Studia Psychologica, Vol. 66, No. 2, 2024, 111-120
https://doi.org/10.31577/sp.2024.02.894

Happiness Underestimated

Zhou Fang[1], Yue-Mei-Tong Chen[2]
[1] Economic Sciences Department, St Mary's College of Maryland, USA
[2] Department of Psychology, University of California, San Diego, USA

In this paper, an exploratory study was conducted to investigate a potential measurement bias in happiness studies using 0-10 Likert scales. A total of 121 college students from a public university in California participated in the study by completing a hypothetical subjective well-being survey. The survey consisted of 23 scenarios with varying levels of happiness or unhappiness, and the students were asked to rate their level of happiness/unhappiness on the 11-point Likert Scales. The results showed that there was an inconsistent understanding of the scale when the expected value was between 7 and 8 on the 0 to 10 Likert scale, which leads to a higher variance and lower observation values. Based on these findings, it is suggested that an alternative scale length or a -5 to 5 scale should be considered to improve data quality.

Key words: Likert scale, inconsistent, happiness, decimal system

Introduction

Psychology studies often use Likert scales to measure people's opinions on various topics, such as stress (Weermeijer et al., 2022), emotional valence (Muller et al., 2022), and subjective well-being (Atalay & Barrett, 2022), among others. Among the different types of Likert scales, the 11-point scale (ranging from 0 to 10) is widely used, particularly in studies on subjective well-being. In fact, among the recent 100 subjective well-being papers published in two leading journals in the field, "The Journal of Happiness Studies" and "The Journal of Positive Psychology," 42 of them used a single term as SWB measurement, and 21 utilized the 0 to 10 Likert scale.

The main reason for this situation may be the researchers' adherence to convention, as many classic measurements of SWB use this standard. For example, the question "All things considered, how satisfied are you with your life?" was used in the famous 1976 report "The Quality of American Life: Perceptions, Evaluations, and Satisfactions" (Campbell et al., 1976), and has since been adopted by numerous sociologists, economists, and

Correspondence concerning this article should be addressed to Zhou Fang, Department of Economics, St Mary's College of Maryland, USA. E-mail: zhou.fang@cgu.edu

Received August 16, 2023

psychologists (Deter, 2021; Gerlich & Wolbring, 2021; Sohier et al., 2021; Wu, 2020). Other examples include the Happiness Index which also consists of just one 11-point scale question: "Do you feel happy in general?" (Abdel-Khalek, 2004).

Despite the widespread use of this scale in positive psychology, the authors recognize a potential misunderstanding, especially for many untrained questionnaire respondents. Unlike other scales, such as the 7-point scale, we are more accustomed to the 0 to 10 scale due to the prevalence of the decimal system in human daily life.

To illustrate the confusion which can arise when using the 0 to 10 Likert scale to rate something positively, consider the following simple question: What does an 8 of this scale mean to you – moderate, positive or very positive? While the correct answer is moderate, many people may interpret 8 as meaning very positive instead. This confusion arises because we are accustomed to using the decimal system, which counts in increments of 10, and thus associate the number 8 with a "greater number." Consequently, we may mistakenly transfer this association to the two-sided 0 to 10 scale used in questionnaires and confuse it with the one-sided 1 to 10 scale we use in our daily lives.

Due to our everyday use of decimal numbers, some individuals may associate the numbers 7 and 8 with fairly "large" quantities. However, on a 0 to 10 Likert scale, the meaning of these numbers on the positive side is counterintuitive for some people. Specifically, 7 indicates "below average," and 8 indicates "moderate," which can lead to confusion when interpreting the scale.

Table 1 provides a mapping of happiness levels in words to the numbers on an 11-point Likert scale. In addition to the more commonly used 0 to 10 version, a less frequently used -5 to 5 version is also presented. While people tend to have similar understandings of most levels, there is a potential issue with the understanding of the levels 7 and 8: some people might match "7 and 8" with "3 and 4", rather than "2 and 3".

Our focus on the 7 and 8 values on the 0-10 Likert scale is based on the unique nature of this scale in representing both positive and negative sentiments. In this scale, values from 0 to 4 indicate negative emotions, which sig-

Table 1 *Levels of Happiness on an 11-point Likert scale*

Level of Happiness	On a 0 to 10 scale	On a -5 to 5 scale
Extremely Unhappy	0	-5
Very Unhappy	1	-4
Unhappy	2	-3
Somewhat Unhappy	3	-2
Slightly Unhappy	4	-1
Neutral	5	0
Slightly Happy	6	1
Somewhat Happy	7	2
Happy	8	3
Very Happy	9	4
Extremely Happy	10	5

nificantly differ from the everyday usage of a 0-10 scale, where the lower end typically represents the absence of a trait or feeling to a lesser extent. Because these lower values (0-4) are conceptually distinct from their common daily usage, respondents are less likely to be influenced by habitual perceptions and more inclined to reflect on their true meanings when encountering unfamiliar scenarios.

In contrast, when responding to positive emotions, the bipolar scale's range of 6 to 10 is conceptually similar to the unipolar use of 6 to 10 in everyday life, where both are used to describe positive elements. This similarity can lead to heuristic processing by respondents, making them more prone to default to their habitual understanding of these numbers. As a result, this can cause confusion and error, particularly in the interpretation of the mid-range values such as 7 and 8.

To address this potential confusion, the right half of the 11-point scale (6 to 10) can be considered as a 5-point scale on the positive side. With the correct understanding, an 8 on the 0 to 10 scale is equivalent to a 3 on a "1 to 5 scale" on the positive side. However, individuals may still be inclined to confuse an 8 on the 0 to 10 scale with a 4 on a "1 to 5 scale" due to their daily intuition. If such inconsistency in understanding exists, the 0-10 scale would bring noises to the results. Therefore, this paper uses a simple experiment to investigate the following research question: is there a significant inconsistency in understanding the 0 to 10 Likert scales around 7 and 8? In other words, do some people understand 8 as "happy" and the others take 8 as "very happy" (see Table 1)?

Despite scientists having realized that dominance of counting in tens can also influence the neural representations of numeric values (Rotondaro et al., 2019), how the scale can be affected by such influences has never been studied. The most relevant research on this topic may be a study that matched the label and rating scales in subjective well-being measurements (Macri, 2017). In Table 4 of their paper, the authors showed that when people rate their well-being regarding family, an 8 is among the last tier of "very happy" using multivariate techniques. However, the authors only divided the results into four parts (not satisfied at all, not very satisfied, quite satisfied, and very satisfied), and this only occurred in one out of seven scales. Therefore, the phenomenon is not very apparent. The existing gap makes our study a novel and valuable contribution to many studies employing an 11-point Likert scale, given that over 20000 papers using such scale were published in 2023 (according to Google scholar search results).

Given the limited availability of prior research in this area, our study adopts an exploratory approach. In our study, we recruited participants and asked them to rate 23 hypothetical experiences. We randomly assigned participants to either the 0-10 scale or the -5 to 5 scale. The reason we used the -5 to 5 scale is that such scale is more intuitive. Positive numbers indicate happiness, negative numbers indicate unhappiness, and 0 means neutral. This scale allows participants to report their subjective well-being without confusion. Notably, a "3" on this scale is equivalent to an "8" on the original 11-point scale. A rating of 3 out of 5 appears more moderate compared to 8, and a rating of "2" is more explicit than "7". If asked to rate a 3 on a 1 to 5 scale, people would likely describe it as moderate, rather than very positive.

The -5 to 5 scale is yet not commonly used: most subjective well-being studies utilize either 0-positive or 1-positive scales, and mainstream scale comparison studies typically compare these two (Maggino & D'Andrea, 2003). While the -5 to 5 scale has been used in a few subjective well-being studies (Bern-

heim, 1999; Livovsky et al., 2021; Mazaheri & Theuns, 2009), the authors of these studies did not discuss why they chose this scale instead of the more commonly used 0 to 10 scale.

Our hypothesis was that if the majority of people rate an experience between 6 to 8 on a 0-10 scale (which is 1 to 3 on a -5 to 5 scale), the 0-10 group responses would have a higher variance due to more diverse understandings about 7 and 8 within the group. This would also lead to a smaller mean as those with false understandings may select smaller numbers.

Methods

For this study, 121 college students (78 female, 42 male, 1 non-binary) from a public university in California were recruited via a student email list. The average age of the students was 19.4 (*SD* = 1.64), with a minimum of 18 and maximum of 23, most of the students were freshmen or sophomores, with a major or interest in psychology.

Participants were compensated with $4 for their time, and the project (#210750SX) was reviewed by the IRB, which waived the need for ethics approval. Informed consent was obtained from all participants, and all methods were conducted in accordance with relevant guidelines and regulations.

Participants were directed to a Qualtrics link to complete a survey, which included 23 happy or unhappy events. They were asked to rate how happy or unhappy they would feel if the event happened to them on an 11-point Likert scale. Participants were randomly assigned to one of two groups: the first group reported their results on an 11-point scale ranging from 0 to 10, while the second group reported their results on a scale of -5 to 5. The entire procedure took approximately 7 minutes to complete. A list of the events can be found in the appendix (Table 5). Participants also completed self-report questionnaires at the end of the survey.

To analyze the results for each event, we first conducted an analysis of variance *F*-test to determine whether the two groups produced results with equivalent variances. The rationale behind this test is straightforward: if people have no difficulty using the 0-10 scale, the variances from both groups should be equal, since there is a perfect mapping between the two scales, and they have identical ranges. However, if the 0 to 10 group has a higher variance, it suggests that there is additional variation, likely arising from measurement error, beyond the variation due to different opinions.

For each question exhibiting significantly different variances between the two groups, we then controlled the relevant data and performed a linear regression to ascertain if the treatment significantly impacted the outcome.

Results

Descriptive Information

Table 2 displays the descriptive statistics of the answers to each question. To aid in comparisons, we converted the right half of the -5 to 5 scale into a 0 to 10 scale. To obtain the original results for mean and quantiles, subtract the values in Table 2 by 5.

Main Results

Table 3 presents the *F*-test results for variances, ordered by the number of 6 or 7 responses from the 0 to 10 group. Such order is based on the hypothesis that if respondents misinterpret values 7 and 8 as excessively high for expressing their happiness, they might conservatively opt for 6 or 7. Consequently,

a higher occurrence of 6 and 7 in responses could indicate a greater variance between the two groups, reflecting the misunderstanding of these scale values. The logic behind the variance test is that the inconsistent understanding will reduce the interrater reliability of the measurement and lead to a higher variance (Tinsley & Weiss, 1975).

The results of our experiment supported this hypothesis, showcasing significant differences in variance as predicted. The events with significantly different variances are: (1) You meet your friend's pet; (5) Go to your best friend's birthday party; (7) Have a delicious sandwich; (9) Get your first car; and (15) Meet your childhood best friend. All five questions have an average response between 6 and 8. If the rate of answering a 6 or 7 for any question is greater than 30%, then the difference between the variances is significant at a level of 0.1 or lower.

Moreover, our hypothesis posited that significant disparities in variance, indicating the presence of measurement error, should result

Table 2 *Descriptive statistics of the responses to the 23 questions*

	0 to 10					-5 to 5				
	Mean	*SD*	Max	Med	Min	Mean	*SD*	Max	Med	Min
Q1	6.84	2.30	10	7	0	7.87	1.75	10	8	4
Q2	8.56	1.82	10	9	0	9.26	1.09	10	10	6
Q3	3.44	1.57	6	4	0	4.10	2.82	10	3	0
Q4	2.76	2.30	9	2	0	3.52	2.45	8	3	0
Q5	7.78	1.80	10	8	1	8.84	1.79	10	10	2
Q6	1.48	2.02	9	1	0	0.77	1.18	5	0	0
Q7	6.49	1.84	10	7	0	7.26	1.32	10	7	5
Q8	9.22	1.41	10	10	4	9.65	0.66	10	10	8
Q9	7.44	2.24	10	8	0	8.74	1.48	10	9	5
Q10	2.57	1.66	6	3	0	3.06	1.65	6	3	0
Q11	6.45	1.74	10	6	1	7.23	1.63	10	7	3
Q12	2.12	2.87	10	1	0	2.06	2.16	6	1	0
Q13	1.71	2.22	10	1	0	1.48	1.29	4	1	0
Q14	2.35	2.11	10	2	0	2.16	1.97	10	2	0
Q15	6.71	2.29	10	7	2	7.97	1.54	10	8	5
Q16	8.00	1.82	10	8	2	8.71	1.37	10	9	6
Q17	4.95	2.47	10	5	0	5.68	2.23	10	5	2
Q18	8.04	1.98	10	8	0	8.74	1.32	10	9	6
Q19	6.53	2.60	10	8	0	7.03	2.14	10	7	0
Q20	3.00	1.77	10	3	0	3.29	1.47	7	3	0
Q21	1.73	2.72	10	1	0	1.16	2.02	8	0	0
Q22	1.60	2.84	10	0	0	1.87	2.03	6	1	0
Q23	9.16	1.66	10	10	2	9.55	0.96	10	10	6

in the experimental group (-5 to 5 scale) exhibiting higher levels of subjective well-being (SWB). This assumption stems from the idea that a clearer understanding of the scale would allow for more accurate representation of actual feelings, thereby reflecting higher SWB scores.

To serve as a robustness check, we conducted a linear regression with the SWB reported as response variables and the scale as the independent variable of interest. In our analysis, we also controlled for gender and age, as these factors, particularly gender, could significantly influence students' perceptions of everyday life scenarios. Gender differences in perceiving and reporting emotions and experiences are well-documented in psychological literature. By controlling for these variables, we aimed to isolate the effect of the scale change on SWB, ensuring that any observed differences were not merely due to demographic variations.

The results in Table 4, which displays the Ordinary Least Squares (OLS) linear regression results, confirmed our hypothesis. The numbers in the cells present the coefficient with their standard errors in parenthesis.

Table 3 *F-test results of the responses to the 23 questions*

	Mean 1	Mean 2	*F*-value	*p*	6/7 Percentage	*η*-squared
Q7	6.49	7.26	3.99	0.049	44.4%	0.051
Q11	6.45	7.23	3.75	0.057	42.2%	0.049
Q5	7.78	8.84	6.42	0.013	35.6%	0.080
Q9	7.44	8.74	7.97	0.006	35.6%	0.097
Q1	6.84	7.87	4.43	0.039	33.3%	0.056
Q16	8	8.71	3.38	0.070	33.3%	0.044
Q18	8.04	8.74	2.95	0.090	33.3%	0.038
Q15	6.71	7.97	7.1	0.009	31.1%	0.087
Q17	4.95	5.68	1.69	0.198	17.8%	0.022
Q2	8.56	9.26	3.7	0.058	13.3%	0.047
Q19	6.53	7.03	0.78	0.380	13.3%	0.010
Q3	3.44	4.1	1.66	0.201	6.7%	0.022
Q8	9.22	9.65	2.41	0.125	6.7%	0.031
Q21	1.73	1.16	0.96	0.329	6.7%	0.013
Q23	9.16	9.55	1.38	0.244	6.7%	0.019
Q4	2.76	3.52	1.91	0.171	4.4%	0.025
Q10	2.57	3.06	1.63	0.206	4.4%	0.021
Q22	1.6	1.87	0.21	0.646	4.4%	0.003
Q12	2.12	2.06	0.01	0.930	2.2%	0.000
Q6	1.48	0.77	3	0.088	0.0%	0.041
Q13	1.71	1.48	0.26	0.610	0.0%	0.003
Q14	2.35	2.16	0.15	0.700	0.0%	0.002
Q20	3	3.29	0.57	0.454	0.0%	0.008

Table 4 *OLS linear regression results for selected questions-age/gender controlled*

	Q1		Q5		Q7		Q9		Q15	
(Intercept)	2.87	5.82*	12.76*	6.72*	8.76	5.72*	7.54*	6.14*	9.23	5.45*
	(3.68)	(0.73)	(3.12)	(0.62)	(2.84)	(0.57)	(3.48)	(0.69)	(3.47)	(0.70)
Scale-N	1.01*	1.03*	0.87*	1.06*	0.81*	0.77*	1.17*	1.29	1.00*	1.26
	(0.50)	(0.49)	(0.43)	(0.41)	(0.39)	(0.38)	(0.48)	(0.46)	(0.47)	(0.47)
Age	0.14	-	-0.3	-	-0.14	-	-0.07	-	-0.19	-
	(0.18)		(0.15)		(0.14)		(0.17)		(0.17)	
Gender-Male	0.7	-	0.06	-	-0.9*	-	0.57	-	1.00*	-
	(0.53)		(0.45)		(0.41)		(0.50)		(0.50)	

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

From the results of the regressions, we observed that in scenarios where the variance in responses was notably different, implying a measurement error, the experimental group using the -5 to 5 scale reported higher SWB. This finding suggests that the adjusted scale provided a more intuitive and accurate medium for respondents to express their levels of happiness and satisfaction.

Discussion

Our study's key finding is the significant variance in the interpretation of the 7 and 8 values on the 0-10 Likert scale. The *F*-test results suggest a potential misalignment between the respondents' perceived and the intended meanings of these scale values. The decimal system's prevalence in daily life might contribute to this discrepancy, as it seems to influence how respondents interpret numerical values in survey contexts. This finding aligns with Rotondaro et al. (2019), who discussed how numeric value representations are influenced by habitual counting systems. Our study extends this concept to the realm of Likert scale interpretation in SWB measurements.

The reason for this can be explained intuitively. When participants feel either extremely unhappy (e.g., Q12, "Your parents divorced") or extremely happy (e.g., Q23, "Win a $1,000,000 lottery"), either scale will not distract them from selecting the lowest/highest number. However, when the statement is about something within the range of "happy," different people may have different understandings of the scale, which can result in higher variance.

For example, consider Question 5: "Your friend invites you to a birthday party." For most college students, going to a friend's party is either happy or very happy. If someone like Josh is not too fond of parties, he may feel moderately happy about joining the party and select 3 on the -5 to 5 scale without much hesitation. However, if he looks at the 0-10 scale, he may feel that 8 is too high to describe his feeling, as he typically uses 8 out of 10 to describe something at a great level. Therefore, without carefully thinking about the math, Josh goes for 7 on the 0-10 scale. This is how the -5 to 5 scale increases the average response: these are true values that were previously underestimated when using the 0 to 10 scale, which is consistent with the OLS results in Table 5 in the Appendix.

The contrast between the 0-10 and -5 to 5 scales in our study provides valuable insights. The -5 to 5 scale's clearer delineation of positive and negative sentiments could explain the more intuitive responses observed. This finding is particularly interesting as it challenges the conventional preference for the

0-10 scale in SWB research, suggesting that alternative scales like -5 to 5 might yield more accurate data in certain contexts.

Our results have significant implications for future SWB research. The tendency to misinterpret the middle range of the 0-10 scale could lead to data inaccuracies, calling for a reevaluation of this scale's effectiveness in accurately capturing respondents' emotional states. This insight encourages a critical review of scale choices in SWB research and highlights the need for more nuanced approaches to measuring subjective experiences.

However, the study is not without limitations. The sample's small size, lack of diversity and the hypothetical nature of the scenarios could have influenced the findings. This study, while adhering to the minimum recommended sample size of 50 plus 8 times the number of measured variables (50+8m), acknowledges that a more substantial sample size could further bolster the robustness and generalizability of the findings. The reliance on a sample comprising solely psychology students also limits the diversity of our participants and, by extension, the applicability of our results across different populations. Future research should aim to replicate these results with a larger, more varied demographic sample and in real-life contexts. Additionally, investigating the impact of cultural backgrounds on scale interpretation would contribute to a more nuanced understanding of how different populations interact with Likert scales.

Conclusion

This study embarked on an exploratory study of how individuals interpret values on the commonly used 11-point Likert scale, focusing particularly on the potential misunderstanding of the scale's mid-range values, 7 and 8. Our findings reveal a significant variance in responses when comparing the traditional 0-10 scale with the less conventional -5 to 5 scale. This variance suggests that the usual interpretation of the 0-10 scale may not align with the intended meaning of its mid-range values, potentially due to the influence of the decimal counting system ingrained in everyday life.

The implications of this discovery are significant for the field of subjective well-being (SWB) research. They highlight the necessity of critically assessing the choice of scales in survey-based research, as traditional scales like the 0-10 might not always accurately capture respondents' true sentiments, especially around the scale's mid-range. The alternative -5 to 5 scale demonstrated a more intuitive response pattern, suggesting that it could be a more accurate tool for gauging SWB in certain contexts.

In light of our findings, we propose that researchers consider the adoption of the -5 to 5 scale for measuring subjective well-being. This scale's intuitiveness, as evidenced by our study, could significantly enhance the accuracy of the results. Unlike the traditional 0-10 scale, the -5 to 5 scale distinctly demarcates positive and negative sentiments, potentially reducing the ambiguity around mid-range values. This clarity could be especially beneficial in complex or nuanced surveys where precise emotional responses are crucial.

Additionally, another approach to improve scale interpretation could involve the use of annotated scales. By attaching descriptive labels to each numerical value, respondents would have a clearer understanding of what each number represents in terms of emotional intensity or satisfaction levels. For example, alongside each number, a brief descriptor such as 'moderately happy' or 'slightly unhappy' could be provided. However, there are also some practical limitations of this method. In paper-based surveys, adding an-

notations to each scale point could lead to cluttered and overwhelming questionnaire designs. Furthermore, this approach might not be compatible with certain online survey tools that use default slider functions, which typically do not allow for detailed annotations alongside each scale point.

In conclusion, our study contributes to the ongoing discussion about the effectiveness and interpretation of Likert scales in psychological research, especially in the context of subjective well-being. It opens the door for further investigations into alternative scaling methods and underscores the need for a nuanced approach in survey design and interpretation.

Authors' ORCID
Zhou Fang
https://orcid.org/0000-0003-2590-3719
Yue-Mei-Tong Chen
https://orcid.org/0009-0004-9528-0531

References

Abdel-Khalek, A. M. (2004). Research note: Happiness among Kuwaiti college students. *Journal of Happiness Studies*, *5*(1), 93–97. https://doi.org/10.1023/B:JOHS.0000021715.45981.40

Atalay, K., & Barrett, G. (2022). Retirement routes and the well-being of retirees. *Empirical Economics*. https://doi.org/10.1007/s00181-022-02213-9

Bernheim, J. L. (1999). How to get serious answers to the serious question: "How have you been?": Subjective Quality of Life (QOL) as an individual experiential emergent construct. *Bioethics*, *13*(3-4), 272–287. https://doi.org/10.1111/1467-8519.00156

Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *Quality of American life, the perceptions, evaluations, and satisfactions*. Russell Sage Foundation. http://www.jstor.org/stable/10.7758/9781610441032

Deter, M. (2021). Hartz and minds: Happiness effects of reforming an employment agency. *Journal of Happiness Studies*, *22*(4), 1819–1838. https://doi.org/10.1007/s10902-020-00297-7

Gerlich, R., & Wolbring, T. (2021). In good times and in bad, in sickness and in health: A longitudinal analysis on spousal caregiving and life satisfaction. *Journal of Happiness Studies*, *22*(3), 1481.

Livovsky, D. M., Barber, C., Barba, E., Accarino, A., & Azpiroz, F. (2021). Abdominothoracic postural tone influences the sensations induced by meal ingestion. *Nutrients*, *13*(2), Article 658. https://doi.org/10.3390/nu13020658

Macri, E. (2017). Label scale and rating scale in subjective well-being measurement. In G. Brulé & F. Maggino (Eds.), *Metrics of Subjective Well-Being: Limits and Improvements* (pp. 185-200). Springer International Publishing. https://doi.org/10.1007/978-3-319-61810-4_9

Maggino, F., & D'Andrea, S. S. (2003). Different scales for different survey methods: Validation in measuring the quality of university life. In M. J. Sirgy, D. Rahtz, & A. C. Samli (Eds.), *Advances in Quality-of-Life Theory and Research* (pp. 233–256). Springer Netherlands. https://doi.org/10.1007/978-94-017-0387-1_14

Mazaheri, M., & Theuns, P. (2009). Structural Equation Modeling (SEM) for satisfaction and dissatisfaction ratings; multiple group invariance analysis across scales with different response format. *Social Indicators Research*, *91*(2), 203–221. https://doi.org/10.1007/s11205-008-9278-8

Muller, N., Nagels, A., & Kauschke, C. (2022). Metaphorical expressions originating from human senses: Psycholinguistic and affective norms for German metaphors for internal state terms (MIST database). *Behavior Research Methods*, *54*(1), 365–377. https://doi.org/10.3758/s13428-021-01639-w

Rotondaro, F., Ponticorvo, M., Gigliotta, O., Pinto, M., Pellegrino, M., Gazzellini, S., Dolce, P., Miglino, O., & Doricchi, F. (2019). The Number Interval Position Effect (NIPE) in the mental bisection of numerical intervals might reflect the influence of the decimal-number system on the Gaussian representations of numerosities: A combined developmental and computational modeling study. *Cortex*, *114*, 164–175. https://doi.org/10.1016/j.cortex.2018.11.015

Sohier, L., Van Ootegem, L., & Verhofstadt, E. (2021). Well-being during the transition from work to retirement. *Journal of Happiness Studies*, *22*(1).

Tinsley, H. E., & Weiss, D. J. (1975). Interrater reliability and agreement of subjective judgments. *Journal of Counseling Psychology*, *22*(4), 358.

Weermeijer, J., Lafit, G., Kiekens, G., Wampers, M., Eisele, G., Kasanova, Z., Vaessen, T., Kuppens, P., & Myin-Germeys, I. (2022). Applying multiverse analysis to experience sampling data: Investigating whether preprocessing choices affect robustness of conclusions. *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01777-1

Wu, F. (2020). An examination of the effects of consumption expenditures on life satisfaction in Australia. *Journal of Happiness Studies*, *21*(8), 2735–2771. https://doi.org/10.1007/s10902-019-00161-3

Appendix

Table 5 *List of events used in the experiment*

#	Question
1	You meet your friend's pet
2	Receive an A for a major required course
3	Did not correctly answer the professor's questions in class
4	Receive a B for a major required course
5	Go to your best friend's birthday party
6	Your pet dies
7	Have a delicious sandwich
8	You get an offer for your designated career
9	Get your first car
10	You eat an awful burger
11	Correctly answered the professor's questions in class
12	Your parents divorce
13	You lose your wallet
14	Have a verbal conflict with your best friend
15	Meet your childhood best friend
16	Have a great conversation with your best friend
17	You tripped over something (did not get injured)
18	You were admitted into college
19	Get a new pet
20	The classroom temperature is too high
21	You are dismissed from your school
22	You fail your major-required course
23	Win a $1,000,000 lottery