university of groningen / faculty of behavioural and social sciences / research master

Master's thesis

The Data is in the Noise: Exploring Variability and Response Behaviour in Ecological Momentary Assessment

Name and initials:	Hesselink, E. S.
Student number:	S2718030
E-mail address:	e.s.hesselink@student.rug.nl
First assessor:	Dr. L. F. Bringmann
Second assessor:	Dr. A. C. Keller
Programme:	Research Master Behavioural and Social Sciences
Theme:	Lifespan Development and Socialization
ECs:	30
Date:	18-5-2024
Word count:	10584

Are there deviations of the Master's thesis from the proposed plan?
☒No
☐Yes, please explain below the deviations

Abstract

The validity in questionnaires is threatened by lack of information of how participants interpret, answer, and are affected by participating in research. Especially in Ecological momentary assessment (EMA) studies, where participants fill out multiple questionnaires a day for a prolonged period of time, understanding response processes and behaviour is vital. This thesis explores the variability within and between participants in Ecological Momentary Assessment (EMA) research, using the subjective concept of happiness. Through a mixed-methods approach combining EMA and semi-structured interviews, the study investigates how participants interpret concepts, choose scores, and are influenced by research participation. In this master thesis, a pilot study of a mixed-method qualitative approach is used to identify these processes. Methods used in this study were EMA surveys supplemented with qualitative questions, combined with qualitative interviews using semi-structured interviews supplemented with cognitive interviewing techniques. A small, but diverse sample of 11 first-year psychology students from 11 different countries (6 different continents) participated in this study, being also diverse in gender and experience in participating in research. This thesis explores the variability within and between participants in Ecological Momentary Assessment (EMA) research, using the subjective concept of happiness. Through a mixed-methods approach combining EMA and semi-structured interviews, the study investigates how participants interpret concepts, choose scores, and are influenced by research participation.

Preface

Before you lies the culmination of my curiosity, enthusiasm, and naivety. This master thesis, like many others, was written with a lot of joy, and an equal amount of tears. At the outset, I was blissfully unaware of the challenges ahead, but now that it is complete, I cannot help but lament that this journey is over.

My curiosity was sparked during my research master's program. Learning about the challenges in science, particularly in behavioural science, such as the complexities of scales, measurement, and interpretation, immediately piqued my interest. I envisioned myself as a knight ready to combat these uncertainties with enthusiasm. Little did I know, I would be more akin to a berserker, but fighting the difficulties nonetheless.

My enthusiasm led me to reach out to my supervisors, Dr. Laura F. Bringmann and Marie Stadel, MSc, who have been nothing but patient, kind, and enthusiastic with me. Like Daedelos, they have given me wings in this project, giving me the freedom to approach this project how I saw fit, but not without necessary guidance and caution.

Despite my initial overconfidence in tackling this ambitious project, their warnings and kindness eventually guided me to refine my focus. They helped me return to the core of my research: interpreting concepts, selecting scores, and understanding the influence of participant engagement in science.

Continuing my gratitude, I want to thank Dr. Marieke Haan and Dr. Dominique F. Maciejewski for their invaluable input in developing the instruments used in this thesis. Their insights were crucial in fine-tuning the instruments, making this project both robust and enjoyable for me and the participants.

Speaking of which, I want to thank the participants for coming to meet me at all the strange locations within the university and for sharing their stories with me, a stranger. I have learned so much from each of you. this study would not have been possible without your contributions. I hope the insights you gained during the study have been helpful, and I wish you happiness in good and bad times.

I also want to thank all the pilot participants. Your input made it possible to see what could and should be changed, while also just giving me a chance to get to know you all better. I hope for all of you, your moderately happy ratings will forever be at a 9.

In this final part of the acknowledgements, I want to thank everyone with whom I have discussed, complained, laughed, and cried during the writing of this thesis. Your support gave me the strength to persevere through the difficult moments. I want to thank two people especially: Sayoni Santara, thank you for being my fellow dancer in sociology and taking me out of the house and the office to work when my motivation was low. You made a lot of the struggles lighter and got me to work on this thesis even if it was only for a little while.

Lastly, I want to thank my fiancé, Robin Bos, for being my rock, my home, my best friend, my believer, and my inspiration. You have seen it all behind the scenes—the good, the bad, the very bad, and the very good. Despite everything, you love me unconditionally. Thank you for all your love and support. I would have been lost without you.

If you've made it through all these heartfelt thank-yous, I thank you as well. I hope you find this thesis an interesting read and that it provides insights useful for your own work, or at least some fun facts to share at parties. I won't keep you any longer.

Officially done,

Mithra Hesselink

AI Statement

This thesis benefited from the assistance of AI tools, specifically ChatGPT 3.5 and Atlas.ti 24.0.0 code suggestions. ChatGPT was employed to enhance the clarity and conciseness of the writing, providing suggestions for refining paragraphs. Additionally, both ChatGPT and Atlas.ti were utilized to propose names for codes by engaging in discussions on relevant concepts. For example, the researcher created the codes “reflective” and “prospective” to denote backward and forward reflection. A discussion concerning whether the term “momentive” could be used for a momentary reflection of happiness was conducted with ChatGPT. Given the positive discussion with the generative AI, the code was incorporated into the codebook. All suggestions were subjected to rigorous evaluation by the researcher. No data was exchanged with either program or its affiliates.

Table of Contents

Introduction

Variety is the spice of life. Each individual's unique journey creates diverse contexts, leading to a variety of confounding factors (Pannucci & Wilkins, 2010). As social and behavioural scientists, understanding this variety in human behaviour is our main point of study. Yet, in our research it is common practice to generalize based on our study results as we want the results to be applicable to a larger group for the findings to be more robust, plausible, and practical (Little, 1993). To promote inference of data, variety is often treated as noise or error in the findings, the unexplainable differences between and within the people participating in our studies (Little et al., 2017; Presser, 1992).

Understanding variety in our studies can lead to more targeted interventions, more concise research practices and more certainty in our findings (Bauer & Schoon, 1993). As the participants in our studies all come from different backgrounds with different cultural, societal, and psychological contexts and frames of reference, it is not always certain if everyone interprets scores and concepts in a similar manner (Von Glasersfeld, 1983). On top of that, how a study might affect someone could also differ greatly due to their personal journey in life (Haynes, 2006). Not only can this differ between people, but also within one person, depending on the (time of) day, the weather, or personal events. In other words: within self-report studies, understanding the subjectivity and reactivity of participants can help reduce the noise in the findings.

This especially applies to longitudinal study formats such as ecological momentary assessment (EMA) (Maher et al., 2024). Also known as Experience Sampling Methods (ESM), EMA has emerged as a powerful tool, allowing researchers to investigate individuals' behaviors, thoughts, and emotions in their natural environments (Shiffman et al., 2008). In EMA studies, participants are requested to fill out multiple questionnaires a day for a prolonged period of time which can span for months or longer. These types of studies allow for in-depth analysis of a participant's emotions, behavioural patterns and the variety within them, making them suitable for a myriad of studies, with or without intervention. As these

studies are conducted over a longer time period and often take place in uncontrolled environments, they are more susceptible to subjectivity and reactivity due to outside influences, or due to the repeated reflection within the study (Maher et al., 2024).

To truly understand the extent of the variety within (EMA) studies, it is important to get a grip on the response processes of the participants. Understanding how participants interpret a concept, what a number on a scale means to them, and how participating in a study influences them not only allows researchers to reduce the noise in their findings, but also to elicit, target, and harness the differences to create more personal interventions and better observational profiles.

In light of these considerations, this master thesis aims to contribute not only to the understanding of measurement reactivity and subjectivity in EMA but also to the broader discourse on the challenges inherent in studying human behavior and experiences. Measurement reactivity and the subjectivity of interpretation and response processes will be observed in a mixed-methods study design combining EMA methods with semi-structured interviews to allow participants to describe their own interpretations, choices, and self-observed influence of participation during this study. This study will serve not only as a master thesis, but also as a pilot study to assess the efficacy to use this design for future studies.

The main subject will be to identify how the context of a participant could affect results in this study. For this purpose, three different areas of EMA survey effects will be assessed, as displayed in the conceptual framework in figure 1. The effects will be studied both as a comparison between different participants, as well as within the same participant over time to allow for a comprehensive overview of measurement reactivity and subjectivity in EMA research.

The first sub subject, placed in the middle, is the interpretation of concepts. From the basis, how a person interprets a concept defines how they will assess and answer any related questions and topics of this concept (Von Glasersfeld, 1983). When differences between participants are large, generalizing the answers of the concept to a larger population will lead

to wrong conclusions, resulting in wrong inferences and possible unwanted effects in policy and interventions. Similarly, within a participant, due to measurement reactivity, the definition might change over the course of the research participation period. This area is understudied, but can have significant impact on a participant's behaviour and therefore, study results.

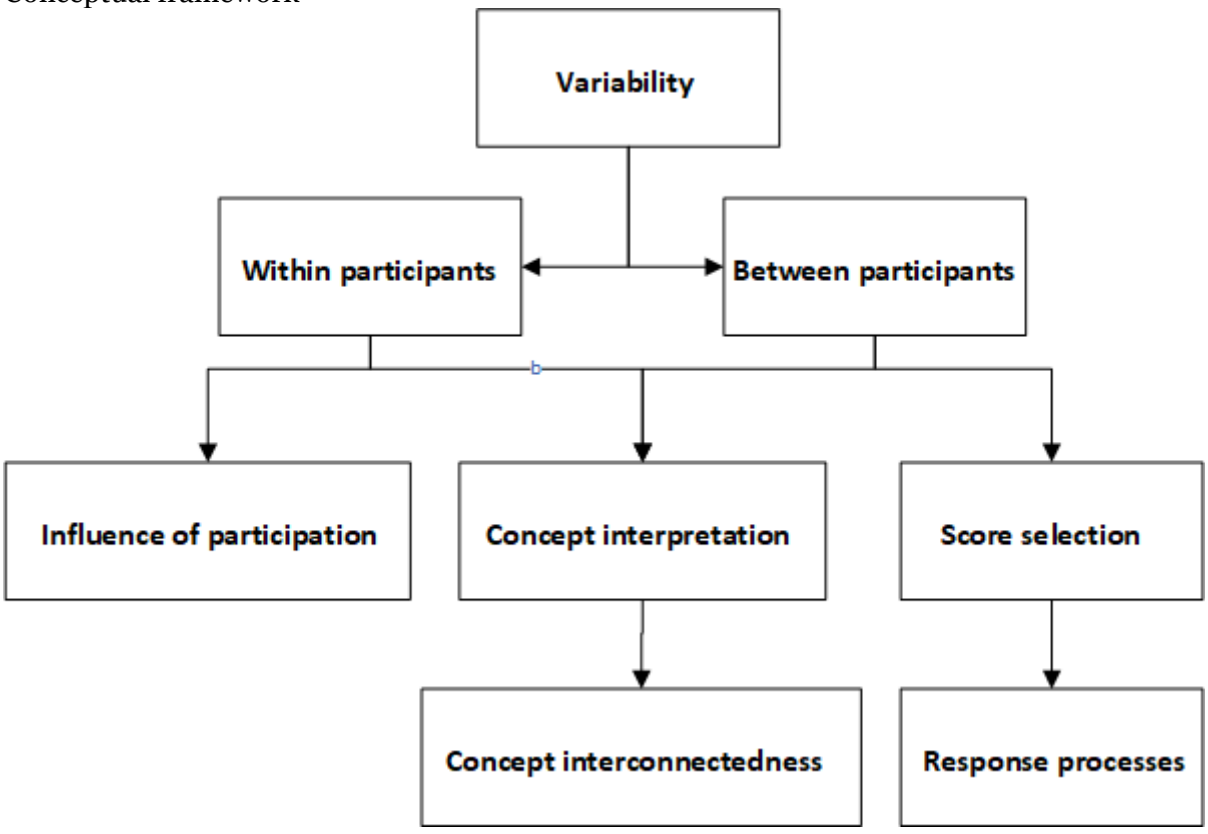
Second, how participants choose an answer, referred to as score selection in this study, will be assessed by identifying what a score means to a participant. Essentially, when two participants select an 8 for a question, does this mean the same thing? Similarly, would different scores receive the same reasoning by different or the same participants? Like concept interpretation, not much has been found in this area. Dawes et al. (2005) found that for perceived exertion scales, participants within different groups, but also in the same group did not fully identify the scale anchors as similar. This sets a precedent for other studies to assess similarity in scales and score selection as well. Within score selection, how participants reason their choice for a score will be assessed from both a behavioural perspective (i.e. how was it written down, retrospective or prospective reasoning) and an attributive perspective. This inductive approach will be tied to the scores themselves, providing insights in participant behaviour and reasoning per score.

Last, a slightly more studied topic is the influence of participation in research. By participating in research and reflecting on a topic the participants' behaviour and interpretation may change. This subject has been studied in specific fields, for example in trauma-related research, finding more reactivity in trauma-related studies (DePrince & Chu, 2008; Ferrier-Auerbach et al., 2009). Participants discussed both negative effects of having to relive trauma, but also increase in awareness of behaviour related to trauma responses, and relief or catharsis by discussing the topic. Similarly, Hayden (2006) found in her discussion with working mothers a sense of gratitude, relief, but also distress by discussing and recalling their experiences after the birth of their child and returning to work. By participating in research, participants can both positively and negatively be affected, giving reason to assess this within other aspects of social science as well.

To assess the subjects, the topic of the study is happiness. Happiness is a concept that is globally known, but not globally interpretated in the same manner. Veenhoven (1984) attempted to collect the known definitions of happiness in his book "*Conditions of happiness*", finding over 20 definitions based on dictionaries, cultural differences, and synonyms. Today, his research has led to a database of happiness (2023) containing more than 16,000 measures of happiness and related concepts. Within these measures, he specifies how complex and subjective happiness can be, making it the perfect candidate topic for this pilot.

On top of the previous discussion, this thesis will also assess briefly whether using happiness as a concept or using a synonym might affect our answers in studies. By using the most commonly used synonyms for happiness in research (well-being, life satisfaction, and feeling happy, Veenhoven, 2023), participants in this study will be able to elaborate on their interpretations of the concepts, allowing an assessment of the interconnectedness of the concepts.

Figure 1
Conceptual framework



Methods

Research design

This mixed-method study with a focus on qualitative analysis consisted of three parts, two interviews and an EMA study to assess variability within and between participants. The study was reviewed and approved by the Ethics committee at the University of Groningen in June 2023 under study code PSY-2223-S-0350.

Materials

In this part, each instrument will be shortly discussed. As the study relied on international participants, all instruments were developed in English. The instruments were developed in an intensive multi-stage pilot phase with multiple pilot testers to assess possible difficulties, misunderstandings, and errors in the instruments. In table 1, an overview of the major outcome variables can be found. A full description of the pilot phases, the full instruments, and a justification per concept within the instruments can be found in the appendix.

Baseline questionnaire

The baseline questionnaire was used as a selection method to obtain a diverse sample of participants. The questionnaire was built in Qualtrics. As Veenhoven's database (2023) indicated differences on different demographical and cultural aspects, selecting participants from different cultural and societal contexts was the main focus while also taking previous participation experience into account. To achieve this, the questionnaire contained questions concerning a participants age, gender, nationality, relationship status, work status, and previous research experience. It started with a question concerning the participants email address for contact in case of selection. Due to the focus on cultural and societal differences and research experience, only age, gender, nationality and previous research experience were used as selection criteria.

Interviews

The interviews conducted were semi-structured interviews as discussed in Hennink et al. (2020), following a guideline but allowing for deviation when participants brought up new topics or discussed a topic earlier than expected. The interviews were conducted by the main researcher, an experienced qualitative interviewer specializing in semi-structured interviews. New to the researcher was cognitive interviewing. Using Willis (1994) guide on cognitive interviewing for questionnaire testing, one section of interview 1 was dedicated to thinking-out-loud interviewing, and one section of interview 3 was dedicated to retrospective probing. Each technique is discussed in its own section.

Interview 1. Interview 1 was a three-part interview concerning experience with or views on research, concept interpretation, and a thinking-out-loud exercise concerning the understanding of the questions in the study.

Views on research. The first part contained questions regarding participants familiarity, preferences, and interpretation with scales and scores. For participants with previous experience in research participation, the focus was on their experience. Participants without experience received questions regarding their views on research.

Concept interpretation. The second part focused on four concepts related to happiness (happiness, well-being, life satisfaction, and feeling happy). Participants were requested to describe how they interpret the concepts individually. This was followed by questions concerning the interconnectedness of the concepts, i.e. how the participant related the concepts to one another, whether they are required for another, and how the participant felt about the concepts being used interchangeably in research.

Thinking-out-loud exercise. The third part concerned the thinking-out-loud exercise to familiarize the participants with the EMA study following the steps described by Willis (1994). First, a short warm-up exercise to familiarize participants with the thinking-out-loud process participants executed a short warm-up exercise to familiarize participants with the thinking-out-loud process. This warm-up exercise consisted of participants visualizing their current living situation and counting the windows within their house, while

discussing any associations and visualizations that came up. After this warm-up exercise, participants were requested to answer all the questions from the EMA study, while speaking their thoughts during the reading of the questions and while selecting an answer. This allowed the participants to see the questionnaire for the first time in a controlled environment, but also gave them the opportunity to ask clarification. The researcher also had the opportunity to provide clarification when a participant misunderstood the question or was unsure whether a certain answer fit within a category.

EMA study

The biggest part of the study was the EMA part.

Schedule. Conducted over a period of seven days, participants were requested to fill out three prompts (in the morning, afternoon, and evening) per day. Participants were able to choose the times most suitable for them to fill out the questionnaires within time frames (morning from waking up until 12:00, afternoon 12:00 – 18:00, and in the evening from 18:00 until sleeping) and as long as at least two hours had passed since the previous prompt. Participants had one hour to fill out the questionnaire and received a reminder after half an hour.

Software. The EMA study in total consisted of four different questionnaires. The questionnaires were built and filled out using m-Path, a platform created by the KU Leuven. Via m-Path the questionnaires can be built online, and then sent to participant's phones using the m-Path app, allowing participants to fill out the questionnaires at any time and location.

Content. During the first interview, participants completed the full questionnaire, while daily prompts were subsets tailored to the time of day. Morning prompts included questions about sleep patterns, evening prompts focused on daily experiences, and afternoon prompts, the shortest, omitted both.

All questionnaires began with a happiness scale question ("At this moment, I feel happy"), followed by an open question for participants to explain their score ("What made you choose the score …?"). This allowed for detailed analysis of their reasoning. Subsequent

questions addressed factors contributing to happiness, such as activities, company, interactions, and location, along with the time-specific queries. Participants also assessed whether completing the questionnaire influenced their initial happiness score ("In the first question, you scored your happiness with .... Do you feel the same?"). If unchanged, the prompt ended; if changed, they assigned a new score on a scale from 1-10, explained the change, and indicated which topics contributed most. This provided insight into the impact of participation and reflection.

Interview 2

The final interview took place after the last day of the EMA study (within 1 – 4 days). This interview also consisted of three parts, a focus on the experience during the EMA study, reflection and possible clarification of answers in the EMA study, and another round of concept interpretation.

In the first part, participants received general questions about their experience in the study, i.e., what questions were difficult or easy to answer, how they went about choosing a score, and whether they felt their week was representative of a typical week in their life. They also received a question concerning the length of the study and whether it was sufficient to gain insights into their happiness.

The second part concerned the given answers during the study. Using retrospective probing (Willis, 1994), participants and the researcher reviewed the trajectory of the EMA scores, and discussed the given answers, reflecting on the occurrences during the study and the factors leading to the scores. Following the reflection, participants were asked whether they would change a score they have given.

The last part once again requested participants to describe and interpret concepts related to happiness (happiness, well-being, life satisfaction, and feeling happy). In this part, the focus was not on whether participants could recall the previously given answers, but to assess whether a participant would describe concepts differently at both a different moment, but also after a week of monitoring their happiness levels.

Both interviews were closed with possibility and space for questions and feedback.

Table 1

Relevant outcome questions per subject

Subject	Instrument	Question	Answer form
Concept interpretation	Interview 1	Thinking-out-loud exercise	Interview question, open
	Interview 2	I1.7/I2.13 What does the concept “Happiness” mean to you?	Interview question, open
		I1.8/I2.14 What does the concept “Well-being” mean to you?	Interview question, open
		I1.9/I2.15 What does the concept “Life satisfaction” mean to you?	Interview question, open
		I1.10/I2.16 What does “Feeling happy” mean to you?	Interview question, open
		I1.11 Based on how you have described these concepts, how do these concepts relate to another for you?	Interview question, open
		I1.13 How do you feel about these concepts being used interchangeably in research?	Interview question, open
Score selection	Interview 1	I1.3 What is your experience with participating in (EMA) survey research?	Interview question, open
	Interview 2	I1.4 When you think of a scale of 1 to 10, what do the numbers mean to you?	Interview question, open
	EMA	I1.5 In your previous school, what did your grading system look like?	Interview question, open
		I1.6/I2.5 How do/did you go about choosing an answer?	Interview question, open
		EQ1. At this moment I feel happy.	Survey scale question, 1-10
		EQ2. What made you choose the score …?	Survey text question, open
Influence of participation	Interview 2	I2.6 How did the questions within the diary study influence you?	Interview question, open
	EMA	I2.7 How did thinking about how happy you were feeling influence you?	Interview question, open
		I2.8 How did filling out the diary study change over time for you?	Interview question, open
		I2.11 How do you feel about the answers you have given?	Interview question, open
		I2.12 Looking back on the week, would there be changes you would like to make to the ratings?	Interview question, open
		EQ11. In the first question, you scored your happiness with …. Do you feel the same?	Survey binary question, yes/no
		EQ11.1 How would you score your happiness now?	Survey scale question, 1-10
		EQ11.2 What made you change your answer?	Survey text question, open
		EQ11.3 Which questions influenced this change?	Survey multi-select question, up to 7 options

Population and recruitment

The focus of the study was to create a diverse sample while maintaining a manageable sample size for the scope of a master thesis (10 participants). The choice to recruit first-year participants from the English psychology bachelor provided the opportunity to get participants from different social and cultural backgrounds, reaching a broad demographic. This also allowed the researcher to provide (non-monetary) rewards for the study considering its intensity. Participants were recruited from first-year psychology students at the University of Groningen using the internal SONA system. As a part of the psychology curriculum, first-year psychology students are required to participate in research activities to acquaint themselves with the research process. For this study, participants were awarded 5.7 points, a significant portion of the 30 points required for their participation in research activities. Through the SONA system, first-year psychology students are able to sign up for studies of their interest. Recruitment utilized a blend of convenience and purposive sampling. Prospective participants self-signed up via the SONA system, forming an availability sample. Selections for diversity were made using the baseline questionnaire. The recruitment period took place from October 2nd until October 6th 2023. Data collection was conducted from October 9th until November 3th 2023.

Prior to participation the participants were informed of the context and topics of the study. They were informed this study was part of a master thesis, but also of a pilot study to assess the efficacy of the methods in the study. All participation was voluntary and participants were able to refuse participation or leave the study at any time. All data was pseudonymized and stored in encrypted data files according to the University guidelines on data storage. Participants were informed of potential risks in the form of an information sheet and were able to provide their consent through an informed consent form. All participants agreed to letting the data be used for the current study, and also that the pseudonymized data can be shared with researchers interested in the topic.

Analytic strategy

This exploratory study combined qualitative and quantitative methods to gain insight in the variability within and between participants. After the data collection phase, all interviews were transcribed and pseudonymized to reduce identifiability of the participants. Similarly, all answers to the open questions concerning the reasoning of choosing an initial and changed score of the EMA questionnaires were extracted from m-Path, pseudonymized, and ranked according to the given scores per participant. These documents were then uploaded into Atlas.ti (version 23, later version 24) for qualitative content analysis. This part will discuss per topic in the study how this was assessed.

Concept interpretation

The data was derived from both interviews, allowing for descriptive analysis of the concepts, and for comparison between and within participants' description. Using deductive coding, the concepts of happiness, life satisfaction, well-being, and feeling happy were identified, compiled, and analyzed. The focus was on creating an overall definition of the descriptions given by the participants, while also discussing differences and similarities. Also assessed were participants' perspectives on similarity and interconnectedness.

Score selection

Using data from the interviews, a focus on participants familiarity and preferences with scales and scores is assessed combined with the experiences in the current study. The main assessment was a qualitative content analysis of the open question what made you choose the score …? tied to the given score per participant. Using this, a description per score will be given to highlight possible patterns in score selection and topics often discussed per score and overall. Also included was an inductive assessment of response processes combined with a co-occurrence analysis of response processes per grade.

Influence of participation

This was assessed using a comparison of the initial score and reasoning and the potential adjusted score and reasoning given in the last parts of the EMA questionnaires. A qualitative content analysis allowed for describing potential factors of influence, providing

insights in how participants are affected by filling out questionnaires. On top of that, participants discussion of whether they would change any of their answers given throughout the EMA study was also assessed. This information was extracted from the second interview.

Results

Sample

The sample for this study comprised 11 participants, selected from 51 applicants. Initially, 10 potential participants were approached, resulting in responses from 9 individuals. To replace the non-respondent, an additional participant was recruited. Additionally, one more participant was recruited due to recording issues during the first interview, bringing the total to 11 participants. Their demographics are described in table 2.

The participants' demographics included an age range of 17-32, representing 11 different countries and 6 different continents. The majority of the participants (6 participants, 54.6%) was from Europe, divided equally over eastern and western European countries. The gender distribution consisted of 5 female participants (45.5%), 4 male participants (36.4%), and 2 non-binary participants (18.2%). In terms of research experience, 7 participants (63.6%) had prior research experience, while 3 (27.3%) did not. Additionally, 1 participant (9.1%) had previous experience specifically in EMA studies. While the sample did not constitute a clinical sample, during the initial interviews some participants discussed having been or still being in therapy or being on medication for mental health complications. Further details were not requested or discussed as this was not part of the study.

Table 2
Demographics of sample

Age	Gender	Ethnicity	Experience	Response	Days
17 (n=1, 9.1%)	Male (n=4, 36.4%)	African (n=1, 9.1%)	No (n=3, 27.3%)	22 (54.6%)	7 (n=7, 63.6%)
18 (n=2, 18.2%)	Female (n=5, 45.4%)	American (North) (n=1, 9.1%	Yes (n=7, 63.6%)	21 (45.4%)	8 (n=2, 18.2%)
19 (n=2, 18.2%)	Non-binary (n=2, 18.2%)	American (South) (n=1, 9.1%)	Yes, EMA (n=1, 9.1%)		11 (n=2, 18.2%)
20 (n=2, 18.2%)		Asian (n=1, 9.1%)			
21 (n=1, 9.1%)		Australian (n=1, 9.1%)			
22 (n=2, 18.2%)		European (East) (n=3, 27.3%)			
32 (n=1, 9.1%)		European (West) (n=3, 27.3%)			

Response rate

Despite recording issues during the initial interview, all subsequent interviews with the 11 participants were conducted satisfactorily for analysis. The EMA study required participants to respond to three prompts daily for one week, resulting in a total of 22 prompts, including the initial prompt during the first interview. 6 participants (54.6%) completed all prompts, while the remainder missed only one, yielding 21-22 prompts for analysis per participant. Most participants filled out the EMA study in the planned time of seven days. Due to scheduling conflicts, for some participants it was eight or eleven days.

Concept interpretation

Participants were requested during the first and second interview to interpret four concepts related to happiness (*happiness*, *well-being*, *life satisfaction*, and *feeling happy*) to compare similarity in understanding of these concepts both between and within participants. In this study, participants described these concepts as multifaceted constructs with

similarities, but also key differences, which we will describe below. To gain understanding of what happiness and related concepts mean to the participants in this study, first a general description of the concepts as described by the participants as given. A full description of the concepts per participant per interview can be found in the appendix.

Happiness. Happiness, as described by participants, encompasses a dynamic emotional state characterized by a predominance of positive feelings and experiences. It is often associated with moments of joy, contentment, and fulfillment, where the positive aspects of life outweigh the negative ones. Happiness may manifest as fleeting moments of pleasure or more enduring states of well-being, influenced by individual experiences, perspectives, and external circumstances.

Well-being. Well-being extends beyond mere happiness to encompass a broader sense of physical, mental, and emotional health. Participants describe well-being as a prolonged state of positivity, incorporating both subjective feelings of happiness and objective indicators of health and vitality. It involves achieving a sense of balance, fulfillment, and contentment across various domains of life, including physical health, mental well-being, and social connections.

Life Satisfaction. Life satisfaction reflects an individual's overall evaluation of their life circumstances, encompassing past achievements, present experiences, and future aspirations. Participants conceptualize life satisfaction as a reflective process, wherein individuals assess their level of fulfillment and contentment across different domains of life, such as work, relationships, and personal goals. It involves a subjective appraisal of one's life trajectory, with the attainment of personal goals and the fulfillment of aspirations playing a central role.

Feeling Happy. Feeling happy, in contrast to the broader constructs of happiness and well-being, refers to momentary experiences of positive emotions and affective states. Participants describe feeling happy as an immediate sensation of joy, contentment, or pleasure, often triggered by specific events, activities, or interactions. While feeling happy

may be fleeting and transient, it contributes to overall well-being and emotional resilience, enriching individuals' daily experiences and enhancing their overall quality of life.

Similarity between participants

Interestingly, in these general descriptions similarities and differences are becoming clear among participants. While these constructs share commonalities and interconnections, they also exhibit distinct nuances that shape individuals' perceptions of their lives. Participants disagree on how clear the distinction is between the concepts mentioned in this study. On one hand happiness and feeling happy were often used interchangeably, as made clear in this quote by participant 9:

"*I'm thinking that my first description of happiness, this feeling happy, as well, as feeling happy to me is doing something fun and being happy with it. Like, that's when I feel happy. I'm not sure how to describe the difference.*"

Similarly, participant 10 described no difference between life satisfaction and feeling happy:

"*When I feel satisfied is the same thing. Similar concept.*"

Yet, participant 4 describes a key difference between happiness and feeling happy, attributing it to the time span in which it takes place:

"*I think more happiness is more the overall, feeling happy is more like the small, like the small sections of your day.*"

Participant 7 adds to this by discussing negative emotions when it comes to the difference between well-being and happiness:

"*Because I think like happiness is probably more of a really, complete absence of like, really negative emotions. And well-being would way more be kind of good balance and the control over those emotions.*"

This indicates that while there is similarity between the concepts, participants are affected by the wording and differences in their own interpretation of the concept. Participant 10 highlights this by stating her scores would have been affected if the concepts would have been used interchangeably:

"*Do I feel both happiness and satisfied? Or if I don't feel one of it? I would maybe pick the, a bit lower score.*"

This is further supported by most participants. Participant 6 however thinks that it is a good thing, as the concepts are connected to one another, making it hard to separate them completely:

"*I think they mean the same thing to a degree. You know, like, I do think well-being is definitely in part happiness and react to, I think it's like the most influential parts, mostly for me too. But I do think that they aren't exactly the same. I think there are a few differences, but then the differences that they have, once again, link to that part. So, it's like, it's all just one humongous spiderweb. They're very interchangeable. And you can't really split them up. I don't think it would be a good study to just do one of them solely, I suppose.*"

One interesting potential cultural difference is that two participants (participant 2 and 3) used the colour yellow to describe happiness. While most participants used a focus on attributes such as food or friends, these participants specifically mentioned yellow as the colour of happiness. Participant 2 also used colours to describe well-being as health as green, and well-being as mental state of mind as purple. Both participants are from Eastern European countries, possibly showing colours there are used more to describe feelings. However, participant 7, also from an eastern European country did not use colours to describe any of the concepts. This may indicate a mere coincidence between participants 2 and 3, or a more detailed study of cultural context and description of feelings might be required.

Similarity within participants

As participants were asked at two timepoints to describe the concepts, it is possible to compare how consistent the definitions of the concepts are in participants. Most participants were quite consistent in their wording, often even saying the same associations in the same order. A difference in their answers was usually that they became more concise in answering the question, giving more of a definition rather than a vague description or keywords. However, a few participants did change their definition quite drastically.

Participant 11 has one of the biggest definition changes in this study. Initially, they defined happiness as a balance between unhappiness and happiness, where a certain level of unhappiness is required for happiness:

"*I guess I would say, balance between, like, unhappiness and happiness. Like you don't really have happiness unless you struggle. So, balance.*"

In the second interview this has changed to the following definition:

"It means, to be in a positive state of mind, that is."

This change indicates a big difference in over only a week time, showing that when a participant is requested to participate, they could be working with a different definition from the week before. A similar big change we see for life satisfaction in participant 5. Their definition of life satisfaction changes from something future-orientated:

"*I think for me, life satisfaction is, it's boring stuff, treating myself with kindness and just being gentle towards myself and gentle towards others. And, you know, it'd be nice if I, if, I'd love to leave the world a little bit better than I found it, but it's about trying your best and it's not about actual, you know, ideally. But trying your best does mean going for the thing that gets the best results, but trying the best that you can to do it. And it's not about necessarily being remembered when I'm dead, but you know it like in a large scale, but I'd love to, you know, need at least somewhat a emotional impact on the people immediately around me. And that's enough for me.*"

To something more momentary:

"Life satisfaction. I see it very much as a in the moment thing, where it's just, I think, especially over the past week, it's very much been the moments where I felt the most life satisfaction, or just when I enjoy these very small moments, like that sandwich, I was just eating that sandwich and like, my life's fucking awesome. Just, yeah, I was just having a good time. And just it's about I think life satisfaction is that, I feel like it can draw a lot of set satisfaction from life, because there's a component of what you managed to accomplish within your life. But it's also just satisfied with life if I have lots of these little enjoyable moments."

Here the participant directly references the events of the past week as a source of their reformulation of their definition of life satisfaction. While the larger, future-orientated aspect isn't fully gone from the new definition, it has taken a much smaller part, expending more on the smaller aspects of life satisfaction.

All in all, participants show the concepts are complex and intertwined, but have distinctions between these concepts that could cause them to interpret a question (and its score) differently. While these concepts remain relatively stable within participants, it is possible for the definition to change within a week's time, implying clarification or a set definition could allow for more stability in a study.

Question interpretation

Using the thinking-out-loud exercise, interpretation of the questions was also observed. In general, most questions were understood as intended. One aspect discussed in almost all interviews (7, 63.6%) concerned the activities question. Foreign participants who had been in Groningen (and the Netherlands) since the start of the semester brought up the question of considering cycling a physical activity. For participants living in the Netherlands for a longer time, cycling was not considered a physical activity but a mode of transport. For this study, participants were requested to fill out the questionnaire according to their interpretation, which can mean the new inhabitants have considered cycling as a physical activity, while others have not, leading to discrepancies in what physical activity can entail based on cultural frame of reference. Similarly, in the activities question, participants required clarification on the 'rested' category. Participants considered their night's sleep as part of this, while the researcher only meant post-night resting, such as a nap during the day. By doing the thinking-out-loud exercise, the researcher was able to clarify this distinction.

Lastly, participants brought into picture a technological advancement in interaction. In the current climate, being in company might not happen while physically being in the same room. During the first interview, participants asked whether online interaction could count as in company:

"That I do have a question about. In company does that have to be physically or can it also be for example, when I'm on Discord with my friend?" (participant 9).

Similar to the physical activity discussion, it was left to the participant's interpretation.

"*My interpretation I would say that's in company." (participant 9)*

This is especially interesting when comparing it to participant 11's statement about interacting with someone physically in the room:

"Sometimes I think it was even in same room. But like, she was reading and I was studying. Yeah. And that's also not in company. Because I'm not like engaging."

Due to communication now happening over multiple modes, being in company has shifted more to the level of engaging with someone, whether this person is within proximity or only digitally connected. While this was not the initial idea of the researcher, it is an indication of a shift in the meaning of connection.

Score Selection

All participants completed their EMA successfully, with 5 participants missing only one prompt. Throughout the week, participants used a wide variety of scores to grade their happiness, as shown in figure 2. EMA plot lines per participant can be found in the appendix. They will not be individually discussed, as the focus for this part is on the scores themselves and their meaning, not the trajectories. This part will discuss participants experience of scales and scores based on previous experience and on the experience in this study.

Prior experience

Prior experience with score selection is divided into two parts. The first part is the grading system in their high school education, the second the previous experience with participating in research. Interestingly, not one participant had the same grading system in high school. The range of the grading systems varied immensely, with the longest being a 0-100 scale, and the smallest a 2-5 scale. Of the 11 participants, 4 had been exposed to multiple scales during their previous education, of varying lengths. Participants who mainly had been exposed to shorter scales (i.e. 1-6, 2-5, 1-5, A-F) indicated more difficulties with filling out the

1 to 10 scale. Prior to the study, participant 3 acknowledged their struggle with the following statement:

"Maybe that's where I have the trouble of seeing those numbers when they giving me from one to 10? Because I got used to 2, 3, 4, 5? Because it's clear what is what yeah, there are no like, as I can as I see them intervals, because I see seven and eight as an interval. So, I can't choose between those two, because I see it as something middle, in between good and very good. So, I can choose and in our [Country's] metric system two, fail, three, okay, four, good, five, excellent. So, it's pretty easy. They're not intervals, so I got used to it. Yeah."

Similarly, participant 11 expressed their aversion for the 1 to 10 scale, stating there are too many options within this scale:

"*Yeah, I guess because you because you can put it in a category. Yeah, like seven or eight. And familiar. The difference between seven or eight is not that big. So, I'll just put down seven for that group. And rather than an eight, yeah.*"

In this participant's view, the scale works more in categories, reducing the scale to a 1 to 5 scale in their own mind. When asked how the participant would make a choice between a 7 and an 8, they said the following:

"It's a 50/50."

Interestingly, this participant also had one of the smaller scales in high school, using the A to F system. Yet for them it had been over 15 years ago. After working with the 1 to 10 scale in this study for one week, the appreciation of the 1-10 did not increase for this participant, remaining steadfast in only using smaller scales.

The opposite is also true, where some participants found the scale too small. After an emotional event, participant 9 found themselves hurting, impacting their happiness. In a slow pace, the participant recovered some of their happiness, however their struggle was not that there had not been improvement, but that the gap between scores was too large:

"Like, it's just, there was an increase, but there were no in-between options. So it's like, yeah, it's not the same as before, but it's also not going towards neutral feelings yet."

From their perspective, scales with decimals or a 1-100 scale could have captured the nuance of the happiness recovery better. It should then also be noted this participant grew up accustomed to a 1-10 scale with decimal options. This could imply that the system we get accustomed to in our high school time can have long-lasting effects on how we appreciate and interpret scales.

The second part of prior experience showing influence on interpretation of scores is more a contextual and interpretation matter. Prior to the EMA study, participant 5 indicated discontent with smaller scales due to the lack of nuance often provided by these scales:

"I remember being really annoyed at it. Because the questions were so unspecific, and then, like the answers were only like, strongly agree, slightly agree, slightly disagree and strongly disagree. And it created such a limited field of giving you an accurate answer that I got three questions in and then stopped the test. Because it was too frustrating?"

As participants in quantitative questionnaires often do not have the ability to provide additional context concerning their thought process, participants can be stuck between several conditions to fully answer the question. After the EMA study, participant expressed their surprise at enjoying the shorter yes and no questions in the current study:

"I think I have to revise my opinion on that. Because this is, I realized if there was, if it was more than that, I just feel like that's too much effort to try and figure out the distinctions between the two and this case, maybe less is more when it's specific questions?"

For participants with similar thoughts concerning short scales, being able to add additional context in the form of an open text box in the questionnaire can give them the ability to make a choice or explain why a choice cannot be made. If this would not be an option within the questionnaire, the short scale questions should be operationalized in such a manner the choices are not dependent on multiple conditions.

Score choices

Throughout the EMA study, participants used a wide range of scores to indicate their happiness. Answering the question "*At this moment, I feel happy*" almost all possible scores were used at least two times during this study. Only the lowest score, 1, was never used. The

least used score in this study following 1, is 10. Only two participants (1 and 7 used this score. An overview of scores per participant can be found in table 3.

The scores used the most in this study are 7 (70 times) and 6 (45 times) and all participants used these scores at least two times. Participants 4 and 6 contributed the most to the use of the score 7, respectively choosing it 12 and 11 times. For most participants, the average score of the week was above 6, indicating in general the week was perceived as positive. Exceptions to this are participants 8 and 10, with scores of 4.76 and 5.05. The highest average belongs to participant 4, with 7.45. This is also due to them having the smallest range of options (6-9, 4 scores). The largest range can be attributed to participant 1, ranging from 3 to 10 (8 scores).

Table 3
Scores per participant

Score	Participant											Total
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	
1												0 (0.0%)
2			1					5	1			7 (2.95%)
3	1	1	1					2	1	6		12 (5.06%)
4	2	4				2	2	2	1	3		16 (6.33%)
5	3	2	3		2	2	5	5	3	3	3	31 (13.08)
6	4	3	5	2	8	4	5	2	3	5	4	45 (18.99%)
7	3	7	7	12	6	11	3	3	8	4	6	70 (29.54%)
8	5	3	4	4	5	2	3		5	1	6	38 (16.03%)
9	2	2	1	4	1		2	2			2	16 (6.75%)
10	2						1					3 (1.27%)
Total*	22	22	22	22	22	21	21	21	22	22	21	238
Mean	6.77	6.27	6.36	7.45	6.77	6.43	6.48	4.76	6.27	5.05	7	6.33
SD	1.95	1.72	1.65	0.91	1.07	1.12	1.69	2.23	1.67	1.65	1.22	1.72
Min – Max	3 – 10	3 – 9	2 – 9	6 – 9	5 – 9	4 – 8	4 – 10	2 – 9	2 – 8	3 – 8	5 – 9	2 – 10
Range	8	7	8	4	5	5	7	8	7	6	5	9
Mode	8	7	7	7	6	7	5/6	2/5	7	3	7/8	7
Scores**	8	7	6	4	5	5	7	7	7	6	5	9

*= Max n of prompts is 22
** = indicates amount of different scores used throughout the EMA study

Response processes

Within the EMA study, the manner in which participants answered can be ascribed to certain response processes. In total, 236 quotes were assessed for response processes. These processes were identified and inductively coded. Following, the processes were divided into three major groups, namely reasoning (5 processes), time perspective (3 processes), and writing style (3 processes). Note, more than one process can be assigned to a quote in this study, providing a comprehensive overview per quote.

The reasoning group of response processes directly refers to the content of why a participant chooses a score. Whenever participants made a direct referral to the reason of feeling happy, one of the processes from this group would be assigned. Within this group, the following processes were identified:

- Attributive: Participant describes an event, person, activity, or other tangible reason for assigning the score that happens in that moment.
- Comparative: Participant compares the current situation to a previous or upcoming situation, or compares themselves to someone else, influencing their score.
- Emotive: Participant describes their reasoning for a score as a feeling, i.e. feeling tired, feeling energized.
- Integrative: Participant actively weights positive and negative factors and integrates both in their answer to come to a score.
- Reactive: Participant describes a reaction to the study as a reason for the score.

Of this group, the attributive was used the most in this study (204 times), followed by emotive (124 times). Reactive was only used once, yet was described as a direct reason for low mood.

The second group concerned time perspective of the participant. Within this group, participants made either direct references to a past or future event or reasoning, or indirectly discussed in the moment aspects. As such, this group consisted of three processes:

- Momentive: The participant describes an in the moment event for their score.
- Prospective: The participant describes a future moment event for their score.

- Reflective: The participant reflects on a past moment event for their score.

Momentive was the most used time perspective in this study (172 times), followed by reflective (93), and then prospective (49). Here it is especially important to note that participants could describe multiple events from different time perspectives in one reasoning, allowing for overlap between the time perspectives, as seen in the co-occurrence table 4. For example, momentive and reflective overlapped 44 times, while reflective and prospective overlapped 17 times.

The last group concerned how participants wrote down their answers, looking at how detailed and clear answers were while also taking the amount of reasons into account. For these, three groups were assigned as well:

- Comprehensive: Participants discuss 2 or 3 reasonings in great detail, or reasoning is clear and concise in such a manner that no additional details are required.
- Satisficive: Participant writes the bear minimum, writing less than 5 words for their reasoning.
- Selective: Participant discusses 1 or 2 reasons in more than 5 words, but additional details are required for the full picture.

Unlike the other processes, each quote in the study could only receive one of these processes, meaning that a satisficive quote could never receive a comprehensive label as well. Selective was used the most, 121 times, followed by comprehensive (87 times), with satisficive being last (22 times).

Response processes per participant

Table 5 provides an overview of the response processes per participant. All participants used attributive reasoning to explain their happiness for the majority of their quotes, with participant 5 using it for each of their quotes. Emotive was the second most used reasoning process, indicating that for most of the participants, happiness was explained both by feeling or by a clear event, person, or object, and considering the high number of quotes per process, both. Only participant 1 diverges from this, having 19 quotes explained by attributes and 2 by emotions., indicating a separation between emotion and attributes. When

it comes to time perspective, only participant 2 never provided future perspectives for their reasoning, while participant 5 tried to equally look to the past and the future, although still referring to the present the most. In writing style, participant 1 was the most satisficive of the participants, writing short answers 7 out of 21 times. Other satisficive participants were 10 and 11 (both 5 satisficive answers), and P7 (4 satisficive answers). On the opposite site, participants 5, 6, and 8 were the most comprehensive writers, providing elaborate answers 16 times, with participant 6 writing comprehensive for all their answers (21 times). Participant 1 was the least comprehensive, only writing an elaborate answer once. The differences in the used processes highlights participants having their own style and perspectives when answering questions.

Table 4

Co-occurrence analysis of response processes

	Attributive n=205	Comparative n=10	Emotive n=124	Integrative n=48	Reactive n=1	Momentive n=172	Prospective n=49	Reflective n=93	Comprehensive n=87	Satisficive n=22	Selective n=123
2 n=7	7	0	3	1	1	5	0	3	7	0	0
3 n=12	9	1	6	0	0	5	3	4	4	3	5
4 n=15	13	2	9	2	0	10	2	4	5	1	8
5 n=31	23	2	20	3	0	28	2	6	9	2	20
6 n=45	39	2	30	14	0	37	13	19	17	8	18
7 n=70	64	3	38	22	0	48	28	32	33	4	33
8 n=36	34	0	12	4	0	23	1	18	9	3	22
9 n=16	13	0	6	2	0	13	0	7	3	0	13
10 n=3	3	0	0	0	0	3	0	0	0	1	2
Attributive n=205	0	8	102	43	1	145	47	90	85	14	101
Comparative n=10	8	0	7	5	0	6	1	6	7	0	3
Emotive n=124	102	7	0	34	1	101	29	44	58	8	56
Integrative n=48	43	5	34	0	0	39	12	22	31	0	16
Reactive n=1	1	0	1	0	0	1	0	1	1	0	0
Momentive n=172	145	6	101	39	1	0	31	44	64	17	86
Prospective n=49	47	1	29	12	0	31	0	17	30	2	16
Reflective n=93	90	6	44	22	1	44	17	0	49	4	39
Comprehensive n=87	85	7	58	31	1	64	30	49	0	0	0
Satisficive n=22	14	0	8	0	0	17	2	4	0	0	0
Selective n=123	101	3	56	16	0	86	16	39	0	0	0

Table 5
Response processes per participant

	Attributive n=205	Comparative n=10	Emotive n=124	Integrative n=48	Reactive n=1	Momentive n=172	Prospective n=49	Reflective n=93	Comprehensive n=87	Satisficive n=22	Selective n=123
P1 n=21	19	0	2	1	0	14	7	6	1	7	13
P2 n=22	21	3	11	5	0	18	0	11	8	0	13
P3 n=22	18	0	9	4	0	15	5	11	5	0	16
P4 n=21	18	0	16	10	0	16	3	5	5	0	16
P5 n=22	22	0	15	4	0	17	11	11	16	0	6
P6 n=21	20	2	19	9	0	20	8	12	21	0	0
P7 n=21	17	2	11	4	0	17	1	7	2	4	15
P8 n=21	19	1	7	4	1	14	4	10	16	0	5
P9 n=22	21	2	12	3	0	11	1	11	7	1	14
P10 n=21	14	0	13	0	0	12	7	4	2	5	14
P11 n=21	16	0	9	4	0	18	2	5	4	5	11

Meaning of scores

In this part, a short discussion of the meaning according to the participants per score is given. This is not assessed individually, but instead as a full overview to get a view of how diverse answers can be within a score. A complete overview with detailed discussions of the meaning per score can be found in table 6.

Starting from the bottom up, the obvious negative to positive association comes up. A two, the lowest used score in this study, is seen as fully negative with no room for nuance or positive belief, while a 10 is always associated with being in company of friends or family engaging in social activities. From a four until 9, the integrative response process is included, indicating that within this study, participants weigh negative and positive against each other within this range. Even with a grade generally considered high (8 or 9) this is the case, showing that participants clearly distinguish what would be required to pick a higher score. Momentive and attributive processes are assigned to all grades (not all quotes!), showing that for the most part, participants are able to identify at least one reason occurring in that moment influencing their happiness. However, when looking at the extreme scores in this study (2 and 10), a 100% of the scores receive an attributive score, indicating an extreme score is associated with a clearer reason for this feeling. For the 2, the writing style also becomes comprehensive (100% of 7 quotes), whereas for 10, the writing style is divided between satisficive (1 time, 33.3%) and selective (2 times, 66.7 %). This does not imply participants wrote more for the two than for the 10. While for the most part this is true, participant 9 attributed the following to selecting a 2:

"*I'm in the middle of post-breaking up with my girlfriend conversation.*"

Although this is only one reason, and perhaps curiosity about the details of the situation could occur, from this quote it is clear what the reasoning for the low score is, assigning this quote a comprehensive label. Similarly, participant 7 wrote for a 10:

"*I'm with my friends and I'm a little drunk.*"

Here, two reasons are discussed, showing attributive and momentive processes, but clarity regarding the activities could be provided to get a clearer view of the situation and the

surrounding context. For this reason, the selective label has been assigned to this quote. It should be noted only 2 participants chose a 10 and only 3 a 2. The participants that chose a 10 had been prone to being selective and satisficive throughout the entire study. Similarly, participants that chose a 2 had been prone to be comprehensive in their answers, with the exception of participant 9, leaning more towards selective answering. The writing style for these scores could therefore both be an effect of the events (more analytic when negative, more distracted when positive) or an effect of the participant's own writing style.

The most difficult discussion in choosing a score seems to be around the true neutral point. As shown in table 6, neutral is discussed from 4 to 8, with participants discussing either not being able to provide reasoning:

"*Nothing, I'm neither happy or sad.*" (participant 10, discussing a 5)

or just being at their base level of happiness:

"*Generally average score.*" (participant 11, discussing an 8).

Other participants due make a clear distinction by putting 5 as their neutral and calculating their happiness based on whether they feel neutral:

"I suppose I don't feel neutral. I feel a little bit above that. Just having a lazy morning downstairs in the living room. Nothing exciting is happening and I'm not doing anything that's really increasing my mood." (participant 6, discussing a 6).

The difference in anchor points (low, neutral, high) might therefore influence when participants consider themselves happy or not. Especially when participants like participant 4 do not want to use scores below a 6:

"But I kind of push myself to look at it in the positive, like, in a positive way. So I'm like, so I kind of, I kind of push myself to like, I mean, I don't, I might have like stress right now and problems or whatever like everyone has, but I'm overall happier than I've been in like a while."

This participant indicates that their life has improved to an extent the lowest current moments do not feel as heavy as the lowest point before the improvement, urging them to not go below a 6. This can be seen as evidence of a response shift, where significant life changes

impact internal standards and values, altering how one rates their life experiences (Schwartz et al., 2007). In this case, a participant might skew their data towards one side of the scale, as is also evident for participant 4. While the short range in their score choices could betray such a shift may have occurred, other participants in this study did not go below a 5 or 4 either. As the study was only for one week, this could imply it has been a positive week for most participants in this study. It could also imply participants see going below a 4 as extreme or undesirable. To understand the standards of the participants anchors could provide clarity on their score choices.

A last trend to discuss in this section is the role of other people in the participants scores. Although the focus is not truly on happiness in this study, it is interesting to dedicate a few words to the impact of one's network and interaction with others on their levels of happiness. For one, all scores over 8 are (with the exception of one quote) associated with interactions with others, whether this is playing games, going for dinner, or seeing pets again after a long time. We also see this reflected in the negative, where uncertainty of connections between friends can cause distress and decreases happiness in this study:

"Because currently socializing is kind of hard. You have like a friend group. But nobody is exactly, friends, like close, close friends that we had friends before. So, like, very small things can make me feel like okay, so maybe they don't want to be friends with me. So, we were planning like, they were planning, like a dinner type of nights with everyone. And I was asking the host, if I can bring my friend because she was here with me. And he did not respond for like two days. And he texted in the group chat. And I said something about it in the group chat, and he did not respond there either. So, I felt like very, very bad." (participant 7)

And although in the end it was not used as a selection criterion, 6 of the participants were in a relationship. Some of the most negatively impactful situations were tied to interactions with the partners of the participants, as four of the participants experienced difficulties in their relationships, with one relationship even ending. Similarly, some of the participants were far from their home country, indicating missing their network from back

home, including partners. While it is not news that the network plays a large role in happiness, this thesis also supports this aspect.

Table 6

Qualitative interpretation per score

Score	Description	Topics	Participants	Response processes
1 (n=0)	A one was never chosen in this study.	-	-	-
2 (n=7)	A two is perceived as strongly negative. Participants discuss negative topics with no nuance, showing there is little positivity. The only exception is when participant 8 mentions liking a room they had seen, but it is followed by the negative belief they won't get the room.	Physical, mental, relationship, financial, housing, academic problems are discussed	3 (n=1, 14.3%%) 8 (n=5, 71.4%) 9 (n=1, 14.3%)	Attributive (n=7, 100%) Comprehensive (n=7, 100%) Emotive (n=3, 42.9%) Integrative (n=1, 14.3%) Momentive (n=5, 71.4%) Reactive (n=1, 14.3%) Reflective (n=3, 42.9%)
3 (n=12)	A three is associated with negative physical and emotional affect states. The participants describe waking up in pain or too early, having an early lecture, but also discuss relationship problems such as breakups or other struggles. For participants 3, 8, and 10, some of the expressed emotions are related to exhaustion and annoyance. Participant 10 also describes academic nervosity and fear, as they were nervous about failing their exam. Participant 2 also describe the loss of their wallet as an attributive reason for the score.	Waking up, annoyance, exhaustion/tiredness, academic and relationship problems, losing items, nervosity/fear concerning exams, sadness	1 (n=1, 8.3%) 2 (n=1, 8.3%) 3 (n=1, 8.3%) 8 (n=2, 16.7%) 9 (n=1, 8.3%) 10 (n=6, 50%)	Attributive (n=9, 75%) Comparative (n=1, 8.3%) Comprehensive (n=4, 33.3%) Emotive (n=6, 50%) Momentive (n=5, 41.6%) Prospective (n=3, 25%) Reflective (n=4, 33.3%) Satisficive (n=3, 25%) Selective (n=5, 41.6%)
4 (n=15)	A four is mainly associated with worries and stressors. Participants express difficulties with needing to study and finding it difficult to find motivation and to stop procrastinating. Participants 2 and 6 explicitly state feeling overwhelmed by the university, whereas participant 10 specifically mentions struggling with statistics. Participants 2, 6, 7, and 9 mention problems in communication, cooperation, and relationship worries as their reasoning. Participant 10 states that *the day wasn't that good, but not that bad either*, introducing this as the first neutral score for a participant. Additionally, participant 2 also mentions physical pain at two moments as the reason for a 4.	Academic responsibilities and stresses, boredom, communication and cooperation issues, waking up, physical pain, lack of motivation, relationship problems, exhaustion/tiredness, perceived lack of productivity, melancholy, procrastination, neutral,	1 (n=2, 13.3%) 2 (n=4, 26.7%) 6 (n=2, 13.3%) 7 (n=2, 13.3%) 8 (n=2, 13.3%) 9 (n=1, 6.7%) 10 (n=2, 13.3%)	Attributive (n=13, 86.7%) Comparative (n=2, 13.3%) Comprehensive (n=5, 33.3%) Emotive (n=9, 60%) Integrative (n=2, 13.3%) Momentive (n=10,66.7 %) Neutral (n=1, 6.7%) Prospective (n=2, 13.3%) Reflective (n=4, 26.7%) Satisficive (n=1, 6.7%) Selective (n=8, 53.3%)

5 (n=31)	A five is used as a neutral the most in this study. Participants tend to use a five when they have just woken up and feel like nothing has happened to influence their mood yet, or if for their feeling nothing exciting is happening. Participants 5, 8 and 11 also mention their planning their studies but not being able to follow the plan due to outside factors (a broken bike, or family distractions). Participants also describe physical factors such as tiredness, hunger, and pain/weakness (hangover) as a reason for a 5. Participants 5 and 9 describe a positive factor being weighed against a negative factor. For participant 5 it was gratifying to fix their bike, but this caused deviations in their study plan. Participant 9 describes still being sad about their breakup, but talking to their ex to finalize the breakup is helping. They also mention slowly improving since the occurrence of their breakup. Participant 1 also describes being stressed about driving, as this is something they do not enjoy.	Neutral, lack of excitement, driving, waking up, uneventfulness, tiredness, hunger, nuance, repairs, planning, physical pain/weakness, boredom, productivity, studying, procrastination, family difficulties, relationship problems	1 (n=3, 9.7%) 2 (n=2, 6.5%) 3 (n=3, 9.7%) 5 (n=2, 6.5%) 6 (n=2, 6.5%) 7 (n=5, 16.1%) 8 (n=5, 16.1%) 9 (n=3, 9.7%) 10 (n=3, 9.7%) 11 (n=3, 9.7%)	Attributive (n=23, 74.2%) Comparative (n=2, 6.5%) Comprehensive (n=9, 29%) Emotive (n=20, 64.5%) Integrative (n=3, 9.7%) Momentive (n=28, 90.3%) Neutral (n=9, 29%) Prospective (n=2, 6.5%) Reflective (n=6, 19.4%) Satisficive (n=2, 6.5%) Selective (n=20, 64.5%)
6 (n=45)	A six is used similarly to a five as in that participants tend to use this grade to indicate nothing exciting is happening. Differently to a five, participants seem more nuanced, weighting positive against negative more often. A six is sometimes also used as a neutral, however participant 6 explicitly states feeling above neutral/average to differentiate from a five. A six also sees the introduction of more food related topics, with participants mentioning dinner, breakfast, or getting food in another way. Here, friends, pets and partners are also mentioned in a more positive light than before, often the reason for the nuancing between the positive and negative. Participants seem to be more on opposite spectrums in their attributive reasoning, with participant 5 stating sleeping well as a reason for a five, while participant 9 mentions bad sleep. The topic of confidence also comes up, with participants 5 and 11 being on opposite sides. Participant 5 states feeling confident in their study plan, while participant 11 discusses not feeling ready for an upcoming exam. Participants also seem to discuss activities more, either preparing for activities, reflecting on activities, or being in the middle of them.	Neutral, friends/partner, food, nuance, responsibilities, activities, studying, early lectures, physical pain/weakness, productivity, calmness/contentment, tiredness, statistics, exhaustion, waking up, relaxation, progress, boredom/repetition, sleep, confidence, lack of excitement, pets, procrastination, stress, outside influences	1 (n=4, 8.9%) 2 (n=3, 6.7%) 3 (n=5, 11.1%) 4 (n=2, 4.4%) 5 (n=8, 17.8%) 6 (n=4, 8.9%) 7 (n=5, 11.1%) 8 (n=2, 4.4%) 9 (n=3, 6.7%) 10 (n=5, 11.1%) 11 (n=4, 8.9%)	Attributive (n=39, 86.7%) Comparative (n=2, 4.4%) Comprehensive (n=17, 37.8%) Emotive (n=30, 66.7%) Integrative (n=14, 31.1%) Momentive (n=37, 82.2%) Neutral (n=2, 4.4%) Prospective (n=13, 28.9%) Reflective (n=19, 42.2%) Satisficive (n=8, 17.8%) Selective (n=18, 40%)

7 (n=70)	A seven is the most used score in this study. It is characterized by a lot of social interaction. All participants note at least one moment where they interact with their family, roommate, partner, pets, or friends. Another connection is the increase in the mention of activities, just like with six in prospective, momentive and reflective perspectives. Another common topic is the weather, mentioned by participants 2, 3, and 5 as a positive factor, and by 4 as a negative factor. Participants 7 and 11 mention travelling as their reasoning. Participant 1 describes their day as neither bad nor good, considering this a neutral score. Similarly, participant 11 calls this their base level of happiness. Participants tend to weigh positive against negative even more in this score in comparison to the other scores in the study, with the main factor lowering the score being statistics, whether it is studying for statistics or being in a statistics practical. Other major topics discussed here are once again food, but also relaxation and peacefulness. Participants also seem to attribute the score to the day of the week in this score, describing excitement due to it being Friday or Saturday, or having a cozy Wednesday or a lazy Sunday. Physical pain and weakness still are present in this score, with participant 6 describing having an ear infection, and participant 11 suffering from back pain. Participant 4 also describes loneliness as a factor weighing down their grade. The most ambiguous answer is given by participant 11, stating "*studying. Bored. Tired.*" as their reasoning.	Neutral, roommates, responsibilities, activities, expectations, nuance, family, friends, waking up, warmth, relaxation, peace, statistics, understanding, weather, partner, food, loneliness, days, stress, tiredness, positivity, productivity, items, sleep, physical pain/weakness, pets, interest, travel, uneventfulness,	1 (n=3, 4.3%) 2 (n=7, 10%) 3 (n=7, 10%) 4 (n=12, 17.1%) 5 (n=6, 8.6%) 6 (n=11, 15.7%) 7 (n=3, 4.3%) 8 (n=3, 4.3%) 9 (n=8, 11.4%) 10 (n=4, 5.7%) 11 (n=6, 8.6%)	Attributive (n=64, 91.4%) Comparative (n=3, 4.3%) Comprehensive (n=33, 47.1%) Emotive (n=38, 54.3%) Integrative (n=22, 31.4%) Momentive (n=48, 68.6%) Neutral (n=2, 2.9%) Prospective (n=28, 40%) Reflective (n=32, 45.7%) Satisficive (n=4, 5.7%) Selective (n=33, 47.1%)
8 (n=36)	An eight is the last number still associated with neutral scores, as participant 11 describes this as their average level of happiness. While this can be interpretated as above neutral, during interview 2 they clarified it as one of their neutral scores. For most participants this score is associated with relaxation and enjoyable activities, good food, and music. Only participant 1 does not mention an interaction with others in this score. This score is also tied to productivity and (academic) success, with participants 2, 4, and 5 feeling productive in their day, and	Activities, courses, home, friends, recovery, productivity, family, partner, food, items, responsibilities, relaxation, peace, physical weakness, relief, absence of negativity, locations, repair, nuance, sun, academic success, neutral,	1 (n=4, 11.1%) 2 (n=3, 8.3%) 3 (n=4, 11.1%) 4 (n=3, 8.3%) 5 (n=5, 13.9%) 6 (n=2, 5.6%) 7 (n=3, 8.3%) 9 (n=5, 13.9%) 10 (n=1, 2.8%)	Attributive (n=34, 94.4%) Comprehensive (n=9, 25%) Emotive (n=12, 33.3%) Integrative (n=4, 11.1%) Momentive (n=23, 63.9%) Neutral (n=1, 2.8%) Prospective (n=1, 2.8%) Reflective (n=18, 50%) Satisficive (n=3, 8.3%)

	participants 6 and 11 mentioning getting a role they wanted and passing their exam respectively. Participants 2 and 6 mention some physical weakness but also how it has recovered or is getting less, providing relief. Participants 1, 4, and 5 attribute this score to being home and participant 7 to being at a cat café, adding location as a reason for happiness. Negativity seems to be getting less in an eight, with only participants 4, 6, and 11 weighing positive against negative factors. Participant 7 states having *nothing to feel unhappy about* as their reasoning.		11 (n=6, 16.7%)	Selective (n=22, 61.1%)
9 (n=16)	A nine is characterized by having a good mood. Participants answers become more selective, describing fewer factors contributing to this score. With the exception of participant 8 who describes more than two reasons, most participants seem to have one clear reason for their nine or can only attribute it to being in a good mood without further explanation. Participant 9 and 11 apply some weighting of negative and positive, stating procrastination of studying and not being in a 100% party mood respectively as their reason for not choosing a 10. While activities are mentioned, social interaction seems to be the key factor.	Home, pets, family, friends, meeting, family, food, activities, partner, days, good mood, productivity, roommates, procrastination	1 (n=2, 12.5%) 2 (n=2, 12.5%) 3 (n=1, 6.3%) 4 (n=4, 25%) 5 (n=1, 6.3%) 7 (n=2, 12.5%) 8 (n=2, 12.5%) 11 (n=2, 12.5%)	Attributive (n=13, 81.3%) Comprehensive (n=3, 18.8%) Emotive (n=6, 37.5%) Integrative (n=2, 12.5%) Momentive (n=13, 81.3%) Reflective (n=7, 43.8%) Selective (n=13, 81.3%)
10 (n=3)	A ten was chosen only three times in this study and is always associated with a social interaction. The answers here are very short and only include one or two reasons that can be seen as just one reason: social connection such as family dinners, conversations with friends, and drinking with friends.	Food, family, friends, conversation, alcohol	1 (n=2, 66.7%) 7 (n=1, 33.3%)	Attributive (n=3, 100%) Momentive (n=3, 100%) Satisficive (n=1, 33.3%) Selective (n=2, 66.7%)

Influence of participation

Overall experience

In general, participants experienced this current study as pleasant and not too heavy. The study was easily filled out and did not demand a lot of the participants. Participants discussed the experience as similar to "gratitude journaling", even indicating they would have liked to continue filling out the questionnaire. This was a shared sentiment among participants, with some expressing disappointment at the ending of the study. While this is nice to hear as a researcher, it also implies that within one week time, participants got accustomed to the regular prompts. It also led to some unwanted effects. Participant 6 indicated she got anxious after the 7 days period was over, as she feared she had missed one of the prompts:

"And then the following day, I was like, wait, why haven't I received the message? And then like, I like I was trying to figure out for 15 minutes, and I was like, did I miss it? And then I realized it's over."

For participant 11, it was a stressful period, and participating in this study contributed to the stress as an extra burden on top of their regular tasks:

"Yeah, it was a first it was nothing crazy. But then after a few days, maybe it was a bit. It just added an extra thing on top of my life.[P SEP]."

Participant 8 adds receiving the prompts was not always an enjoyable experience:

"It was sometimes it was a bit annoying to be honest, like I was doing something and then it popped up and I was like, ah, no, I have to do this questionnaire."

All in all, partaking in this study may not have been experienced as heavy, it could cause some negativity for the participants.

Despite these possible negative effects, most participants indicated they were able to identify some patterns in their happiness behaviour. Most participants stated they got more aware of what brings them happiness and what does not, usually seeing that participants were the happiest when in company. Participant indicated they would continue the reflection

process by keeping a journal, having seen from this week that the small things in life brought them joy.

However, other than these points, participants indicated not being affected more by the continuous reflection. By the time participants had to fill out the next prompt, the thoughts on the previous prompt had already been gone. There was no difference in how questions affected participants, or reflecting on happiness.

Within EMA changes

There are some observed changes in the questionnaire as a consequence of filling out the questionnaire. Out of 236 filled out questionnaires, participants changed their minds 9 times. The overview of the changes can be found in table 6. Out of 9 times, 5 times the grade changed for the better, rising 1 or 2 points. Participants describe participating in the study as a distraction from the negative experiences in their life, or state that reflecting on a day or activity within their day contributed to their happiness, increasing their overall happiness appraisal. The 4 times the grade changed towards the negative, it only changed by 1 point. Similar to the positive change, recalling an experience of the day causes participants to decrease their happiness score. Participant 5 mentions getting more awareness of their current activity as a contributor to the lowered score.

For the questions that contributed to this change, participants indicate mainly the questions where open text answers were given as a contributor to the changed grade. The main culprit has been the experience of the day question (“question”), which asks participants to reflect on the entire day and pick the most memorable moment. In both positive and negative sense, this question has influenced the participants to change their score. The second question influencing the participants concerns the interaction with other people (“question”). This one is only tied to positive changes. Only mentioned once as reasoning are previous activities and reflecting on choosing a score, both leading to positive changes. From a questionnaire point of view, this could imply that purely survey questions with scales or multiple-choice questions would have minimal effect on participants, where open questions could lead to participants being more affected.

Post-EMA changes

After reflecting on the given answers, most participants were quite satisfied with their given answers, regardless of whether they were positive or not. For them, the given score provided a good overview of what their week had been like, given a good overall experience. A shared sentiment within the study was well expressed by participant 1:

"I would say no, but also just because I, what I typed in is what I felt at that moment, if I'm now looking back, I'm surely going to say, oh, it wasn't that bad, or like, yeah, but that's not what I felt at that moment. So, it would be wrong to say I felt differently, because now that I look back, I see it wasn't that bad. But that's not what really was going on. And yeah, like what I was feeling. So, I wouldn't change anything, because I think it kind of represents what I was feeling at the moment."

Out of the 11 participants, 5 indicated wanting to change one or more of their scores, both higher and lower. However, participant 7, initially wanted to change a score to a lower score, but instead changed their mind:

"I think maybe with the hangover part. I was not exactly sure if whether they should be lower, but I wasn't feeling sad exactly, I was just tired. Yeah, now looking at them. I feel like I answered pretty well."

For participant 9, the change was not due to wanting to raise a score because of the current reflection, but because of the lack of nuance in a whole number scale:

"Well, I think two is a very good estimation there. But I think then this would have been a two and a half, this would have been a three and this would have been like a three and a half. Because there was improvement. However, I still wasn't feeling good."

For the most part, this indicates participant reflection was in line with how they felt throughout the week. By providing participants a moment to reflect on their answers, they are provided with an opportunity to more accurately reflect how the affect presented throughout the week.

Table 7
Qualitative analysis of changes in score

Participant	Initial score	Reasoning	Second score	Reasoning	Questions contributing
P1	7	at home playing a game with my roommate	8	reflection of the day made me even happier	Experience of the day, interactions with others
P4	8	I feel generally happy because i feel good now however it isn't 10 because there are certain things in my mind such was responsibilities that are affecting my happiness	9	I had a laugh while answering these questions	Experience of the day, Interactions with others
P5	6	I'm in the process of doing a repetitive only survey for SONA and it's very boring	5	the boredom I'm feeling is really setting in and is affecting what was a really great day so far. This study is like 50 minutes long and I'm only 15 mins in	Experience of the day, previous activities
P6	6	I'd say I feel fairly above average. just had a relaxing morning. lying in bed cuddling my cat right now. had a good breakfast. feeling a little tired but that's okay.	7	I suppose the questionnaire made me think about my day more. feel as though I could have rated it a little bit higher.	Reflecting on choosing a score
	7	feeling relatively happy. Awaiting for dinner to be ready. watching a little show with the family. my day has been fairly average - nothing too exciting.	8	remembering the little painting experience :)	Experience of the day
P8	3	I chose this score because of some struggle in my relationship concerning a phone call i had yesterday evening which is weighing on me.	4	The distraction by and the act of the thinking out loud exercise.	Previous activities, interactions with others
	4	I'm really tired and feeling a bit melancholic.	6	Thinking back to that discussion.	Experience of the day, Interactions with others
	9	I'm having some ice cream, watching a show with my roommate and had a really positive talk with my partner for the first time in a while 2 hours ago. I am procrastinating studying though.	8	Reflecting about the longer call as it wasn't all just great	Experience of the day
P9	5	Not feeling much happiness radiating off of me, quite bored	4	Remembering the argument	My experience of the day

Discussion

This thesis set out to explore the variability within and between participants in EMA research. Using the subjective concept of happiness, how participants interpret concepts, choose a score, and are influenced by research was investigated, providing numerous interesting insights.

Within this study, participants showed that even when a concept is widely known, the interpretation of the concept will differ between all of us. When it comes to the interpretations of the happiness related concepts, participants discussed similarities and an equal amount of differences between them. No one within this study used the exact same definition. Within participants, the definition remained quite stable, with the exception of two participants, completely changing their definition in the span of one week. These two findings highlight the importance of operationalizing our concepts in a manner that is understandable for everyone, for consistency and reliability of the data, but also to make a clear contribution to our concept of choice.

Similarly, how participants choose a score is not only subjective, but can be influenced by our cultural context. For one, if emotions have been taught using colour association, participants are more likely to bring up this colour in relation to the topic. In this study, Eastern European participants referred to happiness as yellow, while participants from other regions did not mention colour. This colour association would help one of the participants get through the second challenge of score selection, the scales used during the previous school period. For participants from school system with short grading scales (i.e. 1-5, A-F), interpreting a 1-10 scale is experienced as more complex, as differentiation between a 7 and an 8 and similar intervals is a new skill to them. This effect is persistent and long-lasting, evident from a participant being out of the system for over a decade. A participant was able to use the colour yellow as a gradient scale to overcome this issue, selecting the score at an imagined hue of yellow. The opposite is also true, where participants used to larger scales may not find the scale precise enough. Understanding the cultural context of participants in a study is therefore vital to creating surveys for populations.

While participants themselves indicate not being to affected by participating in research, some changes did occur in this study. By reflecting on the day, interactions, and activities, participants changed their initial score 9 times, creating a positive or negative difference of 1 or 2 points. Within quantitative studies, this difference can be the catalyst for a significant or non-significant result. Moreover, these changes show reflection does affect a participant in both positive and negative manner, albeit mainly when using open questions urging participant to write out their answers. This finding can especially be helpful for intervention design, allowing for reflective opportunity and creating more awareness of behaviour in both participants and researchers/practioners.

Treating variability in people as unexplainable noise would be to disregard the identifiable context researchers can take into account. While it is difficult to account for everything, a quick understanding of the population should be taken into account whenever new studies and instruments are created. For this reason, the populations culture and school systems should always be considered when the opportunity to discuss is not available. To combat misinterpretation of concepts, the usage of synonyms is not the best approach due to the slight or sometimes large differences in interpretation of the synonyms. Instead, the recommendation from this study would be to provide a definition and the surrounding context as much as possible, or to allow participants to explain the context of their choice in the form of open questions. Even though these open questions are more likely to affect participants, the understanding provided by these questions can be invaluable to both parties, providing opportunities for new research and interventions, or gratitude journaling.

Combining EMA and qualitative interviews was an interesting endeavor that led to a myriad of information. The cognitive methods combined well to allow multiple aspects to be explored prior and post the EMA session, giving additional insights in participant behaviour and understanding of the questions. Using these techniques can help researchers fine tune their instruments, but also correct any misinterpretations prior or during a study. Despite the small sample size, the collected data was rich to the point where many aspects of the data

remain unexplored. For example, the data could also have been used to assess recall of information within this study, as the warm-up exercise already showed how well participants recall the assignment, or whether they focus on all aspects of the assignment. As this was not a major component of this study, for now it has been omitted, but could be an interesting endeavor in the future. Similarly, the focus here was mainly on the qualitative aspects, where quantitative has been mainly used in a descriptive manner. Exploring the EMA data for differences in the participants based on demographics or behaviour from a statistical perspective could proof interesting as well.

In the initial design of this study, participants would have been allowed to answer the open questions of this study by speaking or writing out their answer, instead of just writing. However, due to financial constraints, this had to be omitted from the study. In this study however, some participants indicated a preference to speaking out the longer answers, which could have led to more detailed answers, and less satisficive behaviour in this study. A study allowing for both, or just the speaking version could give new insights.

The most difficult part of this study was dealing with the lack of literature on the response processes. Literature to be found was often from decades ago, or only on qualitative studies, making it complex to create a theoretical framework. Instead, this study for the most part took a more inductive approach, allowing the data to give the direction of the thesis.

In conclusion, this thesis underscores the significance of understanding and accounting for variability in EMA research. By exploring how participants interpret concepts, select scores, and are influenced by research, the study highlights the necessity of considering cultural and contextual factors in survey design. The findings advocate for clear operational definitions and reflective opportunities to enhance research precision and participant engagement. The combination of EMA and qualitative interviews provided rich data and valuable insights, paving the way for future research to further explore these dimensions and improve research methodologies in social and behavioral sciences.

References

Atlas.ti (23.2.1). (2023). [Software]. Atlas.ti. https://atlasti.com

Atlas.ti (24.1.0). (2024). [Software]. Atlas.ti. https://atlasti.com

Bauer, M., & Schoon, I. (1993). Mapping variety in public understanding of science. *Public Understanding Of Science*, *2*(2), 141–155. https://doi.org/10.1088/0963-6625/2/2/004

Bisogni, C. A., Jastran, M., Seligson, M., & Thompson, A. (2012). How people interpret healthy eating: Contributions of Qualitative research. *Journal Of Nutrition Education And Behavior*, *44*(4), 282–301. https://doi.org/10.1016/j.jneb.2011.11.009

Bond, C. E., Shipton, Z., Gibbs, A. D., & Jones, S. (2008). Structural models: optimizing risk analysis by understanding conceptual uncertainty. *First Break*, *26*(6). https://doi.org/10.3997/1365-2397.2008006

Bond, C., Gibbs, A., Shipton, Z., & Jones, S. (2007). What do you think this is? "Conceptual uncertainty" in geoscience interpretation. *GSA Today*, *17*(11), 4. https://doi.org/10.1130/gsat01711a.1

Burke, L. E., Shiffman, S., Music, E., Styn, M. A., Kriska, A., Smailagic, A., Siewiorek, D., Ewing, L. J., Chasens, E., French, B., Mancino, J., Mendez, D., Strollo, P., & Rathbun, S. L. (2017). Ecological Momentary Assessment in behavioral research: Addressing technological and human participant challenges. *JMIR. Journal Of Medical Internet Research/Journal Of Medical Internet Research*, *19*(3), e77. https://doi.org/10.2196/jmir.7138

Cacioppo, J. T., Von Hippel, W., & Ernst, J. M. (1997). Mapping cognitive structures and processes through verbal content: The thought-listing technique. *Journal Of Consulting And Clinical Psychology*, *65*(6), 928–940. https://doi.org/10.1037/0022-006x.65.6.928

Dawes, H. N., Barker, K. L., Cockburn, J., Roach, N., Scott, O., & Wade, D. (2005). Borg's rating of perceived exertion scales: Do the verbal anchors mean the same for different clinical groups? *Archives Of Physical Medicine And Rehabilitation*, *86*(5), 912–916. https://doi.org/10.1016/j.apmr.2004.10.043

DePrince, A. P., & Chu, A. (2008). Perceived benefits in trauma research: Examining methodological and individual difference factors in responses to research participation. *Journal Of Empirical Research On Human Research Ethics*, *3*(1), 35–47. https://doi.org/10.1525/jer.2008.3.1.35

Dequech, D. (2011). Uncertainty: a typology and refinements of existing concepts. *Journal Of Economic Issues*, *45*(3), 621–640. https://doi.org/10.2753/jei0021-3624450306

Dykema, J., Schaeffer, N. C., Garbarski, D., Assad, N., & Blixt, S. (2022). Towards a reconsideration of the use of agree-disagree questions in measuring subjective evaluations. *Research in Social & Administrative Pharmacy/Research in Social And Administrative Pharmacy*, *18*(2), 2335–2344. https://doi.org/10.1016/j.sapharm.2021.06.014

Ferrier-Auerbach, A. G., Erbes, C. R., & Polusny, M. A. (2009). Does trauma survey research cause more distress than other types of survey research? *Journal Of Traumatic Stress*, *22*(4), 320–323. https://doi.org/10.1002/jts.20416

Forman, J., & Damschroder, L. (2007). Qualitative content analysis. In *Advances in bioethics* (pp. 39–62). https://doi.org/10.1016/s1479-3709(07)11003-7

French, D. P., & Sutton, S. (2010). Reactivity of measurement in health psychology: How much of a problem is it? What can be done about it? *British Journal Of Health Psychology*, *15*(3), 453–468. https://doi.org/10.1348/135910710x492341

Gerring, J. (1999). What makes a concept good? A criterial framework for understanding concept formation in the social sciences. *Polity*, *31*(3), 357–393. https://doi.org/10.2307/3235246

Guo, Y., Logan, H. L., Glueck, D. H., & Muller, K. E. (2013). Selecting a sample size for studies with repeated measures. *BMC Medical Research Methodology*, *13*(1). https://doi.org/10.1186/1471-2288-13-100

Haynes, K. (2006). A therapeutic journey? *Qualitative Research in Organizations And Management*, *1*(3), 204–221. https://doi.org/10.1108/17465640610718798

Hennink, M., Hutter, I., & Bailey, A. (2020). *Qualitative research methods*. SAGE.

Hutchinson, S. A., Wilson, M. E., & Wilson, H. S. (1994). Benefits of participating in research interviews. *Image: The Journal Of Nursing Scholarship*, *26*(2), 161–166. https://doi.org/10.1111/j.1547-5069.1994.tb00937.x

Kirtley, O., Hiekkaranta, A. P., Kunkels, Y. K., Eisele, G., Lüken, M., Verhoeven, D., Van Nierop, M., & Myin-Germeys, I. (2018, 4 december). *The Experience Sampling Method (ESM) item repository*. https://doi.org/10.17605/osf.io/kg376

Kwasnicka, D., Kale, D., Schneider, V., Keller, J., Asare, B. Y., Powell, D., Naughton, F., Hoor, G. A. T., Verboon, P., & Perski, O. (2021). Systematic review of ecological momentary assessment (EMA) studies of five public health-related behaviours: review protocol. *BMJ Open*, *11*(7), e046435. https://doi.org/10.1136/bmjopen-2020-046435

Liao, Y., Skelton, K., Dunton, G., & Bruening, M. (2016). A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: An adapted STROBE checklist for reporting EMA studies (CREMAS). *JMIR. Journal Of Medical Internet Research/Journal Of Medical Internet Research*, *18*(6), e151. https://doi.org/10.2196/jmir.4954

Little, D. (1993). On the scope and limits of generalizations in the social sciences. *Synthese*, *97*(2), 183–207. https://doi.org/10.1007/bf01064114

Little, T. D., Widaman, K. F., Levy, R., Rodgers, J. L., & Hancock, G. R. (2017). Error, error in my model, who's the fairest error of them all? *Research in Human Development*, *14*(4), 271–286. https://doi.org/10.1080/15427609.2017.1370965

Ma, Q., Mermelstein, R., & Hedeker, D. (2020). A three-level mixed model to account for the correlation at both the between-day and the within-day level for ecological momentary assessments. *Health Services And Outcomes Research Methodology*, *20*(4), 247–264. https://doi.org/10.1007/s10742-020-00220-w

Macintyre, R. I., Heron, K. E., Dawson, C. A., Filipkowski, K. B., & Arigo, D. (2021). Does assessment alter responses? An examination of measurement reactivity in an ecological momentary assessment of body comparisons. *Journal Of Social And Clinical Psychology*, *40*(4), 304–332. https://doi.org/10.1521/jscp.2021.40.4.304

Maher, J. P., Arigo, D., Baga, K., Salvatore, G. M., Pasko, K., Hudgins, B. L., & König, L. M. (2024). Measurement reactivity in ecological momentary assessment studies of movement-related

behaviors. *Journal For The Measurement Of Physical Behaviour*, *7*(1). https://doi.org/10.1123/jmpb.2023-0035

Marshall, B., Cardon, P., Poddar, A., & Fontenot, R. (2013). Does sample size matter in qualitative research?: A review of qualitative interviews in IS Research. *The Journal Of Computer Information Systems*, *54*(1), 11–22. https://doi.org/10.1080/08874417.2013.11645667

Newman, E., & Kaloupek, D. G. (2004). The risks and benefits of participating in trauma-focused research studies. *Journal Of Traumatic Stress*, *17*(5), 383–394. https://doi.org/10.1023/b:jots.0000048951.02568.3a

Pannucci, C. J., & Wilkins, E. G. (2010). Identifying and Avoiding Bias in Research. *Plastic And Reconstructive Surgery/PSEF CD Journals*, *126*(2), 619–625. https://doi.org/10.1097/prs.0b013e3181de24bc

Pearce, R., & Barr, W. (2014). *Pearce & Stevens' Trusts and Equitable Obligations*. https://doi.org/10.1093/he/9780199644452.001.0001

Peterson, C. H., Peterson, N. A., & Powell, K. G. (2017). Cognitive Interviewing for item development: validity evidence based on content and response processes. *Measurement And Evaluation in Counseling And Development*, *50*(4), 217–223. https://doi.org/10.1080/07481756.2017.1339564

Presser, S., & Traugott, M. (1992). Little white lies and social science models: correlated response errors in a panel study of voting. *Public Opinion Quarterly*, *56*(1), 77. https://doi.org/10.1086/269296

Ronzani, C. M., Da Costa, P. R., Da Silva, L. F., Pigola, A., & De Paiva, E. M. (2020). Qualitative methods of analysis: an example of Atlas.TI™ Software usage. *Revista Gestão & Tecnologia*, *20*(4), 284–311. https://doi.org/10.20397/2177-6652/2020.v20i4.1994

Schaeffer, N. C., & Dykema, J. (2011). Questions for surveys: Current trends and future directions. *Public Opinion Quarterly*, *75*(5), 909–961. https://doi.org/10.1093/poq/nfr048

Schwartz, C. E., Andresen, E. M., Nosek, M. A., & Krahn, G. L. (2007). Response shift theory: Important implications for measuring quality of life in people with disability. *Archives Of*

Physical Medicine And Rehabilitation, *88*(4), 529–536. https://doi.org/10.1016/j.apmr.2006.12.032

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008a). Ecological Momentary assessment. *Annual Review Of Clinical Psychology*, *4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008b). Ecological momentary assessment. *Annual Review Of Clinical Psychology*, *4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Sica, G. T. (2006). Bias in research studies. *Radiology*, *238*(3), 780–789. https://doi.org/10.1148/radiol.2383041109

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In *Advances in psychology* (pp. 233–272). https://doi.org/10.1016/s0166-4115(08)61009-2

Stirling, A. (2003). Risk, uncertainty and precaution: some instrumental implications from the social sciences. In *Edward Elgar Publishing eBooks*. https://doi.org/10.4337/9781843765653.00008

Veenhoven, R. (1984). Conditions of happiness. In *Springer eBooks*. https://doi.org/10.1007/978-94-009-6432-7

Veenhoven, R. (2023). *World Database of Happiness*. Erasmus University Rotterdam. https://worlddatabaseofhappiness.eur.nl

Von Glasersfeld, E. (1983). On the concept of interpretation. *Poetics*, *12*(2–3), 207–218. https://doi.org/10.1016/0304-422x(83)90028-1

Willis, G. B. (2013, juni). *Cognitive interviewing*. ResearchGate. https://www.researchgate.net/publication/241686481_Cognitive_Interviewing

Willis, G. B., & Artino, A. R. (2013). What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *Journal Of Graduate Medical Education*, *5*(3), 353–356. https://doi.org/10.4300/jgme-d-13-00154.1

Appendix

Methods

Reflexivity statement

In undertaking this qualitative study on understanding response behavior in EMA studies among university students, I, Mithra Hesselink, a behavioral scientist with a multidisciplinary background in sociology, neuroscience, and developmental psychology, recognize the importance of reflecting on various aspects of my identity and experiences that may shape the research process.

My academic background provides me with a diverse set of lenses through which I approach the study, but it also introduces potential biases. As a behavioral scientist, I bring perspectives from sociology, neuroscience, and developmental psychology, which may influence the way I frame research questions and interpret findings. Additionally, being both a student and a teacher at the University of Groningen offers a unique vantage point, allowing me to understand student life from multiple perspectives.

However, it's crucial to acknowledge potential sources of bias related to age, skin tone, and cultural familiarity. As a researcher with a different age and skin tone compared to many participants, I am aware that these factors may influence participants' perceptions of both myself and the study. The diverse backgrounds of the participants, representing various nationalities, may introduce biases and stereotypes associated with cultures more familiar to me.

This disparity in age and position as researcher also introduces a power imbalance that may influence participant interactions. Participants might perceive me as an authority figure, potentially affecting the openness of their responses. To mitigate this power imbalance, I will strive to create a supportive and non-authoritative atmosphere during interactions, emphasizing the voluntary and confidential nature of their participation.

Moreover, the choice of the term "interview" might have introduced an element of formality that contrasts with the actual informal nature of the interviews, as revealed by the interview data. This discrepancy may have influenced participant responses. I

commit to acknowledging and addressing any discrepancies between the formal structure and the informal nature of the interviews to ensure the authenticity of the data.

Further complicating the dynamic is my role as a non-first-year psychology student and a former sociology student. I did not experience the first-year psychology curriculum, which may impact my understanding of the pressures and content faced by participants in this specific academic context. To address this limitation, I will remain vigilant in recognizing and addressing potential gaps in my understanding throughout the study.

Additionally, I recognize that my assumption about participants' views on happiness may have influenced the study design. While I assumed that everyone has a perspective on happiness, this assumption may not hold true for all participants. I will remain open to the diverse ways in which individuals conceptualize and express their feelings about happiness, allowing for a more nuanced interpretation of the data.

In presenting this reflexivity statement, I aim to enhance transparency and rigor in this qualitative inquiry by openly addressing potential biases and limitations that may impact the research process and outcomes.

Pilot phases

Due to the mixed-method nature of the study, the pilot phases were set up in four different phases.

During each phase, the pilot participants were informed they were participating in the pilot and requested to provide feedback on their experience and anything that was pleasant, unpleasant, confusing or otherwise unclear, or any mistakes they noted. For the phases done by the supervisors and main researcher, this was discussed face-to-face in meetings. In phase three this was done digitally with the pilot participants. For the fourth and last phase, feedback was given during the first and second interview by the participants. The reasoning for each phase and changes based on the feedback per phase are described below.

The initial pilot phase was conducted with the main researcher and the supervisors. This phase contained only one questionnaire with all questions included as shown in …. The objective of this phase was to assess the questionnaire for mistakes, duration of the questionnaire, and experience of filling out the questionnaire. This phase resulted in minor changes to the wording of questions, removal of redundant questions, and control of the language to assure all answers would be displayed in English rather than native languages to assure all participants would be presented with the same questions and answers.

The second phase had three different questionnaires, one for the morning (including sleep pattern questions), one for the evening (including experience of the day questions), and one for the afternoon (excluding sleep pattern and experience of the day questions). These were tested by the main researcher and supervisors. Based on this phase, minor fixes to the answer options were made (i.e. disabling to option to select multiple answers to the question concerning the current location of the participant), sliders for score questions were set to not start at a certain score (started at 6 prior to this fix), and a neutral option was added to the questions concerning company, sleeping, and day experience (i.e. to the question “I slept well” scores changed from yes/no to yes/neutral/no). This phase lasted 3 days.

Phase 3 included two participants (both female, aged 22 and 23, German and Dutch, one with and one without prior experience in questionnaire research), who filled out the three questionnaires for three days, to assess initial understandability prior to the complete pilot phase. Based on this phase, no extra changes were made.

Phase 4 included 5 participants: 4 female, one male, ages 19-36, two of Dutch nationality, 2 of Indian nationality, and one of mixed Turkish and German nationality. In this phase, both interviews were assessed, the initial full EMA questionnaire was used during the first interview for the thinking-out-loud exercise, and participants filled out the other three questionnaires for three days. A few major changes were made to the first interview guideline:

- Prior to the first pilot interview a warm-up exercise was added to the thinking-out-loud part. Participants were requested to visualize the place where they live, and think about how many windows there are in that place. While visualizing and counting the windows, they described what they were seeing and what came to mind. This is a technique used in cognitive interviewing to warm-up participants to speaking their thoughts out loud, which eases them into the process of doing the same during the questionnaire exercise. In this study, it was also found that the method of describing during this warm-up exercise also showed similar patterns to how participants filled out the questionnaire (i.e. participants that quickly counted all the windows in the room also filled out the open questions of the questionnaire with shorter answers and vice versa).
- Two of the participants indicated they felt pressured to answer each concept differently during the concept interpretation part, despite this not being required for the study. To remedy this, the following statement was added to the introduction of the concept interpretation part: “In this next part I will ask you to tell me what comes to mind for the concepts I will name. You can name anything that comes to mind, be it keywords, stories, vague descriptions, anything is fine. For these concepts you can mention the same things multiple times if these concepts are similar to you, so they

don't have to be different if they mean the same to you." This was added after the second pilot interview. Pilot interview participant 3-5 indicated they did not feel the same pressure.

- Five questions were added to the questionnaire: for participants not familiar with research their general view on research was asked, all participants were asked to describe what the numbers on a scale of 1-10 mean to them, concerning concept interpretation participants were asked to describe how the four different concepts related to each other in their view, how their own personal life and experiences may have affected how they interpret these concepts, and their opinion on whether these concepts can be used interchangeably in research.

Based on this phase only minor wording changes were made to some of the questions in the EMA, and a few options were added to the activities question to include self-care and rested (other than sleeping).

For the second interview only a statement that the concept interpretation was not about recall was added to the introduction of the concept interpretation to specify the purpose of the part: "Now you have done the diary study and we have looked back at your answers, I would like to refer back to the concepts we discussed in the first interview, and I would once again ask you to define what the concepts mean to you. For this, it's not about whether you still remember what you said previous time, so just forget about the first time we did this exercise, and just tell me what comes to mind when you think of these concepts. Here as well, you can mention the same things multiple times if these concepts are similar to you, so they don't have to be different if they mean the same to you."

Post-pilot phase

During the data collection phase three more questions were included in the interview guidelines as it naturally came up with the participants. The first question added was to ask participants to describe the grading system in their previous school in their country of origin. This came up during the interview as the main researcher had made the assumption this might be similar within European countries, but as the first participant described their

system differed even between schools in the same country, it was added to the subsequent interviews as well, including with probes concerning what participants considered a good grade, and how attainable these scores were in general and to them individually. This question was added to the first interview guideline. The second question and third questions were added to the second interview guideline. They pertained to whether the participants would describe their week as a typical week or not, and whether the duration of this study in their opinion was long enough to capture the essence of their happiness. These were asked as for some of the participants they had only recently moved to the Netherlands, and others were already more established, and for some of the participants it was closer to the exam week or even during the exam week in comparison to a regular school week.

Instruments

Justification of questions

Table 8

Justification of questions per concept and instrument

Theme	Concept	Justification	Operationalization	Source	Instrument(s)	Example	Related questions
Demographics	Age	Diversity of sample, selection criteria	Open question, number	-	Baseline questionnaire	What is your age?	B.2
	Gender	Diversity of sample, selection criteria	Multiple choice question, non-binary included	-	Baseline questionnaire	What is your gender?	B.3
	Nationality	Diversity of sample, selection criteria	Open question, text	-	Baseline questionnaire	What is your nationality?	B.4
	Relationship status	Diversity of sample	Multiple choice question	-	Baseline questionnaire	What is your relationship status?	B.5
	Work status	Diversity of sample	Multiple choice question	-	Baseline questionnaire	Are you working?	B.6
	Experience with research	Diversity of sample, selection criteria	Multiple choice question	-	Baseline questionnaire	Do you have experience with participating in (Diary study) research?	B.7
Concept uncertainty	Conceptual similarity	Concept interpretation, view of understanding of participant	Concept description similarity Between participants: Interview 1 Within participants:	Von Glasersfeld, 1983	Interview 1 Interview 2	No direct question	I1.7 I1.8 I1.9 I1.10 I2.13 I2.14

			Interview 1 & Interview 2				I2.15 I2.16
	Conceptual relation	Concept interpretation, view of understanding of participant	Participant interpretation of connectedness and requiredness of happiness related concepts	Von Glasersfeld, 1983	Interview 1	Based on how you have described these concepts, how do these concepts relate to another for you?	I1.11
	Conceptual interchangeability	Concept interpretation, view of understanding of participant	Participant opinion on using happiness related concepts as synonyms in research	Von Glasersfeld, 1983	Interview 1	How do you feel about these concepts being used interchangeably in research?	I1.13
Happiness	Happiness	Concept interpretation, view of understanding of participant	Participant interpretation	Veenhoven, (1984) Veenhoven, (2023)	Interview 1 Interview 2	What does the concept "Happiness" mean to you?	I1.7 I2.13
	Well-being	Concept interpretation, view of understanding of participant	Participant interpretation	Veenhoven, (1984) Veenhoven, (2023)	Interview 1 Interview 2	What does the concept "Well-being" mean to you?	I1.8 I2.14
	Life Satisfaction	Concept interpretation, view of understanding of participant	Participant interpretation	Veenhoven, (1984) Veenhoven, (2023)	Interview 1 Interview 2	What does the concept "Life satisfaction" mean to you?	I1.9 I2.15
	Feeling happy	Concept interpretation, view of understanding	Interviews: Participant interpretation EMA: the interpretation of the	Veenhoven, (1984) Veenhoven, (2023)	Interview 1 Interview 2 EMA	What does "Feeling happy" mean to you?	I1.10 I2.16 EQ1 EQ2 EQ11

		of participant, EMA topic	experience or feeling of momentary happiness				EQ11.1 EQ11.2
Influence on happiness	Activities	Influential on happiness	Multiple select question of activity categories with option to add additional categories	Kirtley et al. (2023) Veenhoven, (2023)	EMA	Since the last prompt, I have socialized	EQ3
	Interaction	Influential on happiness	Being or wanting to be in company, Interaction that stood out	Kirtley et al. (2023) Veenhoven, (2023)	EMA	Are you in company at the moment?	EQ4 EQ4.1 EQ4.2.1 EQ4.2.2 EQ5 EQ5.1
	Location	Influential on happiness	Specification of location and current happiness due to location	Kirtley et al. (2023) Veenhoven, (2023)	EMA	At this moment, I feel happy at this location.	EQ6 EQ6.1
	Sleep	Influential on happiness	Participant interpretation of sleep quality and falling asleep the night before	Kirtley et al. (2023) Veenhoven, (2023)	EMA	I slept well.	EQ7 EQ8
	Day experience	Influential on happiness	Participant evaluation of day & most memorable moment of day description	Kirtley et al. (2023) Veenhoven, (2023)	EMA	Describe the most memorable moment of the day and why it was memorable.	EQ9 EQ10
Influence of participation	Reflection	Process that occurs while letting participating in research	Interview 2: Influence of participation in study, reflection on scores filled out with space for adjustment EMA: Response adjustment in questionnaire with room for reasoning and	Research design of this study Willis & Artino (2013)	Interview 2 EMA	How did the questions within the diary study influence you?	I2.6 I2.7 I2.11 I2.17 EQ11 EQ11.1 EQ11.2 EQ11.3

			indication of influential questions				
	Response adjustment	Providing participant the opportunity to change a score during and post the EMA study	Interview 2: Post-reflection of filled out questionnaires opportunity to change a score EMA: response adjustment during the EMA		Interview 2 EMA	In the first question, you scored your happiness with …. Do you feel the same?	I2.12 EQ11 EQ11.1 EQ11.2 EQ11.3
Experience with research	Experience with prior research	Insight in participant's previous experience with scales and score selection	Discussion of participant's previous experience in research	Research design of this study	Interview 1	What is your experience with participating in (EMA) survey research?	I1.3
	Experience current study	Insight in participant's experience of current study	Discussion of participant's experience in current study	Research design of this study	Interview 2	How did you experience filling out the questionnaires?	I2.1 I2.2 I2.3 I2.4 I2.9 I2.10
Score selection	Response behaviour	Insight in participant reasoning when participating in research	Assessment of reasoning on scores in current study	Research design of this study	EMA	What made you choose the score … ?	EQ2
	Scale	Insight in participant's preferences and interpretation of scales in research	Previous experience with scales, interpretation of 1-10 scale	Research design of this study	Interview 1	When you think of a scale of 1 to 10, what do the numbers mean to you?	I1.3 I1.4

	Scales in school	Possible influence on score selection	Open interview question, grades in previous school systems in own country	First interview of interview 1	Interview 1	In your previous school, what did your grading system look like?	I1.5
	Score selection	Insight in participant score selection	Interview 1: Direct question concerning previous score selection Interview 2: Direct question concerning score selection in this study EMA: Open question to provide reasoning for score	Research design of this study	Interview 1 Interview 2	How do you go about choosing an answer?	I1.6 I2.5
Cognitive interviewing	Thinking-out-loud	Insight in participant response behaviour while filling out the questionnaire	Warm-up exercise: Describing out loud in an unrelated example (counting windows in the house) Thinking-out-loud exercise: Describing out loud the thoughts while filling out the questionnaire	Willis (1994) Willis & Artino (2013)	Interview 1	Instructions: Please read out the question and then say your answer out loud.	No direct questions
	Retrospective probing	Insight in participant reasoning post-surveys	Reflecting on chosen EMA scores	Willis & Artino (2013)	Interview 2	Observing line and answers given in EMA	I2.11 I2.12

Baseline questionnaire

B.1	What is your email address?	[Open question]
B.2	What is your age?	[Open question]
B.3	What is your gender?	[Male Female Non-binary/third gender Prefer not to say]
B.4	What is your nationality?	[Open question]
B.5	What is your relationship status?	[Single In a relationship not living together In a relationship living together]
B.6	Are you working?	[Not working Working parttime Working fulltime]
B.7	Do you have experience with participating in (Diary study) research?	[Yes, with questionnaire research, but not diary studies. Yes, I have experience with diary study questionnaire research. No, I don't have experience with questionnaire research.]

Interview guideline 1: Pre-EMA interview

Introduction
Thank you again for joining my study! This study is part of a master thesis for the Research Master in behavioural science, but is also used as a pilot study to assess the efficacy of this method for other projects, that is to say, I am testing whether it is possible to identify how people's individual choices and interpretations affect research outcomes. In this study I look at what a score on a scale means to a person, how different people interpret the same concepts, and how research participation influences people. Today we are starting with the first interview. This interview consists of three parts: Your experiences with or expectations of research, the concept interpretation, and a thinking-out-loud exercise where you will fill out the diary study for the first time. Before we begin this interview I would like to remind you that you are free to ask questions or clarification if anything is unclear, there are no wrong answers as everything is about your own perception and interpretation, and you are free to refuse a question or stop this interview at anytime. As was mentioned in the information sheet, if you stop the interview you can no longer participate in part 2 and 3 of this study, but you will receive the compensation for today's interview. This interview and the second interview will be recorded, transcribed, and pseudonymized. All data you have given will be saved until 31st of December, after that only pseudonymized data will still be stored for 10 years.

Before we start, do you have any questions?

Opening questions

I1.1 (If no experience) What is your view on research?
Probes: Looks like, questionnaires, outcomes, interpretations, scales

I1.2 What is your expectation of participating in this study?
Probes: Affected/influence, outcomes, worries, learning process

I1.3 What is your experience with participating in (EMA) survey research?
Probes: using scales, duration, types (daily diary, intensive), outcomes, influence, EMA, consistency

I1.4 When you think of a scale of 1 to 10, what do the numbers mean to you?
Probes: Extreme scores, understanding answers/interpretation, preferences

I1.5 In your previous school, what did your grading system look like?
Probes: Scale, pass/fail, desired grade, attainability

I1.6 If experience with questionnaires/EMA: How do you go about choosing an answer?
Probes: Comparison, difficulties, thoughts

Concept interpretation
In this next part I will ask you to tell me what comes to mind for the concepts I will name. You can name anything that comes to mind, be it keywords, stories, vague descriptions, anything is fine. For these concepts you can mention the same things multiple times if these concepts are similar to you, so they don't have to be different if they mean the same to you.

I1.7 What does the concept "Happiness" mean to you?
I1.8 What does the concept "Well-being" mean to you?
I1.9 What does the concept "Life satisfaction" mean to you?
I1.10 What does "Feeling happy" mean to you?

I1.11 Based on how you have described these concepts, how do these concepts relate to another for you?
Probes: Required for each other?

I1.12 How do you think your personal life and experiences have affected how you interpret these concepts?

I1.13 How do you feel about these concepts being used interchangeably in research?

Thinking out loud exercise
For this part you will fill out the Diary study for the first time. For this, I would like you to go through the questions, and while you do that, I would like you to say your thoughts out loud, as to what is going through your mind when you read the question and what comes up when choosing an answer.

Warm up exercise
Try to visualize the place where you live, and think about how many windows there are in that place. As you count the windows, tell me what you are seeing and thinking about.

Now we are going to fill out the diary study for the first time. So you can grab your phone and then I will send you the first prompt. Please read out the question and then say your answer out loud. For the first question you will be asked to choose a score, and then you can go to the second question immediately as there you will be asked to explain why you chose that score. I will also give some clarification for some of the questions so you know what to expect during the study.

I1.14 How was the thinking-out-loud exercise for you?
I1.15 How was the diary study for you?

Instructions for diary study
Tomorrow you will receive a prompt at a time we will set so you can start filling out the diary study for 7 days. You will receive one prompt in the morning, afternoon and evening at a time you will choose. You will have one hour to fill out the prompt. If you find that during the study anything occurs that makes you want to change a time for a prompt or if the study is negatively affecting you, please reach out to me and we can see what changes we can make to make the study work.

I1.16 This was the end of the interview. Before we fully end this, do you have any questions?
I1.17 Any comments?
I1.18 Any feedback you want to give on this initial interview?

Then now you will fill out the diary study for 7 days. After that we will have the second interview where we will reflect on your experience with this study, your answers in the diary study, and another look at the concepts.
Thank you so much for this and I hope it all goes well with the diary study!
I will see you again on …

Interview guideline 2 Post-EMA interview
(include the questionnaire so they don't have to recall the questions)
Welcome back to the last part of the study. You already did the first interview and the diary study, and today we will do the last interview. Just like the first interview, this one is also made up of three parts: the reflection on the diary study, reflection on the answers you have given in the diary study, and another round of concept interpretation. Just like last time this interview will be recorded, transcribed and pseudonymized. You are again free to ask questions or clarification if anything is unclear, there are no wrong answers as everything is about your own perception and interpretation, and you are free to refuse a question or stop this interview at any time.

Before we start this interview, do you have any questions or are there any comments on any of the previous parts of the study you want to make?

Reflection on Diary study

I2.1 How did you experience filling out the questionnaires?

I2.2 What did you notice while filling out the questionnaires?

I2.3 What questions were easy to answer for you?

I2.4 What questions were difficult to answer for you?

I2.5 How did you go about choosing an answer to the questions?

I2.6 How did the questions within the diary study influence you?

I2.7 How did thinking about how happy you were feeling influence you?

I2.8 How did filling out the diary study change over time for you?

I2.9 Was this week a typical week for you?

I2.10 How did you feel about the length of the study?
Reflection on answers (present their answers on the EMA)

I2.11 How do you feel about the answers you have given?
Probes: unexpected, awareness, actions taken/behaviour change

I2.12 Looking back on the week, would there be changes you would like to make to the ratings?

Concept reflection
Now you have done the diary study and we have looked back at your answers, I would like to refer back to the concepts we discussed in the first interview, and I would once again ask you to define what the concepts mean to you. For this, it's not about whether you still remember what you said previous time, so just forget about the first time we did this exercise, and just tell me what comes to mind when you think of these concepts. Here as well, you can mention the same things multiple times if these concepts are similar to you, so they don't have to be different if they mean the same to you.

I2.13 What does the concept "Happiness" mean to you?
I2.14 What does the concept "Well-being" mean to you?
I2.15 What does the concept "Life satisfaction" mean to you?
I2.16 What does "Feeling happy" mean to you?
I2.17 How do you see these concepts reflected in how you filled out the diary study?

EMA Questionnaire

PSY-2223-S-0350 Exploring response processes in EMA research

Full questionnaire questions

*= Only available in full & morning questionnaire

**= Only available in full & evening questionnaire

EQ1.	At this moment I feel happy.	[Scale 1-10]
EQ2.	What made you choose the score … ?	[Open question]
EQ3.	Since the last prompt, I have…	[Multiple select] [Socialized Engaged in physical activities Engaged in passive activities Studied Worked Done my chores Gone somewhere Eaten Rested Engaged in self-care activities Done something else, namely: Done nothing]
EQ4.	Are you in company at the moment?	[Yes/no]
EQ4.1	I would prefer to be in company right now.	[Yes/neutral/no]
EQ4.2.1	With whom are you at this moment?	[Multiple select] [With my partner With my family With my roommates With my friends With my colleagues With my study peers With acquaintances With strangers With my pet(s) With the researcher With others, namely:]
EQ4.2.2	I am enjoying my current company.	[Yes/neutral/no]
EQ5.	Since the last prompt, I had an interaction that stood out to me.	[Yes/no]
EQ5.1	Describe this interaction and why it stood out to you.	[Open question]
EQ6.	At this moment I am…	[Multiple choice At home At school At work At the (university) library At my family's house At a friend's house Somewhere else, namely:]
EQ6.1	At this moment, I feel happy at this location.	[Yes/no/neutral]

EQ7.	I slept well.*	[Yes/no/neutral]
EQ8.	I easily fell asleep last night.*	[Yes/no/neutral]
EQ9.	I had a good day.**	[Yes/no/neutral]
EQ10.	Describe the most memorable moment of the day and why it was memorable.**	[Open question]
EQ11.	In the first question, you scored your happiness with …. Do you feel the same?	[Yes/no]
EQ11.1	How would you score your happiness now?	[Scale 1-10]
EQ11.2	What made you change your answer?	[Open question]
EQ11.3	Which questions influenced this change?	[Multiple select Reflecting on choosing a score My previous activities My interactions with others My feelings about my current location My sleep* My experience of the day** None of the questions]

Codebook Qualitative answers EMA

Table 9

Codebook qualitative answers EMA

Code group	Sub-group	Code	Description	N	Example
Scores		1	Participant provided reasoning for a 1 score	0	-
		2	Participant provided reasoning for a 2 score	7	I've been spending the morning thinking about both my financial and relationship problems.
		3	Participant provided reasoning for a 3 score	12	I'm scared that I'll fail the exam.
		4	Participant provided reasoning for a 4 score	15	My hip is hurting and I feel very overwhelmed by university.
		5	Participant provided reasoning for a 5 score	31	I am feeling a bit low and I didn't do as much work as I wanted to.
		6	Participant provided reasoning for a 6 score	45	Back pain. Big day yesterday. Big day of studying statistics ahead. Generally okay but a bit bored.
		7	Participant provided reasoning for a 7 score	70	I have a lot of work to do but I'm trying to be positive and be able to get through the day.
		8	Participant provided reasoning for an 8 score	36	I am cooking which always makes me happy.
		9	Participant provided reasoning for a 9 score	16	I'm in a good mood but not 100% party mood so doesn't qualify for a 10.
		10	Participant provided reasoning for a 10 score	3	Seeing a good friend and having the best of talks.
Response processes	Reasoning	Attributive	Participant reasons score due to an event, activity, person, etc.at that moment	205	Stressed because I need to drive.
		Comparative	Participant reasons score due to a comparison to an earlier event or other person	10	I feel better than this morning, still not necessarily happy.
		Emotive	Participant reasons score due to an emotion or feeling	124	Nothing much happened so far, but I feel relaxed and peaceful today.
		Integrative	Participant reasons score by weighing positive and negative factors	48	I am excited for seeing friends tonight but feel a bit lonely at the moment.

	Reactive	Participant reasons score due to a reaction to the study	1	It's pretty weird between me and my partner right now and this prompt just made me realize I forgot to call them back from like 2 hours ago.
Neutral	Neutral	Participant considers a score neutral/average	15	I am just neutrally happy to start another day.
Time perspective	Momentive	Participant's reasoning is based on events/feelings in the present	172	I am with some of my friends right now.
	Prospective	Participant's reasoning is based on events/feelings in the future	49	I am about to spend time with my family.
	Reflective	Participant's reasoning is based on events/feelings in the past	93	I lost my wallet and woke up with a headache.
Answering style	Comprehensive	Participant's reasoning is a complete description such that multiple factors are described and/or no immediate additional information is required	87	I had to walk here this morning, cause my bike tire is punctured. This was tiring but its also nice cause its a nice day. Have a lot of stuff to do today, none of it particularly unpleasant, but it will require significant energy and Organisation. This gives me a small amount of stress, which could make my day worse if it got worse. I didn't choose 10 because I'm not filled with utter joy currently, but I feel pretty good.
	Satisficive	Participant's reasoning is only a few words	22	Bad sleep.
	Selective	Participant's reasoning is a short description that includes 1 or 2 reasons that are shortly touched upon	121	I bought a cool new scarf.

Results

Concept interpretation
Description of concept per participant per interview

P1X7LM
Happiness:
Int2: Feeling where mostly everything is good/When everything is positive, absence of negativity, when the positive outweighs the negative

Well-being:
int2: Happiness over a prolonged period of time, positive or negative, being positive most of the time, absence of mental/financial/social/etc troubles

Life satisfaction:
int2: Something in the future, something to reach, reflection on what is already reached, multiple ways of establishing (financially, work satisfaction, etc)

Feeling happy:
int2: Difficult to describe, Feeling positive, feeling good, when they're not bad, euphoria

Connection: -
Required: -
Synonymity: -
Diagram: -

P1Y47F
Happiness:
Int1: Satisfied in the moment with what you have, goal dependent, Momentary things (i.e. relationships), Feel fulfilled for themselves
Int2: When they are satisfied, When they feel good about themselves, when they are productive, when they are in company

Well-being:
int1: Being healthy, getting enough sleep/food, feeling it's good for you
Int2: Physical health, how well they slept

Life satisfaction:
int1: Different for everyone, similar to happiness (when happy they are also satisfied)
int2: When they are happy

Feeling happy:
int1: Listening to happy music
int2: Feeling satisfied

Connection: Happiness, life satisfaction and feeling happy are similar/the same, well-being part of happiness
Required: all for life satisfaction, but happiness possible without well-being
Synonymity: Similar but not the same, would pick different scores
Diagram:

P2X8RN
Happiness:
int 1: Being calm, being fulfilled, Feels warm, Golden honey colour. Something pleasant and relaxing, Safety/security, relaxed
Int2: Sense of calmness and warmth, Yellow colour, not feeling alone/lonely, Feeling content, feeling of belonging

Well-being:
int1: health, something green when it's health related, peace of mind, something purple when peace of mind related, physical health, mental contentment, mentally feeling alright, Calm, secure, not constantly in fight or flight feelings, feeling happy
int2: Green colour, mental well-being, physical well-being, absence of bad things, absence of physical or emotional pain, feeling of calmness, feeling of balance

Life satisfaction:
int1: How all the parts of life sum up and make you feel you're on a good line, no aspects in life requires more work, a connected web, balance, goal orientated (including good relationships with others and themselves)
int2: goal orientated (not just professional or academic goals, but also social goals, personal goals), achieving goals to feel content in current state

Feeling happy:
int1: Like little fireworks or popping candies inside their chest extending towards all ends of their body, a strong, jittery in a good way sensation that floods them with warmth, radiating energy and good light
int2: Something very jittery or explosive in a good way they feel in their chest like candies that pop or little fireworks that spread a warm feeling throughout their body, Makes them want to jump and be very expressive about their feeling

Connection: happiness and well-being required for life satisfaction
Required: Life satisfaction requires happiness and well-being, but the others are not dependent on each other
Synonymity: understands why but dislikes the idea, will answer differently about them
Diagram:

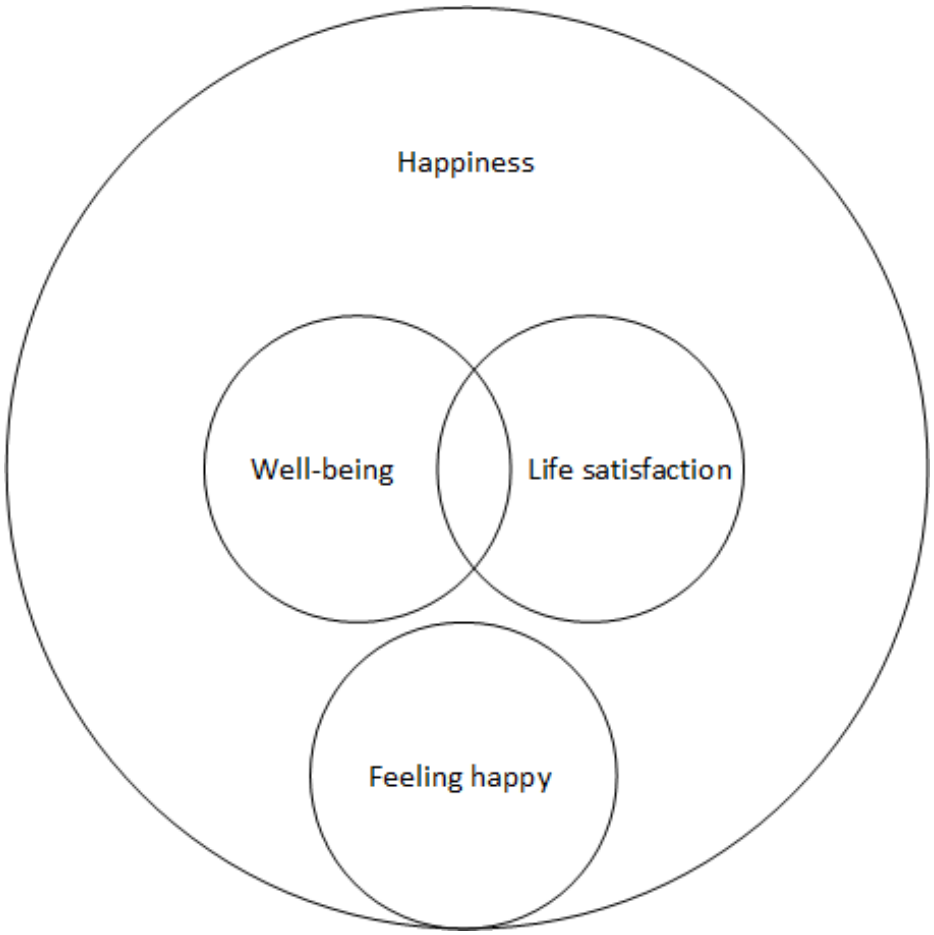
P2Y3AM
Happiness
int1: Balance between happiness and unhappiness, no happiness without unhappiness, things that make them unhappy, cold shower, unhappiness increases appreciation for happiness
int2: To be in a positive state of mind

Well-being:
int1: general term for both physical and mental health of a person
int2: incorporates mental and physical state into one general state of being positive

Life satisfaction:
int1: Tied to well-being, more emphasis on psychological side of things
int2: Well-being measured across your entire life or a longer period (at least across a year)

Feeling happy:
int1: Being happy with being alive, appreciating the moment of the day, becoming more playful and lightening up a bit, not being so serious
int2: A state where you're not focusing on anything else, being in the moment, being happy

Connection: Well-being and life satisfaction are connected, happiness influenced by all, feeling happy a bit more disconnected
Required: No, they are linked but aren't necessary for one another.
Synonymity: They are similar but subjective, hard to find an objective measure for each of the concepts.
Diagram:



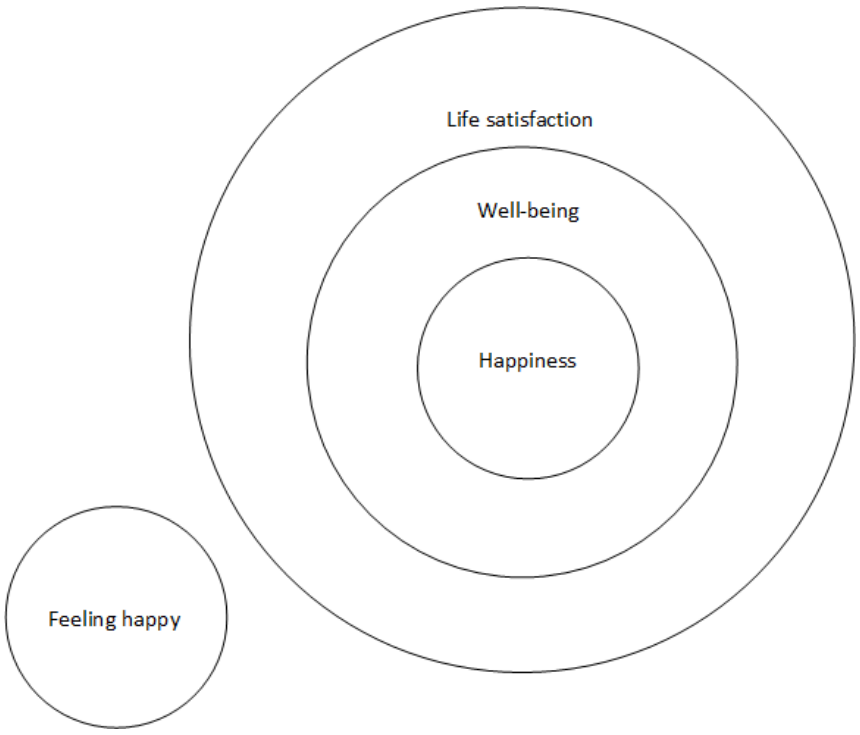
P3X4RF
Happiness:
int1: Success will not lead to happiness, Warmth, yellow (Favorite colour), Food, feeling rested, Content, feeling like there's no weight on you or in life, going with the flow
int2: Something yellow, going with the flow, makes you want to continue

Well-being:
int1: Mental health, general health, good relationship with food, feeling light, not having any pain (unless it's chronic), avoiding self-sabotage, avoiding self-destructive behaviours
int2: physical and mental health, feeling okay, physically little complaints, don't feel something holding you back

Life satisfaction:
int1: Success, happiness, well-being, having great relationships with people you want to keep in your life, avoiding self-destruction, taking care of yourself, listening to yourself, loving yourself, allowing yourself to enjoy life, being physically and mentally healthy and happy,
int2: Having happiness, well-being, good health, having achieved everything or at least something in life, moving towards something, not just existing, having goals, having dreams

Feeling happy:
int1: Walking under the rain, having good food, knowing that you allow yourself to eat it, hanging out with people you like, playing with pets, doing something that makes you happy in the moment
int2: feeling like you have the potential to do a lot more than you're doing, feeling yellow, when you just want to run and jump around, get some nice food, take care of yourself

Connection: Well-being and happiness required for life satisfaction, well-being requires happiness, feeling happy separate from the other concepts
Required: Well-being and happiness required for life satisfaction, well-being requires happiness, feeling happy separate from the other concepts
Synonymity: Does not feel it's correct to use the concepts as synonyms, specification is needed
Diagram:



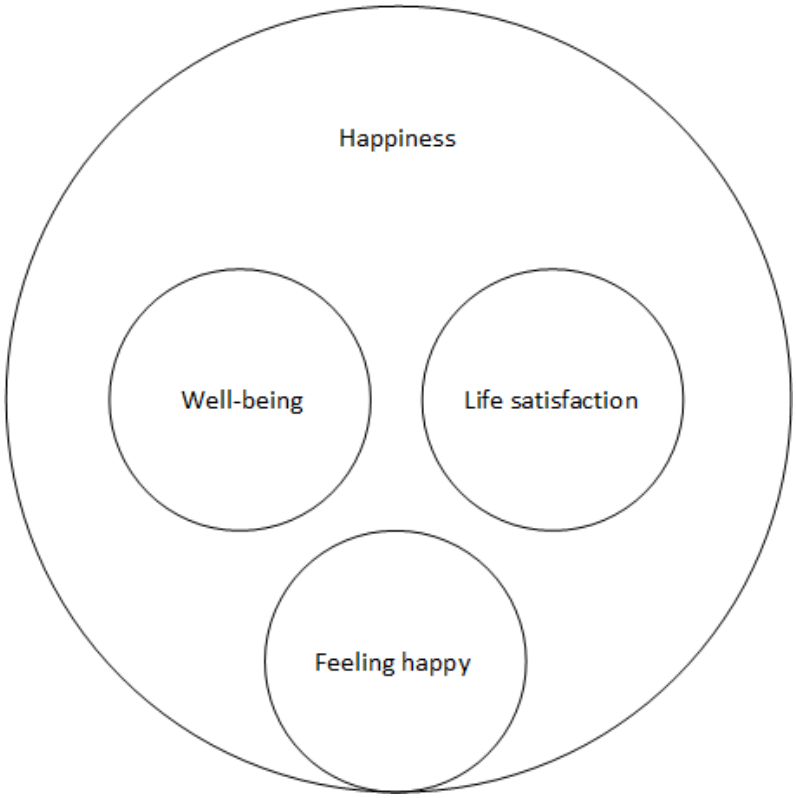
P4XoIF
Happiness:
int1: Comfort, feeling like you have achieved something, short term, persistent, good environment where you are included, (unexpected) Small things, acts of service
int2: Small things that stand out, not constant, feel like you are having a good few days, feel like you're doing what you want to do, feel like you achieved something

Well-being:
int1: Mental health, physical health, feeling safe, having good coping mechanisms, a good environment
int2: Health is a part of it (not necessarily), lower levels of stress, short term stress is good

Life satisfaction:
int1: Achievement, happiness, being where you want to be, doing what you want to be doing,
int2: Feeling like you have achieved something, being satisfied with what you have in that moment, being happy with who you have around you, feeling like you have what you physiologically and emotionally have what you need

Feeling happy:
int1: Smiling, sudden realization of happiness
int2: in the moment thing, doing something you like, small sections of the day

Connection: Happiness as the umbrella
Required: they are not required for another
Synonymity: Concepts should not be used interchangeably
Diagram:



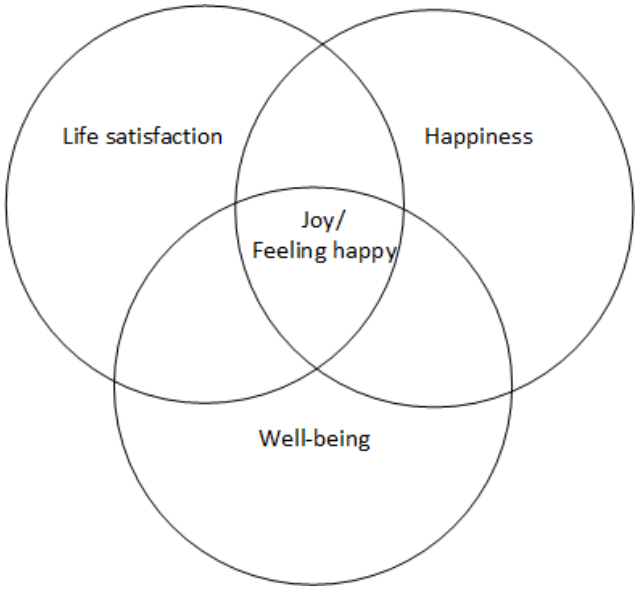
P5X6SM
Happiness:
int1: Generally just feeling good, enjoying life, enjoying being yourself, way of seeing/perceiving life, a feeling, feeling full of energy, being around friends, joy, having a good time
int2: Positive emotion, feeling, a mindset/framework to view the world influenced by the emotions you are currently feeling, feeling good about yourself, feeling good about being with friends, where you are in life.

Well-being:
int1: being in a physically and mentally healthy state, being happy more often than not, feeling of having good social connections, having a good balance of the things you need in life, food that's good for you, food that's bad for you but makes you feel good, activities you enjoy, spending time with friends, doing something you're passionate about, enjoying doing their studies, having a good mix of everything in life.
int2: Being in a state where you're happy more often than you are not, taking care of yourself, feeling physically better more often than not, getting good sleep, doing exercises, seeing friends, having a good balance in life of all those things.

Life satisfaction:
int1: Very subjective, treating themselves with kindness, being gentle to themselves and others, leaving the world a bit better than they found it, going for the thing that gets the best results trying your best that you can do it, leave at least somewhat an emotional impact on the people immediately around them, future-orientated
int2: An in the moment thing, enjoying the small moments, having a good time, what you managed to accomplish within your life

Feeling happy:
int1: Feeling physically good, lack of tiredness, lack of fatigue, satisfaction, joy, feeling physically and mentally good, feeling good about stuff you have accomplished or experienced,
int2: it's easier to feel happy when you have energy but not necessary, feeling good, enjoying things you're currently doing

Connection: Life satisfaction is dependent on the other three,, joy in the center
Required: Not required but connected
Synonymity: Considers that people have different associations with each word, would recommend allowing people to elaborate using open questions in questionnaire research
Diagram:



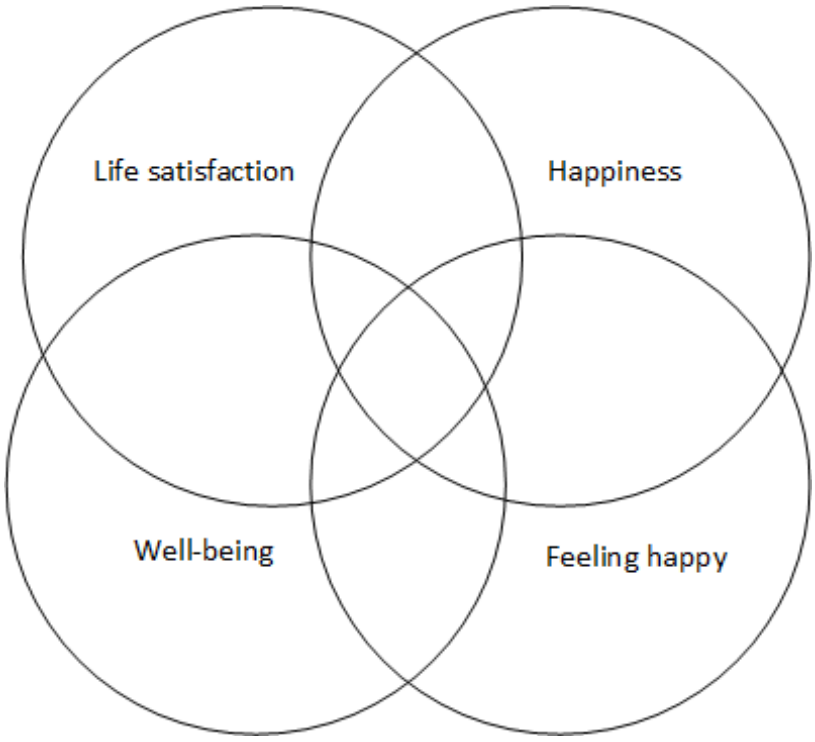
P6X9SF
Happiness:
Int1: Joy, subjective towards each and every person, sense of contentness and peace, feeling calm, feeling immersed in the moment, wouldn't want it any other way, state of mind
int2: Feeling of joy, euphoria, enjoyment, immediate feeling, generally happy, enjoying the direction your life is taking, long term, feeling life is fulfilling and meaningful

Well-being:
int1: General state of a person, general health, general life style, general physical and mental health, ties into happiness (happiness increases well-being), low amount of stressful environmental factors.
int2: One's life, links to health, feeling physically and emotionally well, how happy someone feels

Life satisfaction:
int1: Well-being, happiness, how content an individual feels with their life, tied to the person's environment and the life they live in and how that makes them feel, enjoying their job, relationships with people, general lifestyle and living standards, things that make a person feel safe and secure, getting all of their hierarchy of needs
int2: General attitude towards one's current life and the trajectory of it

Feeling happy:
int1: Something very meaningful, something very important, to have one constant strong flow of happiness, feeling fully safe, not feeling judged, feeling okay, enjoying everything that's happening and everyone around them
int2: Not feeling any form of discomfort, not feeling overwhelmed with anything, when something that brings a small form of joy is happening

Connection: All of these are blurry and interlinked
Required: Feels that it differs per person but personally feels they are required for each other
Synonymity: thinks they mean the same to a degree, the concepts can't really be split up
Diagram:



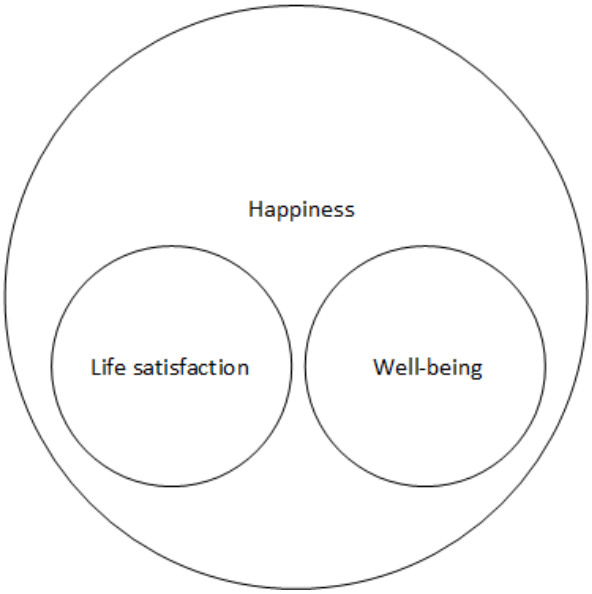
P7X8BF
Happiness:
int1: A feeling, physical feeling, sunny, friends
int2: A feeling within their body, related to experienced things in a specific moment, related to other people

Well-being:
int1: Mental well-being, what they're doing in their own head, what's going on in their own head, something to achieve
int2: Prolonged overall state, persistent happiness

Life satisfaction:
int1: Something an adult has, being satisfied with some things, being happy with something,
int2: Something you achieve later in life, something that requires (more) life experience, something to reflect on

Feeling happy:
int1: Feeling lighter, mainly a feeling in their chest, possibly difficult to recognize, a very quick process of realization,
int2: Happiness and feeling happy are the same

Connection: happiness required for life satisfaction and well-being, life satisfaction required for well-being and happiness
Required: happiness required for life satisfaction and well-being, life satisfaction required for well-being and happiness
Synonymity: Feels the concepts are not the same and shouldn't be used as synonyms
Diagram:



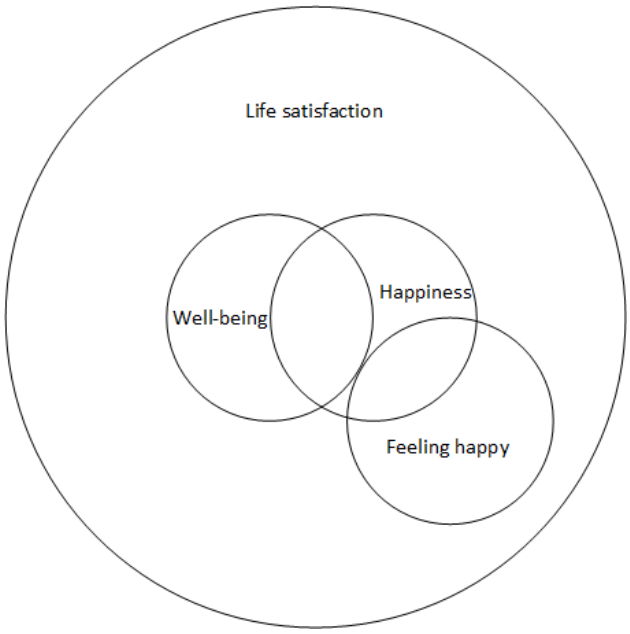
P8X1GN
Happiness:
int1: Type of people they love, social time with people they like, dancing, good food, absence of things that don't make you happy, distraction from unhappiness, fleeting
int2: fleeting, something to strive for, a mood, absence of negative thoughts, other people

Well-being:
int1: Having a bit of structure in life, being able to do the things you want to do, being physically healthy, mental health, absence of negative thoughts and feelings, feeling of belonging, feeling connected to other people, social life
int2: physical and mental health, absence of or control over mental illness/problems, something to strive for, less fleeting and more achievable than happiness, good balance and control over negative emotions

Life satisfaction:
int1: Feeling like you are following some kind of purpose, the ability to do things you want to do, a connection of belonging, being content with yourself and the place you are in, feeling loved, getting to a point where you feel comfortable to settle down,
int2: Includes happiness and well-being, being in a position you feel comfortable in, being on the right track towards an end goal in life where you would like to stand, social life, whether you regret your choices or not, what you do to get your money, having enough money

Feeling happy:
int1: Being in the moment, being energized, being bubbly and expressive, not being afraid or thinking of how others perceive in the moment, being more outward and inward focused
int2: a momentarily thing, feeling energetic and bubbly/outgoing/bouncy, being expressive, not feeling anxious, not filtering what they like to show the outside world or communicate to others, being in a social context with people they feel comfortable with, can shift quickly with small inputs to the opposite direction

Connection: Feeling happy and happiness relate to each other a lot, Happiness and well-being integrate, but well-being and feeling happy not necessarily, life satisfaction is all combined
Required: not always required
Synonymity: feels concepts are quite differently to them
Diagram:



P9X8NM
Happiness:
int1: When they are having a great time, playing an online game with their best friend, the feeling a moment could go on forever and not getting bored, a great experience
int2: Having a good time, feeling happy, wanting to continue a situation

Well-being:
int1: health wise or monetary wise being able to afford what you need and want to afford, being able to do every physical activity you want to do
int2: Being well, physically fit, not sick

Life satisfaction:
int1: Close to well-being, being able to do the things you want to do, having fun with whatever you're doing in life, the amount of happiness you get from everyday activities, the amount of happiness you get from being able to do activities
int2: Being able to do the things you want to do in life

Feeling happy:
int1: doing something fun and being happy with it, hard to differentiate from happiness, part of being happy
int2: Being in a positive mindset, enjoying their time

Connection: Life satisfaction connects well-being and happiness
Required: not required but help each other
Synonymity: feels it's incorrect to use the concepts interchangeably, more elaborate results might be missed
Diagram:

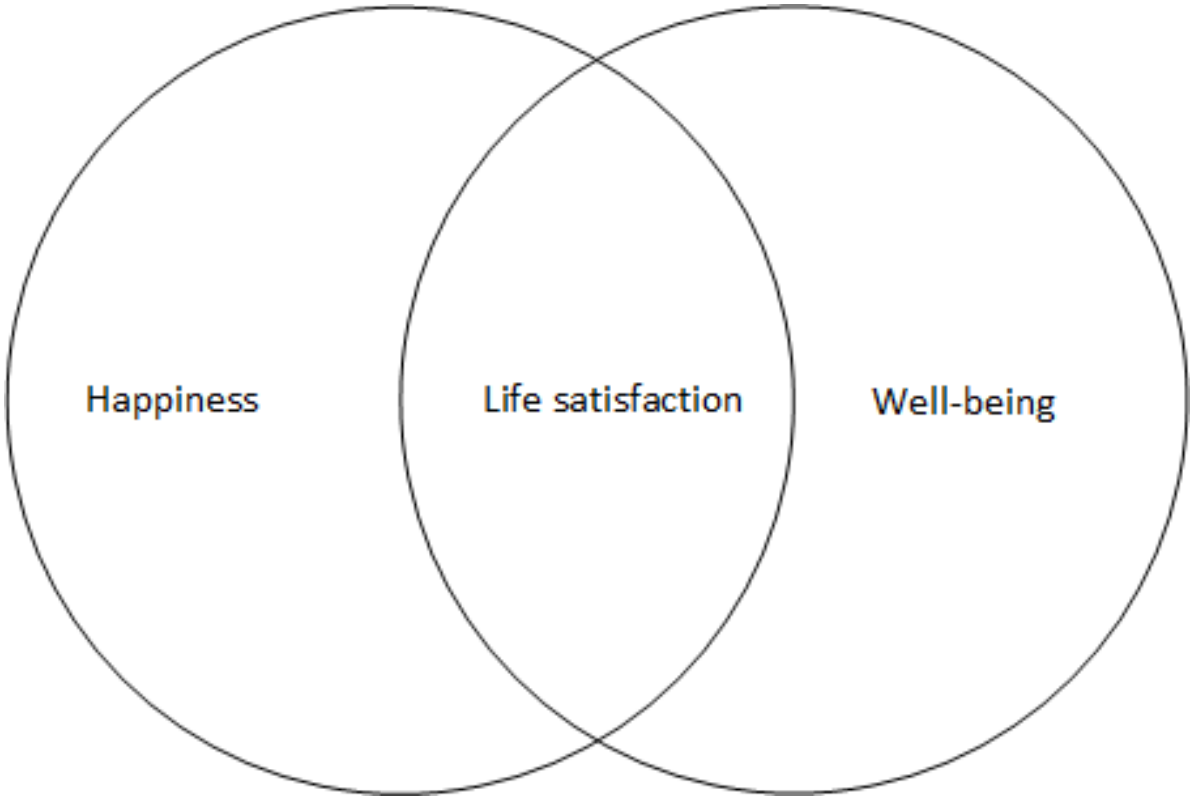


Figure 4
Individual EMA line per participant