Ecological Momentary Assessment Versus Cross-Sectional Survey

An Empirical Study in the Office Environment

- Final Report -

Student information

Student:	Sara Woltjes
Student ID:	930628972130
Study:	Management, Economics and Consumer studies
Specialization:	Management studies
Profile:	Facility Management
Course code:	MST-80433

Supervision information

University:	Wageningen University & Research Centre
First supervisor:	Dr. H.B. Kok
Second supervisor:	Dr. G. Hagelaar
Company supervisor:	M. Meetz

Table of Content

Abstract

Organizations are constantly looking for ways to enhance performance. Currently, most empirical studies on this subject have adopted cross-sectional methods. However, recent literature claims that cross-sectional methods are often biased, and that other methods that collect real-time data are preferred. As there is no empirical evidence that confirms these claims, the current study aims to identify the differences between Ecological Momentary Assessment (EMA) and cross-sectional (CS) methods within the office environment. A case study was performed amongst the employees of a Dutch insurance company. Respondents were divided into two groups: the first group filled out the cross-sectional survey (n=80), and the second group answered the same questions based on the EMA principles using a mobile application (workplace n=95; restroom n=65; restaurant n=28). Overall, it can be concluded that there are indeed significant differences between EMA and CS, in terms of survey response, survey scores and employee-environment relationships. This study confirms that EMA indeed captures within-person fluctuations, which cross-sectional methods are unable to capture. However, results also showed that there are more factors that need to be considered in order to conduct a successful EMA study. Based on the results of this study, it is still thought that EMA is more reliable than CS, when it comes to collecting data within the office environment. Although currently no further claims can be made, this study has opened up the door for further examination of the differences between these two data collection methods.

Key words: Cross-sectional survey, Ecological Momentary Assessment (EMA), Physical Work Environment, Productivity, Well-Being, Job Satisfaction.

Executive summary

Organizations are constantly looking for ways to enhance performance. From the perspective of facilities management, the most effective way to do this is by optimizing the physical work environment in terms of user performance. Currently, most empirical studies on this topic have adopted cross-sectional methods. However, recent literature claims that cross-sectional methods are often biased, and that other methods that collect real-time data are preferred. This method is also referred to as Ecological Momentary Assessment (EMA). Theoretically, EMA counts three major advantages, when compared to a cross-sectional survey. EMA takes into account within-person variability, minimizes recall bias, and maximizes ecological validity.

However, there is no empirical evidence that confirms these claims. Therefore, this report describes and evaluates the differences between two data collection methods: Ecological Momentary Assessment (EMA) and a cross-sectional survey. In order to compare these two methods, this study looks into the relationship between the performance levels of office employees and their working environment. Employee performance is hereby operationalized as: employee productivity, job satisfaction, and employee well-being (EWB). Three touchpoints of the office environment are considered: the restaurant, the restroom, and the employees' personal workspace. These specific touchpoints were selected, as they all serve a different purpose within the office environment.

An exploratory approach has been adopted, as the differences between a cross-sectional survey and EMA are being explored. The results are meant to open up the door for further examination of the differences between these two data collection methods. During the empirical part of this study, a case study approach was taken. Because of this study's exploratory nature, it was decided to focus on one specific organization (i.e. single-case approach). The respondents were selected based on volunteer sampling. The first group filled-out the cross-sectional survey, and the second group answered the same survey questions based on the EMA approach.

A crucial difference between the two methods is that EMA respondents only answer questions about their current location (i.e. one touchpoint), whereas CS respondents evaluate their entire work environment (i.e. multiple touchpoints). Thus, in order to collect enough observations for each touchpoint, EMA requires more respondents than does CS. During this study, the CS survey resulted in 80 observations for each touchpoint (n=80), whereas EMA fluctuated highly across touchpoints (workplace n=95; restroom n=65; restaurant n=28). Although the number of observations for the workplace is sufficient, the number of observations for the restaurant certainly is not. Based on these results it can be concluded that the number of EMA respondents depends at least partially on the amount of time a respondent spends at this particular touchpoint.

After the data collection, the cross-sectional data was analysed using multiple linear regression (MLR) combined with a principal component analysis (PCA). The EMA data was analyzed using Linear Mixed Models (LMM), because this method is particularly suitable for datasets that include repeated-measures. The results of the analysis showed that both datasets identified a significant relationship between the aspects of the workspace and employee complacency. In addition, the cross-sectional data also identified a significant relationship between the restroom and job complacency, and between the interior and acoustics of the workspace and the ability to finish work. These additional relationships may have only been identified within the CS data because the sample size of this group was significantly higher compared to EMA. Thus, these same relationships may be underlying the EMA data, but may currently not be significant due to the limited number of observations.

Based on the current study, it can be concluded that there are indeed differences between EMA and cross-sectional methods, in terms of survey response, survey scores and employee-environment relationships. In addition, this study confirms that EMA captures within-person fluctuations, which cross-sectional methods are unable to capture. However, despite the advantages of EMA that are mentioned in various literature studies, there are more factors that need to be considered in order to conduct a successful EMA study. For example, most people are not familiar with EMA, and do not seem to understand why they have to answer the same survey questions multiple times. This confusion leads to a lack of motivation to fill out the surveys. Consequently, the sample size becomes too small to reveal any meaningful relationships. Therefore, additional attention needs to be paid to the human factor of data collection, to ensure sustainable user engagement. Overall, based on the results of this study, it is still thought that EMA is more reliable than CS, when it comes to collecting data. However, this study does not provide sufficient proof to make hard claims.

Key concepts and definitions

Customer journey: The customer journey is a set of interactions between a customer and a product, a company, or part of its organization (e.g. physical workplace) that spans a longer period of time and consists of multiple components and multiple touchpoints (Lemon & Verhoef, 2016)

Ecological Momentary Assessment (EMA): A collective of a range of research methods which capture momentary behaviours and psychological states in the participant's natural environment, and track those behaviours and states over time (Beal & Weiss, 2003).

Employee productivity: Output per employee hour, quality considered (Sutermeister, 1976).

Employee Well-Being (EWB): The part of an employee's overall well-being, that they perceive to be determined primarily by their work and can be influenced by workplace interventions (Juniper et al., 2009).

Facility management: An integrated approach to operating, maintaining, improving and adapting the buildings and infrastructure of an organisation in order to create an environment that strongly supports the primary objectives of that organisation (Barrett & Baldry, 2003).

Job satisfaction: Pleasurable or positive emotional state, resulting from the appraisal of one's job experience (Locke, 1976).

Open (plan) office: A workspace whose perimeter boundaries do not go to the ceiling (Brill & Weidemann, 2001).

Private (cellular) office: A workspace that has four walls to the ceiling and a door (Brill & Weidemann, 2001).

Subjective well-being (SWB): The overall assessment of an individual's life quality. It is subjective because the assessment is based on one's personal standards, rather than on others' standards (Diener, 1984).

Touchpoint: Touchpoints are all the areas where a person metaphorically "touches" the organization during a during a company visit. Each impression has an impact on the overall evaluation of the customer journey.

Workplace: The entire physical work environment (Brill & Weidemann, 2001).

Workspace: A smaller space within a workplace where an employee sits (mostly) when in the office (Brill & Weidemann, 2001).

1.0 Introduction

In a world with growing international competition, organizations are always looking for new ways to enhance their performance. From the perspective of facilities management, the most effective way to do this is by optimizing the physical work environment in terms of user performance. As the average adult spends about a quarter to a third of their waking life at work (Beal & Weiss, 2003), it is not surprising that the physical workplace has a direct impact on both psychological- and behavioural outcomes. For example, by providing employees with a comfortable and satisfying work environment, employee productivity could increase up to fifteen percent (Leyten et al., 2003). Besides productivity, the physical workplace also has a significant impact on job satisfaction, which is widely recognized as a key indicator of organizational performance. According to Locke (1976), job satisfaction can be defined as a "pleasurable or positive emotional state, resulting from the appraisal of one's job experience". Furthermore, adjustments in the physical work environment can also help to improve Employee Well-Being (EWB). According to existing literature, EWB can be defined as "that part of an employee's overall well-being that they perceive to be determined primarily by work, and can be influenced by workplace interventions" (Juniper et al., 2009). In practice, improved EWB generally leads to reduced levels of e.g. stress, depression and anxiety.

Besides scientific studies, this topic is also receiving increasing attention from organizations. As a result, organizations are investing increasing time and money into the design of their workplace. However, in order to design better buildings, it is essential to know how the various workplace elements are perceived by its users, and how they impact user performance in terms of e.g. productivity, job satisfaction and EWB. A user-centred approach which enables organizations to obtain this information is the "customer journey", and concerns all the interactions between a customer (i.e. office employee) and an organization (Lemon & Verhoef, 2016). These interactions can occur at different points in time, and at different locations, also known as "touchpoints". Examples of touchpoints within the office environment are e.g. the restaurant, restrooms, and an employee's personal workspace. Currently, the most common way to obtain information on employees' views and experiences of each touchpoint, is by conducting a cross-sectional survey. This is generally a self-report questionnaire, which is designed to collect one-time information from a representative sample (cross-section) of the population of interest. Unfortunately, despite its many advantages (e.g. widely accepted, relatively inexpensive, and easy to complete), recent studies claim that cross-sectional methods are unreliable, due to a number of reasons (Fisher & To, 2012).

First, by asking participants to summarize their behaviour and experiences, researchers are missing out on meaningful within-person variability (Beal & Weiss, 2003). For example, when an employee is asked to rate his job satisfaction, his answer could vary from moment to moment, or from day to day, depending on his environmental situation (e.g. mood, workload, work-family conflicts) at the time of retrieval. In case of a cross-sectional survey, such fluctuations would either stay unnoticed, or they would be regarded a measurement error (Engelen et al., 2017). *Secondly*, the outcomes of a cross-sectional survey are often biased because of memory errors (Shiffman et al., 2008). This is because people's memories regarding their past emotions, experiences, and behaviours are in fact poor reflections of the actual history of those emotions, experiences, and behaviours. Much of what people recall is actually a reconstruction, pieced together from fragmentary inputs. Therefore, any delay between an experience and its report also means a loss of information, and reports of current affect and experiences are considered to be more accurate than memory-based reports (Fisher & To, 2012).

In order to capture people's real-time (or very recent) experiences, researchers have introduced an alternative type of data collection. This method is generally referred to as an Ecological Momentary Assessment (EMA), Experience Sampling Method (ESM), or diary studies. EMA is a type of data collection which enables researchers to capture people's real-time experiences and behaviours within their natural environment, and track them over time (Shiffman et al., 2008; Beal & Weiss, 2003). Theoretically, EMA counts three major advantages, as compared to a cross-sectional survey (Shiffman et al., 2008). *First*, EMA minimizes the reliance on retrospective recall. Rather than having to recall or summarize a feeling from the past, respondents of EMA are asked about their current (or very recent) feelings (Freedman et al., 2006). *Secondly*, instead of collecting one-time information, EMA provides a picture of how people's experiences and behaviour varies over time and across situations (Shiffman et al., 2008). This is important because people's moods and perceptions change throughout the day. *Lastly*, EMA data is collected in a real-world environment, as the respondents go about their lives. As a result, this method allows generalization to the respondents' real lives (i.e. ecological validity). In conclusion, EMA takes into account within-person variability, minimizes recall bias, and maximizes ecological validity.

Theoretically, these are all valid arguments, which have therefore caused researchers to believe that EMA is more reliable than a cross-sectional survey (Shiffman et al., 2008; Beal & Weiss, 2003; Alliger & Williams, 1993). However, besides theoretical arguments, there is no empirical evidence which can confirm these claims (Engelen et al., 2017). Until today, solely one study has empirically compared EMA to a cross-sectional survey. This study was conducted by Van den Brink et al. (2001), who looked into the occurrence of recall bias in paediatric headache amongst children. They found that when compared to EMA, adopting a cross-sectional survey leads to the occurrence of recall errors, and to more negative pain complaints (Van den Brink et al., 2001). Although this study did find significant differences in output between EMA and a cross-sectional survey, similar studies have not been conducted within an organisational context. The *objective* of this study is therefore to investigate the differences in output between EMA and a cross-sectional survey, when evaluating the office environment. In order to accomplish the stated objective, the following research question needs to be answered:

"How does an ecological momentary assessment differ from a cross-sectional survey, when evaluating the physical work environment in terms of its contribution to employee performance?"

Sub-questions that will support finding the answer to the main research question are:

Literature study: How can employee performance be conceptualized and measured?
How can the user's experience of the physical work environment best be captured?
What is the nature of ecological momentary assessment versus a cross-sectional survey, in an organizational context?

Empirical study: How do employees evaluate the physical work environment, using EMA and a cross-sectional survey?

By answering the main research question, the present study contributes to the existing body of literature, as it will investigate potential differences in output between EMA and a cross-sectional survey. Doing so will help to find out whether EMA is indeed more suitable for data collection within an organizational context. Moreover, it will lead to an improved understanding of the reliability of cross-sectional methods. This would not only change the decision-making for future studies, but it may also question the reliability of many existing studies who have adopted cross-sectional methods.

With respect to the remainder of this thesis, the structure is as followed. Chapter two, three and four present the literature review on the physical work environment, employee performance, and both data collection methods (i.e. EMA and cross-sectional survey). Based on the literature, the conceptual framework is developed. This framework is presented in chapter five, followed by a detailed explanation of how the data is collected and analysed in chapter six. In chapter seven, the results of the performed data analyses are presented. Finally, chapter eight and nine consist of the conclusion and discussion. Here, the study's findings and limitations are discussed, and directions for future research are given. In addition, this final chapter provides the answer to the main research question, and presents the case organization with a number of practical recommendations.

2.0 Literature – Physical work environment

During this study, three touchpoints of the customer journey are considered: the restaurant, the restroom, and the employee's personal workspace. These specific touchpoints were selected, as they all serve a different purpose within the office environment. *Firstly*, essential to every office building, is the workplace. During office hours, this is where employees spend most of their time. Therefore, it is crucial that the workspace supports performance, and enables employees to concentrate on their work. *Secondly*, besides a workspace for concentration, the office environment should also provide a place for relaxation, where people can go during their break to recharge. Therefore, the office restaurant was selected as the second touchpoint. *Thirdly*, besides the obvious touchpoints where employees spend most of their time, there are also a number of other touchpoints an office simply cannot do without. A typical example is the restroom, which is therefore selected as the third touchpoint. In the remainder of this chapter, the characteristics of each of these three touchpoints (i.e. workspace, restaurant, and restroom) are further discussed, along with possible ways of operationalizing them.

2.1 Restaurant

When visiting a restaurant, a person's meal experience is not just determined by the quality of the food. There are also other components that need to be considered. For example, according to Campbell-Smith (1967), a meal experience is also determined by: the level of service, cleanliness/hygiene, value for money, and ambiance. During a more recent study, Andersson & Mossberg (2004) stated that a meal experience depends upon: food, service, fine cuisine, restaurant interior, good company, and other customers. Similar aspects were identified by Namkung & Jang (2008), who concluded that food is most important, followed by the physical environment, and staff service. What these studies have in common, is that they aim to identify the determinants of a meal experience that takes place within the context of a commercial restaurant. However, these meals differ from the meals provided by organizations, as they are meals for pleasure, rather than meals for necessity (Williams, 2009). The meals that are provided by organizations are generally referred to as "institutional meals". They are desirable, but secondary to the organization's primary goal.

In order to capture the complexity and experience of institutional meals, Gustafsson et al. (2006) developed the Five Aspect Meal Model (FAMM; figure 1). This model consists of three main elements (i.e. the product, the meeting, and the room), which are built upon two "backstage" variables (i.e. management control system and the atmosphere). Gustafsson et al. (2006) state that these five aspects together shape a person's meal experience. The meal takes place in a certain environment (*room*), where customers interact with each other and the staff (*meeting*), and where a variety of food and drinks (*products*) are served. By including the management control system, the model acknowledges that management has an impact on the surrounding atmosphere and on the overall meal experience. However, because these variables cannot be directly observed by guests, the management control system and the atmosphere are considered to be backstage variables.

Product

The first element of FAMM consists of all the characteristics of the food and beverages that are served (i.e. food quality), and has a crucial impact on a person's meal experience (Namkung & Jang, 2008). Not only in restaurants, but also in workplace settings, there is a growing acceptance that the served food has a significant impact on important employee outcomes. Meals at work do not only create welcoming breaks in work routines, but they also constitute important ingredients and provide employees with new energy. As a result, these meals improve both efficiency as well as the quality of employees' work efforts. However, even though consensus is reached on the importance of food quality, researchers have not yet agreed on the individual attributes that constitute food quality. Consequently, food quality has been measured using a number of different attributes. For example, Kivela et al. (2000) studied food quality by means of the tastiness of food, menu variety, and nutrition. In another study, food quality was determined based on food presentation, serving size, menu design, and the variety of food (Raajpoot, 2002). Yet another study selected presentation, variety, healthy options, taste, freshness and temperature to assess food quality (Namkung & Jang, 2008). Thus, consensus is reached on the importance of food quality, but not on identifying its individual attributes.

Meeting
Apart from the food served, a person's evaluation also depends on the social interactions that take place during the meal experience. These interactions could take place between customers, or between customers and staff. Gustafsson et al. (2006) generally refer to these interactions as the "meeting", but it could also be referred to as "service quality". Service quality is the service customers receive in a service setting, and includes variables such as: employee appearance, number of staff, the quality of service, staff friendliness, and service speed (Johns & Pine, 2002). Moreover, service quality also encompasses a customer's perceived waiting time. Here, a wait is defined as the duration from the moment the customer is ready for the service encounter to the moment the encounter actually starts (Johns & Pine, 2002). This is important, as waiting is regarded a waste of time, and it can cause negative emotions, thereby lowering overall satisfaction with the provided service.

Room
The "room", as reported by Gustafsson et al. (2006), is the environment in which the meal is consumed, and entails many factors including: colour schemes, furniture, design, and layout. Together, these factors have a significant impact on a person's meal experience. This is confirmed by scientific studies, as it was demonstrated that identical foods are evaluated differently when consumed in different settings (Edwards et al., 2003). Once people are inside a restaurant, they often spend hours observing (both consciously and subconsciously) the interior of the facility, which in turn impacts their evaluation of the place. An important aspect of the overall evaluation of the room is the level of seating comfort, which depends on the design and condition of the furniture, as well as their arrangement (Sulek & Hensley, 2004). The table placement and the spaces between the seats define personal space, and can regulate both privacy and interaction. For example, if seats are too close to each other, this can cause people to feel crowded. Besides the architectural design, people are also affected by the cleanliness of the walls, floors and furniture. For example, whether or not floors, tables and carpets are clean, or whether garbage cans are overflowing, will affect the perceived quality of a service facility (Barber & Scarcelli, 2009).

Atmosphere
Another aspect considered by the FAMM, is the restaurant's atmosphere. The concept of atmospherics was first introduced by Kotler (1973) in a marketing context, where he defined it as "the effort to design buying environments, to produce specific emotional effects that enhance the chance of purchase". Within the context of a restaurant, atmosphere is generally defined as the entirety of the meal experience, which goes beyond the physical location, and is created by all aspects of the three main elements combined (i.e. room, product, and meeting). In other words, the atmospherics are a group of intangible factors that turn a meal into more than the mere sum of its parts. This includes factors such as the acoustics, temperature, odours, and lighting (Sulek & Hensley, 2004). Atmospheric restaurants can be described as places where guests feel comfortable and at ease (Gustafsson et al., 2006).

To summarize, various tools exist for evaluating the meal experience. What these tools have in common is that besides food quality, they also recognize the importance of both the social- and the physical environment. With respect to institutional meals, the most common model is the Five Aspect Meal Model (FAMM). This model was developed by Gustafsson and colleagues (2006) and identifies three main different elements that impact the overall meal experience: the product, the meeting, and the room. In turn, these three elements are built upon two "backstage" variables: the management control system and the atmosphere (figure 1). Together, these five aspects shape a person's meal experience.

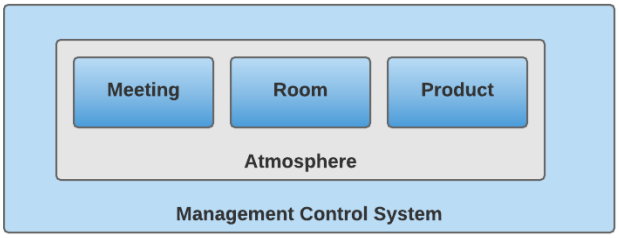

Figure 1: The Five Aspects Meal Model (Gustafsson et al., 2006)

2.2 Workspace

The second touchpoint that is included in this study, is the employee workspace. It is important to note that, even though "workplace" and "workspace" are often used interchangeably, they do represent two distinct concepts. While the workplace is the place where people go to work (i.e. office building), the workspace is where employees actually do their work (i.e. workstation). In order to evaluate workspace quality, buildings were traditionally assessed using a Post Occupancy Evaluation (POE) survey, which is a subcategory of the cross-sectional survey. This method aims to design better buildings, by identifying what people like and dislike about their environment. However, researchers' current interest extends beyond environmental satisfaction, and is more focused on measuring environmental support. To improve the work environment, researchers want to establish to what degree employees can conserve their attention and energy for their work, as opposed to expending it to cope with adverse environmental conditions (Vischer, 2008). For example, Brill & Weidemann (2001) conducted an extensive cross-sectional survey to establish which environmental factors have the strongest effect on productivity and job satisfaction. They identified the ten strongest workspace qualities, which they argue should be every organization's priority (table 1). According to Oseland (1999), all relevant workspace qualities can be divided into four distinct categories: ergonomics, physical conditions, spatial layout and aesthetics. Combined, these four categories constitute what we call the work environment.

Ergonomics

The first dimension, ergonomics, implies studying workspace features, as an extension of the human body. Derived from the Greek words ergo (work) and nomos (natural laws), ergonomics literally means the laws of work (Helander, 2005). Amongst others, the physical design of a workspace has a great impact on ergonomics. By assessing and improving office furniture, ergonomic researchers wish to protect employees from long-term muscular or nerve injury, which can occur due to poor bodily positioning (Vischer, 2007). In practice, if employees have a fixed workstation, specific ergonomic considerations can be given to their desk and chair design. However, as most employees do not have a fixed workstation, organizations are in need of an adaptable design, which provides ergonomic support for all users and all office tasks. Moreover, besides its impact on employee health and safety, ergonomic conditions also affect employee productivity (Dal & Neumann, 2009). This was confirmed by Miles (2000), who found that the additional investment in ergonomic tables and chairs can be earned back in only five months, because of increased productivity levels (Miles, 2000).

Table 1: Ten of the most important workspace qualities (Brill & Weidemann, 2001)

Rank	Workplace qualities
1	Ability to do distraction-free solo work
2	Support for impromptu interactions
3	Support for meetings and undistracted group work
4	Workspace comfort, ergonomics and enough space for work tools
5	Workspace side-by-side work and "dropping in to chat"
6	Located near or can easily find co-workers
7	Workplace has good places for breaks
8	Access to needed technology
9	Quality lighting and access to daylight
10	Temperature control and air quality

Physical conditions

The second aspect of the indoor environment consists of all the physical conditions. These are commonly referred to as the atmospherics, and contain all variables that can be sensed through sight, sounds, scent and touch (Kotler, 1973). It is basically an umbrella variable which covers people's perception of heating, cooling, ventilation, lighting and noise in one overall assessment. Under certain conditions, these variables can generate stress, which may have a negative impact on employee productivity (Clements-Croome, 2006). For example, Hameed & Amjad (2009) found that (a lack of) lighting has a significant impact on employee's sense of comfort and productivity. Moreover, besides lighting, acoustics also play a crucial role in evaluating the office environment. Here, the negative perception of acoustics is mostly referred to as noise, which is considered most irritating when irregular or unpredictable (Sundstrom, 2001). Previous studies found that noise is one of the main causes of dissatisfaction at the workplace, especially in open-plan offices (Brill & Weidemann, 2001). In general, people are in a flow state when they are at work, but when they are distracted they are brought out of this flow state (Mawson, 2002). Thus, noise causes employees to be distracted, which in turn has a negative impact on their productivity (Haynes, 2007).

Spatial layout

The third aspect of the indoor environment concerns the office layout, and its effect on employee privacy, communication and concentration. Amongst others, office layout includes elements such as e.g. workstation size, workstation density, and the distance between co-workers (Vischer, 2007). Based on these characteristics, each office can be categorized as being either an open-plan office, or a private enclosed office. When first introduced, open-plan offices were presumed to increase work efficiency and facilitate communication, while reducing building costs at the same time (Kamarulzaman et al., 2011). Although building costs have indeed proven to be lower for open-plan offices, it also caused people to complain about noise, a lack of privacy, and not being able to concentrate. On the other hand, although office workers want to be able to undertake distraction-free individual work, they also value the opportunity to have an informal conversation with their colleagues (Haynes, 2008; Brill & Weidemann, 2001). Thus, the ideal workspace would have to be flexible in order to provide the best support to employees.

Aesthetics

The fourth and last aspect of the indoor environment concerns the workspace aesthetics. Organizational aesthetics is most broadly defined as the appearance of the workplace, and includes variables such as e.g. décor attractiveness, cleanliness, and colour. Bains et al. (2013) emphasize the importance of a clean workplace, as cleanliness was found to contribute to employee efficiency, and to the room's ambience. Another relevant factor is the environment's colour scheme, as colours may influence people's mood, and even their well-being and productivity (Garris & Monroe, 2005). Therefore, in case of unsuitable interior colours, occupants might be subject to negative psychological impacts such as e.g. stress, depression, or boredom (Kamarulzaman et al, 2011). Garris & Monroe (2005) state that the constant viewing of brightly coloured computer screens creates the need for softer, more restful interior colours. However, a survey conducted by Human Spaces (2015) reported that most people (67%) felt happy walking into a bright office environment that was accented with green, yellow or blue colours.

Overall workspace evaluation

To summarize, the workspace environment can be broken down into four dimensions: ergonomics, aesthetics, spatial layout, and physical conditions. In order to evaluate these dimensions, researchers could ask employees how satisfied they are about each element. Instead, it is also possible to evaluate to what degree each element supports employees in terms of performing certain tasks and activities. According to existing studies, office workers consider the ability to undertake distraction-free individual work to be highly important, but at the same time they value the opportunity to have an informal conversation with another colleague. In order to meet both requirements, the workspace design requires a certain level of flexibility.

2.3 Restroom

This section looks into the various elements of a public restroom, and aims to determine how a person's perception of this touchpoint can best be captured. Within the scientific literature, public restrooms are most frequently discussed within the context of a restaurant. For example, Barber & Scarcelli (2009) studied the importance of restaurant restrooms using a cross-sectional survey. During this study, respondents were asked to rate twelve different aspects of the restroom. First, they were asked about personal hygiene (i.e. broken/clogged toilets, and the availability of toilet paper, soap, hot water, and paper towels), followed by some questions regarding the restroom appearance (i.e. dirty sink, dirty floors, dirty/cracked wall tiles, trash, and restroom odour). Results showed that the most important factors for evaluating a restroom were "no soap" and "toilet clogged or broken". Moreover, restroom cleanliness was perceived to be a direct indicator of kitchen hygiene and food safety. Respondents even indicated they had both chosen not to eat at, as well as not to return to, a restaurant in response to a visit to the restroom. Thus, evaluating the restroom has a direct impact on customers' overall satisfaction with the restaurant.

Besides restaurants, the importance of public restrooms has also been demonstrated within a retail environment. For example, in shopping centres customers perceive restrooms to be very important and have given them very high rankings. Amongst others, Dennis (2005) studied customer behaviour in shopping centres, and found that "the availability of a toilet" is considered the eighth most important factor, when it comes to shopping centre choice. Interestingly, restrooms are considered to be more important than many other factors, including "friendly atmosphere" and "helpfulness of staff". Similarly, Sit et al. (2003) found that the cleanliness, availability, and locational convenience of toilets was considered more essential than most other elements of a shopping centre.

Compared to the retail and restaurant environment, only few studies have assessed (elements of) restrooms within a work environment. However, going to the restroom, just like eating food, and working at a desk, is a necessary element of people's life and work dynamic. Besides toilets being a necessity, people also spend considerable time at this touchpoint. An independent survey of the British research firm "Opinion Matters" found that Dutch employees spend on average 33.5 hours per year on their toilet at work, which is equal to four entire workdays. Nevertheless, more than a quarter of the respondents indicated that the cleanliness of the toilets at work was insufficient. Also in the UK, restroom cleanliness has shown to be a problem. Here, the Association of Plumbing and Heating Contractors found that 16.5 percent of the employees was unsatisfied, because of office toilets not being up to scratch (United Nations, 2016). Research showed that 30 to 40 percent of all complaints are the result of toilets being insufficiently clean (Westerkamp, 2000). According to Mendat et al. (2004), who studied people's perception of public facilities, the greatest problems associated with public restrooms are related to: ventilation, maintenance and cleanliness.

As the activities people undertake in restrooms produce lots of unwanted odour, it is important that restrooms produce sufficient ventilation. Furthermore, besides people's fear of smelling each other, they also prefer not to hear each other while using the restroom. Any sound that is produced while using the restroom is considered embarrassing to both the user as well as the listener. People want a bathroom stall that functions as a site of private refuge, and part of providing this privacy involves creating the right acoustic atmosphere. If this atmosphere is not provided, people may, for example, flush the toilet before using it, to cover any sounds they might make (Vankamamidi, 2004). This kind of behaviour results in double flushing, which in turn increases the average water consumption per user. To avoid this from happening, some facilities choose to play a subtle background music. Alternatively, it is also possible to insulate the walls, floor and ceiling with acoustic material to cut down on the noise.

In addition, restroom décor also plays a role in evaluating restroom quality. For example, it was observed that very light colours, especially white, evoke thoughts of cleanliness, making it a fitting choice for public bathrooms. The downside is however that dirt is easily visible on light colours. Therefore, public restrooms should utilize materials that are less likely to show dirt and are easy to clean. Besides colour use, the number of toilets also determines a person's evaluation of the restroom. Especially if a restroom is located within a work environment, where people come every day and use the same restroom regularly, the number of toilets is crucial. In turn, restroom design also has an impact on other factors, including restroom odour and acoustics.

In conclusion, when evaluating the quality of a public restroom, elements that are considered important are: design, privacy, cleanliness, maintenance, odour, hygiene amenities, acoustics, and location. When it comes to complaints, 30-40 percent is related to restroom cleanliness. Closely related to cleanliness is the restroom odour. As the activities people undertake in restrooms produce lots of unwanted odour, it is important that restrooms produce sufficient ventilation. Furthermore, besides people's fear of smelling each other, they also prefer not to hear each other while using the restroom. People want a bathroom stall that functions as a site of private refuge, and part of providing this privacy involves creating the right acoustic atmosphere.

3.0 Literature - Employee performance

This chapter discusses the nature and characteristics of three aspects of employee performance: employee productivity, job satisfaction, and employee well-being (EWB). What links these variables, is their relation to the environmental comfort model, a system which determines to what degree the environment supports employees to work on specific tasks and activities (Vischer, 2008). This model discriminates between three levels of comfort: physical, functional and psychological comfort (figure 2). Together, these three comfort levels impact important employee outcomes, including well-being and satisfaction. Hence, it was decided to adopt one measure for each level of environmental comfort: job satisfaction represents physical comfort, productivity represents functional comfort, and EWB represents the level of psychological comfort.

Physical comfort and job satisfaction
The first level of Vischer's model concerns the level of physical comfort, which seeks to determine to what extent environmental characteristics affect user (dis)satisfaction. Generally, questionnaires such as Post-Occupancy Evaluation (POE) surveys are used to identify what employees like and dislike about their work environment (Vischer, 2008). Employee satisfaction offers a broad and comprehensive measure of environmental quality, and provides extensive information about employees' needs and preferences. Hence, job satisfaction is included as the first measure of employee performance.

Functional comfort and productivity
According to Vischer (2008), The difference between a supportive and an unsupportive workspace is the degree to which employees can conserve their attention and energy for their tasks, as opposed to expending it to cope with adverse environmental conditions. This is also referred to the "functional comfort", and goes beyond general findings on what people like and dislike, and towards assessing building performance. Functional comfort indicates to what degree features of the work environment support employees to concentrate on their work activities. For example, functional comfort can be measured by asking whether employees can perform their tasks easily, with difficulty, or not at all in their current workspace (Vischer, 2008). Therefore, employee productivity was included as the second measure of employee performance.

Psychological comfort and well-being
Workplace design generally stops at the level of functional comfort. However, there is a third and underdeveloped level – psychological comfort – at the top of the pyramid (figure 2). Amongst others, this level explores individual feelings of attachment and belonging, and relates to psychological factors such as e.g. privacy and environmental control. Moreover, this level links broad notions of employee health and well-being. Consequently, workplace design can be improved if employees' psychological comfort needs are met. Therefore, employee well-being (EWB) is included as the third measure of employee performance.

To summarize, workplace performance is all about creating an environment that supports employees to work on specific tasks and activities (Vischer, 2008). The difference between a supportive and an unsupportive workspace is the degree to which employees can conserve their attention and energy for their tasks, as opposed to expending it to cope with adverse environmental conditions. The three levels of environmental comfort are: physical, functional and psychological comfort (Vischer, 2008). Based on the nature and characteristics of each of these levels, it is decided to study employee performance through 1) employee productivity, 2) job satisfaction, and 3) employee well-being. In the remainder of this chapter, these three constructs are discussed in detail, along with possible ways of measuring them.

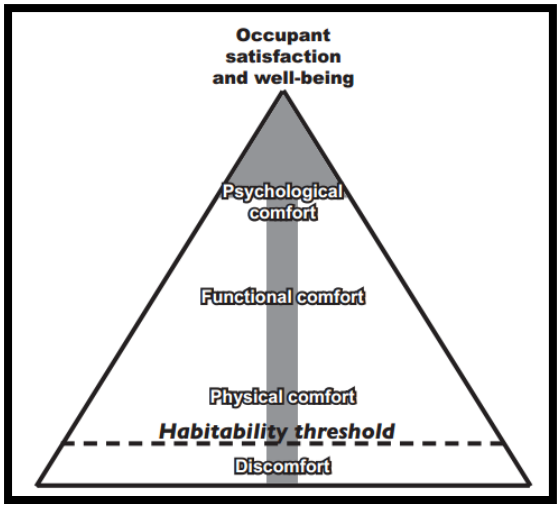

Figure 2: Environmental comfort model of workspace quality (Vischer, 2008)

3.1 Employee productivity

This section will look into the existing literature and theoretical discussion regarding the measurement of employee productivity. It discusses the historical context, the difficulty in defining employee productivity, and the approaches that have been developed to measure employee productivity. Looking at existing literature, various definitions of productivity can be found. For example, Sink (1985) defines it as "a ratio of input to output". Based on this definition, it could be stated that there are two possible ways of increasing productivity: either by increasing the output using the same input, or by achieving the same output with reduced input (Haynes, 2008). Another definition was provided by Sutermeister (1976), who states that productivity is "the output per employee hour, quality considered". Productivity was also defined by Rolloos (1997), who described it as "that which people can produce with the least effort". What all these definitions have in common is that productivity is described as a ratio which aims to measure how well an organization or individual converts input resources into goods and services (Hameed & Amjad, 2009).

In the literature, there is a general consensus that measuring employee productivity is crucial, due to its major impact on organizational outcomes. For example, even a seemingly small increase in employee productivity of 0.1 to 0.2 percent can already have dramatic effects on the profitability of an organization. This is because staff salaries can cover up to 90 percent of an organization's total costs (Clements-Croome, 2006). In addition, measuring productivity enables organizations to evaluate and improve organizational processes. Amongst others, accurate measurements of employee productivity can help to determine the impact of a renewed management approach, or a new technology. Other potential advantages of measuring employee productivity are related to e.g. improved personnel selection, job assignment, rewards and bonuses, strategic planning and performance forecasts. However, despite its importance, measuring employee productivity within office buildings is easier said than done.

Historically, the world's economy was largely based on manufacturing processes. Here, both in- and output could be clearly defined, which made it easy to measure productivity. However, throughout the years there has been a fundamental shift of a manufacturing economy, to a knowledge-based economy, where factories have been replaced with office buildings (Haynes, 2007). Opposed to manual workers, the tasks of office employees are not fixed, and their work cannot be measured in tangible outputs. Office employees carry out so many different tasks of which the output is not easily measurable or observable. It has therefore become notoriously difficult to measure, or even define productivity. Even today, there is no universally accepted measure of office productivity (Mawson, 2002). Nevertheless, researchers agree that any measurement is better than no measurement (Office of Real Property, 1999). Therefore, many researchers attempt to capture office productivity through indirect and surrogate measures. For example, Fitch (2004) suggested to adopt "satisfaction with the office environment" as a surrogate for employee productivity. Clements-Croome (2000) also suggests multiple indirect productivity measures, such as the number of interruptions, absence from workstation, speed and accuracy of work, and volunteer overtime. These and many other surrogate measures have been studied by Van der Voordt (2004), who concluded that all measures of employee productivity can be divided into five different categories (table 2).

Table 2: Five categories of productivity measures (Van der Voordt, 2004)

Method	Explanation and examples
Actual labour productivity	For example, number of phone calls per employee per day at a call centre, or number of man hours per vehicle in an automobile repair shop.
Perceived productivity	By asking people to assess their own productivity, also known as self-assessed productivity.
Amount of time spent	For example, time gained because of new and more efficient processes.
Absenteeism due to illness	A form of non-productivity.
Indirect indicators	For example, ability to concentrate, number of distractions, or satisfaction with the office environment.

One of these categories concerns self-assessed productivity measures (i.e. perceived productivity), where employees are asked to rate their own productivity. Within the organisational literature, this measure is frequently used (see e.g. Humphreys & Nicol, 2007; Clements-Croome, 2000; Oseland, 1999; Leaman & Bordass, 1995). In 1999, Oseland already stated that "self-assessment of productivity has been used in the field for some time, and has provided useful results". Leaman & Bordass (1995) also justify the use of a self-assessed productivity measure, as it is impossible to establish a standardized measure which can capture the actual productivity for all office employees. During their study, Leaman & Bordass (1995) asked participants: "Please estimate how you think your productivity at work is increased or decreased by the environmental conditions in the building". The answer was then given based on a nine-point scale ranging from -40% to +40% (loss and gain). Thus, employees had to report themselves, to what extent their environment impacts their productivity. Nevertheless, the scale was criticized because of the relatively large intervals, as it is generally agreed that only small changes in productivity are significant compared to building operational costs (Oseland, 2004). Humphreys & Nicol (2007) also adopted a self-assessed productivity measure, and asked participants "Do you feel that at present your productivity is being affected by the quality of your work environment, and if so to what extent?". Participants then had to give their answer based on a five-point Likert scale ranging from: "Much lower than normal" (-2) to "Much higher than normal" (+2). Other variants of self-assessed productivity questions are: "What percentage of your time is spent working productively?" and "What percentage of your time is spent working unproductively because of distraction?" (Van der Voordt, 2004).

One of the main advantages of a self-assessed productivity measure, is that you can cover the topic with one single question. This is especially important for surveys that have wider objectives than solely capturing employee productivity (Leaman & Bordass, 2000). However, at the same time this is also the biggest disadvantage, as one could argue that one question is not enough to draw meaningful conclusions. Another concern is that self-assessed productivity is not a quantitative operational measure. It is solely an estimate made by the participant, and may therefore not correlate well enough with employees' actual productivity (Van der Voordt, 2004; Leaman & Bordass, 2000). Moreover, socially desirable answers are highly likely, as no employee likes to admit that he or she has been unproductive (Van der Voordt, 2004). Still, it seems that self-assessed productivity is the most frequently adopted, and the most accepted measure for productivity.

Despite the lack of a universally accepted measure, there is a general consensus that the physical office environment has a significant impact on employee productivity (Haynes, 2007). A study by Hameed & Amjad (2009) on office design factors (i.e. furniture, noise, lighting, temperature and spatial arrangements) revealed that lighting has the greatest impact on employee productivity, followed by the spatial arrangements. Mawson (2002) proposes that in general, there are two major causes of productivity losses in offices: distractions, and a mismatch between the employee's work activities and the provided work environment. Distractions are defined as "anything that takes attention away from the task to be performed" (Mawson, 2002). Distractions can occur due to unexpected stimuli such as noise or visual disturbance, or because of the temperature being too high or too cold. Mawson (2002) states that on average, 70 minutes of productivity is lost in a typical eight-hour day as the result of distractions. According to Clements-Croome (2000), all these factors that impact productivity can be divided into four different groups, based on whether they are personal, social, organisational or workplace related.

To summarize, there are many factors that may impact employee productivity within the office environment, including the furniture, noise, lighting, temperature and spatial arrangements. Besides workplace factors, personal, social and organisational factors may also impact employee productivity. Unfortunately, measuring office productivity is a complex process, for which there is still no universally accepted measure. However, as an alternative, it is common to adopt self-assessed productivity as a surrogate measure of actual productivity. Here, participants have to rate their current productivity, with respect to their average productivity. Adopting this particular measure allows researchers to cover the topic by means of one survey question.

3.2 Employee well-being

This section examines the concept of well-being, and aims to determine how well-being can be measured within a work environment. Overall, it is accepted that well-being can have a major impact on employee performance, and therewith, on organizational performance (Harter et al., 2002). However, because it is such a broad domain, there is also considerable variation when it comes to the meanings and definitions that are used to describe well-being. Roughly, these definitions can be divided into three categories: objective well-being, subjective well-being, and a combination of the two. The first category, Objective Well-Being (OWB), refers to judgments that are independent of people's preferences and feelings (Diener, 1984). For example, within the context of an office environment, employee absenteeism would be a typical OWB measure.

The opposite of OWB is Subjective Well-Being (SWB), which refers to individuals' overall assessment of their life quality (Diener, 1984). More specifically, SWB is "a broad category of phenomena that includes people's emotional responses, domain satisfactions, and global judgements of life satisfaction" (Diener, 1984). These measures are subjective because the assessment is based on one's personal standards, rather than on others' standards. When it comes to the nature and measurement of SWB, research suggests that SWB has three core components: high levels of positive affect, low levels of negative affect, and a cognitive evaluation of one's satisfaction with their life as a whole (Diener, 1984). Thus, to be high on well-being is to be simultaneously low on negative emotion and high on positive emotion. However, it should be noted that this does not mean that individuals have to feel good all the time. The experience of painful emotions (e.g. disappointment, failure, stress) is a normal part of life, and being able to manage these negative or painful emotions is essential for long-term well-being. Only if negative emotions are very long lasting and start to interfere with a person's ability to function in his or her daily life, this will affect SWB (Huppert, 2009).

When it comes to well-being in the office environment, it is important to note that workplace situations differ greatly from general life situations. Therefore, when researchers wish to examine well-being of a specific life domain, such as work, it is important to distinguish between specific forms of well-being, and more general feelings about one's life. Daniels (2000) argues that context-specific measures of well-being are necessary to capture the subtleties, complexities, and variations of people's cognitive and affective experiences. However, measures of well-being should not solely consist of context-specific items, as recent literature suggests that EWB comprises both work and non-work dimensions (e.g. Page & Vella-Brodrick, 2009; Van Laar et al., 2007). Therefore, utilizing both work-related and more general well-being measures can provide a more accurate measurement of EWB, than do context-specific measures alone.

EWB is a broad term, which comprehends a large number of workplace factors (Harter et al., 2002). Still, defining EWB remains rather difficult as is a typical term of which "everyone understands the meaning, but nobody can give a precise definition" (Lyubomirs, 2001). Consequently, no universal definition and measurement of EWB have emerged, and most scholars define well-being for the purposes of their own research, and under their own research frameworks (Page & Vella-Brodrick, 2009). For example, EWB was previously defined as "that part of an employee's overall well-being that they perceive to be determined primarily by their work and can be influenced by workplace interventions" (Juniper et al., 2009). This notion that factors of the workplace impact EWB, has only started to evolve since the past sixty years. Historically, organizations were solely focused on avoiding sickness, opposed to optimizing health. Currently, EWB is receiving greater attention from both managers and scholars (Robertson & Cooper, 2010), as it is now recognized that EWB is linked to employee productivity and organisational performance (Page & Vella-Brodrick, 2008). For example, employees with low levels of well-being have shown to be less productive, make lower quality decisions, and are more prone to be absent from work. Thus, creating a comfortable and supportive working environment can enhance an individual's sense of well-being. As a result, there is a rising number of organizations who are taking a proactive interest in EWB, and wish to evaluate their workplace in terms of EWB, in order to identify new opportunities for effective management action (Juniper et al., 2009).

Since awareness is increasing that workplace characteristics have a direct impact on the physical and mental well-being of employees, a number of context-specific measures and models have been developed to assess EWB. A self-report questionnaire is hereby considered a valid way to measure well-being (Robertson & Cooper, 2011). For example, the 23-item Work Related Quality of Life (WRQoL) scale was developed by Van Laar et al. (2007). This method assesses EWB across six dimensions: job and career satisfaction, general well-being, home-work interface, stress at work, control at work, and working conditions. Respondents were asked to what extent they (dis)agreed with each of the 23 statements, using a 5-point Likert scale. Another EWB survey tool is the 'Shortened Stress Evaluation Tool' (i.e. ASSET; Faragher et al., 2004). This tool counts 86 questions about job perceptions, attitude towards the organization, and personal health. Each item is evaluated based on a 6-point Likert scale ranging from "strongly agree" to "strongly disagree". This tool is specifically focused on stress, as the effect of work-related stress on employee health is often an organization's main concern (Faragher et al, 2004). When the WRQol and ASSET scale are compared, it is striking that this comparison reveals more differences than similarities. This confirms the current literature on EWB, which states that no universal measure of EWB has yet been established.

Opposed to these extensive well-being scales, the World Health Organization (1988) has adopted a more compact scale, which consists of five relatively simple statements that tap into the subjective well-being of the respondents (i.e. The WHO-5 Well-Being Index). The WHO-5 was derived from the WHO-10, which in turn was derived from a 28-item rating scale (Topp et al., 2015). The WHO-5 is among the most widely used questionnaires assessing subjective psychological well-being. Since its first publication in 1998, the WHO-5 has been translated into more than 30 languages and has been used in research studies all over the world (Topp et al., 2015). What is remarkable about the WHO-5, is that the scale only contains positively phrased questions (figure 3).

The WHO-5 questionnaire

Instructions:
Please indicate for each of the 5 statements which is closest to how you have been feeling over the past 2 weeks.

	Over the past 2 weeks...	All of the time	Most of the time	More than half the time	Less than half the time	Some of the time	At no time
1	... I have felt cheerful and in good spirits	5	4	3	2	1	0
2	... I have felt calm and relaxed	5	4	3	2	1	0
3	... I have felt active and vigorous	5	4	3	2	1	0
4	... I woke up feeling fresh and rested	5	4	3	2	1	0
5	... my daily life has been filled with things that interest me	5	4	3	2	1	0

Scoring principle: The raw score ranging from 0 to 25 is multiplied by 4 to give the final score from 0 representing the worst imaginable well-being to 100 representing the best imaginable well-being.

Figure 3: The WHO-5 (Topp et al., 2015)

To summarize, measures of well-being can be divided into three categories: objective well-being (OWB), subjective well-being (SWB), and a combination of both. When researchers wish to examine one particular domain of wellbeing (e.g. at the workplace), it is important to distinguish between specific forms of well-being, and more general feelings about one's life. As there is no universally accepted definition, most scholars define well-being for the purposes of their own research, and under their own research frameworks. Within the office environment, scholars and practitioners generally adopt subjective measures of well-being, generally referred to as Employee Well-Being (EWB). EWB is that part of an employee's overall well-being that they perceive to be determined primarily by work and can be influenced by workplace interventions. With respect to measuring EWB, adopting a self-report questionnaire is preferred, because it is the view of the employee which the researchers are interested in. The most commonly used well-being questionnaire is developed by the World Health Organization (WHO), who have adopted 5 statements which tap into the respondent's subjective well-being.

3.3 Job satisfaction

Within organizational research, job satisfaction is probably the most common, as well as the oldest operationalization of workplace happiness, and has been defined in many different ways. For example, Vroom (1964) defines job satisfaction as "an orientation of emotions that employees possess towards the role they are performing at the work place". Another definition was provided by Adams & Bond (2000), who explain job satisfaction as "the degree of positive affect towards a job or its components". Still, perhaps the most-used definition in organizational research is that of Locke (1976), who defined job satisfaction as "a pleasurable or positive emotional state, resulting from the appraisal of one's job experiences". Despite what definition is adopted, all researchers recognize that job satisfaction is a global concept that comprises a wide range of facets. The most typical categorization was developed by Smith (1969), who states that job satisfaction consists of five different facets: pay, promotions, co-workers, supervision, and the work itself. A couple of years later, Locke (1976) added four additional facets to this categorization: recognition, working conditions, company, and management. Thus, job satisfaction is highly correlated to many factors, some of which are difficult to observe directly. So even though job satisfaction itself is not an absolute measure, it is still an indicator for a range of important job outcomes, including: loyalty, punctuality, cooperation, turnover and performance (Sell & Cleal, 2011; Judge et al., 2001; Spector, 1997). Therefore, the importance of measuring job satisfaction cannot be understated.

When measuring job satisfaction, roughly two different categories can be distinguished: multi-item and single-item measures. Multi-item measures are generally found in psychology and management journals. During these studies, individuals asked to report their satisfaction either with different aspects of their job (e.g. co-workers, workspace, salary) or with different formulations of job satisfaction (e.g. satisfied with my job, enthusiastic about my work). Thereafter, overall job satisfaction is calculated as the sum of these aspects (Skalli et al., 2008). Examples which are based on this method are the Job Descriptive Index (Smith, 1969), and the Job Diagnostic Survey (Hackman & Oldham, 1975). Alternatively, researchers can also decide to adopt a single-item measure to capture job satisfaction. This is *the second method*, which is mainly used by economists. Even though this is a straight forward approach, there are a number of variations. For example, while some researchers might ask respondents to rate their satisfaction with the job, others might ask respondents to rate their satisfaction with the work performed on the job. Moreover, variation is also found in the adopted measurement scales. For example, Sousa-Poza & Sousa-Poza (2007) adopted a scale ranging from 1 (dissatisfied) to 10 (satisfied), whereas Sell & Cleal (2011) only distinguished between four possible answers: "Yes, indeed", "To some extent", "Not so much", and "No or very seldom".

Throughout the years, both methods (i.e. single-item and multiple-item measures) have been widely used, and compared with each other in terms of reliability. Initially, it was assumed that multiple-item measures would have greater validity and higher reliability, because they capture multiple elements of the job, or multiple perspectives of job satisfaction. However, research showed that that single-item measures of job satisfaction perform just as well as composite measures (Dolbier et al., 2005). Some researchers even argue that single-item measures may be preferred, because they define job satisfaction more broadly, and are unconstrained by any specified job facets (Highhouse & Becker, 1993). Single-item measures include all aspects of job satisfaction, also those which multiple-item measures fail to capture. This was confirmed by Scarpello & Campbell (1983), who observed that individual questions about various facets of the job did not correlate well with a global measure of overall job satisfaction. Moreover, single-item measures were found to be less affected by temporal factors. These temporal factors include all emotions or attitudes linked to particular job facets that may vary when, for example, deadlines approach or problems arise (Linz & Semykina, 2012). Other advantages of single-item measures are that they are faster to complete, and easier to interpret by management (Dolbier et al., 2005).

To summarize, job satisfaction is perhaps the oldest operationalization of workplace happiness. However, there is still no consensus about how job satisfaction should be measured or defined. Some studies ask respondents about different aspects of job satisfaction, while others ask about global job satisfaction using different formulations. Moreover, while some studies adopt multiple-item measures, others use a single item to capture job satisfaction. Initially, it was thought that a multiple-item measure would be more reliable than a single-item measure of job satisfaction. However, research has shown that single-item measures are less affected by temporal factors, they are faster to complete, and easier to interpret.

4.0 Literature - Data collection methods

In this chapter, the two data collection methods (i.e. cross-sectional survey and EMA) are discussed. This includes for both methods: an introduction, brief history, (dis)advantages, reliability, and details and considerations regarding the data analysis.

4.1 Cross-sectional survey

The first data collection method of this study, is the cross-sectional survey. What characterizes a cross-sectional survey, is that the data is collected at one point in time. Here, the participants are ought to be a representative sample (cross-section) of the population. This way, the study's findings can be generalized to the entire study population. Most cross-sectional studies are performed using questionnaires, but alternatively participants could also be interviewed. The most common method used in work and organizational psychology is certainly the cross-sectional design based on self-report questionnaires (Ohly et al., 2010). During previous studies, cross-sectional designs have been used to study concepts such e.g. as job attitudes, personality or work characteristics. In general, survey research has three distinct characteristics (Malhotra & Grover, 1998). *First*, it involves collection of information by asking people for information in some structured format. *Second*, survey research is usually a quantitative method that requires standardized information in order to define or describe variables, or to study relationships between variables. *Third*, information is gathered via a sample with is a fraction of the population, with the need to be able to generalize findings from the sample to the population.

This method's popularity is not surprising given that most studies, particularly in (organizational) psychology, investigate individuals' perceptions of and beliefs about themselves, their immediate situation, and the relationship these perceptions and beliefs maintain with behaviour (Goddard & Villanova, 2006). When conducting this kind of research, cross-sectional methods offer various advantages. *First*, cross-sectional methods are frequently employed by researchers, and therefore widely accepted as a reliable method for data collection. *Secondly*, it is a relatively inexpensive method, as only one group is used, data are collected only once, and multiple outcomes can be studied. *Third*, cross-sectional surveys are generally completed in a relatively short period of time (depending on the required sample size, and the access to the study population). Nevertheless, despite all these advantages, authors of leading marketing journals have become increasingly concerned about the validity of cross-sectional surveys (Rindfleisch et al., 2008). This rising concern is an important issue, because cross-sectional methods are the most common form of empirical research in many areas, including organizational research, and thus provides a critical foundation for much of the knowledge on these topics (Jap & Anderson, 2004).

The main reason for concern, is that cross-sectional methods only provide a snapshot of the complex reality. Although researchers assume that their invested constructs have some stability over time, this is not always the case (Ohly et al., 2010). For example, people's mood, feelings and emotions fluctuate throughout the day, and can affect people's decision-making. Thus, a major disadvantage of cross-sectional methods is its inability to capture these short-term fluctuations. In addition, questionnaires rely on the perceptions and memory of the respondents, which also leads to inaccuracies. Recall errors occur in all cross-sectional studies that rely on respondent's self-report. The severity of the recall bias depends on a number of factors: interference, the length of time between the occurrence of an event and the recall of that event, and the respondent's psychological state (i.e. mood). Dependent on their mutual coherence, these factors can lead to an over- or underestimation of the event. When recall errors are distributed non-randomly, this is called recall bias, which is a threat to a study's validity (Van den Brink et al., 2001).

If the researcher finally decides, after weighing the pros and cons, to conduct a cross-sectional survey, it is important to establish the minimum sample size. In the literature, there has been substantial debate about the sample size needed to appropriately conduct tests of statistical significance. As sample size increases, the likelihood of attaining statistical significance increases, and it is important to note the difference between statistical and practical significance (Cohen, 1969). Both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) have shown to be particularly susceptible to sample size effects. The use of large samples assists in obtaining stable estimates of the standard errors to assure that factor loadings are accurate reflections of the true population values (Hinkin, 1998). In most cases, a sample size of 150 observations should be sufficient to obtain an accurate solution in exploratory factor analysis as long as item intercorrelations are reasonably strong (Guadagnoli & Velicer, 1988). In case of a CFA, a minimum sample size of 200 has been recommended (Hoelter, 1983). As the number of items increases, it may be necessary to increase the number of respondents. With larger samples, smaller differences tend to be detectable as more than mere sampling fluctuation.

To summarize, cross-sectional surveys are generally conducted using a self-report survey. This implies that data is collected at one point in time. Collecting data from a representative sample, allows researchers to generalize their findings to the entire population of interest. Cross-sectional methods are widely accepted, relatively inexpensive, and represent the most common form of empirical research. However, since recent years, people have become increasingly concerned about the validity of survey research, as cross-sectional methods are suspected of providing biased data. Cross-sectional methods are incapable of capturing short-term fluctuations, and the data is suspected to be unreliable due to respondents' recall bias.

4.2 Ecological Momentary Assessment (EMA)

This section will focus on the nature and characteristics of Ecological Momentary Assessment (EMA) methods. These methods are alternatively referred to as Experience Sampling Methods (ESM), or Daily Diary methods (DD). Rather than a single method, EMA is the collective of a range of methods (Shiffman et al., 2008). This makes it relatively difficult to provide an exact definition of what EMA entails. However, there are four key features that are common to all EMA approaches: 1) data are collected in real-world environments, 2) assessments focus on participants' current state, 3) moments are strategically selected for assessment, and 4) subjects complete multiple assessments over time (Shiffman et al., 2008). Thus, EMA methods involve the repeated assessment of a participant's momentary state in his or her real-world environment, where data are collected on many different occasions, but always from the same individuals. Rather than having to recall or summarize a feeling they have had in the past, participants are asked about their current (or very recent) feelings. Although EMA offers various advantages, organizational researchers have only recently started to show interest in these methods (Fisher & To, 2012). Beal & Weiss (2003) suggest that this sudden increase in popularity is caused by three factors.

First, there is an increasing recognition that employees' states and behaviours vary meaningfully over time. Everyday experience suggests that we are not always in the same mood and that even job performance may fluctuate from day to day. This is confirmed by scientific studies, who found convincing empirical evidence for the presence of these fluctuations (Ohly et al., 2010). For example, Weiss et al. (1999) measured workers' affective states at multiple times daily, over a 3-week period, and showed that the patterns of within-person variability themselves were predictable. Other examples that confirm the presence of such fluctuations include studies on job performance (Binnewies et al., 2009) and work engagement (Sonnentag, 2003). Besides capturing within-person variability, EMA data is also suitable for between-person analyses. This makes it possible to discover whether experiences are influenced by personal characteristics, and whether certain processes develop differently for certain persons. This is relevant for organizational psychologists, as they study the quality and factors influencing daily work experiences.

The *second* reason for the popularity of EMA, is that there is a greater recognition of the need to consider buildings in the context of business, and from the perspective of end users (Coenen et al., 2013). It is now widely recognized that the physical work environment has a considerable impact on many employee outcomes, including well-being and productivity. So despite the previous focus on cost reductions, practitioners now spend considerable time and money to create a work environment that supports the needs of their employees. When using EMA, the data is collected in a real-world environment, as the participants go about their lives. This is important because people's behaviour and experiences are also affected by the context. Collecting data in participants' natural environment therefore allows the generalization to the participants' real lives (i.e. ecological validity; Shiffman et al., 2008). Furthermore, it allows researchers to examine the processes connecting independent and dependent variables in more detail, which is complementary to information that can be obtained by more traditional methods (Reis, 1994).

The *third* reason for the use of EMA, is that there is an increasing recognition that people's summaries of their prior states, experiences, and behaviours are in fact poor reflections of the actual history of those states, experiences, and behaviours. Much of what people "recall" is in fact a reconstruction, which is pieced together from fragmentary inputs, and highly influenced by the person's mood, and by the time of retrieval (Robinson & Clore, 2002). This is because emotional experiences cannot be retrieved or re-experienced after they occur. The thoughts are available for retrospection, but not the actual feelings on which these thoughts were based. However, recalling contextual details may aid the accuracy of retrospection by allowing the person to recreate an emotional state that is compatible with the emotion experienced at the time of initial occurrence (Robinson & Clore, 2002). The ability to recall contextual details, however, declines quickly with the passage of time. Therefore, any delay between an experience and its report necessarily means a loss of information (Robinson & Clore, 2002). This realization has caused researchers to look beyond traditional methods, to find a way to capture people's current or very recent experiences and behaviours. EMA methods minimize the reliance on retrospective recall. Rather than having to recall or summarize a feeling they have had in the past, participants are asked about their current (or very recent) feelings (Freedman et al., 2006). To summarize, compared to recollective judgments, EMA produces information which is more closely linked to the actual experiences (Robinson & Clore, 2002).

After one has decided to use EMA, the next step involves choosing between three sampling techniques: interval-contingent sampling, event-contingent sampling, and signal-contingent sampling (Reis & Gable, 2000). In the case of *event-contingent sampling*, participants have to report after the occurrence of a particular event. For example, after social interaction, or after experiencing stress. Here, it is important that the participant reports immediately after the event has occurred (Ohly et al., 2010). A common problem with this type of sampling is, however, unclarity about the number of reports. It is therefore important that researchers provide participants with guidelines that clearly indicate which, and how many events should be reported (Ohly et al., 2010). The second type of EMA, *signal-contingent sampling*, refers to the type of data collection in which participants complete self-reports after they are being signalled. The timing of these signals could be at random or at fixed points in time. After each signal, the participants need to complete a questionnaire about their current or very recent thoughts, feelings and behaviours. Lastly, *interval-contingent sampling*, refers to all experiences that occur within a certain time interval (e.g. one hour, day or week). These reports often have an open format, where participants express themselves in their own words. This is in contrast to experience-sampling and event-sampling, which generally use a highly structured survey with mostly standardized questions (Ohly et al., 2010). Moreover, opposed to experience-sampling, the responses of interval-sampling are lagged (Ohly et al., 2010).

Besides choosing the type of sampling technique, there are also a number of other decisions regarding the implementation. As in any research, it is crucial to be very clear on the constructs of interest, and to create measures that assess these constructs accurately. Therefore, it is important to decide on the time frame for each survey item, which should be based on the main research question and the period in which the state or behaviour logically might vary (Fisher & To, 2012). Choices hereby include: the present moment, the time since the previous report, or a specific time interval such as the past hour, today, or this week (Fisher & To, 2012). Next, the desired time frame should be clearly communicated in each item. For example, participants could be asked "How happy do you feel *right now*?" or "How happy do you feel *today*?". Researchers can also decide to use existing items from scales that were originally developed for one-time retrospective reporting (i.e. cross-sectional survey). However, it is possible that some items need to be rephrased in order to make sense in a shorter time context. For example, "How satisfied do you *generally* feel about your job?" should be formulated as "How satisfied have you felt about your job *today*?" Moreover, it is important that the actual text of EMA items is short and simple (while remaining true to the construct), especially if items will be presented on a small screen of a mobile phone (Fisher & To, 2012).

Apart from determining how questions should be asked, it is also important to determine how many questions should be asked, as the response burden for EMA participants is relatively high due to the multiple assessments (Hektner et al., 2007). Thus, in order to motivate participants to respond regularly for multiple days, and to avoid annoyance, previous studies suggest that EMA reports should take no more than three minutes to complete (Hektner et al., 2007). Currently, there are few validated multi-item scales for EMA use (Fisher & To, 2012). Thus, to avoid crossing the time limit of three minutes, it is common for EMA researchers to shorten pre-existing scales. One way to do this, is to choose the items with the highest factor loadings from an pre-existing scale. However, when doing so, researchers should be careful to include all the relevant facets of the original scale. An alternative strategy is to solely include the items that are most likely to fluctuate between reports. For instance, "annoyance" will vary more over time than "anger", as annoyance is a milder state. Items that do not fluctuate much over the time frame of interest are less helpful in measuring within-person change, although they may contribute to assessing stable between-person differences (Shrout & Lane, 2012).

Currently, there is mixed evidence about the effectiveness of single- compared to multi-item measures. While some studies have found that both methods perform equally (Van Hooff et al., 2007), others found that multi-item measures are more reliable (Warren & Landis, 2007). Shrout & Lane (2012) state that at least three items should be used for every EMA construct. However, in EMA research, individuals are often asked to rate very straight forward unidimensional constructs in terms of current or very recent experience. For example, they have to report how hostile they feel right now, or how hard they were working when they were signalled. In these cases, Fisher & To (2012) claim that a single well-chosen item should be sufficient. This was confirmed by Van Hooff and colleagues (2007), who demonstrated that a single-item measure of "current fatigue" rated on a 10-point scale performed just as well as an established 6-item measure, when both were included in an EMA survey. Although, when single items are used to report on continuous constructs, it is desirable to use a larger number of response options, such as a 7- to 10-point scale, or a 0–100 slider scale, to increase variance (Fisher & To, 2012).

Another concern is the medium through which the participants will answer the survey. In the past, pen-and-paper methods were the standard for conducting EMA studies. Back then, participants had to complete the report when they were signalled by devices such as electronic pagers. The alternative, and preferable to pen-and-paper methods, are computerized methods. These include the use of mobile phones, palmtop computers, or Personal Data Assistants (PDAs), which are installed with specialized software. Especially the use of mobile technology offers many advantages, as most people already own a smartphone, know how to use them, and carry them everywhere (Shiffman, 2000). In addition, such mobile applications are both time- and cost-effective and widely available. In recent years, various mobile applications have been developed to collect EMA data (Freedman et al., 2006). However, the majority of these applications focus on EMA for health and psychological purposes only, and do not even mention the possibility of using the software for other contexts, such as the workplace. When a fitting medium has been chosen, and the questionnaire has been designed, it is time to conduct the survey. Fisher & To (2012) state that signal-contingent EMA studies in organizational behaviour often have three to five signals per day for one or two weeks (Fisher & To, 2012; Reis & Gable, 2000). Sadikaj & Moskowitz (2011) suggest that data collection should continue until 30 evaluations of each touchpoint have been reported by each participant.

To summarize, EMA emphasizes the importance of change and context in everyday behaviour and experience. EMA can capture changes in and correlates of employee performance, mood and other states as well as changes in work outcomes (Ohly et al., 2010). Theoretical studies suggest that, opposed to traditional methods, EMA minimizes recall bias, maximizes ecological validity, and creates a realistic view of the situation as it allows micro-processes that influence behaviour in a real-world context (Shiffman et al., 2008). When conducting an EMA study, the use of a mobile application for data collection is recommended, as it is both time- and cost-effective. It is important that the items in the questionnaire are suitable for EMA research, as many existing questionnaires were not designed for assessing momentary states. *First*, it is important to keep the number of questions to a minimum, as the participant burden of EMA studies is relatively high. To accomplish this, researchers are allowed to shorten pre-existing scales. *Second*, EMA questions should contain the right time frame (e.g. today, right now). It is therefore common for EMA researchers to rephrase pre-existing items, to ensure they fit the requirements of EMA. *Third*, EMA questions should not be too long. They need to be straight forward and to the point, especially if they will be presented on a small screen of a mobile phone. Then, depending on the sampling technique (i.e. interval-contingent sampling, event-contingent sampling, or signal-contingent sampling), participants may receive a signal, after which they complete the questionnaire, and the data is collected. In total, EMA data collection takes about 1-2 weeks, during which participants receive 3-5 signals per day.

5.0 Conceptual framework and hypotheses

For the basis of this study's conceptual framework (figure 4), the Stimulus-Organism-Response (S-O-R) framework has been used. This model was originally developed by Mehrabian & Russell (1974), and states that the environment contains certain stimuli (S) that cause changes to people's internal states (O), which in turn cause certain behaviour, or an approach/avoidance response (R). In the context of this study, the stimuli of the office environment are investigated. Consequently, these stimuli may have an impact on employee well-being and/or job satisfaction, which result in a change in employee productivity. The independent variables of this study are all the survey items that concern the restaurant, restroom and workspace. The dependent variables are the questions related to EWB, job satisfaction and productivity.

Based on the literature review, the following hypotheses have been formulated:

H0: The data resulting from EMA captures within-person fluctuations
H1: The standard deviation of the individual items of each construct is larger for EMA, compared to CS
H2: EMA and CS identify different relationships between the environment and employee response

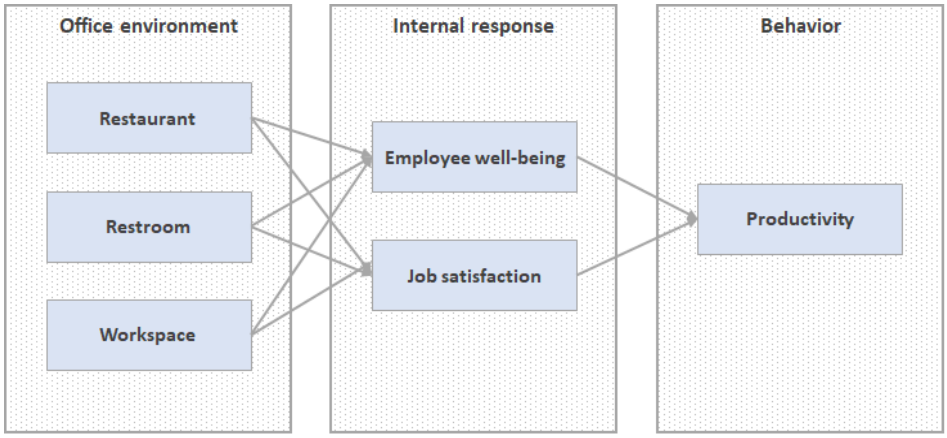

Figure 4: Conceptual framework

6.0 Methodology

This chapter describes the procedures which are used to collect and analyse the data. The first paragraph will elaborate on the case organization, the participants, and the procedure. In the second paragraph, the measures are explained, which have been used to develop the questionnaire. The chapter concludes with a more elaborate explanation of how the data were analysed.

6.1 Participants and procedure

During the empirical part of this study, a case study approach was taken. Case study research allows the exploration and understanding of complex issues, and is particularly suitable for holistic investigation (Zainal, 2007). A case study method enables researchers to closely examine the data within a specific context, and to understand the behavioural conditions through the actor's perspective (Zainal, 2007). Yin (1984) defines the case study research method as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident". In other words, a case study is a unique way of observing any natural phenomenon which exists in a set of data (Yin, 1984). By unique, it is meant that only a very small geographical area or number of subjects are examined in detail.

There are several categories of case studies. Yin (1984) discriminates between three categories, namely exploratory, descriptive, and explanatory case studies. In this study, an exploratory approach has been adopted, as the differences between a cross-sectional survey and EMA are being explored. The results are meant to open up the door for further examination of the differences between these two data collection methods. Because of this study's exploratory nature, it was decided to focus on one specific organization (i.e. single-case approach). The case organization of this study is a.s.r., a Dutch insurance company which is located in Utrecht, the Netherlands. The office building of a.s.r. counts 84.000 square metres, which makes it one of the largest office buildings in the Netherlands. Although the original building originates from the seventies, the environment was completely renovated between 2013-2015. After the renovation, there was a total capacity of 2.800 workplaces. Although a single-case design fits the objective of this study, it should be taken into account that due to the small number of subjects, a single-case provides very little basis for scientific generalisation.

After the case organization was determined, the search for participants started. In order to find enough participants, a message was posted on the local computer network of a.s.r.. Here, employees were informed about the study, and could apply on a voluntary basis. Volunteer sampling is a type of non-probability sampling, as people decide themselves whether they wish to participate or not. Therefore, chance does not play a role. On the one hand, the main strength of volunteer sampling is that it generally attracts motivated individuals who have a strong interest in the main topic of the survey. Moreover, volunteer sampling is convenient, relatively quick and inexpensive. On the other hand, non-probability sampling methods generally result in a sample which is not representative for the population of interest. As the sample is likely to differ from the actual population parameters, this is likely to cause a bias.

After a sufficient number of a.s.r. employees had volunteered to participate, the data collection period started. The *EMA survey* was designed and conducted with the aid of Shign software (Shign, 2019), and the data was collected between 18-02-2019 and 08-03-2019 (i.e. fifteen days, weekend days excluded). The EMA respondents had to download the mobile application of Shign, and create an account before they were able to fill out the survey. In the process of creating an account, respondents had to indicate whether they wanted to receive periodic reminders to fill out the survey, and if so, how often. Depending on their answer, they then received a number of push notifications, which were sent to motivate/remind them to complete the survey on a regular basis. Besides periodic reminders, these messages also included information about the start, the end and the duration of the study. Only the people who indicated that they wanted to be notified, received these push notifications. An overview of the messages other than the automatic periodic reminders, are presented in appendix C. The *cross-sectional survey* was distributed on 18-02-2019, and was designed using Qualtrics software. The respondents of the cross-sectional survey did not receive any reminders or push notifications. In case a respondent did not complete the entire survey, the response was not included in the dataset.

6.2 Measures

One questionnaire was used for both the EMA as well as the cross-sectional respondents. The final questionnaire has been developed on the basis of existing studies and survey instruments. However, some measures were reduced and/or rephrased in order to meet the EMA criteria. In total, the final questionnaire consisted of 31 questions. The items concerning the employee outcomes had a Likert scale ranging from 1 = strongly disagree to 7 = strongly agree. The items concerning the work environment ranged from 1 = very poor to 7 = very good. The questionnaire was originally developed in English, but it was later translated to Dutch to avoid misunderstandings due to unfamiliarity with the English language.

Productivity - Current literature states that measuring office productivity is a complex process, for which there is still no universally accepted measure. However, as an alternative, it is common to adopt self-assessed productivity as a surrogate measure of actual productivity. Hence, it was decided to include two items to measure self-assessed productivity. First, respondents were asked to what degree their current work environment contributes to their productivity (positive formulation), and second whether they struggled to finish their work in the current work environment (negative formulation).

Job satisfaction - According to the literature, global measures of job satisfaction may be preferred, because they are unconstrained by any specified job facets. They include all aspects of job satisfaction, also those which multiple-item measures fail to capture. Hence, it was decided to include two global measures of job satisfaction. The job satisfaction measure of this study was adopted from Dolbier et al. (2005). First, people were asked whether the current work environment contributed to their job satisfaction. Second, they were asked whether the current work environment made their work pleasant.

Employee well-being - Employee well-being was measured using three different items. Respondents were asked to what degree they felt: tired and stressed, calm and relaxed, cheerful and in good spirits. The last two items originate from the WHO-5 Well-Being Index. The original WHO-5 measure counts only five items, which makes it a great match for EMA research. In the original WHO-5 questionnaire, respondents have to base their answers on how they have been feeling during the past 14 days. However, for the purposes of this study, the time frame has been altered to the current moment. Therefore, “In the current work environment…” has been added to each question, to match the criteria of EMA. Besides the two statements that originate from the WHO-5 Well-Being index, a third item was added about feeling “tired and stressed”. It was decided to add this item, because the literature on EWB emphasizes the impact that stress may have on employees and their well-being.

Restaurant - The restaurant was measured using ten items. These specific items were all selected based on the four aspects (i.e. product, meeting, room, and atmosphere) of the Five Aspect Meal Model (FAMM), which has been developed to capture the experience of institutional meals. In this study, food variety and food quality are included to measure the “product”. Furniture, seating options, cleanliness and layout represent the “room”. Staff friendliness and queues capture the “meeting” aspect. Finally, acoustics and atmosphere measure the environment’s “atmosphere”.

Restroom - The restroom was measured using seven items: hygiene amenities, accessibility, cleanliness, privacy, odour, interior, and waiting line. These items were included based on the research that was conducted by Mendat et al. (2004), who studied the negative aspects of public restroom environments. Based on the mean rate of each item, it was decided to include these seven items to measure restroom perception.

Workspace - The workspace was measured using seven items: available workspaces, variety of workspaces, layout, interior design, acoustics, cleanliness, and indoor climate. These seven items were selected based on the study conducted by Oseland (1999), who stated that all relevant workspace qualities can be divided into four distinct categories: ergonomics, physical conditions, spatial layout and aesthetics.

6.3 Data analysis

In the current study, there are two datasets that need to be analysed. Although the underlying questionnaire of both datasets are identical, there are also a number of crucial differences. For example, although the cross-sectional data has independent observations, the EMA dataset contains repeated measures. In other words, the respondents of the EMA group have completed multiple surveys, and therefore the observations are not independent. Moreover, respondents of the cross-sectional survey have evaluated all three touchpoints, whereas the EMA respondents have evaluated one single touchpoint (i.e. their current location). Due to these differences, both datasets require different statistical methods for analysing the data.

For analysing the cross-sectional data, Multiple Linear Regression (MLR) was combined with a Principal Component Analysis (PCA). The EMA data was analysed using Linear Mixed Models (LMM). As multiple responses from the same subject cannot be regarded as independent from each other, LMM suits the nature of the EMA data. The problem of non-independence is solved in LMM by adding a random effect for each subject, which assumes there is a different baseline for each subject. LMM was performed using R Studio (version 3.5.2). Other statistical analyses (e.g. for the descriptives) were performed using the software of IBM SPSS Statistics. This software allowed to check for (multivariate) normality, correlations, outliers, and to inspect the overall distribution of the data.

7.0 Results

This chapter describes the steps that are taken to test the hypotheses, which have been presented in chapter 5.0. First, the survey response is discussed and described in paragraph 7.1. Subsequently, preliminary analyses were conducted to evaluate the validity and reliability of the collected data. The nature of both datasets is described in paragraph 7.2. In paragraph 7.3, the results of the cross-sectional data is discussed, and subsequently the results of the EMA data is presented in paragraph 7.4. The chapter concludes with paragraph 7.5, where the results of the two data collection methods are compared.

7.1 Response analysis

Originally, 83 people volunteered to fill out the cross-sectional survey. At the end of the data collection period, 80 surveys were completed (i.e. response rate of 96,4%). 46 people volunteered to join the EMA group. Out of these 46 people, 33 downloaded the app and completed at least one questionnaire (i.e. response rate of 71,7%). On average, an EMA respondent completed 5.7 surveys during the three weeks (i.e. fifteen workdays) of data collection. In figure 12, the total number of completed EMA surveys is plotted for each of the 33 respondents. The EMA survey resulted in 188 observations, resulting from 33 unique respondents. As the EMA respondents only evaluated their current location, the 188 observations were spread over three different touchpoints: workplace (n=95), restaurant (n=28), and the restroom (n=65).

As this sample size concerns only a small percentage of the entire population of interest, conclusions can be drawn about the dataset, but these conclusions cannot be extrapolated to the entire population. However, as the main interest is to identify the differences between the two data collection methods, and not to draw meaningful conclusions about the population of interest, the sample size is considered to be sufficient for the purpose of this study.

7.1 Descriptive measures

This paragraph numerically describes the features and patterns of the two datasets. This includes for each survey item: the number of respondents, the mean score, and standard deviation (table 3). A visual comparison of the mean scores between the two groups is displayed in figure 5 and 6. It is observed that the mean score of the EMA respondents is structurally higher than the mean score of the cross-sectional survey. Exceptions are however the two items “struggle to finish work” and “tired and stressed”, but these are both negative formulations. Thus, again the EMA respondents evaluate these items more positively. To evaluate whether these mean scores are also significantly different, a Mann-Whitney U test has been performed. Results showed that the scores of 16 out of the 31 items are significantly higher for EMA, when compared to the cross-sectional data. However, no specific pattern is observed.

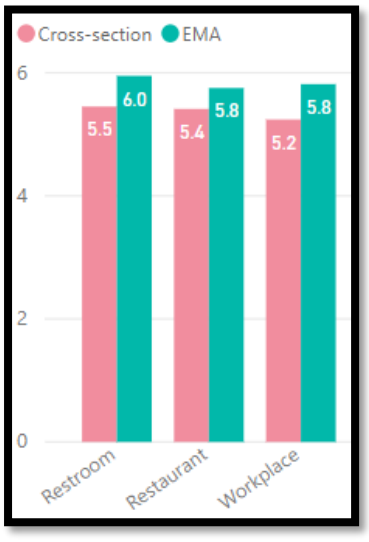

Figure 5: Mean score per touchpoint

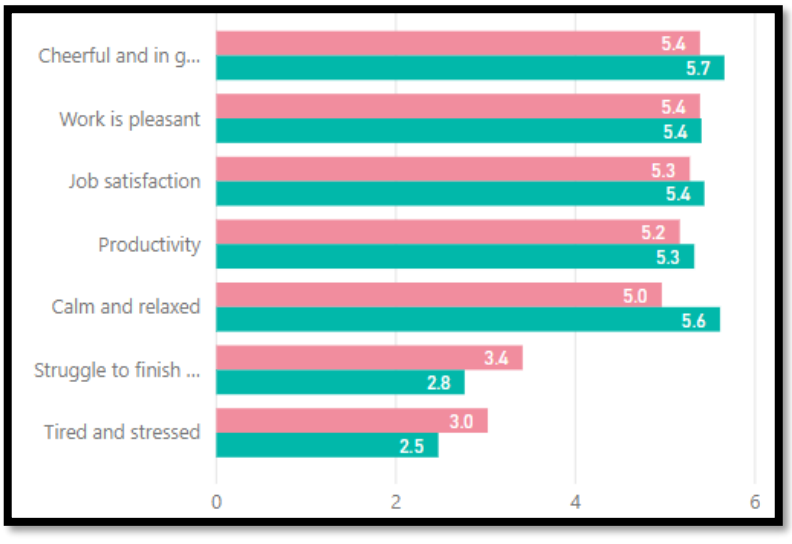

Figure 6: Mean score for employee performance aspects

Apart from the survey questions, the date and time of completion were also documented for each observation. The time of completion for each datapoint is displayed in figure 7. As EMA respondents are expected to answer the survey questions solely about their current location, it was expected that the time of completion would always be during office hours (i.e. between 08:00 and 18:00). However, as shown in figure 7, respondents also completed the survey outside of the office hours. Looking at the respondents of the cross-sectional survey, it was observed that the majority of them completed the questionnaire in the morning, between 08:00 and 10:00. However, the time of completion for this particular group is considered irrelevant, as the cross-sectional survey method does not require respondents to answer the questions about their current situation.

In figure 8, it is displayed at which day people filled out the survey. Here, it was observed that the majority of the cross-section respondents did this shortly after the survey was distributed. The response of the EMA respondents was also relatively high at the beginning of the data collection period. In an ideal situation, the response rates of the EMA respondents would remain constant throughout the entire data collection period. However, in reality the response rates declined. At the end of the data collection period, the response rates were significantly lower compared to the beginning. In figure 8, it is also indicated when the respondents of the EMA group have received a push notification on their mobile phone. Further details regarding these messages are presented in appendix C. In figure 9, the total number of EMA respondents per day is displayed per touchpoint. Here, it was observed that during the first few days of the data collection, most people started off with evaluating their workspace. Overall, the restaurant counts the lowest number of observations, but the distribution of restaurant evaluations remains relatively stable throughout time.

Besides the date and time of completion, the mobile application also requests permission to save the location of the EMA respondents. If permission was granted, the longitude and latitude of the respondent was saved at the time the survey was submitted. Based on the location, it could be checked whether the respondent was present at the office at the time the survey was submitted. An overview of this data is shown in figure 10 on a local level, and in figure 11 on a national level. It was observed that a considerable number of EMA observations has been recorded outside the a.s.r. building. This is considered problematic, as one of the main advantages of EMA is that respondents answer the survey questions based on their current environment. By violating this assumption, there is a possibility of the occurrence of a recall bias, just like with a cross-sectional survey.

At the end of both questionnaires, each respondent was given the opportunity to give some additional remarks. An overview of these comments is presented in appendix D. Comments which were mentioned most frequently, are shown in table 4. During the entire data collection period, most complaints resulted from the indoor climate within the building. In addition, people frequently mentioned that the relationship between the office environment and their productivity was either non-existent or unclear.

According to existing studies, EMA is superior to cross-sectional methods because this method is able to capture short-term fluctuations. To examine whether these short-term fluctuations can really be captured by EMA, table 5, 6 and 7 display the survey answers of one single respondent at different points in time. In table 6, the survey answers of one respondent about the restroom are shown across eight moments in time. What is striking, is that the survey scores regarding the static elements (i.e. restroom privacy, décor and accessibility) remain constant throughout time, whereas the scores regarding the dynamic elements (i.e. restroom odor and cleanliness) do fluctuate. In addition, the survey scores also show that the mood of the respondents is fluctuating throughout time.

Table 3: Descriptive values and Mann-Whitney U test

	Cross-section			EMA			Mann-Whitney U test	
Variable	N	Mean	Std. DV	N	Mean	Std. DV	Δ Mean	Sig.
Restroom - The cleanliness of the restroom	80	5.06	1.118	65	5.66	1.085	0.6	0.000
Restroom - The privacy of the restroom	80	5.41	1.052	65	5.89	0.843	0.48	0.000
Restroom - The hygiene amenities	80	5.41	1.177	65	5.88	1.015	0.47	0.177
Restroom - The odour of the restroom	80	5.24	1.034	65	5.71	1.160	0.47	0.000
Restroom - The décor of the restroom	80	5.55	0.992	65	5.97	0.463	0.42	0.245
Restroom - The queue for the restroom	80	5.85	1.020	34	6.35	0.836	0.5	0.026
Restroom - The accessibility of the restroom	80	5.64	0.903	65	6.38	0.600	0.74	0.000
Workspace - The layout of the workplace	80	5.55	1.124	94	6.11	1.057	0.56	0.081
Workspace - The availability of workspaces	80	5.39	1.258	94	6.20	1.135	0.81	0.000
Workspace - Work at different places	80	5.63	1.118	93	5.96	1.222	0.33	0.801
Workspace - The acoustics	80	4.46	1.484	94	5.17	1.449	0.71	0.000
Workspace - The interior design	80	5.61	1.000	94	6.07	0.970	0.46	0.334
Workspace - The indoor climate	80	4.91	1.314	95	5.55	1.185	0.64	0.146
Workspace - The cleanliness of the room	80	5.14	1.016	95	5.64	1.205	0.5	0.000
Restaurant - The atmosphere	80	5.50	1.079	28	6.00	0.535	0.5	0.028
Restaurant - The layout of the restaurant	80	5.54	0.941	28	5.75	0.950	0.21	0.232
Restaurant - The acoustics in the restaurant	80	4.81	1.213	28	4.79	1.081	-0.02	0.766
Restaurant - The quality of the furniture	80	5.63	0.960	28	5.75	0.987	0.12	0.190
Restaurant - The variation in the catering offer	80	5.49	1.031	27	5.96	1.036	0.47	0.029
Restaurant - The quality of the food	80	5.47	0.981	27	5.63	1.281	0.16	0.183
Restaurant - Staff friendliness	80	5.81	0.858	27	6.41	0.562	0.6	0.003
Restaurant - The cleanliness of the interior	80	5.65	0.828	28	6.11	0.618	0.46	0.009
Restaurant - The seating options	80	5.35	1.020	28	5.75	1.090	0.4	0.068
Restaurant - The queue for (self)service	80	4.89	1.136	26	5.38	1.211	0.49	0.393
Employee performance - Productivity	80	5.16	1.277	188	5.32	1.192	0.16	0.333
Employee performance - Finish work	80	3.41	1.605	188	2.77	1.653	-0.64	0.001
Employee performance - Job satisfaction	80	5.28	1.180	188	5.44	1.251	0.16	0.181
Employee performance - Work is pleasant	80	5.39	1.037	188	5.40	1.109	0.01	0.763
Employee performance - Calm & Relaxed	80	4.96	1.427	188	5.61	1.230	0.65	0.000
Employee performance - Good spirits& Cheerful	80	5.39	1.196	188	5.66	1.172	0.27	0.046
Employee performance - Stressed & Tired	80	3.03	1.653	188	2.47	1.596	-0.56	0.007

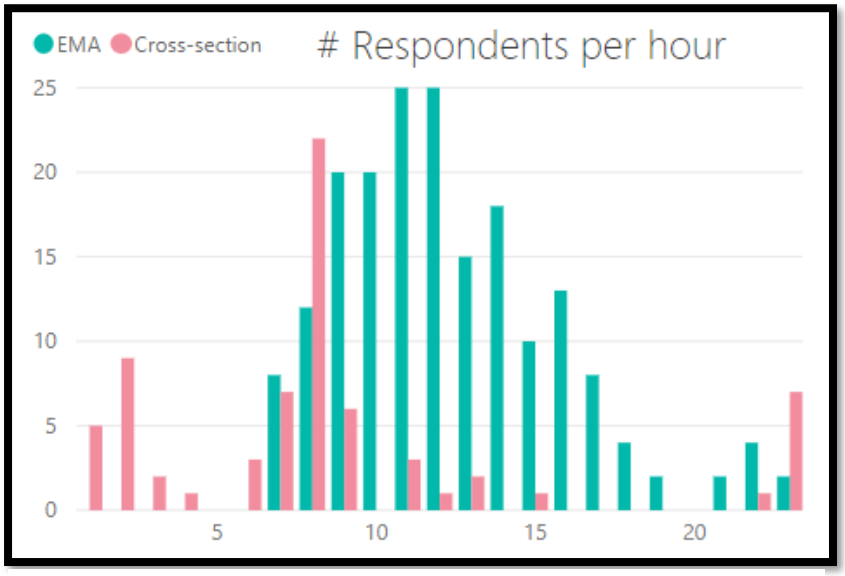

Figure 7: Time of completion for each datapoint

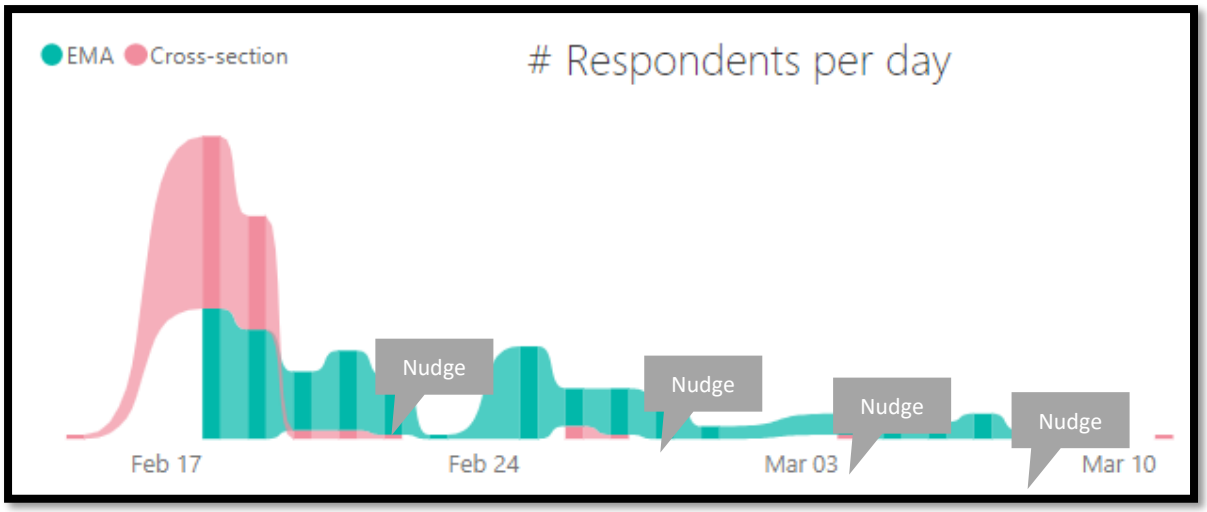

Figure 9: Number of EMA respondents per day, per touchpoint

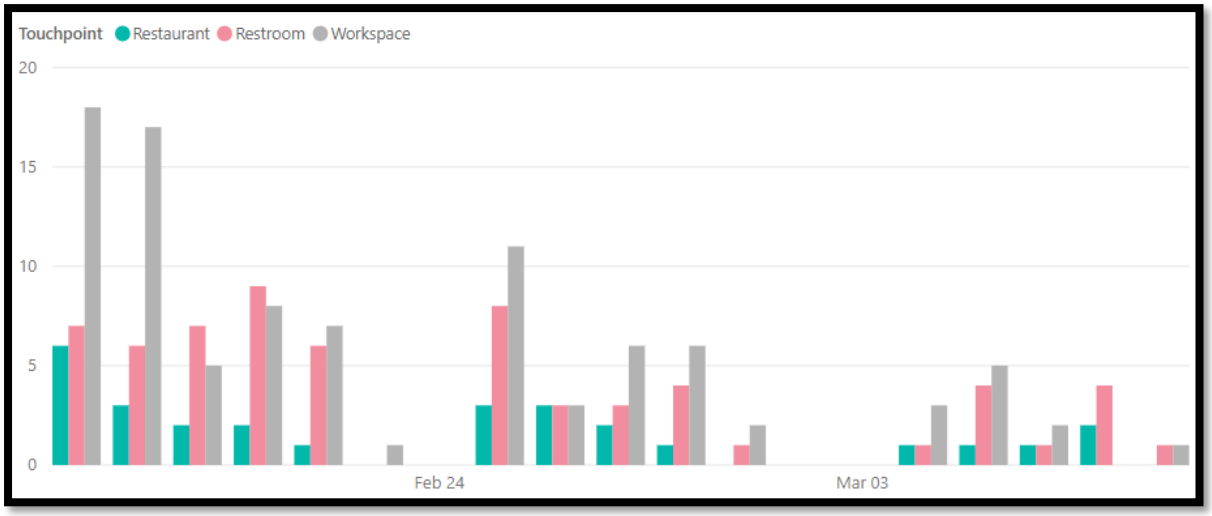

Figure 8: Respondents per day, divided between the three touchpoints

Figure 10: Location of EMA respondents within the a.s.r. office building

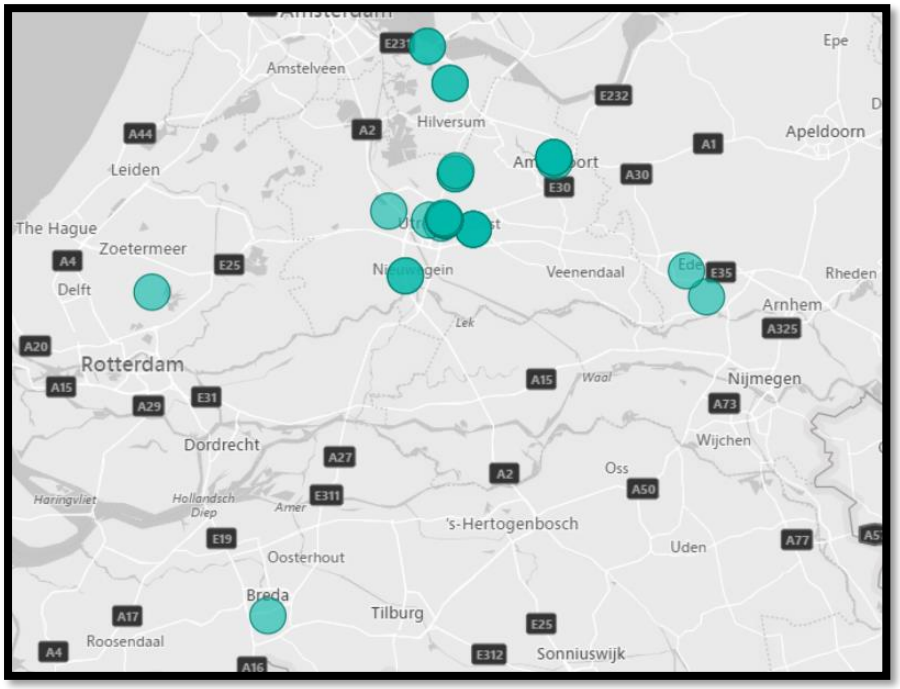

Figure 11: Location of EMA respondents on a national level

Table 4: Additional comments given by the respondents

Subject	Comment	Frequency
Workplace	Workplace cleanliness	4x
	Number of workplaces	3x
	Inability to concentrate (i.e. acoustics)	3x
Restaurant	Food prices	3x
	Restocking food supply	3x
General	Temperature / Air quality	12x
Survey questions	Link between work environment & employee outcomes is unclear	6x

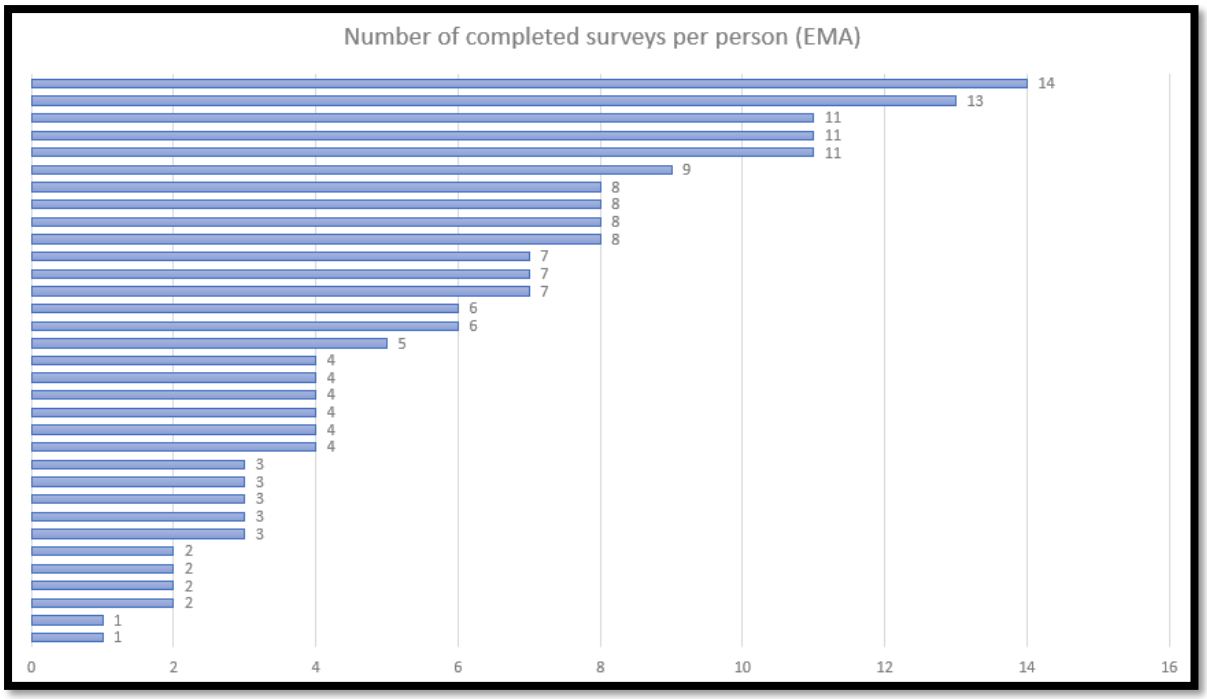

Figure 12: The total number of completed surveys for each of the 33 EMA respondents

Table 5: One respondent evaluating the restaurant at three points in time

	Moment in time		
	One	Two	Three
Restaurant - The variation in the catering offer	7	6	6
Restaurant - The quality of the food	6	6	6
Restaurant - The acoustics in the restaurant	6	6	6
Restaurant - The cleanliness of the interior	7	6	6
Restaurant - Staff friendliness	7	7	7
Restaurant - The quality of the furniture	6	6	6
Restaurant - The seating options	6	7	7
Restaurant - The queue for (self)service	5	7	7
Restaurant - The layout of the restaurant	7	6	6
Restaurant - The atmosphere	6	6	6
Employee performance - Productivity	6	6	6
Employee performance - Finish work	5	5	5
Employee performance - Job satisfaction	6	6	6
Employee performance - Work is pleasant	6	6	6
Employee performance - Stressed & Tired	4	5	5
Employee performance - Calm & Relaxed	5	6	6
Employee performance - Good spirits& Cheerful	4	6	6

Table 6: One respondent evaluating the restaurant at eight points in time

	Moment in time							
	One	Two	Three	Four	Five	Six	Seven	Eight
Restroom - The accessibility of the restroom	6	6	6	6	6	6	6	6
Restroom - The cleanliness of the restroom	4	4	3	3	3	5	3	4
Restroom - The privacy of the restroom	5	5	5	5	5	5	5	5
Restroom - The odour of the restroom	4	5	5	5	5	5	5	2
Restroom - The hygiene amenities	5	5	5	5	5	4	5	5
Restroom - The décor of the restroom	6	6	6	6	6	6	6	6
Restroom - The queue for the restroom	0	0	0	0	0	0	0	0
Employee performance - Productivity	4	5	4	4	4	4	4	4
Employee performance - Finish work	4	4	4	4	4	4	5	4
Employee performance - Job satisfaction	4	4	4	4	4	4	4	4
Employee performance - Work is pleasant	5	6	5	5	4	4	5	4
Employee performance - Stressed & Tired	4	3	3	3	3	3	5	4
Employee performance - Calm & Relaxed	4	5	5	5	5	4	4	4
Employee performance - Good spirits & Cheerful	4	5	5	5	5	5	3	4

Table 7: One respondent evaluating the restaurant at six points in time

	Moment in time					
	One	Two	Three	Four	Five	Six
Workspace - The availability of workspaces	7	7	7	7	5	5
Workspace - Work at different places	6	6	7	7	5	5
Workspace - The layout of the workplace	6	6	6	6	6	6
Workspace - The interior design	6	6	6	6	6	6
Workspace - The acoustics	6	6	6	6	6	6
Workspace - The cleanliness of the room	2	2	2	3	3	3
Workspace - The indoor climate	5	6	6	6	3	2
Employee performance - Productivity	6	6	6	6	5	5
Employee performance - Finish work	1	1	1	1	1	1
Employee performance - Job satisfaction	2	6	6	6	5	5
Employee performance - Work is pleasant	5	6	6	6	5	5
Employee performance - Stressed & Tired	2	2	2	1	2	2
Employee performance - Calm & Relaxed	6	6	6	6	6	6
Employee performance - Good spirits& Cheerful	6	6	6	6	6	6

7.2 Cross-sectional data

The preferred method for analysing the cross-sectional data is Structural Equation Modeling (SEM). SEM is particularly useful in the social sciences, where many if not most key concepts are not directly observable (Westland, 2010). The purpose of SEM is to examine a set of relationships between one or more Independent Variables (IVs) and one or more Dependent Variables (DVs). However, checking for the assumptions underlying SEM, showed that the cross-sectional data is not normally distributed, that the sample size is insufficient, and that the model fit is poor. Thus, although SEM would be the ideal method for analysing this type of data, it is not suitable for this specific dataset. As SEM is no longer an option for analysing the cross-sectional data, Multiple Linear Regression (MLR) and Principal Component Analysis (PCA) are adopted instead. What SEM and MLR have in common is that both methods test causal hypotheses within the dataset. However, opposed to SEM, MLR does not assume multivariate normality and the sample size requirements are significantly lower. The downside is however that MLR is not designed to deal with latent constructs. Luckily, this can be solved by combining MLR with PCA to identify the underlying dimensions (i.e. factors) of the original observed variables. In turn, these factors can be used for MLR, to examine the causal relations between the DVs and the IVs.

Principal Component Analysis

Principal Component Analysis (PCA) is conducted to reduce the original number of correlated variables, by transforming them into a smaller number of components, which still contain most of the information from the original data. First, the various assumptions underlying PCA were tested to make sure that PCA could be used to transform the data. First, PCA assumes there is a linear relationship between all variables. Linearity was tested using a matrix scatterplot. As testing linearity for all variables is quite an overkill, ten random combinations were computed. By analysing the scatterplots, it can be concluded that the relationships between the variables are more or less linear. Another prerequisite for factor analysis is that the variables are measured at an interval level (Field, 2009). Strictly seen, a Likert scale has discrete values. However, Ratray & Jones (2007) argue that a Likert scale is also suitable for PCA.

Secondly, PCA assumes sampling adequacy. In order to get a reliable result, large enough sample sizes are required. Pallant (2010) argues that there should be a ratio of at least five cases for each variable. As the questionnaire consists of 24 questions, the sample size should have been at least 120. Considering the fact that the actual sample size of the cross-sectional survey was 80, this would indicate that the assumption of sampling adequacy was violated. However, MacCallum et al. (1999) argue that the minimum sample size also depends on other aspects of the study design. In short, their study indicated that as communalities become lower, the importance of the sample size increases. They state that, with all communalities above 0.6, relatively small samples (i.e. less than 100) may still be perfectly adequate for PCA. Looking at the communalities of these 24 variables, only one scores slightly below 0.6. The item “acoustics in the restaurant” has a communality of 0.567. Based on this criterion, the assumption of a sufficient sample size would not be violated.

There is also a test which can determine whether the sample size is large enough to reliably extract factors. The Kaiser-Meyer-Olkin (KMO) test was performed to assess the suitability of the data for factor analysis. The score is always between 0 and 1, but Kaiser (1974) recommends accepting values only if they are greater than 0.5. The KMO score of this dataset was 0.832. According to Field (2009), a KMO value between 0.7 and 0.8 is considered to be a good score. The sampling adequacy can also be assessed for each individual variable, by looking at the values on the diagonal of the anti-image correlation matrix. Here, values should be above 0.5, but preferably higher (Field, 2009). The values of the cross-sectional survey range between 0.7 and 0.9, which again indicates that the sample is suitable for factor analysis. All arguments considered, it can be concluded that the assumption of a sufficient sample size is met.

The third assumption of PCA states that there should be adequate correlations between the variables, to reduce them into a smaller number of components. Given that PCA creates clusters of variables, it should be obvious that there are no clusters if variables do not correlate. This assumption can be tested using Bartlett’s test of sphericity, which examines whether the population correlation matrix resembles an identity matrix. Bartlett’s test of sphericity resulted in a value of $\chi 2(276) = 1192.285$, with a p-value of .000. Thus, it can be concluded that the correlations between variables are overall significantly different from zero. Altogether, it can be concluded that the cross-sectional data is suitable for performing PCA.

Table 8: Principal component analysis (CS)

	Component						
	1	2	3	4	5	6	**Communality**
1. Restroom (Cronbach's α = 0.886)							
Restroom - The cleanliness of the restroom	.802						0.730
Restroom - The privacy of the restroom	.792						0.749
Restroom - The hygiene amenities	.771						0.702
Restroom - The odour of the restroom	.722						0.680
Restroom - The décor of the restroom	.600						0.690
Restroom - The queue for the restroom	.547						0.771
Restroom - The accessibility of the restroom	.504						0.623
2. Workspace interior and acoustics (Cronbach's α = 0.841)							
Workspace - The layout of the workplace		.857					0.799
Workspace - The availability of workspaces		.795					0.701
Workspace - Work at different places		.791					0.724
Workspace - The acoustics		.621					0.766
Workspace - The interior design		.520					0.613
3. Restaurant interior and acoustics (Cronbach's α = 0.832)							
Restaurant - Atmosphere			.805				0.786
Restaurant - The layout of the restaurant			.796				0.773
Restaurant - The acoustics in the restaurant			.610				0.567
Restaurant - The quality of the furniture			.582				0.721
4. Restaurant food and service (Cronbach's α = 0.838)							
Restaurant - The variation in the catering offer				.837			0.768
Restaurant - The quality of the food				.830			0.793
Restaurant - Staff friendliness				.639			0.661
Restaurant - Cleanliness				.581			0.700
5. Workspace indoor climate and cleanliness (Cronbach's α = 0.687)							
Workspace - The indoor climate					.752		0.719
Workspace - The cleanliness of the room					.734		0.743
6. Restaurant crowdedness (Cronbach's α = 0.790)							
Restaurant - The seating options						.711	0.792
Restaurant - The queue for (self)service						.668	0.752

Factor interpretation

The results of the PCA are presented in table 8. Stevens (2002) suggests that for a sample size of 100, each loading should be greater than 0.512. As such, values below 0.5 were supressed. The PCA was conducted using a varimax rotation, as this type of rotation allows the variables to correlate. Based on the elbow of the scree plot, either two or six components should be retained. Other methods suggest to look at the total variance accounted for (VAF). This percentage should be larger than 70%, which would mean that six components would be retained (VAF = 72.2%). Based on Kaiser's criterion, which implies looking at the number of components with eigenvalues larger than one, again six components should be retained. Altogether, it was decided to retain a total number of six components.

The *first* component contains all the elements of the restroom, and naturally received the label "Restroom". The *second* component contains five of the original variables related to the workspace interior/acoustics. Therefore, the second component is labelled "Workspace interior and acoustics". The underlying variables of the *third* component are also related to the interior and acoustics, but of a different touchpoint, namely the restaurant. This component was therefore labelled "Restaurant interior and acoustics". The *fourth* component is composed of restaurant cleanliness, food variation, food quality and staff friendliness. They all say something about the level of service and/or food of the restaurant. This component therefore received the label "Restaurant food and service". The *fifth* component has only two underlying variables, resulting in the label "Workspace indoor climate and cleanliness". The *sixth* and last component is a combination of the restaurant's seating options and the queues. Both of these variables depend on the number of people present at the restaurant. Therefore, this component was labelled "Restaurant crowdedness".

Construct reliability

The reliability of the measured items was assessed through Cronbach's alpha, a measure of internal consistency. The value for Cronbach's alpha is $\alpha = 0.922$ for the 24 questionnaire items of the three touchpoints. Although the minimum value of alpha is still a point of discussion, if alpha exceeds 0.8 it can be assumed to be reliable, which means that the questionnaire has a good internal consistency (Field, 2009). To determine how each item individually contributes to the reliability of the questionnaire, it was also inspected what happens to alpha when one of the items is deleted. If alpha increases considerably if a particular item is deleted, one should consider removing this item entirely. In this questionnaire alpha roughly stays the same when one of the items is deleted. Cronbach's α was also computed for the original underlying variables of each of the six principal components. For the first four components, Cronbach's α always exceeded the 0.8 and there was no significant increase of one of the variables would have been deleted. For the last two components, Cronbach's α was below 0.8, but as these components consist of only two items, one of the underlying variables cannot be deleted.

Table 9: PCA outcome variables (CS) with varimax rotation

	Component			
	1	2	3	
1. Employee well-being (Cronbach's α = -1.603)				**Communality**
Well-being - feeling tired & stressed	-.847			.779
Well-being - feeling calm & relaxed	.943			.896
Well-being - feeling cheerful & in good spirits	.838			.718
2. Job complacency (Cronbach's α = 0.829)				
Productivity - employee productivity		.807		.728
Job satisfaction - job satisfaction		.881		.776
Job satisfaction - my work is pleasant		.877		.797
3. Finish work (Cronbach's α = not applicable)				
Productivity – difficult to finish work			.945	.948

Multiple linear regression

Multiple Linear Regression (MLR) is an extension of simple linear regression. It is used to predict the value of a dependent variable (DV) based on the value of two or more independent variables (IVs). MLR allows to determine the overall fit (VAF) of the model and the relative contribution of each of the predictors to the total variance explained. The F-ratio in the ANOVA table tests whether the overall regression model is a good fit for the data. The "R" column represents the multiple correlation coefficient. R can be considered to be one measure of the quality of the prediction of the DV. The R^2 value is the proportion of variance in the DV that can be explained by the IVs. Technically, it is the proportion of variation accounted for by the regression model. Besides R^2, the adjusted R^2 value should also be reported for an accurate interpretation of the data. While R^2 indicates how much variance is been explained by the model, adjusted R^2 only takes into account the variables whose addition in the model are significant. Typically, the more non-significant variables are added to the model, the larger the gap between R^2 and adjusted R^2 becomes. Statistical significance of each of the IVs can be tested with the t-value and the corresponding p-value. This tests whether the coefficients are equal to zero in the population. If $p < .05$ it can be concluded that the coefficients are statistically different to zero.

A multiple linear regression was run to predict the employee outcomes, based on the evaluations of the three touchpoints. The results of this analysis is presented in table 10-12. Based on the MLR of employee well-being in table 10, no causal relationships were found between the work environment and the level of employee well-being. The results of the MLR to predict the difficulty of finishing work is presented in table 11. A multiple linear regression was run to predict the difficulty of finishing work based on the evaluation of the workspace. The variables statistically significantly predicted the difficulty of finishing work, $F(2,77) = 3.444$, $p = .037$. Workspace interior and acoustics contributed significantly to this prediction, $t = -2.494$, $p = .015$. However, the workspace indoor climate and cleanliness did not, $t = -.818$, $p = .416$.

The results of the MLR to predict job complacency are presented in table 12. When trying to predict job complacency, it was found that the variables of the workspace statistically significantly predicted employee productivity and job satisfaction, $F(2,77) = 16.143$, $p < .0005$, adj. $R^2 = .277$. Based on the value of R^2, it can be stated that 27,7% of the variance within job complacency is caused by the quality of the workspace. As the t-values of both variables are significant, it can be concluded that both variables of the workspace contributed significantly to the prediction of job satisfaction and productivity. In addition, the variance within job complacency is also dependent on the quality of the restroom. A simple linear regression was run to predict job complacency, based on the evaluation of the restroom. Restroom statistically significantly predicts job complacency, $F(1,78) = 8.971$, $p = .004$. Based on the value of the adjusted R^2, it can be stated that 9.2% of the variance within job complacency is caused by the quality of the restroom.

Table 10: Multiple linear regression of employee well-being

Well-Being								
Touchpoint	**F-test**	**Sig.**	**R^2**	**Adj. R^2**	**Factor**	**Beta**	**t-test**	**Sig.**
Restroom	F(1,78) = 2.232	.139	.028	.015				
					Restroom	.167	1.494	.139
Restaurant	F(3,76) = .714	.547	.027	-.011				
					Crowdedness	.156	1.376	.173
					Food & Service	-.053	-.466	.642
					Interior & Acoustics	-.020	-.176	.861
Workspace	F(2,77) = 1.375	.259	.034	.009				
					Cleanliness & Indoor climate	.030	.267	.790
					Interior & Acoustics	.183	1.637	.106

Table 11: Multiple linear regression regarding the difficulty of finishing work

Difficult to finish work								
Touchpoint	**F-test**	**Sig.**	**R^2**	**Adj. R^2**	**Factor**	**Beta**	**t-test**	**Sig.**
Restroom	F(1,78) = .328	.569	.004	-.009				
					Restroom	.065	.573	.569
Restaurant	F(3,76) = 1.256	.296	.047	.010				
					Crowdedness	.015	.131	.896
					Food & Service	-.215	-1.917	.059
					Interior & Acoustics	.031	.275	.784
Workspace	F(2,77) = 3.444	.037*	.082	.058				
					Cleanliness & Indoor climate	-.089	-.818	.416
					Interior & Acoustics	-.272	-2.494	.015*

* indicates a significant value with $p < .05$

Table 12: Multiple linear regression of job complacency

Job complacency								
Touchpoint	**F-test**	**Sig.**	**R^2**	**Adj. R^2**	**Factor**	**Beta**	**t-test**	**Sig.**
Restroom	F(1,78) = 8.971	.004*	.103	.092				
					Restroom	.321	2.995	.004*
Restaurant	F(3,76) = 1.034	.382	.039	.001				
					Crowdedness	-.079	-.703	.484
					Food & Service	-.105	-.937	.352
					Interior & Acoustics	.148	1.316	.192
Workspace	F(2,77) = 16.143	.000*	.295	.277				
					Cleanliness & Indoor climate	.240	2.511	.014*
					Interior & Acoustics	.488	5.097	.000*

* indicates a significant value with p < .05

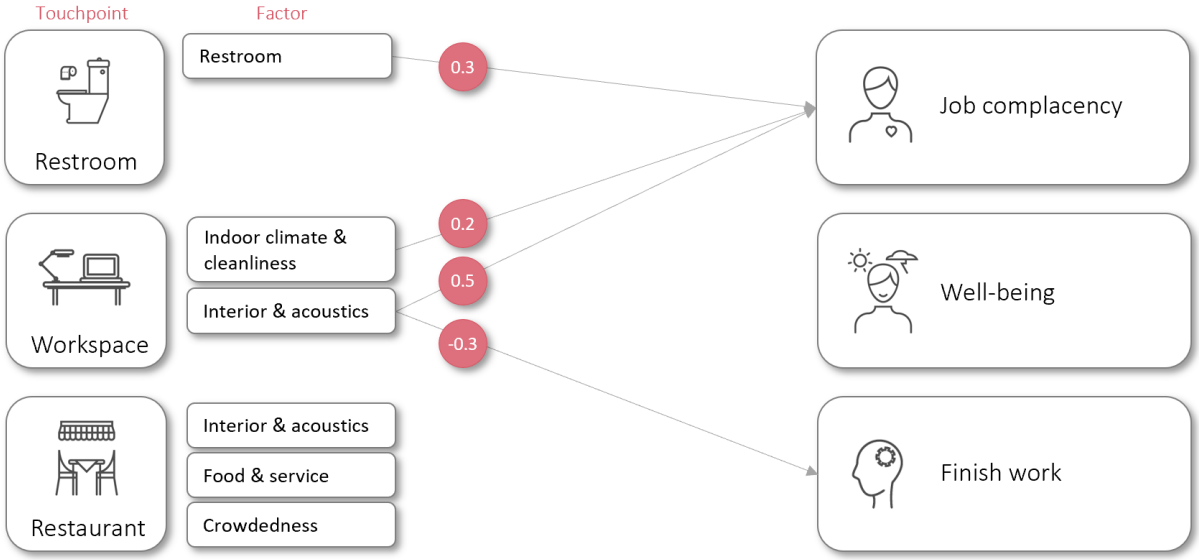

Figure 13: MLR results of the CS data; Confidence interval of 95%

7.3 EMA data

The EMA data was analyzed using Linear Mixed Models (LMM), because this method is particularly suitable for datasets that include repeated-measures. The available procedures in the general-purpose statistical software packages include SAS, SPSS, R and Stata. The current study has used R and lme4, to investigate the relationship between the office environment and employee performance of the EMA data. The name Linear Mixed Model, comes from the fact that these models are linear in the parameters, and that the IVs may involve a mix of fixed and random effects. In contrast to fixed effects, which are represented by constant parameters in an LMM, random effects are represented by (unobserved) random variables, which are usually assumed to follow a normal distribution.

Repeated-measures data may involve measurements made on the same unit over time, or under changing experimental or observational conditions. Measurements made on the same variable for the same subject are likely to be correlated (e.g., measurements of body weight for a given subject will tend to be similar over time). Models fitted to longitudinal or repeated-measures data involve the estimation of covariance parameters to capture this correlation. This is also the case with EMA data. By adding "(1|respondent)" to the regression equation, the by-subject variability is taken into account. The output of a mixed model is a list of explanatory values, estimates and confidence intervals of their effect sizes, p-values for each effect, and at least one measure of how well the model fits.

Assumptions

First, the dataset has been checked for outliers. Outlying data points - those which are well separated from the majority of the data - can have a large influence on an estimated model and its parameters (Orr et al., 1991). However, Likert scale variables have both a floor and ceiling (i.e. 1 and 7). It would be imprudent to discard particular observations, because someone responded at either the low or high end of the narrow spectrum. Therefore, it was decided to include all observations. Moreover, LMM assumes the absence of collinearity. If there is collinearity, the interpretation of the model becomes unstable, as the predictors may steal the explanatory power of each other. If predictors are very similar to each other, it becomes very difficult to decide what in fact is playing a bigger role. As the survey contains many questions which are correlated, it is decided to perform a PCA before proceeding to LMM.

Principal Component Analysis

Dimension reduction through PCA can transform the correlated variables into a smaller set of variables which can then function as fixed effects in an LMM. As respondents of the EMA group evaluated only one touchpoint at a time, it is not possible to conduct PCA for the dataset as a whole. Instead, PCA has been conducted three times, one for each touch point. As the underlying items in the PCA are correlated, it is decided to go with a varimax rotation. The PCA results have been presented in table 13-17.

PCA assumes sampling adequacy. In order to get a reliable result, large enough sample sizes are required. Pallant (2010) argues there should be a ratio of at least five cases for each variable. For the restroom, this implies that there should be 30 observations (i.e. 6 items * 5 cases per variable). With 65 observations, the sample size for the restroom is sufficient. Sample size is also sufficient for the workspace as 93 > 35. For the restaurant, sampling adequacy may not be sufficient, as the restaurant counts solely 26 observations. With 10 variables, 26 is not enough to meet the rule-of-thumb of five cases per variable. Besides this rule-of-thumb of five cases per variable, the KMO can also be computed to determine whether the sample size is large enough to reliably extract factors. Here, values should be above 0.5, but preferably higher (Field, 2009). For each touchpoint, the KMO value was larger than 0.5. The exact KMO values are displayed in table 13. In addition, sampling adequacy can also be determined based on the communalities. MacCallum et al. (1999) argue that with all communalities above 0.6, relatively small samples (i.e. less than 100) may still be perfectly adequate for PCA. Looking at the communalities of the EMA data, only the furniture quality of the restaurant, has a communality score below 0.6. Altogether, it can be concluded that this dataset is suitable for performing PCA.

Construct reliability

The reliability of the measured items was assessed through Cronbach's alpha, a measure of internal consistency. Based on Cronbach's alpha, the item "restroom queue" was excluded from the analysis due to its low reliability. By excluding "restroom queue" from the analysis, Cronbach's alpha increases significantly with α=0.763. Thus, it can be concluded that "restroom queue" does not measure the same underlying construct as the other survey questions about the restroom. As a result, this particular item has been excluded from the analysis. With respect to reliability, the other two touchpoints did not show any problems. The ten items about the restroom had a Cronbach's α=0.799, and the seven items related to the workspace a Cronbach's α=0.767. As deleting one of the individual items did not result to a significant increase in Cronbach's alpha, none of the items were excluded from the analysis.

Table 13: Overview of PCA outcomes (EMA)

PCA analysis	# Components	KMO	Chi-Square	VAF
Restroom	3	0.774	118.529 (p = 0.000)	65.1 %
Restaurant	3	0.595	119.033 (p = 0.000)	67.6 %
Workspace	2	0.769	240.754 (p = 0.000)	66.0 %

Table 14: PCA restroom (EMA) with varimax rotation

	Component 1	Component 2	Communality
1. Restroom design (Cronbach's α = 0.716)			
Restroom - The décor of the restroom	.753		.573
Restroom - The privacy of the restroom	.747		.548
Restroom - The hygiene amenities	.731		.575
Restroom - The accessibility of the restroom	.725		.596
2. Restroom cleanliness (Cronbach's α = 0.617)			
Restroom - The cleanliness of the restroom		.560	.690
Restroom - The odour of the restroom		.956	.922

Table 15: PCA workspace (EMA) with varimax rotation

	Component 1	Component 2	Communality
1. Workspace interior and acoustics (Cronbach's α = 0.848)			
Workspace - The layout of the workplace	.837		.733
Workspace - The availability of workspaces	.826		.693
Workspace - Work at different places	.825		.600
Workspace - The acoustics	.774		.560
Workspace - The interior design	.718		.713
2. Workspace indoor climate and cleanliness (Cronbach's α = 0.531)			
Workspace - The indoor climate		.829	.696
Workspace - The cleanliness of the room		.792	.628

Table 16: PCA restaurant (EMA) with varimax rotation

	Component			
	1	2	3	
1. Restaurant Food, Service & Acoustics (Cronbach's α = 0.833)				**Communality**
Restaurant - Staff friendliness	.833			.729
Restaurant - The queue for (self)service	.772			.672
Restaurant - The acoustics in the restaurant	.715			.641
Restaurant - The quality of the food	.713			.846
Restaurant - The variation in the catering offer	.675			.750
2. Restaurant Atmosphere & Cleanliness (Cronbach's α = 0.600)				
Restaurant - Atmosphere		.755		.612
Restaurant - Cleanliness		.683		.660
3. Restaurant Layout & Furniture (Cronbach's α = 0.576)				
Restaurant - The seating options			.785	.691
Restaurant - The layout of the restaurant			.646	.657
Restaurant - The quality of the furniture			.645	.503

Table 17: PCA outcome variables (EMA) with varimax rotation

	Component			
	1	2	3	
1. Employee well-being (Cronbach's α = -2.601)				**Communality**
Well-being - feeling tired & stressed	-.842			.839
Well-being - feeling calm & relaxed	.931			.886
Well-being - feeling cheerful & in good spirits	.863			.798
2. Job complacency (Cronbach's α = 0.867)				
Productivity - employee productivity		.823		.772
Job satisfaction - job satisfaction		.897		.841
Job satisfaction - my work is pleasant		.923		.867
3. Finish work (Cronbach's α = not applicable)				
Productivity - difficult to finish work			.898	.950

For the EMA dataset, there are three separate Principal Component Analyses, as the respondents have evaluated only one touchpoint at a time. In table 14, the PCA of the restroom is presented. This resulted in retaining two components. As odour is related to the level of cleanliness, it was decided to label the second component as "Restroom cleanliness". Table 15 contains the PCA for the workspace. This resulted in retaining two components. The first component has five underlying variables, which all say something about the workspace interior/acoustics. The second component consists of two variables, and received the label "Workspace indoor climate and cleanliness". Table 16 contains the PCA for the restaurant. It was decided to retain three components. Conducting PCA for the seven employee outcome variables, results in retaining three components (table 17). Remarkable is the value of Cronbach Alpha, which is -2.601 for the three items of employee well-being. Since the formulation of tired and stressed is negative, the average covariance among items has also become negative. This violates the reliability model assumptions of Cronbach's Alpha. As a result, deleting "feeling tired & stressed" would result in a Cronbach's Alpha of 0.861. As the correlation coefficient between these three variables is remarkably high (i.e. Pearson's correlation between .7 and .8 with $p < 0.05$), it is still assumed that these three variables measure the same underlying construct.

Results LMM

In the current design, there are multiple responses per subject. This would violate the independence assumption of MLR: Multiple responses from the same subject cannot be regarded as independent from each other. LMM can deal with this data, by adding a random effect for subject. This allows us to resolve the non-independence by assuming a different "baseline" value for each subject. Through LMM, two significant causal relationships were identified based on a p-value of .05. *First*, workspace indoor climate and cleanliness have a significant impact on job complacency ($\chi^2(1)=3.8427$, $p=0.049$). *Second*, the other component of the workspace, "Workspace interior and acoustics" had an even bigger significant impact on job complacency ($\chi^2(1)=11.727$, $p=0.05$). Together, they explain 20 percent of the variance within productivity and job satisfaction (adj. $R^2 = .2$). An overview of the results is presented in figure 15.

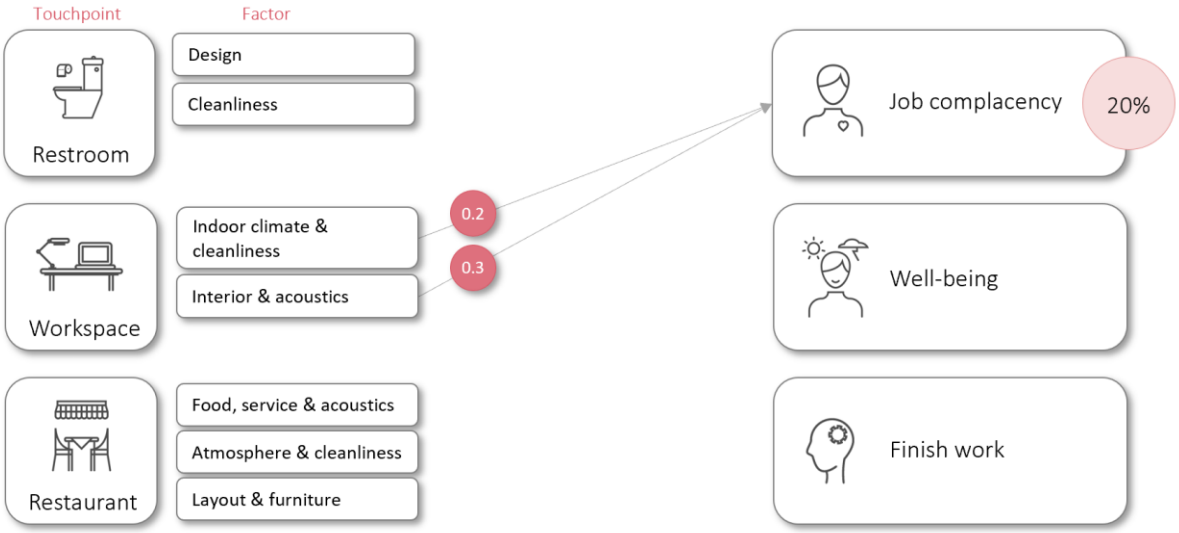

Figure 14: Results LMM of EMA data; Confidence interval of 95%

7.4 Cross-section versus EMA

Principal Component Analysis

For both the EMA as well as the CS data, a PCA was performed. In table 18, the factor interpretation of both methods is compared. For the restroom, the cross-sectional data resulted in retaining one component, whereas EMA distinguished between two different components. Moreover, due to its low reliability, the restroom queue was excluded from the EMA data analysis. With respect to the workspace, both methods resulted in retaining the same number of components, with the same underlying items. For the survey questions regarding the restaurant, the comparison is a bit more complex. Although both methods suggested to retain three components, the underlying items do differ. In table 19, the PCA results for the employee outcomes of both methods is compared. Although factor loadings did differ between the two methods, they did retain the same number of components, consisting of the same underlying items. In addition, for the cross-sectional data it was possible to run PCA on the whole dataset, whereas the EMA data required a separate analysis for each of the three touchpoints. This is because the EMA respondents evaluate solely one location, instead of all three.

Table 18: Factor solution for EMA versus CS

	EMA	Cross-Section
Restroom		
Restroom - The cleanliness of the restroom	1	1
Restroom - The odour of the restroom	1	1
Restroom - The hygiene amenities	1	1
Restroom - The privacy of the restroom	1	1
Restroom - The décor of the restroom	2	1
Restroom - The accessibility of the restroom	2	1
Restroom - The queue for the restroom	*excluded*	1
Workspace		
Workspace - The layout of the workplace	1	1
Workspace - The availability of workspaces	1	1
Workspace - Work at different places	1	1
Workspace - The acoustics	1	1
Workspace - The interior design	1	1
Workspace - The indoor climate	2	2
Workspace - The cleanliness of the room	2	2
Restaurant		
Restaurant - The quality of the furniture	1	1
Restaurant - The layout of the restaurant	1	1
Restaurant - The acoustics in the restaurant	2	1
Restaurant - Atmosphere	3	1
Restaurant - Cleanliness	3	2
Restaurant - The quality of the food	2	2
Restaurant - Staff friendliness	2	2
Restaurant - The variation in the catering offer	2	2
Restaurant - The queue for (self)service	2	3
Restaurant - The seating options	1	3

Table 19: Factor solution for EMA versus CS

Outcome variables	EMA	Cross-Section
Productivity - productivity	1	1
Productivity - difficult to finish work	1	1
Job satisfaction - job satisfaction	1	1
Job satisfaction - my work is pleasant	2	2
Well-being - feeling tired & stressed	3	3
Well-being - feeling calm & relaxed	3	3
Well-being - feeling cheerful & in good spirits	*3*	3

Linear Regression

The cross-sectional data is ideally inspected using SEM. However, because this dataset did not meet the sample size requirements of SEM, the data was analyzed using MLR instead. Because the EMA data contains repeated measures, it was not possible to analyze the data through MLR. Instead, it was decided to use LMM. When comparing the results of both analyses, it was observed that more statistically significant relationships were identified within the cross-sectional data. Where analyzing the EMA data resulted in finding two significant relationships, the cross-sectional data identified four significant relationships.

Conclusion

Based on the literature review, the following hypotheses had been formulated:

H0:	*The data resulting from EMA captures within-person fluctuations.* Within-person fluctuations were studied by zooming in at the survey response of one touchpoint, of one respondent, at different points in time. This indicates that the evaluations of fixed elements remain constant, whereas the evaluation of the dynamic elements fluctuate. In addition, the mood of the respondents also fluctuates across time. Thus, it can be concluded that EMA indeed captures within-person fluctuations, which cross-sectional methods fail to capture.
H1:	*The standard deviation for each construct is larger for EMA, compared to the cross-sectional survey.* No significant differences have been found between the two groups of respondents in terms of standard deviation of the individual survey items.
H2:	*EMA and CS identify different relationships between the environment and employee response* The EMA and CS data do lead to a slightly different factor interpretation. Moreover, more relationships were identified between the environment and employee outcomes within the CS data.

8.0 Discussion and conclusion

This study investigates the differences between EMA and a cross-sectional survey, within an office environment. Researching the characteristics and (dis)advantages of the existing data collection methods is important, because it will aid researchers to choose the right data collection method for their own research. Depending on the way the data is collected, the survey response, survey scores and in turn the study's results and conclusions may also vary. Currently, the majority of empirical studies within social sciences have adopted cross-sectional methods. However, recent literature claims that cross-sectional methods are often biased. Instead, methods that collect real-time data are preferred, such as EMA. EMA counts three major advantages: it takes into account within-person variability, minimizes recall bias, and maximizes ecological validity. Overall, it is claimed that EMA gives a much finer picture of what is going on in an environment throughout time, and that EMA gives a much more detailed picture of reality. However, no empirical studies have been performed to confirm these claims.

In order to accomplish this study's objective, the following research question was examined: *"How does an ecological momentary assessment differ from a cross-sectional survey, when evaluating the physical work environment in terms of its contribution to employee performance?"* Based on the literature review, employee performance was conceptualized as: employee well-being, employee productivity, and job satisfaction. The office environment was conceptualized as: the workspace, the restaurant, and the restroom. To answer the main research question, an exploratory case study was performed amongst the employees of a Dutch insurance company. The first group answered the cross-sectional survey though an online survey tool. The second group answered the same questions based on the principles of EMA. This data was collected through a mobile application.

As expected, there was a difference in survey response between the two data collection methods. The cross-sectional survey resulted in a response of 80 observations for each touchpoint. The EMA survey resulted in 188 observations, resulting from 33 unique respondents. As the EMA respondents only evaluated their current location, the 188 observations were divided between the three different touchpoints: the workplace (n=95), restroom (n=65), and restaurant (n=28). It was observed that the number of observations for the restroom and the restaurant is considerably lower than for the workspace. A logical explanation for this difference is that people spend most of their day behind their desk, and considerably less time at the restroom and the restaurant. Therefore, it is not surprising that the response for the workspace is significantly higher than for the other two touchpoints. It can therefore be discussed that the sample size of EMA depends on the amount of time the respondent spends at that particular location. In conclusion, the sample size of cross-sectional methods is constant, and the sample size of EMA differs between touchpoints.

With the EMA method, people are asked about their current feelings and opinions, instead of the sum of their past experiences. Therefore, it was expected that EMA would lead to a more varied response, and in turn a larger standard deviation, when compared to the cross-sectional data. However, no significant differences have been found between the two groups of respondents in terms of standard deviation of the individual survey items. Instead, when looking at the mean score of each survey item, it was observed that every item is evaluated more positively for EMA than for CS. Consulting the literature did not provide any explanations for this phenomenon. However, it could be discussed that people are more optimistic and positive when they are asked questions about their current situation. With the cross-sectional survey, the respondents were asked to answer the questions about the sum of their past experience. Perhaps, they have answered the questions less intuitively, and thought longer about their answer. Consequently, they may have worried more and they may have taken other factors into consideration as well. This may have lead to a more negative evaluation.

Besides the mean scores of the survey, the survey scores for each individual were also studied throughout time. It was observed that the dynamic aspects of the working environment (e.g. cleanliness and temperature) fluctuated, whereas the static aspects (e.g. layout and interior design) remained constant throughout time. This confirms existing studies on EMA, which state that the spatial conditions change throughout time. Thus, it can be concluded that EMA captures these short-term fluctuations, which cross-sectional methods fail to capture. Instead, cross-sectional surveys are more like a snapshot, an instantaneous photograph.

The results of the factor analyses were also compared between the two data collection methods. For the restroom, the cross-sectional data resulted in retaining one component, whereas EMA distinguished between two different components. Due to its low reliability, the question about the restroom queue was excluded from the analysis of the EMA dataset. With respect to the workspace, both methods resulted in retaining the same number of components, with the same underlying items. Both data collection methods resulted in retaining three restaurant components, but with different underlying items. It should also be noted that for the cross-sectional data it is possible to run one PCA on the whole dataset, whereas EMA requires a separate PCA for each touchpoint. This is because the EMA respondents evaluate only one touchpoint, instead of all three.

After conducting the factor analyses, the employee-environment relationships were studied for both groups. The two datasets were analysed in different ways, because the EMA dataset requires a method which is suitable for analysing repeated-measures. The results of the analyses showed that both datasets identified a significant relationship between the aspects of the workspace and employee complacency. In addition, the cross-sectional data also identified a significant relationship between the restroom and job complacency, and between the interior and acoustics of the workspace and the ability to finish work. These additional relationships may have only been identified within the CS data because the sample size of this group was significantly higher compared to EMA. Thus, these same relationships may be underlying the EMA data, but may currently not be significant due to the limited number of observations. Therefore, no conclusions can be drawn about whether EMA and cross-sectional methods identify more, less, or different relationships between the office environment and employee performance.

Despite the theoretical advantages that EMA may offer, there are more factors that need to be considered to conduct a successful EMA study. People are used to cross-sectional surveys. They know how the method works, and what is expected of them. Switching to EMA has shown to cause some confusion. For example, the additional remarks in the comment section showed that multiple EMA respondents either questioned or did not understand the added value of the repeated measures. Moreover, some respondents may not have understood why they need to answer the questions about their current environment. Consequently, many questionnaires were completed outside of the a.s.r. building. Thus, even if EMA would be superior to cross-sectional methods, it is crucial to pay sufficient attention to the communication with the respondents. Failing to do so might lead to biased data as well as a disappointing number of observations. As a result, the EMA dataset of this study was too small to reveal meaningful relationships between employee performance and their working environment. Thus, it can bed concluded that additional attention needs to be paid to the human factor. For example, it needs to be studied how to communicate the research method to the respondents, and how to ensure sustainable user engagement.

Overall, it can be concluded that there are indeed differences between EMA and CS, in terms of survey response, survey scores and the number of significant employee-environment relationships. Based on the results of this study, it is still thought that EMA is more reliable than CS, when it comes to collecting data. It was confirmed that EMA indeed captures short-term fluctuations which cross-sectional methods fail to capture. However, it is not enough to determine whether EMA is really superior to cross-sectional methods.

9.0 Limitations

In addition, several limitations should also be noted. *First*, the main limitation of this study is considered to be the relatively low survey response. In particular, the survey response of the restaurant and the restroom of the EMA group was considered to be insufficient. Consequently, SEM could not be performed to analyze the data, and other (less suitable) statistical methods had to be adopted instead. Therefore, it is likely that in reality more relationships exist between the work environment and employee performance. However, these relationships could not be identified due to the low survey response.

Second, both LMM and MLR assume that the data is collected from a random sample from the population of interest. However, during this study a non-probability sampling technique was adopted. Instead of a random sample, the respondents were selected on a voluntary basis. Therefore, it is less likely that the sample is a good representation of the population of interest. Moreover, people could choose whether they wanted to be part of the cross-sectional group or the EMA group. As participants were not randomly assigned to a group, there is a possibility that the observed differences between EMA and cross-section can be attributed to differences in personality traits, instead of to the data collection methods.

Third, an analysis on the GPS locations of the EMA respondents showed that some respondents had completed the survey at home, outside of office hours. This indicates that some respondents either had not read the instructions, or they may have ignored the instructions, or they simply did not remember the instructions. Thus, although EMA is based on answering survey questions about your current location, this assumption was not met for every observation. The violation of this assumption is considered a limitation of this study.

Fourth, many respondents indicated that the question regarding the queues at the restroom “does not apply”. Presumably, if there was no waiting line, people said “does not apply” instead of giving a high rating. Due to all these missing values, this particular survey item has become relatively meaningless and unreliable.

References

Adams, A., & Bond, S. (2000). Hospital nurses' job satisfaction, individual and organizational characteristics. *Journal of advanced nursing*, *32*(3), 536-543.

Alliger, G. M., & Williams, K. J. (1993). Using signal-contingent experience sampling methodology to study work in the field: A discussion and illustration examining task perceptions and mood. *Personnel Psychology*, *46*(3), 525-549.

Andersson, T. D., & Mossberg, L. (2004). The dining experience: do restaurants satisfy customer needs? *Food Service Technology*, *4*(4), 171-177.

Barber, N., & Scarcelli, J. M. (2009). Clean restrooms: how important are they to restaurant consumers? *Journal of Foodservice*, *20*(6), 309-320.

Barrett, P.S. & Baldry, D. (2003). *Facilities Management: Towards Best Practice*, 2nd ed., Blackwell, Oxford.

Beal, D. J., & Weiss, H. M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods*, *6*(4), 440-464.

Binnewies, C., Sonnentag, S., & Mojza, E. J. (2009). Daily performance at work: Feeling recovered in the morning as a predictor of day-level job performance. *Journal of Organizational Behavior*, *30*(1), 67-93.

Bollen, K.A. (1989). *Structural Equations with Latent Variables*. Wiley, New York. p. 268

Boomsma, A. (1985). Nonconvergence, improper solutions,and starting values in LISREL maximum likelihood estimation. *Psychometrika*, 50, 229-242.

Brayfield, A. H., & Rothe, H. F. (1951). An index of job satisfaction. *Journal of Applied Psychology, 35*(5), 307-311.

Brill, M., & Weidemann, S. (2001). *Disproving widespread myths about workplace design*. Kimball International.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd Edition). New York, NY: Guilford Press.

Campbell-Smith, G. (1967). *Marketing of the Meal Experience. A Fundamental Approach.*

Clements-Croome, D., & Baizhan, L. (2000). Productivity and indoor environment. In *Proceedings of Healthy Buildings* (Vol. 1, pp. 629-634).

Clements-Croome, D. (2006). *Creating the productive workplace*. Taylor & Francis.

Coenen, C., Alexander, K., & Kok, H. (2013). Facility management value dimensions from a demand perspective. *Journal of Facilities Management*, *11*(4), 339-353.

Daniels, K. (2000). Measures of Five Aspects of Affective Well-being at Work. *Human Relations*, 53 (2), 279-294.

Dennis, C. (2005). *Objects of desire: Consumer behaviour in shopping centre choices*. Basingstoke: Palgrave Macmillan

Diener, E. (1984). Subjective well-being. *Psychological bulletin*, *95*(3), 542.

Dolbier, C. L., Webster, J. A., McCalister, K. T., Mallon, M. W., & Steinhardt, M. A. (2005). Reliability and validity of a single-item measure of job satisfaction. *American Journal of Health Promotion*, *19*(3), 194-198.

Edwards, J. S., Meiselman, H. L., Edwards, A., & Lesher, L. (2003). The influence of eating location on the acceptability of identically prepared foods. *Food Quality and Preference*, *14*(8), 647-652.

Engelen, L., Chau, J. Y., Burks-Young, S., & Bauman, A. (2017). Application of ecological momentary assessment in workplace health evaluation. *Health Promotion Journal of Australia*, *27*(3), 259-263.

Faragher, E. B., Cooper, C. L., & Cartwright, S. (2004). A shortened stress evaluation tool (ASSET). *Stress and Health: Journal of the International Society for the Investigation of Stress*, *20*(4), 189-201.

Fisher, C. D., & To, M. L. (2012). Using experience sampling methodology in organizational behaviour. *Journal of Organizational Behavior*, *33*(7), 865-877.

Field, A. (2009). *Discovering statistics using SPSS*. Sage publications.

Fitch, E. (2004). *Creating and effective office environment within a managed work space*, MSc Dissertation, Sheffield Hallam University.

Freedman, M. J., Lester, K. M., McNamara, C., Milby, J. B., & Schumacher, J. E. (2006). Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. *Journal of substance abuse treatment*, *30*(2), 105-111.

Garris, L.B. & Monroe, L.K. (2005). The Colour Factor. *Journal of Buildings*, 99(10), 72 –73.

Goddard III, R. D., & Villanova, P. (2006). Designing surveys and questionnaires for research. *The psychology research handbook: A guide for graduate students and research assistants*, 114-125.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological bulletin*, *103*(2), 265.

Gustafsson, I. B., Öström, Å., Johansson, J., & Mossberg, L. (2006). The Five Aspects Meal Model: a tool for developing meal services in restaurants. *Journal of foodservice*, *17*(2), 84-93.
Hackman, J. R., & Oldham, G. R. (1975). Development of the job diagnostic survey. *Journal of Applied psychology*, *60*(2), 159.
Hameed, A., & Amjad, S. (2009). Impact of office design on employees productivity: a case study of banking organizations of Abbottabad, Pakistan.
Harter, J. K., Schmidt, F. L., & Keyes, C. L. (2003). Well-being in the workplace and its relationship to business outcomes: A review of the Gallup studies. *Flourishing: Positive psychology and the life well-lived*, *2*, 205-224.
Haynes, B. P. (2007). Office productivity: a theoretical framework. *Journal of Corporate Real Estate*, *9*(2), 97-110.
Haynes, B. P. (2008). The impact of office comfort on productivity. *Journal of Facilities Management*, *6*(1), 37–51.
Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Sage.
Helander, M. (2005). A guide to human factors and ergonomics. Crc Press.
Highhouse, S., & Becker, A. S. (1993). Facet measures and global job satisfaction. *Journal of Business and psychology*, *8*(1), 117-127.
Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational research methods*, *1*(1), 104-121.
Hoelter, J. W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods & Research*, *11*(3), 325-344.
Holman, D. (2004). Employee well-being in call centres. In *Call centres and human resource management* (pp. 223-244). Palgrave Macmillan, London.
Human Spaces (2015). The Global Impact of Biophilic Design in the Workplace. Retrieved from http://www.humanspaces.com/global-report/
Humphreys, M. A., & Nicol, J. F. (2007). Self-assessed productivity and the office environment: monthly surveys in five European countries. *ASHRAE transactions*, *113*, 606.
Huppert, F. A. (2009). Psychological well-being: Evidence regarding its causes and consequences. *Applied Psychology: Health and Well-Being*, *1*(2), 137-164.
Jap, S. D., & Anderson, E. (2004). Challenges and advances in marketing strategy field research. *Assessing marketing strategy performance*, 269-92.
Johns, N., & Pine, R. (2002). Consumer behaviour in the food service industry: a review. *International Journal of Hospitality Management*, *21*(2), 119-134.
Josiam, B., Sohail, M. S., & Monteiro, P. (2007). Curry cuisine: Perceptions of Indian restaurants in Malaysia.
Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job satisfaction–job performance relationship: A qualitative and quantitative review. *Psychological bulletin*, *127*(3), 376.
Juniper, B., White, N., & Bellamy, P. (2009). Assessing employee wellbeing: is there another way? *International Journal of Workplace Health Management*, *2*(3), 220-230.
Kaiser, H.F. (1974). Index of factorial simplicity. *Psychometrika,* 39, 31 – 36.
Kamarulzaman, N., Saleh, A. A., Hashim, S. Z., Hashim, H., & Abdul-Ghani, A. A. (2011). An overview of the influence of physical office environments towards employee. *Procedia Engineering*, *20*, 262-268.
Kivela, J., Inbakaran, R., & Reece, J. (2000). Consumer research in the restaurant environment. Part 3: analysis, findings and conclusions. *International Journal of Contemporary Hospitality Management*, *12*(1), 13-30.
Kotler, P. (1973). Atmospherics as a marketing tool. *Journal of retailing*, *49*(4), 48-64.
Leaman, A., & Bordass, W. (1995). Comfort and complexity: unmanageable bedfellows. In *Workplace Comfort Forum, RIBA, London* (pp. 22-23).
Leaman, A., & Bordass, B. (2000). Keeping occupants "Satisficed". *Energy and Environmental Management*, *2*, 23-7.
Leblebici, D. (2012). Impact of workplace quality on employee's productivity: case study of a bank in Turkey. *Journal of Business Economics and Finance*, *1*(1), 38–49.
Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, *80*(6), 69-96.
Leyten, J. L., & Boerstra, A. C. (2003). Comparison of productivity and absenteeism effects if indoor environmental quality in offices. In *proceedings of Healthy Buildings Conference 2003,* Singapore*,* 3, 319-323.

Linz, S. J., & Semykina, A. (2012). What makes workers happy? Anticipated rewards and job satisfaction. *Industrial Relations: A Journal of Economy and Society*, *51*(4), 811-844.
Locke, E. A. (1976). The nature and causes of job satisfaction. *Handbook of industrial and organizational psychology*.
Lyubomirs, S. (2001). Why are some people happier than others? The role of cognitive motivational processes in well-being. *American Psychology*, 56 (3), 239–249.
MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological methods*, 4(1), 84.
Malhotra, M. K., & Grover, V. (1998). An assessment of survey research in POM: from constructs to theory. *Journal of operations management*, *16*(4), 407-425.
Mawson, A. (2002),The Workplace and its Impact on Productivity, 4.Publication No. 8, Advanced Workplace Associates, London, available at: www.occupier.org
Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. the MIT Press.
Mendat, C. C., Ward, S. J., & Wogalter, M. S. (2004). Human Factors in Public Facilities Design. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 48, No. 8, pp. 1126-1128). Sage CA: Los Angeles, CA: SAGE Publications.
Miles, A.K. (2000). *The ergonomics and organizational stress relationship*. PhD thesis, Florida State University School of Business, micro. 9994574.
Namkung, Y., & Jang, S. (2008). Are highly satisfied restaurant customers really different? A quality perception perspective. *International Journal of Contemporary Hospitality Management*, *20*(2), 142-155.
Nunnally, J. C. (1967). *Psychometric Theory*. McGraw-Hill, New York, p. 355.
Office of Real Property (1999). *Workplace evaluation study: Introducing the cost per person model*, US government Office of Government wide Policy, Washington, DC.
Ohly, S., Sonnentag, S., Niessen, C., & Zapf, D. (2010). Diary studies in organizational research. *Journal of Personnel Psychology*.
Orr, J. M., Sackett, P. R., & Dubois, C. L. (1991). Outlier detection and treatment in I/O psychology: A survey of researcher beliefs and an empirical illustration. *Personnel Psychology*, 44(3), 473-486.
Oseland, N. (1999). Environmental factors affecting office worker performance: a review of evidence. Chartered Institution of Building Services Engineers.
Oseland, N. (2004). Occupant feedback tools of the office productivity network. *Assessed http://www.officeproductivity.co.uk*.
Pallant, J., & Manual, S. S. (2010). *A step by step guide to data analysis using SPSS*. Berkshire UK: McGraw-Hill Education.
Raajpoot, N. A. (2002). TANGSERV: A multiple item scale for measuring tangible quality in foodservice industry. *Journal of Foodservice Business Research*, *5*(2), 109-127.
Rattray, J., & Jones, M. C. (2007). Essential elements of questionnaire design and development. *Journal of clinical nursing*, 16(2), 234-243.
Reis, H. T. (1994). Domains of experience: Investigating relationship processes from three perspectives. *Theoretical frameworks for personal relationships*, 87-110.
Reis, H. T., & Gable, S. L. (2000). Event-sampling and other methods for studying everyday experience. *Handbook of research methods in social and personality psychology*, 190-222.
Rindfleisch, A., Malter, A. J., Ganesan, S., & Moorman, C. (2008). Cross-sectional versus longitudinal survey research: Concepts, findings, and guidelines. *Journal of marketing research*, *45*(3), 261-279.
Robertson, I. T., & Cooper, C. L. (2010). Full engagement: The integration of employee engagement and psychological well-being. *Leadership & Organization Development Journal*, *31*(4), 324-336.
Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological bulletin*, *128*(6), 934.
Rolloos, M. (1997). *Een gezond binnenmilieu betaalt zichzelf terug* Praktijkboek Gezonde Gebouwen.
Ryu, K., & Jang, S. (2008). DINESCAPE: A scale for customers' perception of dining environments. *Journal of Foodservice Business Research*, *11*(1), 2-22.
Sadikaj, G., Moskowitz, D. S., & Zuroff, D. C. (2011). Attachment-related affective dynamics: differential reactivity to others' interpersonal behavior. *Journal of personality and social psychology*, *100*(5), 905.
Sailer, K., Pomeroy, R., & Haslem, R. (2015). Data-driven design—Using data on human behaviour and spatial configuration to inform better workplace design. *Corporate Real Estate Journal*, *4*(3), 249–262.
Scarpello, V., & Campbell, J. P. (1983). Job satisfaction: Are all the parts there?. *Personnel psychology*, *36*(3), 577-600.

Sell, L., & Cleal, B. (2011). Job satisfaction, work environment, and rewards: Motivational theory revisited. *Labour*, *25*(1), 1-23.
Shiffman, S. (2000). Real-time self-report of momentary states in the natural environment: Computerized ecological momentary assessment. *The science of self-report: Implications for research and practice*, 277-296.
Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Reviews Clinical Psychology*, *4*, 1-32.
Shign (2019). *About Shign*. Retrieved from http://www.shign.nl/about-us/
Shrout, P. E., & Lane, S. P. (2012). Psychometrics. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 302-320). New York, NY, US: Guilford Press.
Sink, D. S. (1985). Productivity management: planning, measurement and evaluation, control, and improvement. John Wiley & Sons.
Sirgy, M. J., Efraty, D., Siegel, P., & Lee, D. J. (2001). A new measure of quality of work life (QWL) based on need satisfaction and spillover theories. *Social indicators research*, *55*(3), 241-302.
Sit, J., Merrilees, B., & Birch, D. (2003). Entertainment-seeking shopping centre patrons: the missing segments. *International Journal of Retail & Distribution Management*, *31*(2), 80-94.
Skalli, A., Theodossiou, I., & Vasileiou, E. (2008). Jobs as Lancaster goods: Facets of job satisfaction and overall job satisfaction. *The Journal of Socio-Economics*, *37*(5), 1906-1920.
Smith, P. C. (1969). The measurement of satisfaction in work and retirement: A strategy for the study of attitudes.
Sonnentag, S. (2003). Recovery, work engagement, and proactive behavior: a new look at the interface between nonwork and work. *Journal of applied psychology*, *88*(3), 518.
Sousa-Poza, A., & Sousa-Poza, A. A. (2007). The effect of job satisfaction on labor turnover by gender: An analysis for Switzerland. *The Journal of Socio-Economics*, *36*(6), 895-913.
Spector, P. E. (1997). Job satisfaction: Application, assessment, causes, and consequences (Vol. 3). Sage publications.
Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NS: Erlbaum.
Sulek, J. M., & Hensley, R. L. (2004). The relative importance of food, atmosphere, and fairness of wait: The case of a full-service restaurant. *Cornell Hotel and Restaurant Administration Quarterly*, *45*(3), 235-247.
Sutermeister, R. A. (1976). *People and productivity*. NY: McGraw-Hill Book Company. 2nd ed.
Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 Well-Being Index: a systematic review of the literature. *Psychotherapy and psychosomatics*, *84*(3), 167-176.
Türen, U., Erdem, H., & Çamoğlu, A. (2017). The effects of perceived quality of organizationally provided meal service on employees' job performance and mediating role of organizational identification. *Journal of Workplace Behavioral Health*, *32*(1), 49-71.
United Nations (2016). World toilet day 2016: toilets and jobs. Retrieved from http://www.un.org/en/events/toiletday/assets/img/posters/fact_sheet_toiletsandjobs_EN_3.pdf
Van den Brink, M., Bandell-Hoekstra, E. N. G., & Abu-Saad, H. H. (2001). The occurrence of recall bias in pediatric headache: a comparison of questionnaire and diary data. *Headache: The Journal of Head and Face Pain*, *41*(1), 11-20.
Van Der Voordt, T. J. (2004). Productivity and employee satisfaction in flexible workplaces. *Journal of Corporate Real Estate*, *6*(2), 133-148.
Van Hooff, M. L., Geurts, S. A., Kompier, M. A., & Taris, T. W. (2007). " How fatigued do you currently feel?" Convergent and discriminant validity of a single-item fatigue measure. *Journal of Occupational Health*, *49*(3), 224-234.
Vankamamidi, N. R. (2004). Assessment of user satisfaction of restrooms with existing toilet fixtures and new low consumption fixtures (Doctoral dissertation, Texas A&M University).
Van Laar, D., Edwards, J. A., & Easton, S. (2007). The Work-Related Quality of Life scale for healthcare workers. *Journal of advanced nursing*, *60*(3), 325-333.
Vischer, J. C. (2007). The effects of the physical environment on job performance: towards a theoretical model of workspace stress. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *23*(3), 175-184.
Vischer, J. C. (2008). Towards an environmental psychology of workspace: how people are affected by environments for work. *Architectural science review*, *51*(2), 97-108.
Vroom, V. H. (1964). *Work and motivation*. 1964. *NY: John Wiley & Sons*, *45*.

Warren, C. R., & Landis, R. S. (2007). One is the loneliest number: A meta-analytic investigation on single-item measure fidelity. *Ergometrika*, *4*(1), 32-53.

Weiss, H. M., Nicholas, J. P., & Daus, C. S. (1999). An examination of the joint effects of affective experiences and job beliefs on job satisfaction and variations in affective experiences over time. *Organizational behaviour and human decision processes*, *78*(1), 1-24.

Westerkamp, T. A. (2000). *Two-for-one restroom savings. Maintenance Solutions*. Retrieved from http://www.facilitiesnet.com/ms/Jul00/jul00housekeeping.shtml

Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic commerce research and applications*, *9*(6), 476-487.

Williams, P. G. (2009). Foodservice perspective in institutions. *Meals in science and practice*, 50-65.

Yin, R.K., (1984). *Case Study Research: Design and Methods*. Beverly Hills, Calif: Sage Publications.

Zainal, Z. (2007). Case study as a research method. *Jurnal Kemanusiaan*, 5(1).

Appendices

Appendix A – Questionnaire (English version)

All questions are multiple choice, and have a 7-point Likert scale:
(1) very poor, (2) poor, (3) fair, (4) almost moderate, (5) moderate, (6) good, (7) very good

Restroom

1. The accessibility of the restroom (e.g. proximity, accessibility)
2. The cleanliness of the restroom (e.g. neat, tidy, hygienic)
3. The privacy of the restroom (e.g. hearing or smelling other people)
4. The odour of the restroom (e.g. fresh, unpleasant)
5. The hygiene amenities in the restroom (e.g. availability of soap, toilet paper, towels)
6. The décor of the restroom (e.g. furniture, colour use, atmosphere)
7. The waiting line for the restroom

Restaurant

1. The variation in the catering offer (e.g. today's options, different options throughout the year)
2. The quality of the food (e.g. fresh, healthy, tasty)
3. The acoustics in the restaurant (e.g. background noise)
4. The cleanliness of the interior (e.g. neat, fresh, tidy)
5. The friendliness of the catering staff (e.g. kind, polite)
6. The quality of the furniture (e.g. comfort)
7. The seating options to eat and drink (e.g. number, variety)
8. The queue for service and self-service (e.g. at the cash register and distribution points)
9. The attractiveness of the layout (e.g. table and seating arrangement, spaciousness, view)
10. The atmosphere to enjoy your meal and drinks (e.g. colour use, decoration, lighting, odour)

Workspace

1. The availability of workspaces (e.g. number)
2. The possibility to work at different workplaces for different types of activities (i.e. flexible workplaces)
3. The layout of the workplace (e.g. uncluttered, calm, safe, room to move)
4. The interior design of the workplace (e.g. comfortable furniture, colour use, atmosphere)
5. The sound in the room (e.g. background noise, distractions)
6. The cleanliness of the room (e.g. neat, fresh, tidy)
7. The indoor climate at work (e.g. change temperature, open a window)

Questions asked at the end of each survey

These are all multiple choice questions with a 7-point Likert scale: (1) totally disagree, (2) disagree,
(3) somewhat disagree, (4) neither agree nor disagree, (5) somewhat agree, (6) agree, (7) totally agree.

Productivity

- Overall today's work environment contributes to my productivity
- I am struggling to finish my work when I am in the office

Job satisfaction

- My today's work environment makes me feel satisfied about my job
- My work is pleasant thanks to my today's work environment

Well-being

- At this moment I am tired and stressed
- At this moment I feel calm and relaxed
- At this moment I feel cheerful and in good spirits

Appendix B – Questionnaire (Dutch version)

Het zijn allemaal meerkeuze vragen met een 7-punts Likert schaal:
(1) zeer slecht, (2) slecht, (3) matig, (4) bijna voldoende, (5) voldoende, (6) goed, (7) zeer goed.

Toiletruimte

1. De toegankelijkheid van de toiletruimte (bijv. nabijheid, bereikbaarheid)
2. De netheid van de toiletruimte (bijv. schoon, opgeruimd, hygiënisch)
3. De privacy in de toiletruimte (bijv. hinder van anderen)
4. De geur in de toiletruimte (bijv. fris, aangenaam)
5. De hygiënevoorzieningen (bijv. zeep, toiletpapier, handen drogen)
6. De inrichting van de toiletruimte (bijv. meubilair, kleurgebruik, sfeer)
7. De wachtrij voor de toiletruimte

Restaurant

1. De variatie in het aanbod (bijv. keuze, bereidingswijze, portionering)
2. De kwaliteit van het eten (bijv. vers, gezond, lekker)
3. Het geluid in het restaurant (bijv. lawaai, achtergrondgeluid)
4. De netheid van het restaurant (bijv. netjes, fris, hygiënisch)
5. De vriendelijkheid van het personeel (bijv. aardig, beleefd, behulpzaam)
6. De kwaliteit van het meubilair (bijv. comfort, esthetiek)
7. De zitmogelijkheden in het restaurant (bijv. hoeveelheid, variatie)
8. De wachtrijen in het restaurant (bijv. bij het buffet, voor de kassa)
9. De indeling van het restaurant (bijv. meubelopstelling, ruimtelijk, aantrekkelijk)
10. De sfeer in het restaurant (bijv. kleurgebruik, decoratie, verlichting)

Werkplek

1. De beschikbaarheid van werkplekken (bijv. aantal)
2. De mogelijkheid om op verschillende plekken te werken (bijv. stilteruimte, flexplek)
3. De indeling van je werkplek (bijv. ruimtelijkheid, afstand tot collega's, loopruimte)
4. Het interieur van je werkplek (bijv. meubilair, kleurgebruik, sfeer)
5. De akoestiek op je werkplek (bijv. achtergrondgeluid, lawaai, afleiding)
6. De netheid van de ruimte (bijv. schoon, fris, opgeruimd)
7. Het binnenklimaat op je werkplek (bijv. temperatuur, ventilatie, frisse lucht)

Vragen die zijn gesteld aan het einde van de enquête
Dit zijn allemaal meerkeuze vragen met een 7-punts Likert schaal: (1) helemaal mee oneens, (2) mee oneens, (3) enigszins mee oneens, (4) neutraal, (5) enigszins mee eens, (6) mee eens, (7) helemaal mee eens.

Productiviteit

- De huidige werkomgeving draagt bij aan mijn productiviteit
- Ik heb moeite om mijn werk af te krijgen in de huidige werkomgeving

Tevredenheid

- De huidige werkomgeving draagt eraan bij dat ik tevreden ben met mijn baan
- Mijn werk is aangenaam in de huidige werkomgeving

Welzijn

- Ik ben op dit moment moe en gestrest
- Ik voel me op dit moment kalm en ontspannen
- Ik voel mij op dit moment vrolijk en goedgehumeurd

Appendix C – Communication with respondents

User information EMA
"Deze app is onderdeel van een onderzoek naar de werkplekbeleving binnen a.s.r. In de bijlage vind je de handleiding voor de installatie van de app. Wanneer u de app succesvol heeft geïnstalleerd, is deze klaar voor gebruik! De vragen die u zo meteen gaat invullen gaan over drie specifieke ruimtes binnen het kantoorgebouw: uw werkplek, het restaurant en de toiletruimte. Het is de bedoeling dat u alleen de vragen beantwoordt over uw huidige omgeving. Het is dus niet te bedoeling dat u vragen beantwoordt over het restaurant als u daar de hele dag nog niet bent geweest. Om die reden zal de enquête u eerst vragen welke ruimte u wilt beoordelen.

De data zal worden verzameld tijdens een periode van twee weken (1 maart laatste dag). U beantwoordt dezelfde vragen dus meerder keren. Het zou dus zomaar kunnen dat u tijdens deze twee weken wel 10x het restaurant zal beoordelen. Dit lijkt veel, maar gelukkig duurt het invullen van de enquête maar +/- 1 minuut. Het is namelijk zo: hoe vaker iedereen de enquête beantwoordt, hoe nauwkeuriger de data is. Ik zou u dus willen vragen om de app elke dag te gebruiken. Hoe vaker u de vragen beantwoordt, hoe beter. Dit mag dus ook meerdere keren per dag. Het is hierbij goed om te weten dat er betrouwbaar met uw gegevens om zal worden gegaan en de resultaten geheel anoniem verwerkt worden."

User information cross-sectional survey
"Deze enquête is onderdeel van een onderzoek naar de werkplekbeleving binnen a.s.r. De vragen die u zo meteen gaat invullen gaan over vier specifieke ruimtes binnen het kantoorgebouw: uw werkplek, het restaurant, de toiletruimte en de buitenwerkplekken. Het invullen van de enquête duurt 3-5 minuten. Er zal betrouwbaar met uw gegevens worden omgegaan en de resultaten worden geheel anoniem verwerkt. De enquête kunt u vinden via onderstaande link: https://wur.az1.qualtrics.com/jfe/form/SV_b7dojW5n3MRpIX3"

Introduction of the cross-sectional survey
"Beste meneer/mevrouw, Ik ben een masterstudent aan de Wageningen University & Research en voor mijn afstuderen doe ik onderzoek naar de werkplekbeleving binnen a.s.r. De vragen die u zo meteen gaat invullen gaan over vier specifieke ruimtes binnen het kantoorgebouw: uw werkplek, het restaurant, de toiletruimte en de buitenwerkplekken. Het invullen van de enquête duurt 3-5 minuten. Er zal betrouwbaar met uw gegevens worden omgegaan en de resultaten worden geheel anoniem verwerkt. Alvast bedankt voor het beantwoorden van de vragen. Met vriendelijke groet, Sara Woltjes"

Message to EMA respondents: 22-02-2019 at 13.45
"Beste respondenten, bedankt voor jullie medewerking tot nu toe! Er waren opmerkingen over de afsluitende vragen na het beoordelen van werkplek, restaurant of sanitair. Deze gaan over productiviteit, werktevredenheid en welzijn. Ik wil o.a. weten hoe dit wordt beïnvloed door jullie werkomgeving. Vandaar dat ze telkens terugkomen. Groet, Sara."

Message to EMA respondents: 28-02-2019 at 09.15
"Beste respondenten, bedankt voor de feedback tot nu toe en vooral ook voor de opmerkingen die jullie geven. Deze worden doorgegeven aan Facility. We gaan nog een paar dagen door met het verzamelen van feedback, dus blijf jullie mening geven s.v.p. Met name de fluctuaties per dag zijn interessant. Groet, Sara."

Message to EMA respondents: 04-03-2019 at 09:00
"Beste respondent, deze hele (carnavals)week kun je nog je dagelijkse ervaring met je werkplek, het sanitair en het restaurant beoordelen. Vrijdag 8 maart om 1700 sluit ik de vragenlijst. Groet, Sara."

Message to EMA respondents: 08-03-2019 at 08:55
"Beste respondent, vandaag is de laatste dag dat je je werkplek, het restaurant en het sanitair kunt beoordelen door middel van de Shign app. Vanaf volgende week worden de resultaten verwerkt. Ik zie graag uit naar jullie mening! Groet, Sara."

Appendix D – Additional comments of respondents

Comments of EMA respondents

- Werkplek is te koud
- De laatste vragen komen steeds terug... dat is een irritatiefactor in het onderzoek
- Ik mis de mogelijkheid om mijn bureau op sta stand te zetten
- De toilet, hier wordt niet gewerkt. Dus de helft van de vragen is niet van toepassing
- Ik werk niet in het restaurant dus helft vragen nvt
- Ik neem altijd mijn eigen 'bammetjes' mee.
- ik begon met invullen van de vragen over toiletruimte. Na een aantal vragen ging het ineens over de werplek (of die bijdraagt aan mijn productiviteit etc). Ik denk dat hier overal toiletruimte bedoeld werd ipv werkplek. Foutje in de vragenlijst?
- Ik heb geen flexplek omdat ik op een secretariaat zit
- Toiletruimte automaat papier handen drogen was vastgelopen. Is regelmatig het geval op P1C15
- Toilet was net schoon gemaakt
- Teveel vragen
- Waarom de laatste paar vragen als ik het sanitair wil evalueren?
- Handdoekjes op
- Beetje koud vandaag
- Mijn werkplezier staat los van mijn omgeving en heeft betrekking op mijn werkzaamheden.
- Werk en omgeving staat voor mij los van elkaar
- Hoe dun kan toiletpapier zijn wil het nog goedkoop zijn gebruik nu bijna een halve rol
- Voet van monitor is nog steeds stoffig. Toetsenbord ook. Achterkant van bureau lijkt dus niet te worden schoon gemaakt.
- Mijn werk en mijn omgeving staan los van elkaar.
- Bij de saladebar lopen mensen elkaar geregeld in de weg. Er is geen duidelijke looprichting. Wellicht zou het helpen als er wel een looprichting is (bijv. saladebakjes vooraan ipv in het midden).
- Hing een hele vieze geur op het toilet. Ik gok dat het met de hygiëne boxen te maken had, maar ben we niet zeker van. Gebeurt overigens zelden!
- Maandag. Drukke dag. Ik wilde graag gebruik maken van een sta bureau. Alles bezet. Gezien mijn klachten heb ik dat echt nodig. Moet je vragen of er iemand plaats wil maken. Dat vind ik lastig omdat het al zo vol is
- Banken in restaurant zijn doorgezakt.
- Salade buffet zou meer mogen variëren. Nu iedere dag bijv. aardappelsalade. Wellicht afwisselen met gegrilde groenten, champignons of iets dergelijks?"
- In het restaurant is maar 1 tappunt voor water en de straal is heel dun. Daardoor staat er meestal een rij, dat is een beetje vervelend.
- G3 F7 werkt de unit op het bureau niet om de televisie te bedienen
- Nog steeds stoffige voetstuk monitor en toetsenbord. Achterste deel bureau wordt kennelijk niet schoongemaakt.
- Werkplek nog steeds stoffig. Inmiddels zelf maar schoongemaakt.
- Soms staan er dames te make-up'en en te kletsen op het toilet. Dat was bet ook het geval. Dat is ok hoor, maar maakt dat ik in deze survey bij privacy een 'mee oneens' heb ingevuld.
- De dyson kranen zijn erg fijn.
- Zou fijn zijn als er nog wat meer ruimte is om telefoon of laptop neer te leggen in wc, want je mag ze niet onbeheerd achterlaten
- Vandaag, donderdag, best druk op de toiletten (P1)
- De plek in het restaurant waar de tafels staan en de lampen aan het plafond hangen is niet goed afgestemd. Als je op staat stoot je daardoor je hoofd aan de laaghangende lamp.
- In de fruitschalen ligt standaard teveel appels en peren. Is het mogelijk om de verdeling van fruit aan te passen. Meer fruit wat vaker gegeten wordt en minder van wat altijd overblijft.

Comments of cross-section respondents

- De vergaderzalen op de 6e aan de zuidzijde zijn vanaf het voorjaar veel te heet. Wanneer wordt daar een keer wat aan gedaan? Daarnaast blijft het heel raar dat als het mooi weer is, de zonwering naar beneden gaat en je binnen in een vergaderruimte onder een lamp zit die niet uit kan. Dit is met name vreemd bij vergaderzalen die aan de rand liggen. Het licht is te fel en het komt niet duurzaam over."
- Restaurant zou meer met dieet restricties kunnen doen bv melkeiwitintoleranties en maken op verzoek met weglaten van bepaalde ingrediënten.
- Daar waar ik heb aangegeven gestrest te zijn heeft dit niets met a.s.r. te maken. Lijkt me voor deze enquête zinvol om te weten.
- Inzake de ''wachtrij voor het toilet'' heb ik matig aangegeven omdat de schoonmaakster vaak rond 12:00 uur de wc's gaat schoonmaken terwijl het dan spitsuur is i.v.m. pauze. Is geen handige timing om de toiletten niet beschikbaar te stellen. "
- Gevoelstemperatuur wisselt, sommige plekken (kantine) te koud en soms te warm (werkplek)
- De vragen over met name het restaurant beantwoord ik positief. Dat komt met name omdat ik om 11.45 uur eet. Dat is een rustig tijdstip en het aanbod is dan volledig.
- Ik eet meestal warm. Dan betaal ik eerst bij de kassa en ga naderhand met de kassabon het eten ophalen. Op deze manier wordt het eten niet koud als je in een eventuele wachtrij moet staan.
- We werken met flexplekken, onhandig is dat je iedere werkdag je werkplek opnieuw moet instellen. Het lukt niet altijd om je stoel in de gewenste positie te krijgen. Als je hier gevoelig voor bent, is het wel vervelend.
- Wat ik mis aan mijn werkplek is het standaard aanwezig zijn van kabels om mijn laptop aan vast te leggen.
- Dat ik me nu even niet goed voel heeft niets met mijn werk of werkplek te maken.
- Ik ben niet erg tevreden met de financiële regelingen m.b.t. het thuiswerken, de vergoeding voor de thuiswerkplek is in feite te laag waardoor een goede ergonomie niet gewaarborgd kan worden. Je moet er wel voor tekenen dat het in orde is, anders mag je niet eens thuiswerken. Een beetje krom.
- Ik hoop dat er naar aanleiding van deze enquête goed onderzoek wordt gedaan naar de luchtkwaliteit en temperatuur van afdelingen
- Het is koud
- Het is warm
- Er is een constante luchtstroom (waar komt deze vandaan als de ramen niet open staan)
- In de zomer / warmer weer is de luchtkwaliteit matig"
- Gezond eten goedkoper, snacks duurder.
- Zijn de papieren 'handdoeken' in de toiletten nou echt duurzaam?
- Jammer dat de parkeergelegenheid en het parkeerbeleid niet als vierde onderwerp is meegenomen.
- Waarom aandacht voor bijzondere dingen als buitenwerkplek. Worden maar weinigen blij van terwijl er is/ wordt bezuinigd op dingen waar we dagelijks mee worden geconfronteerd. Denk aan de handdoeken die stuk gaan als je met natte handen er aan trekt, houten vorkjes die niet lekker zijn, weinig bekers bij koffiemachines, betere koffie, etc.
- I.p.v. teams op vaste plekken zetten zou ik graag zelf op verschillende plekken zitten afhankelijk van het weer en het soort onderhanden werk.
- Aanvulling op vragen Restaurant, vind ik te duur!
  Aanvulling op vraag werkplekken, te dicht op elkaar, geen privacy en erg gehorig.
- Mis vragen over de werksfeer van jou(mijn) organisatie. Die vind ik erg zakelijk geworden, er is weinig gevoel meer voor elkaar. Boventalligheid hangt je voortdurend boven het hoofd zonder duidelijke voorafgaande individuele gesprekken van leidinggevende of directie. Mededelingen worden gedaan in een zogenaamd leuke medewerkersbijeenkomst dat vind ik erg onpersoonlijk. Dit creëert onderlinge spanning en speculaties.
- Het zou heel fijn zijn als er meer werkplekken zouden komen zodat we dichter bij elkaar kunnen zitten.
- Restaurant meer en tijdige aanvulling van de groenten. Werkplekken vaker schoonmaken muis en toetsenborg
- De 3 vragen over hoe ik mij voel, zou ik persoonlijk (als iemand die ook onderzoek heeft gedaan voor zijn master) niet koppelen aan de werkplek. Ik heb 3x neutraal gegeven omdat het er voor mij los van staat. Er spelen voor veel mensen hele andere zaken namelijk...
- De temperatuur op kantoor wisselt enorm per dag.

- De variatie in het restaurant is op zich best goed, alleen afhankelijk van tijdstip, vaak zijn er dingen al op en worden dan niet meer bijgevuld."
- In een kantoortuin word je erg afgeleid. Als ik een dag thuis werk, ben ik veel productiever qua verrichting van inhoudelijke werkzaamheden. Anderzijds wordt er in een kantoortuin beter gecommuniceerd met andere collega's (kortere communicatielijnen) en de cultuur/sfeer heeft een meer open karakter, dan in een kantoor met aparte ruimtes.
- Leuke vragen over de werkplek. Over het algemeen is de inrichting strak en modern en soms een beetje te saai. Het mag dus wel wat vrolijker.
- Sinds de renovatie komt er geen warm water meer uit de kraan. Dat mis ik erg. Verder is het een prachtig gebouw met een prettig werkklimaat.
- Het klimaat op P5A#9 is altijd te koud zomer of winter.
- Helaas maar 1 urinoir op herentoiletten (op b.v. Paars 3), waar ruimte voor 2 is, dat zou mijn oordeel in goed veranderen
- De rauwkostsalades in het restaurant zijn vrij prijzig. Het wordt op die manier niet gestimuleerd om de medewerkers gezonder te laten eten.
- Het voordeel van a.s.r. is dat je overal kan werken, dus ook thuis. Thuis ben ik heel productief, maar is minder "gezellig". Het grootste nadeel is momenteel dat er te weinig werkplekken zijn voor iedereen in het team. Als je vroeg komt heb je plek. Op de drukke dagen (dinsdag en donderdag) zit het hele team verspreid. Dan heb je niet het gevoel een team te zijn,
- De kantoortuin waar we inzitten is heel gezellig, maar heeft als nadeel geluidsoverlast van collega's, waardoor concentratie soms een probleem kan zijn."
- De algehele klimaatbeheersing mag wel verbeterd worden. Er zijn zeer grote verschillen in temperatuur en luchtstroom op de werkvloer.
- Ik werk met dubbel beeldscherm wat ik prettig vind. Minder fijn is de warboel aan draden die je in moet pluggen

Appendix E – Lavaan code for SEM

```
#Load the libraries
library("foreign")
library("lavaan")
library("semPlot")
library(dplyr)
library(tidyr)
library(knitr)

#Set working directory
setwd("C:/Users/Sara")

# Load the data
dataset <- read.spss("SPSS cross sectie data.sav", to.data.frame = T, header=T)

# Translate Likert scale from text to numbers
dataset[,4:34]<- sapply(dataset[,4:34],as.numeric)

#Specify the model
workplace.model <- '

# Latent variables (employee)
   Prd =~ Productiviteit + WerkAfkrijgen
   Sat =~ JobSatisfaction + WorkisPleasant
   WB =~ TiredStressed + CalmRelaxed + CheerfulGoodSpirits

# Latent variables (environment)
    Work =~ A2_1 + A2_2 + A2_3 + A2_4 + A2_5 + A2_6 + A2_7
    WC =~ A3_1 + A3_2 + A3_3 + A3_4 + A3_5 + A3_6 + A3_7
    Rest =~ A1_1 + A1_2 + A1_3 + A1_4 + A1_5 + A1_6 + A1_7 + A1_8 + A1_9 + A1_10

# Regressions
   Sat ~ WC + Rest + Work
   WB ~ WC + Rest + Work
   Prd ~ WB + Sat + Work'

#Specify the improved model
workplace.model1 <- '

# Latent variables (employee)
   Prd =~ Productiviteit + WerkAfkrijgen
   Sat =~ JobSatisfaction + WorkisPleasant
   WB =~ TiredStressed + CalmRelaxed + CheerfulGoodSpirits

# Latent variables (environment)
    Work =~ A2_1 + A2_2 + A2_3 + A2_4 + A2_5 + A2_6 + A2_7
    WC =~ A3_1 + A3_2 + A3_3 + A3_4 + A3_5 + A3_6 + A3_7
    Rest =~ A1_1 + A1_2 + A1_3 + A1_4 + A1_5 + A1_6 + A1_7 + A1_8 + A1_9 + A1_10

# Regressions
   Sat ~ WC + Rest + Work
   WB ~ WC + Rest + Work
   Prd ~ WB + Sat + Work

# Model changes
CalmRelaxed ~~ 2.036551*CalmRelaxed
A1_1 ~~ A1_2  # RESTAURANT: Food variation & Food quality
A2_1 ~~ A2_3  # WORKSPACE: Workspace availability & Workspace layout
A1_7 ~~ A1_8  # RESTAURANT: Seating possibilities & Queues
A3_1 ~~ A3_7  # RESTROOM: Accessibility & Queues
A1_9 ~~ A1_10 # RESTAURANT: Layout & Atmosphere
A1_3 ~~ A1_9'  # RESTAURANT: Acoustics & Layout

# CFA
workplace.fit <- cfa(model = workplace.model, data = dataset, std.lv=T)
workplace.fit1 <- cfa(model = workplace.model1, data = dataset, std.lv=T)
```

```
# Summary
summary(workplace.fit1, standardized=T, fit.measures=T, rsquare=T)

# Create the path diagram
semPaths(object = workplace.fit1, edge.label.cex = 0.7, layout="tree", rotation=4, whatLabels = "std")

# Modification indices
modificationindices(workplace.fit1, sort=T)

# Chi-Square comparisson
anova(workplace.fit, workplace.fit1)

# Model fit comparisson
fitmeasures(workplace.fit, c("aic", "ecvi", "TLI", "GFI", "RMSEA"))
fitmeasures(workplace.fit1, c("aic", "ecvi", "TLI", "GFI", "RMSEA"))

# Table with factor loadings
parameterEstimates(workplace.fit1, standardized=TRUE) %>%
  filter(op == "=~") %>%
  select('Latent Factor'=lhs, Indicator=rhs, B=est, SE=se, Z=z, 'p-value'=pvalue, Beta=std.all) %>%
  kable(digits = 3, format="pandoc", caption="Factor Loadings")
```

Appendix F – R code for LMM

```
#Load the libraries
library(lme4)

#Set working directory
setwd("C:/Users/Sara")

# Load the data
dataset <- read.csv("LMM_EMA2.csv", header=T)

# Restroom --> Job satisfaction & productivity
outcome.null <- lmer(OUT_JSenPRD ~ WC_Design + WC_Clean + (1|Nummer), data = dataset)
summary(outcome.null)
outcome.adj <- lmer(OUT_JSenPRD ~ WC_Clean + (1|Nummer), data = dataset)
summary(outcome.adj)
anova(outcome.null, outcome.adj)

# Restroom --> WB
outcome.null <- lmer(OUT_WB ~ WC_Design + WC_Clean + (1|Nummer), data = dataset)
summary(outcome.null)
outcome.adj <- lmer(OUT_WB ~ WC_Clean + (1|Nummer), data = dataset)
summary(outcome.adj)
anova(outcome.null, outcome.adj)

# Restroom --> Finish work
outcome.null <- lmer(OUT_FINISH ~ WC_Design + WC_Clean + (1|Nummer), data = dataset)
summary(outcome.null)

# Restaurant --> WB
outcome.null <- lmer(OUT_WB ~ RST_AtmosClean + RST_LayoutFurn + RST_FoodServAc + (1|Nummer), data = dataset)
summary(outcome.null)
outcome.adj <- lmer(OUT_WB ~ RST_AtmosClean + RST_FoodServAc + (1|Nummer), data = dataset)
summary(outcome.adj)
anova(outcome.null, outcome.adj)

# Restaurant --> Finish
outcome.null <- lmer(OUT_FINISH ~ RST_AtmosClean + RST_LayoutFurn + RST_FoodServAc + (1|Nummer), data = dataset)
summary(outcome.null)
outcome.adj <- lmer(OUT_FINISH ~ RST_AtmosClean + RST_FoodServAc + (1|Nummer), data = dataset)
summary(outcome.adj)
anova(outcome.null, outcome.adj)

# Restaurant --> Productivity and JS
outcome.null <- lmer(OUT_JSenPRD ~ RST_AtmosClean + RST_LayoutFurn + RST_FoodServAc + (1|Nummer), data = dataset)
summary(outcome.null)
outcome.adj <- lmer(OUT_JSenPRD ~ RST_LayoutFurn + RST_FoodServAc + (1|Nummer), data = dataset)
summary(outcome.adj)
anova(outcome.null, outcome.adj)

# Workspace --> Productivity and JS
outcome.null <- lmer(OUT_JSenPRD ~ WRK_ClimClean + WRK_IntAc + (1|Nummer), data = dataset)
summary(outcome.null)
outcome.adj <- lmer(OUT_JSenPRD ~ WRK_ClimClean + (1|Nummer), data = dataset)
summary(outcome.adj)
anova(outcome.null, outcome.adj)

# Workspace --> WB
outcome.null <- lmer(OUT_WB ~ WRK_ClimClean + WRK_IntAc + (1|Nummer), data = dataset)
summary(outcome.null)

# Workspace --> Finish
outcome.null <- lmer(OUT_FINISH ~ WRK_ClimClean + WRK_IntAc + (1|Nummer), data = dataset)
summary(outcome.null)
outcome.adj <- lmer(OUT_FINISH ~ WRK_IntAc + (1|Nummer), data = dataset)
anova(outcome.null, outcome.adj)
```