection techniques split per source	79
C.1	Example of the screen recording tool Snagit	82
C.2	Example of the timesheet within Mercury	84
C.3	Sample estimation (i.e., total number of participants) based on the availability of the researcher, and total time needed for the data capturing and analysis	87
C.4	The PowerPoint presentation used to contact each Helix Team member regarding the purpose of the research	88
C.5	The random number generator used to generate a randomised order for the data analysis	89
C.6	Informed Consent Form	90
D.1	Snippet (filled-in) standardised form used during Observation	94
E.1	Alteryx workflow to check on empty fields	104
E.2	Alteryx workflow to union the data files	105
E.3	Alteryx workflow to (inner) join the observation and screen recording results	106
E.4	Snippet Excel sheet of the comparison between the observation (green) and screen recording (orange) results	107
E.5	Alteryx workflow to examine the richness of the column Field Notes	109
E.6	Pivot Table (results) of the total per Classification (Category)	110
E.7	Pivot Table (results) of the total per Classification (Category) per Participant	111

List of Tables

2.1	A fragment of an event log, adapted from [6]	19
2.2	Summarisation of the data collection techniques including a comparison of the characteristics	21
3.1	The inclusion and exclusion criteria of the study selection process	23
3.2	Main points of the systematic literature review	24
3.3	Research participant profile overview	27

3.4	Predefined analysis order of the data collection techniques and the randomised order of the research participants per data collection technique	30
4.1	Differences between the results from observation and screen recording	32
4.2	(Online and offline) activities only recorded with the observation technique	34
4.3	(Short-lived) activities only recorded with the screen recording technique	34
4.4	The order of the results	35
5.1	Difference in Timestamps of the activities between observation and screen recording	36
5.2	Potential significance of the meetings	38
5.3	The percentage of the offline activity meetings that was not recorded per data collection technique	38
5.4	An overview of the Personal (offline) activities recorded by each data collection technique	39
5.5	An overview of the Tours (offline) activities recorded by each data collection technique	40
5.6	Clients failed to record per data collection technique	41
5.7	Client-related activities not recorded by the timesheet technique	42
5.8	Not recorded client-related (sub-)activities (hours and %)	42
5.9	Total time recorded by each data collection technique	43
5.10	Unable to categorise activities per Classification per data collection technique	44
5.11	Richness information in the Field Notes per data collection technique	45
5.12	The richness of the field notes (min/max) per data collection technique	45
5.13	Richness information in the Field Notes per Research Participant in the Timesheet	46
A.1	Overview of the data collection techniques	69
D.1	The standardised form used to record employee behaviour	92
D.2	The categories used to describe the distribution of work time	96
D.3	The purpose of the activities	96
D.4	The sub-activities used to obtain deeper insights into the purpose of the activities	97
D.5	An overview of the used mediums throughout the working day	98
D.6	Overview of contact persons with whom the employee interacted throughout the working day	99
D.7	Initiation of the activities	100
D.8	An example of the unique (anonymised) clients	101
F.1	Total Sum per Category	113
F.2	Total Sum per Research Participant per Category	113
F.3	Total Sum per Activity	115
F.4	Total Sum per Research Participant per Activity	115
F.5	Total Sum per Sub-Activity	117
F.6	Total Sum per Research Participant per Sub-Activity	118
F.7	Total Sum per Medium	121
F.8	Total Sum per Research Participant per Medium	122
F.9	Total Sum per Participant	125
F.10	Total Sum per Research Participant per Participant	126
F.11	Total Sum per Initiated	128
F.12	Total Sum per Research Participant per Initiated	128
F.13	Number of Words in the Field Notes per Research Participant	129
F.14	Total Sum per Research Participant per Client	131
F.15	Total Sum per Research Participant	133

Chapter 1

Introduction

“*What did you do today?*”. This is a simple question that may be presented to the employee by co-workers, management, or even the employee themselves. The behaviour of employees in the workplace is directly related to the success and operations of an organisation [64]. This has led professional services firms to base many of their key performance indicators (KPIs) on employee behaviour with the assistance of tools such as the timesheet [55, 56]. There is an assumption that there might be a discrepancy between what employees said they have done and what is actually observed throughout the working day [136]. Mills et al. [94] argue that employees tend to omit records that reflect negatively on their behaviour or only record the records that they deem to be important. Therefore, recording employee behaviour whilst carrying out business-related activities has become indispensable [14]. Scholars (see e.g., [66, 78, 86, 127, 140]) have expressed an interest in the study of recording employee behaviour to discover work patterns [86]. There is a variation in how work patterns are recorded depending on the sector (e.g., health, education, or IT) [143]. The commonality between studies is their use of a data collection technique to record “*the (characteristics) of work activities performed by the organisational members to execute specific activities and accomplish practices of interest related to a task*” [104, 107] against a set of predetermined categories.

Nowadays, business processes are increasingly supported by information systems [82] and involve various actors from multiple departments [20]. These information systems record detailed trails of the execution of tasks in databases, system logs, or records [19].

Within Business Process Management (BPM), Process mining has gained a lot of interest, both in research and practice [51]. Process mining exploits automatically recorded data to analyse and optimise business processes [137]. Research on process mining has resulted in a collection of techniques and tools [157]. The aim of process mining is to discover, monitor, and improve business processes by extracting knowledge from event logs, stored in various information systems containing records about the recourse(s) and business process(es) [3]. An event log consists of a set of cases. Each case consists of a sequence of events performed in the context of a single process [3, p. 128]. Event logs are the main input for the process mining techniques [84]. The assumption is that these records are an excellent representation of the actual employee behaviour to discover work patterns [5, 86]. However, Baier et al. [21] suggest that activities are not recorded (properly and/or accurately) by the employee or occur outside of the information system(s). Thus, it is not always evident to what extent the records reflect the actual behaviour of the employee(s).

1.1 Problem Statement

For organisations, it is essential to have a detailed recording of the behaviour of the employees and their activities [52]. A wide variety of data collection techniques can be used to record employee behaviour and provide rich, context-specific insights [63, 81]. Lethbridge et al. [78] categorised the data collection techniques as *direct techniques*, *indirect techniques*, or *independent techniques* based on the required degree of contact between the researcher and participant. Each data collection technique yields unique findings and information. However, the use of each data collection technique comes with its challenges. For example, the downside of a self-reporting technique such as the timesheet (independent technique) [78] is that employees can omit certain records that reflect negatively on their behaviour [94]. As a result, there is a disparity between what employees report

they have done and what actually happens during the course of a working day [26, 136].

There is an abundance of contemporary literature on stand-alone data collection techniques that can be used to record employee behaviour [130]. However, as described in [22, 86], there is a sizeable gap in the literature on recording employee behaviour within an organisational context by using multiple data collection techniques. The issue related to this knowledge gap is that previous studies have almost exclusively focused on one stand-alone data collection technique. This is of particular concern because each data collection technique yields unique results but also has its shortcomings which impact the accuracy or completeness of the results [78, 160]. Moreover, prior research has mainly focused on fields where the use of devices (such as a laptop) is not prevalent [9, 38, 95]. Therefore, there is a significant amount of literature on how to record employee behaviour using data collection techniques such as observation and interviews, whilst the use of screen recording is still an under-explored area [62, 121]. As a result, the existing literature lacks a demonstration of how multiple techniques can be combined to record employee behaviour based on the context of the situation [66].

1.2 Research Aim and Objectives

This study aimed to fill the gap by examining the wide variety of data collection techniques that can be used to record employee behaviour. Specifically, we aimed to discover work patterns within an organisational context where the use of a device (laptop) is prevalent. We combined three data collection techniques *non-participant observation*, *screen recording*, and *the archival analysis (timesheet)* with varying characteristics to record employee behaviour. Our objective was to present the (dis)advantages, and the commonalities and differences between the data collection techniques to examine the level of confidence that should be placed in the analysis of this type of data. We compiled these findings into three realistic scenarios to demonstrate how multiple data collection techniques can be combined depending on the context of the situation.

1.3 Research Question

The aforementioned issues necessitate an answer to the following research question (RQ).

RQ How can different data collection techniques aid in discovering work patterns of employees within an organisation?

In a pursuit to provide a substantiated answer to the main research question (RQ), the following four sub-questions (SQ) have been defined.

SQ₁ How is the phenomenon of work patterns defined in the literature?
SQ₂ What are the existing data collection techniques to record employee behaviour within an organisation?

We conducted a systematic literature review. First, we aimed to gain a deeper understanding of work patterns, their definition, and how we can record the work patterns. Second, we aimed to find data collection techniques that are suitable to record employee behaviour within an organisational setting. We used the categorisation of data collection techniques by Lethbridge et al. [78] to allocate the techniques: *Direct Technique* (interview, focus group, questionnaire, observation, think-aloud, and work diary), *Indirect Technique* (screen recording), and *Independent Technique* (analysis of event logs and archival analysis). For each data collection technique, we examined the following:

- What each data collection technique entails;
- The insights that can be derived;
- The (dis-)advantages, and the impact this may have on our study;
- The commonalities and differences between the data collection techniques.

As a result, we first defined the concept “work patterns” based on the scope of our research. Second, we had an extensive overview of data collection techniques suitable to record employee behaviour. We opted for the following three techniques, one from each categorisation by Lethbridge et al. [78]: *non-participant observation*, *screen recording*, and *the archival analysis (timesheet)*. We used the characteristics of these techniques to design the research method of the study.

SQ₃ Which work patterns can be derived by using multiple data collection techniques to record employee behaviour?

SQ₄ How can the data collection techniques be combined to record employee behaviour?

We conducted a multiple-case study with the aim to discover work patterns. We combined three data collection techniques *non-participant observations*, *screen recording*, and *archival analysis (timesheet)* to record the behaviour of six employees within the Helix Data Delivery Team of Ernst & Young. This resulted in a comparison of the (dis)advantages, commonalities, and differences in work patterns between the data collection techniques and research participants. We translated the insights into three realistic scenarios to demonstrate how multiple data collection techniques can be combined depending on the context of the situation.

1.4 Contributions

This study contributed to the existing body of knowledge by combining three data collection techniques *non-participant observation*, *screen recording*, and *the archival analysis (timesheet)* with varying characteristics to record employee behaviour, and discover work patterns within an organisational context [78]. Our study also contributed to the process mining community. This technique uses automatically recorded data stored in various information systems, to analyse and optimise business processes. We focused on the advantages, disadvantages, and the commonalities and differences between the data collection techniques. The implications for research and practice aid to understand the level of confidence that should be placed in the analysis of this type of data. We translated the insights into practical advice. We present 3 scenarios to illustrate how multiple data collection techniques can be combined to minimise the shortcomings of the techniques and to record employee behaviour based on the context of the situation.

This study also contributed to the (individual) employees, the management of the Helix Data Delivery Team, and similar organisations. Our study showed what can and cannot be recorded by each technique. Our study contributes to a more fair assessment of the performance of the employees. The employees can become more vocal about the rating system and the way that their behaviour is currently recorded (e.g., in the timesheet). Organisations can use the results to identify discrepancies between what employees claim to have done and what actually occurred throughout the working day, detect unwanted or repetitive tasks that can be (partially) automated in the future, and identify the disruptions throughout the working day (e.g., chat messages via Microsoft Teams) [25].

1.5 Thesis Structure

The remainder of this thesis is structured as follows. In **Chapter 2**, we present the Literature Review that introduced the phenomenon of work patterns and the data collection techniques suitable for recording the behaviour of employees within an organisational context. We used these findings to design the Research Method in **Chapter 3**. We conducted a multiple-case study to record the behaviour of six employees within the Helix Team. In **Chapter 4**, we analyse how we obtained and prioritised the results of the multiple-case study. We used these findings to highlight the results that provide key insights to answer the main research question. We describe the case study Results in **Chapter 5**. In **Chapter 6**, we present the implications for research and practice. We used the findings from the Literature Review and the Results of the multiple-case study to conclude to what extent our study has confirmed or refuted these findings. The insights gained during the previous Chapters were translated into practical advice in **Chapter 7**. We discuss 3 scenarios to illustrate how multiple data collection techniques can be combined to record employee behaviour based on the context of the situation. Lastly, **Chapter 8** presents the most notable findings of our study and an answer to our sub-questions and main research question.

Chapter 2

Literature Review

2.1 Content of the Literature Review

The first section of this chapter introduces the phenomenon of work patterns as defined in the literature. The remaining sections provide an overview of data collection techniques suitable for recording the behaviour of employees within an organisational context. We examined what each technique entails, the insights that can be obtained using these techniques, the (dis-)advantages, and the similarities and differences between the techniques. The categorisation and the allocation of each data collection technique are based on the research of Lethbridge et al. [78]. The data collection techniques are categorised based on the required degree of contact between the researcher and the participant. We used this categorisation due to constraints imposed by Ernst & Young that limited the participants' availability, as stated in Section 3.4. We conclude this chapter with Section 2.4 which includes a summarisation of the literature review findings to answer the sub-questions.

2.2 Work Patterns

The study of recording employee behaviour to discover work patterns has gained a lot of interest, both in research and in practice [86]. There is a great amount of variation in how work patterns are recorded depending on the sector (e.g., health or education) or the scope of the research (e.g., managerial behaviour or effective time-management) [143]. The commonality between studies is their use of a data collection technique to record employee behaviour against a set of predetermined classifications [86], by recording “*the observable incident*” [150, p. 82] or to “*uncover the movements and activities*” [49, p. 4] during a specific time frame. Bonazzi [30, p. 223] “*logged, minute by minute, the sequence of activities and events in which they were involved*”.

In general, work patterns are defined as “*the (characteristics) of work activities performed by the organisational members to execute specific activities, and accomplish practices of interest related to a task*” [104, 107]. This means the “*everyday nature of the work activities exhibited by the organisational members*” [53, p. 24]. The characteristics are the *where* and *when* the work activities are performed (location and time), *how* (the tools and documents), but also by *whom* (the involved members) [66].

The classifications introduced by Mintzberg [95] are one of the most often used for recording employee behaviour, and discovering work patterns in order to examine and compare the results [66, 141]. Within this study, Mintzberg [95] recorded the time, activity and duration to examine work patterns using a predetermined classification scheme consisting of (1) desk work, (2) scheduled meetings, (3) unscheduled meetings, (4) telephoning, and (5) tours in the organisation. Mintzberg [95] recommended revising the categories during and after the observation has taken place. Perlow used a variety of data collection techniques to record employee behaviour within an organisation [112] and investigate how the work patterns are controlled [111]. Perlow stated “*I observed everything the individual did, and I wrote down each activity as it occurred*” [111, p. 335]. Perlow specified, for each activity, whether it was an individual working pattern (i.e., one employee is involved), or an interactive working pattern (i.e., multiple employees are involved). The focus of the studies is on identifying work patterns by recording employee behaviour using a data collection technique. This is accomplished by recording the sequence of observable movements, events, and activities performed

by the employees, with a detailed timeline and duration, noting whether it was an individual or interactive activity. Depending on the activities observed by the researcher, a categorisation is created, allowing for the examination and analysis of various activities over time to discover work patterns.

2.3 Categorisation of Data Collection Techniques

Lethbridge et al. [78] categorise the data collection techniques based on the required degree of contact between the researcher and participant.

1. **Direct Technique:** The researcher must have direct access to the participant. We examined six direct techniques *interview (focus group), questionnaire, observation, think-aloud, and work diary*.
2. **Indirect Technique:** The researcher must have direct access to the working environment of the participants (e.g., (home) office). In comparison to the direct technique, the indirect technique does not require the researcher to interact with the participant. We examined one indirect technique *screen recording*.
3. **Independent Technique:** Involves a retrospective study of work artefacts such as event logs or archival sources. The records are not created by or for (the purpose of) the study. We examined two independent techniques *analysis of event logs and archival analysis*.

In the following sections, we discuss each data collection technique according to the aforementioned categories based on the research of Lethbridge et al. [78].

2.3.1 Direct Technique

To gain a deeper understanding of employee behaviour, the researcher must be in direct contact with the participants [78]. The direct data collection techniques provide a real-time depiction of the studied phenomena [127]. The downside is that some techniques are susceptible to the Hawthorne effect which may affect the results [86].

Interviews

One of the most commonly used data collection techniques for the collection of qualitative, descriptive, and in-depth data is the interview [11, 36, 65, 77, 78, 81, 98, 114, 118, 127]. The purpose of an interview is to uncover the participant's views, perspectives, and descriptions of the topics that are addressed [65]. The interview can obtain rich and detailed (historical) data from the participant's mind regarding past decisions and behaviour [114, 127]. Hak [58] and Busetto et al. [33] argue that interviews are used to obtain information on the participant's motivations, thoughts, and subjective experiences. The interview can be one-to-one [35, 81] or carried out in a group setting (focus group [106, 127]). The one-on-one setting allows the researcher to build a stronger connection with the participant [67, 81]. As a result, the researcher is able to address personal or (socially) sensitive topics. This becomes difficult during the focus group.

The qualitative nature of interviews allows for unexpected themes to emerge and be pursued by the researcher [33, 118]. The interactive nature of the interview allows the researcher to observe the visual and auditory cues of the participants [114]. In addition, the interview allows the participants to alternate between moments in time, and seek clarification on specific questions [77, 79]. The interview allows the participants to modify the line of inquiry, therefore assisting to overcome researcher bias [33, 118]. This is an advantage compared to the questionnaires. However, Smagorinsky [134] and Ericsson [45] mention that the interview (including the focus group) are retrospective techniques, hence the obtained data consists of recalling which may lead to inaccurate memories or wishful thinking. In addition, the process of preparing, conducting, transcribing, and analysing the interviews can be time-consuming [36, 67, 77, 78, 118, 121, 127].

Focus Group

An interview conducted in a group setting (i.e., typically consisting of three to 12 participants [106, 127]) is called a focus group [36, 77, 78, 114, 118]. This data collection technique obtains data through group interactions [67] on a topic chosen by the researcher [97]. The debates are supervised and facilitated [33] by the moderator(s) [47, 78, 81, 114]. The discussions begin broadly [35] and gradually focus on the subject of the research [36]. The focus group is becoming increasingly popular in qualitative research [114] because it allows for the simultaneous collection of data (e.g., a variety of facts, beliefs, attitudes, responses, perspectives, and explanations) [114, 118, 127] from multiple sources. The group setting and the direct involvement of the researcher [74, 78, 81] allows

participants to build on each other’s responses, increasing the depth of the obtained data [76]. Therefore, the focus group is a quick and cost-efficient technique for obtaining qualitative data and feedback from the participants [33, 127].

Nonetheless, there appears to be a consensus concerning the drawbacks of the focus group. The claim is that this data collection technique is prone to bias, and therefore, can produce unreliable results due to the group setting, the (lack of) experience of the moderator, and faulty questions [36, 78, 127]. The disadvantage of the group setting is that the focus group can be dominated by certain individuals [76, 114], conflicts can occur between different personality types [118], and the participants may only provide socially accepted responses due to social pressure [127]. Therefore, the focus group is not a suitable data collection technique to acquire personal or (socially) sensitive information [67, 81]. In addition, it can be difficult to organise a focus group to the conflicting schedules of the participants. A focus group focus session consists of three to 12 participants [77, 78], and can last two to three hours [127]. This becomes even more problematic when the participants have limited availability due to a high workload. As indicated in Section 3.4, certain sectors deal with peak periods (“busy season”) throughout the financial year. The limited availability combined with the group setting requires sufficient knowledge, and experience of the participants to be able to discuss all topics [127]. The moderator (researcher) must be well-trained, otherwise, the focus group will become unfocused [67, 78, 118, 127].

Questionnaire

Questionnaires are one of the most often used data collection techniques involving human participants [36, 98, 114, 118]. This technique can measure distinct variables such as behaviour, facts, preferences, attitudes [67], and information regarding a process [78]. The questionnaire is a document containing questions administered in a written format [98], and must be completed by the participants [35, 124]. The procedure and structure of the questionnaire are similar to the structured interview [152]. However, due to its non-interactive and impersonal nature [118], attention must be paid to its design [78, 114]. A limitation of the questionnaire that can affect the validity and reliability is the researcher and response bias [36, 125]. Researcher bias is affected by the characteristics of the interviewer (e.g., motivation, experience, and verbal or non-verbal cues [118]), or the characteristics of the questionnaire (e.g., wording and order of questions [77, 78, 106, 127]) can influence the replies of the participants. The response bias is affected by the characteristics of the participants (e.g., personality, motivation, memory, or experience). The questionnaire and interview (including the focus group) are retrospective techniques [36, 45, 77, 78, 127, 134]. The collected data is based on the self-reports (i.e., the memory) of the participant, which can lead to false memories or wishful thinking [11]. Robson & McCartan [118] add that participants are unlikely to accurately record their behaviour, attitudes, and beliefs due to the need to provide socially acceptable responses.

Observation

The observation technique allows to study of the behaviour of participants within the (working) environment [11, 22]. Observation provides detailed information. The data is based on the researcher’s observation [118]. This is an advantage compared to the work diary technique which consists of self-reported behaviour [33, 58]. The observation provides a real-time representation of the phenomena being studied, in contrast to screen recording which is “retrospective observation” [61, 62, 127]. Patton [108] proposes five dimensions of variation in observational approaches: (1) role of the researcher, (2) portrayal of the researcher’s role towards the participants, (3) portrayal of the purpose of the research to participants, (4) duration of the research observations, and (5) focus of the research observations.

The first dimension is concerned with the role of the researcher. Observations can be categorised on a continuum ranging from participant observation (active) to non-participant observation (passive) [47, 121]. In *participant observation*, the researcher (observer) is part of the observed setting [33], and strives to fully immerse themselves into the lives of the participants [78, 114], to acquire an insider understanding of the work patterns [47]. *Non-participant observations (passive)* involve no level of involvement [35, 114]. The researcher is (physically) present, yet does not partake in the observed setting [22, 33]. The downside of observing in a physical environment shared by the participant is the extent to which an observer affects the observed setting [118, 160]. The risk here is significantly higher during participant observation. First, the observer must assimilate the information gained from their observations before drawing any conclusions [23]. Second, the “Hawthorne” or “observer effect” [86, 106], suggests that participants may change their behaviour due to their awareness of

being observed [67, 125]. Barker [23] and Mintzberg [95] argue that these effects are insignificant because participants revert to their regular work pattern while carrying out familiar activities (i.e., their regular work). Moreover, the researcher can gain knowledge of employee behaviour without directly influencing the work process by performing the observation in a natural setting (e.g., the workplace), allowing the participant to be in contact with colleagues [27], and opting for the non-participant observation technique

The second and third dimensions are concerned with the disclosure of the scope of the research, and the extent to which participants are informed that they are being observed [35]. This can range from *overt observations* (i.e., participants are aware of the observation) to *covert observations* (i.e., participants are unaware of the observation) [108]. Moreover, the participants may receive a full explanation, a partial explanation, no explanation, or a false explanation [35]. Mack [81] states that there are no formal rules for disclosing the researcher’s involvement in research. However, Pickard [114] and Baker [22] stressed that covert observation should have no place in an academic study and that the (lack of) ethical guidelines is a concern. The fourth dimension is concerned with the amount of time spent gathering observational data [108]. The observation periods vary in length. Long-term observations can take place over the course of a full year or even longer, while short-term observations can last a day or less [35]. The duration of the research observations should be determined during the planning stages of the research [114]. The final dimension is concerned with the breadth of focus [114]. The observation can have a *narrow focus* where a single element or component within the observed setting is observed, or a *broad focus* where a holistic view of the entire environment including its elements is sought [35, 108]. Nonetheless, a drawback is that observation results in an overwhelming amount of data, which impacts the time and effort needed for the data analysis [116].

Think-aloud

Nunan [105] noted that the majority of the activities performed by the employees happen within their minds and therefore are imperceptible. The invisible work was once thought to be meaningless because researchers were only interested in the “*observable characteristics of human behaviour*” [105, p. 117]. Recent studies deem this to be insufficient. Thus, researchers seek to identify these mental models (processes) by employing introspective techniques, which involve documenting the thoughts, feelings, and reasoning process of participants to analyse how this affects their behaviour [105].

Think-aloud is an introspective technique [61] in which the participants are asked to concurrently verbalise their thoughts while performing a task [46]. This allows the researcher to access the short-term memory [38], and behaviour of the employees [45] to gather a detailed understanding of the phenomena. The participants are expected to voice all actions, thoughts, and interpretations of certain tasks [78]. The researcher collects the think-aloud on tape to be able to create (mental) models or find (work) patterns of activities [105]. The limitation of the think-aloud is that the participants have to split their attention between the tasks and the narrative [61, 62]. In addition, the participants are not familiar with the process of verbalising their thoughts (to others) whilst performing activities [38]. Ericsson & Simon [46] claim that the inner speech of the participants, and therefore their verbalisation of thoughts, is incomplete and includes speech errors. This accords with Vygotsky’s [148] study, which mentioned that inner speech is almost inaccessible to experimental inquiry because it is only intended for the participant. However, the think-aloud provides reliable data [38] because the thoughts are verbalised from immediate memory whilst the person is in the process of perceiving and retaining the information [134]. This is an advantage compared to the questionnaires, interviews, and focus groups [38, 61] because these are retrospective techniques, and the obtained data consists of recalling which may lead to inaccurate memories or wishful thinking [45, 134].

Work Diary

The work diary technique is used to collect qualitative data regarding participant behaviour, activities, and experiences [29]. The participants are required to self-report various events that occur throughout the working day [78]. The use of work diaries as a data collection within the field of information and computer science has been largely overlooked [114]. The work diary is a self-monitoring report. As a result, the contribution is considered to be of lesser value compared to researcher observations [11]. For this reason, Connaway & Radford [36] argue that it is beneficial to combine self-reporting diaries with other data collection techniques such as (researcher) observations or interviews. The quality and completeness of the data obtained from the work diary technique are dependent on the level of information provided in the participant’s description of their working day

[139]. For instance, Shull et al. [126] asked students to submit weekly progress reports about their work. The majority of the students opted to submit the reports sporadically or not at all. The work diaries rely on self-reports, and the participants are required to recollect events which may lead to significant issues regarding the accuracy (and reliability) of the data [78]. In addition, the need to constantly keep track of the performed activities places a high burden on the participants [77]. This can obstruct their ability to complete the work [160]. As a result, the work diary technique may record fewer details in comparison to the screen recording technique [139]. Lethbridge et al. [78] claim that the self-reporting characteristics of the work diary are an advantage because it records self-reports of occurrences on a continuous basis rather than in retrospect (e.g., the questionnaire and interview). Similar to the discussions of Smagorinsky [134] and Ericsson [45] regarding the think-aloud, the work diary technique reduces the possibility of retrospection. This is obtained by minimising the time between an experience and the report of said experience by the participant [29]. Moreover, the work diary technique allows events to be recorded in their natural setting [67] because the (physical) presence of a researcher is not required. However, Mintzberg [95] criticised the use of the work diary. The participants can omit data that reflect because the employee is likely to record the tasks which they deem to be important. This can lead to partial pictures of the behaviour of employees.

2.3.2 Indirect Technique

The indirect technique requires access to the participants' (working) environment in [77]. However, no interaction between the researcher and the participants is required [78]. The researcher initiates the data collection technique. This means that the participants perform their normal work whilst data is automatically recorded. Therefore, this technique requires a minimal time investment from researchers, except for the data analysis (e.g., transcription and coding) [127].

Screen Recording

A significant number of studies rely on the researchers being psychically present, whilst current work practices often involve communication (e.g., between employees and external parties), or performing activities enabled by a device such as a laptop or even multiple external monitors [62, 139]. Espedal et al. [47] note that the introduction of the screen recording technique is a reaction to the increased use of digital technologies which has challenged the "common" data collection techniques. This means that it has become difficult to observe employee behaviour to discover work patterns when it occurs in a computerised technology environment. However, the use of screen recording as a stand-alone data collection technique is still an under-explored area [62]. Thus far, within the field of computer and information science this technique has been predominantly used to explore mobile application usage ([72]), the interaction between participants and technology ([120]), or for usability testing of websites or computer games. For instance, Salinger et al. [122] used screen recording to investigate how graduate students program in pairs whilst sharing the same device.

Whilst the use of screen recording is *unobtrusive* (i.e., the researcher is not required to be physically present), it remains to be *invasive* (i.e., all interactions including private matters are recorded) [139]. The screen recording technique has raised issues related to privacy [10, 62, 139]. Organisations need to have a justifiable reason to monitor their employees in such a manner [10]. In fact, European privacy regulations such as the General Data Protection Regulation (GDPR) [48] set out strict guidelines and legal standards that organisations must adhere to. As described in [10, 72, 117, 139], the biggest hurdle was to solicit participation in the studies due to the issues related to privacy. Ho [62] comments that the participants from her study expressed concerns about exposing sensitive (private) information, and data related to third parties without their knowledge and informed consent (e.g., communication via email or Microsoft Teams). Another limitation is that the participants are required to install a screen recording application. This may raise issues depending on the capability of the device (e.g., access rights, processing performance, and available hard disk space).

Nevertheless, the advantages of the screen recording technique to unobtrusively obtain rich and empirical data on the behaviour and work patterns of the employees [72] outweigh the issues related to privacy [139]. In addition, in comparison to the observation (or shadowing) technique, the researcher is not required to be physically present in the (working) environment of the participants [77, 78, 127]. This can minimise the Hawthorne effect [125, p. 99]. There are some isolated cases in which the participants altered their behaviour. Tang et al. [139] recorded that some employees warned third parties that their screen is being recorded by stating "*can't be goofy, I'm being taped*". Lastly, the screen recording technique is flexible in terms of where the work behaviour can be recorded [139].

The only requirement is that the screen recording application works. This allows participants to record from the office or their homes (e.g., during COVID-19), which is not a possibility with the observation technique.

2.3.3 Independent Technique

Whereas the previous sections addressed techniques by which the researcher could elicit fresh data (i.e., either recorded by the researcher or the participants themselves), the following techniques deal with the use of existing material which was not created by or for the (study of the) researcher [78, 114, 118].

Analysis of Event Logs

Process mining is a cutting-edge part of Business Process Management (BPM) [51]. Process mining can be defined as “a technique that aims to improve operational processes through the systematic use of event data stored during the execution of the process” [1, 3]. Process mining has been adopted in numerous fields [44] such as insurance (see e.g., [156]), healthcare (see e.g., [119]), and financial services (see e.g., [32]). The family of techniques uses a combination of event data and process models [157], and aims to “exploit event data in a meaningful way” [3, p. 5]. The purpose of process mining is to provide insights about the process(es), identify bottlenecks and deviations, analyse (employee) performance, recommend the automation of repetitive work, monitor key performance indicators (KPIs) such as flow time and utilisation, and streamline the process(es) [4, 83].

Figure 2.1 illustrates the position of process mining, the links between the process models and the actual (operational) processes, and the event data recorded by the information system(s). Van der Aalst [3] differentiates between three types of process mining techniques *discovery*, *conformance*, and *enhancement*. The first type of process mining is (automated) *process discovery*, which requires no a-priori information [2, 5, 7]. It extracts knowledge from the event log. Its goal is to produce a model that matches the observed behaviour within an event log [2, 5, 7]. The second type of process mining is *conformance checking*. This technique compares the behaviour of an (existing) process model and the behaviour recorded within an event log (of the same process), to detect discrepancies, bottlenecks, or resource optimisation [1, 43]. The third type of process mining is *model enhancement*. The goal is to enhance or improve an (existing) process model by using the additional information in the event log. For example, the model can be extended by including timestamps (from the event log) to showcase bottlenecks or throughput times [2, 5].

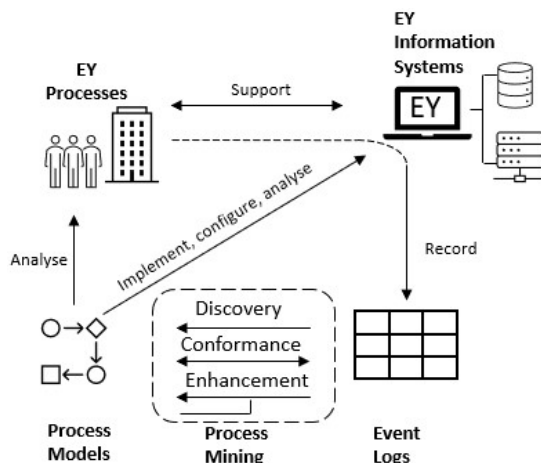


Figure 2.1: The three types of process mining: *discovery*, *conformance*, and *enhancement*, adapted from [1, 102]

The prerequisite of process mining is event data. The frequent use of computers and tools increases the availability of event logs [44]. The modern information systems record detailed trails of the supported business processes including the execution of tasks in databases, system logs, or records [19, 82]. An event log, also known as an audit trail or system log [137], consists of a set of timestamped cases [43]. Each case consists of a sequence of events performed in the context of a single process (i.e., task) [3, p. 128]. Event logs are the main input for process mining techniques [84]. An event log typically records the following information for each event: (1) an identification for each case (“Case ID”), (2) the activities that took place in each case (“Activity”), and (3) a reference to a point in

time i.e., the start and end of each activity within a case (“Timestamp”) [84]. In addition, the event log can contain additional properties such as the resource (i.e., person or device) who initiated or executed the activity, or domain-specific data recorded with the event (e.g., role, costs, or quantity) [3, 43]. Table 2.1 shows a fragment of an event log.

Table 2.1: A fragment of an event log, adapted from [6]

Case ID	Activity	Timestamp	Resource
...
6	Activity A	18-05-2022 09:56	nlsinik01
7	Activity A	18-05-2022 09:57	nlsinik01
8	Activity A	18-05-2022 09:58	gbrick01
8	Activity B	18-05-2022 09:59	gbrick01
6	Activity B	18-05-2022 10:00	nljans01
...

A weakness of this data collection technique is that the researcher is limited to the types (and quality) of the provided data [114]. The assumption is that recorded event logs, are an excellent representation of the activities in the real-life process(es) (i.e., cases) [5, 82]. In some cases, the tasks are not (properly) recorded or happen outside of the information system(s) [21]. As a result, it is not always clear to what extent the event log of the information system reflects the actual behaviour of the employees [57]. The issue of event log quality has been highlighted in several research studies (see e.g., [3, 69]), therefore it is deemed to be necessary to review the quality of the event logs prior to proceeding with the project mining project [1]. The Process Mining Manifesto [5] defined a star rating ranging from one star, meaning that the event log is of poor quality and is not suited for use in a process mining analysis. To five stars, which characterises the event log as complete, trustworthy, and of excellent quality. In previous research (see e.g., [31, 84, 137]), four categories of data quality issues in event logs have been identified: *missing data* (i.e., different kinds of information are missing from the event log); *incorrect data* (i.e., the data is logged incorrectly); *imprecise data* (i.e., the logged entries are too coarse); *irrelevant data* (i.e., the data is irrelevant for the context of the project, but can be used to derive relevant information through pre-processing and filtering). Process mining provides objective information based on factual data. This is an advantage compared to the observation technique, and the timesheet technique. The former technique is based on the researcher’s observation that consists of subjective information. The latter technique depends on the employee to provide an accurate, complete, and honest description of their behaviour in a given context [114]. According to Nicholas et al. [103, p. 251], event logs are a “*direct and immediately available record of what people have done: not what they say they might, or would, do; not what they were prompted to say; not what they thought they did*”.

Archival Analysis

The archival analysis consists of a record or a set of records [114], which sheds light on the past whilst aiming to impact the present [152]. The existing material was not created by or for the (study of the) researcher [159], allowing the researcher to unobtrusively collect relevant data [160]. Mills et al. [94] argue that archival sources can only be used as evidence if the researcher is aware of the (historical) context in which the record was created. An archival analysis can contain a wide variety of materials such as in-house documents, e-mail exchanges, and previously collected measurements in an organisation (e.g., timesheets) [118, 121, 159]. The archival analysis provides a broader picture of the studied phenomena [47]. The data is often readily available (in electronic format) [77], which simplifies the data collection, coding, and analysis compared to the direct and indirect techniques which require an investment of time by the researcher [78]. Most organisations require employees to digitally record information about the performed work, behaviour, and activities [29] using timesheets via (in-house enterprise) systems [127]. The self-reporting characteristic of the timesheet is similar to the work diary technique [86, 127]. However, the work diary is a direct technique, which means that the researcher obtains new data (recorded by the researcher or participants) [78]. The timesheet (an archival source) is an independent technique [78], which means that the archival records were not developed to provide data for research [114], therefore may restrict the quality of the research. The timesheets are mandatory for the organisation, hence must be filled in according to the guidelines provided by the organisation [118]. The (in-house enterprise) system requires the employees to supplement their daily records with activity codes or descriptions [13, 118]. A limitation is that researchers have no control over the quality of the manually entered information (e.g., in the

descriptions) about employee behaviour. In fact, Lethbridge et al. [78] found that descriptive fields were often not filled in (adequately) to amass sufficient information. Moreover, each organisation tends to use a different system (in-house or bought) resulting in dissimilarity between templates and data formats. This complicates the data collection, coding, and analysis at multiple organisations [78, 127]. Lastly, the timesheet is a self-report, therefore the employees can omit certain records that reflect negatively on their behaviour [94]. These limitations can be mitigated by accumulating additional evidence from sources other than employee self-reports [94, 160].

2.4 Synthesis

We used the findings from the literature review to provide an answer to the sub-questions.

Work patterns are defined as “*the (characteristics) of work activities performed by the organisational members to execute specific activities, and accomplish practices of interest related to a task*” [104, 107]. In this context, the work patterns are focused on the “*everyday nature of the work activities exhibited by the organisational members*” [53, p. 24]. The characteristics are the *where* and *when* the work activities are performed (location and time), *how* (the mediums and documents), but also by *whom* (the involved members) [66]. The data collection technique is used to record employee behaviour against a set of predetermined classifications. We used this information to answer the first sub-question *How is the phenomenon of work patterns defined in the literature?*

We then examined the data collection techniques suitable to record employee behaviour in an organisational context, to provide an answer to the second sub-question *What are the existing data collection techniques to record employee behaviour within an organisation?* The data collection techniques are structured according to the categorisations by Lethbridge et al. [78]. There are three categorisations based on the required degree of contact between the researcher and participant: *direct, indirect, and independent technique*. The selection of a data collection technique from each categorisation provides the advantage of gaining insights from different perspectives for this research. Table 2.2 presents the different characteristics, advantages, and disadvantages of the selected data collection techniques. The purpose is to demonstrate that the combination of the three data collection techniques will have a complementary effect. We found a significant amount of literature regarding the data collection techniques used to record employee behaviour e.g., interviews (focus group), questionnaires, observations, work diaries, and the analysis of event logs and archival analysis. We noted that the use of screen recording as a stand-alone data collection technique is still an under-explored area [62, 121]. This was quite surprising considering the increased use of digital technology in organisational settings. The use of screen recording as a data collection technique is expected to grow in importance [47]. Therefore, the use of screen recording in an organisational context can have a scientific contribution. We used Patton’s [108] five dimensions of variation in observational approaches to design the approach: *role of the researcher, portrayal of the researcher’s role towards the participants, portrayal of the purpose of the research to participants, duration of the research observations, and focus of the research observations*. We considered the impact of each dimension in order to design a suitable approach within the context and restrictions of the research.

We did not use process mining (analysis of event logs) during the multiple-case study. However, this technique has gained a lot of interest, both in research and in practice [51]. The aim of process mining is to discover, monitor, and improve business processes by extracting knowledge from event logs, stored in various information systems containing records about the recourse(s) and business process(es) [3]. We examined what the requirements are of a complete (event) log [84], and the impact on the recording once activities are not (properly) recorded or happen outside of the information system(s) [5, 82].

The literature review insights and the characteristics of the data collection techniques presented in Table 2.2 were used throughout the entire study. First, to design the research method of the study, as discussed in Chapter 3. Second, to select, prioritise, and highlight the results that provide key insights to the main question in Chapter 4. Third, to supplement the results in Chapter 5 with the strengths, weaknesses, differences, and impact on the quality of the records for the data collection techniques. Fourth, in Chapter 6 to present the implications to research and practice, and to conclude to what extent our study has confirmed or refuted these findings. And lastly, to translate the insights into practical advice based on 3 scenarios in Chapter 7. Table A.1 (Appendix A) presents a summary of the data collection techniques including a description, an example case, the (dis-)advantages, and the sources split by the synonyms found during the literature review.

Table 2.2: Summarisation of the data collection techniques including a comparison of the characteristics

Characteristic	Non-participant Observation	Screen Recording	Archival Analysis (timesheet)
Categorisation	This <i>direct technique</i> requires direct contact between the researcher and participant [78]. Non-participant observation allows a real-time representation of the studied phenomena [61, 127].	This <i>indirect technique</i> requires direct access to the participants' (working) environment [78]. Interaction between the researcher and participant is not required [61, 62]. Screen recording allows for a retrospective observation of the studied phenomena [62].	The previous techniques allow the researcher to elicit "fresh" data either recorded by the researcher or the participants themselves. This <i>independent technique</i> [78, 127] consists of existing material not created by the researcher or for the purpose of the research [114, 118].
Perspective	Observation consists of <i>subjective</i> information [73]. The data is based on the researcher's observation which is influenced by their perspective and what they deem to be important [118].	Screen recording is <i>more objective</i> compared to observation because this technique allows the researcher to repeatedly view the recordings which minimise the possibility of missing certain information [118].	The timesheet consists of <i>objective</i> information [159] because the participant records their own behaviour. This self-report consists of evidence-based information extracted from the timesheet [94].
Obrusiveness	Observation is <i>obtrusive and invasive</i> towards the participant [61]. Patton [108] stated that this is caused by the physical presence of the researcher (e.g., the researcher is seated next to the participant during the entire observation). Observation provides <i>detailed information</i> because both the online setting (i.e., on the screen) and the offline setting (i.e., physical environment) can be recorded. The downside is that not all activities are recorded (on time) due to unfamiliarity or the fast-paced environment [66].	Screen recording is <i>less obtrusive</i> because no contact between the researcher and participant is required [61, 62], yet <i>highly (privacy-)invasive</i> because everything on the screen(s) (both work-related and private) is recorded [72, 139]. Screen recording provides <i>highly detailed information</i> of the online setting because the recordings are permanent records of interactions that can be viewed repeatedly for later analysis [118]. This is an advantage compared to the observation, especially in a fast-paced environment or if the participant has multiple screens [66].	Archival analysis is an <i>unobtrusive technique</i> because the researcher is not present whilst the participant creates the archival records [78]. It may be invasive, depending on the kind of information that is shared [118]. This technique <i>does not provide detailed information</i> [62]. The data is not created for the research [78]. Moreover, the researcher has no control over the details of the data (e.g., the timesheet must be filled in according to the guidelines provided by the organisation) [118].
Richness of information	The "Hawthorne effect" [86, 106] suggests that participants may change their behaviour due to their awareness of being observed. However, these effects become insignificant because participants revert to their regular work patterns whilst carrying out familiar activities (i.e., their regular work) [23, 95] in a natural setting (e.g., the workplace) [85, 131].	Gaining access to a company can be challenging [66], especially if confidential information (e.g., client data) is involved. Moreover, the risk of using a screen recording application is that the recordings are not complete [73]. For example, it may be that the recording is not saved properly or that only a partial recording is available.	The timesheet is a self-report. Employees can omit records that reflect negatively on their behaviour [94] or only record the records that they deem to be important. The field notes are often not filled in (adequately) to amass sufficient information [78], providing only a partial picture of employee behaviour [95]. This risk can be mitigated by accumulating additional evidence from sources other than employee self-reports [160].
Risks			

Chapter 3

Research Method

This chapter discusses the research method, the used data collection techniques, and the threats to validity along with the tactics used to mitigate the threats.

3.1 Triangulation

We used triangulation as recommended by Denzin & Lincoln [39]. Patton [109, p. 1192] defined triangulation as “the use of multiple methods or data sources in qualitative research to develop a comprehensive understanding of phenomena”. Within our study, we used multiple methods and sources of evidence to increase the validity of the research, understanding of the topic, and confirmation of findings [26]. We applied the following methods:

1. **Literature review:** We conducted a systematic literature review to analyse the existing literature (Chapter 2).
2. **Multiple-case study:** We conducted a multiple-case study to record the behaviour of six employees within the Helix Team. Our aim was to discover work patterns by combining three data collection techniques *non-participant observations, screen recording, and archival analysis (timesheet)*.

Figure 3.1 illustrates the five phases of our research approach. Each phase will be explained in detail.

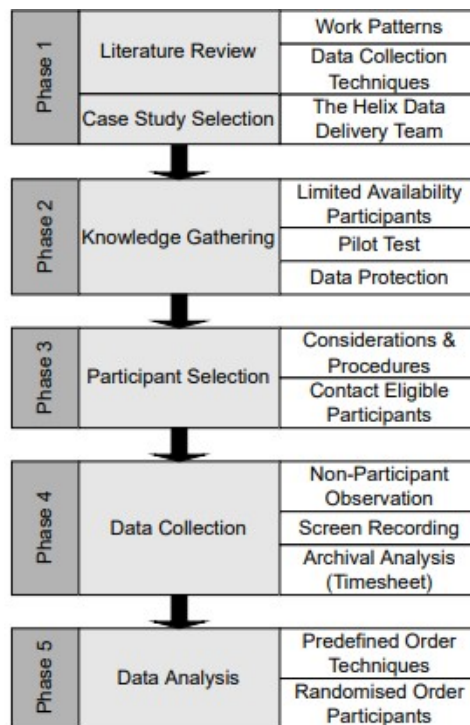


Figure 3.1: Research approach phases

3.2 Literature Review

We conducted a systematic literature review. First, to gain a deeper understanding of work patterns and how we can record them. Second, to find data collection techniques suitable to record employee behaviour in an organisational setting. Kitchenham [70] defined a systematic literature review as “a means of evaluating and interpreting all available research relevant to a particular research question, topic area, or phenomenon of interest” [70, p. 1]. We based the systematic literature review protocol on the guidelines presented by Kitchenham [70, 71] and Wohlin [154]. The aim of creating the literature study protocol was to improve the reproducibility of the research and minimise bias. Greenhalgh & Peacock [54] stated that a systematic literature review should not be solely based on a protocol-driven approach, due to the risk of missing relevant sources. To minimise this risk, we applied a hybrid search strategy in which we combined the use of a database search and snowballing (backward and forward snowballing) [100]. Throughout our search in digital databases, we used the following search terms:

- Data collection technique(s)
- Research method(s)
- Work pattern(s)
- Data collection method(s)
- Screen recording
- Process mining technique(s)

To exclude sources in which the search terms are not relevant to our study (e.g., a paper has a section named “data collection method” which is about the methods used for that particular research), we used the filter function to display sources that mention the search terms within the title. The search was restricted to literature written in English or Dutch; this criterion was in accordance with the language skills of the researcher. No restrictions were placed on the year of publication or geographical location. Moreover, to complement the database search, we applied backward and forward snowballing to search for additional papers. Wohlin [154] defined forward snowballing as “identifying new papers based on those papers citing the paper being examined”, and backward snowballing as “using the reference list [of the paper that is being examined] to identify new papers to include”. We proceeded with the search until theoretical saturation was reached and we were unable to find new (information regarding the) data collection techniques [80].

As stated by Kitchenham [70], the inclusion and exclusion criteria aim to identify sources that contribute to achieving the objective of the systematic literature review and to answer the research questions. Table 3.1 presents an overview of the inclusion and exclusion criteria of the study selection process.

Table 3.1: The inclusion and exclusion criteria of the study selection process

Inclusion criteria	Exclusion criteria
I1: Peer-reviewed studies and academic books	E1: Non peer-reviewed studies (e.g., magazine articles)
I2: Studies must be relevant to the research and stated goals of the systematic literature review	E2: Studies that are not focused on the goals of the research and systematic literature review
I3: Studies must be digitally accessible	E3: Studies are not digitally accessible
I4: Studies must be written in Dutch or English	E4: Studies are published in a language other than Dutch or English
I5: Studies may be published within any possible time frame or geographical location	E5: Duplicates are removed

We analysed the title, abstract, keywords, and conclusion (and when needed, the entire paper) of the selected papers within the long-list. We assigned each paper one of the following values: included, excluded, or doubtful (also according to the defined inclusion and exclusion criteria). The aim of assigning values to the papers was to create a short-list. Papers marked as “included” were immediately included in the short-list. Papers marked as “excluded” were immediately excluded from the short-list. Papers marked as “doubtful” were read in their entirety to see if they were relevant to the research. If they were deemed to be relevant, they would be marked as “included”, otherwise, they would be marked as “excluded” [153].

We documented the found sources in an Excel sheet, as can be seen in Figure B.1 (Appendix B). Each source marked as “included” was thoroughly read and coded. To systematically document the information, we created an overview in Word that illustrates the data collection techniques

(including synonyms) per source, as shown in Figure B.2 (Appendix B). This information was used to supplement Table A.1 (Chapter 2), which presents an extensive overview of the available data collection techniques. We present the main points from the systematic literature review in Table 3.2.

Table 3.2: Main points of the systematic literature review

<i>Goal</i>	First, to gain a deeper understanding of work patterns, their definition, and how they can be recorded. Second, to find data collection techniques suitable to record employee behaviour in an organisational context. We examined: <ul style="list-style-type: none"> • What each data collection technique entails; • The insights that can be derived; • The (dis-)advantages and potential impact on the study; • The commonalities and differences between techniques.
<i>Databases</i>	Google Scholar and Scopus
<i>Search terms</i>	Data collection technique(s), data collection method(s), research method(s), screen recording, work pattern(s), and process mining technique(s)
<i>Search strategy</i>	We used search terms via databases (i.e., Google Scholar and Scopus). To complement the database search, we applied backward and forward snowballing until theoretical saturation was reached.
<i>Method</i>	Analysed the title, abstract, keywords, and conclusion of the papers within the long-list. Each paper was marked with one of the following values: included, excluded, or doubtful (also according to the defined inclusion and exclusion criteria) to create the short-list.
<i>Documentation</i>	Synthesis (Chapter 2) and systematic literature review protocol (Appendix B)

3.3 Case Study

We conducted a multiple-case study and combined three data collection techniques *non-participant observation, screen recording, and the archival analysis (timesheet)* to record employee behaviour. We aimed to discover work patterns within an organisational context and to find (dis-)advantages displayed by each technique. The case study is a suitable method since it investigates and provides a deeper understanding of a contemporary phenomenon [159], in its own real context (e.g., the organisational context) [60, 121, 155], by using multiple sources of evidence [118]. An advantage of the case study approach is its flexibility and adaptability [34, 113]. This allowed us to use multiple sources of evidence (triangulation) and study the phenomenon from different angles to strengthen the evidence [121, 159].

The case study used a multiple-case design. The requirement is that the cases are similar [135, p. 6]. This requirement is met because we used three data collection techniques (non-participant observation, screen recording, and archival analysis) on six employees within the Helix Team to record their behaviour whilst looking at the same process (their behaviour throughout the working day). We opted for a multiple-case study to examine the (dis-)similarity within and between the cases and to discover work patterns by using multiple data collection techniques. We compared the results among the cases to increase the external validity of our findings [159]. The case study protocol can be found in Appendix C.

Setting

The case study was performed within the Helix Data Delivery Team at Ernst & Young Accountants (hereafter referred to as “EY”) from May 2022 to February 2023. Within this section, we present a brief overview of the organisation and the department.

Ernst & Young

EY is a globally operating service-oriented organisation whose purpose is to “*build a better working world*” [40]. At the moment, EY is operating in more than 150 countries and employs over 312,000 people through four integrated service lines. Within **Consulting** the teams are transforming businesses through the power of people, technology and innovation. **Tax** assists organisations and individuals by advising on tax and regulatory obligations. **Strategy and Transactions** provides clients with capital and transaction strategy through to execution to enable fast-track value creation

for inclusive growth. And lastly, **Assurance** supports clients in building stakeholders’ confidence and trust, managing regulatory responsibilities and driving long-term, sustainable growth.

Helix Data Delivery Team

The context of this research was set within the Helix Team, which is part of the Assurance service line. The Helix Team is divided into four self-managing squads (hereafter referred to as “squads”). Each squad is composed of squad members and one squad lead who is assigned to the planning activities of its squad. The Helix Team is responsible for the entire Extraction-Transformation-Loading (ETL) process [147]. The main tasks within the data delivery process include:

- **Planning:** The Helix Team works with a ticketing system, in which each ticket represents a deliverable of a specific client engagement. Once a ticket is assigned to a squad, the squad lead is required to assign the ticket to a member within their squad, and plan the needed hours for that particular deliverable. For each engagement, the squad member must confirm the scope (e.g., books closed date, business unit(s), and delivery date) with the audit team.
- **Data Extraction:** The squads extract the needed data (e.g., trial balance and journal entries) from the clients.
- **Transformation & Validation:** The squads transform the extracted raw data to be able to validate the data.
- **Delivery & Documentation:** The squads load the validated data into a Helix medium (e.g., Helix Platform or PrepEY, see Table D.5) and deliver the data cube to the audit team. The squads are required to prepare an extraction and validation memo, which covers the performed extraction methods and observations made throughout the validation process.

3.4 Knowledge Gathering

Phase two – the knowledge gathering phase – took place in May and June of 2022. During this phase, we gathered knowledge about the constraints and restrictions set forth by EY. We were faced with the following constraints: the limited availability of the Helix Team members, the stringent guidelines for using applications and software, and the strict rules about data protection.

Limited Availability

The Helix Team is part of the assurance service line, therefore is subject to an identical schedule. The workload within the assurance service line is seasonal and depends on the client’s fiscal year-end [41]. There are two peak moments (“busy season”) within the year: (1) from January to April, and (2) from September until the end of November. During the busy season, it is not uncommon for the Helix Team members to work (a minimum of) 60 hours per week.

Our data collection (phase 3) took place in September. We had to find data collection techniques that were appropriate to use during the height of the busy season. As noted in the literature review, the majority of the data collection techniques (e.g., interview (focus group), work diary, and think-aloud analysis) require active participation and engagement from the participants which can be time-consuming. As an example, the focus group or interview can last two to three hours. Due to the limited availability of the Helix Team, we decided to forgo the use of the aforementioned techniques. We favoured the following data collection techniques *non-participant observations, screen recording, and archival analysis (timesheets)* due to two reasons. First, other than initiating and ending the screen recording and sharing the content (i.e., recordings and timesheets) these techniques did not involve any active participation from the participants. Second, the Helix Team members perform the majority of their activities on their laptops. The use of screen recording allowed us to record the *online* behaviour.

Pilot Test

We faced a constraint regarding the recording software. We had to use software that EY had either created or approved. We used the application *Snagit* that we had to request via the *EY AppStore*¹. Snagit is a screen capture and recording software created by TechSmith². This application allowed the participants to record (a part of) the screen or even multiple screens. We conducted a pilot test, to confirm that the application can be requested via the EY AppStore and that multiple screens can be recorded. We recorded our two screens, including sound, for an entire working day.

¹The EY AppStore can only be accessed by EY employees and consists of software that has been approved by EY.

²Website TechSmith: <https://www.techsmith.com/snagit-features.html>

Data Protection

For the purpose of this study, we were allowed to observe the research participants and record their screens. We also had access to archival sources (i.e., timesheets). The collected data consisted of sensitive information such as client-specific data (e.g., the name of the client and their financial statements). Moreover, the rating system within the Helix Team is based on performance. This information could have been derived from the collected data. Therefore, we had the responsibility to protect the privacy and data of the clients and the research participants. We facilitated this by creating a set of ethical considerations and guidelines to be certain that all data is treated as confidential, as shown in Appendix C. The ethical considerations and guidelines were part of our presentation to the Helix Team members, as mentioned in Phase three (Section 3.5).

3.5 Participant Selection

Phase three – the participant selection phase – took place in June and July of 2022.

Considerations & Procedures

We aimed at attaining an appropriate sample size in terms of composition and size (i.e., the number of participants). To increase the generalisability, we selected a Helix Team member from each squad, with a range of roles (squad lead), ranks, years of work experience, and employment types, to create a diversified group of research participants. According to Sandelowski [123], the sample size should be large enough to foster new and rich data on the studied phenomenon, yet small enough to allow for in-depth case-oriented analysis. We used factors such as the accessibility and willingness of the participants [99], and the resource constraints (i.e., the available time of the researcher to perform the data collection and data analysis) to pre-determine our sample size (i.e., the number of research participants) [146]. To justify our decisions thoroughly and transparently, we estimated the minimum and the maximum number of research participants based on our availability and the time required for the data collection (September) and the data analysis (October) as illustrated in Figure C.3 (Appendix C). The Helix Team is made up of 10 employees. To ensure that the sample size is large enough to foster new and rich data, we set the minimum to be 50% of the total number of employees (five participants). Due to the sensitive nature of the data (i.e., client data), the data collection and analysis were solely performed by the researcher, which limited our resources. This meant that the sample could not correspond to the total number of employees (10). Based on the literature, we concluded that the data analysis would be most time-consuming due to the large amount of collected data needed to be transcribed, coded, and analysed [127]. We calculated the maximum number of participants based on the total availability of the researcher to conduct the data analysis (seven participants). Thus, the range of participants was set at five to seven participants.

Research Participants

During the case study, the data collection techniques were applied in a real-life setting (i.e., the working day) of the squads within the Helix Team. To be eligible to participate, the Helix Team member had to meet the following criteria:

- The participant had to be part of the Helix Team.
- The participant had to be available for one entire working day in September 2022.
- The participant had to be willing to support the research and consent for the researcher to use all three data collection techniques on the same (working) day.
- The participant had to agree to record their screen(s) and to share the recordings afterwards.
- The participant had to agree to share the timesheet of the working day the study took place.

Each Helix Team member was contacted separately. We created a PowerPoint presentation to inform the Helix Team members of the purpose of the research, the scientific and practical contributions, the three data collection techniques used to record the employee behaviour, and we included the set of guidelines used to protect the privacy and data of the clients and the research participants. An example of the PowerPoint presentation is illustrated in Figure C.4 (Appendix C). Each research participant was required to sign the informed consent form. Table 3.3 presents a broad overview of the participant profile including their rank and work experience (in years). Any information that could have reasonably identified the research participant (i.e., (squad) names) was anonymised.

Table 3.3: Research participant profile overview

ID	Squad	Squad Lead (Y/N)	Rank	Work experience (YRS)	Employment
P1	C	No	Staff	1,5 years	Part-time
P2	B	No	Staff	2,5 years	Full-time
P3	C	No	Staff	2,5 years	Full-time
P4	B	Yes	Senior	2,5 years	Full-time
P5	D	Yes	Senior	4 years	Full-time
P6	A	No	Staff	2 years	Full-time

Note: (Y/N = Yes or No) and (YRS = in years).

3.6 Data Collection

Phase four – the data collection phase – took place in September 2022.

Comparison of Techniques

Our data collection rests on the triangulation method because we used three data collection techniques: non-participant observation, screen recording, and archival analysis (timesheet). The main benefit of triangulation is the increased reliability of the findings of our study [159]. There are three main reasons for combining these data collection techniques:

1. The data collection took place during the busy season. As a result, our research participants had limited availability. Our three data collection techniques were appropriate to use during the height of the busy season because they do not require any active participation from the research participants.
2. The use of screen recording to record employee behaviour can have a scientific and practical contribution. First, the use of screen recording as a data collection technique for recording employee behaviour is an under-explored area [62, 121]. Therefore, our results can have a scientific contribution. Second, the Helix Team members perform most tasks on their laptops. Screen recording allowed us to record online activities.
3. We used the categorisations by Lethbridge et al. [78] to structure the data collection techniques. The categorisations *direct*, *indirect*, and *independent technique* are based on the required degree of contact between the researcher and participant. Each categorisation yields unique results and has different characteristics, and (dis-)advantages, as shown in Table 2.2 (Section 2.4). We opted for a data collection technique from each categorisation to gain insights from a variety of perspectives. The combination of these techniques had a complimentary effect. We also met the requirements of a case study because we used several complementary sources of evidence (data collection techniques) to obtain multiple perspectives on the phenomenon (recording employee behaviour to discover work patterns) [159].

Standardised Form

We used a standardised form for all data collection techniques per research participant to preserve uniformity, and allow the comparison and analysis of the findings across the techniques [66]. The standardised form used to categorise and analyse the data was based on the study by Mintzberg [95] because the classifications have been tested in a multitude of studies and allow comparison [141]. We tailored the classifications to the context of our study because Mintzberg’s research was published over 50 years ago and did not capture the reliance on technology. Figure D.1 (Appendix D) illustrates a snippet of the filled-in standardised form.

The standardised form used to record employee behaviour consisted of the following classifications:

- **Time:** We recorded the time when we observed the activity. We used a time frame of 09.00 to 17.00 (i.e., standard working day) with a time interval of one minute. We used the time to calculate the duration of each activity.
- **Category:** The category described the distribution of time throughout the working day e.g., Desk Work (i.e., all activities performed using a device excluding telephoning), or Personal (i.e., use of the personal mobile phone).
- **Activity:** This classification described the (purpose of the) activity e.g., Organisational Work (i.e., extraction, transformation, and validation of a dataset), or Scheduling & Administration (i.e., administration of the planning, scheduling meetings and data extractions, or updating the status of tickets in JennEY³).

³JennEY is an EY tool used for the management of tickets (e.g., planning)

- **Sub-Activity:** Similar to Mintzberg’s [95] contact record, we included the sub-activities to gain a deeper understanding of the purpose of the activities. For example, we distinguished between incoming and outgoing telephone calls, messages, and emails to show if the process is disrupted.
- **Medium:** The Medium consists of the mediums (or tools) that the employee used to perform the (sub-)activity. For example, the employee used Microsoft Teams to communicate with other participants, and Alteryx to clean, transform, and validate the client data.
- **Participants:** We recorded with whom the employees interacted (“participants”) throughout the working day e.g., Audit Team or Helix Team.
- **Initiated:** We recorded who initiated a certain activity e.g., the employee or the other party.
- **Duration:** We calculated the duration of each activity based on the classification “time”.
- **Field Notes:** We took field notes during the observation and screen recording. We used the descriptions provided by the employees (in their timesheets) for the timesheet field notes. The purpose of the field notes was to recall an activity after the observation had occurred so that it could be transcribed and coded.
- **Client:** We included this classification to record much time is spent on (non-)client-related work.
- **Participant:** Each research participant was given a unique identifier (e.g., P1 or P2) to generate the analysis as a whole and for each participant

Data Collection

The data collection took place in September 2022. We observed six Helix Team members, as shown in Table 3.3. Each research participant was subjected to three data collection techniques *non-participant observation, screen recording, and archival analysis (timesheets)* on one working day of their choosing.

Hawthorne Effect

The risk of a direct technique [78] such as observation is the extent to which we (“the observer”) affect the setting [118, 160]. The “Hawthorne” or “observer” effect suggests that participants can change their behaviour due to their awareness of being observed [67, 125]. Barker [23] and Mintzberg [95] argued that these effects are insignificant because participants revert to their normal work pattern after a certain period when carrying out familiar activities (i.e., their regular work). However, McCall [85] and Sirris et al. [131] suggested that the participant’s awareness of being observed is offset by factors such as the natural setting (e.g., being in the workplace) and contact with colleagues [27]. We conducted the study on the fourteenth floor of the EY Amsterdam office. We deliberately chose this floor because this is the natural setting for the employees and due to its open-plan design (the colleagues are seated next to each other).

Non-participant observation

We observed the participant between 09.00 and 17.00 except during breaks (e.g., lunch or restroom breaks). As proposed by Pickard [114], we defined the observation duration upfront. This allowed us to record the employees for the same amount of time. To record the employee behaviour, we used Mintzberg’s [95] standardised form. We revised the standardised form while the observation took place. During the observation, we used two screens. On the screen turned towards the participant, we opened a medium unrelated to our study e.g., Microsoft Outlook, Microsoft Teams, or the Internet Browser. We opened the standardised form on the screen turned away the participant. Our goal was to minimise the observer effect by pretending that we are not actively observing the activities performed by the participant. At the end of the working day, we would ask the participant to stop recording and share the recordings. Additionally, we briefly evaluated the participant’s experience of the observation.

Screen Recording

While the observation took place, the participant recorded their screen(s) (i.e., laptop and external monitor) using the application *Snagit*. Figure C.1 (Appendix C) illustrates the recording application. The participant would record the screen(s) *and* sound (microphone and system audio). The participant initiated the recording at 09.00 and stopped the recording at 17.00. To share the recordings, the participants used *Media Shuttle*⁴ which is an encrypted data-sharing tool approved by EY. We verified if we had received all recordings and if the recordings were complete (e.g., the recording was saved properly without any disruptions). We stored the recordings on our device and we made a backup on OneDrive (encrypted by EY).

⁴Website Signiant: <https://www.signiant.com/products/media-shuttle/>

Archival Analysis (Timesheet)

Each participant recorded their behaviour in a timesheet using an in-house system called Mercury. All EY employees are required to digitally record information about the performed activities throughout their working day. The timesheet consists of an engagement ID (e.g., client work), activity ID (e.g., planning activity), the total duration spent on an activity, and a short description of each activity filled in according to the guidelines provided by EY. An example of the timesheet is illustrated in Figure C.2 (Appendix C). We used their description as the field notes. The participants exported their timesheets the week after the observation. We converted the format of the timesheet from PDF to Excel.

Data Storage

We stored the data in separate folders split per participant and data collection technique to prepare the data for the data analysis (transcription and coding). We stored the data on our device and we made a backup on OneDrive (encrypted by EY). Any data collected during the non-participant observations, screen recording, and archival analysis (timesheet) that could reasonably identify the research participant or client were anonymised before the results were shared. We used gender-neutral pronouns (e.g., “they/them” instead of “he/him” or “she/her”) and we assigned a unique code to the research participants (e.g., P1 or P2) and clients (e.g., 1A or 2D).

3.7 Data Analysis

Phase five – the data analysis phase – took place in October 2022. The likelihood that we could have been influenced while analysing the collected data of each research participant per data collection technique was minimal. This implies the risk of the researcher being influenced by using the data gathered from one data collection technique to supplement the data from another data collection technique. However, to ensure the validity of the research, we split the data analysis into two steps, which will now be discussed:

(1) Predefined the data analysis order of the data collection techniques:

We analysed the collected data based on the following order. First, we analysed the data collected from the non-participant observation. We recorded our observations in the standardised form. Therefore, we already transcribed (a part of the data) during the data collection [22, 33].

We used the findings from the literature to decide which data collection technique would influence us the least. The archival analysis is a self-report consisting of evidence-based information extracted from the timesheet [94]. The screen recording captures all activities depicted on the screen [62]. Therefore, we concluded that the data from the screen recording is more comprehensive compared to the data from the archival sources (timesheet). To minimise the risk of being influenced, we decided to first analyse the timesheets and then the screen recording.

Second, we analysed the data collected from the archival analysis (timesheets) that provided a broader picture of the studied phenomena [47] from the perspective of the research participant (i.e., self-report) [86, 127]. Lastly, we analysed the data from the screen recording that presented rich empirical data regarding the behaviour and work patterns of the employees [72]. This technique allowed us to (re-)watch the recordings until the entire working day was transcribed and coded.

(2) Randomised the data analysis order of the research participants per data collection technique:

We randomised the analysis order of the collected data of the research participants per data collection technique (archival analysis and screen recording). To minimise the risk for the researcher to create connections between the data of a research participant across the data collection techniques and use this information to supplement the data (from another data collection technique). For the non-participant observation, we used the order taken during the data collection. We used a random number generator tool⁵ to generate a unique and randomised order for the archival analysis and screen recording. Table 3.4 illustrates the predefined analysis order of the data collection techniques (see column “Order”) and the randomised order of the research participants per data collection technique (see column “Participant order”).

⁵Website Random Generator: <https://www.randomlists.com/random-numbers?dup=false&qty=6&max=6&min=1>

Table 3.4: Predefined analysis order of the data collection techniques and the randomised order of the research participants per data collection technique

Order	Data collection technique	Participant order					
1	Observation	1	2	3	4	5	6
2	Archival Analysis	2	5	1	3	6	4
3	Screen Recording	4	1	2	6	3	5

We used two mediums for the data analysis. First, we used Alteryx⁶, which is a data analytics solutions desktop tool used to transform and validate the data [12]. We have used Alteryx to decrease the number of manual actions (e.g., merging files by copy-pasting) to reduce errors and improve the reliability of the data analysis (see Figures E.1, E.2, E.3, E.4, and E.5 in Appendix E). Second, to summarise the large amounts of data, we generated pivot tables in Microsoft Excel (see Figures E.6 and E.7 in Appendix E) We describe the analysis performed with each medium in Appendix E. In Chapter 4, we continue the discussion on how the data was generated and analysed.

The multiple-case study allowed us to collect data from six participants using three different data collection techniques [121]. We transcribed, coded, and grouped the data gathered during the non-participant observation, screen recording, and archival analysis (timesheet). By using the standardised form, we facilitated the comparison of commonalities and differences in work patterns between the data collection techniques and participants [26]. For example, there could have been a difference between what employees said they had done (timesheet) and what was observed throughout the working day [136].

3.8 Threats to Validity

Yin [159] distinguishes four types of validity threats that can be assessed to evaluate the validity of the research. We used the definitions of the threats as introduced by Yin [159]. The threats to validity including the tactics to mitigate the threats were formulated before the research was conducted to minimise the possible threats to validity.

Construct validity: This aspect of validity is concerned with identifying correct operational measures for the concepts being studied. Several measures were undertaken to enhance the construct validity. First, we conducted a systematic literature review and we proceeded with the search until theoretical saturation was reached and we were unable to find new data collection techniques [80]. The insights gained in the extensive literature review were used to design the research approach. Second, we used multiple sources of evidence (non-participant observations, screen recording, and archival analysis) to record employee behaviour. The main benefit of triangulation is the increased reliability of the findings because “multiple sources of data provide multiple measures of the same phenomenon” [158, p. 133]. Moreover, we noted that the archival analysis (timesheet) is a self-report, which means that the employees can omit certain records that reflect negatively on their behaviour [94]. Mills et al. [94] and Zwijze-Koning et al. [160] stated that this risk can be mitigated by accumulating additional evidence from sources other than employee self-reports. Therefore, we combined the archival analysis (timesheet) with two techniques that are not based on employee self-reports (i.e., non-participant observation and screen recording).

Third, due to the sensitive nature of the data (i.e., client data), we were unable to involve multiple evaluators. Therefore, the data collection and analysis were performed by one researcher. The use of “investigator triangulation” [109, 158] could have minimised bias [118, 121]. To mitigate this threat, we split the data analysis into two steps. First, we predefined the data analysis order of the data collection techniques. We have used the findings from the literature review to determine the data analysis order i.e., we started with the technique that would influence us the least. Second, we randomised the data analysis order of the research participants per data collection technique (for the archival analysis and the screen recording). By doing so, we mitigated the risk of creating a connection between the data of a research participant across all data collection techniques and using this information to supplement the data (from another data collection technique).

Lastly, the use of screen recording raised issues related to privacy [10, 62, 139]. We created a set of guidelines which have been approved by EY, this can be found in Appendix C. We presented these

⁶Website Alteryx: <https://www.alteryx.com/>

guidelines to the Helix Team. Moreover, each research participant had to sign an informed consent form. We anonymised any information that could identify the client or research participant. In the study of Ho [62], the participants expressed concerns about installing screen recording software. Therefore, we only used software that is approved by EY and can be requested via the EY AppStore.

Internal validity: This aspect of validity is concerned with seeking to establish a causal relationship, whereby certain conditions are believed to lead to other conditions, as distinguished from spurious relationships. First, to be certain that each data collection technique is applied identically for each participant throughout the entire research, we used the insights gained from the literature review to standardise the set-up of the three sources of evidence (non-participant observations, screen recording, and archival analysis) [118, 158]. We used a single standardised form for all data collection techniques to preserve uniformity and allow the comparison and analysis of the findings across data collection techniques [66]. Second, during the observation, we were within the same (physical) environment as the participant. This could have introduced the Hawthorne effect [86, 106], which means that the participant can change their behaviour due to their awareness of being observed [67, 125]. Barker [23] and Mintzberg [95] argued that the Hawthorne effect is insignificant because participants quickly revert to their normal work pattern. To minimise the awareness of being observed, we conducted the observation in the natural setting of the participant (i.e., the fourteenth floor in the EY Amsterdam office), as suggested by McCall [85] and Sirris et al. [131]. Moreover, we used two screens during the observation. On the screen turned towards the participant, we opened a medium unrelated to our study e.g., Microsoft Outlook. We opened the standardised form on the screen turned away from the participant. As proposed by Pickard [114], we set the duration of the observation period between 09.00 and 17.00 which is a standard working day within the Helix Team. This allowed us to record each participant for the same amount of time.

External validity: This aspect of validity is concerned with to what extent the results of the study can be generalised. We conducted a multiple-case study, in which three data collection techniques (non-participant observation, screen recording, and archival analysis) were applied in a real-life setting on six employees (from different squads within the Helix Team). By recording their behaviour while looking at the same process throughout one working day, we were able to compare our results and examine the (dis-)similarity within and between cases [158].

Reliability validity: This aspect of validity is concerned with demonstrating that the operations of a study can be repeated with the same results. The risk is that the research cannot be repeated in the same way or produce the same results. Therefore, we used (case study) protocols and guidelines (see e.g., [70, 71, 154, 159]) to mitigate this threat. We explained each performed step, including our reasoning, in an explicit and detailed manner. We aimed to be transparent about how the findings from the literature review, constraints, and decisions have impacted our research design. For example, the constraints given by the stakeholders from EY, the justification of our sample size, and how the participants were selected and informed. We also explained how we created the standardised form, and how we performed the data collection and analysis.

In the next Chapter, we present how we prepared, analysed, and prioritised the large amount of data generated during the multiple-case study. A thorough understanding of the (relevance of the) collected data was required before we were able to discuss the results.

Chapter 4

Analysis

In this Chapter, we discuss how we obtained, prepared, aggregated, selected, and prioritised the results of the multiple-case study. For all 6 participants over the course of 6 working days, we collected 136:04 hours of data during the non-participant observation technique (43:49 hours), screen recording technique (43:25 hours), and the timesheet technique (48:50 hours). In total, the techniques recorded 849 activities. The analysis was necessary to be able to prioritise the results because the multiple-case study generated a large amount of data.

4.1 Data Preparation

We used Alteryx to prepare the transcribed and coded data, as discussed in Appendix E. First, we checked if our data files contained empty fields, which would influence our data analysis. Once we established that there were no empty fields, we merged the 18 separate data files (one file per participant per technique) into 3 files (one file per data collection technique containing all participants). We performed two additional analyses in Alteryx. First, we compared the results between the observation technique and the screen recording technique. This will be explained in the next Section. Second, we generated the minimum, maximum, and average number of words in the field notes per data collection technique. The aim was to determine whether there is a discrepancy between the richness and completeness of the column *Field Notes* between the data collection techniques. We then used Microsoft Excel for the remaining analysis. To summarise the large amounts of data, we generated two pivot tables per classification: total per classification, and total per classification per participant. We repeated this for each classification per data collection technique. The results can be found in Appendix F.

4.2 Comparison

In Alteryx, we used the Inner Join tool to join the records of the observation and screen recording. As a result, our Excel sheet contained both the observation records (left in green) and the screen recording records (right in orange), as shown in Figure E.4 (Appendix E). This allowed us to make a one-on-one comparison between the recorded activities of the observation and the screen recording technique. We compared the records of all classifications in the standardised form (e.g., Time, Category, Activity, Duration, and Field Notes). In total, we found 58 differences between the observation and screen recording results, as presented in Table 4.1. We categorised these differences into the *accuracy, quality, and particularities and shortcomings*. In the following Sections, we will describe the differences found in each category.

Table 4.1: Differences between the results from observation and screen recording

Differences between observation and screen recording	
Description	Count
<i>Accuracy</i>	
Difference when the activity was recorded	15
Continued on next page	

Table 4.1 – continued from previous page

Differences between observation and screen recording	
<i>Quality</i>	
Activity <i>only</i> recorded with observation	21
Activity <i>only</i> recorded with screen recording	11
<i>Particularities and shortcomings</i>	
Difference in categorisation	2
Categorisation was <i>not</i> possible with screen recording	4
Categorisation was <i>only partially</i> possible with screen recording	5
<i>Total differences</i>	
Total	58

Due to a difference in the format of the standardised form, we were unable to compare the timesheet records in a similar manner. To be able to compare and include the results from the timesheet technique, we used the results presented in Appendix F.

4.3 Accuracy

We analysed how accurate the reference was to the point in time that the activity was recorded. We compared when the activities were recorded (i.e., the classification “Time”) of the observation technique and screen recording technique. We found 15 differences. This is an important insight because a difference in timestamp impacts (1) the duration of the activity, (2) the possibility to reconstruct the employee behaviour, and (3) the possibility to identify the discrepancies and bottlenecks [114]. As discovered in the literature review, process mining aims to discover, monitor, and improve processes by extracting knowledge from event logs stored in various information systems [3]. Event logs are the main input for process mining techniques and should contain at least the following three attributes: (1) an identification for each case (“Case ID”), (2) the activities that took place in each case (“Activity”), and (3) a reference to a point in time i.e., the start and end of each activity within a case (“Timestamp”) [84].

4.4 Quality

Data quality issues such as missing data (i.e., different kinds of information are missing from the records or event log) [31, 84, 137] can result in large errors or inadequate results [42]. We analysed the quality of the observation technique and screen recording technique. In total, we found 32 differences. We compared the standardised form records to find the type of differences for each data collection technique.

4.4.1 (Offline) Activity only recorded with observation technique

We compared the standardised form of the observation technique and the screen recording technique. We found 21 activities that were only recorded with the observation technique. In Table 4.2, we present the category, type, duration, and the number of times (count) that the activity has been recorded. We noted that the 21 activities fell under four categories *Meeting (offline)*, *Personal (offline)*, *Tours (offline)*, and *Desk Work (online)*. The 10 activities related to Desk Work (online) are activities that were only recorded by the observation technique because the process was “disrupted” by one of the 11 activities of the aforementioned categories. Therefore, we analysed the (offline) activities related to the categories *Meeting*, *Personal*, and *Tours*.

The remaining activities that are only recorded with the observation technique are *offline activities*. This means that these activities did not occur on the screen. We used Table F.1 (Appendix F) to compare the total sum (in duration and count) of the categories recorded by the screen recording technique and the timesheet technique. We concluded that there is a discrepancy in the records. We noted that the screen recording technique completely failed to record the category “Meeting” and that there were some discrepancies in the remaining categories “Personal and Tours”. The timesheet technique partially failed to record the category “Meeting” and completely failed to record the categories “Personal and Tours”.

Table 4.2: (Online and offline) activities only recorded with the observation technique

Category	Type	Duration	Count
Meeting	Offline	00:55 hours	3
Personal	Offline	00:30 hours	5
Tours	Offline	00:28 hours	3
Desk Work	Online	01:41 hours	10
<i>Note:</i> The duration is shown in hours.			

4.4.2 (Short-lived) Activity only recorded with the screen recording technique

We compared the standardised form of the observation technique and the screen recording technique. We found 11 activities that were only recorded with the screen recording technique. In Table 4.3, we present the sub-activity, duration, and the number of times (Count) that the activity has been recorded. We noted that 10 out of 11 activities are *short-lived activities*. Only one activity had a duration of 26 minutes (Sub-Activity “Data Transformation & Validation”). We compared the standardised forms again. We noted that 1 minute prior to this activity, the process was “disrupted” by one of the 10 remaining activities that were only recorded with the screen recording technique (Sub-Activity “Incoming Message”). Therefore, we analysed the (short-lived) activities.

Table 4.3: (Short-lived) activities only recorded with the screen recording technique

Description (Sub-Activity)	Duration	Count
Incoming Mail	1 minute	1
Incoming Message	1 minute	3
Outgoing Message	1 minute	3
Outgoing Message	2 minutes	1
Outgoing Call	1 minute	1
Technical Issues	1 minute	1
Data Transformation & Validation	26 minutes	1
<i>Note:</i> The duration is shown in minutes.		

We selected the 10 activities that were only recorded with the screen recording technique. To analyse the type of activity, we inspected each classification in the standardised forms of the observation technique and the screen recording technique. We noted that a short-lived activity (Sub-Activity “Incoming Message”) was related to client “5M”. We used Table F.14 (Appendix F) to compare the total sum (in duration and count) of clients per research participant. We concluded that there is a discrepancy in the records. We noted that the observation technique failed to record 2 clients, the screen recording technique failed to record 1 client, and the timesheet technique failed to record 21 clients. We discovered that the 21 clients that the timesheet technique failed to record, were all recorded by the screen recording technique. Therefore, we used the standardised form (in Microsoft Excel) of the screen recording technique to filter the 21 clients. This produced a record of 60 activities that the timesheet technique failed to record. We used these records to examine the type of sub-activities that were not recorded by the timesheet technique.

4.4.3 (Overtime) Activity only recorded with the timesheet technique

Once we had analysed the activities that were recorded either with the observation technique or the screen recording technique. We wanted to examine if there were activities that were only recorded with the timesheet technique. We were unable to conduct a one-on-one comparison, identical to the observation technique and screen recording technique since the standardised form of the timesheet technique has a different format. To compare the records of the three techniques, we used the Tables in Appendix F. The variance in the total time (“Grand Total”) recorded by each technique is one factor that we noticed in each Table. First, we noticed a 24-minute difference between the records of the observation technique and the screen recording technique. For the analysis of this difference, we examined the 21 activities that were only recorded with the observation technique, as shown in Table 4.2. We found 1 (offline) activity that falls under the category Tours that has a duration of 24 minutes. We then compared the standardised forms of the observation technique and the screen

recording technique. We noted that this activity was only recorded by the observation technique. Based on the observation records, we noted that participant 5 started their working day at 10:25. Participant 5 stood outside of the office for 24 minutes due to a fire drill. The next activity recorded by both the observation technique and the screen recording technique was at 10:49. This explains the 24-minute difference. Second, we noticed a difference of more than 5 hours between the observation and screen recording technique, and the timesheet technique. We categorised this difference as *overtime activities* because only the timesheet technique was able to record these activities. This difference is caused by 5 participants (P2 - P6). Based on Table F.15 (Appendix F), we noted that the variance of participants P2 - P5 are ranging from a few minutes to 01:32 hours. However, P6 had a difference of 02:30 hours. We examined the total sum per Research Participant per Activity, (Table F.4), the total sum per Research Participant per Sub-Activity (Table F.6), and the total sum per Research Participant per Clients (Table F.14).

4.5 Particularities and shortcomings

We analysed the particularities and shortcomings of the data collection techniques, and how this impacted the records in the standardised form. We found 12 differences between the records of the observation technique and the screen recording technique. We noted that there were three types of differences: *difference in categorisation*, *categorisation was not possible*, and *the categorisation was only partially possible*.

We found 2 differences in categorisation between the observation technique and the screen recording technique. Based on the observation records, these activities were categorised as offline activities (Tours) with a sub-activity “Socialising”. However, the screen recording technique categorised these activities as Desk Work (online activities). The field notes of the observation and screen recording technique indicate that the employees were socialising with participants while simultaneously performing desk work-related activities. The difference between the observation technique and screen recording technique to record online and offline activities has already been examined. In total, there were 9 differences regarding the ability of the screen recording technique to categorise the activities. We noted that the screen recording technique failed 4 times to *completely* categorise an activity. The remaining 5 times, the technique failed to *partially* categorise an activity. Once we had established that there is a discrepancy in the ability of the techniques to categorise an activity, we aggregated the results from Tables F.1, F.3, F.5, F.7, F.9 (Appendix F), to calculate the percentage of the activities that the observation technique, screen recording technique, and timesheet technique were unable to categorise. The observation technique did not fail to categorise any activities. We used the field notes of the screen recording technique and timesheet technique to examine what caused the discrepancy. We also looked at the design of the timesheets provided by the employees. The screen recording technique was unable to categorise the activities due to a technical issue with the recording tool Snagit. The timesheet technique was unable to categorise the activities due to a lack of information that can be extracted from the timesheet. This is caused by the current design and set-up of the timesheet in Mercury, and by the limited information provided by the participants in the description of the timesheet.

We used these findings to select, prioritise, and highlight the results that provide key insights to answer the main research question. We describe the results in more detail in the next Chapter. As shown in Table 4.4, the results in Chapter 5 are presented in the same order as the findings presented in this Chapter.

Table 4.4: The order of the results

Accuracy	
Timestamp	Section 5.1
Quality	
Online versus Offline Activities	Section 5.2
Meeting	Sub-Section 5.2.1
Personal	Sub-Section 5.2.2
Tours	Sub-Section 5.2.3
Short-lived Activities	Section 5.3
Overtime Activities	Section 5.4
Particularities and shortcomings	
Uncategorised Activities	Section 5.5

Chapter 5

Results

The previous Chapter described how we obtained, prepared, analysed, selected, and prioritised the results. We used these findings to select the key insights to answer the main research question. The objective of this chapter is to describe the results of the multiple-case study. We used three data collection techniques *non-participant observation*, *screen recording*, and *archival analysis (timesheets)* to record the behaviour of six employees within the Helix Team. The results of the multiple-case study are presented in Appendix F. We generated two Tables per classification: total per classification, and total per classification per research participant. This Chapter discusses the results related to the timestamp, online versus offline activities, short-lived activities, overtime activities, and un-categorised activities. Each Section ends with a recap of the strengths, weaknesses, differences, and impact on the quality of the records per data collection technique. We used these findings for the implications in Chapter 6.

5.1 Timestamp

The timestamp is one of the minimum requirements of an event log [84]. This information can be used to calculate the duration of an activity, discover bottlenecks, or analyse the performance of the resources (e.g., employees) [1, 3]. We compared the records in the standardised form of the observation technique and the screen recording technique. We found 15 differences between the moment in time that the activity was recorded (“timestamp”). As shown in Table 5.1, the observation technique recorded the activity *1 minute before*, *1 minute later*, *2 minutes later*, or *3 minutes later* than the screen recording technique.

Table 5.1: Difference in Timestamps of the activities between observation and screen recording

Description	Count
Observation recorded the activity <i>1 minute before</i> than screen recording	2
Observation recorded the activity <i>1 minute later</i> than screen recording	9
Observation recorded the activity <i>2 minutes later</i> than screen recording	2
Observation recorded the activity <i>3 minutes later</i> than screen recording	2

We looked at the recording in the standardised form to see what caused this *variance* in differences between the moment in time that the activity was recorded using the observation technique and the screen recording technique. We first looked at the differences where the observation technique recorded the activity *1 minute before and 1 minute later* than the screen recording technique. The activities or the Field Notes did not indicate what caused this difference. However, during the literature review, we established that the observation technique consists of *subjective* information [73]. The observation records are based on the researcher’s observation. This is influenced by their perspective, observation, and what they deem to be important [118]. Moreover, there is also a *time lag* between the moment that the researcher observes that an activity is performed by the employee and the moment that the researcher looks at the time that the activity is performed. The screen recording technique is able to accurately record the (online) activities.

We then looked at the differences where the observation technique recorded the activity *2 and 3*

minutes later than the screen recording technique. Based on the Field Notes, we noted that the differences are caused by *overlapping (parallel) activities*. We present two examples of overlapping activities. We first provide the purpose of the activity that occurred prior to the activity where we discovered the difference. We then present the field note to clarify the context of the activity.

Example 1: At 13:06, the employee (participant 1) opened an email via Microsoft Outlook. Their squad lead had sent them an email about the validation of client 1F. The records from the screen recording technique indicate that the employee opened the medium Alteryx on their second screen at 13:09. Based on the screen recording field notes, we noted that the employee alternated between the email and Alteryx. The observation technique recorded this activity at 13:11.

Screen Recording Field Notes: “The employee opened Alteryx. They are creating an Alteryx workflow for client 1F. They are briefly looking at the email regarding the needed data transformation and validation steps on the second screen.”

Example 2: At 16:05, the employee (participant 1) initiated a telephone call via Microsoft Teams. They joined the weekly call with the entire Helix Team. Based on the observation technique and screen recording records, we noted that the employee continued the call until 16:30. However, the records from the screen recording technique indicate that the employee opened the medium Alteryx on their second screen at 16:07. The observation technique recorded this activity at 16:10.

Screen Recording Field Notes: “The employee is working in Alteryx on screen two for client 1G. The employee is still on the call but they are not actively participating anymore. They are only working in the workflow.”

A disadvantage of the timesheet technique is that the records do not contain a timestamp of the moment that the activity was recorded. The participants combine the duration of the activities that belong together. For example, based on the records of the screen recording we noted that participant 6 performed 9 activities related to client 6B. These records have been individually recorded and contain a timestamp. The timesheet technique recorded only 1 activity for client 6B and combined the duration of the related activities.

Recap

The timesheet technique has a different standardised form compared to the observation technique and screen recording technique. The records of the timesheet technique do not contain a timestamp. This restricts the ability to conduct additional analyses (e.g., the performance of the resources). The disadvantage of the timesheet technique is that it aggregates and combines the duration of activities that belong together.

An advantage of the observation and screen recording technique is that the records contain a timestamp of the moment in time that the activity was recorded. A timestamp is a minimum requirement of the event log. The timestamp can be used to calculate the duration of the activity, discover bottlenecks, or analyse the performance of resources (e.g., the employees). However, there is a discrepancy in the ability of the techniques to accurately record the point in time (“timestamp”) that the activity was performed. The records of the observation technique are less accurate compared to the records of the screen recording technique. This is caused by the fact that the observation technique consists of *subjective information*. The observation records are based on the researcher’s perspective and the moment that the researcher observes the activity. The latter is accompanied by the *time lag*. This means that there is a delay between the moment that the activity is recorded, and the moment that the time is recorded. In addition, the observation technique produces less accurate records compared to the screen recording technique once the employee performs *overlapping (parallel) activities*. The (calculation of the) duration of the activities and the quality of the analyses are affected by the discrepancy in the ability of the observation techniques to accurately record the timestamp caused by the overlapping (parallel) activities and the time lag.

5.2 Online versus Offline Activities

We initially compared the records in the standardised form of the observation technique and screen recording technique. We found 11 activities that were only recorded with the observation technique. These were *offline activities* that fell under the category *Meeting, Personal, and Tours*. We then compared the records in the standardised form of the observation technique, screen recording technique, and timesheet technique. The aim was to examine to what extent the techniques are capable of recording online and offline activities.

5.2.1 Meeting

We noted that the offline category Meeting can be *completely* recorded with the observation technique and *partially* recorded with the timesheet technique. The screen recording technique is *unable* to (partially or completely) record the offline activities related to the category Meeting.

The observation technique recorded 3 meetings with a total duration of 55 minutes. For all 6 participants over the course of 6 working days, the observation technique recorded 43:49 hours. This means that the 6 participants spent 2,09% (55 minutes of 43:49 hours) on meetings. We must emphasise that although this percentage (2,09%) may seem insignificant, the meeting *was* a substantial part of the daily activities. The case study was conducted while (some) COVID-19 restrictions were still in effect. The majority of EY employees still work hybrid and have not yet fully returned to the office. Due to COVID-19, the employees switched from *offline meetings* to *online “telephoning”* via Microsoft Teams.

Table 5.2 compares the duration and count of the meeting and telephoning. We had already established that the meeting was recorded 3 times by the observation technique with a total duration of 55 minutes. This is 2,09% of the total duration of 43:49 hours recorded for all 6 participants over the course of 6 working days. During the same period of time, the observation technique recorded telephoning 32 times with a total duration of 09:50 hours. This means that the 6 participants spent 22,44% of their time telephoning. Based on this calculation, we can conclude that once the employees return to their five-day workweeks in the office (or client site), the meeting will become a larger component of the daily activities. If the participants would opt for meetings instead of telephoning, the total time spent on meetings would be 10:45 hours. This means that the 6 participants would spend 24,53% of their time on meetings.

This indicates the significance of this category, and the impact it will have if this category is not (properly) recorded by the techniques.

Table 5.2: Potential significance of the meetings

	Meeting	Telephoning	Combined total
Duration Recorded (in hours)	00:55	09:50	10:45
Count Recorded	3	32	35
Total (%)	2,09%	22,44%	24,53%
<i>Note:</i> We used the 43:49 hours recorded by the observation technique for the calculations, taken from Table F.1 (Appendix F).			

As shown in Table 5.3, the screen recording technique recorded 0 minutes of activities related to meetings. This means that the screen recording technique failed to record 100% of the activities related to meetings. We conclude that the screen recording technique is not able of recording offline activities related to meetings.

Table 5.3: The percentage of the offline activity meetings that was not recorded per data collection technique

	Observation	Screen Recording	Timesheet
Duration Recorded (in hours)	00:55	00:00	00:30
Count Recorded	3	0	1
Total Not Recorded (%)	0,00%	100,00%	45,45%
<i>Note:</i> We used the 55 minutes recorded by observation as the total amount of time that should have been recorded by the remaining techniques.			

The timesheet technique recorded 1 meeting with a total duration of 30 minutes. For all 6 participants over the course of 6 working days, the observation technique recorded 55 minutes of meetings. This indicates that out of a total of 55 minutes, the timesheet technique failed to record 25 minutes (45,45%). The timesheet technique only allows for a *partial recording*. The current design and set-up of the timesheet in Mercury do not take the category Meeting into consideration. As illustrated in Figure 5.1, there is a limited number of available Activity IDs that can be used in the timesheet. The Activity IDs cover the main tasks of the Helix Team employee e.g., “Plan & Request”, “Extract & Transfer”, “(Review of) Memo”, and “Transform & Validate”. The timesheet does not have

an Activity ID related to the meeting. The employees are unable to categorise their activities as meetings in their timesheets.

Second, the timesheet technique is a self-reporting technique. Certain activities may not be recorded if the participants do not find them important. In addition, there is a lack of information that can be extracted from the timesheet. This is due to the limited information provided by the participants in the description of the timesheet. We used the description as the Field Notes in the standardised form of the timesheet technique. Only 1 participant (“participant 1”) mentioned “meeting” in the description of their timesheet. Based on the records of the observation technique, it is evident that participant 4 should have mentioned “meeting” in two activities of their timesheet.

Figure 5.1: The Activity IDs in the Mercury timesheet

5.2.2 Personal

The activities that fall under the category of Personal are *the use of a personal mobile phone, socialising (not work-related), and the use of the internet browser (not work-related)*. We noted a discrepancy in the ability of the techniques to record the activities related to the category Personal, as shown in Table 5.4. This difference is caused by the ability of the techniques to record online and/or offline activities, and by the employee *choosing* not to record this in their timesheet.

The advantage of the observation technique is its ability to record both online and offline activities. This means that the category Personal (offline and online) can be completely recorded with the observation technique. The screen recording technique failed to record 3 activities (15/18). The advantage of this technique is that it provides an accurate recording of online activities. However, the disadvantage is that offline activities such as “the use of a personal mobile phone” or “socialising” cannot be recorded.

Table 5.4: An overview of the Personal (offline) activities recorded by each data collection technique

	Observation	Screen Recording	Timesheet
Duration Recorded (in hours)	01:24	01:44	00:00
Count Recorded	18	15	0

The timesheet technique cannot (partially or completely) record the offline and online activities related to the category Personal. The timesheet is a self-reporting technique. This means that the

employees can omit records that reflect negatively on their behaviour, or only record the activities that they deem to be important. We looked at the observation records in Table F.5 (Appendix F) and noted that participant 3 spent 58 minutes on activities related to using the internet browser, their personal mobile phone, and socialising. However, based on the timesheet records, we conclude that participant 3 *chose* not to include this in their timesheet. The self-reporting characteristic allows employees to omit certain activities which is a disadvantage.

5.2.3 Tours

The activities that fall under the category of Tours are *set-up time and breaks (e.g., restroom, coffee, and lunch breaks)*. The category Tours consists of offline activities.

An advantage of the observation technique is that it can record all activities related to the category Tours. For example, based on the observation records and field notes, we noted that participant 5 started their working day with a fire drill at 10.25. They stood outside of the Amsterdam office for 24 minutes. This activity was recorded by the observation technique but not by the screen recording technique or the timesheet technique.

However, the sound recording (microphone and system audio) allowed the screen recording to *partially* record the activities related to Tours. For example, based on the sound recording, the technique was able to record that participant 1 went on a (smoke) break. This is also indicated in the field notes of the screen recording technique *“The employee took a break to smoke. Based on the recording we would not have been able to categorise this because we only saw the locked screen. However, based on the audio we heard the team member (Helix) ask the employee to ‘go for a smoke’”*. As shown in Table 5.5, for all 6 participants over the course of 6 working days, the observation technique recorded 08:23 hours of activities related to Tours. Not a single minute was recorded by the timesheet technique. The timesheet technique cannot (partially or completely) record the activities related to the category Tours. We established that the timesheet is a self-reporting technique. This means that the employees can omit records that reflect negatively on their behaviour. They can *choose* to only record the activities that they deem to be important. As a result, the records of the timesheet technique provide a partial picture of employee behaviour.

Table 5.5: An overview of the Tours (offline) activities recorded by each data collection technique

	Observation	Screen Recording	Timesheet
Duration Recorded (in hours)	08:23	07:42	00:00
Count Recorded	26	21	0

Recap

The observation technique is able to record online and offline activities. This technique can record the activities that occur on the screen, in the hallway, or even outside of the building. This is an advantage because offline activities such as meetings can be a large component of the daily activities of the employee. Therefore, this technique is able to produce a complete picture of (online and offline) employee behaviour.

However, there is a discrepancy in the ability of the screen recording technique and the timesheet technique to record online and offline activities. The ability of the screen recording technique to (partially) record these activities is caused by whether the participant recorded their screen(s) and audio (microphone and system audio). The screen recording technique was not able to record certain activities because there was no sound or the activities did not occur on the screen (e.g., meetings, socialising, and/or use of personal mobile phones). The screen recording technique was able to partially record the offline activities related to Tours (e.g., smoke breaks) because sound recording documented that the participant verbally announced the reason for their break. The online personal activities such as the use of the internet browser (not work-related) can be completely recorded. The screen recording technique produced a partial recording of the offline activities. As a result, the records contain a partial picture of the behaviour of the employees. It is not possible for the technique to completely record offline activities. However, recording the screen(s) and audio (microphone and system audio) can aid in recording certain offline activities.

Compared to the screen recording technique, the timesheet technique was able to record part of the meetings. However, the timesheet technique completely failed to record the activities related

to the category Personal and Tours. These discrepancies are caused by three factors. First, the timesheet technique is a self-reporting technique. This allows employees to omit records that reflect negatively on their behaviour (e.g., the use of a personal mobile phone) or the records that they do not deem important [94]. This is a disadvantage compared to the observation and screen recording technique. The records of these techniques are more objective because they are not the result of a self-report. Second, the current design of the timesheet has a limited number of available Activity IDs that can be used to record the activities in the timesheet. As a result, the employees cannot categorise activities related to meetings, tours, or even personal. And lastly, the employees do not provide sufficient information in the descriptions of their timesheets.

5.3 Short-lived Activities

We initially compared the records in the standardised form of the observation technique and screen recording technique. We found 10 activities that were only recorded with the screen recording technique. These were *short-lived activities* with a duration of 1 or 2 minutes. The screen recording technique was able to record the short-lived activities related to the use of the internet browser to resolve technical issues, or incoming and outgoing emails, (chat) messages, and telephone calls. However, the observation technique failed to record these short-lived activities. The observation records are based on the researcher’s observations. The disadvantage is that certain short-lived activities are not recorded. This can be the result of the researcher missing the activity, or still categorising and/or creating field notes of the previous activity. The screen recording technique produces permanent records of the activities (the screen recording and sound) that can be viewed repeatedly [118]. This aids in the recording of the short-lived activities. The “shortest” activity recorded by the timesheet technique had a duration of 10 minutes. This shows the inability of the timesheet technique to record short-lived activities (with a duration of 1 or 2 minutes).

Based on the records of the screen recording technique, we noted that there were 133 activities with a duration of 1 minute and 55 activities with a duration of 2 minutes. In total, 47,36% of the activities had a duration of 1 or 2 minutes. This means that it is vital for a data collection technique to be able to record short-lived activities.

We noted that a short-lived activity (Sub-Activity “Incoming Message”, see Table 4.3) was related to client “5M”. This client was recorded by the screen recording technique. However, the observation technique failed to record this client. We then compared the records in the standardised form of the observation technique, screen recording technique, and timesheet technique. The aim was to examine to what extent the techniques are capable of recording short-lived activities, and what the impact is once the short-lived activities cannot be recorded. We used Table F.14 (Appendix F) for the comparison of the clients per data collection technique. After cross-checking the clients recorded for each participant and per technique, it became evident that each technique should have recorded 45 clients. Table 5.6, presents the total number of clients that the data collection techniques failed to record.

Table 5.6: Clients failed to record per data collection technique

	Observation	Screen Recording	Timesheet
Total Recorded (Count)	43	44	24
Total Not Recorded (Count)	2	1	21
Total Not Recorded (%)	4,44%	2,22%	46,67%
<i>Note:</i> The calculation is based on 45 being the total number of recorded Clients. This total is calculated based on the 44 recorded and 1 not recorded Clients (screen recording).			

The screen recording technique recorded 44 clients but failed to record 1 client (“1C”). The observation technique also failed to record this client. The only technique that recorded the activities related to this client, is the timesheet technique. This means that this was an overtime activity (occurred either before 09:00 or after 17.00). The observation technique recorded 43 clients and failed to record 2 clients. In addition to the overtime activity related to client 1C, it also failed to record the activity related to client 5M. This was a short-lived activity with a duration of 1 minute.

The timesheet technique recorded 24 clients. This means that the technique failed to record 21 clients, which is 46,67% of the total (of 45 clients) recorded for all 6 participants over the course of 6 working days. According to the records of the screen recording technique, there were 60 activities

related to the 21 clients that the timesheet technique failed to record. Table 5.7, shows the client-related activities that were not recorded by the timesheet technique. We distributed the activities based on the following time ranges: *more than 10 minutes*, *5 to 10 minutes*, *3 or 4 minutes*, and *1 or 2 minutes*.

Table 5.7: Client-related activities not recorded by the timesheet technique

	More than 10 minutes	5 to 10 minutes	3 or 4 minutes	1 or 2 minutes
Activities (Count)	2	7	9	42
Activities (%)	3,33%	11,67%	15,00%	70,00%
<i>Note:</i> The calculation is based on 60 being the total number of client-related activities.				

Out of the 60 activities, only 2 activities (3,33%) had a duration of more than 10 minutes. The activity related to client 3A had a duration of 17 minutes and was related to the sub-activity data transformation & validation. The activity related to client 1A with a duration of 29 minutes, was related to an incoming call. There were 7 activities (11,67%) that had a duration between 5 and 10 minutes. These activities were related to incoming and outgoing calls, and the review of the extraction memo (571). There were 9 activities (15,00%) that had a duration of 3 or 4 minutes. These activities were related to planning, outgoing emails and messages, and one 3-minute activity related to data transformation & validation. However, the majority of the activities that the timesheet technique failed to record were *short-lived activities*. In total, 42 activities (70%) had a duration of 1 or 2 minutes. The majority of the activities were related to incoming and outgoing messages or emails.

In Table 5.8, we present the type of sub-activities that the timesheet technique failed to record. In total, 37,22% of the activities that were not recorded by the timesheet technique were related to incoming and outgoing messages or emails. These were mostly *short-lived activities* with a duration of 1 or 2 minutes. One of the characteristics of the timesheet is that it is a self-reporting technique. This means that the participants must record their own activities. The disadvantage is that the employees have the choice not to record certain activities because they do not find them important, or they do not want to disturb their process to record each short-lived activity with a duration of 1 or 2 minutes.

Table 5.8: Not recorded client-related (sub-)activities (hours and %)

Sub-Activity	Duration	Percentage
Incoming & Outgoing Messages and Emails	01:07	37,22%
Incoming & Outgoing Call	00:54	30,00%
Data Transformation & Validation	00:40	22,22%
Planning	00:10	5,56%
(Review) Extraction Memo (571)	00:09	5,00%
<i>Note:</i> The calculation is based on 03:00 hours being the total duration of client-related activities recorded with screen recording.		

We established that the short-lived activities are not recorded by the timesheet technique. However, based on Table F.14 (Appendix F), it seems that most employees compensate for this by rounding up to the nearest (two) quarter(s) of an hour.

Recap

There is a discrepancy in the ability of the techniques to record short-lived activities with a duration of 1 or 2 minutes. The screen recording technique is able to record all online short-lived activities. The technique produces permanent records of the activities (the screen recording and sound) that can be viewed repeatedly [118].

The observation technique only produced a partial recording because it cannot record all short-lived activities. The technique failed to record 10 short-lived activities with a duration of 1 or 2 minutes such as the use of the internet browser to resolve technical issues, or incoming and outgoing emails, (chat) messages, and telephone calls. This technique does not produce a “permanent record of the activities” similar to the screen recording technique. The disadvantage of the observation technique is that the records are based on the researcher’s observations. The researcher may have missed

the activity (e.g., still in the process of categorising and/or creating field notes of the previous activity). The failure of the observation technique to record the short-lived activities had an impact on the duration and count (the number of times that the activity was recorded) of activities, and the completeness of the employee behaviour. We provided an example of the client (5M) that the observation technique failed to record.

The timesheet technique failed to record the majority of the short-lived activities. This can be caused by the self-reporting characteristic of the technique. The drawback is that employees can opt not to record certain activities if they do not deem them to be important (e.g., a chat message received via Microsoft Teams from the audit team who inquires about the planning from a client). Moreover, the employees do not want to constantly disturb their process to record each short-lived activity with a duration of 1 or 2 minutes. The impact is that this technique only produces a partial picture of the behaviour of the employee.

5.4 Overtime Activities

As presented in Table 5.9, there is a discrepancy in the total time recorded by the data collection techniques during the course of the 6 working days of the 6 participants. The observation technique recorded 43:49 hours, the screen recording technique recorded 43:25 hours, and the timesheet technique recorded 48:50 hours.

Table 5.9: Total time recorded by each data collection technique

	Observation	Screen Recording	Timesheet
Total Time Recorded	43:49	43:25	48:50
<i>Note:</i> The total time recorded is shown in hours.			

There is a 24-minute difference between the observation technique and the screen recording technique. As described in Section 5.2.3, this difference is caused by 1 offline activity (Tours) that was only recorded by the observation technique, and not by the screen recording or timesheet technique. The observation technique is a *direct technique*. The technique requires direct contact between the researcher and the participant. A disadvantage of the observation technique is that it cannot record any activities performed outside of the time frame of the study because the physical presence of the researcher is required. The screen recording technique did not record any activities performed before 09.00 or after 17.00. However, the screen recording technique would be able to record *online overtime activities*. The screen recording technique is an indirect technique and does not require any interaction between the researcher and the participant. The technique only requires a functioning screen recording application. The downside is that the technique may produce a partial picture of employee behaviour because it can only record *online* overtime activities. The technique cannot record offline activities such as meetings, socialising, or the use of a personal mobile phone.

There is a difference of more than 5 hours between the observation and screen recording technique, and the timesheet technique. We categorised this difference as *overtime activities* because only the timesheet technique was able to record these activities. This means that the employee performed the activities outside the study’s time frame (before 09.00 or after 17.00). This difference is caused by 5 participants (P2 - P6), as shown in Table F.15 (Appendix F). Participants P2 - P5 had a difference ranging from a few minutes to 01:32 hours. The differences are related to the Activities “Organisational Work”, and “Scheduling & Administration”. Participant P6 had a difference of 02:30 hours caused by the activities “Technical Problem Solving” (01:30 hours), and “Organisational Work”. The advantage of the timesheet technique is its self-reporting characteristic. This allows employees to record their own activities performed before, during, and after working hours.

Recap

There is a discrepancy in the ability of the techniques to record overtime activities. These are activities that occurred outside of the time frame of the study (before 09:00 or after 17:00). The observation technique failed to record the overtime activities. The observation technique is a direct technique [78], meaning that it requires the presence of the researcher (“observer”). The disadvantage is that activities that occur outside of the time frame cannot be recorded. As a result, the technique cannot produce a complete picture of employee behaviour. The screen recording technique did not record the overtime activities. However, screen recording is an indirect technique. This means that the technique only requires direct access to the working environment (i.e., laptop). The screen

recording technique is more flexible compared to the observation technique in terms of where the *online* behaviour can be recorded [139]. The participants can record all online activities, from the office or their homes, as long as the screen recording application works. However, the disadvantage is that the *offline* overtime activities (e.g., meetings or use of a personal mobile phone) cannot be recorded by the screen recording technique.

The participants could have performed the overtime activities at the office, on the train, or at home. This cannot be deducted from the timesheet records. However, it is certain that the timesheet technique is able to record overtime activities. The timesheet is an independent self-reporting technique. The advantage is that the participants can record their own activities. As a result, this technique produces a record including overtime activities.

5.5 Uncategorized Activities

There is a discrepancy in the ability of the data collection techniques to categorise the activities. Table 5.10, presents the percentage of the activities that the techniques were unable to categorise split per classification. The observation technique was able to categorise all activities. The advantage of the observation technique is that it is able to record both online and offline activities. The first (partial) categorisation takes place while the observation occurs. The categorisation is accompanied by a field note that provides detailed information about the context of the activity. This allows for a second (review) round to complete the categorisation.

We established that the screen recording technique can accurately record the *online activities*. The technique provides highly detailed information about the online setting because the technique produces permanent records of the activities that can be viewed repeatedly [118]. This aids in the categorisation of the activities, and the creation of the field notes. This is an advantage compared to the observation technique. We used the field notes from the screen recording technique to examine what caused the discrepancy. The technique was unable to categorise certain activities due to technical issues with the recording tool Snagit. Participant 3 experienced technical issues during the recording of their working day which caused the tool to produce a partial screen and sound recording. As a result, the technique was unable to categorise 1,01% of the activities. The dependency on a recording tool to provide a complete recording (of the screen and sound) is a disadvantage of the screen recording technique.

Table 5.10: Unable to categorise activities per Classification per data collection technique

Classification	Observation	Screen Recording	Timesheet
Category	0,00%	1,01%	0,00%
Activity	0,00%	1,01%	0,00%
Sub-Activity	0,00%	1,01%	7,84%
Medium	0,00%	1,01%	68,63%
Participant	0,00%	1,01%	60,78%
Initiated	0,00%	1,01%	74,51%
Client	0,00%	1,01%	0,00%
<i>Note:</i> The results are aggregated from Tables F.1, F.3, F.5, F.7, F.9, F.11, F.14 (Appendix F).			

Compared to the observation and screen recording technique, the percentage of the activities that the timesheet technique failed to categorise is higher. This is due to a lack of information that can be extracted from the timesheet, caused by the design of the timesheet and the limited information provided in the description. The current design and set-up of the timesheet in Mercury do not take the majority of the classifications into consideration. Based on Figure 5.1 (Section 5.2.1), we had established that there is a limited number of available Activity IDs that can be used in the timesheet. In addition to the Activity ID, the employee can only fill in the Engagement ID (e.g., billable hours charged to a client or non-billable hours such as leave or holiday), the hours worked, and a description. Figure C.2 (Appendix C) illustrates an example of the timesheet in Mercury. This reduces the employee’s ability to self-report their activities [94].

As a result, the timesheet technique failed to categorise the activities based on the information provided in the Activity ID or the Engagement ID. This highlights the need for a detailed description because this is used for the field notes in the standardised form. However, we noted a discrepancy

in the richness of the information provided in the field notes per data collection technique. In Table 5.11, we present the minimum, maximum, and the average number of words recorded in the Field Notes per technique.

Table 5.11: Richness information in the Field Notes per data collection technique

	Observation	Screen Recording	Timesheet
Minimum	5	5	1
Maximum	79	164	21
Average	20,30	32,58	4,63

The observation field notes had a minimum of 5 words, a maximum of 79 words, and an average of 20,30 words. The field notes provided detailed information to categorise all activities. An advantage of this technique is that it provides *detailed* field notes of the online setting *and* the offline setting. An example of a field note of the offline setting “*The employee is socialising with a Helix colleague. They are talking about the induction that they had for the past two weeks*”. However, the observation field notes are created while the observation takes place. The researcher (“observer”) divides their attention between observing the employee, categorising the activity in the standardised form, and writing their observation in the field notes. As a result, the field notes of the observation technique are not as detailed as the field notes of the screen recording technique. The field notes of the screen recording technique had a minimum of 5 words, a maximum of 164 words, and an average of 32,58 words. The screen recording provides *highly detailed* field notes. An advantage of the technique is that the recordings can be viewed repeatedly to create a highly descriptive field note of the *online* activities.

The observation technique and the screen recording technique provide rich information in the field notes. However, the difference is that the observation technique provides *detailed* information, while the screen recording technique provides *highly detailed*, as shown in Table 5.12.

Table 5.12: The richness of the field notes (min/max) per data collection technique

<i>Observation</i>	
Maximum	The employee received an incoming call on Microsoft Teams from their Squad Lead. They are speaking about roll-forward differences found for client 1A. The employee is sharing their screen to present an Excel sheet consisting of the found differences. The employee is also briefly showing the Alteryx workflow that they used to transform and validate the data. They are discussing the Alteryx workflow and making adjustments where needed. The employee is alternating between Microsoft Excel and the Alteryx workflow.
Minimum	The employee is uninstalling Python.
<i>Screen Recording</i>	
Maximum	The employee opened Microsoft Outlook because they received an email from the client with the squad lead and two audit team members in the CC. The email is about the data extraction of client 1B. The client states that the extraction will be similar to the previous year, with the exception that another business unit will need to be extracted. Moreover, the client stated that the extraction of October 5th will be too early because the books are not closed yet. The employee scrolled back to their initial email and highlighted the business units that were in the scope. The business unit as suggested by the client was not included in the email. The employee then briefly went to JennEY, they searched for the client’s name and clicked on the ticket. The scope of the ticket also did not include this business unit. Based on the audio it seems that the employee spoke to a nearby colleague about this email and what to do next.
Minimum	The employee is installing Python.
<i>Timesheet</i>	
Maximum	228669: Special request from the audit team and SAS to see if we can additional data in the GLA, Client 6C ¹ .
Minimum	Learning

¹We copied this field note from the timesheet description without making any changes (i.e., grammar).

The timesheet technique does not provide rich or detailed field notes. The field notes had a minimum of 1 word, a maximum of 21 words, and an average of 4,63 words. More than 50% of the timesheet field notes consist of 3 words (14/51 records) or 4 words (17/51 records). Some field notes consist of one or two words such as “*Learning*”, “*Meeting*”, or “*Squad Call*”.

We noted that there is only a universal template for client-related activities. These activities all start with a 6-digit ticket number e.g., “000000²: *client 5B GLA*”. For the remaining activities, there is no template that the employees must use for their descriptions. As a result, there is a difference in how the activities are described. We also noted that certain participants combine multiple activities into 1 single record. For example, participant 1 created a record with a duration of 1 hour and described it as “*Helix midweek checkup / Squad meet central (casual session)*”. The “*Helix midweek checkup*” and the “*Squad meet central (casual session)*” are two different activities, and therefore should have their own record in the timesheet. Based on this timesheet, the duration of each activity cannot be deducted.

In addition, we noted a discrepancy in the richness of the descriptions per participant, as shown in Table 5.13. This difference is found between the participant who did not record their activities while or right after completing the activities and the participant who did record their activities while or right after completing the activities. Participants P1 - P3 did not record their activities while or right after completing the activities. This means that the timesheet records and descriptions were created in retrospect, relying on memory recall. Participants P 4 - P6 recorded their activities while or right after completing the activities using. The participants used additional mediums (e.g., Microsoft Excel, Notepad or Microsoft Outlook) to record their activities.

Table 5.13: Richness information in the Field Notes per Research Participant in the Timesheet

Participant	Field Notes (Average)
P1	4,38
P2	3,25
P3	2,00
P4	3,53
P5	4,47
P6	10,33
<i>Note: The results are summarised from Table F.13 (Appendix F).</i>	

Recap

There is a discrepancy in the ability of the techniques to categorise the activities. The observation technique was able to categorise all activities. The advantage of this technique is that categorisation occurs twice. The first (partial) categorisation takes place while the observation takes place. The categorisation is accompanied by a field note that provides information about the context of the activity. This information is used for the second round of categorisation. The advantage of the screen recording technique is that it produces permanent records of the activities (the screen and sound recording) that can be viewed repeatedly [118]. This aids in the categorisation of the activities. However, the disadvantage of the technique is that it depends on a recording tool to provide a complete recording of the screen (including the sound). This technique does not produce field notes while the recording takes place, compared to the observation technique. As a result, it is not possible to categorise the activities if the recording is incomplete due to technical issues. The timesheet technique failed to categorise most activities. The timesheet has a limited number of Activity IDs that can be used to categorise the activities. Therefore, the design of the timesheet does not allow the participant to self-report and categorise their activities. This emphasised the need for a detailed field note. However, the quality of the field note is subpar. Minimal information can be extracted from the field notes because more than 50% of the field notes consisted of 3 or 4 words. In addition, the lack of a universal template used to describe certain activities complicates the analysis. As a result, the timesheet technique only produces a partial, and *very* limited picture of the behaviour of the employees.

In the next Chapter, we present the implications for research and for practice that we created based on the results.

²A 6-digit ticket number is used to link the activity to a client.

Chapter 6

Implications

This Chapter presents the implications for research and for practice. We structured the implications based on the order of the results (Chapter 5). We used the findings from multiple sources to create the implications. First, we looked at the Analysis (Chapter 4) and the Results of the multiple-case study (Chapter 5) to identify the differences and similarities between the techniques *non-participant observation, screen recording, and the archival analysis (timesheet)*. Second, we used the findings from the Literature Review (Chapter 2) and the characteristics of the techniques (Table 2.2) to conclude to what extent our study has confirmed or refuted these findings.

6.1 Timestamp

Implications for Research

We have shown that a weakness of the observation technique and timesheet technique has to do with the ability to (accurately) record the timestamps. The timesheet technique is unable to record timestamps because the records consist of aggregated activities and the total time spent on these activities. The literature acknowledged that the timesheet technique does not provide detailed information [62, 78]. However, the impact of the limited data analysis and monitoring potential was not acknowledged. Our study confirmed that the observation technique records are based on the researcher’s observation, their observation, and what they deem to be important [73, 118]. However, the literature failed to mention that the observation technique can produce inaccurate timestamps due to the following two factors identified during our study. First, we noted a *time lag* between the moment that the researcher observes that an employee has performed an activity and the moment that the researcher looks at the time that the activity is performed. As a result, the timestamps in the observation records lag behind as opposed to the screen recording records. Second, we noted that the observation records contained *overlapping (parallel) activities*. Activities can be completed faster because there are fewer restrictions on the execution order (i.e., parallel rather than strictly sequential) [43]. However, the overlapping activities do not allow us to easily distinguish when an activity starts and ends. The failure of the observation and timesheet technique to (accurately) record the timestamp has an impact on the ability to use the records for analysis and monitoring purposes. Moreover, we noted that process mining aims to discover, monitor, and improve business processes by extracting knowledge from event logs [3]. An event log should at least contain the following three attributes: (1) an identification for each case (“Case ID”), (2) the activities that took place in each case (“Activity”), and (3) a reference to a point in time i.e., the start and end of each activity within a case (“Timestamp”) [84]. The lack of (accurate) timestamps limits the possibility of using additional techniques such as process mining.

Implications for Practice

General Implications for Organisations

The inability of the observation and timesheet techniques to (accurately) record the timestamps inhibits the ability of the management of an organisation to use the records for monitoring and analysis. First, timestamps can be used for legal purposes, such as providing a (digital) signature on a document. Second, data management relies on timestamps to ensure the quality of the data. Third, the timestamp is required to be able to calculate the duration of an activity, discover bottlenecks, or analyse and monitor the performance of the resources (e.g., employees) [1, 3]. And lastly, an accurate and complete record of the timestamps is required to use process mining. The inability of

the observation technique and timesheet technique to (accurately) record the timestamps means that the management of an organisation must rely on the screen recording technique or even alternative techniques (e.g., extraction of logs) to supplement the records.

Implications for EY

The management of EY already uses the records for analyses and monitoring purposes. First, the management already uses Power BI dashboards to monitor the KPIs (e.g., the productivity of employees). An (accurate) timestamp is one of the requirements to be able to create these dashboards. Second, a timestamp (consisting of a date and time) is used to track the status of a ticket. Lastly, the management has expressed their want to use techniques such as process mining to reconstruct the behaviour of employees [113].

6.2 Online versus Offline Activities

Implications for Research

We have shown that a weakness of the screen recording technique and timesheet technique has to do with *online versus offline activities* that cannot be (completely) recorded. This study confirmed that the screen recording technique provides *highly detailed* records of the online setting [72]. This means that all activities that occur on the screen(s) can be recorded, regardless of the participant's location (e.g., at the office or at home). Offline activities such as meetings, tours around the office, or the use of a personal mobile phone cannot be recorded. However, the importance of recording the screen(s) *and* audio (microphone and system audio) was not discussed in the literature. We discovered that the audio allowed the screen recording technique to record certain offline activities.

The timesheet technique consists of objective information because the participant records their behaviour [159]. However, our study confirmed that the self-reporting characteristic of the timesheet technique has drawbacks [86, 127]. The self-reporting technique allows the employees to omit records that reflect negatively on their behaviour (e.g., socialising or the use of their personal mobile phones), or only record the activities they deem to be important [94]. Furthermore, our study confirmed that the quality of the records is limited because the purpose of the timesheet technique is to offer data for the organisation rather than for research [114, 159]. We noted that the quality and completeness of the records depend on the design of the timesheet. The employees digitally record information about the performed work, behaviour, and activities [29] using timesheets via in-house enterprise systems [127]. Our study confirmed that organisations have unique Engagement IDs and Activity IDs [13], and guidelines regarding how to fill in the descriptions of the timesheet [118]. This can limit the ability to use the records for analysis and monitoring purposes, and make it challenging to compare different organisations, or even different teams within the same organisation [118]. The timesheet technique is implied to be imprecise and inaccurate [140]. However, it can provide a fair view of the total workload and its distribution from an employee's perspective, as long as additional evidence from sources other than employee self-reports is accumulated [160].

Implications for Practice

General Implications for Organisations

The observation technique can record online and offline activities. However, the observation technique involves intensive interaction between the researcher and participant(s) [108]. This is a costly technique because it requires plenty of time and effort. Therefore, it is not practical for an organisation to completely rely on this technique. The screen recording technique and the timesheet technique are unable to (completely) record the online and offline activities. This means that organisations will only have a partial recording of the behaviour of their employees. The online activities can be recorded with the screen recording technique. However, for the screen recording technique to record a small portion of the offline activities, recording the screen(s) and audio (microphone and system audio) is necessary. Activities such as meetings (offline) cannot be recorded by the screen recording technique. As a result, to assess the impact, an organisation's management must examine the distinction between online and offline activities. The timesheet technique can (partially) record the online and offline activities. The accuracy and completeness of the records vary per organisation because it depends on the design of the timesheet, and the guidelines and rules set forth by the organisation.

Implications for EY

At the moment, the majority of the activities occur on the screen (online). Due to COVID-19, the employees switched from offline activities (e.g., meetings) to online "telephoning" via Microsoft

Teams. Offline activities are expected to become a larger component of daily activities once employees return to their five-day workweeks in the workplace (or client site). This means that a large portion of the employee behaviour will not be recorded if EY only relies on the screen recording to record the offline activities. Furthermore, the employees must use a timesheet in Mercury (an in-house enterprise system). These timesheets provide limited options for the Activity ID. Therefore, the employees are restricted from properly recording their activities. In addition, the purpose of the timesheet is to record work-related activities. This technique is not designed to record (offline) activities related to the category Tours and Personal. As a result, the timesheet technique will provide a partial picture of the behaviour of the employees.

6.3 Short-lived Activities

Implications for Research

Our study confirmed the findings of previous studies (see e.g., [16, 66, 95, 140]) that the behaviour of employees is characterised by many short-lived activities, with the aim to solve short-term problems, at an unrelenting pace with frequent interruptions. The literature failed to mention that a weakness of the observation technique and timesheet technique has to do with short-lived activities that cannot be recorded. Our study did confirm that the perspective of the observation technique consists of *subjective* information [73]. The observation records are based on the researcher's observation. This is influenced by the perspective of the researcher, and what they deem to be important [118]. The disadvantage is that a large portion of short-lived activities is not recorded. This can be the result of the researcher missing the activity, or still categorising and/or creating field notes of the previous activity. In addition, we noted the time interval used for the standardised form differs per study (e.g., due to context, setting, the number of researchers, and sector). Arman et al. [16] had 4 researchers who used a 30-second time interval while Mintzberg [95] did not record the exact number of minutes for activities with a time interval of 3 minutes or less during his study [66]. However, the impact of the time interval used in the standardised form, and the implications of the number of researchers who are tasked to observe the employees, were not discussed in the literature review. The observation of (short-lived) activities can be easier if multiple researchers (i.e., two or more) are involved.

A factor that was not mentioned by the literature is that the timesheet technique is not suitable for recording short-lived activities. Our belief is that this is caused by the following two characteristics. First, the timesheet consists of the aggregated activities and the total time spent on these activities. For example, the "shortest" activity recorded in the timesheet had a duration of 10 minutes. Second, the timesheet is a self-reporting technique. This means that the participants must record their own activities. The disadvantage is that the employees do not want to constantly disturb their process to record each short-lived activity with a duration of 1 or 2 minutes. The impact is that this technique only produces a partial picture of the employee's behaviour.

Implications for Practice

General Implications for Organisations

The observation technique and the timesheet technique are unable to record the short-lived activities. This has an impact on the organisation(s) and their employee(s) because the techniques provide a partial picture of the behaviour of the employees (e.g., the billable hours cannot be charged or the utilisation rate (%) will be lower). The inability of the observation technique and timesheet technique to record the short-lived activities means that the management of an organisation must rely on the screen recording technique or even alternative techniques (e.g., extraction of logs) to supplement the records.

Implications for EY

Almost 50% of the activities that occur throughout the working day are short-lived activities with a duration of 1 or 2 minutes. This means that numerous short-lived activities that occur during an employee's working day must be recorded. The timesheet technique has three additional objectives that require a complete record. First, the timesheet records are used to charge the billable hours to a client. The failure of this technique to record all short-lived activities has an impact on the profit. Second, the timesheet is used to measure the performance of the employees. These insights are used during the annual rating of the employees. The fact that the employee behaviour is not correctly registered in the timesheet means that the employees are not rated in a fair and equal manner. And lastly, various accounting scandals have resulted in laws and policies, which have been enacted to ensure the independence of auditors and those who have access to client data. This means that accounting firms, partners, employees, and their immediate or close family members

must be independent while working on a client engagement. The timesheet is used to track the independence of the employees. The failure of the timesheet to record short-lived activities related to a client engagement may result in severe consequences e.g., discredit of the image of EY or hefty fines imposed by the Dutch Authority for the Financial Markets (AFM) [8].

6.4 Overtime Activities

Implications for Research

We have shown that a weakness of the observation technique and screen recording technique has to do with overtime activities that cannot be (completely) recorded. Our study confirmed the literature findings which indicated that the categorisation of techniques (i.e., direct, indirect, and independent techniques) impacts the quality of the records. Lethbridge et al. [78] categorised the data collection techniques based on the required degree of contact between the researcher and participant. The observation technique is a *direct technique*. The researcher must have direct access to the participant which means that the observation technique will not be able to record the overtime activities. The screen recording technique is an *indirect technique*. The researcher must have direct access to the working environment (i.e., device) and a working screen recording application. The literature failed to mention that the screen recording technique can be used to record *online* overtime activities. Within our study, we did not use the screen recording technique to record the online overtime activities. However, based on our findings from recording the online activities (within the time frame of our case study), we believe that the screen recording technique would be able to record the online overtime activities. The timesheet is an independent (self-reporting) technique which means that the participants can record their own (overtime) activities. A previous study even commented that self-reporting techniques are not as precise and accurate [140]. Our study confirmed that the observation technique and the screen recording technique provide accurate records with detailed information. However, its inability to (completely) record the overtime activities means that the timesheet technique must be included otherwise the proportion of overtime activities will decrease or be non-existent [140].

Implications for Practice

General Implications for Organisations

We categorised the activities that occurred outside the study's time frame (before 09.00 or after 17.00) as *overtime activities*. Therefore, the discrepancy in the ability of the techniques to record overtime activities will have an impact on the completeness of the records. Employees have a company device (i.e., laptop or external monitor) at their disposal, that spans both work and home to support a wide range of activities. This increases the flexibility of an employee to perform their activities. The employee can work at the office using their laptop. However, the employee might also choose to work on the train, or at home late at night. These kinds of activities occur outside of the time frame and therefore will not be included in the records if only the observation technique is used. Organisations can use the timesheet technique and/or the screen recording technique to record overtime activities.

Implications for EY

The employees already use the timesheet to record their own activities. This means that the employees can provide a weekly record that consists of the (overtime) activities. However, we have observed that, depending on the type of activity, there is a discrepancy in the accuracy and completeness of the timesheet record for each research participant. This is a result of the self-reporting characteristic of the timesheet technique. The inability of the observation technique and screen recording technique to (partially) record the overtime activities can have an effect because the timesheet records cannot be compared and/or supplemented. As a result, the employees can lose a portion of their overtime activities.

6.5 Uncategorized Activities

Implications for Research

The observation technique can categorise the (online and offline) activities. The literature failed to mention that an advantage of this technique is that classification occurs twice. The first (partial) categorisation, including the creation of the field notes, takes place during the observation. This allows for a second round to review the field notes and to complete the categorisation. We have shown that a weakness of the screen recording technique and timesheet technique has to do with *uncategorised activities*. Data quality issues such as missing data (i.e., different kinds of information

are missing from the records or event log) [31, 84, 137] can result in large errors or inadequate results [42]. The literature does mention that the timesheet technique provides limited information [94]. In fact, Lethbridge et al. [78] found that descriptive fields were often not filled in (adequately) to amass sufficient information. Our study confirmed that the timesheet descriptions do not provide sufficient information to be able to categorise the activities. In addition, our study confirmed that the screen recording technique provides highly detailed information about the online setting [72, 139]. However, only one study (see e.g., [73]) mentioned that using a screen recording application carries the risk that the recording will not be complete. This means that using any kind of technology poses a risk in terms of data loss “*an incident where data is destroyed, deleted, corrupted, or made unreadable by users or software applications*” [37]. Past research (see e.g., [94, 160]) recommended that the limitations of the timesheet technique can be mitigated by accumulating additional evidence from sources other than employee self-reports. Our study confirmed that it is necessary to collect data from multiple sources due to the drawbacks of the timesheet technique [78, 95, 118]. However, the literature failed to mention that the screen recording technique should also be combined with an additional technique to minimise the risk of data loss, categorise the activities, and provide a complete overview of the behaviour of the employees.

Implications for Practice

General Implications for Organisations

The discrepancy in the ability of the techniques to categorise the records will impact the completeness of the records and the ability of an organisation’s management to use these records for further monitoring and analysis purposes. The categorisation of the records serves a variety of purposes for the organisation’s management. First, it provides the context of the activity. The categorised activities provide information about the classifications used in the standardised form e.g., the (purpose of the) activity, mediums, and clients. Second, categorised activities allow for the monitoring and comparing of records (across periods and employees). And lastly, the lack of categorised activities limits the possibility of using additional techniques such as process mining. An event log should contain sufficient information (about each activity) to be characterised as complete, trustworthy, and of excellent quality [5].

Implications for EY

We concluded that the timesheet technique is unable to properly categorise the classifications. This is caused by the following factors. First, the design of the timesheet in Mercury does not take all classifications into consideration. The timesheet only consists of the columns *Engagement ID*, *Activity ID*, and *Description*. In addition, the description does not provide rich information. This is a result of the lack of universal templates that the employees must use to describe their activities. The failure of the timesheet technique to categorise the majority of these classifications has an impact on the ability of the management of EY to use the timesheet records for analyses and monitoring purposes. The EY stakeholder highlighted the significance of three classifications: *Medium*, *Participant*, and *Initiated*. The timesheet technique is unable to record the usage time of the mediums per activity. As a result, the timesheet records cannot be used to detect unwanted or repetitive tasks that can be (partially) automated in the future. In addition, the timesheet technique is unable to record with whom the employees interacted (“participants”) throughout the working day, and who initiated the activity. Therefore, the timesheet records cannot be used to identify the cause, and the involved participants of the disruptions throughout the working day (e.g., who sent the chat messages via Microsoft Teams or emails via Microsoft Outlook). Lastly, due to the risk of data loss, the management cannot solely rely on the screen recording technique.

In the next Chapter, we translate the insights into practice. We present 3 scenarios to illustrate how multiple data collection techniques can be combined to minimise the shortcomings of the techniques and to record employee behaviour based on the context of the situation. We also propose alternative techniques to supplement the records.

Chapter 7

Translation of Insights Into Practice

In this Chapter, we translate the insights gained during the previous Chapters into practical advice. We discuss 3 scenarios to illustrate how multiple data collection techniques can be combined to minimise the shortcomings of the techniques depending on the context of the situation.

Our recommendation is based on the findings from multiple sources. First, we used the Results (Chapter 5), and the Implications for research and practice (Chapter 6) to determine the strengths, weaknesses, differences, and impact on the quality of the records per data collection technique. Second, we looked at which techniques are already being used or are most likely to be used by the organisation. The timesheet technique and the screen recording technique are already being used within an organisation. However, it is not common for the observation technique to be used (for an entire working day) to record employee behaviour. The observation technique involves intensive interaction between the researcher and participant(s) [108], therefore is a costly technique because it requires plenty of time and effort (which also impacts the availability of the employees). For this reason, we are unable to recommend the use of the observation technique to record employee behaviour. And lastly, even after combining the screen recording and timesheet techniques, certain work patterns will not be (completely and/or accurately) recorded. We used the Literature Review findings (Chapter 2), internal documentation, and information gained from discussions with the EY stakeholder, to propose alternative techniques that can be used to supplement the records.

7.1 Tick Tock, What Time Is It

Scenario 1: Record the Timestamps

Organisation TimeCorp wants to start recording the behaviour of their employees. Their goal is to use the records for monitoring and analysis purposes. TimeCorp's main objectives are to monitor the KPIs, and to discover, monitor, and improve their business processes. TimeCorp wants to make its own Power BI dashboards e.g., for financial or legal analysis. In addition, TimeCorp has decided to implement process mining to be able to reconstruct the behaviour of their employees. TimeCorp's only requirement is that the records contain accurate timestamps of each recorded activity.

At the moment, TimeCorp is not using any techniques to record employee behaviour. Therefore, TimeCorp wants to examine which technique is most suitable to record employee behaviour with an accurate timestamp. TimeCorp has a budget, but would prefer to set it aside for the acquisition of a process mining tool (e.g., Celonis or Disco) as well as a thorough training on how to use it. In addition, TimeCorp is an IT organisation that only offers remote jobs. Their employees are digital nomads who work at home, and even around the world. This means that TimeCorp does not have direct access to its employees.

The event logs are the main input for process mining techniques [84]. Therefore, we must be certain that the technique produces records containing the following information: (1) an identification for

each case (“Case ID”), (2) the activities that took place in each case (“Activity”), and (3) a reference to a point in time i.e., the start and end of each activity within a case (“Timestamp”).

We recommend that TimeCorp uses the screen recording technique to record the behaviour of the employees. This technique is able to record all online activities. In addition, the screen recording technique produces records that contain accurate timestamps of each recorded activity. Due to the remote nature of the jobs, we do not need to record offline activities. Therefore, by using the screen recording technique, TimeCorps would be able to record all online activities.

We identified a potential data quality issue (i.e., different kinds of information is missing from the event log) [31, 137] if TimeCorps would solely depend on the screen recording technique. The use of a screen recording application carries the risk that the recording will not be complete. This means that using any kind of technology poses a risk in terms of data loss. There is a plethora of available recording tools that are free to use. As a result, the screen recording technique will have no effect on the budget. However, the remote nature of the job has the following disadvantage. The employees can experience technical issues due to a lack of a stable internet connection and/or (internet) censorship in a certain country [151]. As a result, the recording tool will not produce complete records.

To ensure that the records are complete, trustworthy, and of excellent quality [31, 137], we recommend supplementing the screen recording records with logs extracted from the information systems. The screen recording technique and the extraction of the logs from the information systems are both suitable techniques within the context of this scenario. The implementation of these techniques will not be costly. In addition, by combining these techniques TimeCorps will not be dependent on a recording tool or the regulations of a country.

7.2 Back to the Office

Scenario 2: Record the meetings (offline activities)

Due to COVID-19, the organisation Back2Work was compelled to transition from working in the office and on a client site (offline), to working from home (online). From March 2020 to October 2022, the employees exclusively worked from home. During this period, Back2Work was keen to record the behaviour of their employees to see if working from home had impacted the performance of their employees. Even before COVID-19, the employees had to fill in their timesheets on a weekly basis. However, Back2Work noted that there were some discrepancies in the quality of the timesheet that limited their ability to use the records for analysis. First, the employees provided minimal information in their timesheets. Second, the timesheets consisted of aggregated records instead of separate records for each activity.

For this reason, Back2Work initiated a project in which employees could sign up voluntarily to record their screens. However, Back2Work failed to gain support for this project. The employees mentioned that the screen recording technique was too intrusive because it recorded their screen(s) and the audio (e.g., personal conversations with family members).

From November 2022, Back2Work has noticed a shift to hybrid working. Back2Work observed that their employees meet in the office three to five days per week. This transition means that the employees conduct offline meetings instead of telephoning via Microsoft Teams. Depending on the number of days that the employee is in the office, Back2Work estimates that meetings consume 20% to 50% of their time.

We observed a reluctance of the employees to be part of the study. Therefore, our recommended technique(s) are low-effort.

The timesheet technique is already being used within the organisation Back2Work. This technique provides a view of the workload and its distribution from an employee’s perspective. In addition, the timesheet is an independent self-reporting technique. As a result, the employees can record the activities performed at the office, on the train, or at home. For this reason, we recommend continuing the use of the timesheet technique to record employee behaviour. However, due to a lack of information in the timesheet and inconsistent recording of the activities, the timesheet’s potential is not realised. To minimise the shortcomings of the timesheet technique, we recommend creating guidelines that all employees must adhere to (e.g., each activity should be recorded separately). In addition, Back2Work can create templates on how to fill in the timesheet descriptions e.g.,

“Meeting: with TimeCorp about IT support services”. However, the timesheet technique is implied to be imprecise and inaccurate [140]. Therefore, the timesheet technique must be combined with additional (low-effort) techniques to supplement the records.

We inquired about the techniques that are already available within the organisation. The employees use the calendar function in Microsoft Outlook to schedule activities. We propose that Back2Work uses the calendar in Microsoft Outlook to record the meetings [95]. Moreover, a separate registration can be introduced to record how the meeting rooms are used (e.g., when a meeting is taking place, for how long, and how many employees are attending). Due to COVID-19 guidelines, desks and meeting rooms must be reserved beforehand. These registrations can be used to supplement the records. Lastly, to generate additional data that cannot be recorded by the aforementioned techniques, we propose to explicitly inquire if and how the employee performed these activities [66].

Back2Work should consider change management due to employee resistance. To help employees cope with change and resistance, Back2Work must keep them informed and involved throughout the project [43].

7.3 Double the Work, and Half the Rating

Scenario 3: Record the Overlapping (Parallel) and Short-Lived Activities

Organisation OverlapS received several complaints from its employees after the annual rating. Some employees questioned why their rating was lower than expected. OverlapS explained that the employees are evaluated based on their initiative shown during projects and their performance. OverlapS uses the timesheets to determine the performance.

The employees became more vocal about the rating system and questioned the accuracy and completeness of the timesheets. The employees believe that their behaviour is not properly recorded by the timesheet technique. In fact, some employees stated that the timesheet only provides a “partial picture” of their behaviour. OverlapS took these complaints seriously and wondered which technique(s) would be best suited to record the behaviour of their employees. The employees want to be evaluated in a fair manner based on the work that is actually performed. As a result, OverlapS gained full support from its employees for this initiative. Employees continue to work from home (online). In addition, employees devote 25% of their time to data extractions, which are carried out via (video) calls on Microsoft Teams. OverlapS only criterion is that the technique does not intrude on their clients’ privacy.

The employees complained that their timesheet does not accurately record their behaviour. This means that there is a discrepancy between what employees claim to have done and what actually occurred throughout the working day. As a result, the employees are not assessed in a fair manner.

The employees stated that their timesheets only provided a “partial picture” of their behaviour. The timesheet is a self-reporting technique. This means that the employees have (unintentionally) omitted certain records [78, 94]. After further inquiring, the employees stated that they are multi-tasking on two screens while constantly being interrupted by brief disruptions such as chat messages via Microsoft Teams. These individual disruptions may be short but the duration mounts up at the end of the working day. We conclude that the timesheet technique failed to record *overlapping (parallel) activities* (“*multi-tasking*”) and *short-lived activities* (“*brief disruptions*”) with a duration of 1 or 2 minutes.

The employees still work from home, therefore all activities occur on the screen. In addition, OverlapS gained full support from its employees because they want to be assessed in a fair manner. Therefore, we can recommend an invasive technique such as screen recording to record multiple screens and sound (microphone and system audio). The screen recording technique provides highly detailed information about the online setting and allows the employees to (accurately) record the overlapping (parallel) and short-lived activities [72, 139]. However, screen recording is an invasive technique, therefore we provide two alternatives to preserve the privacy of clients during data extractions. First, employees can choose to record one screen without audio. Second, the employees can choose to stop the recording during video calls with clients. We propose to supplement the screen recording records with logs extracted from the information systems. Each employee has their own device (laptop), therefore we can use their employee ID as a unique identifier.

Chapter 8

Conclusion

This final chapter presents the most notable findings of our study and an answer to our main research question. In addition, we discuss the limitations, challenges in the research context, and future work that emerged during this study.

We identified that a detailed recording of the behaviour of the employees and their activities is crucial because this is directly related to the success and operations of an organisation [64]. This data is used to monitor the KPIs (e.g., the productivity of employees), measure the performance of employees (e.g., for the annual rating), or detect unwanted or repetitive tasks that can be (partially) automated in the future. We found abundant contemporary literature on stand-alone data collection techniques that can be used to record employee behaviour [130]. However, previous studies have almost exclusively focused on one stand-alone data collection technique [22, 86]. This was concerning because each data collection technique yields unique results but also comes with its own challenges which can impact the accuracy or completeness of the results [78, 160].

Therefore, the purpose of this study was to combine multiple data collection techniques with different characteristics to record employee behaviour and discover work patterns within an organisational context. The goal was to present the advantages, disadvantages, and the commonalities and differences between the data collection techniques to examine the level of confidence that should be placed in the analysis of this type of data. Specifically, we aimed to identify the shortcomings of each data collection technique and the impact it can have on the quality of the records. Finally, we wanted to demonstrate how multiple data collection techniques can be combined to record employee behaviour based on the context of the situation.

8.1 Answers to Sub-Questions

We formulated the following four sub-questions to answer the main research question.

SQ₁ How is the phenomenon of work patterns defined in the literature?

We introduced the phenomenon of work patterns as defined in the literature in Chapter 2. Work patterns are defined as “*the (characteristics) of work activities performed by the organisational members to execute specific activities, and accomplish practices of interest related to a task*” [104, 107]. This means the “*everyday nature of the work activities exhibited by the organisational members*” [53, p. 24]. The characteristics are the *where* and *when* the work activities are performed (location and time), *how* (the mediums and documents), but also by *whom* (the involved members) [66]. We noted that previous studies used a data collection technique to record employee behaviour against a set of predetermined classifications. The classifications introduced by Mintzberg [95] are one of the most often used for recording employee behaviour and discovering work patterns. We identified that Mintzberg’s classifications are outdated and do not include the reliance on technology in the workplace. We introduced new categories based on the literature and the requests made by the EY stakeholder.

SQ₂ What are the existing data collection techniques to record employee behaviour within an organisation?

Through a systematic literature review (Chapter 2), we examined the data collection techniques suitable to record employee behaviour in an organisational context. We found 9 techniques *interview (focus group), questionnaire, observation, think-aloud, work diary, screen recording, analysis of event logs, and archival analysis (timesheet)* that were structured according to the categorisations by Lethbridge et al. [78]. The three categorisations *direct technique, indirect technique, and independent technique* are based on the required degree of contact between the researcher and the participant. We learned that each technique provides different results, and has its own advantages and disadvantages that can impact the completeness and accuracy of the records. We compiled the characteristics in Table 2.2, which we used to conclude to what extent our study has confirmed or refuted these findings.

SQ₃ Which work patterns can be derived by using multiple data collection techniques to record employee behaviour?

We conducted a multiple-case study to record the behaviour of 6 employees within the Helix Team during the course of 6 working days. The techniques produced 136:04 hours of data and 849 activity records. We analysed (Chapter 4) the accuracy, quality, and particularities and shortcomings. We identified 58 differences related to timestamps, online versus offline activities, short-lived activities, overtime activities, and uncategorised activities. In the results (Chapter 5), we discussed these differences, strengths, weaknesses, and impact on the quality of the records per data collection technique. In Chapter 6, we presented the implications for research and for practice (for organisations, EY, and its employees). Additionally, we concluded that our study confirmed the literature findings about the characteristics of the data collection techniques. However, we were able to present implications for research and practice that were not mentioned in the literature such as the impact of overlapping (parallel) activities, time lags, and the hardship of recording short-lived activities.

SQ₄ How can the data collection techniques be combined to record employee behaviour?

We presented 3 scenarios (Chapter 7) to illustrate how multiple data collection techniques can be combined to minimise the shortcomings of each technique. The literature review and the multiple-case study showed that it is not practical for the organisation to completely rely on the observation technique. This technique is costly and requires plenty of time and effort due to intensive interaction between the researcher and participant(s). The timesheet technique and the screen recording technique are already being used or are most likely to be used by the organisation. Therefore, we examined how these techniques can be combined. We acknowledge that even after combining the screen recording and timesheet techniques, certain work patterns will not be (completely and/or accurately) recorded. For that reason, we also proposed alternative techniques that can be used to supplement the records.

8.2 Answer to Main Research Question

The insights from the sub-questions allowed us to provide an answer to the main research question.

How can different data collection techniques aid in discovering work patterns of employees within an organisation?

To conclude, our study showed that each data collection technique yields unique results, and has its advantages and disadvantages. Due to the shortcomings of the techniques, we have seen that the use of only one technique **will not** produce a complete and accurate record of the activities that occur on the screen (online), in the hallway (offline), and in the wee hours (overtime). Therefore, it is necessary to combine multiple techniques with different characteristics to be able to (completely and accurately) record employee behaviour and discover work patterns within an organisational context. Even after combining the techniques, the records will not be a representation of the actual employee behaviour to discover work patterns. Certain activities are not recorded (properly and/or

accurately) by the employee or technique, or occur outside of the information system(s). However, using different data collection techniques will minimise the shortcomings of the technique.

8.3 Limitations

We recognise the following limitations of our study.

Availability

This is a Master’s thesis project. We were restricted due to the available time to conduct the research, our availability to collect and analyse the data, and the availability of the participants (due to the busy season). This had an impact on the number of research participants and on the choice of data collection techniques. Previous studies observed the participants for a longer period of time. For example, Mintzberg [95] observed five participants for one week each while Tengblad [140] observed eight participants for four weeks each. Our study had 6 research participants. We were unable to observe the participants for more than one working day. However, this study differentiates itself from other studies because we combined three different data collection techniques that we applied to each research participant.

Technical Issues

Participant 3 faced technical issues during the use of the Snagit tool. We became aware of this once the working day had ended. We analysed the data and concluded that we missed 01:38 hours of the screen recording. We did have sound (microphone and system audio). We considered the following options. First, to accept the partial recording if we can demonstrate the partial data is still reliable and valid for our data analysis based on the audio recording or our observation notes. Second, if the reliability of the data cannot be ensured, then we would plan a new observation (either a whole or partial working day). The reason why we accepted the partial recording is that we had the sound recording, and we did not want to risk the participant exhibiting different work patterns from those noted during the initial observation.

8.4 Challenges in Research Context

Double-Hat

In this study, we served two roles simultaneously (“double-hat”). We had the role of researcher and the role of a Helix Team member. We feel that this needs to be addressed because this has an impact on how easily someone can replicate this research. It is difficult to access these kinds of organisations because they work with client data, and do not wish for outsiders to share the internal process(es). We have been part of the Helix Team for almost five years. Being part of the team under study has had a variety of advantages and challenges. The advantage is that we were not an “outsider”. We already had a relationship with the management and the members of the Helix Team. We believe that this had a positive effect on the *almost* effortless acceptance of the research requirements. We had to make agreements about the protection of the client and data. However, we feel that this was more of a formality. We had full access to the archival sources (i.e., timesheets), and approval to collect and analyse data about the behaviour of the employees. The main challenge that we experienced was the tension involved in navigating between two worlds - being a researcher and being a member of the Helix Team. We worked full-time alongside the research, thus time management was essential for us. Even though we did *try* to block certain days for our research, we still encountered issues that had to be dealt with immediately.

Unsettling Presence of the Researcher

The literature review findings indicated that participants can change their behaviour due to their awareness of being observed. We used these findings to design the research method and minimise the Hawthorne effect. However, we remained to be concerned that our presence would unsettle the participants. We created our own approach to omit this risk. We used two screens during the observation. On the screen turned towards the participant we opened a medium unrelated to our study e.g., Microsoft Outlook, Microsoft Teams, or the Internet Browser. On the screen turned away the participant we opened the standardised form. Surprisingly, by using this approach we were able to minimise the participant’s awareness of being observed. One participant even asked us “when and how” the observation would take place. This was during a coffee break at 14.00.

8.5 Future Work

The nature of our study raised a number of opportunities for future research. The main opportunity is to conduct the study without the limitations caused by COVID-19 and the busy season. COVID-19 caused the shift from offline activities (e.g., meetings) to online telephoning. Our study examined the ability of the techniques to record online and offline activities, and we presented the potential significance of the meetings. We believe that it is important to assess the ability of the techniques to record online and offline activities once the employees have fully returned to the office. The busy season limited the duration of the case study to one working day. For future work, we could set the case study duration to one week per participant. This would allow us to record the employee behaviour over a longer period of time (to record more data) and minimise the Hawthorne effect. This study is the first step toward realising the importance of combining multiple techniques. Future research may also focus on the alternative techniques as proposed in our scenarios (Chapter 7), such as the extraction of the logs from information systems, including the separate registration of the use of meeting rooms, and asking the employees to elaborate on the activities that were performed during the working day.

Bibliography

- [1] Van der Aalst, W.M.P.: Process Mining: Discovery, Conformance and Enhancement of Business Processes (2011). <https://doi.org/10.1007/978-3-642-19345-3>
- [2] Van der Aalst, W.M.P.: Extracting Event Data from Databases to Unleash Process Mining. In: BPM - Driving Innovation in a Digital World, pp. 105–128. Springer (2015)
- [3] Van der Aalst, W.M.P.: Process mining: Data Science in Action, vol. 2. Springer (2016). <https://doi.org/10.1007/978-3-662-49851-4>
- [4] Van der Aalst, W.M.P.: Process Mining: A 360 Degree Overview, pp. 3–34. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-031-08848-3_1
- [5] Van der Aalst, W.M.P., Adriansyah, A., Karla Alves de Medeiros, A., Arcieri, F., Baier, T., Blickle, T., P Jagadeesh Chandra Bose, R., van den Brand, P., Brandtjen, R., CAM Buijs, J., et al.: Process Mining Manifesto. In: International Conference on Business Process Management. pp. 169–194. Springer (2011), https://www.win.tue.nl/ieeetfpm/doku.php?id=shared:process_mining_manifesto
- [6] Van der Aalst, W.M.P., Reijers, H.A., Weijters, A., van Dongen, B., Alves de Medeiros, A., Song, M., Verbeek, H.: Business process mining: An Industrial Application. Information Systems **32**(5), 713–732 (2007). <https://doi.org/10.1016/j.is.2006.05.003>
- [7] Van der Aalst, W.M.P., Weijters, T., Maruster, L.: Workflow mining: Discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering **16**(9), 1128–1142 (2004). <https://doi.org/10.1109/TKDE.2004.47>
- [8] AFM: AFM imposes a fine on audit firm EY in connection with a violation of independence rules (October 2018), <https://www.afm.nl/en/nieuws/2018/okt/boete-ey>
- [9] Ahmadi, E., Macassa, G., Larsson, J.: Managers’ work and behaviour patterns in profitable growth SMEs. Small Business Economics **57**(2), 849–863 (2021). <https://doi.org/10.1007/s11187-020-00386-0>
- [10] Alario-Hoyos, C., Rodríguez-Triana, M.J., Scheffel, M., Arnedillo-Sánchez, I., Dennerlein, S.M., Frey, T., Broos, T., Papamitsiou, Z., Calleja, A., Tammets, K., et al.: Addressing Global Challenges and Quality Education: 15th European Conference on Technology Enhanced Learning, EC-TEL 2020, Heidelberg, Germany, September 14–18, 2020, Proceedings **12315** (2020). <https://doi.org/10.1007/978-3-030-57717-9>
- [11] Allen-Meares, P., Lane, B.A.: Social Work Practice: Integrating Qualitative and Quantitative Data Collection Techniques. Social Work **35**(5), 452–458 (1990), <http://www.jstor.org/stable/23715281>
- [12] Alteryx: Why Alteryx (2022), <https://www.alteryx.com/why-alteryx>
- [13] Anda, B., Benestad, H.C., Hove, S.E.: A multiple-case study of software effort estimation based on use case points. In: 2005 International Symposium on Empirical Software Engineering, 2005. pp. 10–pp. IEEE (2005). <https://doi.org/10.1109/ISESE.2005.1541849>
- [14] Anderson-Gough, F., Grey, C., Robson, K.: Tests of time: organizational time-reckoning and the making of accountants in two multi-national accounting firms. Accounting, Organizations and Society **26**(2), 99–122 (2001). [https://doi.org/10.1016/S0361-3682\(00\)00019-2](https://doi.org/10.1016/S0361-3682(00)00019-2)

- [15] Andrews, A.A., Pradhan, A.S.: Ethical Issues in Empirical Software Engineering: The Limits of Policy. *Empirical Software Engineering* **6**(2), 105–110 (2001). <https://doi.org/10.1023/A:1011442319273>
- [16] Arman, R., Dellve, L., Wikström, E., Törnström, L.: What health care managers do: Applying Mintzberg’s structured observation method. *Journal of Nursing Management* **17**(6), 718–729 (09 2009). <https://doi.org/10.1111/j.1365-2834.2009.01016.x>
- [17] Askeland, H.: Managerial Practice In Faith-based Welfare Organizations. *Nordic Journal of Religion and Society* **28**, 37–58 (05 2015). <https://doi.org/10.18261/ISSN1890-7008-2015-01-03>
- [18] Askeland, H., Espedal, G., Løvaas, B.J., Sirris, S.: *Understanding Values Work*. Palgrave Macmillan Cham (2020). <https://doi.org/10.1007/978-3-030-37748-9>
- [19] Augusto, A., Conforti, R., Dumas, M., Rosa, M.L., Maggi, F.M., Marrella, A., Mecella, M., Soo, A.: Automated Discovery of Process Models from Event Logs: Review and Benchmark. *IEEE Transactions on Knowledge and Data Engineering* **31**(4), 686–705 (2019). <https://doi.org/10.1109/tkde.2018.2841877>
- [20] Bade, F.M., Vollenberg, C., Koch, J., Koch, J., Coners, A.: The Dark Side of Process Mining. How Identifiable Are Users Despite Technologically Anonymized Data? A Case Study from the Health Sector. In: *Business Process Management. Lecture Notes in Computer Science*, vol. 13420, pp. 219–233. Springer International Publishing, Cham (2022). https://doi.org/10.1007/978-3-031-16103-2_16
- [21] Baier, T., Di Ciccio, C., Mendling, J., Weske, M.: Matching events and activities by integrating behavioral aspects and label analysis. *Software & Systems Modeling* **17**(2), 573–598 (2018). <https://doi.org/10.1007/s10270-017-0603-z>
- [22] Baker, L.: Observation: A Complex Research Method. *Library Trends* **55**(1), 171–189 (2006). <https://doi.org/10.1353/lib.2006.0045>
- [23] Barker, K.N.: Data Collection Techniques: Observation. *American Journal of Hospital Pharmacy* **37**(9), 1235–1245 (1980). <https://doi.org/10.1093/ajhp/37.9.1235>
- [24] Bartkowiak-Theron, I., Sappey, J.R.: The methodological identity of shadowing in social science research. *Qualitative Research Journal* **12**(1), 7–16 (2012). <https://doi.org/10.1108/14439881211222697>
- [25] Bedard, J.C., Deis, D.R., Curtis, M.B., Jenkins, J.G.: Risk Monitoring and Control in Audit Firms: A Research Synthesis. *AUDITING: A Journal of Practice & Theory* **27**(1), 187–218 (2008). <https://doi.org/10.2308/aud.2008.27.1.187>
- [26] Bekhet, A., Zauszniewski, J.: Methodological Triangulation: An Approach to Understanding Data. *Nurse Researcher* **20** **2**, 40–43 (November 2012). <https://doi.org/10.7748/nr2012.11.20.2.40.c9442>
- [27] Bell, E., Harley, B., Bryman, A.: *Business Research Methods*. Oxford University Press (2011)
- [28] Bezemer, J., Mavers, D.: Multimodal transcription as academic practice: A social semiotic perspective. *International Journal of Social Research Methodology* **14**(3), 191–206 (2011). <https://doi.org/10.1080/13645579.2011.563616>
- [29] Bolger, N., Davis, A., Rafaeli, E.: Diary Methods: Capturing Life as it is Lived. *Annual Review of Psychology* **54**(1), 579–616 (2003). <https://doi.org/10.1146/annurev.psych.54.101601.145030>
- [30] Bonazzi, G.: Between Shock Absorption and Continuous Improvement: Supervisors and Technicians in Fiat ‘Integrated Factory’. *Work, Employment and Society* **12**(2), 219–243 (1998). <https://doi.org/10.1177/0950017098122002>
- [31] Bose, R.J.C., Mans, R.S., Van der Aalst, W.M.P.: Wanna improve process mining results? In: *Proceedings of the 2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, Singapore, 16–19 April 2013. pp. 127–134 (2013). <https://doi.org/10.1109/CIDM.2013.6597227>

- [32] Buijs, J.C., Bergmans, R.F., El Hasnaoui, R.: Customer journey analysis at a financial services provider using self-service and data hub concepts. In: Proceedings of the Industry Forum at BPM 2019 co-located with 17th International Conference on Business Process Management (BPM 2019), Vienna, Austria, September 1-6, 2019. CEUR Workshop Proceedings, vol. 2428, pp. 25–36. CEUR-WS.org (2019), <http://ceur-ws.org/Vol-2428/paper3.pdf>
- [33] Busetto, L., Wick, W., Gumbinger, C.: How to Use and Assess Qualitative Research Methods. *Neurological Research and Practice* **2**(1), 1–10 (2020). <https://doi.org/10.1186/s42466-020-00059-z>
- [34] Cavaye, A.L.M.: Case study research: a multi-faceted research approach for IS. *Information Systems Journal* **6**(3), 227–242 (1996). <https://doi.org/10.1111/j.1365-2575.1996.tb00015.x>
- [35] Cln, L.I.S.: Data collection techniques a guide for researchers in humanities and education. *International Research Journal of Computer Science and Information Systems (IRJCSIS)* **2**(3), 40–44 (2013)
- [36] Connaway, L.S., Radford, M.L.: *Research Methods in Library and Information Science*. Library and information science text series, Libraries Unlimited (2016)
- [37] Corporate Finance Institute: Data loss: When data is destroyed, deleted, corrupted, or made unreadable by users and software applications (July 2021), <https://corporatefinanceinstitute.com/resources/knowledge/data-analysis/data-loss/>
- [38] Cowan, J.: The potential of cognitive think-aloud protocols for educational action-research. *Active Learning in Higher Education* **20**(3), 219–232 (2019). <https://doi.org/10.1177/1469787417735614>
- [39] Denzin, Norman K and Lincoln, Yvonna S: *The SAGE Handbook of Qualitative Research*. Sage Publications (2011)
- [40] Di Sibio, C.: EY value realized: Global impact in 2021 (September 2021), https://www.ey.com/en_gl/global-review/value-realized-2021
- [41] Diem, T.: The effect of auditors’ workload pressure and compression on audit quality. Master’s Thesis, University of Oulu (2016), <http://jultika.oulu.fi/files/nbnfioulu-201605121710.pdf>
- [42] Dramski, M.: Missing Data Problem in the Event Logs of Transport Processes. In: *Smart Solutions in Today’s Transport*. vol. 715, pp. 110–120. Springer International Publishing, Cham (2017). https://doi.org/10.1007/978-3-319-66251-0_9
- [43] Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: *Fundamentals of Business Process Management*. Springer Berlin Heidelberg, 2nd edn. (March 2018). <https://doi.org/10.1007/978-3-662-56509-4>
- [44] Eden, R., Syed, R., Leemans, S.J., Buijs, J.A.: A Case Study of Inconsistency in Process Mining Use: Implications for the Theory of Effective Use. In: *International Conference on Business Process Management*. pp. 363–379. Springer (2021)
- [45] Ericsson, A.: Valid and Non-Reactive Verbalization of Thoughts During Performance of Tasks: Towards a Solution to the Central Problems of Introspection as a Source of Scientific Data. *Journal of Consciousness Studies* **10**(9-10), 1–18 (2003)
- [46] Ericsson, K.A., Simon, H.A.: *Protocol Analysis: Verbal Reports as Data*. The MIT Press (1984). <https://doi.org/10.7551/mitpress/5657.001.0001>
- [47] Espedal, G., Jelstad Løvaas, B., Sirris, S., Wæraas, A.: *Researching Values: Methodological Approaches for Understanding Values Work in Organisations and Leadership*. Springer Nature (2022)
- [48] European Union: DIRECTIVE (EU) 2016/680 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons 2008/977/JHA OJ L 119 (2016), <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC>

- [49] Fenton, N., Bryman, A., Deacon, D., Birmingham, P.: ‘Sod off and find us a boffin’: journalists and the social science conference. *The Sociological Review* **45**(1), 1–23 (1997). <https://doi.org/10.1111/1467-954X.00051>
- [50] Flewitt, R.: Using video to investigate preschool classroom interaction: Education research assumptions and methodological practices. *Visual Communication* **5**(1), 25–50 (2006). <https://doi.org/10.1177/1470357206060917>
- [51] Ghasemi, M., Amyot, D.: Process mining in healthcare: a systematised literature review. *International Journal of Electronic Healthcare* **9**(1), 60–88 (2016). <https://doi.org/10.1504/IJEH.2016.078745>
- [52] Glueck, W.F., Snyder, N.: How Managers Plan-The Analysis of Managers’ Activities. *Long Range Planning* **13**(1), 70–76 (1980). [https://doi.org/10.1016/0024-6301\(80\)90058-8](https://doi.org/10.1016/0024-6301(80)90058-8)
- [53] Gratton, L.: *The Shift: The Future of Work is Already Here*. William Collins (2014)
- [54] Greenhalgh, T., Peacock, R.: Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ* **331**(7524), 1064–1065 (2005). <https://doi.org/10.1136/bmj.38636.593461.68>
- [55] Greenwood, R., Empson, L.: The Professional Partnership: Relic or Exemplary Form of Governance? *Organization Studies* **24**(6), 909–933 (2003). <https://doi.org/10.1177/0170840603024006005>
- [56] Greenwood, R., Hinings, C.R., Brown, J.: “P2-Form” Strategic Management: Corporate Practices in Professional Partnerships. *Academy of Management Journal* **33**(4), 725–755 (1990). <https://doi.org/10.5465/256288>
- [57] Günther, C.W., Rozinat, A., Van der Aalst, W.M.P.: Activity Mining by Global Trace Segmentation. In: *International Conference on Business Process Management*. pp. 128–139. Springer (2009)
- [58] Hak, T.: Waarnemingsmethoden in kwalitatief onderzoek. In: *Kwalitatief onderzoek*, pp. 13–25. Springer (2007)
- [59] Heath, C., Hindmarsh, J.: *Analysing Interaction: Video, Ethnography and Situated Conduct*, pp. 99–121. Sage (2002). <https://doi.org/10.4135/9781849209656>
- [60] Hevner, A., Chatterjee, S.: Design Science Research in Information Systems. In: *Design Research in Information Systems*, pp. 9–22. Springer (2010)
- [61] Ho, W.Y.J.: *Translanguaging in Online Language Learning: Case Studies of Self-Directed Chinese Learning of Multilingual Adults*. Ph.D. thesis, UCL (University College London) (2018), <https://discovery.ucl.ac.uk/id/eprint/10044705/>
- [62] Ho, W.Y.J.: ‘I knew that you were there, so I was talking to you’: the use of screen-recording videos in online language learning research. *Qualitative Research* **21**(1), 120–139 (2021). <https://doi.org/10.1177/1468794119885044>
- [63] Hodson, R.: Organizational Ethnographies: An Underutilized Resource in the Sociology of Work. *Social Forces* **76**(4), 1173–1208 (1998). <https://doi.org/10.1093/sf/76.4.1173>
- [64] Ind, N.: Inside out: How employees build value. *Journal of Brand Management* **10**(6), 393–402 (2003). <https://doi.org/10.1057/palgrave.bm.2540136>
- [65] Jentoft, N., Olsen, T.S.: Against the flow in data collection: How data triangulation combined with a ‘slow’ interview technique enriches data. *Qualitative Social Work* **18**(2), 179–193 (2019). <https://doi.org/10.1177/1473325017712581>
- [66] Johnson, B.M.: *Understanding the work of top managers: A shadowing study of Canadian healthcare CEOs*. Ph.D. thesis, University of Warwick (2013), <https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.589919>
- [67] Kabir, S.M.S.: *Basic Guidelines for Research: An Introductory Approach for All Disciplines* pp. 168–180 (2016)

- [68] Kawaf, F.: Capturing digital experience: The method of screencast videography. *International Journal of Research in Marketing* **36**(2), 169–184 (2019). <https://doi.org/10.1016/j.ijresmar.2018.11.002>
- [69] Kherbouche, M.O., Laga, N., Masse, P.A.: Towards a better assessment of event logs quality. In: 2016 IEEE Symposium Series on Computational Intelligence (SSCI). pp. 1–8. IEEE (2016)
- [70] Kitchenham, B.: *Procedures for Performing Systematic Reviews*. Keele, UK, Keele University **33**, 1–26 (2004)
- [71] Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering—A systematic literature review. *Information and Software Technology* **51**(1), 7–15 (2009). <https://doi.org/10.1016/j.infsof.2008.09.009>
- [72] Krieter, P.: Can I record your screen? Mobile screen recordings as a long-term data source for user studies. In: *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*. pp. 1–10 (2019)
- [73] Lammers, W.J., Badia, P.: *Fundamentals of Behavioral Research Textbook*. Wadsworth Publishing (2004), <https://uca.edu/psychology/fundamentals-of-behavioral-research-textbook/>
- [74] Langford, J., McDonough, D.: *Focus groups: Supporting Effective Product Development*. CRC press (2003). <https://doi.org/10.4324/9780203302743>
- [75] Lau, A.W., Newman, A.R., Broedling, L.A.: The Nature of Managerial Work in the Public Sector. *Public Administration Review* **40**(5), 513–520 (1980), <http://www.jstor.org/stable/3110205>
- [76] Lederman, L.C.: Assessing educational effectiveness: The focus group interview as a technique for data collection. *Communication education* **39**(2), 117–127 (1990). <https://doi.org/10.1080/03634529009378794>
- [77] Lethbridge, T.C., Sim, S.E., Singer, J.: Software Anthropology: Performing Field Studies in Software Companies. Special Issue on Empirical Studies of IEEE Transactions on Software Engineering pp. 1–63 (1998)
- [78] Lethbridge, T.C., Sim, S.E., Singer, J.: Studying Software Engineers: Data Collection Techniques for Software Field Studies. *Empirical Software Engineering* **10**(3), 311–341 (2005). <https://doi.org/10.1007/s10664-005-1290-x>
- [79] Lincoln, Y.S., Guba, E.G.: *Naturalistic inquiry*. Sage Publications (1985)
- [80] Low, J.: A Pragmatic Definition of the Concept of Theoretical Saturation. *Sociological Focus* **52**(2), 131–139 (2019). <https://doi.org/10.1080/00380237.2018.1544514>
- [81] Mack, N.: *Qualitative Research Methods: A Data Collector’s Field Guide* (2005)
- [82] Mannhardt, F., de Leoni, M., Reijers, H.A., Van der Aalst, W.M.P., Toussaint, P.J.: From Low-Level Events to Activities - A Pattern-Based Approach. In: *Business Process Management - 14th International Conference, BPM 2016, Rio de Janeiro, Brazil, September 18-22, 2016. Proceedings. Lecture Notes in Computer Science*, vol. 9850, pp. 125–141. Springer (2016). https://doi.org/10.1007/978-3-319-45348-4_8
- [83] Mans, R.S., Schonenberg, M.H., Song, M., Van der Aalst, W.M.P., Bakker, P.J.M.: Application of Process Mining in Healthcare – A Case Study in a Dutch Hospital. In: *Biomedical Engineering Systems and Technologies*. pp. 425–438. Springer Berlin Heidelberg (2009). https://doi.org/10.1007/978-3-540-92219-3_32
- [84] Marin-Castro, H.M., Tello-Leal, E.: Event Log Preprocessing for Process Mining: A Review. *Applied Sciences* **11**(22), 10556 (2021). <https://doi.org/10.3390/app112210556>
- [85] McCall, M.: Systematic Field Observation. *Annual Review of Sociology* **10**, 263–282 (1984), <http://www.jstor.org/stable/2083176>

- [86] McDonald, S.: Studying actions in context: a qualitative shadowing method for organizational research. *Qualitative Research* **5**(4), 455–473 (2005). <https://doi.org/10.1177/1468794105056923>
- [87] Meunier, D., Vasquez, C.: On Shadowing the Hybrid Character of Actions: A Communicational Approach. *Communication Methods and Measures* **2**(3), 167–192 (2008). <https://doi.org/10.1080/19312450802310482>
- [88] Microsoft: When to use Microsoft Project, Planner, To Do, or the Tasks app in Teams (2021), <https://support.microsoft.com/en-us/office/when-to-use-microsoft-project-planner-to-do-or-the-tasks-app-in-teams-8f950d32-d5f4-40db-a8b7-4d1b82b55e17#:~:text=Microsoft%20Planner,-Quickly%20create%20plans&text=Planner%20is%20a%20light%20weight,charts%20of%20your%20team's%20progress.>
- [89] Microsoft: Hulp en educatie voor Excel (2022), <https://support.microsoft.com/nl-nl/excel>
- [90] Microsoft: Hulp en educatie voor Word (2022), <https://support.microsoft.com/nl-nl/word>
- [91] Microsoft: Notepads App (2022), <https://apps.microsoft.com/store/detail/notepads-app/9NHL4NSC67WM?hl=nl-nl&gl=nl>
- [92] Microsoft: Outlook help & learning (2022), <https://support.microsoft.com/en-us/outlook>
- [93] Microsoft: What is Microsoft Teams? (2022), <https://support.microsoft.com/en-us/topic/what-is-microsoft-teams-3de4d369-0167-8def-b93b-0eb5286d7a29>
- [94] Mills, A.J., Durepos, G., Wiebe, E.: *Encyclopedia of Case Study Research*. Sage Publications (2009). <https://doi.org/10.4135/9781412957397>
- [95] Mintzberg, H.: Structured Observation as a Method to Study Managerial Work. *Journal of Management Studies* **7**(1), 87–104 (1970). <https://doi.org/10.1111/j.1467-6486.1970.tb00484.x>
- [96] Mintzberg, H.: Henry MINTZBERG, The Nature of Managerial Work (1973) & Simply Managing: What Manager Do-And Can Do Better (2013). *Mngement* **18**(2), 186–188 (2015). <https://doi.org/10.3917/mana.182.0186>
- [97] Morgan, D.L.: *Focus Groups as Qualitative Research*, *Qualitative Research Methods*, vol. 16. Sage Publications (1996). <https://doi.org/10.4135/9781412984287>
- [98] Morgan, G.A., Harmon, R.J.: Data collection techniques. *Journal-American Academy Of Child And Adolescent Psychiatry* **40**(8), 973–976 (2001). <https://doi.org/10.1097/00004583-200108000-00020>
- [99] Morse, J.M.: *Determining Sample Size*, vol. 10. Sage Publications (2000). <https://doi.org/10.1177/104973200129118183>
- [100] Mourão, E., Kalinowski, M., Murta, L., Mendes, E., Wohlin, C.: Investigating the Use of a Hybrid Search Strategy for Systematic Reviews. In: *2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*. pp. 193–198. IEEE (2017). <https://doi.org/10.1109/ESEM.2017.30>
- [101] Mozilla: What is a web browser? (2022), <https://www.mozilla.org/en-US/firefox/browsers/what-is-a-browser/#:~:text=A%20web%20browser%20takes%20you%20anywhere%20on%20the%20internet,.are%20transmitted%20on%20the%20web.>
- [102] Munoz-Gama, J., Martin, N., Fernandez-Llatas, C., Johnson, O.A., Sepúlveda, M., Helm, E., Galvez-Yanjari, V., Rojas, E., Martinez-Millana, A., Aloini, D., Amantea, I.A., Andrews, R., Arias, M., Beerepoot, I., Benevento, E., Burattin, A., Capurro, D., Carmona, J., Comuzzi, M., Dalmas, B., de la Fuente, R., Di Francescomarino, C., Di Ciccio, C., Gatta, R., Ghidini, C., Gonzalez-Lopez, F., Ibanez-Sanchez, G., Klasky, H.B., Prima Kurniati, A., Lu, X., Mannhardt, F., Mans, R., Marcos, M., Medeiros de Carvalho, R., Pegoraro, M., Poon, S.K., Pufahl, L., Reijers, H.A., Remy, S., Rinderle-Ma, S., Sacchi, L., Seoane, F., Song, M., Stefanini, A., Sulis, E., ter Hofstede, A.H., Toussaint, P.J., Traver, V., Valero-Ramon, Z., van de Weerd, I., van der Aalst, W.M., Vanwersch, R., Weske, M., Wynn, M.T., Zerbato, F.: *Process Mining*

- for Healthcare: Characteristics and challenges. *Journal of Biomedical Informatics* **127**, 1–15 (2022). <https://doi.org/10.1016/j.jbi.2022.103994>
- [103] Nicholas, D., Huntington, P., Watkinson, A.: Scholarly journal usage: the results of deep log analysis. *Journal of Documentation* **61**(2), 248–280 (2005). <https://doi.org/10.1108/00220410510585214>
- [104] Nicolini, D.: *Practice Theory, Work, and Organization*. Oxford University Press (2012)
- [105] Nunan, D., David, N., Swan, M.: *Research Methods in Language Learning*. Cambridge Language Teaching Library, Cambridge University Press (1992)
- [106] Olsen, W.: *Data Collection: Key Debates and Methods in Social Research*. Sage Publications (2011). <https://doi.org/10.4135/9781473914230>
- [107] Orlikowski, W.J.: Sociomaterial Practices: Exploring Technology at Work. *Organization Studies* **28**(9), 1435–1448 (2007). <https://doi.org/10.1177/0170840607081138>
- [108] Patton, M.Q.: *How to use qualitative methods in evaluation*. Sage Publications (1987)
- [109] Patton, M.Q.: Enhancing the Quality and Credibility of Qualitative Analysis. *Health Services Research* **34**, 1189–1208 (1999)
- [110] Pavlov, I.: The main features of 7-Zip (2022), <https://www.7-zip.org/>
- [111] Perlow, L.A.: Boundary Control: The Social Ordering of Work and Family Time in a High-Tech Corporation. *Administrative Science Quarterly* **43**(2), 328–357 (1998)
- [112] Perlow, L.A.: The Time Famine: Toward a Sociology of Work Time. *Administrative science quarterly* **44**(1), 57–81 (1999)
- [113] Pervan, G., Maimbo, M.: Designing a Case Study Protocol for Application in IS research. In: *Proceedings of the Ninth Pacific Asia Conference on Information Systems*. pp. 1281–1292 (2005)
- [114] Pickard, A.J.: *Research Methods in Information*. Facet Publishing (2013). <https://doi.org/10.29085/9781783300235>
- [115] Project Jupyter: Free software, open standards, and web services for interactive computing across all programming languages (2022), <https://jupyter.org/>
- [116] Quinlan, E.: Conspicuous Invisibility: Shadowing as a Data Collection Strategy. *Qualitative Inquiry* **14**(8), 1480–1499 (2008). <https://doi.org/10.1177/1077800408318318>
- [117] Reeves, B., Ram, N., Robinson, T.N., Cummings, J.J., Giles, C.L., Pan, J., Chiatti, A., Cho, M., Roehrick, K., Yang, X., et al.: Screenomics: A Framework to Capture and Analyze Personal Life Experiences and the Ways that Technology Shapes Them. *Human–Computer Interaction* **36**(2), 150–201 (2019)
- [118] Robson, C., McCartan, K.: *Real World Research: A Resource for Social Scientists and Practitioner-Researchers*. Blackwell Publishing (2002)
- [119] Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: A literature review. *Journal of Biomedical Informatics* **61**, 224–236 (2016). <https://doi.org/10.1016/j.jbi.2016.04.007>
- [120] Ruhleder, K., Jordan, B.: Capturing Complex, Distributed Activities: Video-Based Interaction Analysis as a Component of Workplace Ethnography. In: *Information Systems and Qualitative Research*, pp. 246–275. Springer (1997)
- [121] Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering* **14**(2), 131–164 (2009). <https://doi.org/10.1007/s10664-008-9102-8>
- [122] Salinger, S., Plonka, L., Prechelt, L.: A Coding Scheme Development Methodology Using Grounded Theory For Qualitative Analysis Of Pair Programming. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments* **4** (2008). <https://doi.org/10.17011/ht/urn.200804151350>

- [123] Sandelowski, M.: Sample size in qualitative research. *Research in Nursing & Health* **18**(2), 179–183 (1995). <https://doi.org/10.1002/nur.4770180211>
- [124] Saunders, M., Lewis, P., Thornhill, A.: *Research Methods for Business Students*. Pearson Education (2009)
- [125] Shipman, M.D.: *The Limitations of Social Research*. Aspects of Modern Sociology, Longman (1997)
- [126] Shull, F., Lanubile, F., Basili, V.R.: Investigating Reading Techniques for Object-Oriented Framework Learning. *IEEE Transactions on Software Engineering* **26**(11), 1101–1118 (2000). <https://doi.org/10.1109/32.881720>
- [127] Shull, F., Singer, J., Sjøberg, D.I.: *Guide to Advanced Empirical Software Engineering*. Springer (2007). <https://doi.org/10.1007/978-1-84800-044-5>
- [128] Signiant: Why You Need Media Shuttles (2022), <https://www.signiant.com/products/media-shuttle/>
- [129] Singer, J., Vinson, N.G.: Ethical Issues in Empirical Studies of Software Engineering. *IEEE Transactions on Software Engineering* **28**(12), 1171–1180 (2002). <https://doi.org/10.1109/TSE.2002.1158289>
- [130] Singh, R.: *Research Methods in Library and Information Science*. University News **36**, 7–11 (November 1998)
- [131] Sirris, S., Lindheim, T., Askeland, H.: Observation and Shadowing: Two Methods to Research Values and Values Work in Organisations and Leadership. In: *Researching Values*, pp. 133–151. Palgrave Macmillan, Cham (2022)
- [132] Slater, M.: *Research Methods in Library and Information Studies*. Library Association (1990)
- [133] Smagorinsky, P.: The Reliability and Validity of Protocol Analysis. *Written Communication* **6**(4), 463–479 (1989). <https://doi.org/10.1177/0741088389006004003>
- [134] Smagorinsky, P.: Rethinking protocol analysis from a cultural perspective. *Annual Review of Applied Linguistics* **21**, 233–245 (2001). <https://doi.org/10.1017/S0267190501000149>
- [135] Stake, R.E.: *Multiple Case Study Analysis*. Guilford Publications (2006)
- [136] Stickdorn, M., Hormess, M.E., Lawrence, A., Schneider, J.: *This is Service Design Doing: Applying Service Design and Design Thinking in the Real World*. O’Reilly (2018)
- [137] Suriadi, S., Andrews, R., ter Hofstede, A.H., Wynn, M.T.: Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs. *Information Systems* **64**, 132–150 (2017). <https://doi.org/10.1016/j.is.2016.07.011>
- [138] Suriadi, S., Wynn, M.T., Ouyang, C., ter Hofstede, A.H., Dijk, N.J.v.: Understanding process behaviours in a large insurance company in Australia: A case study. In: *International Conference on Advanced Information Systems Engineering*. pp. 449–464. Springer (2013)
- [139] Tang, J.C., Liu, S.B., Muller, M., Lin, J., Drews, C.: Unobtrusive but Invasive: Using Screen Recording to Collect Field Data on Computer-Mediated Interaction. In: *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*. pp. 479–482. Association for Computing Machinery (2006). <https://doi.org/10.1145/1180875.1180948>
- [140] Tengblad, S.: Is there a ‘New Managerial Work’? A Comparison with Henry Mintzberg’s Classic Study 30 Years Later. *Journal of Management Studies* **43**(7), 1437–1461 (2006). <https://doi.org/10.1111/j.1467-6486.2006.00651.x>
- [141] Tengblad, S.: *The Work of Managers: Towards a Practice Theory of Management*. Oxford University Press (2012). <https://doi.org/10.1093/acprof:oso/9780199639724.001.0001>
- [142] The Python Software Foundation: What is Python (2022), <https://docs.python.org/3/faq/general.html>
- [143] Tulowitzki, P.: Shadowing school principals: what do we learn? *Educational Management Administration & Leadership* **47**(1), 91–109 (2019). <https://doi.org/10.1177/1741143217725325>

- [144] Urquhart, C.: Observation research techniques. *Journal of the European Association for Health Information and Libraries* **11**(3), 29–31 (2015), <http://ojs.eahil.eu/ojs/index.php/JEAHIL/article/view/61>
- [145] Utrecht University: Research Institute of Information and Computing Sciences - Ethics and privacy informed consent examples (2022), <https://www.uu.nl/en/research/institute-of-information-and-computing-sciences/ethics-and-privacy>
- [146] Vasileiou, K., Barnett, J., Thorpe, S., Young, T.: Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC medical research methodology* **18**(1), 1–18 (2018)
- [147] Vassiliadis, P., Simitsis, A., Skiadopoulou, S.: Conceptual Modeling for ETL Processes. In: *Proceedings of the 5th ACM International Workshop on Data Warehousing and OLAP*. p. 14–21. DOLAP '02, Association for Computing Machinery, New York, NY, USA (2002). <https://doi.org/10.1145/583890.583893>
- [148] Vygotsky, L.S.: *Thought and language*. MIT press (2012). <https://doi.org/10.1037/11193-000>
- [149] W3Schools: JSON - Introduction (2022), https://www.w3schools.com/js/js_json_intro.asp
- [150] Walker, C.R., Guest, R.H., Turner, A.N.: *The Foreman on the Assembly Line*. Routledge (2017)
- [151] Wang, D., Mark, G.: Internet Censorship in China: Examining User Awareness and Attitudes. *ACM Transactions on Computer-Human Interaction* **22**(6), 1–22 (2015). <https://doi.org/10.1145/2818997>
- [152] Ward, J., Shortt, H.: *The SAGE Handbook of Qualitative Business and Management Research Methods: Methods and Challenges*. SAGE Publications (2018)
- [153] Weidt, F., Silva, R.: *Systematic Literature Review in Computer Science - A Practical Guide 1* (2016). <https://doi.org/10.13140/RG.2.2.35453.87524>
- [154] Wohlin, C.: Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering. In: *Proceedings of the 18th international conference on evaluation and assessment in software engineering*. pp. 1–10 (2014). <https://doi.org/10.1145/2601248.2601268>
- [155] Wohlin, C., Runeson, P., Höst, M., Ohlsson, M., Regnell, B., Wesslén, A.: *Experimentation in Software Engineering*. Springer, Germany (2012). <https://doi.org/10.1007/978-3-642-29044-2>
- [156] Wynn, M.T., Suriadi, S., Eden, R., Poppe, E., Pika, A., Andrews, R., ter Hofstede, A.H.: Grounding Process Data Analytics in Domain Knowledge: A Mixed-Method Approach to Identifying Best Practice. In: *International Conference on Business Process Management*. pp. 163–179. Springer (2019). https://doi.org/10.1007/978-3-030-26643-1_10
- [157] Yang, H., Van Dongen, B.F., Ter Hofstede, A.H.M., Wynn, M., Wang, J.: Estimating Completeness of Event Logs. *BPM Center Report* **12** (2012)
- [158] Yin, R.K.: *Case Study Research Design and Methods*, vol. 5. Sage Publications (2014)
- [159] Yin, R.K.: *Case Study Research and Applications: Design and methods*. Sage Publications (2018)
- [160] Zwijze-Koning, K.H., De Jong, M.D.: Auditing Information Structures in Organizations: A Review of Data Collection Techniques for Network Analysis. *Organizational Research Methods* **8**(4), 429–453 (2005). <https://doi.org/10.1177/1094428105280120>

Appendix A

Literature Review

Table A.1 presents the data collection techniques including a description, an example case in which the technique has been used, the (dis-)advantages, and the sources split by the synonyms found during the literature review. During the review, it became apparent that various synonyms are used to define each data collection technique.

Table A.1: Overview of the data collection techniques

Technique	Description	(Dis-)advantages	Sources
<i>Direct Technique(s)</i> Interview	<p>The interview is one of the most commonly used data collection techniques for the collection of qualitative, descriptive, and in-depth data [11, 36, 81]. The purpose of the interview is to uncover rich and detailed data about the participant's behaviour [127], views, perspectives [65], motivations, and subjective experiences [33, 58]. Out of the three types of interviews, the fully structured interview is most similar to the questionnaire due to the pre-established questions. However, the interview has an interactive nature due to the face-to-face setting which allows the researcher to observe visual and verbal clues [114]. An interview carried out in a group setting (three to 12 participants [106]) is a focus group.</p> <p>Example: Askeland [17] interviewed managers in faith-based welfare organisations regarding their behaviour throughout the working day as a manager in order to compare managerial roles across organisations.</p>	<p>Advantages: The one-on-one setting of the interview is seen as an advantage because it allows the researcher to build a stronger relationship with the participant [36], therefore, allowing to address (socially) sensitive topics [67, 81]. The advantage of the focus group is the positive group dynamic [74, 78, 127], which can stimulate conversation amongst the participants [81]. This technique allows the collection of data (e.g., a variety of facts, beliefs, attitudes, responses, perspectives, and explanations) from multiple sources at once [114, 118].</p> <p>Disadvantages: The interview and focus group are retrospective techniques meaning that the obtained data can consist of recalling which results in inaccurate memories [45, 134]. The focus group is prone to bias, and therefore, can produce unreliable results due to the group setting (i.e., because the participants provide socially acceptable answers), faulty questions, or an inexperienced moderator [36, 78, 127].</p>	<p>Interview [11, 18, 33, 35, 36, 47, 58, 65, 67, 76–79, 81, 97, 98, 106, 118, 121, 127, 131, 152]</p> <p>Focus Group [18, 33, 35, 36, 47, 58, 67, 74, 76–78, 81, 97, 98, 106, 114, 118, 121, 127, 131, 152]</p> <p>Discussion Group [47, 58, 67, 76–78, 97, 106, 114, 127, 131, 152]</p>
			Continued on next page

Table A.1 – continued from previous page

Technique	Description	(Dis-)advantages	Sources
Questionnaire	<p>The technique that is most frequently employed involving human participants is the interview and questionnaire [36, 98, 114, 118]. Kabir [67] claims that questionnaires allow the measurement of distinct variables such as behaviour, facts, preferences, and attitudes. Lethbridge et al. [78] concur, stating that the technique can be used to gather details on the process. There are two types to administer the questionnaire. First, it can be self-administered, meaning that the questionnaire is transmitted electronically (e.g., by e-mail) and completed by the participant. Second, the researcher conducts the questionnaire via telephone or face-to-face during the interviewer-completed questionnaire.</p> <p>Example: Ahmadi et al. [9] distributed the questionnaire amongst 133 managers from small- and medium-sized Swedish companies to investigate their behavioural and work patterns.</p>	<p>Advantages: The self-administered questionnaire is transmitted electronically, meaning that a large and geographically dispersed population can be obtained in a (cost-)efficient manner whilst providing complete confidentiality [67]. The advantage of the interviewer-completed questionnaire is the direct contact between the researcher and participant, resulting in a higher response rate (50 – 70%) compared to the self-administered questionnaire (30 – 50%) [11, 118, 124].</p> <p>Disadvantages: The retrospective nature of the questionnaire, interview, and focus group has a limiting effect [36, 45, 77, 78, 127, 134]. According to Allen-Mearns et al. [11], the collected data is based on the participant’s self-reports (i.e., memory) which can result in misleading memories. Robson & McCartan [118] add that participants are unlikely to accurately record their behaviour, attitudes, and beliefs due to the need to provide socially acceptable responses.</p>	<p>Questionnaire [11, 33, 35, 36, 58, 67, 77, 78, 98, 106, 114, 118, 124, 125, 127, 152] Survey [33, 35, 36, 58, 67, 77, 78, 98, 106, 114, 118, 124, 125, 127, 152]</p>
		Continued on next page	

Table A.1 – continued from previous page

Technique	Description	(Dis-)advantages	Sources
Observation	<p>The diary method is criticised because it consists of self-reported behaviour, therefore, allowing the employees to record the tasks which they deem to be of importance resulting in an incomplete picture of their behaviour [33, 95]. Hence, observation was offered as an alternative form of research to record employee behaviour and develop an understanding of work patterns. The observation allows researchers to record the observed employee behaviour instead of self-reported behaviour (work diary) as a real-time representation of the studied phenomena [36, 61, 97, 127], in contrast to screen recording which is a “retrospective observation” [62]. Patton [108] proposed five dimensions of variation in observational approaches including the role of the researcher. This dimension is categorised from participant observation (active), in which the researcher is part of the observed setting, to non-participant observation (passive), in which there is no involvement between the researcher and participant.</p> <p>Example: Mintzberg [95] observed five chief executives for one week each to record their behaviour to discover work patterns.</p>	<p>Advantages: The observation is considered to be a valuable data collection technique used to study employee behaviour [22]. According to Allen-Mearns et al. [11], the researcher records the observed behaviour, compared to the work diary which is created by the participant. Connaway & Radford [36] and Morgan & Harmon [97] claim that the use of this technique allows the researcher to compare the observed behaviour with the self-reported behaviour (work diary) in order to find similarities and discrepancies.</p> <p>Disadvantages: Within the literature, a recurring disadvantage is the “Hawthorne” or “observer effect” caused by the researcher being physically present in the observed environment [118, 160]. The participant may alter their behaviour due to the awareness of being observed [67, 86, 106, 125]. However, the effects become insignificant because participants quickly revert to their normal work pattern whilst carrying out familiar activities (i.e., regular work) [23, 95], in a natural setting [85, 131], and surrounded by colleagues [27]. A significant drawback is that observation results in an overwhelming amount of data, impacting the time and effort needed for the data analysis [116]. Lastly, gaining access to a company can be challenging [66], especially if confidential information (e.g., client data) is involved.</p>	<p>Observation [11, 22, 23, 35, 36, 47, 58, 61, 66, 67, 77, 78, 86, 95, 97, 106, 108, 114, 118, 121, 125, 127, 160] Shadowing [47, 66, 77, 78, 86, 106, 127]</p>
		Continued on next page	

Table A.1 – continued from previous page

Technique	Description	(Dis-)advantages	Sources
Think-aloud	<p>According to Numan [105], the difficulty is that most of the actions performed by employees take place within their heads and are thus imperceptible. Traditionally, the main concerns were related to the observable qualities of human behaviour, rendering the invisible work insignificant. However, researchers deem this to be insufficient, resulting in the need to identify the mental models (processes) of the participants by using the think-aloud. This is an introspective data collection technique [61], in which participants are asked to document their thoughts, feelings, and reasoning process whilst performing a task [46].</p> <p>Example: Cowan [38] used the think-aloud to study the observed activities and report on the learning process.</p>	<p>Advantages: Cowan [38] and Lethbridge et al. [77] argue that the think-aloud is a useful technique to provide rich insights into the participant’s thought process. Smagorinsky [133] adds that think-aloud is used to seek if the participants follow the sequential step-by-step reasoning, or identify the feelings that are being generated by performing certain activities. The questionnaires, interviews, and focus groups are retrospective techniques meaning that the obtained data consists of recalling, which can lead to inaccurate memories [45, 134]. The think-aloud is an introspective technique because all thoughts are verbalised from immediate memory whilst the participant is in the process of perceiving and retaining the information [134] therefore is a source of reliable data [38]. Lethbridge et al. [78] add that the think-aloud provides a unique view of the participant’s mental models.</p> <p>Disadvantages: Shull et al. [127] state that the think-aloud is a relatively easy technique to deploy because the only requirement is for the participants to verbalise their thoughts. However, the unfamiliarity of simultaneously narrating, thinking and performing activities [46, 61] may impact the participant’s articulation of their thought process [38], which results in incomplete or inaccurate data [148].</p>	<p>Think-aloud Protocol Analysis [38, 61, 62, 77, 78, 105, 127, 134] Verbal Report [38, 45, 46, 61, 134] Think-aloud Method [38, 61, 62, 105] Verbal Data [38] Verbal Thinking [148]</p>
Continued on next page			

Table A.1 – continued from previous page

Technique	Description	(Dis-)advantages	Sources
<p>Work Diary</p> <p>The work diary is a technique used to collect data about participant behaviour, activities, and experiences [29]. The participant is required to self-report details regarding the performed activities that occur throughout the working day [78]. Allen-Mearns et al. [11] argued that the contribution of the work diary is of lesser value due to the self-monitoring report characteristic compared to the researcher’s observation. However, Lethbridge et al. [78] disagreed with this statement, claiming that the self-reporting characteristics should not be seen as a limitation because participants record their activities continuously rather than in retrospect (as in questionnaires and interviews).</p> <p>Example: Shull et al. [126] used the work diary technique on 15 students who were asked to submit a self-report of the performed work every week. The self-report contained the number of hours spent per activity including a description.</p>	<p>Advantages: The diary requires participants to record occurrences continuously as opposed to retrospective techniques such as the questionnaire and interview. Similar to the think-aloud, the work diary is an introspective data collection technique [61, 78]. This technique minimises the time between the experience and the reporting (by the participant) of thus said experience [29, 45, 134]. This has a positive effect on the reliability of the data.</p> <p>Disadvantages: The work diary is a self-report in which the participant must record the events that occur throughout the working day [11, 78]. This places a high burden on the participants, due to the constant need to keep track of the performed activities [160]. This burden may even obstruct the participant’s ability to complete their work [78]. Moreover, according to Tang et al. [139], the quality and completeness of the data depend on the commitment of the participant and their ability to describe their working day.</p>	<p>(Work) Diary [29, 36, 67, 77, 78, 86, 114, 127, 132, 139, 160]</p> <p>Self(-monitoring) Report [11, 77, 78, 127, 132, 139]</p> <p>Participant Diary [114]</p> <p>Self-observation [114]</p>	
Continued on next page			

Table A.1 – continued from previous page

Technique	Description	(Dis-)advantages	Sources
<p><i>Indirect Technique</i></p> <p>Screen recording</p>	<p>The use of video recording during an interview and observation has become a common practice [121] for the researcher to capture the interaction of participants. However, the use of screen recording as a stand-alone technique is still an under-explored area [62]. Espedal et al. [47] state that the introduction of (remote) screen recording is a reaction to the increased use of technologies within the workplace. This challenged the use of common techniques (e.g., interview and observation) to observe employee behaviour and work patterns because the activities occur in a computerised technology environment.</p> <p>Example: As stated, the use of screen recording as a stand-alone technique is still an under-explored area. Within the field of computer and information science, this technique has predominantly been used to explore mobile application usage (see e.g., [72]), the interaction between participants and technology (see e.g., [120]), or to unobtrusively collect data on how team members utilise their computers to coordinate their work (see e.g., [139]).</p>	<p>Advantages: Screen recording allows the researcher to capture the interaction (e.g., (non-)verbal cues) of the participants with the (working) environment [62], resulting in rich and empirical data [72, 139]. The screen recording produces a permanent record of interactions that the researcher can view repeatedly for later analysis [118].</p> <p>Disadvantages: Screen recording is considered to be <i>unobtrusive</i> because the researcher does not need to be physically present which is a necessity during observations. However, the technique is deemed to be highly <i>privacy-invasive</i> because all interactions including private messages are recorded [72, 139]. Data privacy is the main concern [10, 62, 139]. Alario-Hoyos et al. [10] mention that organisations must have a justifiable reason to monitor their employees. This must adhere to European privacy regulations such as the General Data Protection Regulation (GDPR) [48]. Moreover, according to Krieter [72], the downside of a permanent record of rich and empirical data is that the manual transcription and analysis of the recording(s) is time-consuming. Lastly, gaining access to a company can be challenging [66], especially if confidential information (e.g., client data) is involved.</p>	<p>Screen Recording [10, 62, 72, 121, 122, 139] (Video) Ethnography [28, 47, 50, 59, 62, 78, 114, 117, 120, 121] Fly on the Wall [77, 78, 127]</p>
<p><i>Independent Technique(s)</i></p>	<p>Continued on next page</p>		

Table A.1 – continued from previous page

Technique	Description	(Dis-)advantages	Sources
<p>Analysis of Event Logs</p>	<p>Modern information systems record detailed trails of the supported business processes including the execution of tasks [19, 82]. The event log contains information about the participant’s usage of systems and applications [1]. Process mining aims to extract this information [137] allowing the researcher to reconstruct the employee behaviour [114]. Applications of process mining include the <i>automated process discovery</i> that produces a model from an event log, <i>conformance checking</i> allows the comparison between the behaviour in a process model and event log, and <i>model enhancement</i> that extends the existing model [3].</p> <p>Example: Suriadi et al. [138] conducted a six-month case study within an insurance company in Australia regarding the application of process mining. To research the behaviours, problems, and performance of existing processes to improve the efficiency of the processes.</p>	<p>Advantages: Process mining allows researchers to exploit (historical) information on the execution of tasks. The technique can be used to identify and monitor employee behaviour [7]. This allows the identification of discrepancies or bottlenecks within the process and monitors key performance indicators (KPIs) [44]. Process mining is not dependent on the participant’s self-reports (i.e., work diary) or the observations made by the researcher.</p> <p>Disadvantages: The underlying premise is that event logs provide an accurate depiction of the actions taken in the actual processes (i.e., cases) [5, 82]. This means that the researcher depends on the quality of the data [114]. Several studies (see e.g., [21, 57, 69]) highlighted the issue of event log quality, in which tasks were not (properly) recorded or happened outside of the information system(s).</p>	<p>Analysis of Event Logs [31, 43, 84, 114, 137]</p> <p>Analysis of Tool Logs [77, 78, 127]</p> <p>Process Mining [1–3, 5, 7, 19, 21, 31, 32, 43, 44, 57, 69, 78, 82, 84, 119, 127, 137, 156, 157]</p>
Continued on next page			

Table A.1 – continued from previous page

Technique	Description	(Dis-)advantages	Sources
Archival Analysis	<p>The archival analysis is a set of existing records [114] not created by or for the (study of the) researcher [159], allowing the researcher to unobtrusively collect relevant data [160]. An archival record can be extracted from in-house documents, e-mails, or timesheets. [118, 121, 159].</p> <p>Example: Anda et al. [13] applied the use of timesheets as a form of archival analysis. For this research, a multiple-case study was conducted at four organisations to study effort estimation based on use cases.</p>	<p>Advantages: Archival records are readily available within the organisation. For example, the timesheet is a digital record of the work performed by the participant and its behaviour and activities [29]. Therefore, this technique requires no time-intensive data collection compared to the observation or questionnaire.</p> <p>Disadvantages: The archival records were not created for the (study of the) researcher. The researcher was not involved in the data collection, and therefore, depends on the guidelines and (quality) standards given by the organisation. Moreover, Mills et al. [94] argue that archival sources can only be used as evidence if the researcher is aware of the (historical) context in which the record was created [94].</p>	<p>Archival Analysis [47, 118, 127, 159, 160]</p> <p>Archival Data [47, 94, 118, 121, 152, 159, 160]</p> <p>Archival Records [94, 121, 127, 159, 160]</p> <p>Document Analysis [47, 77, 78, 94, 114, 118, 127, 152, 159]</p> <p>Secondary Data [94, 114, 118, 152]</p> <p>Analysis of Electronic Database of Worked Performed [78, 94, 127]</p> <p>Archival Source [47, 152, 159]</p> <p>Work Records [78, 127]</p>

Appendix B

Systematic Literature Review Protocol

This appendix is an extension of the Research Method (Chapter 3) about the Systematic Literature Review (SLR).

Throughout our search in digital databases, we used the following search terms:

- Data collection technique(s)
- Research method(s)
- Work pattern(s)
- Data collection method(s)
- Screen recording
- Process mining technique(s)

Each search term yielded a list of sources. To systematically keep track of the found sources, we exported the lists from the digital database (e.g., Scopus) into an Excel file. We analysed the title, abstract, keywords, and conclusion (and when needed, the entire paper) of the found papers within the long-list. We assigned each paper to one of the following values: *included*, *excluded*, or *doubtful* (according to the defined inclusion and exclusion criteria in Table 3.1). The aim of assigning the values was to create a short-list. Figure B.1 presents the information which we collected for each source [71]:

- Title
- Author(s)
- Year of publication
- Database
- Keyword(s) ¹
- Snowballing (forward or backward)/Search Engine
- Included (included, excluded or doubtful)
- Reasoning
- Topic (Label)
- Summary (or abstract)

Each source marked as “included” was thoroughly read and coded. To systematically keep track of which data collection techniques are mentioned in which source. We created an overview in Word that breaks down the aforementioned data collection techniques (including synonyms) by source, see Figure B.2. This information was used to supplement Table A.1, which presents an extensive overview of the data collection techniques.

¹When applicable, both the singular and plural form of the keywords were used.

ID	Title	Author(s)	Year	Database	Keyword(s)	Technique (e.g., Snowballing/Search Engine)	Snowballing	Document Type	Included	Reasoning	Topic (label)	Abstract
1	Studying Software Engineers: Data Collection Techniques for Software Field Studies	Janice Singer, Timothy C. Lethbridge, Susan Elliott, Sim	2005	Scopus	Data collection techniques	Search Engine		Article	Included		Data collection techniques	Software engineering is an intensively people-oriented activity, yet too little is known about how designers, maintainers, requirements analysts and all other types of software engineers perform their work. In order to improve software engineering tools and practice, it is therefore essential to conduct <i>field studies</i> , i.e. to study real practitioners as they solve real problems. To do so effectively, however, requires an understanding of the techniques most suited to each type of field study task. In this paper, we provide a taxonomy of techniques, focusing on those for data collection. The taxonomy is organized according to the degree of human intervention each requires. For each technique, we provide examples from the literature, an analysis of some of its advantages and disadvantages, and a discussion of how to use it effectively. We also briefly talk about
2	Guide to Advanced Empirical Software Engineering	Forrest Shull, Janice Singer, Dag I.K. Sjøberg	2008	Google Scholar		Forward snowballing	From 1	Book	Included		Data collection techniques, case study, research methods	
3	Research Methods in Information	Alison Jane Pickard	2013	Google Scholar	Research methods	Search Engine		Book	Included		Triangulation, case study, interview	
4	Against the flow in data collection: How data triangulation combined with a 'slow' interview technique enriches data	Jentoft N, Olsen T.S.,	2019	Scopus	Data collection techniques	Search Engine		Article	Doubtful	Uncertain if source is related to topic.	Triangulation, interview	Capturing the voices of the 'disadvantaged' in society may often be difficult to achieve. In this article, we argue that the combination of data triangulation and a 'slow' interview technique may be a good method when the target group is vulnerable people who may find the interview challenging. Data triangulation can reveal a social phenomenon's complexity by providing a fuller picture, while in-depth interviews and a 'slow' interview technique can enhance data quality. We present an alternative to what we perceive as being dominant and tacit support for use of voice recordings in interview-based studies, thus our research approach goes against the current flow in the research community. In the absence of a voice recording, the quality of the interview can be assured through cooperation between the researcher and the informant. We also discuss the importance of establishing trust when meeting the informant in order to ensure rich data.
5	Economical 600 percent CDP data collection techniques	Logel J.D., Cumerlato C.L.,	1982	Scopus	Data collection techniques	Search Engine		Conference Paper	Excluded	Not digitally accessible	N/A	No abstract available
6	Techniques of data collection and reduction in systematic field observation	Sykes R.E.,	1977	Scopus	Data collection techniques	Search Engine		Article	Excluded	Not related to the topic	N/A	Two specific instances of data collection under field observation conditions are noted: (1) the sampling of police activities and interactions, and (2) the influence of proximity and similarity on informal group formation. Coding and recording of behaviors are described and the management of data detailed.

Figure B.1: Excel file containing the information documented for the Systematic Literature Review documentation

Title	Author(s)	Techniques
Guide to Advanced Empirical Software Engineering	Forrest Shull, Janice Singer, Dag I.K. Sjøberg	Chapter 1: Lethbridge et al. – Screenshot 2 Chapter 2: participant observation, interview Chapter 3: survey (2 types: cross-sectional and longitudinal) Chapter 4: focus group (definition, strengths and weaknesses) Chapter 5: simulation – analyse process behaviour - Survey - Interview - Questionnaire - Observation - Diary - Focus Group
Research Methods in Information	Alison Jane Pickard	- Participant Observation - In-depth interview - Focus groups - Diary study
Qualitative Research Methods: A Data Collector's Field Guide	NATASHA MACK • CYNTHIA WOODSONGKATHLEEN M. MACQUEEN • GREG GUEST • EMILY NAMEY	- Observation (conflicting definitions) - Direct structured observations - Direct non-structured observations - Interviews - Questionnaires - Self-monitoring reports
Diary Methods: Capturing Life as it is Lived	Niall Bolger, Angelina Davis, Eshkol Rafaeli	
Observation: A Complex Research Method	Lynda Baker	
Social work practice: Integrating qualitative and quantitative data collection techniques	Allen-Meares P., Lane B.A.,	- Data collection techniques, therefore, allow the researcher to systematically collect information about their object of study (people, objects, phenomena) and about the setting they occur.
Data collection techniques a guide for researchers in humanities and education	Sani lawal iro	- Observation - Questionnaires - Interview
Data collection techniques	Morgan G.A., Harmon R.J.,	- Direct Observation - Questionnaires - interview

Figure B.2: Word file containing the overview of data collection techniques split per source

Appendix C

Case Study Protocol

C.1 Preamble

The purpose of the case study was to combine multiple data collection techniques with different characteristics to record employee behaviour and discover work patterns within an organisational context. The goal was to present the advantages, disadvantages, and the commonalities and differences between the data collection techniques. In addition, we focused on the implications for research and practice (i.e., organisations, employees, and EY).

*Disclaimer: Our objective was **not** to encourage or promote (illegal) employee monitoring, or recording the employee behaviour (without the consent of the employees) for the advantage of the organisation. Organisations need to have a justifiable reason to monitor their employees in such a manner [10]. In fact, European privacy regulations such as the General Data Protection Regulation (GDPR) [48] set out strict guidelines and legal standards that organisations must adhere to. In addition, (illegal) monitoring will have an impact on the company image and the relationship between the employer and other parties (e.g., employees and clients).*

The case study was conducted between May 2022 and February 2023 at Ernst & Young. We used three different data collection techniques (non-participant observation, screen recording, and archival analysis) on six employees within the Helix Team from different squads to record their behaviour while looking at the same process (the behaviour of the employees throughout the working day).

We have seen that previous studies (see e.g., [22, 24, 86, 87, 144]) have almost exclusively focused on one stand-alone data collection technique such as observation or interview. This is concerning because every technique yields unique results but also has its own shortcomings which may lead to incomplete or inaccurate results. Moreover, the main focus was set on fields where the use of a device such as a laptop is not common. This means that screen recording is still an under-explored area. As a result, there is not much information available on how to record employee behaviour by using screen recording. Therefore, the result of the case study was (1) a rich overview of the data collection techniques available to record employee behaviour, including what each technique entails, and the (dis-)advantages and commonalities, (2) an overview of similarities and differences in insights and patterns displayed by each technique, and (3) a record and rich understanding of employee behaviour within the Helix Team.

The case study results are part of the Master's thesis of the Master of Business Informatics at Utrecht University. We stored, processed, and analysed data gathered from six research participants using three data collection techniques: non-participant observation, screen recording, and archival analysis (timesheet). This data included client-specific information and information that could impact the rating of the research participant. Therefore, we had an obligation toward the research participants, EY, and its clients to protect the data. We anonymised any information collected during the non-participant observations, screen recording, and archival analysis (timesheet), that could reasonably identify the research participant or client. The data will be deleted one year after successfully concluding the Master thesis. Each research participant was required to sign the informed consent form stating that they understand and agree to the terms of the study.

C.2 Research Instruments

We used three data collection techniques: non-participant observation, screen recording, and archival analysis (timesheets). According to Lethbridge et al. [78], the techniques are categorised into three different groups (i.e., direct, indirect, and independent technique) based on the required degree of contact between the researcher and participants. Moreover, each technique has different characteristics, advantages, and disadvantages. Therefore, the combination of these techniques had a complementary effect.

We also used Patton’s [108] five dimensions of variation in observational approaches to design the approach: *role of the researcher*, *portrayal of the researcher’s role towards the participants*, *the portrayal of the purpose of the research to participants*, *duration of the research observations*, and *focus of the research observations*. We considered the impact of each dimension in order to design a suitable approach within the context and restrictions of the research.

The first dimension is concerned with the role of the researcher. We could have opted between *participant observation* (i.e., the researcher is part of the observed setting and aims to fully immerse themselves into the lives of the participants) [78, 114], and *non-participant observation* (i.e., the researcher is physically present yet does not partake in the observed setting) [22, 33]. Both techniques are considered to be obtrusive and invasive [61]. However, we opted for non-participant observation because this technique is less susceptible to the Hawthorne effect [23, 95, 118, 160]. To minimise the awareness of being observed, we conducted the observation in the natural setting of the participant (i.e., EY Amsterdam office), as suggested by McCall [85] and Sirris et al. [131].

The second and third dimensions are concerned with the disclosure of the scope of the research and the extent to which participants are informed that they are being observed [35]. This can range from *overt observations*, where participants are aware that observations are being made. To *covert observations*, where participants are unaware that observations are being made [108]. Mack [81] stated that there are no formal rules for disclosing the researcher’s involvement in research. However, Pickard [114] adds that covert observation should have no place in an academic study. The participants were fully informed about the intent of the research due to the obtrusive nature of the research (instruments). Moreover, Baker [22] argued that the (lack of) ethical guidelines is a concern. The collected data consisted of client-specific data (e.g., the name of the client and their financial statements) and information regarding the performance of the Helix Team member, which could influence their rating. Therefore, to be certain that the data is treated as confidential, we created a set of ethical considerations and guidelines which we presented to the Helix Team members.

The fourth dimension is concerned with the amount of time spent gathering observational data [108]. The observation period can be *long-term* which takes place over the course of a full year or even longer, or *short-term* which can last a day or less [35]. Pickard [114] stated that the duration of the research observations should be determined during the planning stages of the research. Therefore, set the time frame between 09.00 and 17.00 which is a standard working day within the Helix Team. The fifth dimension is concerned with the breadth of focus [108]. The observation can have a narrow focus where a single element within the observed setting is observed or a broad focus where a holistic view of the entire environment is observed. The focus of this study was on identifying the work patterns by using the categories as proposed by Mintzberg [95].

C.2.1 Non-participant Observation

The following decisions have been made to standardise the process (of non-participant observations) amongst all research participants [73]:

1. **Applications:** We used Mintzberg’s [95] standardised form, as discussed in Appendix D. For each activity, we recorded the time, category, activity, sub-activity, medium, participants, who initiated the activity, the duration, field notes, client, and the participant.
2. **Setting & Time frame:** We observed the participants for one working day between 09.00 and 17.00, except during lunch and restroom breaks. The observation took place in the EY Amsterdam office on the fourteenth floor. We used two screens during the observation. On the screen turned towards the participant, we opened a medium unrelated to our study e.g., Microsoft Outlook, Microsoft Teams, or the Internet Browser. We opened the standardised form on the screen turned away the participant. Our goal was to minimise the observer effect by pretending that we are not actively observing the activities performed by the participant.

3. **Interaction:** As stated in Chapter 2, observation can be categorised on a continuum ranging from participant observation (active) to non-participant observation (passive) [47, 121]. The Helix Team is part of the assurance service line, therefore is subjected to two peak moments (“busy season”) within the year. During the busy season, it is not uncommon to work (a minimum of) 60 hours per week. Therefore, the research participants had limited availability. We decided to use non-participant observation which does not require any active participation from the research participants. We were (physically) present in the same room (we sat next to the research participant), however, we did not partake in the observed setting.

C.2.2 Screen Recording

The following decisions have been made to standardise the process (of screen recording) amongst all research participants [68]:

1. **Applications:** We used the screen recording application *Snagit* because we had to use an application approved by EY. Snagit is a screen capture and recording application created by TechSmith¹. The advantage of the application is that it allowed the participants to record (a part of) a screen or even multiple screens. The members of the Helix Team all use two screens (i.e., their laptop and an external monitor), therefore it was a must that the application can capture and record multiple screens. The participants recorded their screen(s) and the sound (microphone and system audio). Figure C.1 illustrates the recording tool. We also used the same standardised form as during the observation [95], to preserve uniformity, and allow the comparison and analysis of the findings across the techniques [66].
2. **Setting & Time frame:** The research participants recorded their screen(s) on one working day between 09.00 and 17.00, except during lunch and restroom breaks.
3. **Modes in screen recording:** Screen capture and recording applications have several options e.g., partial screen recording, full-screen recording, and multiple screen recording. As stated, the members of the Helix Team use two screens. We wanted to avoid the loss of data that occurred outside of the selected recording area. Therefore, the participants had to capture and record both of their screens in their entirety.
4. **Interaction:** The non-participant observation and screen recording techniques were performed simultaneously. This means that we were (physically) present in the same room (i.e., we sat next to the research participant), however, we did not partake in the observed setting. At the same time, the participant would record their screen(s).
5. **Share:** We stopped the recording at 17.00. The participants shared the recording(s) via the Media Shuttle portal created by Signiant². This portal is approved by EY and allowed the research participants to safely share large files (i.e., video recordings). We were present while the participant shared the files, to avoid the risk that the participant alters and/or deletes the files. Moreover, we performed two checks. First, to see if the participant had shared all files made throughout the working day. Second, to see if the recordings were complete.

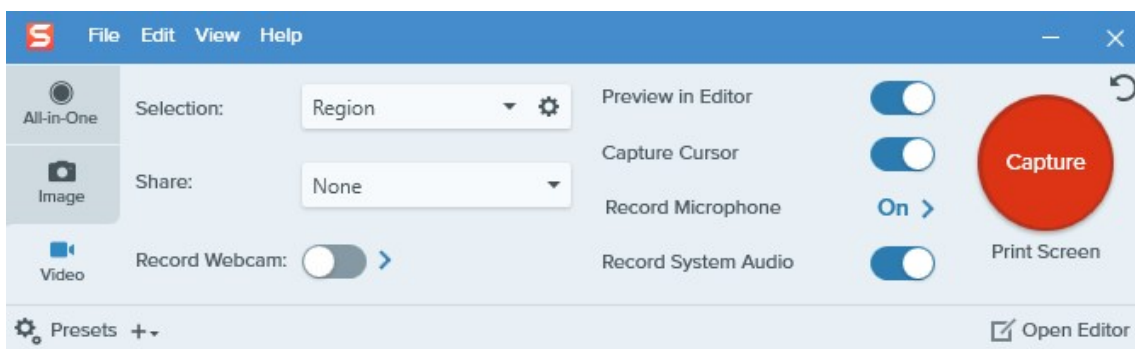


Figure C.1: Example of the screen recording tool Snagit

¹Website TechSmith: <https://www.techsmith.com/snagit-features.html>

²Website Signiant: <https://www.signiant.com/products/media-shuttle/>

C.2.3 Archival Analysis

The following decisions have been made to standardise the process (of the archival analysis) amongst all research participants:

1. **Applications:** Mercury is an integrated enterprise solution that allows the recording of time and claiming of business expenses. Mercury is used by the Helix Team to record and submit the weekly timesheets. The timesheet displays details for the entire working week. Every employee must enter the Engagement ID (e.g., billable hours charged to a client or non-billable hours such as leave or holiday), Activity ID, hours worked, and a description. Figure C.2 illustrates an example of a timesheet within Mercury. The research participants exported their timesheets into a PDF file. We converted the PDF into an Excel file and used the same standardised form as for the aforementioned data collection techniques.
2. **Setting & Time frame:** The research participants would record and submit the timesheet of the working day that we observed. However, the timesheet also included overtime activities.
3. **Share:** The participant shared their timesheet the week after the observations had taken place via the Media Shuttle portal.

Figure C.2: Example of the timesheet within Mercury

C.3 Data Collection & Analysis

Only one researcher conducted the data collection and data analysis. Therefore, to attain an appropriate sample size in terms of size (i.e., the number of participants). We made an estimation of the minimum and the maximum number of research participants. We based this estimation on our availability, the time required for the data capturing (September), and the time required for the data analysis (October). Figure C.3 illustrates the sample estimation (i.e., the total number of participants) based on our availability and the total time needed for the data capturing and analysis.

The data collection techniques such as the non-participant observation, and the screen recording are considered to be obtrusive [61, 62] and (privacy-)invasive (i.e., all interactions including private matters such as messages or emails are recorded) [72, 139]. We found it to be important to contact the Helix Team members one-on-one, to be able to convey the purpose of the research, the scientific and practical contributions, information regarding the three data collection techniques, and the set of guidelines used to protect the privacy and data of the clients and the research participants. We created a PowerPoint presentation to share this information with the participants, as illustrated in Figure C.4.

As stated, due to the sensitive nature of the data (i.e., client data) it was not possible to involve multiple researchers tasked to analyse the data. Therefore, the data analysis was solely performed by the researcher. The likelihood that we could have been influenced while analysing the collected data of each research participant per data collection technique was minimal. However, to ensure the validity of the research, we split the data analysis into two steps. First, we predefined the data analysis order of the data collection techniques. We used literature to decide which data collection technique would influence us the least. We analysed the collected data based on the following order: non-participant observation, archival analysis, and screen recording. Second, we randomised the data analysis order of the research participants per data collection technique. We aimed to minimise the risk of creating a connection between the data from each participant between each technique. We used a random number generator tool³ to generate a unique and randomised order for the archival analysis and screen recording. Figure C.5 illustrates the random number generator used to generate a randomised order for the data analysis of the archival sources and screen recording.

³<https://www.randomlists.com/random-numbers?dup=false&qty=6&max=6&min=1>

C.4 Ethical Considerations & Guidelines

Andrews & Pradhan [15] stated that research should be built on trust between the researcher and research participants. However, it is still necessary to take explicit measures and ethical considerations to prevent any issues which may harm the involved parties (the researcher, research participants, Ernst & Young, or Utrecht University) [129]. The case study results will be published as part of the Master (Business Informatics) thesis at Utrecht University. We have listed the following ethical considerations & guidelines to be certain that all data is treated as confidential [124].

1. Before the start of the research, the research participants had to sign the informed consent form which is shared as part of the formal invitation. The informed consent form includes all information and guidelines which has been discussed during the presentation, illustrated in Figure C.4. The quality of the research depends on our integrity in terms of being open, honest, and truthful about the set-up of the research and how the data will be collected and used. Section C.5 presents the informed consent form that has been shared with the research participants.
2. Participation is completely voluntary. Research participants have the right to withdraw from participation. If you wish to withdraw from participation, please inform the researcher via mobile phone or email. Any information or data gathered up to this moment will be deleted and will not be used during this study.
3. Data collection techniques such as non-participant observation and screen recording are considered to be obtrusive and (privacy-)invasive. Even the archival analysis (timesheets) includes sensitive information. We have an obligation toward the research participants, EY, and its clients to protect the data.
 - We will only use applications that have been approved by EY and can be requested via the EY AppStore (e.g., the Media Shuttle and Snagit).
 - The data will be carefully stored on our device and a backup will be made on OneDrive (this is the encrypted OneDrive of Ernst & Young).
 - The data will be deleted one year after the research has been concluded.
 - We are the only ones who will have access to the collected data of each research participant per data collection technique (each research participant has access to their own screen recording and timesheet).
 - Any information collected during the non-participant observations, screen recording, and archival analysis (timesheet), that could reasonably identify the research participant or client will be removed or anonymised before the data and files are shared with others (e.g., the manager of the Helix Team or the supervisors of Utrecht University).
 - The performance and rating of the research participants will not be impacted. The manager of the Helix Team will not have access to the complete data (i.e., (squad) names) nor will these results be used to rate the Helix Team members. The research findings will only be used for scientific purposes.

C.5 Informed Consent Form

Figure C.6 illustrates the informed consent form including the ethical considerations & guidelines. We based this consent form on information provided by Utrecht University [145]. The research participants were required to sign the informed consent form prior to the start of the study.

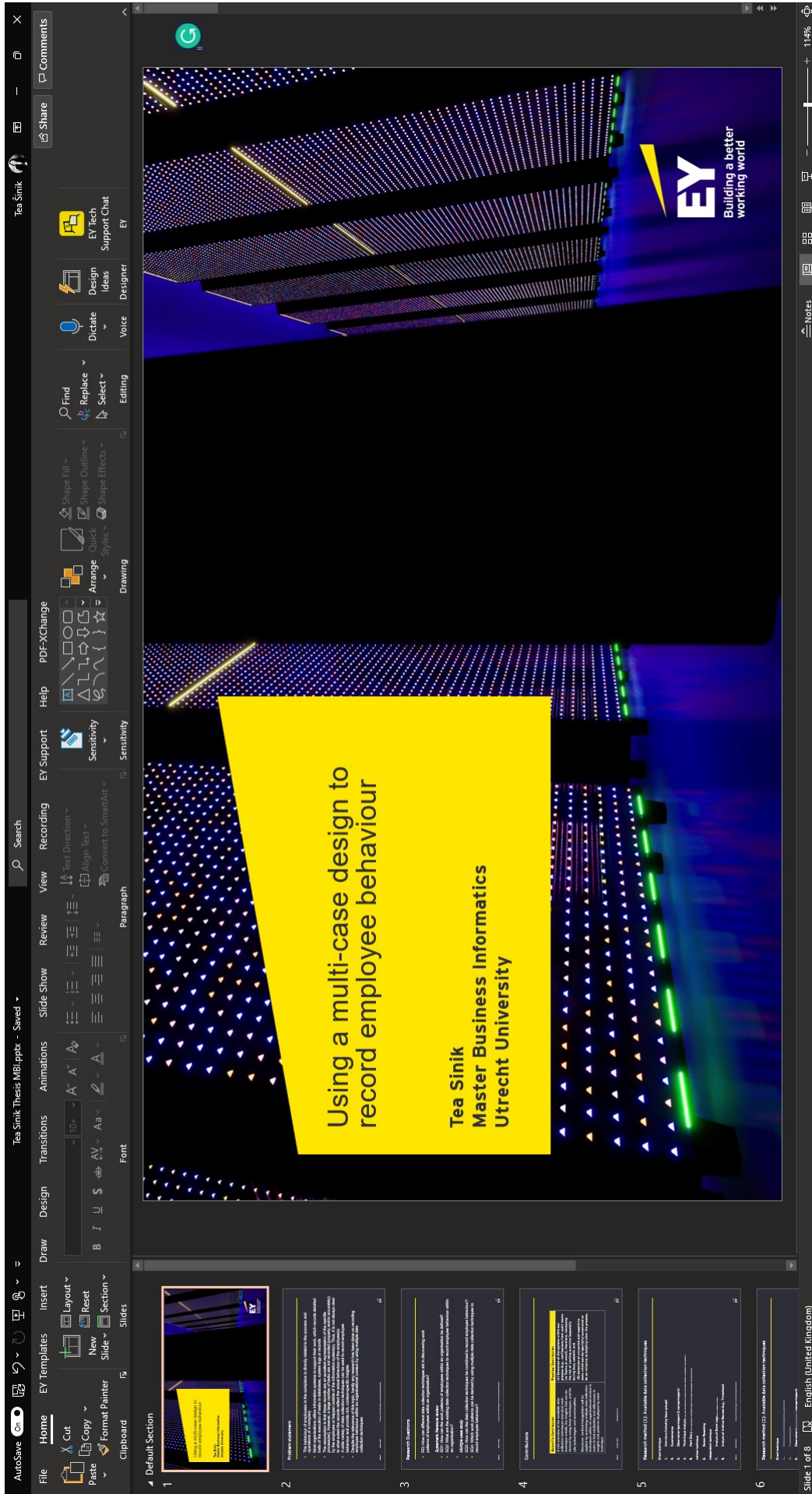


Figure C.4: The PowerPoint presentation used to contact each Helix Team member regarding the purpose of the research

Random Numbers:

1 5 2 3 3 2 4 4 5 6 6 1

REERUN OPTIONS

REERUN OPTIONS SHARE

Reserveer nu
Hertz

Ga relaxed op reis.
Plan vooruit.

Random Number Generator

Its the core of all randomness. Pick a number or generate a whole sequence of numbers within a minimum and maximum value (inclusive) while including or suppress duplicates. Your device is used to quickly generate these numbers, completely random and unique to you every time.

Change the quantity to one if you just want it to pick a number.

You can switch the presentation to [roll some dice](#) instead. Or change gears completely with the [phone number generator](#) or [random letter generator](#).

Edit Settings

Min Number

Max Number

Quantity

Duplicates

REERUN

Figure C.5: The random number generator used to generate a randomised order for the data analysis



**Utrecht
University**



Informed Consent Form

Research Project: Thesis Master Business Informatics
 Title: Using a multi-case design to record employee behaviour
 Institutions: Ernst & Young Accountants
 Utrecht University

Dear research participant,

You have been invited to participate in the study in which we will use three different data collection techniques (non-participant observation, screen recording and analysis of archival sources) on research participants within the Helix Team (from different squads) to record their behaviour whilst looking at the same process (the behaviour of the employees throughout the working day). This research is conducted by Tea Šinik as part of the Master's thesis of the Master Business Informatics at Utrecht University.

The study will take place in the EY Amsterdam office. We will use three different data collection techniques (non-participant observation, screen recording and analysis of archival sources) to record your behaviour throughout one working day between 09.00 and 17.00. We will observe you for the entire working day except during lunch and restroom breaks. At the same time, you will record both of your screens (laptop and external monitor) using the application Snagit. At the end of the working week, you will make an export from your timesheet in Mercury.

All information collected throughout the study will be considered confidential. The researcher will be the only person who has access to the complete data. Any information collected during the non-participant observations, screen recording and analysis of archival sources (timesheet), that could reasonably identify the research participant or client will be removed or anonymised before it can be shared with others. Even Tijmen Baas (manager of the Helix Team) will not have access to the complete data. Therefore, this will not influence your performance or rating. The research findings will only be used for scientific purposes!

Participation is completely voluntary, research participants have the right to withdraw from participation. If you wish to withdraw from participation, please inform the researcher via mobile phone or email. Any information or data gathered up to this moment will be deleted and will not be used during this study. Please see the next page for the list of ethical considerations and guidelines that apply to this study.

Thank you for your participation!

In case of any questions or concerns - prior, during or after your participation - please contact the researcher via mobile phone or email:

Researcher: Tea Šinik
 Mobile phone: 0031621252271
 Email: tea.sinik@nl.ey.com

The signature below means that the research participant has given their consent (voluntary and freely) and understands and accepts the requirements of this research.

X

 Research Participant:

Figure C.6: Informed Consent Form

Ethical considerations and guidelines

The case study results will be published as part of the Master (Business Informatics) thesis at Utrecht University. We have listed the following ethical considerations & guidelines to be certain that all data is treated as confidential.

1. Before the start of the research, the research participants must sign the informed consent form which is shared as part of the formal invitation. The informed consent form includes all information and guidelines which has been discussed during the presentation. The quality of the research depends on our integrity in terms of being open, honest and truthful about the setup of the research and how the data will be collected and used.
2. Participation is completely voluntary, research participants have the right to withdraw from participation. If you wish to withdraw from participation, please inform the researcher via mobile phone or email. Any information or data gathered up to this moment will be deleted and will not be used during this study.
3. Data collection techniques such as non-participant observation and screen recording are considered to be obtrusive and (privacy-)invasive. Even the analysis of archival sources (timesheets) includes sensitive information. We have an obligation toward the research participants, EY and its clients to protect the data:
 - ▶ We will only use applications that have been approved by EY and can be requested via the EY AppStore (e.g., the Media Shuttle and Snagit).
 - ▶ The data will be carefully stored on our device and a backup will be made on OneDrive (this is the encrypted OneDrive of Ernst & Young).
 - ▶ The data will be deleted one year after the research has been concluded.
 - ▶ We are the only ones who will have access to the collected data of each research participant per data collection technique (each research participant has access to the screen recording of their working day seeing as they record it onto their device).
 - ▶ Any information collected during the non-participant observations, screen recording and analysis of archival sources (timesheet), that could reasonably identify the research participant or client will be removed or changed/anonymised before the data and files are shared with others (e.g., the manager of the Helix Team or the supervisors of Utrecht University).
 - ▶ The performance and rating of the research participants will not be impacted. The manager of the Helix Team will not have access to the complete data (i.e., (squad) names) nor will these results be used to rate the Helix Team members. The research findings will only be used for scientific purposes

Appendix D

Classification Scheme

This Appendix presents the classification scheme used to categorise and analyse the observed activities of the employees. The classification scheme was predetermined in line with the study by Mintzberg [95]. This was recommended by Tengblad [141] because the classification scheme has been tested in a multitude of studies and allows for comparison.

Mintzberg’s [95] research was published over 50 years ago and was solely focused on the use of one data collection technique (structured observation) to study managerial work. We noted that this research was not focused on a field where the use of devices (such as a laptop) is prevalent. Therefore, the reliance on technology throughout the working day was not investigated [66]. In an effort to capture workplace technology use and establish a basis for future comparisons, we introduced new categories (e.g., “technical problem solving”) while maintaining and tailoring the categories to the context of this research. The categories that were not relevant to our research (e.g., “location”) were excluded.

D.1 Standardised Form

To preserve uniformity and allow the comparison and analysis of the findings across data collection techniques [66]. We used a single standardised form for all data collection techniques. Table D.1 presents the standardised form used to record employee behaviour. We used one form for each data collection technique per participant. In total, we applied the standardised form 18 times.

During the data collection of the observation and screen recording, we recorded a new activity “*at any point at which there was a change in activity, the basic participants and/or the medium*” [96]. Within the standardised form we included information such as the category, the purpose and duration of each activity, and the involved participants [95, 141]. In contrast to Mintzberg’s [95] classification scheme, our standardised form did not include “location” (e.g., driving to the client’s location). This is not applicable to the context of our research because the employees are rarely on-site due to COVID-19.

We included the following three classifications: *sub-activity*, *client*, and *participant*. First, we included *sub-activity* and *client* because we wanted to gain a deeper understanding of the purpose of the activities and how much time is spent on client-related work. This was also a request made by the EY stakeholder. Second, we included *participant* as a unique identifier to generate the analysis as a whole and for each participant.

Table D.1: The standardised form used to record employee behaviour

Classifications	Description	Sources
Time Section D.2	This classification is concerned with the time that the activity was observed. We used a time frame of 09.00 to 17.00 (standard working day) with a time interval of one minute.	[66, 95, 141]
Category Section D.3	This classification is concerned with the category used to describe the distribution of time throughout the working day.	[95, 141]

Continued on next page

Table D.1 – continued from previous page

Classifications	Description	Sources
Activity Section D.4	This classification is concerned with the (purpose of the) activities that were observed throughout the working day.	[16, 66, 95, 141]
Sub-Activity Section D.5	This classification is concerned with the sub-activities that were observed throughout the working day.	[16, 66, 75, 95]
Medium Section D.6	This classification is concerned with the medium(s) that were used to perform the (sub-)activity. For example, the employees used Alteryx to clean and transform the data, and Microsoft Teams to communicate with other participants.	[66]
Participants Section D.7	This classification is concerned with the (multiple) member(s) the employee had contact with.	[16, 96]
Initiated Section D.8	This classification indicates who initiated the activity.	[16, 95]
Duration Section D.9	This classification indicates the duration of each observed activity (shown in minutes).	[16, 96]
Field Notes Section D.10	This classification consists of the researcher’s notes. The purpose of the notes is to aid in recalling this particular activity after the observation has taken place (during transcription and coding).	[16, 66, 95, 141]
Client Section D.11	This classification consists of the clients for whom the employee carried out activities.	
Participant Section D.12	This classification indicates the (research) participant and is used as a unique identifier.	

Figure D.1 visualises a snippet of the filled-in standardised form used during the non-participant observation of participant P1. Due to its length, only a small portion of the field notes are displayed in this snippet.

Time	Category	Activity	Sub-Activity	Medium	Participants	Initiated	Duration	Field Notes	Client	Participant
10:12	Desk Work	Giving Information	Message Outgoing	Microsoft Teams	Helix Team - One	Employee	0:03	The employee opened Microsoft Teams because they received a chat message from	1A	P1
10:13										
10:14										
10:15	Desk Work	Giving Information	Message Outgoing	Microsoft Teams	Audit Team - Two	Other party	0:02	The employee received chat messages via Microsoft Teams from two different audit team participants. The first audit team	1A	P1
10:16										
10:17	Desk Work	Requests & Solicitations	Mail Incoming	Microsoft Outlook	Combination - Helix, Audit Team and Client	Other party	0:04	The employee opened Microsoft Outlook because they received an email from the client with the squad lead and two audit team members in the CC. The email is	1B	P1
10:18										
10:19										
10:20										
10:21	Desk Work	Requests & Solicitations	Message Outgoing	Microsoft Teams	Audit Team - One	Employee	0:02	The employee opened Microsoft Teams, they initiated a chat message to the audit	1B	P1
10:22										
10:23	Desk Work	Organisational Work	Validation memo (572)	Microsoft Excel	Independent	Employee	0:02	The employee opened Microsoft Excel and is looking at the validation memo (572) of	1D	P1
10:24										
10:25	Desk Work	Organisational Work	Data Transformation & Validation	Alteryx	Independent	Employee	0:03	The employee opened Alteryx, they are looking at the workflow of client ID. They	1D	P1
10:26										
10:27										
10:28	Desk Work	Scheduling & Administration	Planning	JennEY	Independent	Employee	0:01	The employee opened JennEY, they opened the "My tickets" list. The employee	1E	P1
10:29	Desk Work	Organisational Work	Data Transformation & Validation	Microsoft Excel	Independent	Employee	0:04	The employee opened a validation memo	1F	P1

Figure D.1: Snippet (filled-in) standardised form used during Observation

D.2 Time

This study recognised the challenge of “short-lived activities” lasting a few seconds. We noted that within the time interval of one minute, the employee performed multiple activities, some short-lived and some longer. In response, and in line with Mintzberg’s [95] study, we noted all activities in the “field notes”. The most dominant activity was categorised and the decision has been made to round the activity to the minute. As stated by Mintzberg [96] and Johnson [66], this reduces the burden on the researcher because there is no need to carry a stopwatch or check each second. This means that the research can fully focus on the research participant.

D.3 Category

Table D.2 presents the categories used to describe the distribution of time throughout the working day. They were initially based on the study by Mintzberg [95]. In light of the context of the study, we have made the following changes to the categories:

Personal: The focus of the study was set on work-related activities, and on the individual as an employee and not on the individual person. Personal activities (e.g., use of the personal phone) that occurred throughout the working day were categorised but not explicitly noted in the field notes.

Desk Work: There is a major difference in the definition for “desk work”. In Mintzberg’s study “desk work” was defined as “*time the manager spends at his desk, processing mail, scheduling activities, and communicating with his secretary*” [95, p. 92]. In 1970, devices such as the laptop did not play a role in the workplace, therefore Mintzberg only distinguished desk work from in-person encounters (meetings) and telephone calls. In this research, desk work included all activities that were performed using technology (e.g., data extraction or transformation, planning tickets in JennEY, reviewing documents, or sending messages via Microsoft Teams). We acknowledge that *telephone* conversations take place via Microsoft Teams. However, we consider this to be a separate form of communication (compared to sending a chat message via Microsoft Teams) and therefore should be a separate category [66].

Telephone: The employees do not use their mobile work phones. Every *telephone* conversation takes place via Microsoft Teams. Therefore, this category refers to all incoming and outgoing “telephone” calls via Microsoft Teams.

Meeting: In line with early studies, we included “meeting” as a categorisation. However, in contrast to Mintzberg’s study, no record was kept of whether the meetings were scheduled or unscheduled. Mintzberg defined unscheduled meetings if they were “*arranged hastily, as when someone just ‘drops in’*” [95, p. 92]. To augment this data, a variety of additional sources were collected such as printed calendars, meeting agendas, or the participants were asked to verbally indicate whether the meeting was scheduled or not [66]. Our research did not include this data, because we opted for the non-participant observation (passive) which prevented us from involving the participant.

Interactions with researcher: As proposed by Patton [108], during the non-participant observation the researcher is (physically) present yet does not partake in the observed setting. Therefore, any interaction between the participant and researcher must be noted. As stated in Appendix C, prior to the study we informed the research participants what each data collection technique entails.

Unable to categorise: We used three data collection techniques with different characteristics. Based on the literature review, we concluded that each technique has its (dis-)advantages. For example, screen recording allows capturing of the interaction between the participant and their device [62, 139]. To be able to categorise the activity, there must be a screen recording and audio recording (i.e., system audio and microphone).

Transportation: The Helix Team employees are rarely in the office or on-site at the client because of COVID-19. Therefore, we have excluded the following category from our study “transportation”.

Table D.2: The categories used to describe the distribution of work time

Category	Description	Source
Personal	This category is concerned with all personal (private) activities observed and noted but not explicitly included in the research e.g., private email and private messages.	[66]
Tours	This category refers to activities that do not take place behind the desk e.g., informal talks in the hallway, (coffee) breaks.	[16]
Desk Work	This category refers to the time that the employee spends at their desk e.g., processing emails, extracting data, and planning activities.	[66]
Telephone	This category refers to incoming or outgoing “telephone” calls via Microsoft Teams.	[66]
Meeting	This category refers to meetings that happen offline (i.e., without the use of Microsoft Teams).	[66]
Interactions with researcher	This category is concerned with all interactions between the employee and researcher.	[16]
Unable to categorise	We do not have sufficient information to categorise this activity.	[16]

D.4 Purpose of the Activities

This classification described the (purpose of the) activities, as presented in Table D.3. The activities are based on the literature, and the requests made by the EY stakeholder

Table D.3: The purpose of the activities

Purpose of activities	Description	Source
Organisational Work	Any activity concerned with the work (e.g., extraction, transformation, and validation of a dataset), projects, or other initiatives that require the specialised skills and knowledge of the employee.	[95]
Scheduling & Administration	Any activity concerned with the administration of the planning or scheduling of meetings and data extractions, compiling to-do lists, and updating the status of tickets in JennEY.	[16]
Requests & Solicitations	Times when the employee approaches or is approached by participants (e.g., the audit team, CAD, Helix Team, or others) who make requests for status updates or actions.	[16]
Tours	Times when the employee leaves their desk to greet someone in the hall, grab a coffee, and take a break.	[95]
Receiving Information	Any activity concerned with receiving information e.g., receiving information updates, attending briefings and presentations, and responding to inquiries.	[95]
Giving Information	Any activity concerned with giving information e.g., giving information updates, giving advice, and speeches and presentations held by the manager.	[95]
Socialising	An activity concerned with (informal) socialising.	[16]
Technical Problem Solving	Any activity concerned with technical problem-solving. The employee relies heavily on technology, which sometimes may cause technical issues.	[66, 75]
Set-up time	Any activity concerned with the time spent preparing for the day (e.g., turning on the laptop).	[16]
Personal	Times when the employee performed private activities (e.g., use of the personal mobile phone).	[16]
Breaks	Times when the employee took a break (e.g., lunch, coffee break, or smoke break).	[16]

Continued on next page

Table D.3 – continued from previous page

Purpose of activities	Description	Source
Interactions with researcher	Times when the employee interacts with the researcher.	[16]
Unable to categorise	We do not have sufficient information to categorise this activity.	[16]

D.5 Sub-Activity

Similar to Mintzberg’s [95] *contact record*, we included the sub-activities to gain a deeper understanding of the purpose of the activities. Our EY stakeholder requested that we distinguish between incoming and outgoing telephone calls, messages, and emails to show whether and who is disrupting the process. Table D.4 presents the sub-activities used to examine the activities in greater detail. The sub-activities are created based on the literature (see e.g., [16, 66, 75, 95]) and the researcher’s (practical) experience and knowledge.

Table D.4: The sub-activities used to obtain deeper insights into the purpose of the activities

Sub-Activity	Description
Private Matter - Not Observed	Times when the employee performed private sub-activities (e.g., use of the personal mobile phone)
Break - Not Observed	Times when the employee took a break (e.g., lunch, coffee break, or smoke break).
Socialising	A sub-activity concerned with (informal) socialising.
Set-up time	A sub-activity concerned with the time spent preparing for the day.
Planning	A sub-activity concerned with planning-related tasks (e.g., planning their work in JemEY).
Coordination	A sub-activity concerned with coordination (e.g., involving colleagues in an activity in an organised way).
Data Extraction	A sub-activity concerned with the extraction of client data.
Data Transformation & Validation	A sub-activity concerned with the transformation and validation of client data.
Data Loading	A sub-activity concerned with the data loading of the dataset(s) in the Helix Platform.
Data Delivery	A sub-activity concerned with the delivery of the dataset(s).
Extraction memo (571)	A sub-activity concerned with the creation of the extraction memo (571).
Review extraction memo (571)	A sub-activity concerned with the review of the extraction memo (571).
Validation memo (572)	A sub-activity concerned with the creation of the validation memo (572).
Review validation memo (572)	A sub-activity concerned with the review of the validation memo (572).
Meeting	A sub-activity concerned with a meeting that occurs offline (i.e., without the use of Microsoft Teams).
Call Incoming	A sub-activity concerned with incoming <i>telephone</i> calls (via Microsoft Teams).
Call Outgoing	A sub-activity concerned with outgoing <i>telephone</i> calls (via Microsoft Teams).
Mail Incoming	A sub-activity concerned with incoming emails (via Microsoft Outlook).
Mail Outgoing	A sub-activity concerned with outgoing emails (via Microsoft Outlook).
Message Incoming	A sub-activity concerned with incoming (chat) messages (via Microsoft Teams).

Continued on next page

Table D.4 – continued from previous page

Sub-Activity	Description
Message Outgoing	A sub-activity concerned with outgoing (chat) messages (via Microsoft Teams).
Technical Issues	A sub-activity concerned with technical problem-solving.
Learning	A sub-activity concerned with learning (e.g., online courses).
Unable to categorise	We do not have sufficient information to categorise this sub-activity.

D.6 Medium

The employees rely heavily on technology in the workplace because the majority of their activities are performed on a laptop. They use a variety of mediums depending on the activity and personal preference. Our EY stakeholder requested that we provide a detailed list of all mediums that are used throughout the working day. This is presented in Table D.5. The mediums that are developed within and owned by EY are marked as “EY”.

Table D.5: An overview of the used mediums throughout the working day

Medium	Description	Source
Microsoft Teams	Chat-based collaboration platform developed by Microsoft.	[93]
Microsoft Outlook	Email and task management application developed by Microsoft.	[92]
Redacted	EY tool used for the management of tickets (e.g., planning).	EY
Redacted	Redacted	EY
Redacted	Redacted	EY
Microsoft Excel	Spreadsheet program developed by Microsoft. Mostly used for transforming and validating the data and validation memos (572).	[89]
Microsoft Word	Word processing software developed by Microsoft. Mostly used for the extraction memos (571).	[90]
Alteryx	Alteryx is a data analytics solutions desktop tool used to transform and validate data.	[12]
Media Shuttle	Media Shuttle is an encrypted data-sharing tool approved by EY.	[128]
Redacted	Redacted	EY
Redacted	Redacted	EY
Redacted	Redacted	EY
Jupyter	The Jupyter Notebook is a platform used to create and run Python scripts.	[115]
Internet browser	An internet browser is used to search for web pages on the device.	[101]
Microsoft Planner	Web-based application to create plans, assign tasks, and to-do lists developed by Microsoft.	[88]

Continued on next page

Table D.5 – continued from previous page

Medium	Description	Source
Redacted	Redacted	EY
7-Zip	A program to archive and (de-)compress files.	[110]
JSON	JavaScript Object Notation (JSON) is a “ <i>lightweight data-interchange text format used for storing and transporting data</i> ”.	[149]
Notepad	A text editor application developed by Microsoft.	[91]
Redacted	Redacted.	EY
Python	“Python is an interpreted, interactive, and object-oriented programming language”. Python can be used to solve many different kinds of issues (e.g., cleaning and transformation of data).	[142]
No Medium	No medium was used.	
Unable to categorise	We do not have sufficient information to categorise this activity.	[16]

D.7 Participants

Table D.6 presents an overview with whom the employees interacted (“participants”) throughout the working day. Mintzberg [95] developed categories such as director, peer, client, supplier, or the government. These categories are not applicable to this study because the employees are not in contact with these types of participants. To make the categories relevant for this study, we examined the participants the employee interacts with e.g., the Audit Team, Helix Team, and Central Audit Delivery (CAD). This study recognised that employees frequently communicate with several participants at once. As such, the categories were expanded by including the number of participants *one, two, and three or more* [66]. Moreover, this study recognised that the employee can be in contact with multiple participants from various departments. As such, the categories were further expanded by including a combination of multiple participants e.g., “Combination - Helix, Audit Team, and Client”. Lastly, this study recognised that the employee can be working independently, meaning without the presence of a participant. Or the employee can indirectly be in contact with the board of directors of EY, Young EY, and Office Assistants (i.e., the employees receive a company-wide email from the aforementioned participants). As such, the categories were further expanded by including “Independent” and “Other”.

Our EY stakeholder requested that we provide an overview with whom the employees interacted (“participants”) throughout the working day.

Table D.6: Overview of contact persons with whom the employee interacted throughout the working day

Participants	Description
Helix Team - One Helix Team - Two Helix Team - Three or more	This entails one or multiple participants of the Helix Team.
GDS - One GDS - Two GDS - Three or more	This entails one or multiple participants of Global Delivery Services (GDS). They are the offshore team of the Helix Team.
Audit Team - One Audit Team - Two Audit Team - Three or more	This entails one or multiple participants of the audit team.
Client - One Client - Two Client - Three or more	This entails one or multiple participants of the client.

Continued on next page

Table D.6 – continued from previous page

Participants	Description
CAD - One CAD - Two CAD - Three or more	This entails one of the multiple participants of the Central Audit Delivery (CAD). They offer support services to the audit team e.g., journal entry testing and preparation of documentation.
Combination - Helix and Audit Team Combination - Helix and GDS Combination - Helix Team and CAD Combination - Audit Team and GDS Combination - Helix, Audit Team, and Client Combination - Helix, Audit Team, and GDS	This entails a combination of multiple participants from various departments such as from the Helix Team, Audit team, client, and GDS.
Other	This entails company-wide emails from participants such as the board of directors of EY, Young EY, Global Chairman, EY NL Leadership, and BSC Office Assistants
Independent	This entails only the employee who is working independently (i.e., no other participants are involved).
Unable to categorise	We do not have sufficient information to categorise this activity.

D.8 Initiation of the Activities

We recorded who initiated a certain activity [16, 95]. This is presented in Table D.7. Our EY stakeholder requested that we include this in the standardised form.

Table D.7: Initiation of the activities

Initiated by	Description
Employee	The activity is initiated by the employee.
Other party	The activity is initiated by the other party (i.e., not the employee).
Both	The activity was initiated by both parties.
Unable to categorise	We do not have sufficient information to categorise this activity.

D.9 Duration

Mintzberg [95] measured the duration of each activity with a stopwatch and rounded to the closest 30-second interval. The duration is recorded to the nearest tenth of an hour [16, 96].

Within our standardised form, we used a time frame of 09.00 to 17.00 (i.e., standard working day) with a time interval of one minute. We used the clock in the toolbar of our laptop rather than a stopwatch. We rounded the activity to the minute and categorised the most dominant activity within that time interval [66]. We calculated the duration in Excel. We set the number format to “Time” for the columns “Time” and “Duration”. We then subtracted the time recorded during the new activity, from the time recorded during the old activity.

D.10 Field Notes

The *Field Notes* consist of the descriptions given by the employees (in their timesheet) and the researcher’s notes (made during the observation and screen recording).

During the data collection of the observation, we made notes about the purpose of the (sub-)activity, who initiated the activity, the medium that was used, the participant(s) with whom the employee

interacted, and if applicable the client (name). Due to the fast-paced environment, we used keywords or unstructured notes to record this information. The purpose of the field notes was to aid in recalling this particular activity after the observation has taken place. For the screen recording, we meticulously recorded each observed activity.

During the data analysis, we anonymised all elements that could reasonably identify the research participant, the involved participants (colleagues), or the client. For the research participant and colleagues with whom they interacted, we used the pronouns “they/them” instead of “he/him” or “she/her”. We assigned a unique code to the research participants and clients, as shown in the following sections.

D.11 Client

Our EY stakeholder requested that we include the classification *Client* to record how much time is spent on (non-)client-related work. For each participant, we created a unique (anonymised) identifier for the client. Table D.8 presents an example of how the clients per participant are coded. This table is meant to serve as an example. These are not actual clients of EY (that were observed throughout the working day). The categorisation “Non-Client” is used for activities that are not related to client-specific work (e.g., technical issues solving).

Table D.8: An example of the unique (anonymised) clients

Participant	Client Name	Client Code
P1	Utrecht University	1A
P1	Eindhoven University of Technology	1B
P2	TIAS School for Business and Society	2A
P2	Erasmus University Rotterdam	2B
P2	Non-Client work-related activities (e.g., technical issues)	Non-Client

D.12 Participant

We included the classification *Participant* as a unique identifier to generate the analysis as a whole and for each participant. We used the participant ID as presented in Table 3.3 (Section 3.5).

Appendix E

Data Analysis

We used the following two mediums for our data analysis: Alteryx and Excel. We will describe the analysis performed with each medium.

E.1 Alteryx

We used Alteryx to decrease the number of manual actions (e.g., merging files by copy-pasting) to reduce errors and improve the reliability of the data analysis. This section presents the Alteryx workflows along with a description.

E.1.1 Check Empty Fields

Figure E.1 presents the Alteryx workflow that we used to check if the data files contain empty fields. We first **imported** the data files. The use of only one filter on one specific column was identified as a threat because we did not know that this column did not contain any empty fields. To minimise the threat of filtering out any data and/or using data with empty fields, we used two separate filters on the columns *Category* and *Sub-Activity*. We used the **filter** query to identify records where the columns *Category and Sub-Activity is not null* (i.e., does not contain empty fields). We then used the **summarise** tool to count the empty fields. We performed this check on both filter queries. We found two empty fields in the column “Participant” in the observation file of participant two, which we manually filled in.

E.1.2 Union Data Streams

The risk of using Excel to import data from multiple data streams is that we can unintentionally delete data or change the (number) format or position of the columns. Figure E.2 presents the Alteryx workflow that we used to combine the six data streams per data collection technique. We first **imported** the six data files per participant per data collection technique. We used the **select** tool to check if the data types are identical in all incoming data streams. The **union** tool combined the six data streams based on the “column position”. Lastly, we **exported** the data into separate Excel files for each data collection technique.

E.1.3 Comparison Results Observation & Screen Recording

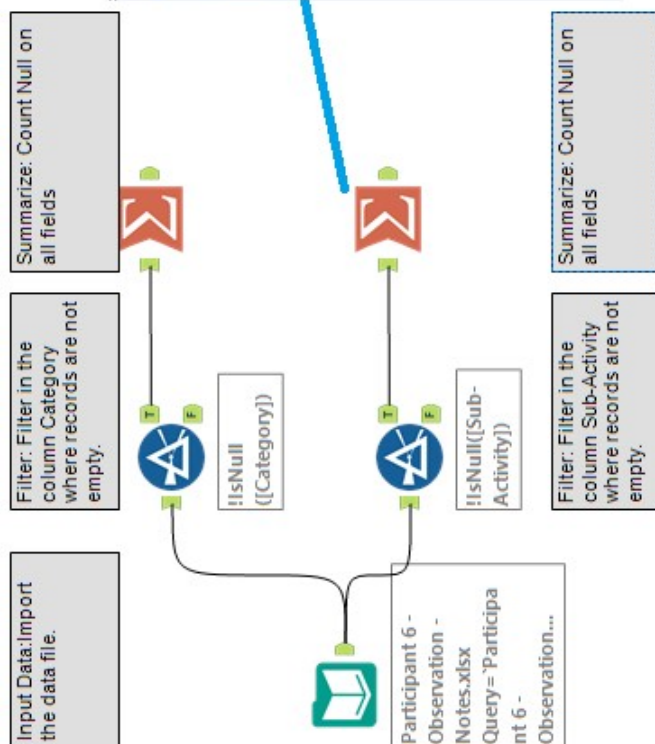
Once we were certain that all files are complete, we wanted to compare the results of the observation and screen recording per participant. We can only compare the files of these two techniques because of their similarity (i.e., for both techniques we observed and categorised the activities per minute from 09.00 to 17.00). Figure E.3 presents the Alteryx workflow that we used to (inner) join the data files for the observation and screen recording per participant. We first **imported** the participant observation and screen recording data files. We then used the **select** tool to check if the data types are identical in both data files. The **join** tool combined the two data streams based on record position. We opted for the **inner join** with a **J anchor** to have an output that contains records that are joined from the *L input (observation)* and the *R input (screen recording)*. Lastly, we **exported** the data into an Excel file which we used for further analysis. Figure E.4 shows a snippet of the Excel sheet that contains the results of the observation (left in green) and the results of the screen

recording (right in orange). We added a column named "Differences" (middle in blue) to note the differences between the results of the two data collection techniques.

E.1.4 Richness of the Field Notes

We performed this analysis to determine whether there is a discrepancy between the richness and completeness of the column *Field Notes* between the data collection techniques. We aimed to see if the differences between the categorisations that can be observed per data collection technique and/or the inability to observe a certain categorisation can be linked to the minimal provided information in the column *Field Notes*.

Figure E.5 presents the Alteryx workflow that we used to count the minimum and the maximum number of words in the "Field Notes" per data collection technique. We first **imported** a data file that contained the records of all participants of a data collection technique. This file was created with the union tool, as stated in the previous Section E.1.2. We used the **filter** query to identify records where the column *Field Notes* is not null. If the column is empty then no activity has taken place (i.e., we only wrote down field notes for each observed activity). We then used the **select** tool to change the data type for the column *Field Notes* from *string* to *double* to be able to use the min/max function in the **summarise** tool.



Actions:

Field	Action	Output Field Name
Activity	Count Null	CountNull_Activity
Sub-Activity	Count Null	CountNull_Sub-Activity
Medium	Count Null	CountNull_Medium
Participants	Count Null	CountNull_Participants
Initiated	Count Null	CountNull_Initiated
Duration	Count Null	CountNull_Duration
Field notes	Count Null	CountNull_Field notes
Client	Count Null	CountNull_Client
Participant	Count Null	CountNull_Participant
Data Collection ...	Count Null	CountNull_Data Collection Technique

Figure E.1: Alteryx workflow to check on empty fields

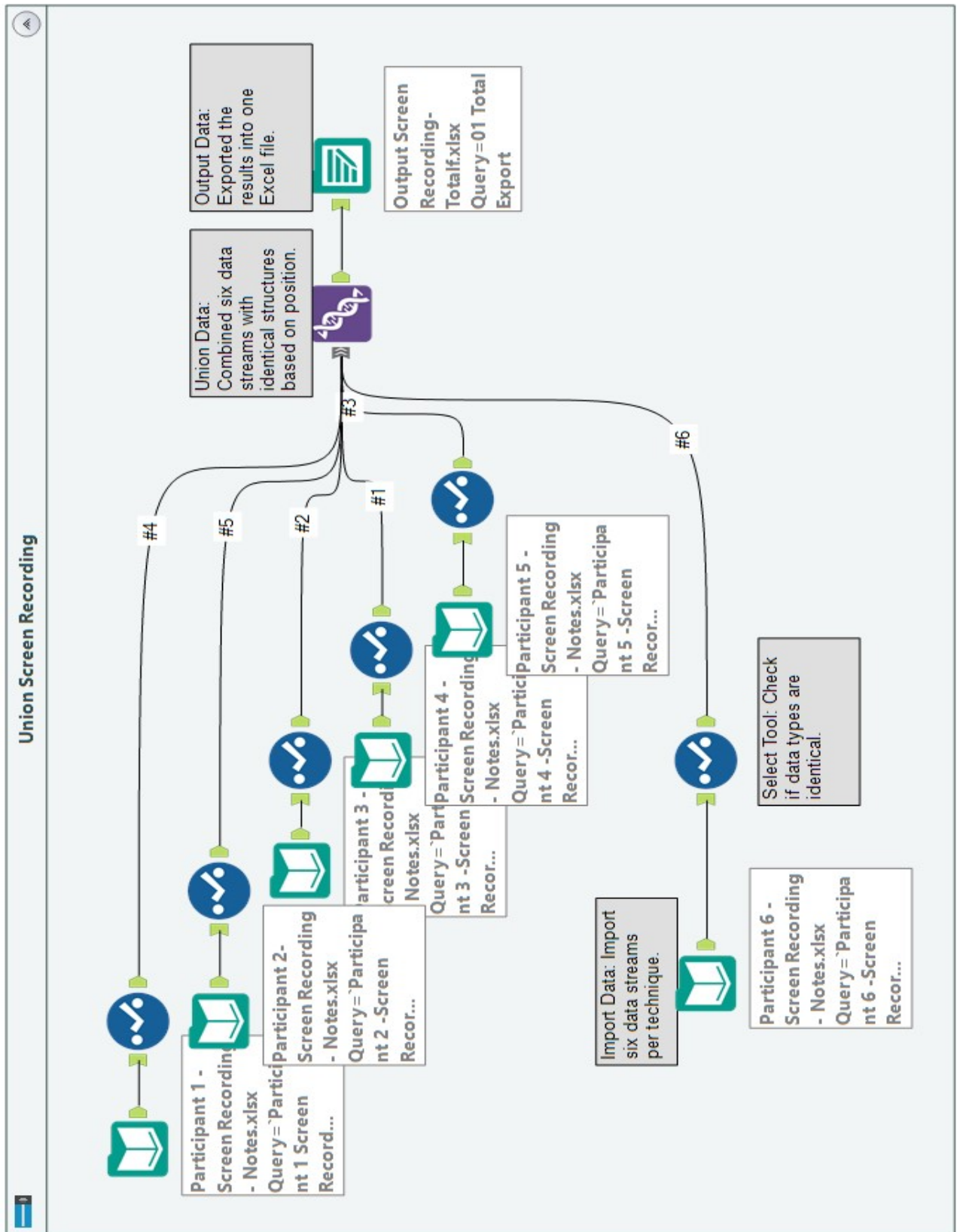


Figure E.2: Alteryx workflow to union the data files

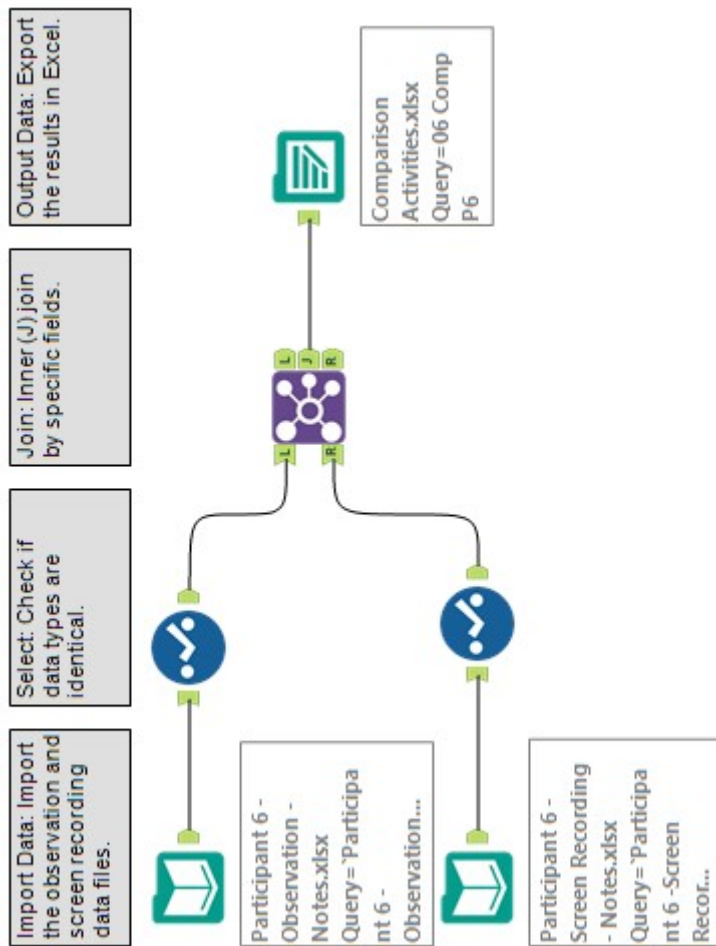


Figure E.3: Alteryx workflow to (inner) join the observation and screen recording results

Category (Screen Recording)	Activity (Screen Recording)	Sub-Activity (Screen Recording)	Medium (Screen Recording)	Participants (Screen Recording)	Initiated (Screen Recording)	Duration (Screen Recording)	Field notes (Screen Recording)	Client (Screen Recording)	Participant (Screen Recording)	Data Collection Technique (Screen Recording)
Telephone	Scheduling & Administration	Outgoing Call	Microsoft Teams	Helix Team - Three or more	Employee	00:30:00	The employee opened Microsoft Teams to join the weekly team meeting. The entire Helix Team has joined this call. The participants are discussing the planning and any issues they are facing related to the planning. The employee has informed the team about their planning and a client that has caused quite some		P1	Screen Recording
09:00:00										

Figure E.4: Snippet Excel sheet of the comparison between the observation (green) and screen recording (orange) results

Category (Observation)	Activity (Observation)	Sub-Activity (Observation)	Medium (Observation)	Participants (Observation)	Initiated (Observation)	Duration (Observation)	Field Notes (Observation)	Client	Part Data	Differences
Telephone	Scheduling & Administration	Outgoing Call	Microsoft Teams	Helix Team - Three or more	Employee		The employee joined the weekly team meeting via Microsoft Teams. The purpose of this meeting is to socialise, talk about new developments and innovation projects, share information and talk about the planning and issues related to work. During this call, there was a presentation from a fellow colleague regarding the new WIKI page. The employee did not present during the meeting but was part of the group that prepared the presentation.			
09:00:00						00:30:00			P1	Observation

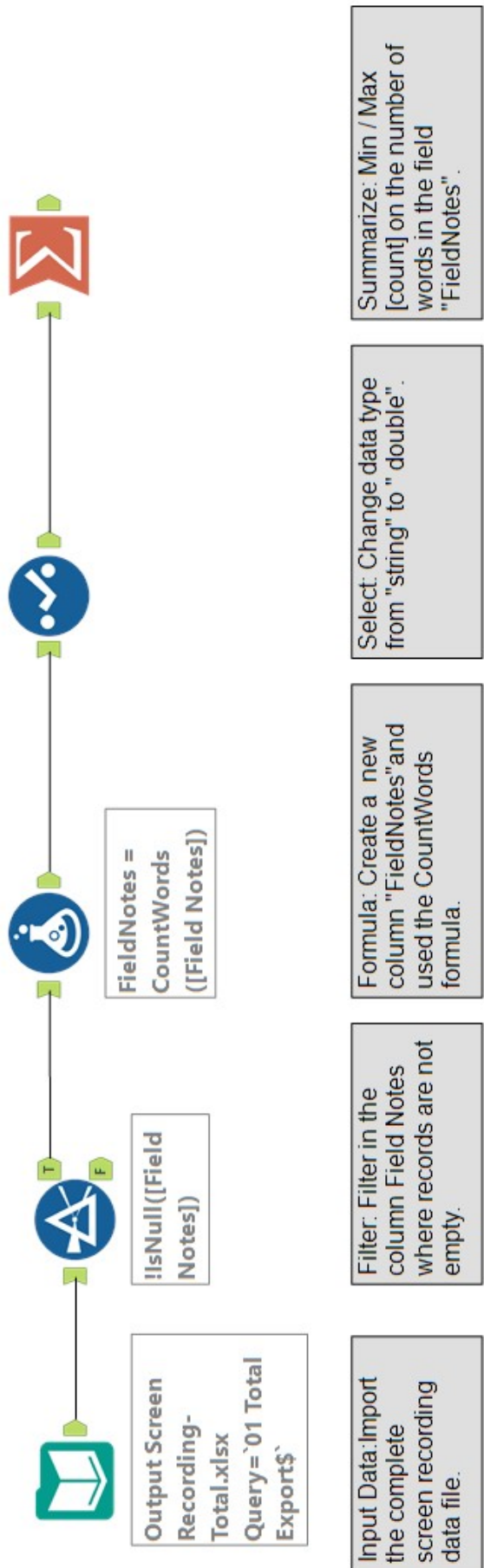


Figure E.5: Alteryx workflow to examine the richness of the column Field Notes

E.2 Microsoft Excel

We used Microsoft Excel for the remaining analysis. To summarise the large amounts of data, we generated two pivot tables per classification: total per classification, and total per classification per participant. We repeated this for each classification per data collection technique. We then combined the overviews of each data collection technique in one Excel sheet to compare the results.

E.2.1 Total per Classification

First, we generated a pivot table per classification as mentioned in our standardised form. Figure E.6 shows the results of the total per classification (Category) recorded using the non-participant observation. Within the pivot table (left) we selected Category as rows, the sum of Duration, and the count of the Category as the values. This generated the overview of the results of the total per Category (right).

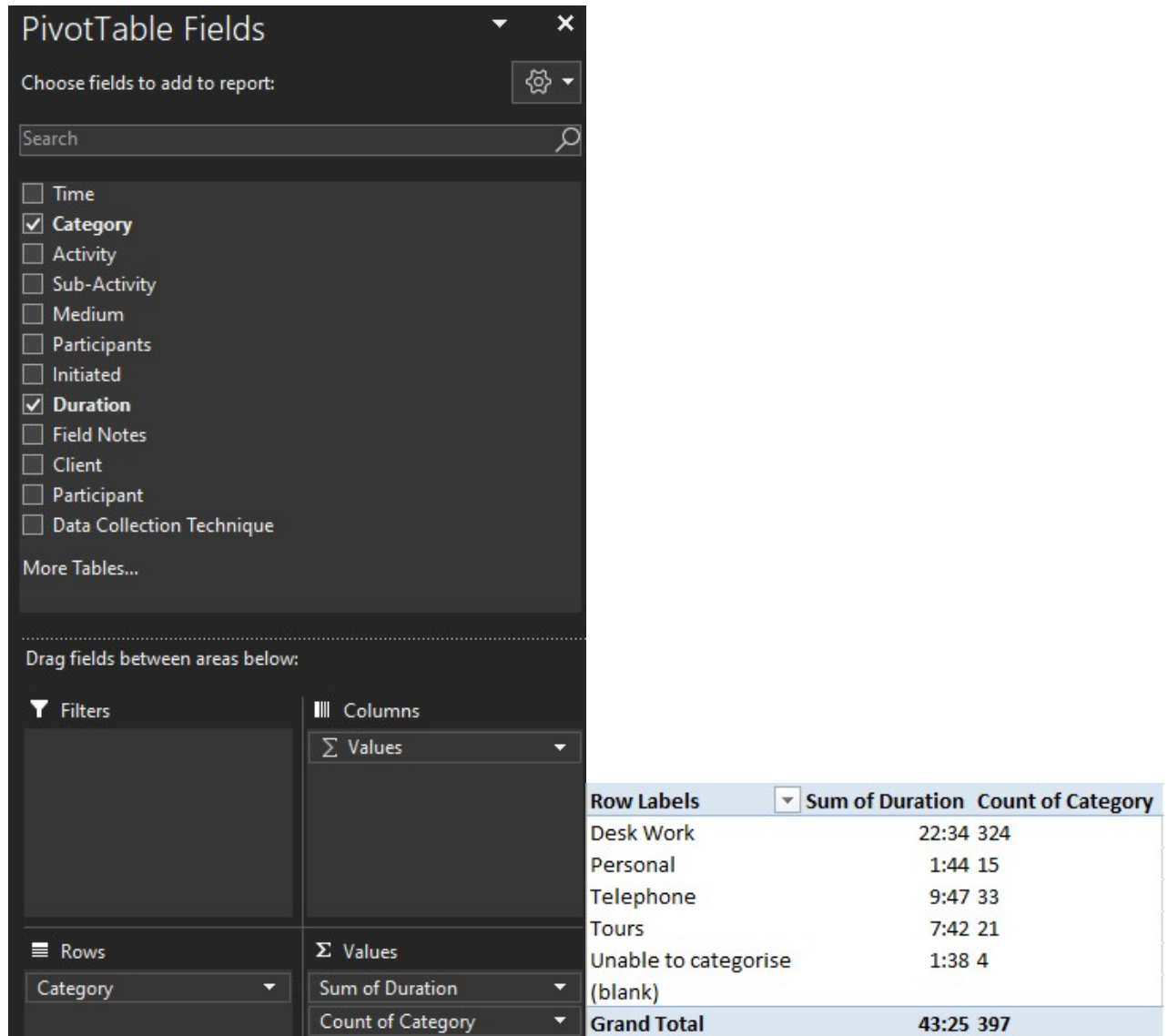


Figure E.6: Pivot Table (results) of the total per Classification (Category)

E.2.2 Total per Classification per Participant

Second, we generated a pivot table per classification per participant. Figure E.7 shows the results of the total per classification (Category) per Participant recorded using the non-participant observation. Within the pivot table (left) we selected the participant and the Category as rows and the sum of

Duration and the count of the Category as the values. This generated the overview of the results of the total per Category per participant (right).

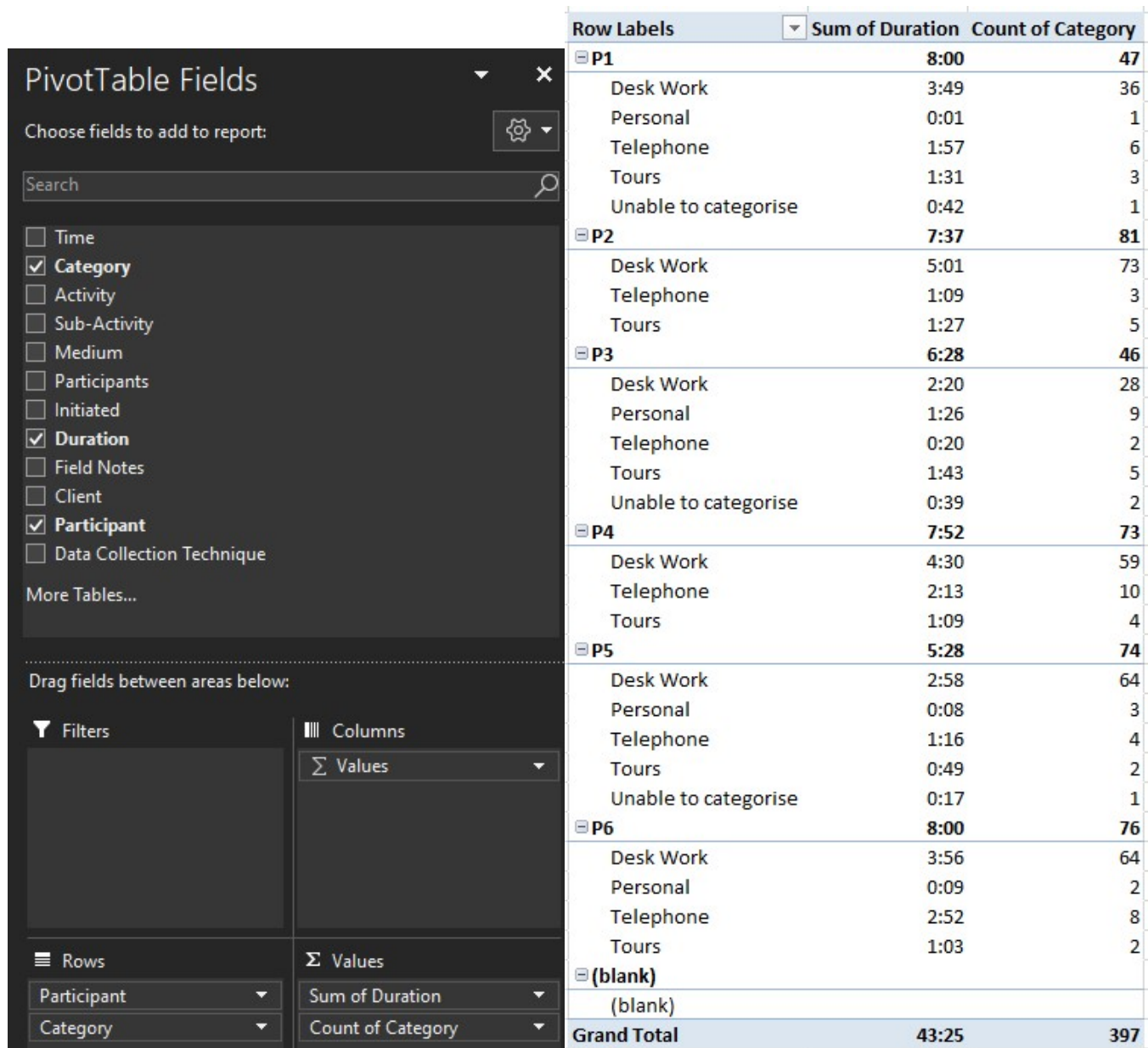


Figure E.7: Pivot Table (results) of the total per Classification (Category) per Participant