

Evie Project Example Walk Through for HKS

Contents

Overview	1
1) Challenge	1
2) Aim	1
3) Approach	2
4) Method	2
5) Results	3
A more detailed and formal report of the project.	8
Design and Protocol	8
Procedure	10
Data cleaning	10
Data analysis	10
Results	12

Overview

This study utilizes a new and robust method to assess well-being in a real-world setting, with a large sample size (Sample = 821; Prompts = 7357). Its method and insight could be generalized to assess subjective experience of any sort, with careful investment into the questions asked.

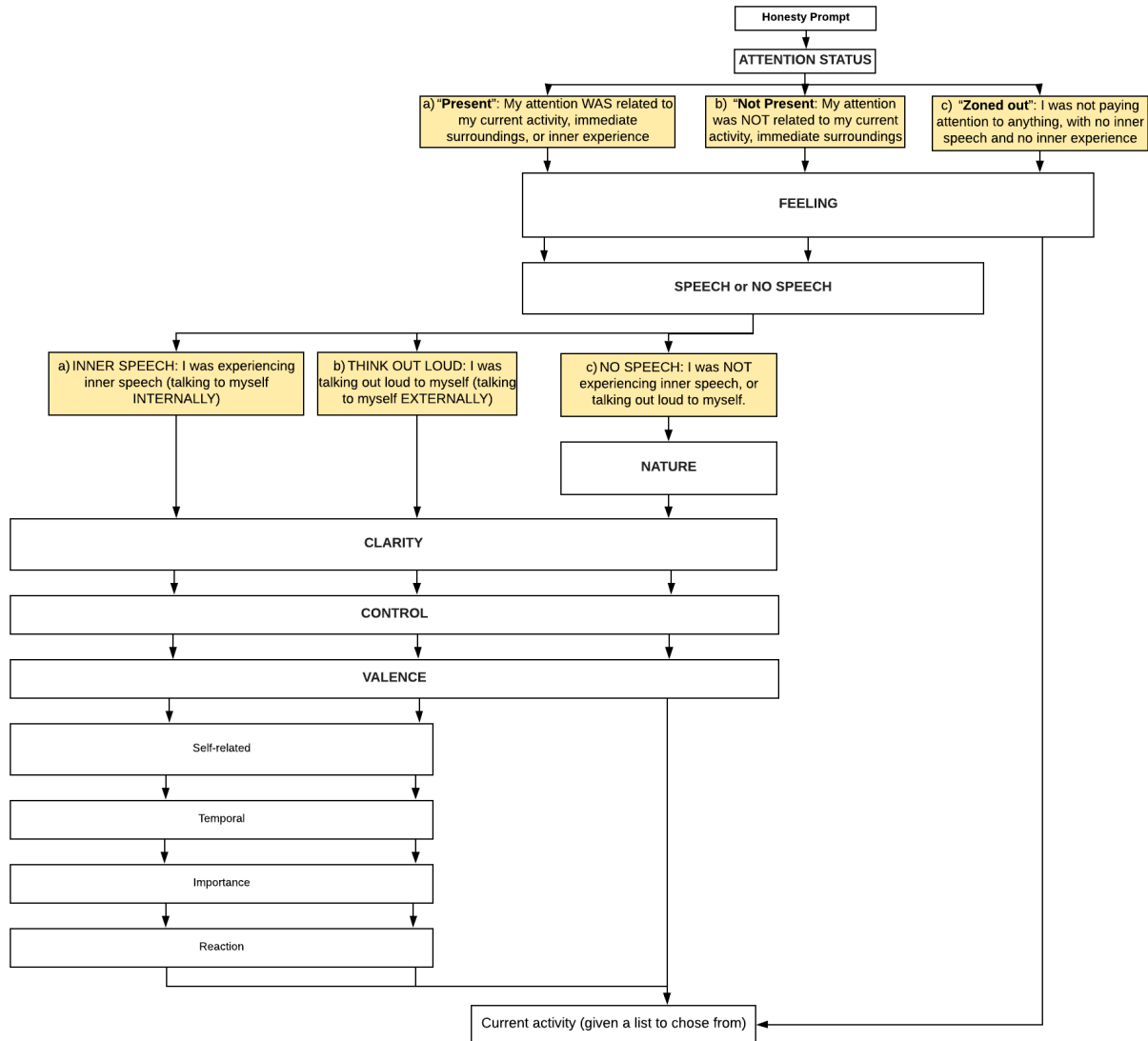
1) Challenge

- Lack of research on micro-level or momentary human experience
- Past result of intervention on momentary well-being might be contaminated by momentary valence-of-thought (e.g. the participants might be happy just because she's thinking about pleasant things that the moment, which is not due to the factor or intervention of interest)
- A mind-wandering mind is an unhappy mind?

2) Aim

- Use experience sampling method to study how momentary well-being is influenced by other subjective factors.
- Include a valence-of-thought question and later take it into control. Do a more robust assessment on the effect of attention state on momentary well-being by taking valence-of-thought into consideration.
- Explore the effect of activity and other subjective experience on momentary well-being.

3) Approach



4) Method

Generalized linear mixed effect modeling

- Attention State (and Controllability, and their interaction) as the fixed effect(s)
- Participant ID and their Current Activity are crossed random effects

Approach of reporting

- Significance test of model comparison (a full model vs. a lesion model without the effect of interest)
- Partial eta-square

Preparation before analysis:

A. Importing R Libraries (Tools)

B. Importing raw datasets

C. Cleaning raw datasets

- For instance, exclude incomplete data, recoding categorical data into numeric data.

D. Excluding suspicious participants from the dataset. The participants were suspicious if they fall into at least one of the following categories.

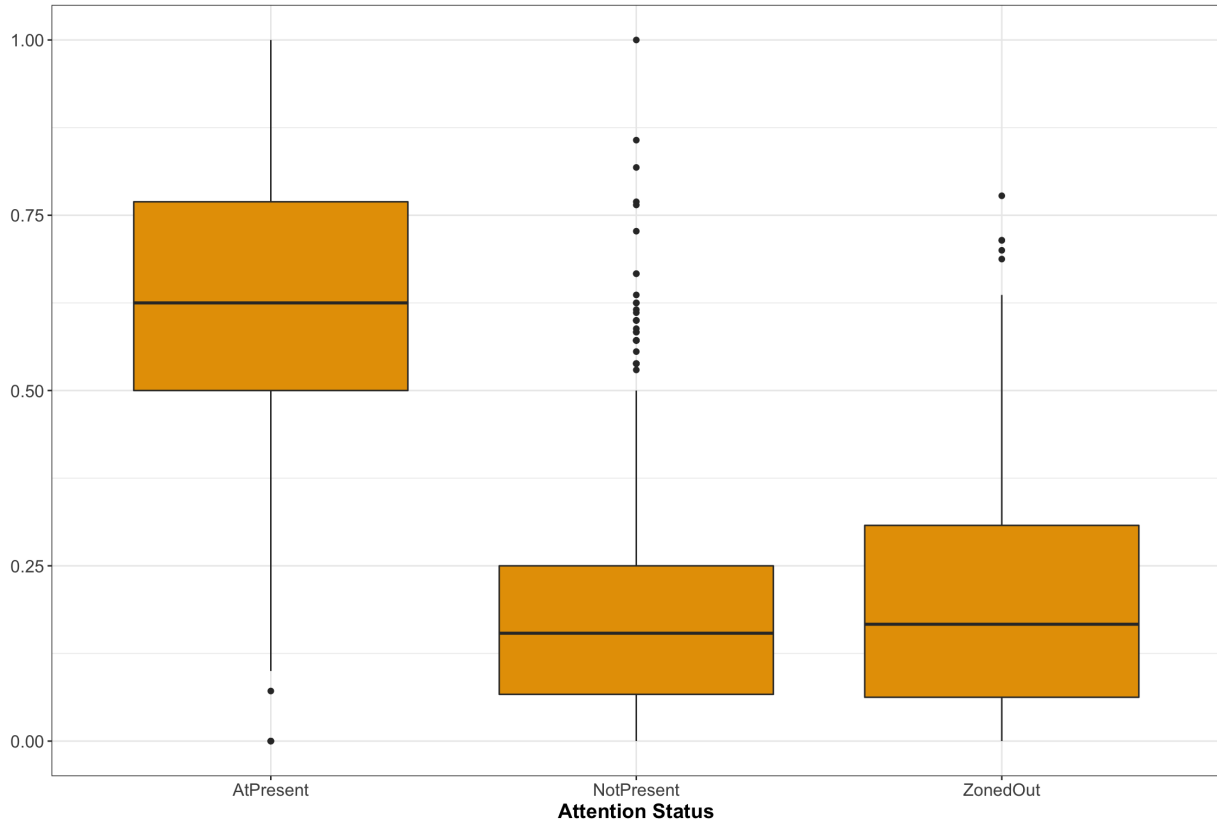
- Did not respond to at least 9 prompts.
- Chose “Zoned-out” for too many times (three SDs above the mean).

E. Combine the ESM data with Pre-ESM survey, which has the demographic information and trait-level happiness.

5) Results

Descriptive findings:

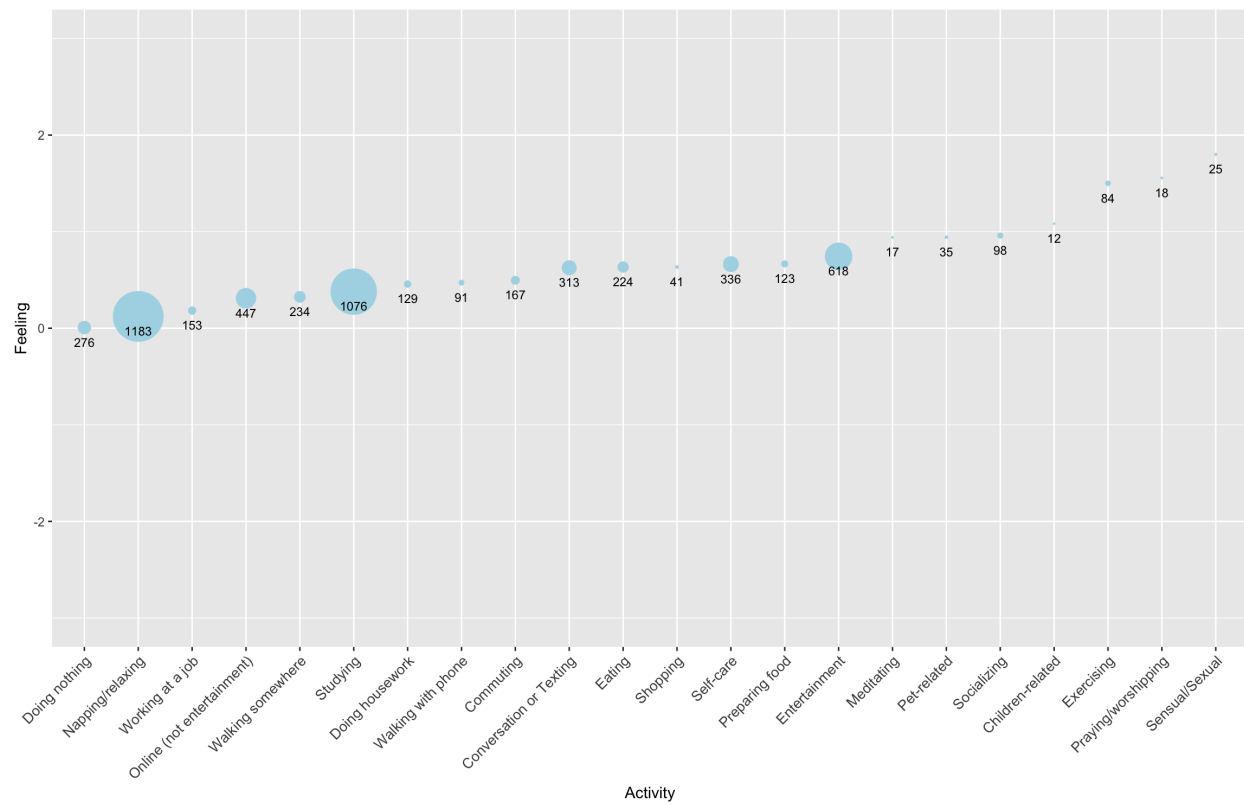
5.1) How often do we pay attention to the here-and-now, not attending to the here-and-now, versus no thoughts at all during waking time? - More than half of our waking time are at-present, phew!



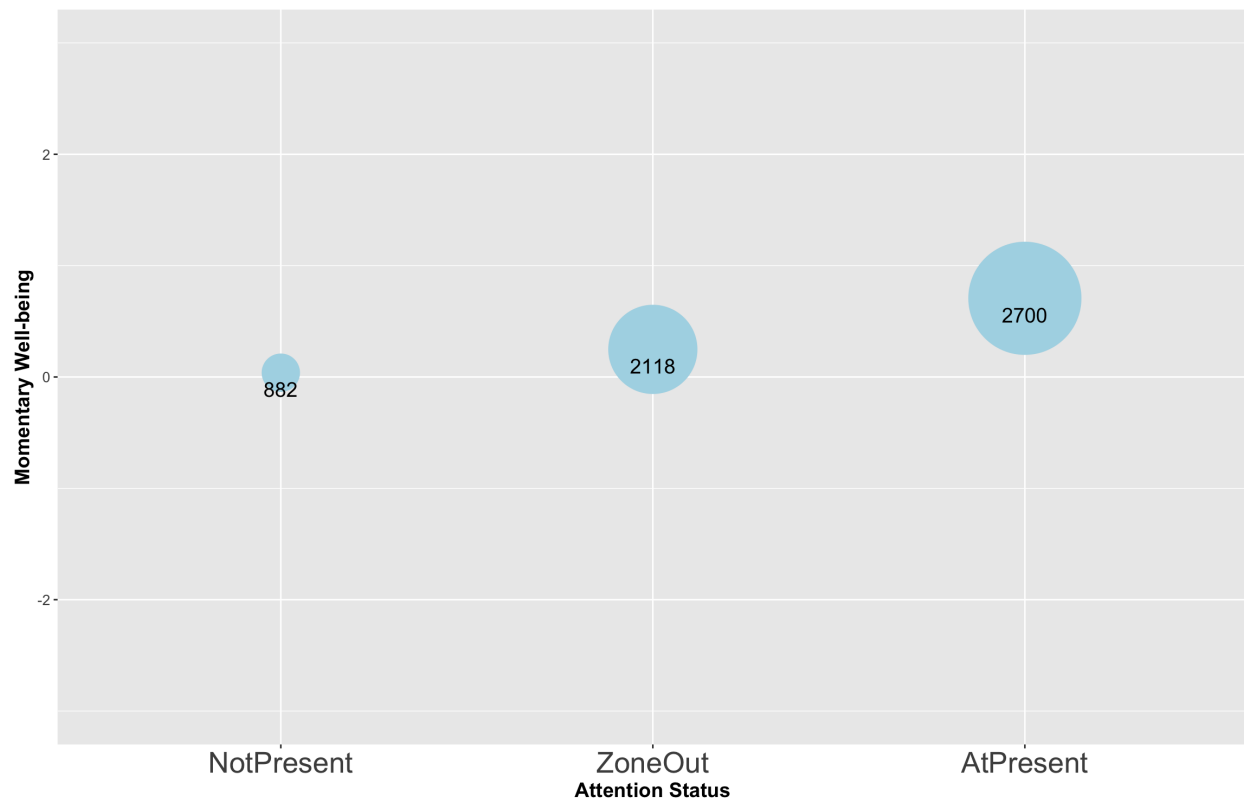
5.2) If we are interested in the modalities of thought, we can further distribute the prompts based on attention state and modality. We can zoom in to focus on one modality (in this case, verbal thoughts) if that would deepen our understanding of something more specific.

5.3) Does attention state matter more to momentary well-being or current activity? - Hmmm, the descriptive result votes activity as a stronger influence than attention states.

5.3.1) Visualize the range of well-being as a function of activity



5.3.2) Visualize the range of well-being as a function of attention state



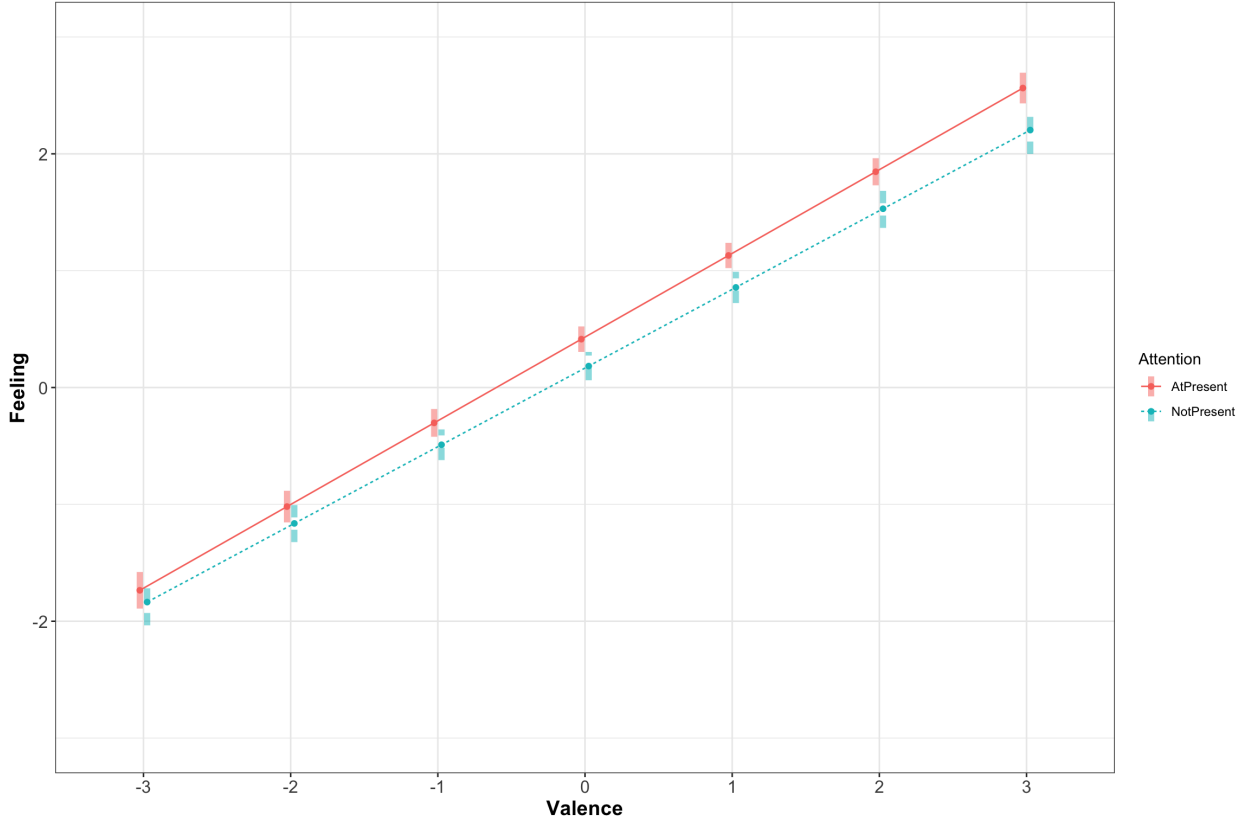
5.3.3) Does attention state contribute to well-being over-and-beyond the pleasantness of thought? - Yes!

Table 1: Table 1. Effect of Attention and Control on Dependent Variables Without Valence As A Covariate

Dependent Variables	Main Effects of Attention Status	Main Effects of Controllability	Interaction between Attention Status and Controllability
Well-being - Model Comparison	Chi ² = 123.014 , p = 0 , partial eta squared = 0.039	Chi ² = 73.937 , p = 0 , partial eta squared = 0.023	Chi ² = 0.004 , p = 0.947 , partial eta squared = 0
Well-being - R ² for LMM	0.033	0.022	0
Clarity - Model Comparison	Chi ² = 61.768 , p = 0 , partial eta square = 0.022	Chi ² = 261.37 , p = 0 , partial eta square = 0.08	Chi ² = 2.578 , p = 0.108 , partial eta square = 0.001
Clarity - R ² for LMM	0.031	0.034	0
Self-relevancy - Model Comparison	Chi ² = 0.197 , p = 0.657 , partial eta squared = 0	Chi ² = 51.598 , p = 0 , partial eta squared = 0.016	Chi ² = 1.451 , p = 0.228 , partial eta squared = 0
Self-relevancy - R ² for LMM	0	0.021	0.001
Importance - Model Comparison	Chi ² = 1.07 , p = 0.301 , partial eta square = 0	Chi ² = 97.459 , p = 0 , partial eta square = 0.03	Chi ² = 4.667 , p = 0.031 partial eta square = 0.001
Importance - R ² for LMM	0	0.026	0.001
Reaction - Model Comparison	Chi ² = 96.519 , p = 0 , partial eta square = 0.031	Chi ² = 84.721 , p = 0 , partial eta square = 0.026	Chi ² = 0.403 , p = 0.525 partial eta square = 0
Reaction - R ² for LMM	0.027	0.026	0

Table 2: Table 2. Effect of Attention and Control on Dependent Variables With Valence As A Covariate

Dependent Variables	Main Effects of Attention Status	Main Effects of Controllability	Interaction between Attention Status and Controllability
Well-being - Model Comparison	Chi ² = 29.694 , p = 0 , partial eta square = 0.009	Chi ² = 5.929 , p = 0.015 , partial eta square = 0.002	Chi ² = 0.674 , p = 0.412 , partial eta square = 0
Well-being - R ² for LMM	0.006	0.001	0
Clarity - Model Comparison	Chi ² = 30.762 , p = 0 , partial eta squared = 0.011	Chi ² = 204.553 , p = 0 , partial eta squared = 0.063	Chi ² = 2.121 , p = 0.145 , partial eta squared = 0.001
Clarity - R ² for LMM	0.009	0.055	0.001
Self-relevancy - Model Comparison	Chi ² = 0.087 , p = 0.768 , partial eta square = 0	Chi ² = 51.118 , p = 0 , partial eta square = 0.016	Chi ² = 1.463 , p = 0.226 , partial eta square = 0
Self-relevancy - R ² for LMM	0	0.021	0.001
Importance - Model Comparison	Chi ² = 0.266 , p = 0.606 , partial eta square = 0	Chi ² = 90.495 , p = 0 , partial eta square = 0.028	Chi ² = 4.587 , p = 0.032 , partial eta square = 0.001
Importance - R ² for LMM	0	0.024	0.001
Reaction - Model Comparison	Chi ² = 13.551 , p = 0 , partial eta square = 0.004	Chi ² = 10.718 , p = 0.001 , partial eta square = 0.003	Chi ² = 1.611 , p = 0.204 , partial eta square = 0
Reaction - R ² for LMM	0.003	0.002	0



A more detailed and formal report of the project.

Design and Protocol

Characterizing Inner Speech The first goal of the current study was to characterize the nature and frequency of inner speech, which we refer to as the “Inner Experience” study. To this end, we used the Experience Sampling Method (ESM) with an app called “ExpiWell” (downloaded on a mobile device), which prompted participants to answer questions about their real-time inner experience. Participants were prompted three times a day, for six days. The prompts were distributed at semi-random times from 9 am to 10 pm, with the first of the three daily prompts in the morning, the second in the afternoon, and the third prompts in the evening. ExpiWell would send a notification 15 minute after each prompt if there was no response from the participant up till that time point. A prompt would expire if the participant did not reply to it within the predesigned one hour lapse. We estimate that answering the questions for each prompt took as long as three minutes, such that if a participant responded to all three prompts, on each of the six days, this would amount to 18 prompts per participant, and a total of 54 minutes. Note that participants received an instructional prompt from ExpiWell on their phone one day before experience sampling start. The instructional prompt asked them to not respond to prompts, under the following conditions: 1) they were driving (for safety reasons), 2) they were in class (because we did not want to disrupt their class time).

Each time the prompt went off, a series of questions was presented, with the specific flow of questions depending on the participants’ answers (see Figure 1 for the flowchart). Participants were shown the flowchart in the instructional prompt one day before the experience sampling started so that they knew what to expect during each prompt. The opening prompt always began with an honesty prime saying “Please be honest about your experience, it’s really important to us. Thanks!”). Then, the first question was about their “Attentional Status”: “PRESENT: My attention WAS related to my current activity, immediate surroundings, or inner experience”, which was internally coded as “At-Present”; 2) “NOT PRESENT: My attention was NOT related to my current activity, immediate surroundings”, which was internally coded as “Not-present”, 3) “ZONED OUT: I was not paying attention to anything, with no inner speech and no inner experience”, which was internally coded as “Zoned-out”. Note that we added this third option, because we reasoned that there may be times that participants feel neither Present nor Not-Present, and we wanted to give them that option. If participants chose (3) in an episode, then they would be asked about their momentary well-being (ranging in a 7-point scale from “Very bad” to “Very good”, internally coded as Feeling), and then they would skip over the majority of the questions asked if they were chosen (1) and (2) in the episode, going straight to a question about their Current Activity, see below). Note that we also used option (3) as a way to screen for participants who may have been overly choosing this option as a way to skip questions and finish quickly, see Data Cleaning, below. Given that the participant chose (1) or (2), they were then asked questions about the phenomenology of their inner experience, in the following order:

After the Feeling question, Participants were then asked about whether their experience involved talking to themselves or not. They were prompted “SPEECH or NO SPEECH: Which best describes what you were experiencing?” and were provided with the three options “INNER SPEECH: I was experiencing inner speech (talking to myself INTERNALLY)”; “THINK OUT LOUD: I was talking out loud to myself (talking to myself EXTERNALLY)”; “NO SPEECH: I was NOT experiencing inner speech, or talking out loud to myself”. If they chose “NO SPEECH” or they were not talking to themselves, they will be asked about the representative format of the experience (although we referred to this as “nature” for the participant), with choices that included: “BODY: I was experiencing body sensations, e.g. hunger”; “EMOTION: I was experiencing emotions, e.g. sadness”; “ENVIRONMENT: I was noticing the environment, e.g. looking at the trees”; “MUSIC/SOUNDS: I was listening to music/podcast, e.g. with my earphones in, or imagining music/sounds”; “VISUAL IMAGERY: I was experiencing visual imagery, e.g. imagining my dog”; “ANOTHER PERSON: I was experiencing another person, e.g. holding hands”; “Other (click and specify)”

The self-talk (or the nature question for No Speech prompts) question was followed by questions about the phenomenology of their inner speech or inner experience. Note that the exact wording depends on the answer to the self-talk question. If participants indicated they were talking to themselves (either inner speech or think-out-loud), the wording of questions are as follows:

- “CLARITY: What was the clarity of your inner/external speech?”, ranging on a 7-point scale with the two extremes being “Not at all clear” and “Extremely clear”, and the middle being “Moderately clear”
- “CONTROL: Do you feel you were in control of the inner/external speech?”, on a binary scale of “YES: I CHOSE to engage in inner/external speech” and “NO: The inner/external speech I had was SPONTANEOUS”
- “VALENCE: What was the valence of your inner/external speech? If a stranger saw the content of your inner/external speech (i.e., just the actual words), how would they rate it?”, ranging on a 7-point scale with the two extremes being “Very negative” and “Very positive”, and the middle being “Neutral”
- “SELF-RELATED: Was your inner/external speech related to yourself in any way?” on a binary scale, and the options were “YES: For example, analyzing myself, thinking/talking about getting something done, or thinking/talking about myself in relation to other people” and “NO: For example, thinking/talking about other people I know or don’t know, or thinking about something unrelated to others (e.g. wondering if it will rain)”
- “PAST/FUTURE: What was the temporal distance of the thought?” ranging in a 9-point scale, shown in Table 1. This question on the Temporal Distance of the thought was planned to be a dependent measure, but later Temporal Distance response was combined with Attention Status responses to serve as a data-cleaning rule, see data-cleaning for detail. Because the data-cleaning pertained to the distribution of Attention Status on Temporal Distance, the effect of Attention Status on Temporal Distance will not be analyzed and therefore, Temporal Distance was not a dependent variable. Nonetheless, the histogram of Temporal Distance of Not-Present prompts were presented in Figure 2.
- “IMPORTANCE: How important was the inner/external speech?”, ranging on a 7-point scale with the two extremes being “Not at all important” and “Extremely important”, and the middle being “Moderately important”.

If No Speech was chosen in an episode, only Clarity, Control, and Valence would be asked. And the wording of these question was altered to tailor to the No Speech experience, by changing “inner/external speech” to “inner experience”. Moreover, No Speech episodes did not get to answer Self-related, Temporal Distance and Importance questions since we believe these three features are pertain to self-talk but not other representative format of thought.

Theoretically, the Feeling question enables us to concetually replicate the original results of Killingsworth and Gilbert (2010). As described in the Introduction, this previous study showed that momentary well-being is better when people are Present (in their study defined as “not thinking about something other than what they are currently doing” vs. Not-Present (in their study defined as “thinking about something other than what they are currently doing”), at least when the Not-Present thoughts are either “neutral” or “negative”. The data of the current study allowed us to investigate the same question, noting that there are six methodological differences between the current study and that of Killingsworth and Gilbert (2010), as follows: 1) The current study provided a third option for Attention Status, i.e., “Zoned-Out”, 2) The current study (but not the Killingsworth & Gilbert study) asked participants if they experienced inner speech, 3) The current study used a 7-point scale for rating well-being , whereas the Killingsworth and Gilbert study used a sliding scale from 0 – 100, 4) The current study had participants answer many more questions than the Killingsworth and Gilbert study, which might have contaminated their penultimate response about well-being. 5) The Current Activities used were slightly different between the two studies, and 6) The current study collected data from a relatively small (801 participants) group of college students, whereas Killingsworth & Gilbert collected data from a very large and diverse sample (2250 participants, recruited from the lab website, 18 – 88 years, from different countries). All of these differences, and their potential effects on the results, are addressed in the Discussion.

The very last question asked at the end of all prompts was about the participants’ Current Activity, for which they were given a list of 21 possible activities (see Table 2). Our list overlapped with, but was not identical to, those employed in Killingsworth and Gilbert (2010), as well as Kahneman, Krueger, Schkade, Schwarz, and Stone (2004). These current activity responses were used for three purposes. First, we asked whether the distribution of activities differed across different attention statuses. Second, we directly asked

whether well-being differs across different activities (mirroring Killingsworth & Gilbert, 2010). Third, given the possibility that the distribution of activities differs across attentional states, and that current activity affects some of our dependent measures (e.g., well-being, Clarity, Controllability, etc.), we included current activity as a random effect in our statistical models when comparing effects across different attention statuses. Note that the current study is only testing the link between Current Activity and well-being, although a future analysis might ask whether other dependent variables, e.g., Clarity, Controllability, differs across current activity.

For the current study, we restricted our analyses to prompts where participants reported experiencing Inner Speech, which accounted for 40.4% of all prompts (reported in more detail in the Results). With these data, the current study asked whether inner speech differs for moments when people are Present vs. Not-Present and for inner speech that is Controlled vs. Spontaneous, with respect to: Clarity, Importance, and Self-related thought. Valence was used as a predictor variable and/or covariate (see Data Analysis, below).

Procedure

Data cleaning

We removed participants who did not appear to be adequately engaged in the protocol, which was determined in two ways. First, participants were removed if they chose “Zoned-Out” too frequently. As explained above, participants knew that choosing “Zoned-Out” led to them skipping over the majority of the questions regarding their inner experience, and therefore may have been choosing it as a way to finish quickly. If a participant chose Zoned-Out too frequently, we considered this suspicious. To this end, for each participant, we calculated the proportion of prompts for which they chose the “Zoned-Out” option. If the Zoned-Out proportion for a participant was greater than three standard deviations above the mean (across participants), this participant was removed from the analysis ($N = 12$). Second, participants were removed if they provided insufficient numbers of prompts; participants who completed less than nine prompts were removed from the analysis ($N = 156$). Further, we also removed experience sampling episodes that are not internally consistent. Specifically, we removed experience sampling episodes if “At-present” was chosen to be the Attention Status of the episode, but “At the present moment” was not chosen as the Temporal Distance of the episode. After the extensive data-cleaning, we found that participants responded to 13.2546816 prompts on average, which corresponds to a rate of 73.6371203%. In total, the ESM data consisted of 10617 prompts from 801 participants. The median lapse between the time the App sent those prompts and the time participants responded was 360. The median time for participants to finish a prompt 60.

Data analysis

We used R [Version 3.6.2; @R-base] for all our analyses.

Question 1: What is the distribution of different attentional states (Present, Not-Present, Zoned-Out), with and without inner speech? Using data from all 10617 prompts (across 801 participants), we report descriptive statistics of the proportions for 5 different prompt outcomes (which add up to 100%): 1) Present (with Inner Speech), 2) Not-Present (with Inner Speech), 3) Present (without Inner Speech), 4) Not-Present (without Inner Speech), and 5) Zoned-Out. Note that for (1) and (2), we included any prompt for which the participant reported experiencing Inner speech, even if they had also selected one of the Non-Inner speech representative formats. For (3) and (4), we included any prompt for which the participant did not report experiencing Inner speech, and we collapsed across prompts from different representative formats. Only data from (1), (2) and (5) were analyzed further in the remaining study questions.

Question 2: Does momentary (state) well-being differ when people are Present vs. Not-Present? This question served as a conceptual replication of the Killingsworth & Gilbert (2010) study, which reported that people are happier when they are Present vs. Not-Present. Like Killingsworth & Gilbert (2010), we provide visual depictions of mean well-being (averaged across prompts) for the different Current

Activities, collapsing across Attention Status (with proportion of prompts represented by the size of the data point). Likewise, we present mean well-being across the different Attention Statuses, collapsing across the different Current Activities (with proportion of prompts represented by the size of the data point).

We then used statistical models to analyze the data as follows. First, we were curious as to whether the distribution of Current Activities differed across the three main Attentional Statuses (At-Present, Not-Present and Zoned Out). To this end, we conducted a simple Chi-square (3 Attention Statuses x 22 Activities). (If the distributions do differ, and if current Activity affects well-being, it is of course imperative to include Current Activity in any model that looks at the effects of Attention status on well-being). Second, we used a linear mixed effect model to determine whether well-being differs across the three main attentional states (At-Present, Not-Present and Zoned-Out). Here, we used prompt as the unit of analysis, with the dependent measure being well-being (score from -3 to 3), the predictor variable being attention status (At-Present, Not-Present and Zoned-Out) and the random effects being Participant and Current Activity. Although current Activity was a random effect, the model will nonetheless determine its unique contribution (in terms of variance accounted for) to well-being. Third, we then investigated the joint effects of Attention Status and Valence (since Valence was an important variable in Killingsworth & Gilbert, 2010), with the same linear mixed model, but excluding Zoned-Out prompts, and including Valence as another predictor variable (ranging continuously from -3 to 3). We also investigated the interaction between Attention Status and Valence, to see if Valence might have a different effect on well-being when one is At-Present vs. Not-Present.

Question 3: How does the nature of inner speech differ across different attention statuses, and level of control? Here, we asked whether Attention Status (At-Present vs. Not-Present) and Controllability (Controlled vs. Spontaneous), or a combination of the two, are related to the phenomenological nature of Inner Speech. We were driven to include Controllability as a predictor variable because we hypothesized that this might affect the inner experience as much as, if not more, than whether one is At-Present vs. Not-Present. Testing the effects of Controllability also allows us to join the discussion of the definition of mind-wandering, and whether it is deleterious, which is under debate in the research field. In line with Seli et al., it may be that in the current study, the combination of Not-Present and Spontaneous constitutes Mind-wandering, which may be its own unique category. If so, this combination may stand out in a statistical analysis as an interaction between Attention Status and Controllability.

The following five phenomena were investigated: Clarity, Valence, Feeling, Importance, and Self-relevancy. Each of these were included as the dependent measure in a linear mixed effect model, where prompt was the unit of analysis, the random effects were Participant and Current Activity, and the two predictor variables were Attention Status (At-Present, vs. Not-Present, data collapsed across Valence) and Controllability (Controlled vs. Spontaneous). An additional interaction term (Attention Status x Controllability) was included in the model to address the possibility that the two effects interact. Finally, in a separate analysis, we included Valence as a covariate for Clarity, Feeling, Importance, and Self-relevancy. Since we learned from results of Question 2 that Valence had a strong relationship with well-being, and we reasoned that it might also have an effect on some of our other dependent measures, which should be controlled in our model. In Table 4 of the results, we present statistical results with Valence included vs. excluded as a covariate, just to show that our findings are robust either way. When there was either a difference between models that included vs. excluded Valence or when we found an effect of the covariate Valence, we emphasize that in the Results. Otherwise, we present just the results when Valence was not included as a covariate.

There are a few things to note about the analyses associated with Question 3. First, note that, with the exception of Self-Relevancy (which is binary and therefore used a logistic regression), all the dependent measures are continuous. Second, momentary well-being is one of the dependent measures we investigate in Question 3, even though it is also investigated in Question 2. The difference is that Question 2 asks about well-being as a function of Attention status and Valence (for a conceptual replication of Killingsworth & Gilbert), whereas Question 3 asks about well-being as a function of Attention Status and Controllability (and includes Valence as a “covariate”, since we found it to have effects on well-being in Question 2). Third, because both Attention Status and Controllability were included in the same model, the model accounts for any relatedness between the two, allowing us to determine the unique contribution of each (if there is one) to the dependent measure. In fact, one of our research questions was whether there is any relatedness between Attention Status and Controllability (i.e., is there a greater chance to be in control of inner speech when one

is At-Present vs. Not-Present)? To investigate this interrelatedness, we conducted a generalized linear mixed model, similar to those described above, with the difference being that the binary variable Controllability became the dependent variable and Attention Status was the predictor variable. (Note that we chose not to conduct a simple Chi-square to address this question, as it cannot account random effects of Participant and Current Activity). Forth, we acknowledge we are conducting models on multiple dependent measures, which of course, leads to an increased chance of Type I errors. For this reason, we also reported the effect sizes through partial eta-square and the R-square for linear mixed-models (Nakagawa & Schielzeth, 2013), and not just p values. Even then, the current study should be considered exploratory.

Results

Q1: Distribution of different attention statuses The percentage of the 10617 prompts is broken down in Figure 3, for five categories: 1) At-Present (with Inner Speech), 2) Not-Present (with Inner Speech), 3) At-Present (without Inner Speech), 4) Not-Present (without Inner Speech), 5) Zoned-Out. By collapsing Categories (1) and (3), and collapse Categories (2) and (4), we find that At-Present, Not-present, and Zoned-Out prompts accounts for 62.4074519% ($SD = 20.4522471\%$), 17.9658077% ($SD = 15.4836622\%$), and 19.6267404% ($SD = 16.7814146\%$) of the prompts, respectively. These percentages of At-Present vs. Not-Present are within the range reported in previous (within and outside the laboratory) studies of college students (range of Not-Present percentage = 15%-50%, McVay, Kane & Kwapil, 2009, Song & Wang, 2012). The current study is the first to report that college students report feeling “Zoned-Out” a large proportion of the time, although we acknowledge that this percentage may be artificially inflated in the Discussion. Only data from categories (1), (2) and (5) were analyzed further in the remaining study questions.

Q2: Does momentary (state) well-being differ when people are At-Present vs. Not-Present? As a conceptual replication of Killingsworth and Gilbert (2010), in Figure 4 and 5, we present visual depictions of mean momentary well-being (averaged across prompts) for the different Current Activities (collapsing across Attention Status) and for the different Attention Statuses (collapsing across the different Current Activities), with proportion of prompts represented by the size of the data point. Note that we do perform statistics on these descriptive data, yet come back to a statistical model of well-being later in the Results. In Figure 4, it can be seen that “meditation” and “exercising” are Current Activities that are associated with more positive feeling, which in line with previous studies showing that meditation and exercising elevates well-being (e.g., Archer, Josefsson & Lindwall, 2014; Brown & Ryan, 2003; Menezes & Bizarro, 2015; Wolkin, 2015). The other activities associated with better feeling include “Sensual/sexual” and “worshiping/praying” which were not surprising. By contrast, “Doing Nothing”, which we interpret as not doing anything meaningful, was associated low momentary well-being. It has been documented that bored participants during dry and meaningless lab activities would rather self-administer shocks than continue the activities. It is unsurprising that participants feel that doing nothing is a miserable way to spend time. “Online activities” was associated with low average momentary well-being, which is in line with previous studies showing that high amounts of time using social media is associated with anxiety and depressive symptoms (e.g., Rosen et al., 2013; Evren et al., 2018).

In Figure 4, it can be seen that the two Attention Statuses associated with highest momentary well-being were Zoned-Out and At-Present, with lower values for Not-Present. Our results provide a conceptual replication of Killingsworth & Gilbert (2010), in showing that people (college students) are happier when they are At-Present vs. Not-Present. Unlike Killingsworth & Gilbert (2010), who suggested that there is more variation in well-being across different Attention Statuses than across different Current Activities, this was not obvious in the current study. This difference between studies could be due to methodological differences between the two (see Methods).

In our next analysis, the results of a Chi-square showed that the distribution of Current Activities differed across the three main Attentional Statuses (At-Present, Not-Present and Zoned Out) ($\chi^2(42) = 1469.4220741$, $p < .001$). Post-hoc analysis showed that this difference was driven by differences across each of the three pairings: At-Present vs. Not-Present ($\chi^2 = 338.5606239$, $p < .001$), At-Present vs. Zoned-Out ($\chi^2 = 1294.378046$, $p < .001$) and Not-Present vs. Zoned-Out ($\chi^2 = 245.8224621$, $p < .001$). These differences are driven mainly by a few activities. For example, napping/relaxing are particularly high frequency activities when participants

report being Zoned-Out, yet particularly low when being At-Present. And, working/studying are particularly high frequency activities when participants report being At-Present, yet particularly low when participants are Zoned-Out and Not-Present.

Because the distributions of Current Activity were found to differ across Attention Status, and given that Current Activity affects well-being (as can be seen in Figure 3), it is of course imperative to include Current Activity in any model that looks at the effects of Attention status on well-being, which we turn to next. The results of our linear mixed effect model (see Methods) showed that Attention Status is a significant predictor of well-being ($\chi^2 = 125.2579216$, $p < .001$). However, when Zoned-Out was removed from the model, which more closely mimics the statistical analysis of Killingsworth & Gilbert (2010) that compared only At-Present vs. Not-Present, well-being was found to be significantly higher for At-Present vs. Not-Present prompts ($z = 10.4716624$, $p < .001$). Interestingly, Feeling of Not-Present was found to be lower than Feeling of Zoned-Out ($z = -3.7356512$, $p < .001$). As might be expected, Zoned-Out was associated with lower feeling than At-Present ($z = -7.6461094$, $p < .001$).

In the next model, we once again excluded Zoned-Out, and added another predictor variable, Valence (participants were asked to rate, from “very negative” to “very positive”, the valence of their inner speech content, from an objective perspective) as well as an interaction term between Attention Status and Valence. The model-comparison (Valence on Feeling w/ Attention vs. w/o Attention in as a predictor) shows that including Attention Status as a predictor significantly improves model fits (main effect of Attention Status: $\chi^2 = 29.693679$, $p < .001$). Further, revealed significantly higher well-being for At-Present vs. Not-Present prompts ($z = 5.6313296$, $p < .001$). There was no interaction between the Attention Status and Valence ($\chi^2 = 1.9039182$, $p = 0.1676404$), indicating that the effect of Valence on momentary well-being does not differ between At-Present and Not-Present prompts. To visualize these effects, in Figure 5 we present regression lines of well-being vs. Valence, separately for At-Present vs. Not-Present prompts, where it can be seen that the line for At-Present prompts is elevated that for Not-Present prompts. In sum, the results from these analyses suggest that participants are happier when they are At-Present and when Inner Speech is more positive and that these two effects can be considered be additive with each other on Well-being.

Question 3: Nature of inner speech as a function of attention status and controllability Before investigating how Attention Status and Controllability are related to the phenomenological nature of Inner Speech, the interrelatedness between Attention Status and Controllability was examined. A generalized linear mixed effect model with Control as the reactor factor and Attention Status as the predictor factor (see Methods) showed that there was a significantly greater chance that inner speech was controlled (vs. spontaneous) when participants were At-Present vs. Not-Present ($\chi^2 = 87.0184878$, $p < .001$). Because Attention Status and Controllability are interrelated, and because Attention Status is likely to affect at least some of our dependent measures (as was seen in Question 2 showing effects of Attention Status on Feeling), it is of course imperative to include both Attention Status and Controllability in the same model to determine the unique contribution of each (if there is one) to our dependent measures, which we turn to next.

Linear mixed effect models were performed for 5 different dependent measures: Clarity, Feeling, Valence, Self-relevancy, and Importance. For each dependent measure, we investigated the effects of Attention Status, Controllability, and their interaction. Analyses were done both when Valence was included vs. excluded as a covariate, for exploratory reasons. Results of when Valence as not included as the covariate are presented in table format in Table 4. Obviously, the analysis when Valence is a covariate is not possible when Valence is also a dependent measure. The results of the rest of the dependent measures of when Valence was included as a covariate were presented in Table 5. To prevent redundancy, we only present in text the result of when Valence was excluded as the covariate, unless its inclusion changed the results substantially (see Methods).

We found the significance levels of p -values of model-comparisons for Attention Status, Control and their interaction in Table 4 and Table 5 are the same, although there was a general reduction for these three factors’ Chi-squares, partial eta-squared, and R-squared. For Clarity, Attention Status and Control are both significant predictors, but not their interaction term. Specifically, compared with not-present and out-of-control, being at-present and in-control is associated with higher Clarity (At-Present vs. Not-Present: $z = 4.899$, $p < .001$; In-Control vs. Spontaneous: $z = 15.155$, $p < .001$). Similarly, for Valence and Feeling, Attention Status and Control are both significant predictors. But the interaction between Attention Status

and Control was not a significant predictor for Valence or Feeling. Specifically, at-present was associated with greater Valence ($z = 8.54, p < .001$) and Feeling ($z = 9.508, p < .001$). In-control was also associated with greater Valence ($z = 9.849, p < .001$) and Feeling ($z = 8.23, p < .001$) than Spontaneous. For Self-relevancy of the inner-speech content, neither Attention Status nor the interaction term between Attention Status and Control was a significant predictor. However, the main effect of Control was significant: In-control is associated with significantly higher Self-relevancy than Spontaneous (In-control vs. Spontaneous: $z = 6.44, p < .001$). For Importance of Inner-Speech content, the main effect of Attention Status was not significant. But there was a significant main effect of Control and the interaction between Attention Status and Control on Importance. Specifically: In-Control was associated with higher than Spontaneous across Attention Statuses, but the difference of Importance was greater in At-Present than the Importance difference in Not-Present.