



OPEN

On the relationship between mind wandering and mindfulness

Angelo Belardi¹✉, Leila Chaieb², Alodie Rey-Mermet¹, Florian Mormann², Nicolas Rothen¹, Juergen Fell² & Thomas P. Reber^{1,2}

Mind wandering (MW) and mindfulness have both been reported to be vital moderators of psychological wellbeing. Here, we aim to examine how closely associated these phenomena are and evaluate the psychometrics of measures often used to quantify them. We investigated two samples, one consisting of German-speaking unpaid participants (GUP, $n = 313$) and one of English-speaking paid participants (EPP, $n = 228$) recruited through MTurk.com. In an online experiment, we collected data using the Mindful Attention Awareness Scale (MAAS) and the sustained attention to response task (SART) during which self-reports of MW and meta-awareness of MW were recorded using experience sampling (ES) probes. Internal consistency of the MAAS was high (Cronbachs α of 0.96 in EPP and 0.88 in GUP). Split-half reliability for SART measures and self-reported MW was overall good with the exception of SART measures focusing on Nogo trials, and those restricted to SART trials preceding ES in a 10 s time window. We found a moderate negative association between trait mindfulness and MW as measured with ES probes in GUP, but not in EPP. Our results suggest that MW and mindfulness are on opposite sides of a spectrum of how attention is focused on the present moment and the task at hand.

Waking experience can be described as a stream of thoughts, perceptions, and emotions that come in and out of the focus of our conscious awareness. Mind wandering (MW) refers to our thoughts becoming decoupled from an ongoing task and coupled to thoughts and feelings not being subject to the task at hand or our surroundings¹. In comparison, mindfulness refers to the mental act of intentionally resting the focus of awareness on a particular subject of experience in the present moment without judgment². These constructs appear to emphasise aspects which lie on opposite sides of a spectrum of how intentional, focused, and self-aware one is regarding the thoughts and perceptions that make up one's conscious experience³.

In light of these conceptual considerations, it seems surprising that statistical associations between measures of MW and mindfulness are rather low^{3–5}. One possible explanation for this may be the low reliability of the psychometric tools used to measure these constructs. Another possibility could be that meta-awareness of MW^{6,7}, i.e., awareness of the fact that one's contents of consciousness is decoupling from an ongoing task, moderates the relationship between mindfulness and MW. To investigate these questions, we first assessed the psychometrics of a well-established mindfulness questionnaire and self-report measures of MW and meta-awareness thereof in a large sample from an online study. Then we estimated the associations between measures of MW, mindfulness, and meta-awareness.

Evidence for the importance of both MW and mindfulness for psychological wellbeing has been reported numerous times in the literature. Increased propensity to MW was associated with reduced affect⁸ and in its extreme form MW can result in persistent negative and repetitive thoughts leading to rumination. Such rumination is at the heart of neurocognitive models of depression^{9–11}. Furthermore, distraction due to MW can potentially cause physical harm e.g. when driving¹², operating heavy machinery¹³, or when working as a medical professional¹⁴. Excessive MW may also interfere with career goals by affecting work and educational performance¹⁵. While a majority of studies focus on negative consequences, MW may also facilitate future planning, goal setting, and aid creative problem solving^{16–18}. For example, Medea and colleagues¹⁸ found that self-generated cognition during an episode of MW may allow the development of more concrete personal goals.

In contrast, mindfulness has been associated predominantly with an increased feeling of wellbeing. The concept of mindfulness has its origins in eastern philosophy and is closely linked to processes of awareness and attention. Mindfulness describes a state in which a person willingly chooses the focus of conscious experience and takes constant notice of their contents of consciousness. Practicing to achieve this mindful state has been a central tenet of traditional Buddhist meditation, and has been introduced in western cultures as a secular form

¹Faculty of Psychology, UniDistance Suisse, 3900 Brig, Switzerland. ²Department of Epileptology, University of Bonn, 53127 Bonn, Germany. ✉email: angelo.belardi@fernuni.ch

of mental practice and flavours in psychotherapy, such as e.g., the mindfulness-based stress reduction (MBSR) program or acceptance and commitment therapy^{2,19,20}.

A widely used task to experimentally elicit MW is the sustained attention to response task (SART)²¹. Participants view, for example, a stream of numbers from 0 to 9 appearing in a random sequence and at a constant rate. The participants' task is to press a button in response to all non-target digits (Go trials) except for one – the target, where they are required to withhold their button press (i.e., the Nogo trial, such as the number 7). Several dependent variables have been used in the SART, such as the performance of the task (i.e., the error rate on Go trials and Nogo trials), the mean reaction time (RT) in Go trials and the variance of these RTs, as well as scores combining performance and RT (e.g. a skills index, calculated as accuracy/RT)²². Variants of the task include querying participants intermittently in defined intervals as to whether their mind was 'on task' or 'off task' using experience sampling (ES) probes to measure MW. Furthermore, meta-awareness of MW is queried after ES of MW in some studies immediately after 'off task' responses^{23,24}. In addition to the self-reports from ES, low performance^{25–27} as well as long and widely dispersed RTs¹⁶ in the SART are considered evidence for low sustained attention and potentially for MW. Several versions of the paradigm combining ES probes and SART have been used in previous research. For example, some studies restricted performance and RT analyses to short time windows immediately preceding appearance of the ES probe^{16,27}. Other studies varied SART difficulty by either adding auditory noise²⁸, by making the number stream predictable²⁹, or by increasing the inter-stimulus interval (ISI)³⁰. Taken together, there is a variety of ways in which the SART is used to elicit and assess MW.

Tools to measure mindfulness, on the other hand, consist predominantly of self-report questionnaires. One of the most commonly used questionnaires is the Mindful Attention Awareness Scale (MAAS)³¹. Previous assessments of the MAAS found that it has a single factor structure and overall robust reliability (Cronbach's α between 0.8 and 0.87)³¹. External validity was evaluated with numerous questionnaires assessing a variety of related constructs such as everyday attention, personality traits and anxiety^{31,32}. Because of the high importance of this questionnaire in mindfulness research, we explored the possibility of shorter versions of the MAAS, based on only 5 and 3 items, which would be quicker to implement in future research.

Despite the close conceptual relationship between MW and mindfulness, estimates of the strength of their association have been surprisingly low^{3–5}. Furthermore, none of these previous studies reported an estimate of reliability for ES of MW, making the interpretation of this association difficult. See Table 1 for a detailed summary of previous findings. Together, there is only weak evidence to suggest that a direct measure of MW such as ES during the SART correlates with MAAS scores. Moreover, when these associations were reported, they were moderate at best.

Additional evidence for the relationship comes from a related line of research that investigates whether mindfulness training impacts direct and indirect measures of MW (for a review, see³³). Such intervention studies found that the practice of mindfulness usually improved SART performance^{3,34–36,37,38} and reduced the frequency of self-reported MW in some cases^{36,39} but not in others^{34,38}. Moreover, one study reported higher MAAS scores after mindfulness training³⁵. Similarly to the findings of those correlation studies reported before, in these mindfulness training studies the associations between the direct MW measure and mindfulness is not as strong as one might expect.

One possible explanation of low associations between ES of MW and MAAS scores could be that querying participants for whether they were on or off task alone conflates over two forms of MW that are oppositely linked to mindfulness, namely MW with and without meta-awareness⁶. This hypothesis has been put forward by Smallwood and Schooler⁷, and initial empirical evidence for the importance of considering meta-awareness was gathered by the same authors in an ensuing study⁹. Here, 'zone outs' (MW without awareness) were linked to higher inhibition errors in an ongoing task while 'tune outs' (MW with awareness) were not. How these 'zone out' and 'tune out' propensities are linked to trait mindfulness, however, seems unclear in the previous literature. Deng et al.⁴ found no significant relationship between either the 'zone out' or the 'tune out' rate with trait mindfulness as measured by the MAAS. A more recent study⁵ found both rates to be negatively associated with MAAS scores. Together there is inconsistent evidence on the role of meta-awareness as potential mediator between MW and trait mindfulness. Another possible explanation for low correlations between SART, ES, and MAAS is insufficient reliability of measures derived from these instruments. Reliability is an often overlooked quality metric in cognitive tasks while it is routinely reported for questionnaires⁴⁰. Reliability estimates are important as they determine an upper limit of how large correlations between two measures can be. For all the individual measures for mindfulness and MW discussed above, robust psychometric properties have been reported before, though rarely combined and sometimes in small samples: MAAS^{31,32}, specific SART measures with and without ES for MW^{41–45}. Table 1 lists all referenced studies that measured the MAAS and/or the SART with or without ES of MW. The table depicts sample sizes, reliability estimates and estimates of association. Most importantly, this table shows that none of the previous studies employed all three measures (MAAS, SART, and ES of MW) and reported both, reliabilities of all measures as well as correlations between all of them. The present study aims to fill this gap and offers data from two new large samples.

Overall the aim of this study is to assess the psychometric quality of several measures for MW and mindfulness from the SART, MAAS, ES of MW and ES of meta-awareness. In a second step, we want to gain an estimate for the statistical association between these constructs. We combined ES of MW during the SART with an established measure of mindfulness in an online study in two large samples collected in an online experiment and by doing so add psychometric estimates for these measures gained in an online study and assessed together.

Citation	Measures used				N	Reliability estimates				Estimates of association					
	MAAS	SART (acc)	SART (RT)	ES of MW		MAAS	SART (acc)	SART (RT)	ES of MW	SART (acc) and MAAS		SART (RT) and MAAS		ES of MW and MAAS	
Mrazek et al. 2012 ³	✓	✓	×	✓	113	0.85	–	–	–	–0.23	*	–	–	–0.22	*
Smilek et al. 2010 ⁴⁶	✓ ^a	✓	✓	×	363	–	–	–	–	0.22 ^a	**	–0.16 ^a	***	–	–
Cheyne et al. 2006 (Exp. 1) ⁴⁷	✓	×	×	×	449	0.87	–	–	–	–	–	–	–	–	–
Cheyne et al. 2006 (Exp. 2) ⁴⁷	✓	✓	✓	×	504	0.88	0.86	0.98	–	–0.31	***	0.17	***	–	–
Brown and Ryan 2003 ³¹	✓	×	×	×	74–327	0.80–0.87	–	–	–	–	–	–	–	–	–
Park et al. 2013 ³²	✓	×	×	×	Review	0.87–0.92	–	–	–	–	–	–	–	–	–
Deng et al. 2014 ⁴	✓	✓	×	✓ ^c	23	–	–	–	–	–0.45	*	–	–	0.44 ^c	*
Nayda et al. 2021 ⁵	✓	✓	×	✓ ^c	200	0.85	–	–	–	–0.23	**	–	–	0.31 ^c	***
Sofuoglu et al. 2008 ⁴¹	×	✓ ^e	✓	×	11	–	0.94–0.98	–	–	–	–	–	–	–	–
O’Connell et al. 2009 ⁴²	×	✓	✓	×	13	–	0.87, 0.89	–	–	–	–	–	–	–	–
McVay and Kane 2009 ⁴³	×	✓ ^b	✓	✓	244	–	0.95 ^b	0.93	0.89	–	–	–	–	–	–
Unsworth and McMillan 2014 ⁴⁴	×	✓	✓	✓	252	–	0.83	0.92	–	–	–	–	–	–	–
Kane et al. 2016 ⁴⁵	×	✓	✓	✓	472	–	0.96	0.98	0.93	–	–	–	–	–	–
McKillop et al. 2007 ⁴⁸	✓	×	×	×	727	0.89	–	–	–	–	–	–	–	–	–
Michalak et al. 2008 ⁴⁹	✓	×	×	×	469	0.83	–	–	–	–	–	–	–	–	–
Medvedev et al. 2016 ⁵⁰	✓	×	×	×	250	0.87	–	–	–	–	–	–	–	–	–
This study (GUP)	✓	✓ ^d	✓	✓	313	0.96	0.65	0.98	0.91	0.03 ^d	n.s.	0.06	n.s.	–0.29	***
This study (EPP)	✓	✓ ^d	✓	✓	228	0.88	0.71	0.99	0.89	0.13 ^d	*	–0.06	n.s.	0.041	n.s.

Table 1. Literature summary on MAAS, SART, and ES of MW measures. Letters a–e indicate measures with inverse direction from those used by most other studies, to help with the comparison of association estimates. a = MAAS-LO used (a value of attention lapses which is inverse to the MAAS score that gives a value for mindfulness); b = used signal-detection sensitivity as SART measure; c = used on-task rate for ES of MW, while most other studies use an error rate; d = proportion of correct Nogo trials and not the error rate; e = Go trials error rate. Significance markers: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. n.s. = not significant. N = samples size(s). GUP = German-speaking unpaid participants. EPP = English-speaking paid participants.

Results

We recruited two samples of participants for a German and an English version of the experiment. In our first recruitment phase we targeted German-speaking participants through the participant pool of our institution, made up of students and volunteers from the public. Throughout the study, we refer to this sample as German-speaking unpaid participants (GUP). In a second phase, we recruited and paid participants predominantly through Amazon Mechanical Turk (AMT, mturk.com) for an English version of the experiment. We refer to this sample as English-speaking paid participants (EPP). All participants first answered a questionnaire on demographics and the MAAS, then they performed a 20 min version of the SART during which ES probes of MW and meta-awareness were obtained (see “Methods” section).

Sample differences. We initially planned to report our findings as one sample, since the online experiment was identical except for the language. However, after finding significant differences between our two samples in the SART and ES data, we decided post-hoc to report all findings separately for GUP and EPP (see Table 2 for sample differences between all main measures). Most strikingly, EPP reported significantly less than half as often to be ‘off task’ than the GUP ($t(519.01) = -10.06, p < .001, d = 0.81, M_{EPP} = 0.09, M_{GUP} = 0.25$). This indi-

Measurement	EPP		GUP		Mean difference test				Variance difference test		
	Mean	SD	Mean	SD	Stat	df(s)	<i>p</i>	Cohen's <i>d</i>	Stat	df(s)	<i>p</i>
Reaction time											
<i>All trials</i>											
Mean Go	659	172	560	116	7.54	372.13	< .001	0.7	36.1	1.539	< .001
Mean Nogo	586	219	470	161	6.56	369.35	< .001	0.62	23.5	1.507	< .001
SD Go	251	77	211	59	6.46	410.36	< .001	0.59	15.64	1.539	< .001
SD Nogo	238	183	142	128	6.2	324.23	< .001	0.62	40.37	1.439	< .001
Accuracy											
<i>All Trials</i>											
Go	0.95	0.07	0.97	0.05	- 4.45	387.75	< .001	- 0.41	42.1	1.539	< .001
Nogo	0.83	0.13	0.86	0.11	- 3.3	426.81	.001	- 0.3	14.1	1.539	< .001
ES MW probes											
Attention Off	0.09	0.14	0.25	0.23	- 10.06	519.01	< .001	- 0.81	72.1	1.539	< .001
Meta-Awareness Off	0.36	0.36	0.26	0.28	2.81	238.1	.005	0.31	17.43	1.426	< .001

Table 2. Sample difference tests. Samples sizes were $n = 228$ and $n = 313$ for the EPP and GUP sample, respectively. For brevity, we only report parametric tests here even though several of the variables show a non-normal distribution. Non-parametric tests (Mann–Whitney U tests instead of *t* tests for mean differences and Brown–Forsythe tests instead of the Levene's tests for variance differences) showed a similar overall pattern of results, see Table S8 in the supplementary materials at <https://osf.io/8kg6z>. SD = standard deviation, EPP = English-speaking paid participants, GUP = German-speaking unpaid participants.

cates much lower variance in ES data in the EPP. There were also significant differences on all measures derived from the SART directly (RT, accuracy) albeit in a lower magnitude (see Table 2).

Factorial structure and reliability of the MAAS. We first checked the correlation matrices of the individual items on the questionnaire and the total score, separately for each of the two samples. In the GUP sample, item 6 had low item-to-total correlation ($r = 0.05$) and correlations below $r = 0.2$ with most other items. For that reason, we excluded item 6 from further analyses for the GUP. Thus, our total MAAS score for the EPP contained all 15 initial items, while the score of the GUP contained only 14 items.

We then conducted an exploratory factor analysis (EFA) for the MAAS responses for each of the two samples (factor loadings for one-factor EFAs are presented in Table 3). Figure 1 depicts scree plots for the EPP and GUP; these plots suggest that a single latent factor drives responses in the MAAS. Further EFAs also revealed that two-factor models only explain little additional variance (EPP: 3% and GUP: 5%), in comparison to that explained by one-factor models (EPP: 63% and GUP: 36%). However, the Kaiser rule (selecting the factors with an eigenvalue above 1; indicated by the dotted line in the scree plots) is also in accordance with a two-factor solution in our GUP.

The model fit statistics from confirmatory factor analyses (CFA) were estimated using the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), and the Root mean square error approximation (RMSEA). We compared the values against common standards for an acceptable fit (CFI/TLI > 0.9, RMSEA < 0.06)⁵². For one-factor models, the fits are acceptably high in the EPP (CFI = 0.954, TLI = 0.946). The fits were poorer, however, for the GUP (CFI = 0.858, TLI = 0.832). The RMSEA, which is an absolute fit statistic, indicates a poor approximate fit for both models, in the EPP (RMSEA = 0.08) and GUP (RMSEA = 0.096). However, the use of a fixed threshold for the RMSEA is questionable^{53,54}. The full fit statistics of these two models and of an alternative two-factor model for the GUP can be found in the supplementary materials at <https://osf.io/8kg6z>. Together, EFA and CFA are mostly consistent with the notion of one single factor driving responses to the MAAS, even though some fit statistics for the CFA were below the threshold for an acceptable fit.

Reliabilities of the MAAS score (mean of individual items) were overall high. For the full MAAS the standardized Cronbach's α was 0.88 in the GUP sample and 0.96 in the EPP. We created and then investigated shorter versions of the questionnaire consisting of the three or five items with the highest loadings in the EFAs. In the EPP these items were 7, 8, 10, 1, 11, and in the GUP items 14, 8, 9, 10, 7, in order of decreasing loading (see also Table 3). We refer to these shortened scales as the MAAS-5 and MAAS-3. The Cronbach's α s of the scales are given in Table 4 and further descriptives of the scores are available in the supplementary materials (Table S7). Correlations between short and full MAAS scores were reasonably high (between $r = 0.79$ and $r = 0.97$, see full correlation matrices in the supplementary materials (Figs. S9 and S10).

Reliability of MW measures taken from the SART and ES. Estimates of reliability of the measures derived from the SART and ES probes are presented in Table 4. They are split-half reliabilities derived using a permutation-based approach with 5000 random splits^{40,57}. For further descriptives of the measures, see Table S7 in the supplementary materials. From the SART, we report these measures: accuracy, the mean (M) and standard deviation (SD) of RTs during all trials and also in only those trials preceding the ES probes within a 10-s time window, a measure used in MW neuroimaging studies²⁷. SART values are reported separately for correct Go trials and incorrect Nogo trials. From ES probes, we report the proportion of all ES probes in which participants

Items	Loading	Communality
German MAAS (used with GUP)		
1. Ich könnte ein Gefühl haben und mir dessen erst irgendwann später bewusst werden.	0.45	0.20
2. Ich zerbreche oder verschütte Dinge aus Achtlosigkeit, ohne den Dingen Aufmerksamkeit zu schenken oder weil ich an anderes denke.	0.39	0.16
3. Ich finde es schwierig, auf das konzentriert zu bleiben, was im gegenwärtigen Augenblick passiert.	0.64	0.41
4. Ich neige dazu, schnell zu gehen, um dorthin zu kommen, wo ich hingehere, ohne darauf zu achten, was ich unterwegs erlebe.	0.53	0.28
5. Ich neige dazu, Gefühle körperlicher Anspannung oder Unwohlsein nicht wahrzunehmen, bis sie meine Aufmerksamkeit vollständig in Anspruch nehmen.	0.52	0.27
7. Es sieht so aus, als würde ich "automatisch funktionieren", ohne viel Bewusstsein für das, was ich tue.	0.66 ^a	0.44
8. Ich hetze durch Aktivitäten, ohne wirklich aufmerksam für sie zu sein.	0.74 ^{a,b}	0.55
9. Ich bin so auf das Ziel konzentriert, das ich erreichen möchte, dass ich den Kontakt dazu verliere, was ich hier und jetzt tue, um dieses Ziel zu erreichen.	0.69 ^{a,b}	0.47
10. Ich erledige Aufträge oder Aufgaben automatisch, ohne mir bewusst zu sein, was ich tue.	0.68 ^a	0.47
11. Ich bemerke, wie ich jemandem nur mit einem Ohr zuhöre, während ich gleichzeitig etwas anderes tue.	0.50	0.25
12. Ich fahre zu Orten wie von einem "Autopiloten" gesteuert und frage mich dann, wie ich dorthin gekommen bin.	0.60	0.36
13. Ich bemerke, dass ich gedankenverloren der Zukunft oder der Vergangenheit nachhänge.	0.53	0.28
14. Ich merke, wie ich Dinge tue, ohne auf sie zu achten.	0.78 ^{a,b}	0.61
15. Ich esse eine Kleinigkeit, ohne mir bewusst zu sein, dass ich esse.	0.51	0.26
English MAAS (used with EPP)		
1. I could be experiencing some emotion and not be conscious of it until some time later.	0.82 ^a	0.67
2. I break or spill things because of carelessness, not paying attention, or thinking of something else.	0.79	0.62
3. I find it difficult to stay focused on what's happening in the present.	0.80	0.63
4. I tend to walk quickly to get where I'm going without paying attention to what I experience along the way.	0.81	0.65
5. I tend not to notice feelings of physical tension or discomfort until they really grab my attention.	0.79	0.62
6. I forget a person's name almost as soon as I've been told it for the first time.	0.75	0.56
7. It seems I am "running on automatic," without much awareness of what I'm doing.	0.86 ^{a,b}	0.74
8. I rush through activities without being really attentive to them.	0.86 ^{a,b}	0.74
9. I get so focused on the goal I want to achieve that I lose touch with what I'm doing right now to get there.	0.79	0.63
10. I do jobs or tasks automatically, without being aware of what I'm doing.	0.85 ^{a,b}	0.72
11. I find myself listening to someone with one ear, doing something else at the same time.	0.82 ^a	0.66
12. I drive places on 'automatic pilot' and then wonder why I went there.	0.80	0.64
13. I find myself preoccupied with the future or the past.	0.67	0.45
14. I find myself doing things without paying attention.	0.80	0.63
15. I snack without being aware that I'm eating.	0.78	0.61

Table 3. MAAS EFA results for models with one factor. Item 6 ("Ich vergesse den Namen einer Person fast sofort nachdem er mir erstmals gesagt wurde.") was excluded in the German MAAS due to low correlations ($r < 0.2$) with other items. ^aFive highest loading items used in the short scales MAAS-5. ^bThree highest loading items used in the short scales MAAS-3. Communality refers to a variable's variance accounted for by all factors of the model⁵¹.

answered that they were off-task (Attention Off) and the proportion of meta-awareness probes in which participants answered that they were unaware that their attention was off task (Meta-Awareness Off). The sample sizes for the meta-awareness probes were smaller, because they exclude participants who reported that they were always on task. Split-half reliabilities for measures from Go trials in the SART and for ES probes are generally high. Reliabilities for Nogo trials were markedly lower, and were further reduced when restricting the analyses to the 10-s time windows immediately preceding ES probes. It is noteworthy that the sample sizes varied for these different measures due to the structure of the data and restrictions for the split-half calculations: Each participant needed at least four valid data points for the split-half procedure, as each split required two data points to calculate a mean or standard deviation. Furthermore, only 10.6% of all trials were Nogo trials and participants only reacted to 15.2% of Nogo trials, making Nogo trials with participant reaction somewhat scarce.

Estimates of association between the MAAS, SART, and ES. In a next step, we assessed the hypothesized negative association of MW with mindfulness. To this aim, we correlated measures derived from the SART and ES with the MAAS (Fig. 2). For the link between the direct measure of MW and mindfulness, we found ES probes (Attention Off) were moderately negatively associated with the MAAS in GUP ($r = -.29$, $p < 0.001$) but not in EPP ($r = 0.04$, $p > 0.1$). Between indirect measures of MW and mindfulness, there was no indication for an association between the SART and the MAAS in GUP. In EPP, however, there were small correlations

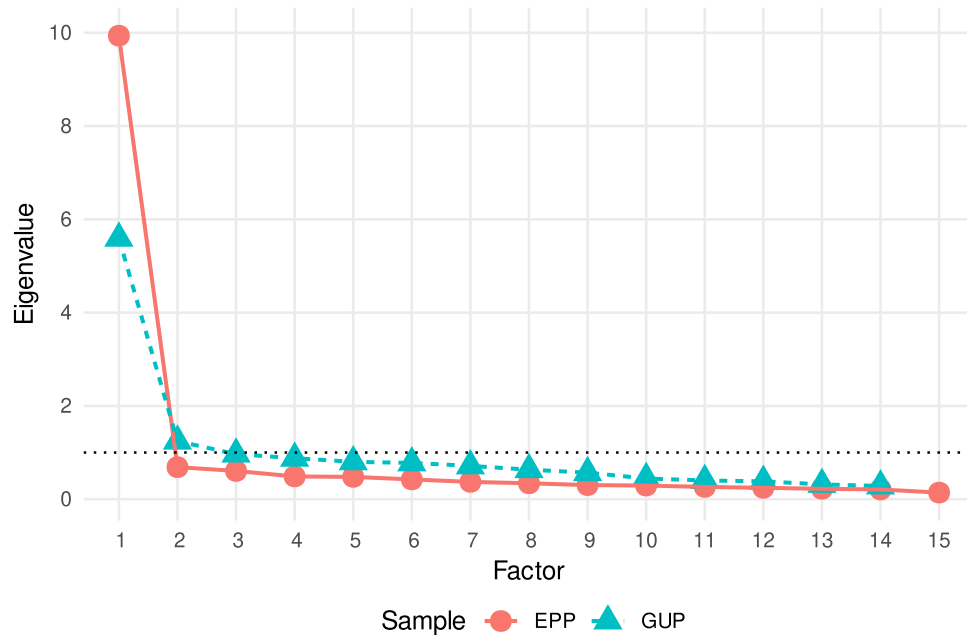


Figure 1. Scree plot for MAAS for EPP and GUP samples. This figure was created using R (v. 4.02)⁵⁵ with package ‘ggplot2’ (v. 3.3.5)⁵⁶.

between MAAS total score and SD of RTs in the Go trials during the 10 s window before ES probes ($r = -.23$, $p < 0.05$), between MAAS total score and accuracy in all Nogo trials ($r = .13$, $p < 0.05$), and a medium association between MAAS total score and accuracy of Nogo trials in the 10 s window before ES probes ($r = -.43$, $p < 0.01$). The pattern is mostly consistent with the idea of a negative association of MW and mindfulness. There was no association between meta-awareness probes and MAAS scores in both samples. All pairwise correlations for both samples are available in Tables S1 and S2 in the supplementary materials at <https://osf.io/8kg6z>.

To check whether these correlations might have been heavily influenced by outliers or non-normally distributed data, we additionally bootstrapped the correlation coefficients and 95% confidence intervals (CIs) for these pairwise correlations (1000 iterations, 100 random participants sampled in each). In addition, we compared the Pearson product-moment correlations to Spearman rank correlations. These analyses showed a similar pattern of results from the Pearson correlations reported above in the GUP, but in the EPP the three reported associations with ES probes were not significant in the Spearman correlations. This further indicates the different answer patterns in self-reported MW between our two samples. The detailed results of these additional versions of the correlations are available in Tables S3–S6 in the supplementary materials.

This study entailed between-subject manipulations hypothesized to affect MW that are out of the scope of the current work. Briefly, we investigated whether exposing participants to auditory stimuli (5 Hz monaural or binaural auditory beats, silence, 440 Hz sine tone) could reduce their propensity to MW. Since such a finding has been reported earlier, in particular for participants exhibiting high proportions of MW²⁴, we experimentally manipulated the occurrence of MW in three different ways. First, we varied the inter-stimulus-interval (1 vs. 2 s). Second, we implemented the stimuli in the SART predictably or unpredictably. Third, a creative problem-solving task was executed for a second time after the SART, and participants were either informed before the SART about the second execution or they were not informed.

These between-subject manipulations may have affected our estimates of associations between MW and mindfulness. To investigate this possibility, we first calculated ANOVAs with the experiment’s main manipulations (and all pairwise interactions) as predictors and measures from SART and ES as outcome variables. We then added the MAAS score as covariate to these, to create a set of comparable ANCOVAs. To evaluate whether our associations were affected by the experimental manipulations, we then checked two things. First, we compared the effect sizes (η^2) of the total MAAS score in these ANCOVAs with the coefficient of determination (r^2) between the MAAS score and SART and ES measures. Second, we calculated model comparisons between the ANOVAs and ANCOVAs using likelihood-ratio tests (Table 5).

The effect sizes were for most combinations very similar in the correlations and the ANCOVAs. In all but one case, adding the MAAS score as covariate did not significantly improve the model fit. Only in the ES MW variable in GUP did adding the MAAS score as covariate significantly improve the model fit. There the estimate of association between ES MW and the MAAS score slightly increased when accounting for experimental manipulations. This result provides confirmatory evidence that MAAS and ES MW are weakly negatively associated in the GUP sample.

Measurement	GUP				EPP			
	Reliability ^a			n	Reliability ^a			n
	Estimate	95% CI			Estimate	95% CI		
		Lower	Upper		Lower	Upper		
MAAS								
MAAS total ^b	0.88			313	0.96			228
MAAS-5	0.92			313	0.93			228
MAAS-3	0.84			313	0.91			228
SART								
<i>Reaction time</i>								
<i>All trials</i>								
Mean Go	0.98	0.98	0.99	313	0.99	0.99	0.99	228
Mean Nogo	0.58	0.46	0.69	150	0.59	0.49	0.69	134
SD Go	0.89	0.87	0.91	313	0.94	0.92	0.95	228
SD Nogo	0.31	0.09	0.46	150	0.31	0.12	0.46	134
<i>10 s Window</i>								
Mean Go	0.98	0.97	0.98	313	0.98	0.97	0.99	96
Mean Nogo	0.33	- 0.22	0.82	13	0.36	- 0.07	0.79	9
SD Go	0.82	0.79	0.85	313	0.88	0.84	0.91	96
SD Nogo	0.26	- 0.56	0.90	13	0.24	- 0.31	0.89	9
Accuracy								
<i>All Trials</i>								
Go	0.95	0.94	0.97	313	0.96	0.95	0.97	228
Nogo	0.65	0.58	0.70	313	0.71	0.65	0.76	228
<i>10 s Window</i>								
Go	0.89	0.85	0.92	313	0.89	0.85	0.93	96
Nogo	0.54	0.40	0.66	137	0.36	0.11	0.56	53
ES MW Probes								
Attention Off	0.91	0.90	0.93	313	0.89	0.86	0.92	228
Meta-Awareness Off	0.74	0.68	0.79	182	0.80	0.72	0.87	47

Table 4. Reliabilities for MAAS, SART, and ES measures in both samples. ^aReliability estimates are Cronbach's alphas in case of MAAS scores and split-half reliabilities for SART and ES measures. Split-half estimates are Spearman-Brown corrected and were derived using a permutation-based approach with 5000 random splits⁵⁷. ^bMAAS total score included all 15 items in EPP, but only 14 items (excluding item 6) in the GUP. 95% CI = 95% confidence interval, SD = standard deviation.

Discussion

We examined the psychometrics of MW, meta-awareness of MW, and trait mindfulness, as well as the associations between these constructs. Overall, we found reasonably good psychometrics of all measures, and evidence that MW and trait mindfulness are indeed moderately negatively correlated. This association was not moderated by meta-awareness of MW. Neither the psychometrics nor moderating effects of meta-awareness can therefore readily explain that associations between MW and mindfulness are of a rather low magnitude.

In keeping with previous studies, we found overall good psychometric properties and evidence mostly consistent with a single-factor structure for the MAAS questionnaire. Our estimates of reliability of the MAAS were slightly higher than those reported in earlier studies, in both the EPP and GUP. For the English MAAS, the original publication reported internal consistencies in the range of [0.8, 0.87]³¹, and a further study reported 0.89⁴⁸, but this value was 0.96 in our EPP. For the German MAAS, a Cronbach's α of 0.83 was reported in the initial publication on the psychometric properties of the questionnaire⁴⁹, while the value in our GUP was 0.88. Very high internal consistencies might indicate redundancy in a questionnaire, suggesting some items are superfluous and can be removed, which would lead to a more efficient assessment⁵⁹. Results on the proposed shorter versions of the MAAS (MAAS-5 and MAAS-3) outlined in this study support this notion and may provide researchers with tools to optimize data collection.

One peculiarity we observed in the MAAS data for the GUP was item 6 ("I forget a person's name almost as soon as I've been told it for the first time."³¹), which correlated very poorly with all other items and the total score. Interestingly, the authors of the German MAAS also observed complications with this item but decided to include it to ensure international comparability⁴⁹. Specifically, they found an item-to-total correlation of $r = 0.18$ for item 6 while the next-lowest correlation was for item 1 ($r = 0.26$) and those for all other items ranged from 0.33 to 0.65. We did not observe, however, such a low item-to-total correlation of item 6 in EPP. Nevertheless, we assume that cultural differences or mere issues related to translation cannot account for low item-to-total correlation for

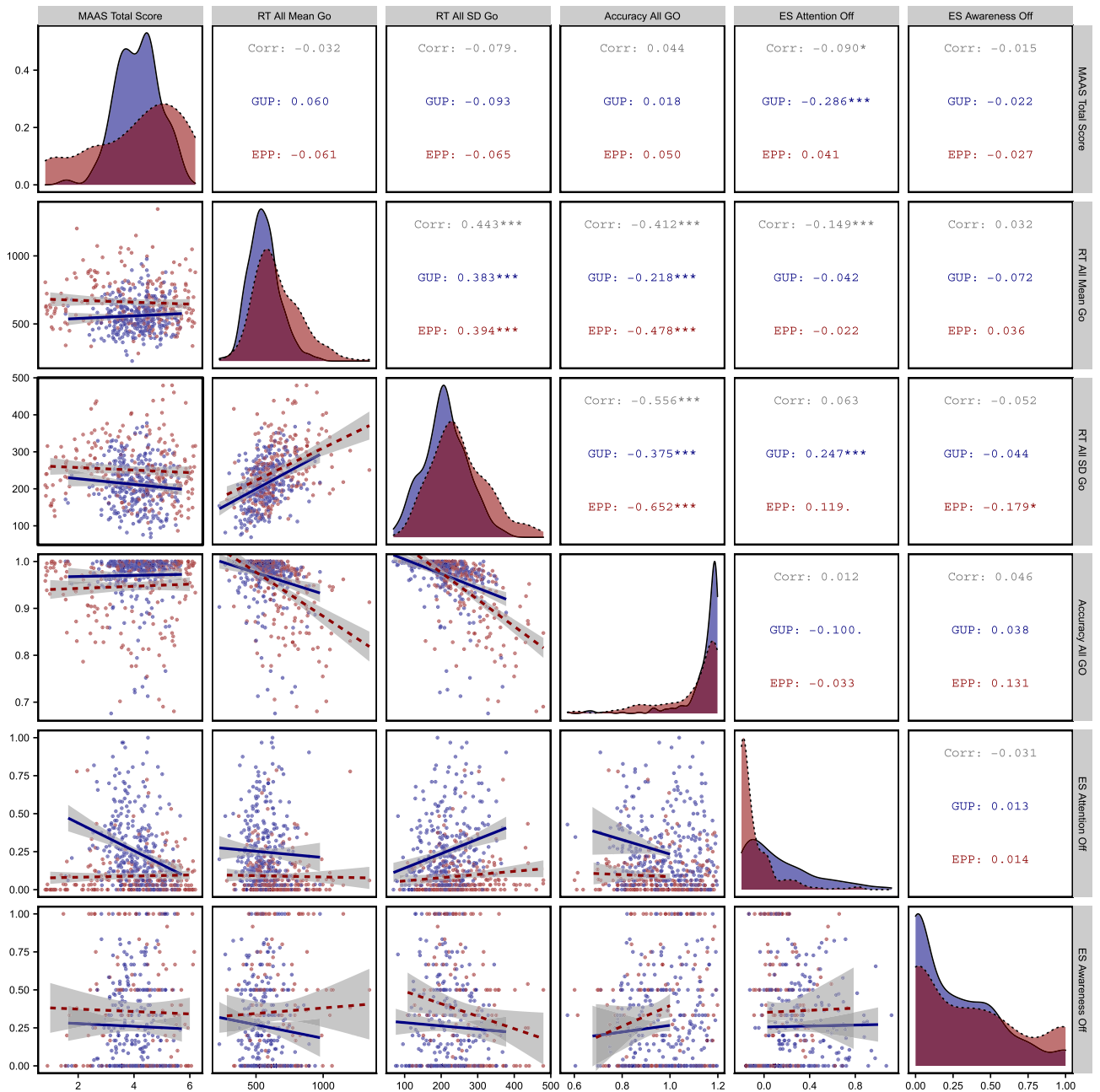


Figure 2. Pairwise Pearson correlations for MAAS, SART, and ES measures. Correlation coefficients are reported for whole sample ('Corr'), and for EPP and GUP samples separately. Individual plots below the diagonal are scatter plots with regression lines for the two variables intersecting at this cell, those on the diagonal show density distribution plots for the two samples. Significance markers: . = $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. This figure was created using R (v. 4.02)⁵⁵ with packages 'ggplot2' (v. 3.3.5)⁵⁶ and 'GGally' (v. 2.1.2)⁵⁸.

this item, as it was also observed in a study with English-speaking participants from New Zealand⁵⁰. Moreover, item 6 was also one of the most poorly correlated items in the original English article detailing the MAAS³¹. We suggest item 6 may only occasionally be problematic as its meaning is ambiguous, and can be understood in two different ways. First, it could—probably as intended by the authors of the scale—measure attention usually directed to a person introducing themselves, or it can be understood as asking for self-report on one's long-term memory abilities, which is arguably an altogether different trait than mindfulness.

While reliability is routinely reported for questionnaires such as the MAAS, they are less common for cognitive behavioral measures, e.g. for the MW measures derived from the SART and ES⁴⁰. Still, earlier studies generally reported high reliabilities also for the SART: e.g. between 0.83 and 0.89 for overall accuracy in the SART^{42,44}, between 0.92 and 0.98 for SDs of RT^{44,45}, and even as high as 0.94 to 0.98 for the accuracy of Nogo trials⁴¹ (see Table 1). Some of these studies, however, used a shorter stimulus-onset asynchrony (SOA) and much

Measurement	GUP						EPP					
	ANCOVA ^a		Correlation ^b		LR test	n	ANCOVA ^a		Correlation ^b		LR test	n
	η^2	<i>p</i>	r^2	<i>p</i>	<i>p</i>		η^2	<i>p</i>	r^2	<i>p</i>	<i>p</i>	
SART												
<i>Reaction time</i>												
<i>All Trials</i>												
Mean Go	0.004	0.248	0.004	0.293	0.232	313	0.002	0.449	0.004	0.362	0.427	228
Mean Nogo	0.008	0.123	0.010	0.086	0.110	296	0.013	0.100	0.013	0.101	0.083	213
SD Go	0.006	0.160	0.009	0.102	0.146	313	0.010	0.124	0.004	0.327	0.107	228
SD Nogo	0.005	0.245	0.008	0.151	0.225	250	0.001	0.648	0.001	0.699	0.629	191
<i>Accuracy</i>												
<i>All Trials</i>												
Go	< 0.001	0.971	< 0.001	0.751	0.970	313	0.001	0.603	0.002	0.456	0.586	228
Nogo	0.003	0.281	0.001	0.538	0.265	313	0.007	0.186	0.017	0.048	0.165	228
ES MW Probes												
Attention Off	0.080	< 0.001	0.082	< 0.001	< 0.001	313	0.0005	0.739	0.002	0.541	0.727	228
Meta-Awareness Off	0.001	0.684	0.0005	0.708	0.673	283	0.002	0.572	0.001	0.748	0.542	145

Table 5. MAAS total score as covariate compared to correlation coefficients and model comparison ANOVA versus ANCOVA. ^aEffect sizes for the MAAS total score covariate in ANCOVAs. ^bPearson product-moment correlations between MAAS total score and the SART and ES measures. LR test = Likelihood-ratio test between ANOVA and ANCOVA models. η^2 = Eta-squared effect size as proportion of variance explained by the predictor. r^2 = r-squared coefficient of determination based on the correlation coefficients, as proportion of shared variance among the two correlated variables.

smaller sample sizes (13⁴² and 12⁴¹ participants). Also, earlier studies reporting SART reliabilities were usually laboratory studies with more controlled environments. These factors might have led to even higher reliability estimates for measures derived from Nogo trials. Our study adds further reasonably high reliabilities with alphas ranging from 0.84 to 0.99, on measures derived from the Go trials of the SART. In contrast to previous studies, reliability estimates for measures derived from Nogo trials were markedly lower (between 0.24 and 0.71) in our samples. These were probably low in our study due to only a small fraction of the SART trials that can be used to derive these measures as we increased the SOA from the original version in order to foster MW. Overall reliabilities are further reduced when restricting the analyses to a short time window preceding ES probes. Filtering the usable trials to a specific time window seems predominately appropriate for neuroimaging studies looking to isolate brain activity patterns of MW, which is where this analysis strategy originated²⁷. Researchers focusing on Nogo trials and segmenting the data accordingly, should therefore take care to ensure that the number of trials analyzed remains reasonably large, and bear in mind that reliability of measures derived with these strategies is likely limited. Our reliability estimates for the ES MW probes during the SART (0.91 in GUP and 0.89 in EPP) were within the range of what earlier studies reported (e.g., 0.89⁴³ and 0.93⁴⁵). Together with the reliability estimates of the MAAS, our study demonstrates that high reliabilities of the MAAS, SART, and ES for MW can also be obtained in an online study setting.

Our results also stress notions of caution related to recruiting participants via crowdsourcing platforms such as—as in our case—Amazon Mechanical Turk (AMT, mturk.com). We noticed that the two samples behaved differently in the SART and ES, in that AMT participants (the EPP) were less likely to respond that their attention had been ‘off task’ but at the same time showed lower accuracy rates and slower, more varied RTs during the SART. This is likely to have also affected the estimate of association between self-reports of MW in ES probes and the MAAS score. A significant correlation was found in GUP, but not in EPP. The absence of a significant correlation could be due to lower variance in the ES probes of EPP versus GUP. We suggest the different patterns of results relating to the ES probes is not simply due to cultural or language differences, but rather due to differences in motivation to participate. Requesters at AMT are allowed to withhold payment if they are not satisfied with the performance of the participant. It thus seems reasonable to assume that some participants recruited through AMT reported being on task even when they were not. Our data underlies arguments made earlier that caution is warranted when recruiting via AMT and similar platforms, especially when using measures that are susceptible to the issues discussed above^{60,61}. It might help to explicitly ensure participants that they will experience no disadvantages when they report being off task.

Our results support the hypothesis of a negative link between trait mindfulness and MW. Associations, however, were scattered over different measures and differed between our two samples: There was a moderate correlation of the MAAS with the self-report measure of MW (ES probes during the SART) in one of our samples (GUP) and with SART SDs of RTs and SART accuracy in the SART in the other sample (EPP). Low and absent associations between MW and mindfulness cannot be explained by low reliabilities of the measures we used, as reliabilities were generally satisfyingly high, with the exception of measures derived from SART Nogo trials. With that in mind, the associations based on measures with high reliabilities are only two: that between MAAS total score and ES MW in the GUP, and between MAAS total score and SDs of RTs in SART Go trials during the 10 s window before ES probes in the EPP. One potential explanation for finding the clear association

between MAAS and ES MW only in the GUP might be a lack of variance in the EPP data as mentioned above. The lack of variance was due to a large proportion of participants who answered that they were rarely or never 'off task' during the experiment.

Despite good psychometrics of our measures, the link between trait mindfulness and MW was only moderate. A further explanation for rather low associations could be that meta-awareness of MW moderates the hypothesized associations. Our finding that meta-awareness of MW is not linked to mindfulness goes against such a hypothesis and some empirical evidence^{7,23}. However, our results are in accordance with more recent papers that also do not find a moderating effect of meta-awareness on the association between MW and mindfulness^{4,5}. Nayda et al.⁵ reported negative associations between both, the propensity to 'tune out' (meta-aware MW) with mindfulness, and the propensity to 'zone out' (meta-unaware MW). An earlier publication by Deng et al.⁴ found insignificant correlations between trait MW and both 'zone out' and 'tune out' propensities. It seems noteworthy that both correlations of the Deng et al.⁴ study are in the same range and direction as in Nayda et al.⁵ but do not reach statistical significance likely due to the low sample size ($N = 23$). A potential caveat here is that these rates are calculated using the total of MW probes, rather than the total of meta-awareness probes only. These estimates are therefore biased in that the sum of the 'tune out' and 'zone out' rates is perfectly inverse proportional to the 'on-task' rate. In our analyses, we calculated the meta-awareness rate as proportion of the total of meta-awareness probes instead of the total of MW probes. We found no significant correlation between meta-awareness of MW and mindfulness. Thus, further research seems needed to isolate a potentially moderating effect of meta-awareness on the correlation between MW and mindfulness.

A further reason for low associations between MW and mindfulness could result from the difference in the trait versus transient nature of the constructs. Mindfulness is conceived and measured as a general personality trait. However, MW is a much more transient and fluctuating phenomenon during an ongoing and often boring task. Moreover, boredom itself may explain the low associations between MW and mindfulness. In MW research, the SART is often chosen as an ongoing task, because it is boring and therefore is thought to facilitate MW. The notion that boredom is an enabling factor for MW is supported by two findings. First, boredom has been shown to correlate with attentional lapses as measured with the SART⁶². Second, positive correlations between boredom and MW have been recently reported¹⁶³. In contrast, when participants respond to the mindfulness questions of the MAAS, it is unclear to what extent participants consider boring ongoing tasks (e.g., "I rush through activities without being really attentive to them." see Table 3 for the complete list of items of the MAAS). Therefore, while boredom seems a relevant aspect of MW when measured with the SART, this is not assessed with the MAAS. Together, this emphasizes the necessity of investigating the role of boredom in the relation between MW and mindfulness in future studies.

One may argue that a further reason for low associations between MW and trait mindfulness could be that the on-task state is more heterogeneous than previously thought. Heterogeneous on-task states were identified by assessing ongoing thought with multidimensional experience sampling (MDES), i.e., extending ES with several questions inquiring about the thoughts' content and nature⁶⁴. Principal component analysis (PCA) of MDES data revealed several components taxing into the on-task state, which were associated with distinct neural correlates^{65–68}. One component was related to self-focused off-task thoughts while another component indicated detailed task focus. This task-focused component was common in cognitively demanding tasks like tasks measuring working memory, task switching, and gambling. However, it was less observed in low-demand tasks like the SART, where self-focused off-task thoughts prevail⁶⁴. Together, these studies suggest that being more mindful might be linked to how people engage with tasks, perhaps by doing so in a more focused way. The possibility of multiple on-task states may therefore, contribute to the relatively low estimate of the association between mindfulness and ongoing thought.

Finally, low associations between MW and mindfulness could be due to insufficient validity, rather than reliability of the measures we used. While our current study focuses on reliability others have focused on issues related to validity, especially concerning the questionnaires to measure mindfulness⁶⁹. On the one hand, the MAAS in particular has been shown to correlate reasonably well with other questionnaires measuring mindfulness such as the Five Facet Mindfulness Questionnaire (FFMQ)⁷⁰. Further evidence for converging validity with, e.g., positive affect or attention, as well as evidence for discriminant validity, e.g., with anxiety and rumination, has been found in studies reporting correlations with MAAS scores^{31,32}. On the other hand, questionnaires rely on introspective capabilities and may be subject to bias. A recent study by Isbel et al.⁷⁰ questioned especially the discriminant validity of the MAAS and the FFMQ as these measures increased following both a mindfulness training intervention and a control training intervention not aimed at mindfulness. Rather, objective accuracy of breath counting has been found to respond selectively to the mindfulness training intervention⁷⁰. A potential reason why the breath counting task responded selectively to the mindfulness training is that mindfulness training itself often consists of exercises to guide one's attention specifically on the breath. It is hence a rather near transfer from mindfulness training to an increase in accuracy in breath counting. Nevertheless, more research exploring the practical validity of mindfulness measures is required.

Recent methodological developments in MW research highlight limitations in our findings and offer advice for future research. Contemporary studies of ongoing thought that utilized MDES show that different tasks used in MW research elicit several distinct thought patterns to varying degrees^{64,67}. Our study is consequently limited by the fact that we only used the SART to investigate the association between individual variation in mindfulness and MW instead of several tasks. The SART also has the limitation that it does not lead to much detailed task focus and tends to stimulate self-focused MW⁶⁴. Due to that, it is unclear whether our findings generalize to other tasks or whether they are specific to the SART and thus to those types of ongoing thoughts more likely to be evoked by the SART like self-focused MW.

Besides the heterogeneity of ongoing thoughts, the relationship between MW and mindfulness is likely modulated by various other factors. A recent study has highlighted MW as a complex phenomenon that warrants a

Variable	Category	GUP	EPP
		<i>n</i>	<i>n</i>
Sex	Male	116	123
	Female	196	104
	Trans	0	1
	Other	1	0
Education	Compulsory Education	15	3
	High School Diploma or equivalent	55	29
	Professional Degree/Apprenticeship	70	1
	Associate Degree	0	13
	Bachelor Degree	83	118
	Master Degree	68	63
	Doctorate / PhD	10	1
	Other	12	0

Table 6. Participant demographics. SD = standard deviation. GUP = German-speaking unpaid participants. EPP = English-speaking paid participants.

multi-faceted approach that includes a) dispositional traits, like conscientiousness, agreeableness, or mindfulness, b) contextual factors, like motivation or alertness, and c) cognitive abilities, like working memory capacity⁷¹. If the relationship between MW and mindfulness is embedded within such a multi-faceted approach, the association between these two factors might be diluted by other potential confounding factors that were not accounted for. In this regard, future research will benefit from assessing MW and mindfulness with a broad set of tools including MDES and multiple tasks with variable demands that elicit different patterns of ongoing thoughts.

Methods

Participants. A total of 715 participants performed or started our online experiment between October 2019 and January 2021. We excluded participants from the data analysis for these reasons and in this order: Repeated participation ($n = 11$), incomplete data due to technical issues ($n = 1$), incomplete or delayed participation in the experiment (time in experiment < 23 min or > 120 min [$n = 59$]), low number of correct SART trials (proportion of correct Go trials $< 2/3$ [$n = 51$], or proportion of correct Nogo trials $< 1/2$ [$n = 22$]), and outliers who took a long time to answer the ES probes ($n = 30$). For this last point we established a cutoff based on the interquartile range (IQR) due to the highly skewed distribution of these values. Cutoff was the 75th percentile plus three times the IQR. We based our data analyses on a total sample of 541, separated into 313 GUP (aged between 16 and 85, $M = 38.78$, $SD = 12.95$) and 228 EPP (aged between 19 and 68, $M = 34.27$, $SD = 11.39$). Further demographic characteristics are given in Table 6.

We recruited participants for two different language versions of the experiment through various routes. The GUP ($n = 313$) consists of: (a) 97 participants recruited by the students of two classes in the autumn 2019 and spring 2020 semesters at UniDistance Suisse; (b) 200 students and members of the public interested in participating in experimental research from our institute's pool of research participants; and (c) 16 participants who followed links in an information email to university employees and on different websites. The EPP ($n = 228$) contains: (a) 217 participants recruited through AMT, (b) 10 who were PhD students at the Department of Epileptology at the University of Bonn, and (c) 1 who followed a link from an external website.

Those recruited through AMT were paid USD 3.50 when they had completed the whole experiment. Students in the Bachelor's program in Psychology at the UniDistance Suisse received course credits for their participation. Other participants received no compensation. All participants gave informed consent by reading information provided online and then checking off tick boxes in an online form before they proceeded to the experiment. The study was carried out following all the relevant guidelines and regulations. The study and its compliance with relevant guidelines and regulations was approved by the ethical review committee of the Faculty of Psychology at UniDistance Suisse (<https://distanceuniversity.ch/research/ethics-commission/>). In particular, all procedures are in accordance with the Declaration of Helsinki.

Design. The data reported here was collected in a study also investigating the effects of experimental manipulations on MW. Participants performed the SART with intermittent ES probes to directly obtain self-reports of episodes of MW. These experimental manipulations are outside the scope of the current work as they focus on potential effects of auditory beat stimulation on MW^{24,72} and will be reported elsewhere. Briefly, experimental manipulations were performed in a $4 \times 2 \times 2 \times 2$ between-subjects design and included the independent variables Auditory Beat Stimulation (5 Hz binaural, 5 Hz monaural; 440 Hz pure tone; no sound), SART ISI (1 or 2 s), Predictability of the Number Sequence in the SART (random or ascending), and Expectancy of an ensuing creativity task (expected or unexpected). Dependent variables are RTs and Accuracy during the SART and ES MW probes. It was for the purpose of this study, that we collected data using the MAAS.

Instruments. To assess trait mindfulness we applied the MAAS, a 15-item questionnaire that determines attention to the present in everyday experiences³¹. For the German version of the experiment, we used the validated German translation available from the Leibniz Institute for Psychology Information (ZPID)⁷³.

To measure MW indirectly through lapses in sustained attention in a deliberately monotonous task, we used the SART²¹. The SART is a Go/No-go task that uses digits as stimuli which are presented individually on screen with a fixation cross shown between stimuli. Participants are asked to react to all digits (Go trials) except for the number 7 (Nogo trials). We adapted the original SART with the intention to make it more monotonous, in order to elicit more MW. Specifically, we displayed each stimulus longer (2 s instead of 250 ms), had a longer ISI (1 or 2 s instead of 900 ms), and used a fixed font size (instead of randomly varying font sizes) to present our stimuli²¹.

We assessed self-reported MW using ES probes during the SART. In intervals between 25 and 35 s, participants were asked: “Immediately before this question appeared, was your attention focused ON the task or OFF task?” with a dichotomous forced-choice answer. When “OFF task” was selected, a second question appeared: “Were you aware that your attention was OFF task?” with a dichotomous forced choice answer again (yes or no). There was no time limit to answer these probes.

To further increase MW by adding a cognitive distraction during the SART and to assess participants’ creativity, we implemented a short task for divergent thinking based on the alternative/unusual uses concept originally introduced by Guilford⁷⁴. In this unusual uses task (UUT), participants were given 20 s to find alternative uses for a brick, with the original use described as “building houses”. Participants entered their answer in a large text field and were asked to enter one answer per line.

We implemented the MAAS and SART with ES as an online experiment using the JavaScript-based online experiment builder “lab.js” (<https://lab.js.org>)⁷⁵. Participants were required to wear headphones during the experiment. We included a headphone test before the SART to ensure participants had correctly placed the headphones and could listen to the stimulation. Runnable files and code for both language versions of the experiment can be found in the supplementary materials at <https://osf.io/8kg6z>.

Procedure. The online experiment started with information about the experiment, data processing, and informed consent request. This was followed by a short demographic questionnaire, the MAAS, the headphone test, the UUT, and 20 min of the SART. After the SART, a summary page informed the participants about their performance and a debriefing page gave further background information about the study.

Data processing, analysis and creation of figures and tables were done in R (v 4.0.2)⁵⁵, using the following packages in addition to base R: ‘tidyverse’⁷⁶ for various data wrangling and processing tasks and for data visualizations via ‘ggplot2’⁵⁶, ‘Ggally’⁵⁸ for data visualizations, ‘e1071’⁷⁷ for kurtosis and skewness calculations, ‘lubridate’⁷⁸ for handling of date and time data, ‘lavaan’⁷⁹ for confirmatory factor analyses, ‘stargazer’⁸⁰ to create and export LaTeX tables, ‘splithalf’⁵⁷ for permutation-based split-half calculations.

Data availability

The datasets generated and analysed for the current study are available in the Open Science Framework (OSF) repository, <https://osf.io/wg9q5>. Tables, figures, and other supplementary materials specifically for this publication are available in a different repository at OSF, <https://osf.io/8kg6z>.

Received: 7 September 2021; Accepted: 26 April 2022

Published online: 11 May 2022

References

- Smallwood, J. & Schooler, J. W. The science of mind wandering: Empirically navigating the stream of consciousness. *Annu. Rev. Psychol.* **66**, 487–518. <https://doi.org/10.1146/annurev-psych-010814-015331> (2015).
- Kabat-Zinn, J. *Full Catastrophe Living: Using the Wisdom of Your Body and Mind to Face Stress, Pain, and Illness* (Batam Books Trade Paperbacks, 2013).
- Mrazek, M. D., Smallwood, J. & Schooler, J. W. Mindfulness and mind-wandering: Finding convergence through opposing constructs. *Emotion* **12**, 442–448. <https://doi.org/10.1037/a0026678> (2012).
- Deng, Y.-Q., Li, S. & Tang, Y.-Y. The relationship between wandering mind, depression and mindfulness. *Mindfulness* **5**, 124–128. <https://doi.org/10.1007/s12671-012-0157-7> (2014).
- Nayda, D. M. & Takarangi, M. K. The cost of being absent: Is meta-awareness of mind-wandering related to depression symptom severity, rumination tendencies and trauma intrusions? *J. Affect. Disord.* **292**, 131–138. <https://doi.org/10.1016/j.jad.2021.05.053> (2021).
- Dunne, J. D., Thompson, E. & Schooler, J. Mindful meta-awareness: Sustained and non-propositional. *Curr. Opin. Psychol.* **28**, 307–311. <https://doi.org/10.1016/j.copsyc.2019.07.003> (2019).
- Smallwood, J. & Schooler, J. W. The restless mind. *Psychol. Bull.* **132**, 946–958. <https://doi.org/10.1037/0033-2909.132.6.946> (2006).
- Killingsworth, M. A. & Gilbert, D. T. A wandering mind is an unhappy mind. *Science* **330**, 932–932. <https://doi.org/10.1126/science.1192439> (2010).
- Smallwood, J., O’Connor, R. C., Sudbery, M. V. & Obonsawin, M. Mind-wandering and dysphoria. *Cogn. Emot.* **21**, 816–842. <https://doi.org/10.1080/02699930600911531> (2007).
- Malhi, G. S. *et al.* Mood disorders: Neurocognitive models. *Bipolar Disord.* **17**, 3–20. <https://doi.org/10.1111/bdi.12353> (2015).
- Fell, J. Is the hippocampus a potential target for the modulation of mind wandering in major depression? *Front. Psychiatry* **9**, 10. <https://doi.org/10.3389/fpsy.2018.00363> (2018).
- Yanko, M. R. & Spalek, T. M. Driving with the wandering mind: The effect that mind-wandering has on driving performance. *Hum. Factors* **56**, 260–269. <https://doi.org/10.1177/0018720813495280> (2014).
- Di Nocera, F. *et al.* Attentional control in accidents involving agricultural tractor operators. *Ergon. Des. Q. Hum. Factors Appl.* **26**, 17–23. <https://doi.org/10.1177/1064804617737444> (2018).
- Smallwood, J., Mrazek, M. D. & Schooler, J. W. Medicine for the wandering mind: mind wandering in medical practice. *Med. Educ.* **45**, 1072–1080. <https://doi.org/10.1111/j.1365-2923.2011.04074.x> (2011).

15. Smallwood, J., Fishman, D. J. & Schooler, J. W. Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychon. Bull. Rev.* **14**, 230–236. <https://doi.org/10.3758/BF03194057> (2007).
16. Leszczynski, M. *et al.* Mind wandering simultaneously prolongs reactions and promotes creative incubation. *Sci. Rep.* **7**, 10197. <https://doi.org/10.1038/s41598-017-10616-3> (2017).
17. Mooneyham, B. W. & Schooler, J. W. The costs and benefits of mind-wandering: A review. *Can. J. Exp. Psychol.* **67**, 11–18. <https://doi.org/10.1037/a0031569> (2013).
18. Medea, B. *et al.* How do we decide what to do? Resting-state connectivity patterns and components of self-generated thought linked to the development of more concrete personal goals. *Exp. Brain Res.* **236**, 2469–2481. <https://doi.org/10.1007/s00221-016-4729-y> (2018).
19. Hayes, S. C., Levin, M. E., Plumb-Villardaga, J., Villatte, J. L. & Pistorello, J. Acceptance and commitment therapy and contextual behavioral science: Examining the progress of a distinctive model of behavioral and cognitive therapy. *Behav. Ther.* **44**, 180–198. <https://doi.org/10.1016/j.beth.2009.08.002> (2013).
20. Gu, J., Strauss, C., Bond, R. & Cavanagh, K. How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clin. Psychol. Rev.* **37**, 1–12. <https://doi.org/10.1016/j.cpr.2015.01.006> (2015).
21. Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T. & Yiend, J. 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia* **35**, 747–758. [https://doi.org/10.1016/S0028-3932\(97\)00015-8](https://doi.org/10.1016/S0028-3932(97)00015-8) (1997).
22. Jonker, T. R., Seli, P., Cheyne, J. A. & Smilek, D. Performance reactivity in a continuous-performance task: Implications for understanding post-error behavior. *Conscious. Cogn.* **22**, 1468–1476. <https://doi.org/10.1016/j.concog.2013.10.005> (2013).
23. Smallwood, J., McSpadden, M. & Schooler, J. W. The lights are on but no one's home: Meta-awareness and the decoupling of attention when the mind wanders. *Psychon. Bull. Rev.* **14**, 527–533. <https://doi.org/10.3758/BF03194102> (2007).
24. Chaieb, L., Dermer, M., Leszczynski, M. & Fell, J. Modulation of mind wandering using auditory beat stimulation: A pilot study. *J. Cogn. Enhanc.* **4**, 40–48. <https://doi.org/10.1007/s41465-019-00137-4> (2019).
25. Manly, T., Robertson, I. H., Galloway, M. & Hawkins, K. The absent mind: Further investigations of sustained attention to response. *Neuropsychologia* **37**, 661–670. [https://doi.org/10.1016/S0028-3932\(98\)00127-4](https://doi.org/10.1016/S0028-3932(98)00127-4) (1999).
26. Smallwood, J. *et al.* Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Conscious. Cogn.* **13**, 657–690. <https://doi.org/10.1016/j.concog.2004.06.003> (2004).
27. Christoff, K., Gordon, A. M., Smallwood, J., Smith, R. & Schooler, J. W. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. In *Proceedings of the National Academy of Sciences*, Vol. 106 8719–8724, <https://doi.org/10.1073/pnas.0900234106> (2009).
28. Smucny, J., Rojas, D. C., Eichman, L. C. & Tregellas, J. R. Neuronal effects of auditory distraction on visual attention. *Brain Cogn.* **81**, 263–270. <https://doi.org/10.1016/j.bandc.2012.11.008> (2013).
29. Seli, P., Risko, E. F. & Smilek, D. On the necessity of distinguishing between unintentional and intentional mind wandering. *Psychol. Sci.* **27**, 685–691. <https://doi.org/10.1177/0956797616634068> (2016).
30. Birnie, L. H. W., Smallwood, J., Reay, J. & Riby, L. M. Glucose and the wandering mind: Not paying attention or simply out of fuel? *Psychopharmacology* **232**, 2903–2910. <https://doi.org/10.1007/s00213-015-3926-x> (2015).
31. Brown, K. W. & Ryan, R. M. The benefits of being present: Mindfulness and its role in psychological well-being. *J. Pers. Soc. Psychol.* **84**, 27. <https://doi.org/10.1037/0022-3514.84.4.822> (2003).
32. Park, T., Reilly-Spong, M. & Gross, C. R. Mindfulness: A systematic review of instruments to measure an emergent patient-reported outcome (PRO). *Qual. Life Res.* **22**, 2639–2659. <https://doi.org/10.1007/s11136-013-0395-8> (2013).
33. Feruglio, S., Matiz, A., Pagnoni, G., Fabbro, F. & Crescentini, C. The Impact of Mindfulness Meditation on the Wandering Mind: A Systematic Review. *Neurosci. Biobehav. Rev.* **131**, 313–330. <https://doi.org/10.1016/j.neubiorev.2021.09.032> (2021).
34. Giannandrea, A. *et al.* Effects of the Mindfulness-Based Stress Reduction Program on Mind Wandering and Dispositional Mindfulness Facets. *Mindfulness*. **10**(1), 185–195. <https://doi.org/10.1007/s12671-018-1070-5> (2019).
35. Bennike, I. H., Wieghorst, A. & Kirk, U. Online-based Mindfulness Training Reduces Behavioral Markers of Mind Wandering. *J. Cogn. Enhanc.* **1**(2), 172–181. <https://doi.org/10.1007/s41465-017-0020-9> (2017).
36. Morrison, A. B., Goolsarran, M., Rogers, S. L. & Jha, A. P. Taming a wandering attention: Short-form mindfulness training in student cohorts. *Front. Hum. Neurosci.* **7**, 897. <https://doi.org/10.3389/fnhum.2013.00897> (2014).
37. Rahl, H. A., Lindsay, E. K., Pacilio, L. E., Brown, K. W. & Creswell, J. D. Brief mindfulness meditation training reduces mind wandering: The critical role of acceptance. *Emotion* **17**(2), 224–230. <https://doi.org/10.1037/emo0000250> (2017).
38. Cantone, D., Feruglio, S., Crescentini, C., Cinot, S. & Matiz, A. A multilevel approach to explore the wandering mind and its connections with mindfulness and personality. *Behav. Sci.* **11**(9), 125. <https://doi.org/10.3390/bs11090125> (2021).
39. Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B. & Schooler, J. W. Mindfulness Training Improves Working Memory Capacity and GRE Performance While Reducing Mind Wandering. *Psychol. Sci.* **24**(5), 776–781. <https://doi.org/10.1177/0956797612459659> (2013).
40. Parsons, S., Kruijt, A.-W. & Fox, E. Psychological science needs a standard practice of reporting the reliability of cognitive-behavioral measurements. *Adv. Methods Pract. Psychol. Sci.* **2**, 378–395. <https://doi.org/10.1177/2515245919879695> (2019).
41. Sofuoglu, M., Waters, A. J., Mooney, M. & Kosten, T. Riluzole and d-amphetamine interactions in humans. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **32**, 16–22. <https://doi.org/10.1016/j.pnpbp.2007.05.003> (2008).
42. O'Connell, R. G. *et al.* Two types of action error: Electrophysiological evidence for separable inhibitory and sustained attention neural mechanisms producing error on go/no-go tasks. *J. Cogn. Neurosci.* **21**, 93–104. <https://doi.org/10.1162/jocn.2009.21008> (2009).
43. McVay, J. C. & Kane, M. J. Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *J. Exp. Psychol. Learn. Mem. Cogn.* **35**, 196. <https://doi.org/10.1037/a0014104> (2009).
44. Unsworth, N. & McMillan, B. D. Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta Physiol. (Oxf.)* **150**, 14–25. <https://doi.org/10.1016/j.actpsy.2014.04.001> (2014).
45. Kane, M. J. *et al.* Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *J. Exp. Psychol. Gen.* **145**, 1017. <https://doi.org/10.1037/xge0000184> (2016).
46. Smilek, D., Carriere, J. S. A. & Cheyne, J. A. Failures of sustained attention in life, lab, and brain: Ecological validity of the SART. *Neuropsychologia* **48**, 2564–2570. <https://doi.org/10.1016/j.neuropsychologia.2010.05.002> (2010).
47. Cheyne, J. A., Carriere, J. S. & Smilek, D. Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Conscious. Cogn.* **15**, 578–592. <https://doi.org/10.1016/j.concog.2005.11.009> (2006).
48. MacKillop, J. & Anderson, E. J. Further psychometric validation of the mindful attention awareness scale (MAAS). *J. Psychopathol. Behav. Assess.* **29**, 289–293. <https://doi.org/10.1007/s10862-007-9045-1> (2007).
49. Michalak, J., Heidenreich, T., Ströhle, G. & Nachtigall, C. Die deutsche Version der Mindful Attention and Awareness Scale (MAAS) Psychometrische Befunde zu einem Achtsamkeitsfragebogen. *Z. Klin. Psychol. Psychother.* **37**, 200–208. <https://doi.org/10.1026/1616-3443.37.3.200> (2008).
50. Medvedev, O. N. *et al.* Measuring trait mindfulness: How to improve the precision of the mindful attention awareness scale using a Rasch model. *Mindfulness* **7**, 384–395. <https://doi.org/10.1007/s12671-015-0454-z> (2016).

51. Lewis-Beck, M., Bryman, A. & Futing Liao, T. The SAGE encyclopedia of social science research. *Methods*. <https://doi.org/10.4135/9781412950589NV-0> (2004).
52. Hooper, D., Coughlan, J. & Mullen, M. R. Structural equation modelling: Guidelines for determining model fit. *Electron. J. Bus. Res. Methods* **6**, 53–60. <https://doi.org/10.21427/D79B73> (2008)
53. Kline, R. B. *Principles and Practice of Structural Equation Modeling* (Guilford publications, 2015).
54. Chen F, Curran, P. J., Bollen, K. A., Kirby, J. & Paxton, P. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociol. Methods Res.* **36**, 462–494. <https://doi.org/10.1177/0049124108314720> (2008).
55. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2020).
56. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, New York, 2016).
57. Parsons, S. splithalf—robust estimates of split half reliability. <https://doi.org/10.6084/m9.figshare.5559175.v5> (2020).
58. Schloerke, B. *et al.* GGally: Extension to ggplot2. R package version 2.1.2. <https://CRAN.R-project.org/package=GGally> (2021).
59. Tavakol, M. & Dennick, R. Making sense of Cronbach's alpha. *Int. J. Med. Educ.* **2**, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd> (2011).
60. Ward, P., Clark, T., Zabriskie, R. & Morris, T. Paper/pencil versus online data collection: An exploratory study. *J. Leis. Res.* **46**, 84–105. <https://doi.org/10.1080/00222216.2014.11950314> (2014).
61. Chmielewski, M. & Kucker, S. C. An M-Turk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science* **11**, 464–473. <https://doi.org/10.1177/1948550619875149> (2020).
62. Malkovsky, E., Merrifield, C., Goldberg, Y. & Danckert, J. Exploring the relationship between boredom and sustained attention. *Exp. Brain Res.* **221**, 59–67. <https://doi.org/10.1007/s00221-012-3147-z> (2012).
63. Martarelli, C. S., Bertrams, A. & Wolff, W. A personality trait-based network of boredom, spontaneous and deliberate mind-wandering. *Assessment* **28**, 1915–1931. <https://doi.org/10.1177/1073191120936336> (2021).
64. Konu, D. *et al.* Exploring patterns of ongoing thought under naturalistic and conventional task-based conditions. *Conscious. Cogn.* **93**, 103139. <https://doi.org/10.1016/j.concog.2021.103139> (2021).
65. Sormaz, M. *et al.* Default mode network can support the level of detail in experience during active task states. *Proc. Natl. Acad. Sci.* **115**, E11198–E11198. <https://doi.org/10.1073/pnas.1817966115> (2018).
66. Konu, D. *et al.* A role for the ventromedial prefrontal cortex in self-generated episodic social cognition. *Neuroimage* **218**, 116977. <https://doi.org/10.1016/j.neuroimage.2020.116977> (2020).
67. Smallwood, J. *et al.* The neural correlates of ongoing conscious thought. *iScience* **24**, 102132. <https://doi.org/10.1016/j.isci.2021.102132> (2021).
68. Turnbull, A. *et al.* Left dorsolateral prefrontal cortex supports context-dependent prioritisation of off-task thought. *Nat. Commun.* **10**, 3816. <https://doi.org/10.1038/s41467-019-11764-y> (2019).
69. Grossman, P. On measuring mindfulness in psychosomatic and psychological research. *J. Psychosom. Res.* **64**, 405–408. <https://doi.org/10.1016/j.jpsychores.2008.02.001> (2008).
70. Isbel, B., Stefanidis, K. & Summers, M. J. Assessing mindfulness: Experimental support for the discriminant validity of breath counting as a measure of mindfulness but not self-report questionnaires. *Psychol. Assess.* **32**, 1184–1190. <https://doi.org/10.1037/pas0000957> (2020).
71. Robison, M. K., Miller, A. L. & Unsworth, N. A multi-faceted approach to understanding individual differences in mind-wandering. *Cognition* **198**, 104078. <https://doi.org/10.1016/j.cognition.2019.104078> (2020).
72. Chaieb, L., Wilpert, E. C., Reber, T. P. & Fell, J. Auditory beat stimulation and its effects on cognition and mood states. <https://doi.org/10.3389/fpsy.2015.00070> (2015).
73. Michalak, J., Heidenreich, T., Ströhle, G. & Nachtigall, C. MAAS: Mindful attention and awareness scale—deutsche version [Verfahrensdokumentation aus PSYNDEX Tests-Nr. 9006040 und Fragebogen]. In *Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID), Elektronisches Testarchiv*, <https://doi.org/10.23668/psycharchives.393> (Trier: ZPID, 2011).
74. Runco, M. A. & Acar, S. Divergent thinking as an indicator of creative potential. *Creat. Res. J.* **24**, 66–75. <https://doi.org/10.1080/10400419.2012.652929> (2012).
75. Henninger, F., Shevchenko, Y., Mertens, U., Kieslich, P. J. & Hilbig, B. E. lab.js—a free, open, online experiment builder. <https://doi.org/10.5281/zenodo.3767907> (2020).
76. Wickham, H. *et al.* Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686. <https://doi.org/10.21105/joss.01686> (2019).
77. Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-4. <https://CRAN.R-project.org/package=e1071> (2020).
78. Grolemond, G. & Wickham, H. Dates and times made easy with lubridate. *J. Stat. Softw.* **40**, 1–25. <https://doi.org/10.18637/jss.v040.i03> (2011).
79. Rosseel, Y. lavaan: An R package for structural equation modeling. *J. Stat. Softw.* **48**, 1–36. <https://doi.org/10.18637/jss.v048.i02> (2012).
80. Hlavac, M. Stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2.2. <https://CRAN.R-project.org/package=stargazer> (2018).

Acknowledgements

The authors would like to thank all students of the following two classes at the UniDistance Suisse, who recruited participants for the experiment: “Methoden III: Experimentelle Übungen” during the fall semester 2019, “Wissenschaftliches Arbeiten” during the spring semester 2020.

Author contributions

Using the CRediT contributor roles taxonomy (casrai.org/credit/). A.B.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing—original draft, Writing—review & editing. L.C.: Conceptualization, Resources, Writing—review & editing. A.R.-M.: Conceptualization, Writing—review & editing. F.M.: Resources, Writing—review & editing. N.R.: Resources, Writing—review & editing. J.F.: Conceptualization, Writing—review & editing. T.P.R.: Conceptualization, Resources, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Supervision, Writing—original draft, Writing—review & editing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-11594-x>.

Correspondence and requests for materials should be addressed to A.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022