

Running head: Testing the Attention-Distractibility Trait

Testing the Attention-Distractibility Trait

Matt E. Meier

Western Carolina University

Please address correspondence to:

Matt E. Meier

Western Carolina University

Department of Psychology

Room 302I, Killian Bldg.

Western Carolina University

Cullowhee, North Carolina 28723

mmeier@wcu.edu

### Abstract

Forster and Lavie (2014, 2016) found that task-irrelevant distraction correlated positively with a measure of mind-wandering and a report of Attention Deficit Hyperactivity Disorder (ADHD) symptomology. Based primarily on these results, Forster and Lavie claimed to establish an attention-distractibility trait. Here, I tested if these associations could be distinguished from associations with working memory capacity and task-relevant distraction (measured with an antisaccade task). With data collected from two hundred and twenty-six subjects (*ns* differ among analyses), the results from the current study suggest that the measures of task-irrelevant distraction and working memory capacity were not (or only very weakly) associated with measures of mind wandering (measured both with a standalone questionnaire and in-task thought probes) and ADHD symptomology. Task-relevant interference (i.e., antisaccade accuracy) was associated with mind-wandering reports from in-task thought probes (presented in a separate task) but not the standalone mind wandering questionnaire or ADHD symptomology.

Additionally, the measure of irrelevant-distraction exhibited low internal consistency suggesting that (as measured) it may not be a suitable individual difference measure. [Preregistration, data, analysis scripts and output are available via the Open Science Framework: <https://osf.io/bhs24/>].

*Keywords: attention, distraction, individual differences, cognitive trait, replication*

Forster and Lavie (2014, 2016) provided evidence that interference from task-irrelevant distractors relates to mind wandering propensity and attention-deficit/hyperactivity disorder (ADHD) symptoms. Based on this evidence, Forster and Lavie (2016) claimed to establish an attention-distractibility trait. However, another construct, working memory capacity also associates with performance on many attentional tasks and measures (e.g., Colom et al., 2008; Heitz & Engle, 2007; Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2003; Meier & Kane, 2013; Meier, Smeekens, Silvia, Kwapil, & Kane, 2018; Redick & Engle, 2006; Unsworth, Schrock, & Engle, 2004; including mind wandering; e.g., McVay & Kane, 2009, 2012a, 2012b; Robison & Unsworth, 2018; Unsworth & McMillan, 2013, 2014) suggesting that working memory capacity is also an attention-distractibility trait. Here, I sought to replicate Forster and Lavie's reported associations with irrelevant distractor interference (with mind wandering and ADHD symptoms) on the way to testing if these associations are distinct from those found among working memory capacity, task-relevant distraction, mind wandering, and ADHD symptomology.

In their Experiment 1, Forster and Lavie (2014) had 94 subjects complete a task where the goal was to indicate if an *X* or an *N* appeared amongst a circular array. On 80% of the trials, a letter appeared outside of the array. This letter was either response congruent (e.g., the letter *X* when an *X* was target) or response incongruent (e.g., the letter *N* when an *X* was target). This manipulation provided task-relevant interference. On 10% of trials, a task-irrelevant cartoon character (a static image) appeared outside of the array creating task-irrelevant interference. Interference of both types (i.e., task-relevant and task-irrelevant) was operationalized as a difference score between trials where distraction was present and trials where there was no distraction. After finishing this task, subjects completed a mind wandering self-report

questionnaire (the Daydreaming Frequency Scale of the Imaginal Processes Inventory; Singer & Antrobus, 1970).

Mind wandering scores significantly positively correlated ( $r = .26$ ) with task-irrelevant interference but nonsignificantly negatively correlated with task-relevant interference ( $r = -.13$ ). Although not reported in Forster and Lavie (2014), these correlations significantly differed from each other ( $t = 2.8, p = .01$ ). From this, Forster and Lavie claimed support for their hypothesis that “mind wandering propensity is driven by overall differences in susceptibility to irrelevant distraction”. In their Experiment 3, Forster and Lavie provided more support for this claim by replicating these associations with 40 subjects and a different task (this time experimentally controlling for the novelty, saliency, visual complexity, and semantic meaningfulness of distractors). Task-irrelevant distraction was again significantly positively associated with questionnaire-measured mind wandering ( $r = .38$ ) and task-relevant distraction was again nonsignificantly negatively associated with mind wandering ( $r = -.12$ ). These correlations were also significantly different from one another,  $t = 2.9, p = .01$ .

Building on their previous work, Forster and Lavie (2016; Experiment 1) had 93 subjects complete a task where subjects again determined if an *X* or an *N* appeared in a circular array. This time there were two blocked conditions. One condition was high perceptual load (hereafter referred to as high-load) where the other (nontarget) positions on the circular array were occupied by capital letters. The other condition was low perceptual load (hereafter referred to as low-load) where like in Forster and Lavie (2014), the nontarget items were small *O*s. A task-irrelevant cartoon character distractor appeared outside of the circular array on 25% of the trials in both conditions. After completion of this task, subjects completed a self-report attention deficit hyperactivity disorder (ADHD) self-report questionnaire (Childhood Symptoms Scale—

Self-Report Form; Barkley & Murphy, 1998) on childhood symptoms (the subjects were students from a college course [age range 18-39]).

Critically, in the low-load condition, irrelevant-distractor interference (i.e., the response time difference between trials with the distractor present and trials with the distractor absent) was significantly positively correlated with ADHD symptoms ( $r = .32$ ; this analysis used data from 75 subjects because of task-performance exclusions). Experiment 2 in this paper produced the exact same estimate of the correlation between irrelevant-distractor interference and ADHD symptoms ( $r = .32$ ;  $N = 74$ ). To isolate task irrelevance as the association's determinant, Forster and Lavie included task-relevant distractors that matched the task-irrelevant distractors in salience, familiarity, meaning, and presentation frequency. The task-relevant distractors still were not significantly correlated with ADHD symptoms ( $r = .03$ ), suggesting that task relevance was indeed the crucial difference between types of distractors and their relation to ADHD symptoms.

Taken together, Forster and Lavie (2014; 2016) provided evidence that subjects who experience greater task-irrelevant distraction are more likely to report having a high propensity for mind wandering and experiencing more ADHD symptoms as children. Forster and Lavie (2016) speculated that susceptibility to irrelevant distraction (i.e., the attention-distractibility trait) may be the “common determinant” of both mind wandering propensity and ADHD. Forster and Lavie provided some evidence that task-irrelevance is responsible for this association by experimentally controlling for some characteristics that were different between task-relevant and task-irrelevant distraction in their tasks. But this is not enough evidence to be confident in the establishment of a *new* trait. More pointedly, although promising, Forster and Lavie’s evidence is limited from a construct validation perspective (Chronbach & Meehl, 1955). Forster and Lavie provide convergent validity with task-irrelevant distraction correlations with only one ADHD

questionnaire and one daydreaming questionnaire, and divergent validity primarily comes from their findings of statistically significant associations between task-irrelevant distraction, mind wandering, and ADHD symptoms. In contrast, task-relevant distraction was not found to associate (with mind wandering, and ADHD symptoms).

Considering that working memory capacity is a trait-like cognitive construct (Meier & Kane, 2017) that predicts performance on a wide array of cognitive tasks and has been tied explicitly to ADHD (e.g., Alderson, Kasper, Hudec, & Patros, 2013; Kasper, Alderson, Hudec, 2012; Kofler et al., 2019) and mind wandering (e.g., McVay & Kane, 2009; Unsworth & McMillan, 2013), a new trait must be distinguished from this already existing trait. Also, the initial evidence provided by Forster and Lavie requires buttressing by independent investigators to increase credence in the claim of a new trait. Moreover, because Forster and Lavie used sample sizes demonstrated to provide imprecise estimates of association (Schönbrodt & Perugini, 2013), future investigations must use larger sample sizes.

### **The Current Study**

The goals of this study were: 1) to replicate Forster and Lavie's associations between irrelevant distractor interference and mind wandering, and between irrelevant distractor interference and ADHD in a single study with a larger sample size; 2) to test if the association between irrelevant distractor interference and mind wandering extends to task-based (rather than questionnaire) measures of mind wandering; 3) to test the divergent validity of irrelevant distractor interference by including working memory capacity measures and a measure of attention that includes only task-relevant interference; 4) to test if irrelevant distractor interference is a common determinant of ADHD symptomology and mind wandering. To this end, I had a relatively large sample of subjects (compared to Forster & Lavie, 2014, 2016)

complete the irrelevant distractor task, two working memory capacity tasks, an antisaccade task, a sustained attention to response task (SART) with embedded thought probes, an ADHD questionnaire, and a mind wandering questionnaire.

### **Method**

I report how I determined the sample size, all data exclusions, all manipulations, and all measures in the study (Simmons, Nelson, & Simonsohn, 2011). This study was preregistered on February 22, 2018 (<https://aspredicted.org/ng4ar.pdf>).

### **Subjects**

Two hundred and thirty-five subjects from Western Carolina University (WCU) completed the informed consent for this study (mean incoming student SAT scores range from 1085 to 1118 for cohorts entering Fall 2016 through Fall 2018). I collected demographic data from 233 of these subjects (data from 2 subjects [of the 235 who completed the informed consent] were lost because of technical errors). Of these 233 subjects, 65% were female. Subjects had a mean age of 19 ( $SD = 1$ ); one subject reported an age of 1999, I inferred that this subject reported their birth year and converted this value to 19. Of the subjects who gave ethnicity information (3 subjects declined), 187 identified as white, 17 as black, 13 as multiracial, 5 as other, and 4 as Asian. Subjects received partial credit for a course requirement as compensation for their participation. The stopping rule for data collection was the end of the semester in which data was collected from at least 220 subjects. This sample size was chosen on the basis that correlations as weak as  $\rho = .10$  stabilize within a narrow window when approaching 250 subjects (Schönbrodt & Perugini, 2013) thus allowing precise estimates. Data collection terminated at the end of the Fall 2018 semester.

### **General Procedure**

Subjects completed one two-hour session in groups of up to 2, with the following task order: Irrelevant Distractor task, IPI daydreaming questionnaire, operation span, antisaccade, ADHD questionnaire, SART, symmetry span and then a demographic questionnaire. Experimenters read all on-screen instructions aloud while subjects read along silently. An experimenter was always present in the room. Subjects were tested in a room alone or with one other subject. The testing room was small and windowless except for a small window in the door that faces a hallway. For sessions with two subjects, one subject sat at a desk facing a wall opposite the door and the other subject sat at a desk facing a wall 90 degrees to the left of the desk facing opposite the door (in sessions with only one subject the subject sat at the desk facing to the left). The experimenter was inside the room to the door's immediate right side (not in view of the subjects when they were performing tasks). There was a white noise machine that was on for the entire session blocking potentially distracting noises. All tasks were programmed and administered with E-prime software (Psychology Software Tools, Pittsburg, PA). The session protocol is available at <https://osf.io/bhs24/>.

### **Irrelevant Distractor Task**

This task was provided by the lead author of Forster and Lavie (2014, 2016), and is the same task as that used in Forster and Lavie (2016) Experiment 1. Here, the task was presented on a 22-inch widescreen liquid crystal display monitor. This monitor's aspect ratio was set to 4:3 (making the display area more comparable to the non-widescreen monitors used in Forster and Lavie [2014, 2016]). The goal of the task is for the subjects to quickly and accurately press the “0” key if they see an “X” or press the “2” key if they see an “N” among a circular array of six letters. These target letters (i.e., “X” and “N”) were presented within two blocked conditions of the task. In the low-load condition, the target letter is presented among lower case *os* and in the



high-load condition these target letters are presented among a subset of the following letters selected at random: *K*, *V*, *W*, *Z*, *M*, and *H*. The eight blocks' order was counter-balanced between subjects using the 2 following orders: ABBAABBA or BAABBAAB.

The radius of the circular array was approximately 1.9 cm. The target and distractor letters in the high load condition measured approximately  $.75 \times .91$  cm. The lower case *os* were approximately  $.32 \times .32$  cm. Tape was stuck to the floor 60 cm from the screen on each side of each subject's chair. Experimenters instructed subjects that their eyes should stay at this distance from the screen throughout the task.

A 500 ms screen with a fixation dot began every trial. This fixation was followed by the presentation of the array for 100 ms. If the subject responded erroneously or failed to respond within a 2000 ms window, the task played a tone (through headphones). Each block contained 48 trials. Before beginning the 48-trial blocks, subjects completed 3 practice trials each for the high and low-load conditions followed by a block of 12 practice trials in the high and low-load conditions where to proceed into the 48-trial blocks subjects had to achieve 65% accuracy. If subjects did not achieve 65% accuracy, they were presented with the instructions again and then given another opportunity to meet the criterion. This process went on until the criterion was met (all subjects met this criterion).

On 25% of the trials (in the 48-trial blocks), a task-irrelevant distractor was presented either above or below the circular array. The task-irrelevant distractor was a color-image of one of the following cartoon characters: Donald Duck, Mickey Mouse, Pikachu, Spider-Man, SpongeBob SquarePants, and Superman. The distractors were approximately 3.2-4.6 cm wide and 4.6-5.7 cm tall (the dimensions differed for different characters). These characters appeared 4.5 cm from fixation and at least 1.5 cm outside the nearest item in the array. These distractors

remained on-screen until a response was made or the trial was terminated due to time. The dependent variables from this task were response times (RT) and accuracy. More specifically, the difference in RT and accuracy between trials with and without the irrelevant distractor were computed for each load condition. As in Forster and Lavie (2014, 2016) the first three trials of the 48-trial blocks were not included in these calculations and RTs were only taken from correct trials.

**Imaginal Processes Inventory Daydreaming Subscale (IPI daydreaming scale; Kane et al., 2020 modified from Singer & Antrobus, 1970)**

On 12 items, subjects indicated how strongly they agreed or disagreed with statements related to their self-perceived propensity to daydream on a 5-point Likert-type scale (1 = Strongly disagree; 5 = Strongly agree). The original scale used frequencies for response options. Here, items were revised to fit an agree-disagree scale. For example, subjects responded to the item “When I am not paying close attention to my job, a book, or TV, I tend to be daydreaming”. This revision produced ten regularly-scored items and two reverse-scored items. In addition to the 12 IPI-DS items, subjects completed four interleaved infrequency items (e.g., “I enjoy the music of Marlene Sandersfield”). Consistent with the preregistration, these infrequency items were not used in the analyses presented here<sup>1</sup>. The dependent variable was the sum (after accounting for reverse scoring) of the 12 IPI-DS item responses.

**ADHD Rating Scale-IV (Dupaul, Bower, Anastopoulos, & Reid, 1998)**

Subjects completed a scale of 16 items by clicking on a response that best described their behavior when they were a child age 5 to 12 years. The original and intended scale is 18 items, but because of a programming error the final two questions were not administered (the erroneously not-presented items were: Interrupt or intrude on others; Forgetful in daily

activities). The four answer options were: never or rarely, sometimes, often, and very often. Subjects received a score of 0 for never or rarely, 1 for sometimes, 2 for often, and 3 for very often. Subjects' scores for all items were summed for a total score. Of the 16 items, eight correspond to the inattentive subtype and the other eight to the hyperactive subtypes. Scores for these subtypes were summed.

### **Antisaccade**

Subjects identified a target letter (*B*, *P*, or *R*) on one side of the screen that was cued by a flash on the opposite side (modified from Kane et al., 2001). Subjects first saw a central-fixation array of three asterisks over one of five fixation-cue delay intervals that unpredictably ranged from 200–1800 ms (200, 600, 1000, 1400, or 1800 ms) followed by a flashing cue (“=”) presented 8.6 cm to the left or right of fixation. The flashing cue was presented by having the cue displayed for 100 ms and then followed by a 50 ms blank screen, with the sequence happening twice on every trial (i.e., cue-blank-cue-blank). The target letter appeared in the opposite screen location from the cue (8.6 cm from fixation) and was pattern-masked after 100 ms with the letter *H* for 50 ms then the digit 8 until the subject responded. Subjects responded via keys on the number keypad labeled *B*, *P*, and *R* with stickers. Stimuli were presented in white (12-point Courier New font) on a black background. The dependent measure was accuracy on 90 test trials.

### **Sustained Attention to Response Task (SART)**

The task instructed subjects to press the space bar for words from one category (animals; 89% of trials) while withholding response to another (vegetables; 11% of trials; see Kane et al. 2016). Each of 675 trials presented a word for 300 ms, then a mask for 1500 ms. Trials were divided into five seamless blocks (i.e., these blocks are in the computer program but there is nothing to indicate these blocks to subjects), each comprising 3 miniblocks of 45 trials that

presented 40 unique animal names and 5 unique vegetable names. The dependent measures were  $d'$  (i.e., hit rate to animals minus false alarm rate to vegetables) and SD of RTs to “go” (animals) trials.

During this task, subjects responded to 45 thought probes. Nine thought probe trials were pseudo-randomly placed in each of the five blocks (of 135 trials) following target trials. Prior to beginning the task, subjects were instructed to respond based on what they were thinking just before the screen appeared and not reconstruct what they were thinking during the preceding words on the screen. The thought probe presented the question “What are you thinking about?”. The answer options were (response options as seen by subjects are italicized; they are followed by brief descriptions that were part of the task instructions): 1. *The task*, on-task thoughts about the stimuli or response; 2. *Task experience/performance*, evaluative thoughts about one’s task performance; 3. *Everyday things*, thoughts about routine things that have happened or may happen; 4. *Current state of being*, thoughts about one’s current physical or emotional state, such as being sleepy, hungry, or cheerful; 5. *Personal worries*, thoughts about one’s concerns or worries; 6. *Daydreams*, fantastic thoughts disconnected from reality; 7. *External environment*, thoughts about something task-unrelated in the immediate environment; 8. *Other*, only those thoughts that do not fit the other categories. The mind wandering dependent measure for this task was the proportion of probe responses 3 to 8.

### **Working Memory Capacity Complex Span Tasks.**

I assessed working memory capacity with two shortened complex span tasks (Foster et al., 2015; Unsworth, Heitz, Schrock, & Engle, 2005). Foster et al. provided evidence that using one block each of two complex span tasks (the contents of these blocks are explained below) is a more valid assessment of working memory capacity than the more popular approach of three

blocks of one complex span task. Specifically, they found that using one block each from an operation and symmetry span task predicts more variance in a fluid intelligence composite than a three-block operation span task alone (the most commonly used complex span task).

In these tasks, subjects memorized short sequences of items while completing an interleaved processing task. Following trial sequences of unpredictable length, subjects recalled the memorial items in serial order. Before beginning the task's scored portion, subjects practiced memorizing small sets, practiced the processing task alone, and then practiced both task components together. From the processing-only practice, an individualized response deadline was set. If on any processing-task portion of a trial, a response was not made within 2.5 standard deviations of the subject's processing-only practice RT mean, the program skipped the subsequent memory stimulus and recorded a processing error. All subjects were instructed that if they did not achieve 85% accuracy on the processing portion of the task their data would not be used in analyses, but as per the preregistration this processing criterion was not enforced.

**Operation Span.** Subjects memorized sequences of 3–7 letters (each letter was presented for 1 s). These letters appeared in alternation with an arithmetic equation to verify [e.g.,  $(3 \times 2) - 1 = 4$ ; half were true]. At recall, all 12 letters (used in the task, but not necessarily in any one specific trial) appeared in a grid; subjects selected recalled letters with a computer mouse click. Each set length of 3–7 occurred once in a random order for each subject. The variable used in analyses was the total number of letters recalled in correct serial position (of 25; i.e., partial-credit scoring [Conway et al., 2005]).

**Symmetry Span.** Subjects memorized sequences of 2–5 red squares appearing within a matrix. Each red square appeared (for 650 ms) in alternation with a black-and-white pattern made from an  $8 \times 8$  grid to verify if it was vertically symmetrical (half were symmetrical). At

recall, subjects saw an empty  $4 \times 4$  matrix and mouse-clicked the red square locations. When a square was clicked upon, a number indicating the serial position of the square's presentation appeared in that square. Each set length of 2–5 occurred three times in a random order for each subject. Each subject's score was the total number of red-square locations recalled in correct serial position (of 14).

### **Data Analysis**

Any analyses performed in this manuscript that are not included in the preregistration were made in response to the data and should be judged as such. I performed analyses in the R system for statistical analysis (R Core Team, 2019). Data, analysis code, and outputs are available at the following link: <https://osf.io/bhs24/>.

### **Data Loss**

I made all data exclusions in accordance with the preregistration<sup>2</sup>. Because of computer or experimenter error, I am missing data from two subjects in the irrelevant distractor task, one subject in the SART task, and one subject in the symmetry span. One subject did not meet the accuracy criterion in irrelevant distractor task (the subject's accuracy was below 55%). I dropped all data for six subjects who were deemed by experimenters as noncompliant with instructions across tasks. These decisions were made without consulting the subjects' data. Thus, I included data from 226 subjects from the Irrelevant Distractor, ADHD-IV, IPI-DS, antisaccade, and operation span tasks and data from 225 subjects for the SART and symmetry span in the analyses. In the irrelevant distractor task, all trials that had RTs faster than 100 ms and slower than 1500 ms were removed. This resulted in a loss of 2.3% of trials. Descriptive statistics for all dependent variables can be seen in Table 1. Intercorrelations for measures can be seen in Table 2.

**Results**

Table 1. Descriptive Statistics

Measure	M	SD	Min	Max	Skew	Kurtosis	N
Low Load RT Interference	18.52	27.93	-68.56	116.38	0.46	1.66	226
High Load RT Interference	-12.28	44.05	-160.49	104.68	-0.46	0.85	226
Low Load Accuracy Interference	0.00	0.06	-0.16	0.15	-0.09	-0.16	226
High Load Accuracy Interference	0.02	0.07	-0.16	0.23	0.09	0.07	226
ADHD Total	18.33	9.29	3	48	0.65	0.00	226
ADHD Inattention	8.50	4.95	1	24	0.87	0.32	226
ADHD Hyperactivity	9.83	5.16	0	24	0.34	-0.63	226
IPI Daydreaming	40.16	8.37	21	60	0.04	-0.48	226
SART d'	2.17	0.89	-0.44	4.28	-0.21	-0.13	225
SART Nontarget RT SD	191.81	57.31	88.42	424.85	1.06	2.00	225
SART Mind Wandering	0.54	0.25	0	1	-0.17	-0.82	225
Antisaccade Accuracy	0.59	0.15	0.23	0.93	0.11	-0.71	226
Shortened Operation Span	15.23	5.94	0	25	-0.35	-0.52	226
Shortened Symmetry Span	8.35	3.28	0	14	-0.40	-0.36	225
Working Memory Capacity	0.00	0.82	-2.17	1.69	-0.40	-0.47	225

*Note.* RT = response time; IPI-Daydreaming = Imaginal Processes Inventory Daydreaming Subscale; ADHD = attention deficit hyperactivity disorder; SART = sustained attention to response task; Interference measures are raw difference scores.

Table 2. Correlations Among Measures (With Reliability Point-estimates for Key Measures on the Diagonal)

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Low Load RT Interference	.26														
2. High Load RT Interference	0.10	.05													
3. Low Load Accuracy Interference	-0.02	0.13	-												
4. High Load Accuracy Interference	0.16	0.00	0.10	-											
5. ADHD Total	-0.08	0.09	0.01	0.01	.90										
6. ADHD Inattentive	-0.02	0.08	0.06	0.02	0.92	.85									
7. ADHD Hyperactivity	-0.13	0.08	-0.04	0.00	0.92	0.69	.81								
8. IPI Daydreaming	-0.13	0.05	0.10	-0.06	0.37	0.41	0.27	.72							
9. SART d'	0.08	0.07	0.06	0.08	-0.19	-0.21	-0.14	-0.23	.94						
10. SART Nontarget RT SD	-0.07	-0.08	0.05	-0.06	0.20	0.21	0.16	0.09	-0.52	.96					
11. SART Mind Wandering	-0.05	0.02	0.00	-0.11	0.26	0.25	0.23	0.26	-0.18	0.37	.92				
12. Antisaccade Accuracy	-0.01	0.00	0.04	0.03	-0.05	-0.08	-0.02	-0.11	0.35	-0.37	-0.17	.89			
13. Operation Span	0.05	0.05	-0.03	0.06	-0.01	-0.09	0.06	-0.08	0.02	0.01	-0.02	0.18	.72		
14. Symmetry Span	0.02	0.08	-0.10	0.11	-0.02	-0.04	0.01	-0.06	-0.02	-0.10	-0.01	0.24	0.33	.62	
15. Working Memory Capacity	0.05	0.08	-0.08	0.11	-0.02	-0.08	0.05	-0.08	0.00	-0.05	-0.02	0.26	0.82	0.82	-

*Note.* RT = response time; IPI-Daydreaming = Imaginal Processes Inventory Daydreaming Subscale; ADHD = attention deficit hyperactivity disorder; SART = sustained attention to response task; Interference measures are raw difference scores.



## Experimental Effects of Irrelevant Distractor Task

As can be seen in Table 3, the descriptive statistics of the experimental effects from the current study are strikingly close to the values reported by Forster and Lavie (2016).

Table 3. Mean Response Times, Error Rates (Standard Errors in Parentheses), and Difference Scores in the Irrelevant-Distractor and No-Distractor Conditions

Load Condition and Measure	<u>Forster and Lavie, 2016</u>			<u>Current Study</u>		
	<u>Distractor Condition</u>		Difference	<u>Distractor Condition</u>		Difference
	Irrelevant distractor	No distractor		Irrelevant distractor	No distractor	
Low Load						
Response Time	516 ms (6)	491 ms (6)	25 ms	508 ms (5)	490 ms (4)	19 ms
Error	14%	13%		13%	12%	
High Load						
Response Time	692 ms (13)	699 ms (14)	-7 ms	703 ms (8)	715 ms (8)	-12 ms
Error	29%	26%		27%	24%	

**Response Times.** Following from Forster & Lavie (2016), I conducted a repeated-measures analysis of variance (ANOVA) with factors of load condition (low versus high) and distractor (present versus absent) predicting response time (RT). Like Forster and Lavie, the main effect of condition was statistically significant,  $F(1, 225) = 1355.60$ ,  $MSE = 0.79$ ,  $p < .001$ ,  $\eta^2_p = .86$ , as was the most critical interaction between load and distractor conditions,  $F(1, 225) = 86.67$ ,  $MSE = 0.67$ ,  $p < .001$ ,  $\eta^2_p = .28$ . Unlike Forster and Lavie (2015) the main effect of distractor was not statistically significant,  $F(1, 225) = 2.96$ ,  $MSE = 741.71$ ,  $p = .086$ ,  $\eta^2_p = .01$ .

**Accuracy.** Although accuracy was not the focus of this study, to best understand performance on the irrelevant distractor task, accuracy (in addition to RTs) requires consideration. I conducted a parallel analysis to the repeated-measures ANOVA on RTs, but the dependent variable was accuracy. Like Forster and Lavie (2016), the main effect of condition (high vs. low load) was statistically significant,  $F(1, 225) = 399.06$ ,  $MSE = 0.010$ ,  $p < .001$ ,  $\eta^2_p = .64$ , as was the main effect of distractor (present vs. absent),  $F(1, 225) = 14.56$ ,  $MSE = 0.002$ ,  $p < .001$ ,  $\eta^2_p = .06$ . In contrast to Lavie and Forster (but still theoretically consistent with their previous claims), the interaction between condition and distractor presence was also statistically significant,  $F(1, 225) = 13.66$ ,  $MSE = 0.002$ ,  $p < .001$ ,  $\eta^2_p = .06$  with accuracy favoring distractor absent trials in the high load condition and no difference between distractor absent or present trials in the low load condition.

### **Individual Differences**

Bayes factors (BFs) for correlations were computed with the BayesFactor package (Morey & Rouder, 2018). The BF allowed us to assess if the correlation estimate is more likely from a point-null distribution (i.e., null hypothesis) or a Cauchy distribution where 50% of the distribution lies between  $-.33$  and  $.33$  (i.e., the alternative hypothesis). This specification of the alternative model was chosen because it best represented the magnitude of the estimates provided by Forster and Lavie (2014, 2016). Numbers greater than one supported the alternative hypothesis, and numbers less than one supported the null hypothesis (of no association). In these correlations, I used raw RT difference scores in the correlations. In the supplemental materials, I also conducted correlations with normalized interference scores (normalized by dividing the difference score by irrelevant distractor absent trials, as was done in Forster and Lavie, 2016)<sup>3</sup>. As can be seen in Figure 1, inconsistent with Forster and Lavie (2016; Experiment 1,  $r(75) = .32$ ,

$p = .004$ ; Experiment 3,  $r(72) = .32$ ,  $p = .005$ ), low-load irrelevant distraction was not associated with ADHD total,  $r(224) = -.08$ ,  $p = .204$ ,  $CI [-.21, .05]$ ,  $BF_{10} = .34$ , and ADHD inattentive,  $r(224) = -.02$ ,  $p = .76$ ,  $CI [-.15, .11]$ ,  $BF_{10} = .16$ , but was with ADHD hyperactivity,  $r(224) = -.13$ ,  $p = .045$ ,  $CI [-.26, -.002]$ ,  $BF_{10} = 1.1$ . But this estimate was negative rather than positive as in Forster and Lavie (2016). Correlations with the normalized differences scores resulted in slight differences in the estimate of association (i.e., low-load irrelevant distraction  $\times$  ADHD total,  $r = -.06$ ; irrelevant distraction  $\times$  ADHD hyperactivity,  $r = -.11$ ; irrelevant distraction  $\times$  ADHD inattentive,  $r = -.004$ ) but were not close to suggesting different inferences.

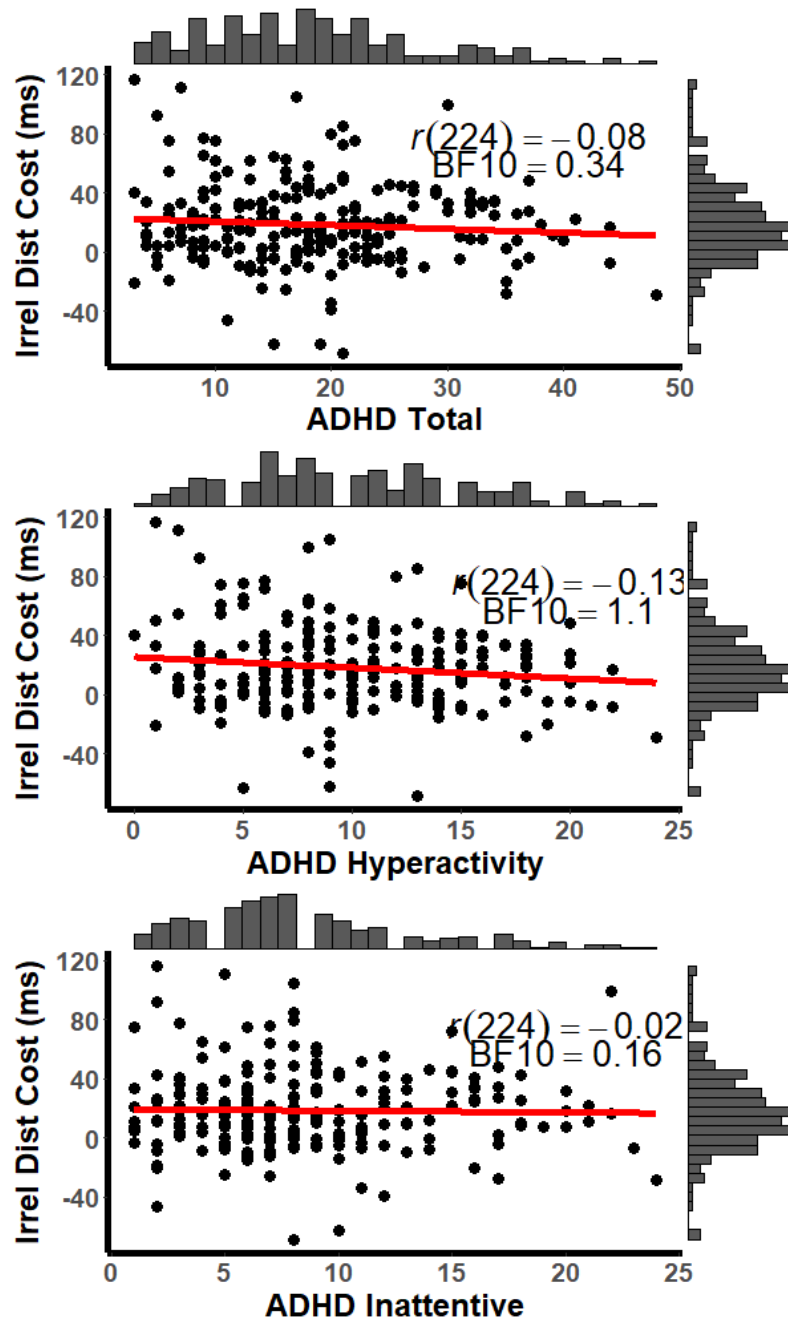


Figure 1. Scatterplots (with best-fitting regression lines in red) showing the associations between irrelevant distractor interference and three scores derived from the ADHD Rating Scale-IV (Total, Inattentive, Hyperactivity). Histograms for each variable are presented across from each axis. Note. Irrel Dist = Irrelevant Distractor;  $BF_{10}$  = Bayes Factor with numbers less than one favoring the null hypothesis and numbers greater than one favoring the alternative hypothesis.

And as can be seen in Figure 2, low-load irrelevant distraction was not associated ( $p > .05$ ) with the questionnaire measure of mind wandering and the estimate indicated a negative

association (inconsistent with the statistically significant positive association [ $r(92) = .26, p = .012$ ] reported by Forster and Lavie [2014]),  $r(224) = -.13, p = .052, CI [-.26, -.0006], BF_{10} = .99$ . The probe-caught measure of mind wandering was also not associated with low-load irrelevant distraction,  $r(223) = -.05, p = .427, CI [-.18, .08], BF_{10} = .21$ .

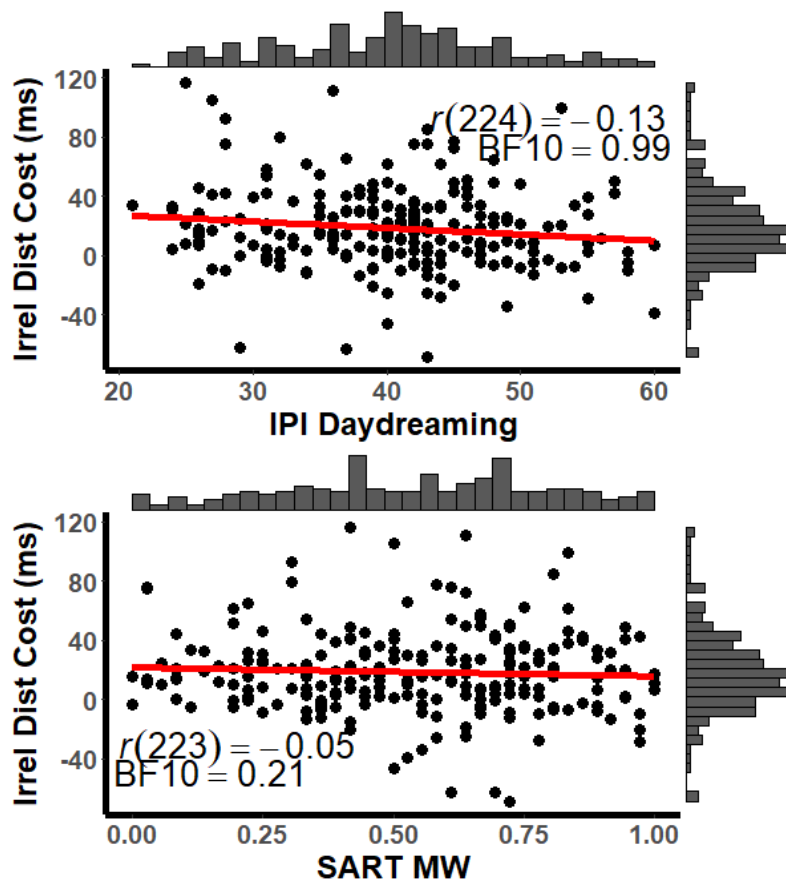


Figure 2. Scatterplots (with best-fitting regression lines in red) showing the associations between irrelevant distractor interference and mind wandering measured by the IPI daydreaming scale and by the proportion of SART probe responses scored as mind wandering. Histograms for each variable are presented across from each axis. Note. Irrel Dist = Irrelevant Distractor;  $BF_{10}$  = Bayes Factor with numbers less than one favoring the null hypothesis and numbers greater than one favoring the alternative hypothesis.

### Testing for Unique Contributions of Task-irrelevant Distraction

Although the bivariate correlations above did not provide evidence for the critical associations among irrelevant distractor interference, ADHD symptoms, and mind wandering

propensity, I still carried out the preregistered hierarchical regressions. I used these analyses to test if irrelevant distractor interference predicts unique variance in ADHD symptoms and mind wandering propensity over and above variance accounted for by working memory capacity and task-relevant interference. Working memory capacity and antisaccade accuracy were entered as predictors in the first step, and then irrelevant distractor interference was entered in the second step. The dependent variables in this series of models were the three ADHD sum scores (i.e., total, hyperactivity, and inattentive) and mind wandering rates (questionnaire and probe caught). All variables were converted to Z-scores before running the models. The models were assessed by examining R squared's change from step one to step two (i.e.,  $\Delta R^2$ ) and by computing a BF assessing if the data favors the model with or without irrelevant distractor interference.

Parameter estimates for these models can be seen in Table 4. In the models predicting Total ADHD, Inattentive ADHD, and mind wandering, none of the changes in R squared (due to the adding of the irrelevant distractor interference in the 2<sup>nd</sup> step) reached statistical significance: Total ADHD,  $\Delta R^2 = .007$ ,  $F = (1, 221) = 1.6$ ,  $p = .21$ ; Inattentive ADHD,  $\Delta R^2 = .0003$ ,  $F = (1, 221) = .06$ ,  $p = .81$ ; SART mind wandering,  $\Delta R^2 = .0003$ ,  $F = (1, 220) = .67$ ,  $p = .41$ . The BFs comparing the model with and without the irrelevant distraction (for Total ADHD, Inattentive ADHD, and mind wandering) all modestly favored the models without irrelevant distractor interference over the models with irrelevant distractor interference: Total ADHD BF = 1.8, ADHD inattentive BF = 3.6, mind wandering BF = 2.9, suggesting that the models without irrelevant distraction offer a better account of the data. For the model predicting the score on IPI-daydreaming, the change in R squared was not statistically significant,  $\Delta R^2 = .01$ ,  $F = (1, 221) = 2.7$ ,  $p = .10$ , and the BF did not discriminate between the models, BF = 1.1. The change in R squared from step one to step two of the model predicting ADHD hyperactivity was statistically

significant,  $\Delta R^2 = .02$ ,  $F = (1, 221) = 4.2$ ,  $p = .04$ , but as can be seen in Table 4, the parameter estimate for irrelevant distractor interference was negative and portrayed a relation in the opposite direction as found by Forster and Lavie (2016). The BF for this model was slightly in favor of the model with irrelevant distraction versus the model without irrelevant distraction,  $BF = 1.8$ .

Table 4. Hierarchical Multiple Regression Model Parameter Estimates

DV	Step	Predictor	<i>B</i>	t value	<i>p</i>
ADHD Total	1	WMC	-0.01	-0.06	.952
		Antisaccade Accuracy	-0.05	-0.74	.463
	2	WMC	0.00	0.00	.999
		Antisaccade Accuracy	-0.05	-0.76	.451
		Irrelevant Distractor Interference	-0.09	0.07	.207
ADHD Hyperactivity	1	WMC	0.07	0.77	.440
		Antisaccade Accuracy	-0.03	-0.45	.657
	2	WMC	0.07	0.88	.381
		Antisaccade Accuracy	-0.03	-0.48	.633
		Irrelevant Distractor Interference	-0.14	-2.06	.041
ADHD Inattentive	1	WMC	0.07	0.77	.440
		Antisaccade Accuracy	-0.03	-0.45	.657
	2	WMC	0.07	0.88	.381
		Antisaccade Accuracy	-0.03	-0.48	.633

		Irrelevant Distractor Interference	-0.14	-2.06	.041
SART Mind Wandering	1	WMC	0.03	0.32	.746
		Antisaccade Accuracy	-0.18	-2.60	.010
	2	WMC	0.03	0.36	.718
		Antisaccade Accuracy	-0.18	-2.60	.010
		Irrelevant Distractor Interference	-0.06	-0.82	.413
IPI-Daydreaming	1	WMC	-0.07	-0.86	.392
		Antisaccade Accuracy	-0.10	-1.40	.164
	2	WMC	-0.07	-0.78	.437
		Antisaccade Accuracy	-0.10	-1.43	.155
		Irrelevant Distractor Interference	-0.11	-1.64	.102

Note. ADHD = attention deficit hyperactivity disorder; SART = sustained attention to response task; IPI-Daydreaming = Imaginal Processes Inventory Daydreaming Subscale; WMC = working memory capacity composite;

In another model, I tested if irrelevant distractor interference predicts variance in an attention problem composite over and above a general (i.e., broad and task-relevant) attention factor. The attention problem composite was a composite of ADHD symptoms and mind wandering scores (each score converted to Z-scores and then averaged across the two scores) while the general attention factor was a composite of working memory capacity, antisaccade accuracy, SART  $d'$ , and SART response time standard deviation (each score converted to Z-scores and then averaged across the four scores; SART response time standard deviation was subtracted from the observed maximum value so a 0 equaled the highest value and the greater the



positive value the lesser the response time standard deviation, so higher scores represented better performance on each measure). This model was assessed the same way as the hierarchical models above. Parameter estimates for this model are in Table 5. In this model, there was not a statistically significant change in R squared when adding irrelevant distraction to the model,  $\Delta R^2 = .009$ ,  $F = (1, 221) = 2.1$ ,  $p = .15$ . The BF favored the model without irrelevant distraction over the model with irrelevant distraction by a factor of 3.3.

Table 5. Parameter Estimates from Hierarchical Multiple Regression Model Predicting Attention Problem Composite

Step	Predictor	<i>B</i>	t value	<i>p</i>
1	General Attention Factor	-0.46	-4.66	.000
2	General Attention Factor	-0.45	-4.57	.000
	Irrelevant Distractor Interference	-0.09	-1.45	.150

Note. General Attention Factor = Z-score average of working memory capacity, antisaccade accuracy, SART response time standard deviation (reverse-scored), and SART d'

### Estimating Reliability

Because I did not detect the patterns of individual differences found in Forster and Lavie (2014, 2016), I inspected the reliability of the key measures in this study (I did not preregister these reliability analyses). Where appropriate I used a permutation-based split-half correlation approach to estimate internal consistency because estimates produced this way are more stable than estimates from a more traditional split-half approach (Parsons, Kruijt, & Fox, 2019). These estimates were produced with the splithalf package (Parsons, 2019). All estimates produced in this way were constructed from 10000 permutations. Internal-consistency reliability was estimated separately for the different conditions of the Irrelevant Distractor task. The estimate for

the irrelevant distraction dependent variable from the low-load condition was .26 with a 95% *CI* of [.10, .41]. The estimate for the high-load condition was .05, 95% *CI* [-.14, .23]. The reliability estimate for antisaccade accuracy was .89, 95% *CI* [.87, .91], and the estimate for responses to the SART thought probes was .92, 95% *CI* [.90, .93].

Reliability estimates for complex span tasks were based on the percentage of correctly recalled memoranda for each trial. The internal consistency of the operation span was estimated as .72 and the symmetry span as .62. Chronbach's alpha was used as the estimate for the ADHD scores and IPI daydreaming scale. The estimates for ADHD hyperactivity, ADHD inattention, and ADHD total were .81, .85, and .90, respectively. Cronbach's alpha for the IPI daydreaming scale was .72. For SART nontarget RT SD and  $d'$ , estimates were derived from splitting the trials into two sets, calculating RT SD and  $d'$ , correlating these measures, and then applying the Spearman-Brown formula (Brown, 1910; Spearman, 1910). The estimate for nontarget RT SD was .96 and for  $d'$  was .94.

## Discussion

This study assessed the attention-distractibility trait. The fundamental individual difference findings on which the claim of a new trait was based were inconsistent between the current study and the studies of Forster and Lavie (2014, 2016). I found no evidence of a relation between how a subject performs in the face of task-irrelevant distraction and the propensity to mind wander (measured by questionnaire and within-task thought probes) or ADHD symptoms. In this study, irrelevant distraction interference (in the low-load condition) was only correlated with one variable (other than high-load RT irrelevant distractor interference) where the estimate qualified as statistically significant (ADHD hyperactivity,  $r = -.13$ ,  $p = .045$ ), and that correlation estimate was in the opposite direction from what Forster and Lavie (2014) reported. Considering

the evidence provided here along with the evidence by Forster and Lavie (2014, 2016), the claim of a new attention-distractibility trait is premature.

Recent work has highlighted the importance of psychometric properties in measuring individual differences (e.g., Cooper, Gonthier, Barch, & Braver, 2017; Hedge, Powell, & Sumner, 2018). Motivated by the lack of associations between irrelevant distractor interference and other measures in this study, I examined the key measures' internal consistency. The measure of irrelevant distractor interference used in the current study and Forster and Lavie (2014, 2016) produced very low internal consistency estimates (.26 in the critical low-load condition). This low estimate of internal consistency is the most parsimonious explanation for the lack of associations between irrelevant distractor interference and other variables in this study.

Although the validity of working memory capacity as an individual differences construct is well established (e.g., Kane et al., 2004; Meier & Kane, 2017; Unsworth & Engle, 2007), some expected correlations with the working memory capacity composite were not realized. For example, the correlation between working memory capacity and SART task mind wandering propensity was  $r = -.02$ ,  $p = .73$ ,  $CI [-.15, .11]$ ,  $BF_{10} = .16$ . Statistically significant negative zero-order correlations for this association have been reported multiple times (e.g., McVay & Kane, 2009, 2012b; Meier, 2019), but these correlations are usually small and only seem to reliably arise when latent variable techniques are used (e.g., McVay & Kane, 2012a; Robison, Gath, & Unsworth, 2017; Smeekens & Kane, 2016; Unsworth & McMillan, 2013). Additionally, working memory capacity often associates with SART task performance measures (e.g., Kane et al., 2016, McVay & Kane, 2009, 2012a, 2012b), here it did not (response time standard deviation,  $r = -.05$ ,  $p = .41$ ,  $CI [-.18, .08]$ ,  $BF_{10} = .22$ ;  $d'$ ,  $r = .005$ ,  $p = .95$ ,  $CI [-.13, .14]$ ,  $BF_{10} = .16$ ). The

working memory capacity composite did, however, associate in the anticipated way with the antisaccade task (e.g., Kane et al., 2001; Unsworth, Schrock, & Engle, 2004),  $r = .26$ ,  $p < .001$ ,  $CI [.13, .37]$ ,  $BF_{10} = 253$ , the individual complex span tasks correlated with each other at a magnitude in line with a recent estimates of the association between shortened span tasks (Goller, Banks, & Meier, 2020;  $r = .36$ ) and longer three-block versions of the tasks (Kane et al., 2017,  $r = .40$ ),  $r = .33$ ,  $p < .001$ ,  $CI [.21, .44]$ ,  $BF_{10} = 57049$ , and provided reasonable internal consistency estimates for shortened measures (operation span = .72; symmetry span = .62) suggesting that the measures did represent working memory capacity.

At first sight, the null associations among ADHD questionnaire scores, working memory capacity measures, and antisaccade performance may seem strange and perhaps cause doubt about the primary findings and interpretations reported here. Confidence in the ADHD scores reported here is bolstered by their reliability estimates and the findings of the expected associations with IPI Daydreaming, SART  $d'$ , SART RTSD, and SART mind wandering. In the introduction, I cited three meta-analyses as evidence of an association between working memory capacity and ADHD. Notably, two of these studies assessed the relationship in children. Some recent work suggests that ADHD in children and adults are distinct syndromes (Agnew-Blais et al., 2016; Caye et al., 2016; Moffitt et al., 2015); thus, findings in children may not apply here. The meta-analysis in adults that focused on this association contained no complex span tasks. Multiple recent investigations with young adult samples have reported null associations between ADHD self-reports and complex span task performance (Franklin et al., 2017; Jonklin et al., 2017; Unsworth et al., 2019), suggesting that the conceptually inconsistent results reported here (i.e., the null association between working memory capacity and ADHD symptomology) may be the product of the task used to measure working memory capacity.

Reports of positive associations between antisaccade accuracy and ADHD in adults are common (e.g., Carr, Nigg, & Henderson, 2006; Feifel et al., 2004; Schwerdtfeger et al., 2013). Here, I reported a conceptually inconsistent null association. A potential critical difference between this study and (at least the majority) of prior studies reporting a positive association is that prior studies used eye-tracking to assess saccades. I used a behavioral antisaccade that required subjects to press a button indicating the identity of a quickly-masked target. This behavioral antisaccade task may lack the sensitivity to pick up on ADHD-related performance differences. The null findings reported here line up with other work using the same behavioral antisaccade task. Null associations were also found between ADHD symptomology and antisaccade performance (measured as it was in the current study) in Unsworth et al. (2019; personal communication, July 22, 2020) and in a recent large-sample study by Kane et al. (in press; N = 822) suggesting that the findings reported here are not idiosyncratic.

### **Potential Reasons for Differences Between Studies**

There are many reasons why estimates produced by different studies may vary. One reason is that the current study may not have replicated the conditions of the original study. This concern is diminished because of the following design choices: I used the irrelevant distractor task from Forster and Lavie (2016, Experiment 1) and performance on the irrelevant distractor task lines up well between the current study and Forster and Lavie (see Table 3). Also, following Forster and Lavie (2014), I used the IPI-daydreaming scale to measure mind wandering. Here, the IPI-daydreaming scale was modified (details are in the methods section) to enhance the measure's construct validity. These changes did not seem to injure the construct validity of the measure because the estimated correlations among the IPI-daydreaming scale and ADHD, SART

performance measures, and most critically, probe-caught mind wandering are all consistent with what one would expect from the unmodified scale.

Although I did not use the same ADHD measure by name, the measure I used is essentially identical in content to the one used by Forster and Lavie (2016). Both measures assess the DSM-IV-TR ADHD symptom list experienced by the subjects in childhood. The current administration of the ADHD measure was incomplete and did not present one item each from the inattentive and hyperactivity subscales. However, because of these scales' high internal consistency (both reported here and in prior work), this error is most likely not responsible for the current studies discrepant finding with the prior study. In contrast to Forster and Lavie (2016), I did not exclude subjects for taking ADHD medication. Although medication adherence for the observed differences between studies cannot be ruled out, it stands to reason that medication adherence would also affect the associations between ADHD and other measures in this study. This was not the case. Positive associations among ADHD, SART Nontarget RTSD, SART Mind Wandering, and IPI Daydreaming and the negative association between ADHD and SART d' were detected. It is also challenging to explain how medication adherence could impair the irrelevant distraction measure's internal consistency.

In the current study, during the irrelevant distractor task, the viewing distance between the subject and the computer display was maintained by having reference marks on the floor, instructing subjects to keep their head at this distance, and having an experimenter monitor this distance. The current study also used white noise machines to block potentially distracting sounds. Some research has suggested that white noise can remedy cognitive task deficits in children with ADHD (Bajot et al., 2016), others have reported benefits not specific to those with ADHD (i.e., helping those with and without ADHD; Söderlund, Björk, & Gustafsson, 2016), and

yet others report results consistent with white noise impairing performance in higher-ability subjects while aiding lower-ability subjects (Helps et al., 2018). If the inclusion of white noise was responsible for the near-zero correlation between ADHD symptomology and irrelevant distractor interference, it did so in a selective way, i.e., while maintaining the theoretically-consistent associations among ADHD, SART Nontarget RTSD, SART Mind Wandering, SART  $d'$  and IPI Daydreaming. It is also unclear how white noise would affect the criterion measure's reliability, which seems to be an issue in the current study.

Another way the conditions between this study and the original may have differed is in the composition of the sample. Notably, scores on the ADHD questionnaire were higher for the WCU sample, even with one less item per subscale (ADHD inattention  $M = 9$ ; ADHD hyperactivity  $M = 10$ ) than the University College London (UCL) sample (in both Experiment 1 and 2, ADHD inattention  $M = 6$ ; ADHD hyperactivity  $M = 7$ ). It is also possible that the irrelevant distractors were less familiar and thus less distracting to subjects from WCU than they were to UCL subjects that participated in Forster and Lavie (2014, 2016). The size of the interference effect in the low-load condition was only six milliseconds smaller (i.e., 24% smaller) in the current study using the same task than in Forster and Lavie (2016; other work from Forster and Lavie [2014, 2016] using different samples and task variants has produced larger effect sizes), suggesting that this (i.e., the distractors being less salient and disruptive to the Western Carolina sample) is likely not a driving factor in the inconsistent findings. Here, I reported characteristics of the sample in the case that future research explores this finding's generalizability.

Even if a study perfectly replicates a previous study's conditions, it may not achieve the same results as the original (in terms of statistical significance) if the replication failed to have

sufficient statistical power to reject the null hypothesis. Although I planned the sample size for precise estimates rather than statistical power, with a sample size of 226, I had a power of .998 (calculated with the pwr package; Champley et al., 2018) to correctly reject the null hypothesis for association between irrelevant distractor interference and ADHD symptoms with the effect size provided by Forster and Lavie (2016;  $r = .32$ ) and I had a power of .675 to correctly reject the null hypothesis for an effect that was half of that effect size ( $r = .16$ ). Regarding the association between irrelevant distractor interference and the IPI-daydreaming scale, I had a power of .978 to detect an effect of the same size as reported by Forster and Lavie (2014) and a power of .498 to detect an effect of half this size. (Notably, my estimate for this association was in the opposite direction compared to the original).

Finally, there is a chance through multiple avenues (i.e., sampling error, researcher degrees of freedom; Simmons, Nelson, & Simonsohn, 2011) that the original findings were false positives or that the current finding is a false negative. Because I have no basis with which to judge the original or current studies with regards to sampling error (besides sample size) and differences in researcher degrees of freedom between the original and the current work are unknown, additional pre-registered, well powered studies by independent researchers are needed to discern the (conditional) veracity of the current and previous estimates.

### **Conclusion**

Forster and Lavie (2014, 2016) provided evidence that a measure of task-irrelevant distraction was associated with ADHD symptomology and the propensity to mind wander. They claimed this as evidence for a *new* trait. This study sought to replicate these associations on the way to testing if irrelevant distraction propensity was truly a new and distinct construct. The present study produced estimates of association that were inconsistent with those found by



Forster and Lavie. That is, the data from this study supported a conclusion of no (or extremely weak) association between irrelevant distraction propensity and the other measured variables (rendering the goal of determining the uniqueness of these associations pointless). Moreover, the measure of irrelevant distraction used by Forster and Lavie produced a very low estimate of internal reliability, suggesting that irrelevant distraction (measured in this way; see Draheim et al. [2019] for a review of issues using reaction time difference scores) lacks the psychometric properties desirable in an individual differences variable. Taken together, these results suggest that claims of a *new* trait are not yet warranted.

#### Open Practices Statement

The data and analyses for this project are available here: <https://osf.io/bhs24/>. This study was preregistered. The preregistration is available here: <https://aspredicted.org/ng4ar.pdf>. The SART and antisaccade tasks used here are available at: <https://osf.io/bhs24/>. The irrelevant distractor task is available from the original author. Shortened complex span tasks are available from Randy Engle's lab at Georgia Institute of Technology: <http://englelab.gatech.edu/taskdownloads>. The questionnaires used here are available from the original authors.

## References

- Alderson, R. M., Kasper, L. J., Hudec, K. L., & Patros, C. H. (2013). Attention-deficit/hyperactivity disorder (ADHD) and working memory in adults: a meta-analytic review. *Neuropsychology, 27*(3), 287-302.
- Agnew-Blais, J. C., Polanczyk, G. V., Danese, A., Wertz, J., Moffitt, T. E., & Arseneault, L. (2016). Evaluation of the persistence, remission, and emergence of attention-deficit/hyperactivity disorder in young adulthood. *JAMA psychiatry, 73*(7), 713-720.
- Baijot, S., Slama, H., Söderlund, G., Dan, B., Deltenre, P., Colin, C., & Deconinck, N. (2016). Neuropsychological and neurophysiological benefits from white noise in children with and without ADHD. *Behavioral and Brain Functions, 12*(1), 11.
- Barkley, R. A., & Murphy, K. R. (1998). A clinical workbook: Attention-deficit hyperactivity disorder. *New York: Guilford*.
- Brown, W. (1910). Some experimental results in the correlation of mental health abilities. *British Journal of Psychology, 1904-1920, 3*, 296–322.
- Carr, L. A., Nigg, J. T., & Henderson, J. M. (2006). Attentional versus motor inhibition in adults with attention-deficit/hyperactivity disorder. *Neuropsychology, 20*(4), 430-441.
- Caye, A., Rocha, T. B. M., Anselmi, L., Murray, J., Menezes, A. M., Barros, F. C., ... & Swanson, J. M. (2016). Attention-deficit/hyperactivity disorder trajectories from childhood to young adulthood: evidence from a birth cohort supporting a late-onset syndrome. *JAMA psychiatry, 73*(7), 705-712.
- Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., ... & De Rosario, M. H. (2018). Package 'pwr'. *R package version, 1-2*.

- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281-302.
- Colom, R., Abad, F. J., Quiroga, M. Á., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why?. *Intelligence*, 36(6), 584-606.
- Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, 12(5), 769-786.
- Cooper, S. R., Gonthier, C., Barch, D. M., & Braver, T. S. (2017). The role of psychometrics in individual differences research in cognition: A case study of the AX-CPT. *Frontiers in psychology*, 8, 1482.
- Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, 145(5), 508-535.
- DuPaul, G. J., Power, T. J., Anastopoulos, A. D., Reid, R. *ADHD Rating Scale-IV: Checklists, Norms, and Clinical Interpretation*. New York: The Guilford Press; 1998.
- Feifel, D., Farber, R. H., Clementz, B. A., Perry, W., & Anllo-Vento, L. (2004). Inhibitory deficits in ocular motor behavior in adults with attention-deficit/hyperactivity disorder. *Biological psychiatry*, 56(5), 333-339.
- Forster, S., & Lavie, N. (2014). Distracted by your mind? Individual differences in distractibility predict mind wandering. *Journal of experimental psychology: learning, memory, and cognition*, 40(1), 251-260.

- Forster, S., & Lavie, N. (2016). Establishing the attention-distractibility trait. *Psychological science*, 27(2), 203-212.
- Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & cognition*, 43(2), 226-236.
- Goller, H., Banks, J. B., & Meier, M. E. (2020). An individual differences investigation of the relations among life event stress, working memory capacity, and mind wandering: A preregistered replication-extension study. *Memory & Cognition*, 1-13.
- Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, 50(3), 1166-1186.
- Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General*, 136(2), 217-240.
- Helps, S. K., Bamford, S., Sonuga-Barke, E. J., & Söderlund, G. B. (2014). Different effects of adding white noise on cognitive performance of sub-, normal and super-attentive school children. *PloS one*, 9(11), e112768.
- Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of experimental psychology: General*, 130(2), 169-183.
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of experimental psychology: General*, 132(1), 47-70.

- Kane, M. J., Gross, G. M., Chun, C. A., Smeekens, B. A., Meier, M. E., Silvia, P. J., & Kwapil, T. R. (2017). For whom the mind wanders, and when, varies across laboratory and daily-life settings. *Psychological science*, *28*(9), 1271-1289.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of experimental psychology: General*, *133*(2), 189-217.
- Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General*, *145*(8), 1017-1048.
- Kane, M. J., Smeekens, B. A., Meier, M. E., Welhaf, M. S., & Phillips, N. E. (in press). Testing the Construct Validity of Competing Measurement Approaches to Probed Mind-Wandering Reports. *Behavior Research Methods*.
- Kasper, L. J., Alderson, R. M., & Hudec, K. L. (2012). Moderators of working memory deficits in children with attention-deficit/hyperactivity disorder (ADHD): a meta-analytic review. *Clinical psychology review*, *32*(7), 605-617.
- Kofler, M. J., Irwin, L. N., Soto, E. F., Groves, N. B., Harmon, S. L., & Sarver, D. E. (2019). Executive functioning heterogeneity in pediatric ADHD. *Journal of abnormal child psychology*, *47*(2), 273-286.
- McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 196-204.

- McVay, J. C., & Kane, M. J. (2012a). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of experimental psychology: general*, *141*(2), 302-320.
- McVay, J. C., & Kane, M. J. (2012b). Drifting from slow to “d'oh!”: Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3), 525-549.
- Meier, M. E., & Kane, M. J. (2013). Working memory capacity and Stroop interference: Global versus local indices of executive control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 748-759.
- Meier M.E., & Kane, M. J. Attentional control and working memory capacity. In: Egner T, editor. *The Wiley Handbook of Cognitive Control*. Chichester, West Sussex, UK: John Wiley and Sons (2017). p. 50–63.
- Meier, M. E., Smeekens, B. A., Silvia, P. J., Kwapil, T. R., & Kane, M. J. (2018). Working memory capacity and the antisaccade task: A microanalytic–macroanalytic investigation of individual differences in goal activation and maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(1), 68 -84.
- Moffitt, T. E., Houts, R., Asherson, P., Belsky, D. W., Corcoran, D. L., Hammerle, M., ... & Poulton, R. (2015). Is adult ADHD a childhood-onset neurodevelopmental disorder? Evidence from a four-decade longitudinal cohort study. *American Journal of Psychiatry*, *172*(10), 967-977.
- Morey, R. D., & Rouder, J. N. (2018). BayseFactor: computation of bayes factors for common designs.
- Parsons, S. (2019). Splithalf: Robust estimates of split half reliability.

Parsons, S., Kruijt, A. W., & Fox, E. (2019). Psychological Science Needs a Standard Practice of Reporting the Reliability of Cognitive-Behavioral Measurements. *Advances in Methods and Practices in Psychological Science*, 2(4), 378-395.

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from <http://www.pstnet.com>.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20(5), 713-721.

Robison, M. K., Gath, K. I., & Unsworth, N. (2017). The neurotic wandering mind: An individual differences investigation of neuroticism, mind-wandering, and executive control. *The Quarterly Journal of Experimental Psychology*, 70(4), 649-663.

Robison, M. K., & Unsworth, N. (2018). Cognitive and contextual correlates of spontaneous and deliberate mind-wandering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(1), 85-98.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609-612.

Schwerdtfeger, R. M. H., Alahyane, N., Brien, D. C., Coe, B. C., Stroman, P. W., & Munoz, D. P. (2013). Preparatory neural networks are impaired in adults with attention-deficit/hyperactivity disorder during the antisaccade task. *NeuroImage: Clinical*, 2, 63-78.

- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, 22(11), 1359-1366.
- Singer, J. L., & Antrobus, J. S. (1970). Imaginal Processes Inventory. Princeton, NJ: Educational Testing Service.
- Smeekens, B. A., & Kane, M. J. (2016). Working memory capacity, mind wandering, and creative cognition: An individual-differences investigation into the benefits of controlled versus spontaneous thought. *Psychology of Aesthetics, Creativity, and the Arts*, 10(4), 389-415.
- Söderlund, G. B., Björk, C., & Gustafsson, P. (2016). Comparing auditory noise treatment with stimulant medication on cognitive task performance in children with attention deficit hyperactivity disorder: results from a pilot study. *Frontiers in psychology*, 7, 1331.
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 1904-1920, 3, 271–295.
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: active maintenance in primary memory and controlled search from secondary memory. *Psychological review*, 114(1), 104-132.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior research methods*, 37(3), 498-505.
- Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 832.



Unsworth, N., & McMillan, B. D. (2014). Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta psychologica, 150*, 14-25.

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1302-1321.

## Footnotes

1. The decision to not exclude any subjects based on this infrequency items was because no criterion was preregistered. This initial intent was to have one but because of experimenter error it was never determined. Only one subject who was included in the analyses provided answers to more than one of the questions that may have been considered infrequent and therefore may have been eliminated from the study. The post-hoc criterion for single items being counted as an infrequent response was providing an answer on the other side of neutral on the Likert-type scale from the “correct” response (only six responses overall [i.e., across all subjects] would have been deemed infrequent with this criterion). For example, a response to the statement “I don’t like getting speeding tickets” was deemed infrequent if the subject responded “disagree” or “strongly disagree”.
2. All correlation analyses were also conducted with the more stringent exclusion criteria from Forster and Lavie, 2016. With these criteria, subjects were excluded from all analyses if they achieved under 55% accuracy in either condition (i.e., low or high-load; rather than overall) or if their mean response time was greater than two standard deviations from the mean in either condition. The change in estimates was minimal. The largest magnitude of change was for the ADHD hyperactivity by low-load RT interference association which went from  $-.13$  to  $-.04$ . Using these exclusion criteria, the strongest association between an irrelevant-distractor-interference variable and another non-interference variable was the association between the symmetry span score and high-load accuracy interference, this association was  $r(200) = .13$  ( $p = .055$ ,  $BF = .97$ ). A document with both the correlation matrix using the preregistered exclusion criteria and a matrix using the exclusion criteria from Forster and Lavie can be found at the OSF site for this project (filename: Correlation Matrices from differing exclusion criteria.pdf). Analysis scripts and detailed correlation analyses for the correlations with Forster and Lavie’s (2016) exclusion criteria are reported in the files named “AD notebook with Forster and Lavie exclusions” on this project’s OSF page: <https://osf.io/bhs24/>.
3. All analyses were also run using an accuracy interference measure in place of the response time interference measure. None of these analyses provided evidence in line with the associations reported by Forster and Lavie (2014, 2016). These analyses can be found in the supplemental materials.

## Author Contributions

M.E. Meier developed the study concept and design, performed the data analysis and interpretation, and drafted the manuscript. I wish to acknowledge the help of Spencer Acker, Kaitlynn Divine, Natalia Torres Wong, and Madison Pruitt for their work in collecting the data for this project.