Investigating the Role of Response Formats on Affect Dynamics in Experience Sampling Questionnaires for Positive affect items: A Focus on VAS versus Likert Scales

Nina Zarrin Tigh

Faculty of Behavioural Management and Social Sciences

Department of Psychology, University of Twente

Bachelor Thesis

BSc. Psychology

June 26, 2023:

Supervisor: Dr Jannis T. Kraiss Second Supervisor: Dr Thomas Vaessen

Abstract

Background: The Experience Sampling Method (ESM) gathers real-time data in naturalistic settings, enabling the capture of affect dynamics, including emotional variability, instability, and inertia. While Likert and VAS scales are commonly used in ESM studies, their association with different outcomes suggests that they may impact affect dynamics. However, the influence of response formats on affect dynamics remains unclear. This study aims to examine whether the use of VAS or Likert scales in ESM studies influences findings related to emotional variability, instability, and inertia of PA.

Method: Participants were recruited through convenience sampling, and PA was measured using ESM questionnaires. Participants were randomly assigned to either the Likert (N=35) or the VAS (N=36) condition. Participants completed ten ESM questionnaires per day for seven days. Fixed-effect and linear mixed models were used to analyse the impact of the response formats on the emotional variability, instability, and inertia of PA.

Results: No significant effect of condition was found for emotional variability (b= -0.05, p=0.34, 95% CI [-0.15, 0.05]), emotional instability (b= -0.-0.13, p=0.25, 95% CI [-0.36, 0.09]) and emotional inertia (b= -0.05, p=0.19, 95% CI [-0.03, 0.13]).

Conclusion: The findings of this study suggest that the response format used, whether Likert or VAS, has no significant impact on affect dynamics in ESM studies. Overall, this study provides useful insights into the influence of response formats on affect dynamics in ESM for PA items. However, further research is needed to investigate the impact of response format on ESM measures and thoroughly investigate the potential significance of habituation in this context.

Introduction

ESM is a widely used approach that allows researchers to collect real-time data in reallife settings over an extended time (Van Berkel et al., 2017). Participants in ESM studies can complete daily surveys, for example on their mobile devices, allowing them to respond in their natural surroundings. Change within a person can be measured over time using ESM (Csikszentmihalyi & Larson, 1987). Thus, ESM participants' daily self-reports are collected in a comparable way to traditional diary studies (Van Berkel et al., 2017). Unlike traditional diary research, however, participants in ESM studies are pro-actively prompted at predefined times throughout the day (Van Berkel et al., 2017). Thus, using this method, self-reports are gathered in context, reducing the need for participants to reconstruct earlier events and allowing the data to be focused on events of interest to the researcher (Van Berkel et al., 2017).

Many biases can arise while reconstructing previous events, one of which is hindsight bias, in which people believe that specific results were expected at the very beginning (Stahlberg & Maass, 1997). According to Trull and Ebner-Priemer (2009), retrospective approaches may be biased, raising concerns about the data's reliability and validity. This is a result of discrepancies between retrospective and real-time self-reports. However, the fact that ESM collects data in everyday situations, as opposed to traditional laboratory conditions, improves its external validity (Trull & Ebner-Priemer, 2009). ESM offers realtime data that can lessen recall bias, increasing to accurately represent participants' presentmoment emotional state (Trull & Ebner-Priemer, 2009). ESM additionally simplifies questions and gives real-time data, which helps people comprehend psychological events in practical contexts more precisely (Trull & Ebner-Priemer, 2009). As a result, ESM is a useful study approach that can successfully address the complexity of psychological phenomena in real-world contexts (Myin-Germeys et al., 2018).

Nowadays, there is an increase in the availability of mobile devices, allowing for autonomous observation and hence new opportunities in fields such as psychological study (Raento et al., 2009). This can lead to new opportunities in methodology and analysis of the collection of data. Nevertheless, since mobile devices are widely available, the integration of ESM with new technologies can increase data collection even further and make it more comfortable for participants to participate in an ESM study (Van Berkel et al., 2017).

Response Formats

To measure psychological phenomena, it is required to quantify such subjective thinking in a reliable and validated technique that can be examined using a variety of response formats (Joshi et al., 2015). The Likert scale and the visual analogue scale (VAS) are two response scales that are frequently used in ESM studies. The Likert scale is the most used response format in survey research (Krebs, 2012 & Allen, 2017). The Likert scale allows participants to show their level of agreement by selecting one of several answer options, each of which is accompanied by a distinct descriptor and numerical value (Borg, 2019). On a 5point Likert scale, response options may include "strongly disagree," "disagree," "neutral," "agree," and "strongly agree." (Borg, 2019). Another type of response format is VAS, which allows participants to rate their responses on a scale with only two endpoints, such as "extremely true" and "extremely false (S. Park & Xiong, 2021). The VAS presents the respondents with a horizontal line with two ends reflecting two opposite extreme values, thereby giving the participants a response continuum. This suggests that it might be less cognitively demanding than fully labelled scales since it is more precise and easier to remember (Moors et al., 2014). It's worth mentioning that VAS scales often span from 0 to 100, reflecting a continuous scale that, theoretically, offers more options for evaluating one's current state than Likert scales (Visuelle Analogskala Im Dorsch Lexikon Der Psychologie, 2019).

Furthermore, there are contradictories about whether Likert or VAS scale is effective and more reliable. According to Hasson and Arnetz (2005) and Kuhlmann et al. (2017), VAS scales are more effective in measuring accurate variable results. Studies examining the testretest reliability of response formats, particularly Likert scales, have revealed inconsistencies (Weng, 2004). Another study found no statistically significant difference in response format and no evidence that the Likert or VAS scale is superior (Alkadi et. al., 2022).

Response Styles in Likert and VAS Scales

It is worth noting that research on response styles in the VAS versus Likert conditions is limited and not uniform. As a result, the section that follows discusses various response styles, potential underlying reasons, and their relationship with VAS and Likert scales. However, overall, it has been demonstrated that there is a strong correlation between VAS and Likert scale responses, indicating the presence of relationships and variations in participants' responses across those scales (Hasson & Arnetz, 2005).

Response style refers to the tendency to respond in a somewhat consistent manner, regardless of the question content, which can contribute to systematic error (Allen, 2017). Moreover, when a response style is present, it can result in inaccurate measurements and alter the interpretation of the score (Park & Wu, 2019).

Response bias is noticeable in responses to Likert scales, which ask respondents to express their agreement or disagreement with a variety of statements (Smith, 2014). Specifically, in the context of Likert scales, respondents tend to exhibit a tendency known as central tendency, meaning they often avoid selecting extreme endpoints and instead favour responses closer to the midpoint (Douven, 2017). However, according to Park and Wu (2019), different response styles can be manifested when individuals respond to Likert scales and commonly reported response styles include extreme response style (ERS). Additionally, According to Weijters et al. (2021), respondents consider the visual midpoint of a rating scale to be the conceptual midpoint. As a result, visually distant endpoint categories on a scale are perceived as more extreme, leading to a higher likelihood of selecting those categories. Studer (2012) demonstrated that, when compared to Likert scales, VAS scale responses have a higher prevalence of ERS patterns.

While long questionnaires are commonly associated with ERS due to poor cognitive structure (Baumgartner & Steenkamp), Kieruj and Moors (2010) discovered that other factors such as demographics and personality traits may also have an impact on ERS. Furthermore, previous research has found that the length of the questionnaires, rather than their frequency, influence the burden experienced by participants (Eisele et al., 2020). Because of cognitive effort, heuristics such as "satisficing" and "anchoring" are therefore used (Kieruj & Moors, 2010). Satisficing refers to participants responding with the least amount of effort required, whereas anchoring refers to respondents relying on anchoring points rather than their genuine response (Kieruj & Moors, 2010).

Furthermore, when compared to Likert scales, VAS scales that rely solely on anchors may result in anchoring effects and thus ERS. However, because of the descriptions of the response option, satisficing, in which respondents opt for less cognitive effort, may be more prevalent in Likert scales, potentially leading to less genuine responses. However, the current literature lacks an in-depth investigation of the different impacts of Likert and VAS scales on response style, highlighting the need to investigate how these response formats influence participants' tendencies toward biases including extreme responding or midpoint preference.

Affect Dynamics

One often-used outcome in ESM studies is emotional dynamics. affect dynamics can be used to follow a person's shifting and fluctuating emotional and affective state over time, and ESM is often used to measure them (Hipson & Mohammad, 2021). There are recurring links between affect dynamics and psychological well-being. For instance, lower psychological well-being is associated with less stable, unstable, and inert emotions (Houben et al., 2015). However, it has been shown that PA indeed influences physical health. Including an association between PA and various health benefits, including reduced morbidity, decreased pain symptoms, and increased longevity among older individuals living within the community. Consequently, PA is strongly associated with protective responses that promote overall well-being (Pressman & Cohen, 2005). Three emotional dynamic patterns are often used: emotional variability, emotional instability, and emotional inertia (Houben et. al, 2015).

The range of swings in the emotional state across time can be understood as emotional variability. People with high emotional variability tend to experience extreme and more deviations from their usual emotional stage. Emotional variability is typically quantified using within-person standard deviation (Houben et. al, 2015). Emotional instability can be used to compare feelings from one moment to the next. As a result, people with higher levels of emotional instability vary their feelings from moment to moment, creating an unpredictable emotional life. The mean squared successive difference between consecutive emotion ratings is frequently used to assess emotional instability (Houben et. al, 2015). *Emotional inertia* is the tendency for one emotion to be carried from one instant to the next, or the maintenance of an emotional state. Thus, it indicates how predictable an emotional state might be from one moment to the next. Emotional inertia is usually operationalized as the autocorrelation of emotions across time. There is reason to believe that the choice of response scales could impact ESM study findings and emotional dynamics. As aforementioned, response styles can be influenced through the response scales and can lead to ERS. Therefore, the current study will focus on the influence of response scales on participants regarding PA items in ESM studies.

As previously stated, the format of response can influence how participants respond, potentially leading to misinterpretations of affect dynamics. A response style triggered by a response format may result in inflated or deflated data. When assessing affect dynamics, the Likert scale may result in a lower standard deviation (SD) compared to VAS. This, particularly in ESM studies, may have an impact on emotional variability. Since of the increased SD, the results of a VAS scale may result in higher emotional variability than a Likert scale. Participants in the Likert condition have a tendency toward responses closer to the midpoint, which may result in less extreme values of emotional variability, instability, and inertia. On the other hand, the length and cognitive effort involved in ESM studies can introduce extreme response styles, such as anchoring, which are more commonly associated with VAS scales. This, in turn, can lead to more extreme levels of emotional variability, inertia, and instability.

Current Study

The following randomized controlled trial aims to investigate the impact of VAS and Likert Scale response formats on participants' response styles, which will be analysed using PA composite scores from an ESM study. The following research question (RQ) will be the focus and will be tested in this thesis.

RQ. Is there an influence of VAS versus Likert scales on findings obtained from ESM studies regarding emotional variability, instability, and inertia of PA?

This study proposes three hypotheses to address the RQ based on existing literature that investigates the varying impacts of Likert and VAS scales in studies and their subsequent interpretations:

H1: Emotional variability will be higher in VAS compared to the Likert condition.

H2: Emotional instability will be higher in VAS compared to the Likert condition.

H3: Emotional inertia will be higher in VAS compared to the Likert condition.

Methods

The study was approved by the ethics committee with request number 22144 of the University of Twente. It is important to note that relevant data from two additional studies were included in this study, forming distinct cohorts.

Participants

The study aimed to recruit 136 participants. To accomplish this, each member of the two research groups was responsible for recruiting at least 15 participants, for a total of 120 participants. In this study, convenience sampling was used to select participants from the researchers' immediate surroundings because they were easy to reach. As a result, the researchers used personal contacts to find potential participants. To make the study more appealing, it was decided to have a small contest at the end of the study in which all participants were picked at random and the winner received a 40€ Amazon gift card. Furthermore, additional participants should be recruited using SONA (https://utwente.sona-systems.com/Default.aspx?ReturnUrl=%2f). Furthermore, the study will recruit additional participants through SONA, a platform available at (https://utwente.sona-systems.com/Default.aspx?ReturnUrl=%2f). This allows psychology students at the University of Twente to take part in research studies. Participation in SONA is mandatory for earning credits, as students are required to complete a minimum of 15 credits as part of their bachelor's degree requirements in psychology. To determine the required sample size for a

linear model, a priori power analysis was performed using G*Power (Faul et al., 2009). The analysis revealed that a sample size of 128 participants would be required to achieve a power of 8 with an effect size of F = 0.25 and a significance level of = 0.05. All participants were required to be 18 years old or older and have sufficient knowledge of English as well as a mobile device on which to download the ETHICA app ((https://ethicadata.com/) to fill out the ESM questionnaires. Additionally, before taking part in the study, the participants were required to provide informed consent.

Randomization

After recruiting participants, each researcher used stratified randomization to assign them to one of two conditions. This randomization procedure was carried out independently for each researcher's group of recruited participants. Each researcher made nine separate lists for this purpose. These lists were used to randomly assign participants to one of two conditions. This process was carried out using an online tool

(https://www.sealedenvelope.com/). The researcher divided the participants to ensure an even distribution, ensuring a balanced 50/50 ratio for each condition in both this study and the first cohort. It is important to note, however, that the second cohort was randomized with a 70/30 ratio between the conditions. This was done as a new condition was included to ensure equal allocations throughout the study. However, this did not work out, which is why the sample sizes for the two conditions differ slightly (see Figure 1). After participants provided their SONA number and email address, they were randomly assigned.

Participants recruited through SONA were assigned to one of two conditions at random.

Figure 1



Flow Chart of the samples across different cohorts.

Note: Participants who completed less than 33% of the questionnaires were excluded.

Design and Procedure

Three to four days before the study began, participants received an email corresponding to either the VAS or Likert condition. The email contained all the relevant information including the description, procedure, and purpose of the study, which can be found in Appendix A. The email also included a unique registration code as well as a link to the ETHICA platform, which allowed participants to register and participate in the study. Furthermore, before the start of the study, participants provided informed consent (see Appendix C).

As aforementioned, the data in this study consists of three distinct cohorts. The first cohort took place between Monday, November 11, 2022, and Sunday, November 13, 2022. This cohort specifically compared an endpoint-only labelled Likert scale to a fully labelled Likert scale. The second cohort began on Monday, February 13, 2023, and ended on Sunday, February 19, 2023, with the same conditions as the current study. Finally, the current study began on Monday, April 17, 2023, at 07:30 a.m. and ended on Sunday, April 2023 at 10:30 p.m.

All participants received the same questionnaire at the start of the study and were first required to fill out a lengthy questionnaire about their well-being and demographics. Following that, depending on whether they were in the Likert scale or VAS condition, they received much shorter surveys for the next seven days. The participants would be notified per assessment ten times per day to complete the questionnaire over the course of seven days, for a total of 70 surveys. The notification to complete the questionnaire was delivered at random within ten pre-defined intervals of 90 minutes. However, the questionnaire expired when participants were unable to finish it 15 minutes after being notified (see Table 1).

Measures

This study is part of a larger study, and only components that are directly relevant to this thesis were included. Specifically, demographics from the baseline questionnaire and PA measures from the ESM questionnaire.

Demographics

At the start of the study, all participants were asked to provide information that included their gender, occupation, age, nationality, email address, and educational background. Participants who received SONA points for their participation were required to provide their SONA identification number.

Baseline questionnaire

On Monday, April 17, 2023, at 7:30 a.m., all participants received the same questionnaire and were first required to fill out a lengthy questionnaire about their well-being and demographics.

ESM questionnaire

The PANAS scale (Positive and Negative Affect Schedule) was used in this ESM questionnaire to assess PA and NA, consisting of four items each. Four items (cheerfulness, enthusiasm, satisfaction, and relaxation) were used to measure PA.

All PA items were assessed using either a Likert scale or a VAS scale. In the Likert condition, participants were given seven answer options ranging from "strongly agree" (1) to "strongly disagree" (7), and they had to select the option that best represented their response. Participants in the VAS condition, on the other hand, used an analogue slider ranging from 0 to 100 to indicate their response on a continuum.

Table 1

Each day's study schedule included the type of questionnaire, relevant variables, specific time points, and expiration times.

Day	Questionnaire	Relevant	Time points	Expire time
		variables		
1	Baseline	All	At the	No
			beginning of	
			the study	
1-7	ESM	PA	7:30 - 9:00	After 15
			9:00 - 10:30	minutes
			12:00 - 13:30	
			13:30 - 15:00	
			15:00 - 16:30	
			16:30 - 18:00	
			18:00 - 19:30	
			19:30 - 21:00	
			21:00 - 22:30	

Data Analysis

For data analysis, the data was imported into RStudio version 4.3.0 (<u>https://www.rstudio.com/tags/homepage/</u>). The data analysis is divided into two steps: data cleaning and statistical analysis.

Data Cleaning

First, both the Likert and VAS datasets were combined, and a condition variable was added (0= Likert condition, 1= VAS condition). Then, an additional variable was created, the time variable, which indicated the specific assessment point for each data collection. Participants who had completed less than 33% of the questionnaires were excluded (i.e., fewer than 23 questionnaires). Furthermore, data with potential errors, including participants with an unusually large number of data points, were removed. To specifically analyse the data for PA, a composite score was calculated by aggregating the scores of all four items. Finally, a z-score standardization for PA was computed. The within-person standard deviation was computed to assess the influence of variability, the Mean Square of Successive Differences (MSSD) for instability, and inertia a lagged variable of PA was created.

Additionally, to obtain an understanding of the participant's demographic characteristics, descriptive statistics were calculated on the baseline questionnaire (see Table 2).

Statistical Analysis

Following that, a series of linear mixed models were used to analyse the data. These models took into account the nested structure of the data where individual observations were nested within participants by adding a random effect for participants. For the statistical analysis the package lme4 and for the visualizations the package ggplot was used.

To assess emotional variability and determine whether it is higher in the Likert condition compared to the VAS condition, a linear fixed-effect model was used. For this analysis, the independent variable was condition (VAS vs. Likert), while the dependent variable was emotional variability. To avoid a unit of analysis error, only one row per participant was considered for analysis.

A linear fixed-effect model and a significance test were used to assess emotional instability. Emotional instability was the dependent variable, while the condition (VAS vs. Likert) was the independent variable.

A linear mixed model was used to investigate emotional inertia as it effectively handles the nested structure of the data, where observations are nested within participants. The condition + PA(t-1) + PA(t-1) * condition was the independent variable, while PA(t) was the dependent variable. Notably, the model included an interaction of the lagged PA variable and condition. The covariance structure was set to autoregressive (1), and the estimation was done with restricted expectation maximum likelihood. The random effect accounted for random intercepts of participants. Furthermore, a multiple intercept model with PA as the dependent variable and the time-lagged variable as the independent variable was created. The model also accounted for the interaction effect of the condition variable (VAS vs. Likert) and the time-lagged variable, as well as the nested structure of the data with random intercepts for each unique participant (Device. ID). A boxplot was also created using the dependent variable PA and the independent variable condition. Statistical significance was determined using a p-value of 0.05.

Results

Demographics

This study initially involved a total of 162 participants. However, due to incomplete responses (at least 33% of all questionnaires), missing data due to application errors, and participants with an unusually high number of data points, resulting from a technical issue in the data collection app ETHICA, the final number of participants included in the analysis is 80. Most participants in this study were female (60.7%), with an average age of 28. Furthermore, most participants were German (92.6%), students (38.3%), and had completed their education up to the high school level (53.1%) (see Table 2). The distribution of relative frequencies is illustrated in Figure 2. It demonstrates that the distributions in both cases have identical kurtosis and skewness. This means that the distributions are comparable and show similar patterns. Table 3 shows the relative frequencies of standardized PA scores for each condition.

Table 2

Sample Descriptives

Variable	Likert Condition (<i>nBaseline</i> = 28), n	VAS Condition (<i>nBaseline</i> = 35),
	(%)	n (%)
Mean Age (SD)	29.04 (12.7)	27.17 (10.8)
Initial sample size n	67	95
Gender		
Female	17 (60.7)	23 (65.7)
Male	11 (39.3)	12 (34.9)
Nationality		
German	27 (96.4)	34 (97.1)
Dutch	1 (3.6)	1 (2.9)
Occupation		
Working	12 (42.9)	10 (28.6)
Student	8 (28.6)	12 (34.3)
Studying and Working	8 (28.6)	11 (31.4)
Self-employed	0	1 (2.9)
Other	0	1 (2.9)
Education		
High School	9 (32.1)	16 (45.7)
Bachelor	13 (46.4)	11 (31.4)
Master	5 (17.9)	8 (22.9)
Other	1 (3.6)	0
PA Composite	Likert (<i>nESM</i> = 47)	VAS (nESM=33)
Mean PA (SD)	4.05 (1.30)	51.1 (19.7)
Median PA	4	53.1

Table 3

Relative Frequencies of Observations of the PA Composite Scores for the Likert and VAS Condition

	Li	Likert (<i>n</i> =2014)		VAS (<i>n</i> = 1371)		
Category	Count	%	Count	⁰∕₀		
1	42	2.24	52	4.07		
2	236	12.59	132	10.33		
3	288	15.37	221	17.29		
4	661	35.27	318	24.88		
5	357	19.05	362	28.33		
6	254	13.55	164	12.83		
7	36	1.92	29	2.27		

Visualizations

Figure 2

Histogram of SD PA scores Measures. On the left side is the Likert condition (0), and on the right side is the VAS condition (1).

Histogram of Positive Affect (Likert Condition) Histogram of Positive Affect (VAS Condition) 300 400 -300 -200 -Leduency Frequency Frequency 100 -100 -0 0 -1 1 -3 2 -3 0 2 0 1 -2 -1 -2 Positive Affect (Likert) Positive Affect (VAS)

Figure 3

Multiple intercept model of PA SD and the lagged PA. The colour red illustrates the Likert condition (=0) and the colour blue the VAS condition (=1).



The multiple intercept model doesn't show significant differences in affect dynamics between the VAS and Likert conditions indicating similar distributions. (See Figure 3).

Figure 4

Boxplot of PA composite (t) Measures of emotional inertia. The Likert condition is on the left side (=0) and the VAS condition is on the right side (=1).



Figure 4 shows that the boxplots of both conditions are very similar and thus nonsignificant, but it is worth noting that there is a minor difference in the median of both conditions, indicating similar distributions and spread of the data in the central tendency.

Hypothesis Testing

H1: Emotional variability will be higher in VAS compared to the Likert condition.

There was no significant effect of condition on emotional variability of PA (b= -0.05, p = 0.34) (see Table 4).

Table 4

Coefficients Regression Analysis PA for emotional variability

Term	Std. Estimate	SE	CI (95%)	t-value	р
Condition (Likert	-0.05	0.05	[-0.15, 0.05]	-0.95	0.34
=0, VAS=1)					

Note. $N = CI = confidence interval, 80, R^2adjusted = 0.01, CI = Confidence interval, Degrees of freedom=78$

H2: Emotional instability will be higher in VAS compared to the Likert condition.

There was no significant effect of condition on emotional instability of PA (b= -0.13, p = 0.25) (see Table 5).

Table 5

Coefficients Regression Analysis PA for emotional instability

Term	Std. Estimate	SE	CI (95%)	t-value	р
Condition (Likert	-0.13	0.11	[-0.36, 0.09]	-1.17	0.25
=0, VAS=1					

Note. CI = confidence interval, N = 80, R^2 adjusted = 0.02, Degrees of freedom=78

H3: Emotional inertia will be higher in VAS compared to Likert condition.

There was no significant effect of condition on emotional inertia of PA (b= 0.05, p = 0.19) (see Table 6). Figure 3 shows a multiple intercept model and Figure 4 illustrates boxplots of the two conditions.

30 0	2 0				
Term	Std. Estimate	SE	CI (95%)	t-value	р
Condition (Likert=0,	-0.05	0.05	[-0.20, 0.11]	-0.57	0.5paP7
VAS=1)					
PA(t-1)	0.46	0.03	[0.41, 0.51]	18.08	0.00
PA(t-1)	0.05	0.04	[-0.03, 0.13]	1.29	0.19
*Condition					

 Table 6

 Coefficients Regression Analysis PA for emotional inertia

Note. CI = confidence interval, N = 80, R^2 adjusted = 0.01, Degrees of freedom=1960

Discussion

This thesis aimed to examine the influence of Likert and VAS scales on participant responses in ESM studies, with a particular focus on PA measures. The results of the linear mixed model analysis indicated that neither the Likert nor VAS scales had significant effects on the emotional variability, emotional instability, or emotional inertia of PA. These findings indicate that the scale used (Likert or VAS) had no significant impact on affect dynamics as typical outcomes of ESM.

Main Findings

According to the findings, there were no significant differences in emotional variability, emotional instability, or emotional inertia based on response format. These findings contradict the existing reviewed literature. As a result, all three hypotheses must be rejected: H1 (Emotional variability will be higher in the Likert condition compared to the VAS condition), H2 (Emotional instability will be higher in the Likert condition compared to the the VAS condition), and H3 (Emotional inertia will be higher in the VAS condition compared to the VAS condition).

As previously stated, the choice of response scales in research studies varies greatly, and no consensus exists on which scale is superior for the selection of ESM design in order to obtain genuine responses. Even though many researchers select a response format based on their personal preferences, the choice of a response scale should be determined by the specific study and its objectives (Gries et al., 2018). Depending on the needs and goals of the study, different response scales may be required. Previous studies have shown that the response format influences affect dynamics. Specifically, it has been found that certain response styles, such as anchoring or satisficing, may occur as a result, which are heuristic behaviours that can result in less genuine responses (Kieruj & Moors, 2010). However, the findings of this study suggest that using the Likert scale or the VAS scale to analyse affect dynamics has no significant impact on the results. This is beneficial because it allows researchers to choose the scale that is best suited to the context of their study.

However, it is important to recognize that several factors may have contributed to this study's non-significant findings, which may have outweighed the influence of the response format. To begin, it is crucial to remember that the current study has primarily focused on PA, implying that the results may differ significantly when examining other emotions. This would mean that the findings could have been different for negative affect items (NA). The nature of ESM itself may contribute to different response styles. Previous research suggests that it is not the frequency of the questionnaires that causes an increased burden, but rather the length of the questionnaires (Eisele et al., 2020). Furthermore, other variables, such as cultural background or individual characteristics, can interact with response formats, indicating the presence of confounding or moderating variables that can influence participants' response styles and perceptions of various response formats. Cultural factors, for instance, could have an impact on ERS. According to Hui and Triandis (1989), Hispanics have higher ERS than non-Hispanics. It should be noted, however, that there was no significant relationship between ERS and age (Light et al., 1965). Furthermore, Arce-Ferrer (2006) identified other factors that can influence response styles beyond the choice of response scale. These factors include familiarity with the response scale, communication style, and social norms.

Overall, the findings of this study suggest that the response format does not have an impact on affect dynamics. This goes align with the findings of Kuhlmann et al. (2017). Kuhlmann et al. (2017) conducted a study comparing a VAS with a Likert scale in a within-subject design, and the results revealed: The researchers found a high correlation between the two scales, indicating consistency in participants' responses. Furthermore, the internal reliabilities, means, and standard deviations did not differ significantly between the two scales. Furthermore, the Personality measure included in the study had no significant effect on the results. These findings imply that both the VAS and Likert scales may yield similar outcomes and provide reliable measures for assessing variables.

Strengths and Limitations

One limitation of this study is that the researchers used the convenience sampling method, which involved recruiting participants from their immediate environment. As a result, there was a greater likelihood of recruiting students and individuals with a high school education level. As a result, the generalizability of this thesis' findings may be limited.

Furthermore, because students are more familiar with completing questionnaires, they may have been less likely to be influenced by the questionnaire design, which could have impacted the hypothesis's significance. A further limitation was the low compliance rate among participants, which resulted in the exclusion of 143 people during the data-cleaning phase. The frequency of the questionnaires, according to feedback from these participants, was a significant factor contributing to their inability to complete them. Consequently, the low compliance rate resulted in a smaller sample size than initially anticipated. Another limitation of this study is the repetition of questionnaires, which may cause participants to respond out of habit or with decreased attentiveness, which is a feature of ESM studies.

However, one of the research's strengths is that it provides valuable insights into research choices, particularly in ESM studies, as one of the first studies that focused specifically on PA and the influence of Likert and VAS scales. Another strength of this study is the use of a randomized controlled trial (RCT), which is a relatively uncommon methodological approach in the design choices in ESM studies. The use of an RCT increases the research's accuracy and validity by randomly assigning participants to different conditions, wing for a better assessment of the causal effects being investigated.

Future Directions

The potential influence of scale design on participants is a crucial aspect to consider. Therefore, it is important for future directions to include a more diverse range of social groups in the sample. By doing so, it becomes possible to account for any potential impact that the occupation of being a student or having a high school education level may have on the study results. This broader inclusion of participants with different backgrounds will help to ensure a greater variability and improve the overall quality of the study, it may be beneficial to investigate sampling methods other than convenience sampling. Furthermore, a better understanding of moderating variables could help in better understanding the impact of response styles.

Furthermore, the nature of the ESM study, which involves repeated questionnaires, must be considered. This characteristic may have contributed to a non-significant effect by causing non-attentive responses. Individuals may engage in habitual responding, which means they might not employ the entire range of possible responses, according to Scollon et al. (2003). To account for the potential influence of habitual responding, future research should consider changing the sequencing of the questionnaires. This modification could account for the possibility of participants becoming familiar with the questionnaires while also improving the general reliability of the findings. Furthermore, researchers use pilot testing and refining

procedures to address this issue and ensure more accurate data collection (Scollon et al., 2003). Overall, gaining a better understanding of moderating variables would contribute to understanding the impact of response styles. Researchers can learn more about how these variables interact with response formats and provide insight into the nuanced factors that shape participants' responses by identifying and examining them. This could be researched effectively using Experimental Manipulation. Thus, designing experiments in which different variables (e.g., demographic factors, emotional states) and response formats (e.g., Likert scale, VAS) are systematically manipulated to examine their combined effects on participants' responses. This enables a controlled investigation of how variables interact with response formats.

Implications

This study provides new insight into the impact of scales, particularly Likert and VAS scales, on participants' responses to PA measures in ESM studies. The results of this study show that the type of scale used, whether a Likert or a VAS scale, has no significant effect on PA when studying emotional variability, instability, and inertia related to PA in ESM studies. This implies that researchers are free to use either scale without the fear of obtaining different mean values. As a result, scale selection is more flexible, allowing researchers to consider factors such as participant preferences, study design, and ease of administration when choosing between the two scales. These non-significant findings allow researchers to make scale choices based on other factors rather than solely on their potential impact on participant responses.

Conclusion

The purpose of this ESM study was to compare VAS and Likert scales and investigate their impact on emotional variability, instability, and inertia using PA items. The findings of this study did not reveal any significant influence of response formats on affect dynamics, which partially contradicts the existing literature. Future research should aim to replicate these findings and highlight the benefits of ESM in collecting real-time data, even though questionnaire repetition may lead to habituation. Therefore, it is important to note that future research should broaden its focus to include other aspects of affect dynamics and take participant engagement into account to improve the study design.

References

- Alkadi, L. T., Masuadi, E., Mohamed, T. A., Mohamud, M. S., & Farook, F. F. (2022). A Likert Scale Versus a Visual Analogue Scale and the Participant Response: A Cross Sectional Study. *Journal of International Dental and Medical Research*, 15(1), 255-262.
- Allen, M. (2017). The SAGE Encyclopedia of Communication Research Methods. SAGE Publications, Inc. eBooks. <u>https://doi.org/10.4135/9781483381411</u>
- Arce-Ferrer, A. J. (2006). An Investigation Into the Factors Influencing Extreme-Response Style. *Educational and Psychological Measurement*, 66(3), 374–392. <u>https://doi.org/10.1177/0013164405278575</u>
- Berry, D. S., & Hansen, J. S. (1996). Positive affect, negative affect, and social interaction. *Journal of Personality and Social Psychology*, 71(4), 796–809. <u>https://doi.org/10.1037/0022-3514.71.4.796</u>
- Borg, I. (2019). Likert-Skala im Dorsch Lexikon der Psychologie. <u>https://dorsch.hogrefe.com/stichwort/likert-</u> <u>skala#search=721af43403aeda24457d1e7eab0f1f08&offset=0</u>
- Csikszentmihalyi, M., & Larson, R. W. (1987). Validity and Reliability of the Experience-Sampling Method. *Journal of Nervous and Mental Disease*, *175*(9), 526–536. https://doi.org/10.1097/00005053-198709000-00004
- Dockray, S., & Steptoe, A. (2010). Positive affect and psychobiological processes. *Neuroscience &Amp; Biobehavioral Reviews*, 35(1), 69–75. <u>https://doi.org/10.1016/j.neubiorev.2010.01.006</u>
- Douven, I. (2017). A Bayesian perspective on Likert scales and central tendency. *Psychonomic Bulletin & Review*, 25(3), 1203–1211. <u>https://doi.org/10.3758/s13423-017-1344-2</u>
- Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2020). The Effects of Sampling Frequency and Questionnaire Length on Perceived Burden, Compliance, and Careless Responding in Experience Sampling Data in a Student Population. *Assessment*, 29(2), 136–151. https://doi.org/10.1177/1073191120957102
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power
 3.1: Tests for correlation and regression analyses. *Behaviour Research Methods*, 41(4), 1149–1160. https://doi.org/10.3758/brm.41.4.1149
- Folkman, S. (1997). Positive psychological states and coping with severe stress. *Social Science* &*Amp; Medicine*, 45(8), 1207–1221. <u>https://doi.org/10.1016/s0277-9536(97)00040-3</u>
- Furnham, A. (1986). Response bias, social desirability, and dissimulation. *Personality and Individual Differences*, 7(3), 385–400. <u>https://doi.org/10.1016/0191-8869(86)90014-0</u>

- Gries, K. S., Berry, P., Harrington, M., Crescioni, M., Patel, M. C., Rüdell, K., Safikhani, S., Pease, S., & Vernon, M. (2018). Literature review to assemble the evidence for response scales used in patient-reported outcome measures. *Journal of Patient-Reported Outcomes*, 2(1). https://doi.org/10.1186/s41687-018-0056-3
- Hasson, D., & Arnetz, B. (2002). Validation and Findings Comparing VAS vs. Likert Scales for Psychosocial Measurements. *International Electronic Journal of Health Education*, 8, 178–192.
- Hipson, W. E., & Mohammad, S. M. (2021). Emotion dynamics in movie dialogues. *PLOS ONE*, 16(9), e0256153. https://doi.org/10.1371/journal.pone.0256153
- Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, 141(4), 901–930. <u>https://doi.org/10.1037/a0038822</u>
- Hui, C. H., & Triandis, H. C. (1989). Effects of Culture and Response Format on Extreme Response Style. Journal of Cross-Cultural Psychology, 20(3), 296–309. https://doi.org/10.1177/0022022189203004
- Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert Scale: Explored and Explained. British Journal of Applied Science and Technology, 7(4), 396–403. <u>https://doi.org/10.9734/bjast/2015/14975</u>
- Kieruj, N. D., & Moors, G. (2010). Variations in Response Style Behavior by Response Scale Format in Attitude Research. *International Journal of Public Opinion Research*, 22(3), 320–342. https://doi.org/10.1093/ijpor/edq001
- Kuhlmann, T., Dantlgraber, M., & Reips, U. (2017). Investigating measurement equivalence of visual analogue scales and Likert-type scales in Internet-based personality questionnaires. *Behavior Research Methods*, 49(6), 2173–2181. https://doi.org/10.3758/s13428-016-0850-x
- Krebs, D. (2012). The impact of response format on attitude measurement. *VS Verlag Für Sozialwissenschaften eBooks*, 105–113. <u>https://doi.org/10.1007/978-3-531-18898-0_14</u>
- Light, C., Zax, M., & Gardiner, D. H. (1965). Relationship of age, sex, and intelligence level to extreme response style. Journal of Personality and Social Psychology, 2(6), 907–909. https://doi.org/10.1037/h0022746
- Matejka, J., Glueck, M., Grossman, T., & Fitzmaurice, G. (2016). The Effect of Visual Appearance on the Performance of Continuous Sliders and Visual Analogue Scales. *Human Factors in Computing Systems*. <u>https://doi.org/10.1145/2858036.2858063</u>
- Morishita, S., & Aoki, S. (1999). Clonazepam in the treatment of prolonged depression. *Journal of Affective Disorders*, 53(3), 275–278. <u>https://doi.org/10.1016/s0165-0327(98)00122-0</u>

- Moors, G., Kieruj, N. D., & Vermunt, J. K. (2014). The Effect of Labeling and Numbering of Response Scales on the Likelihood of Response Bias. *Sociological Methodology*, 44(1), 369– 399. https://doi.org/10.1177/0081175013516114
- Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry*, 17(2), 123–132. <u>https://doi.org/10.1002/wps.20513</u>
- Park, M., & Wu, A. D. (2019). Item Response Tree Models to Investigate Acquiescence and Extreme Response Styles in Likert-Type Rating Scales. Educational and Psychological Measurement, 79(5), 911–930. <u>https://doi.org/10.1177/0013164419829855</u>
- Park, S., & Xiong, S. (2021). The effect of slider design and length on user performance and preference of smartphone versions of the visual analogue scale. Applied Ergonomics, 97, 103521. <u>https://doi.org/10.1016/j.apergo.2021.103521</u>
- Pressman, S. D., & Cohen, S. (2005). Does positive affect influence health? *Psychological Bulletin*, 131(6), 925–971. <u>https://doi.org/10.1037/0033-2909.131.6.925</u>
- Raento, M., Oulasvirta, A., & Eagle, N. (2009). Smartphones. *Sociological Methods & Research*, 37(3), 426–454. <u>https://doi.org/10.1177/0049124108330005</u>
- Scollon, C. N., Kim-Prieto, C., & Diener, E. (2003). Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. Journal of Happiness Studies, 4(1), 5–34. https://doi.org/10.1023/a:1023605205115
- Shek, D. T. L., & S, C. M., MA. (2011). Longitudinal Data Analyses Using Linear Mixed Models in SPSS: Concepts, Procedures and Illustrations. *The Scientific World Journal*, 11, 42–76. https://doi.org/10.1100/tsw.2011.2
- Smith, P. (2014). Response Bias(es). In Springer eBooks (pp. 5539–5540). https://doi.org/10.1007/978-94-007-0753-5_2503
- Stahlberg, D., & Maass, A. (1997). Hindsight Bias: Impaired Memory or Biased Reconstruction? European Review of Social Psychology, 8(1), 105–132. <u>https://doi.org/10.1080/14792779643000092</u>
- Studer, R. (2012). Does it matter how happiness is measured? Evidence from a randomized controlled experiment. *Journal of Economic and Social Measurement*, 37(4), 317–336. <u>https://doi.org/10.3233/JEM-120364</u>

- *Tendenz zu Extremwerten im Dorsch Lexikon der Psychologie*. (2019). https://dorsch.hogrefe.com/stichwort/tendenz-zuextremwerten#search=9a6fc875b1a5f19491136aaf1b9e5046&offset=1
- Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment*, 21(4), 457–462. https://doi.org/10.1037/a0017653
- Van Berkel, N., Ferreira, D., & Kostakos, V. (2017a). The Experience Sampling Method on Mobile Devices. ACM Computing Surveys, 50(6), 1–40. https://doi.org/10.1145/3123988
- Van Berkel, N., Ferreira, D., & Kostakos, V. (2017b). The Experience Sampling Method on Mobile Devices. ACM Computing Surveys, 50(6), 1–40. <u>https://doi.org/10.1145/3123988</u>
- Weijters, B., Millet, K., & Cabooter, E. (2021). Extremity in horizontal and vertical Likert scale format responses. Some evidence on how visual distance between response categories influences extreme responding. *International Journal of Research in Marketing*, 38(1), 85– 103. https://doi.org/10.1016/j.ijresmar.2020.04.002
- Wendt, L. P., Wright, A. G. C., Pilkonis, P. A., Woods, W. G., Denissen, J. J. A., Kühnel, A., & Zimmermann, J. (2020). Indicators of Affect Dynamics: Structure, Reliability, and Personality Correlates. *European Journal of Personality*, 34(6), 1060–1072. <u>https://doi.org/10.1002/per.2277</u>
- Weng, L. (2004). Impact of the Number of Response Categories and Anchor Labels on Coefficient Alpha and Test-Retest Reliability. *Educational and Psychological Measurement*, 64(6), 956– 972. <u>https://doi.org/10.1177/0013164404268674</u>

Appendix A

Enrolment Information Sheet: Likert Condition

Dear participant,

Thank you for your participation in the study on mental health in daily life. We are contacting you because you kindly agreed to participate in this study for the bachelor psychology at the University of Twente.

A brief summary of the project

The study you are participating in is a daily diary study. With this study, we want to investigate how people feel and react to events in their day-to-day lives. By asking a few questions at several moments throughout the day, we get an insight into behavior of people in their everyday environment, which is necessary if we want to understand how people behave and feel in daily life. You will receive a notification on 10 random moments a day to answer a short questionnaire which will take about 1 minute to complete. We ask you to do this for 7 days in a row. The first questionnaire will be send on Monday morning, the 17th of April. Of course, there are situations in which it is not possible to fill it out (such as when you are driving), but to get a good overview of your daily life <u>it is important that you fill out as many of these questionnaires as possible.</u> In addition to these short questionnaires, you will receive one questionnaire in the beginning of the study that takes about 20 minutes to complete. It's important that you complete this questionnaire as well.

How to get ready to participate

Before continuing, make sure to download the Ethica application on your smartphone. Clicking on the following links on your smartphone will bring you the app store.

Android:

https://play.google.com/store/apps/details?id=com.ethica.logger&hl=en_US&gl=US&pli=1 IOS: https://apps.apple.com/nl/app/ethica/id1137173052

Then follow the these steps:

- Open the Ethica application on your phone. Please make sure to allow push notifications for the Ethica app on your phone!
- Click on "Sign up" and create an account.

- After you signed up in Ethica, login in to the Ethica application using your username and password.
- After logging in, click on the following link on your phone:

https://ethicadata.com/study/2349/

- Alternatively, you can also directly enter the registration code **2349** in the Ethica application.
- On the next window click on "Register" to enroll in the study.
- The study should now be set up and you will receive the first questionnaire next Monday.

Contact details

This study is part of a larger project with many researchers involved. If you have any questions, you can contact one of the following students who are involved in data collection or the supervisors. The contact details can be found below.

Students

Simon Brune (<u>eMail</u>) Nick Delventhal (<u>eMail</u>) Jan Derksen (<u>eMail</u>) Gina Haccou (eMail) Samuel Pietsch (<u>eMail</u>) Aleksandra Popovic (eMail) Lea Staudigel (eMail) Nina Zarrin Tigh (eMail)

Supervisors

Jannis Kraiss: eMail Thomas Vaessen: eMail

Thank you for participating in this study. Your contribution is greatly appreciated.

Kind regards, also on behalf of the whole study team,

Enrolment Information Sheet: VAS Condition

Dear participant,

Thank you for your participation in the study on mental health in daily life. We are contacting you because you kindly agreed to participate in this study for the bachelor psychology at the University of Twente.

Brief summary of the project

The study you are participating in is a daily diary study. With this study we want to investigate how people feel and react to events in their day-to-day lives. By asking a few questions at several moments throughout the day, we get an insight in behavior of people in their everyday environment, which is necessary if we want to understand how people behave and feel in daily life. You will receive a notification on 10 random moments a day to answer a short questionnaire which will take about 1 minute to complete. We ask you to do this for 7 days in a row. The first questionnaire will be send on Monday morning, the 17th of April. Of course, there are situations in which it is not possible to fill it out (such as when you are driving), but to get a good overview of your daily life <u>it is important that you fill out as many of these questionnaires as possible.</u> In addition to these short questionnaires, you will receive one questionnaire in the beginning of the study that takes about 20 minutes to complete. It's important that you complete this questionnaire as well.

How to get ready to participate

Before continuing, make sure to download the Ethica application on your smartphone. Clicking on the following links on your smartphone will bring you the app store.

Android:

https://play.google.com/store/apps/details?id=com.ethica.logger&hl=en_US&gl=US&pli=1 IOS: https://apps.apple.com/nl/app/ethica/id1137173052

Then follow the these steps:

- Open the Ethica application on your phone. Please make sure to allow push notifications for the Ethica app on your phone!
- Click on "Sign up" and create an account.
- After you signed up in Ethica, login in to the Ethica application using your username and password.

• After logging in, click on the following link on your phone:

https://ethicadata.com/study/1296/

- Alternatively, you can also directly enter the registration code **1296** in the Ethica application.
- On the next window click on "Register" to enroll in the study.
- The study should now be set up and you will receive the first questionnaire next Monday.

Contact details

This study is part of a larger project with many researchers involved. If you have any questions, you can contact one of the following students who are involved in data collection or the supervisors. The contact details can be found below.

Students

Simon Brune (<u>eMail</u>) Nick Delventhal (<u>eMail</u>) Jan Derksen (<u>eMail</u>) Gina Haccou (eMail) Samuel Pietsch (<u>eMail</u>) Aleksandra Popovic (eMail) Lea Staudigel (eMail) Nina Zarrin Tigh (eMail)

Supervisors

Jannis Kraiss eMail Thomas Vaessen eMail

Thank you for participating in this study. Your contribution is greatly appreciated.

Kind regards, also on behalf of the whole study team,

Appendix B Informed Consent

Dear participant,

Thank you for your participation in this study.

Brief summary of project

The study is using the Experience Sampling Method (ESM) to obtain data. This means that 10 times a day there will be a prompt to answer a questionnaire containing about 20 items, which will take about 1 minute to complete. The questions regard your psychological well-being in the specific moment you are receiving the questionnaire and the time in-between questionnaires. It is important to fill out as many questionnaires as possible to ensure the success of the project.

To participate in this study, we need to ensure that you understand the nature of the research, as outlined in the participant information sheet. Please confirm at the bottom of the page to indicate that you understand and agree to the following conditions:

- I confirm that I have read the participant information sheet for this study. I have had the opportunity to consider the information, ask questions, and have had these answered satisfactorily
- I understand that to take part in this study, I should
- Be at least 18 years old
- Possess a basic level of English
- I understand that personal data about me will be collected for the purposes of the research study including age, gender, nationality, level of education, current studies, and primary occupation, and this data will be processed completely anonymous and in accordance with data protection regulations.
- I understand that taking part in this study involves that I will be filling in 10 questionnaires every day for one week.
- I am voluntarily taking part in this research, and I know that I can stop the research at any time without giving any reason, without my rights being affected
- I don't expect to receive any benefit or payment for my participation.
- I understand that I am free to contact the researchers or supervisor with any questions I may have in the future.

- I understand that the data collected in this study will be anonymized, and only be used for academic purposes i.e., writing a thesis for the bachelor and/or master.
- I understand that personal data that will be collected within this study will not be shared with anyone other than the study team.
- I agree to take part in this study.

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by (eMail)

Students

Simon Brune (<u>eMail</u>) Nick Delventhal (<u>eMail</u>) Jan Derksen (<u>eMail</u>) Gina Haccou (eMail) Samuel Pietsch (<u>eMail</u>) Aleksandra Popovic (eMail) Lea Staudigel (eMail) Nina Zarrin Tigh (eMail)

Supervisors

Jannis Kraiss eMail Thomas Vaessen eMail

Appendix C

	R	lange
Category	Likert	VAS
1	1.00-1.25	0.00-14
2	1.50-2.50	14.75-28
3	2.75-3.25	28.25-42
4	3.50-4.50	42.25-57
5	4.75-5.25	57.25-71
6	5.50-6.50	71.25-85
7	6.75-7.00	85.50-100

Table A1

Ranges of Values for each Category for the Likert and VAS Condition