Baseline Pupil Size Seems Unrelated to Fluid Intelligence, Working Memory Capacity, and Attentional Control

Ruuskanen, Veera; Hagen, Thomas; Espeseth, Thomas; Mathôt, Sebastiaan

Published in:
Journal of Cognition

DOI:
10.5334/joc.365

journal of cognition

Baseline Pupil Size Seems Unrelated to Fluid Intelligence, Working Memory Capacity, and Attentional Control

RESEARCH ARTICLE

VEERA RUUSKANEN
THOMAS HAGEN
THOMAS ESPESETH
SEBASTIAAN MATHÔT
*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

Over the past few years, several studies have explored the relationship between resting-state baseline pupil size and cognitive abilities, including fluid intelligence, working memory capacity, and attentional control. However, the results have been inconsistent. Here we present the findings from two experiments designed to replicate and expand previous research, with the aim of clarifying previous mixed findings. In both experiments, we measured baseline pupil size while participants were not engaged in any tasks, and assessed fluid intelligence using a matrix task. In one experiment we also measured working memory capacity (letter-number-sequencing task) and attentional control (attentional-capture task). We controlled for several personal and demographic variables known to influence pupil size, such as age and nicotine consumption. Our analyses revealed no relationship between resting-state pupil size (average or variability) and any of the measured constructs, neither before nor after controlling for confounding variables. Taken together, our results suggest that any relationship between resting-state pupil size and cognitive abilities is likely to be weak or non-existent.

CORRESPONDING AUTHOR:
Veera Ruuskanen
Department of Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712TS Groningen, The Netherlands
v.h.s.ruuskanen@rug.nl

KEYWORDS:
Baseline pupil size; fluid intelligence; working memory capacity; attentional control

TO CITE THIS ARTICLE:
Ruuskanen, V., Hagen, T., Espeseth, T., & Mathôt, S. (2024). Baseline Pupil Size Seems Unrelated to Fluid Intelligence, Working Memory Capacity, and Attentional Control. *Journal of Cognition*, 7(1): 41, pp. 1–18. DOI: https://doi.org/10.5334/joc.365

Changes in pupil size are not only brought about by changes in luminance, but also by cognitive activity. For instance, viewing arousing images (Hess & Polt, 1960), memorizing more items during working memory tasks (Kahneman & Beatty, 1966), and performing increasingly difficult mental arithmetic (Hess & Polt, 1964) are all associated with pupil dilation. These so-called psychosensory pupil responses are small in size, but well-established, reliable markers of mental effort or arousal (for recent reviews, see Laeng et al., 2012; Mathôt, 2018), and may play a functional role in optimizing vision (reviewed in Vilotijević & Mathôt, 2023).

When measuring pupil size, there is an important distinction between task-evoked and baseline pupil size. Task-evoked changes in pupil size are measured during task performance, in response to a specific event, such as the appearance of a stimulus. Much of the research on pupil size has focused on relating the strength of task-evoked changes in pupil size to cognitive factors, such as the amount of mental effort invested in a task. However, more recently the possible relationship between baseline pupil size (as opposed to task-evoked pupil size) and individual differences in cognitive abilities has attracted considerable attention. Baseline pupil size can refer either to a pre-trial baseline, measured during task performance immediately before each trial; or to a pre-experimental baseline, measured in the absence of a task, during a resting state. Here, we will use baseline pupil size to refer to the latter: a pre-experimental, resting-state baseline.

Cognitive abilities that have been suggested to be related to baseline pupil size include working memory capacity (WMC; Aminihajibashi et al., 2019; Heitz et al., 2008; Tsukahara et al., 2016), fluid intelligence (Tsukahara et al., 2016; Tsukahara & Engle, 2021b; van der Meer et al., 2010), and attentional control (Unsworth & Robison, 2017). Working memory refers to a cognitive system involved in maintaining, manipulating, and retrieving information relevant to the task at hand, and its capacity is limited (Baddeley & Hitch, 1974; Cowan, 1988; Miller, 1994; Sperling, 1960). Fluid intelligence refers to reasoning and problem-solving abilities (Cattell, 1963; Fukuda et al., 2010). Attentional control refers to the ability to direct attentional resources to the task at hand, and (here) to prevent attentional capture by irrelevant stimuli.

The notion that there might be a relationship between baseline pupil size and cognitive abilities follows mainly from the association between pupil size and activity of the locus coeruleus (LC), such that large pupils indicate high levels of LC activity (Alnæs et al., 2014; Joshi et al., 2016; Murphy et al., 2011, 2014). The LC is a neuromodulatory nucleus located in the brainstem that has projections to the neocortex, including prefrontal areas, and has been suggested to play a modulating role in various cognitive functions, such as attention and behavioral control (Aston-Jones et al., 1999; Aston-Jones & Cohen, 2005; Chandler et al., 2014). Key support for this role of the LC comes from findings showing that people tend to perform better at moments when their pupils are relatively dilated (Keene et al., 2022), presumably because at these moments the LC is relatively active.

Some authors have also linked individual differences in cognitive abilities to the neuromodulatory activity of the LC, both during task performance and at rest. For instance, it has been suggested that differences in performance on tasks involving attention and working memory are due to differences in LC functioning, such that individuals with a dysregulated LC will exhibit poorer performance, and thus, score lower on WMC and attentional control. Conversely, better performance (indicating higher WMC and attentional control) is suggested to be due to more optimal LC functioning (Unsworth & Robison, 2017). Moreover, because of the link between LC activity and baseline pupil size, it has even been hypothesized that differences in these cognitive abilities may be reflected in baseline pupil size as well; in essence, it has been noted that there is a correlation between the size of a person's pupils at rest and their (fluid) intelligence and working-memory capacity (Tsukahara et al., 2016). This is suggested to be driven by differences in the strength of functional connectivity between brain areas at rest, due to the neuromodulatory role of the LC (Tsukahara & Engle, 2021a). However, investigations into the relationship between baseline pupil size and cognitive abilities have produced mixed results (see also Unsworth et al., 2021 for a comprehensive review).

The first empirical indications of a relationship between baseline pupil size and WMC (Heitz et al., 2008) as well as between baseline pupil size and fluid intelligence (van der Meer et al., 2010) emerged as incidental findings in studies where the focus was on measurements obtained during task performance. Since baseline pupil size was not the focus of these studies, possible confounding factors such as age and alcohol consumption (Birren et al., 1950; Tryon, 1975) were not considered. Therefore, although these results were suggestive of a relationship, they cannot be taken as decisive evidence.

Following these incidental findings, several studies have provided further evidence for a relationship between baseline pupil size and cognitive abilities. The first large-scale study was conducted by Tsukahara, Harrison & Engle (2016). In a series of three experiments, they found a significant positive correlation between average pupil size and both WMC and fluid intelligence. WMC was measured with a composite score of the operation span, rotation span, and symmetry span tasks; fluid intelligence was measured with Raven's Advanced Progressive Matrices (RAPM), Letter Sets, and Number Series. They accounted for several possible confounding factors such as mental effort, familiarity with the environment, demographic variables (e.g., age), and substance use. When statistically controlling for fluid intelligence, they found that the relationship between pupil size and WMC was non-significant; conversely, when controlling for WMC, the relationship between pupil size and fluid intelligence persisted. Based on this the authors concluded that it is fluid intelligence, rather than WMC, that is uniquely related to baseline pupil size (Tsukahara et al., 2016). However, despite this initial positive finding, the few studies that have since investigated the link between baseline pupil size and fluid intelligence have generally failed to replicate the relationship (Coors et al., 2022; Robison & Brewer, 2022; Robison & Campbell, 2023; Robison et al., 2023).

Most subsequent studies have focused on the relationship between baseline pupil size and WMC. However, a number of them, including Aminihajibashi et al. (2019), Coors et al. (2022), Robison et al. (2022; 2023), and Unsworth et al. (2019), have failed to replicate the finding of a correlation between average baseline pupil size and WMC. Additionally, a recent meta-analysis concluded that the correlation between baseline pupil size and WMC is not robust (Unsworth et al., 2021). One exception was the study by Aminihajibashi et al. (2019), which found a positive correlation between *variability* in pupil size (indexed by the coefficient of variation [CoV]) and WMC.

Similarly, the few studies that have measured attentional control have generally failed to find a correlation with resting-state baseline pupil size (Robison et al., 2022; Robison & Brewer, 2022; Unsworth et al., 2019). Some studies have demonstrated a relationship with pre-trial baseline pupil size instead, whereby *variability* in baseline pupil size was related to attentional control (but average baseline pupil size was not; Unsworth & Robison, 2017). Overall however, this relationship has received less attention than the correlation between pupil size and both WMC and fluid intelligence.

In spite of the several failed attempts to replicate relationships between baseline pupil size and individual differences in cognitive abilities, methodological differences between studies make it difficult to draw firm conclusions. Tsukahara & Engle (2021b) investigated how these differences may influence results, in an effort to clarify why in their lab they consistently find relationships while many others fail to do so. One crucial issue they identified was that different studies measured pupil size under different luminance conditions; this is relevant because a bright laboratory may result in small pupils for everyone, thus reducing variance in pupil size between individuals. With little variance between individuals it becomes difficult to find relationships between variables. Based on both a reanalysis of their own data and a new experiment where luminance conditions were manipulated, Tsukahara & Engle (2021b) showed that low variance in the sample of pupil sizes can indeed result in a weak and non-significant correlation between pupil size and WMC.

In addition, based in part on their first study, Tsukahara & Engle (2021b) emphasized the importance of the relationship between baseline pupil size and fluid intelligence over other cognitive abilities, such as WMC and attentional control. As mentioned above, many studies have focused on WMC rather than fluid intelligence; however, it seems that only fluid intelligence is uniquely related to pupil size, and that this direct relationship mediates indirect relationships with both WMC (Tsukahara et al., 2016; Tsukahara & Engle, 2021b) and attentional control (Tsukahara & Engle, 2021b). More generally, all three constructs are highly correlated with each other, with some evidence even suggesting that attentional control largely accounts for the shared variance between WMC and fluid intelligence (Unsworth et al., 2014). Nevertheless, they are still partly distinct (Kane et al., 2005) and could potentially have unique relationships with pupil size.

Following these considerations, here we present the results of two experiments that investigated the relationship between pupil size on the one hand, and fluid intelligence (two experiments), WMC (one experiment) and attentional control (one experiment) on the other. To account for the relative paucity of studies on this topic, as pointed out by Tsukahara & Engle (2021b), our main analysis focuses on fluid intelligence. Furthermore, the conditions in which

Journal of Cognition
DOI: 10.5334/joc.365

our baseline pupil size measurements were obtained closely follow their recommendations. In addition to measuring cognitive abilities, we assessed several possible confounds that are known to affect pupil size (e.g., age, nicotine, alcohol, and caffeine consumption etc.). Our goal was to determine whether previous results on the relationship between pupil size and cognitive abilities are replicable, to help resolve previously observed inconsistencies.

METHODS

PARTICIPANTS

The results presented here are based on two separate samples of participants, one collected at the University of Groningen (The Netherlands) and the other at the University of Oslo (Norway). The final combined sample consisted of 224 participants. As detailed below, we did not specify a minimum number of participants prior to the experiments. However, other authors have cited 200 as a minimum viable sample for similar studies (Robison et al., 2023).

Groningen

The sample collected in Groningen consisted of 104 participants, most of whom were undergraduate students at the University of Groningen. We had not specified a minimum sample size in advance, and intended to collect as many participants as was feasible during the course of the academic year. These participants were recruited through a web-based system, and they received partial course credit in exchange for participation. The mean age in the sample was 20.05 years, with a standard deviation of 1.9. There were 72 females and 32 males. The majority of the participants were Dutch (35.29%) or German (32.35%), while the rest came from various other, mainly European, countries. Participants indicated informed consent by signing a form before the beginning of the study. The study was approved by the local ethics review board (study approval code: PSY1920-S-0178).

Oslo

The sample collected in Oslo consisted of 122 participants that were recruited through social media. These participants were collected for another project which did not aim to investigate baseline pupil size. So, a minimum required sample size for the analyses presented here had not been determined prior to data collection. The mean age was 25.56 years, with a standard deviation of 4.11. There were 81 females and 41 males. The study was approved by the local ethics review board (study approval code: 1439337).

MATERIALS, APPARATUS AND PROCEDURE

Procedure

Groningen
The experiment began with information and instructions given in both written and verbal form, followed by participants signing the informed consent form. After this, participants filled in a questionnaire on demographic information and possible confounds. Next, a two-minute recording of baseline pupil size was obtained, followed by participants completing the three cognitive tasks described below. Each task took approximately 10 minutes, and the order of the tasks was counterbalanced across participants. Finally, another pupil-size recording was obtained, followed by debriefing. The whole experiment took approximately 45 minutes to complete (Figure 1).

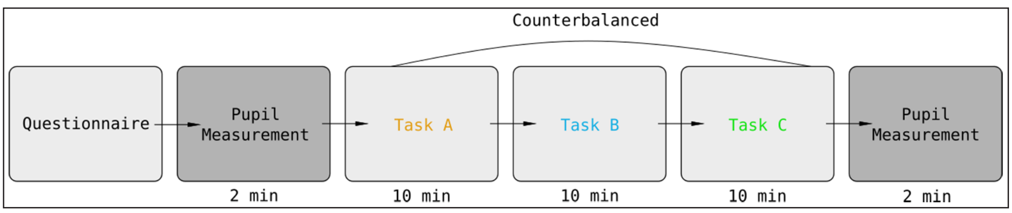

Figure 1 An overview of the experimental procedure in Groningen.

Oslo

The experiment began with informed consent and instructions. Participants then filled out a questionnaire on demographic information and possible confounds. Next, a two-minute recording of baseline pupil size was obtained. Participants looked at a fixation cross presented in the middle of a blank screen. Subsequently, participants completed a battery of tasks and questionnaires on cognition and motivation, which will be reported elsewhere. The whole procedure took approximately two hours.

Pupil measures

Groningen

Participants' pupils were measured twice: at the beginning and end of the experiment. Each measurement was conducted in the absence of a task, and lasted for two minutes. A video-based binocular eye-tracking device (The Eye Tribe, n.d.) was used to record pupil size at a sampling rate of 30 Hz. Pupil size was recorded in arbitrary units. No calibration procedure was applied prior to recording since spatial eye movements were not of interest. A chinrest was used to stabilize head movements and to keep the eye-to-monitor distance constant at 60 cm. All participants were tested in a windowless booth with constant illumination (approx. 130 lx measured at the chinrest). During recording participants were asked to fixate on the reflection of their own eyes in the eye tracker. This gives a frontal view of the eye, thus providing the best pupil recording. Further, fixating on the eye tracker provides a dark background for fixation, as recommended by Tsukahara and colleagues (2021b) for baseline pupil size measurements. Participants who wore glasses were allowed to keep them on during recording; however, some chose to take them off.

Oslo

One measurement was taken, again in the absence of a task, and lasting for two minutes. A binocular Remote Eye Tracking Device (R.E.D.; SMI-SensoMotoric Instruments, Teltow, Germany) was used to record pupil size at a sampling rate of 60 Hz. Pupil size was recorded in millimeters of diameter. A chinrest was used to stabilize head movements and to keep the eye-to-monitor distance fixed at 57 cm. Illuminance in the room as measured at the chinrest was 42 lx. Participants fixated on a black fixation cross presented on a gray background (RGB: 106, 106, 106).

Assessment of confounds

Groningen

The participants filled out a short questionnaire assessing several demographic variables that may influence pupil size. Specifically, participants reported their age, sex, nationality, handedness, and whether they wore contact lenses (hard or soft) or glasses. In addition, similar to the study by Tsukahara et al. (2016), participants reported: how much sleep they got the night before (in hours); their use of nicotine in the last 10 hours; consumption of alcohol in the last 24 hours; consumption of caffeine in the last 8 hours; and, optionally, use of medication in the past 24 hours. Lastly, participants reported any other factors (if any) they could think of that might influence their pupil size or cognitive abilities, such as memory or attention. Further, the time of the day the experiment took place was recorded.

Oslo

Participants verbally reported their age and sex, and whether they had consumed any nicotine or caffeine on the day of the experiment.

Fluid-intelligence assessment

Both samples of participants completed a matrix task as an assessment of fluid intelligence.

Groningen

The task was developed by a research group at the Goethe University in Frankfurt, and is modeled after Raven's advanced progressive matrices (RAPM; Raven et al., 2000). In this task, participants were shown a three-by-three matrix consisting of eight abstract figures and one empty box (Figure 2a). Underneath the matrix, eight response options were presented. Participants selected the response option that logically belongs in the empty box, thus completing the matrix. The response options were numbered, and participants gave their response by pressing the key corresponding to the number of the chosen option. 10 minutes were given to complete 16 matrices.

Oslo

The Hagen Matrices Test (HMT; Heydasch, 2014) was used. The HMT follows the same format as the RAPM, with participants being presented with an incomplete matrix and eight response options. Participants completed eight matrices, and were given maximally two minutes to complete each matrix.

Other cognitive tasks

Groningen

Participants completed assessments of working memory capacity (WMC) and attentional control. All three tasks (including the matrix task described above) were programmed with OpenSesame (version 3.2.8: Kafkaesque Koffka; Mathôt et al., 2012) and presented on a computer screen at a viewing distance of approximately 60 cm. The order of the tasks was counterbalanced such that each possible ordering was completed an equal number of times. The duration of each task was 10 minutes.

WMC was estimated with a letter-number-sequencing (LNS) task (Figure 2b). The LNS is a subset of the Wechsler Adult Intelligence Scale Third Edition (WAIS-III; Wechsler, 1997). In the task, participants see unsorted strings of letters and numbers. They memorize each string and report it with the numbers organized in numerical order, followed by the letters organized in alphabetical order. The length of the strings increases from two to eight. The test is traditionally administered orally (Mielicki et al., 2018). However, to increase ease of administration and analysis, we opted for visual presentation. The procedure and stimuli were the same as in experiment 3-A of a study by Mielicki et al. (2018) that assessed differences in visual and oral administration of the LNS, focusing on language background (bilingual vs. monolingual). They found that visual administration reduces biased results especially for bilingual participants (Mielicki et al., 2018). Given that most participants in the current study speak more than one language, this is ideal. Each character was presented in the center of the screen, in a black monospace font with a font size of 82 px, on a white background, for one second. After the presentation of the entire string, the participants gave their response using the corresponding keys on the keyboard. The time allocated to give a response increased from 7 to 16 seconds (with more time reserved for longer strings), after which the response was considered incorrect. The response times were determined at the author's discretion. All strings and associated response times can be found in Appendix A. The total number of presented strings was 30.

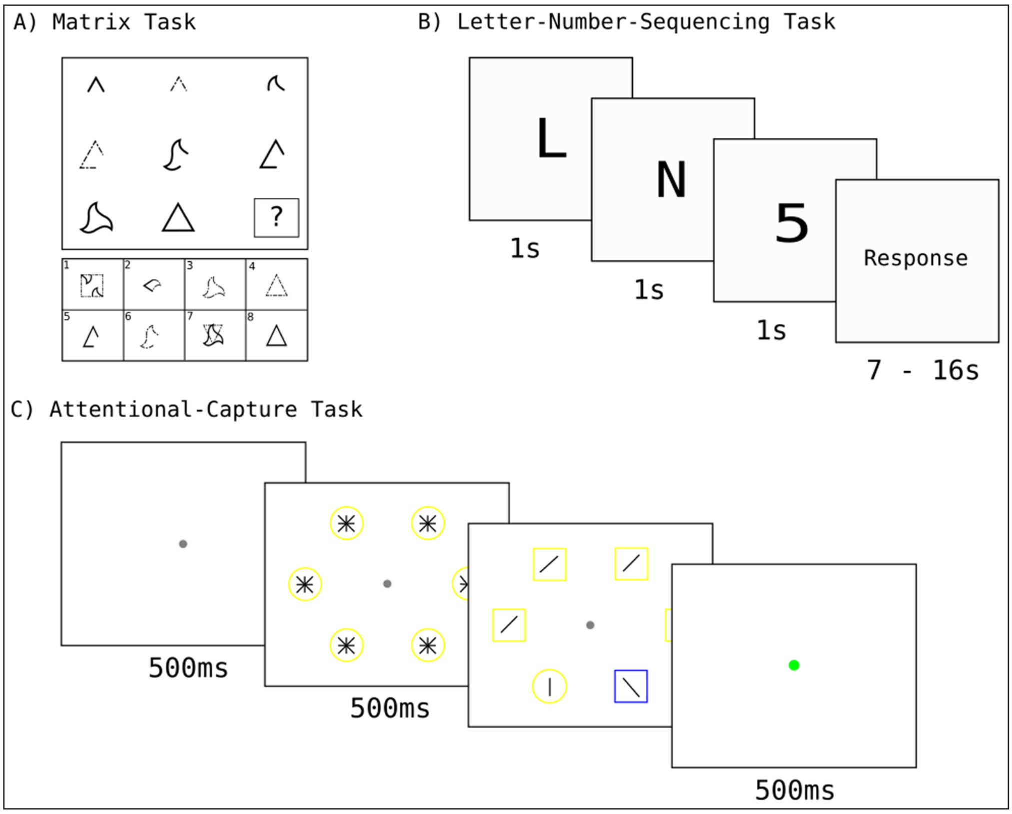

Figure 2 The three tasks completed by participants in Groningen. **A)** An example of a trial of the matrix task used to assess fluid intelligence. **B)** An example of a trial of the Letter-Number-Sequencing task (LNS) used to assess working memory capacity (WMC). **C)** An example of the attentional-capture task used to assess attentional control. *Note*: In the actual task the background was dark and the line segments in the shapes were white. The colors are reversed here for illustration purposes.

Attentional control was measured with an attentional-capture task that was loosely based on Theeuwes (1992). Participants are presented with six shapes symmetrically organized on a virtual circle on a black background (Figure 2c). On each trial, one of the six shapes is different from the others (a circle among squares or a square among circles). The shapes are blue or yellow in color and contain a line segment. Participants are asked to indicate the orientation of the line in the unique shape by pressing one key on the keyboard for a horizontal line and another for a vertical line. Crucially, on some trials, a distractor shape with a unique color is present, and on others all shapes have the same color. Attentional capture is indicated by the difference in reaction times between distractor-present and distractor-absent trials.

Participants completed 1 block of practice trials, and 8 blocks of experimental trials. A block consisted of 20 trials, making a total of 20 practice trials and 160 experimental trials. Each trial began with a fixation dot presented in the center of the screen for 500 ms. Following the fixation, a premask canvas was presented for 500 ms. The premask canvas had the target shape and color, with all possible line segments in each shape. Following the premask, the target canvas was presented until a response was given or until three seconds had passed, after which the response was considered incorrect. After each trial, feedback was given in the form of a green (correct) or red (incorrect) dot that was presented for 500 ms. After each block, feedback was given on accuracy and average reaction time during the block.

ANALYSIS AND RESULTS

DATA PROCESSING

Pupil data

Groningen

The preprocessing steps of the pupil data included smoothing with an 11- sample Hanning window and applying a blink-correction algorithm (Mathôt & Vilotijević, 2022). One participant was excluded from the analysis due to a technical measurement error that resulted in an impossible average pupil size value (–1.0). Following preprocessing, the mean and standard deviation over time was determined for each participant separately. Since the Eye Tribe reports pupil size in arbitrary units (as opposed to mm), pupil size was normalized across participants by z-scoring. This was done so that the two samples could be combined for analysis. Further, a coefficient of variation (CoV) of pupil size was computed from the raw data for each participant separately with the following formula: (SD/Mean) * 100. The same formula was used by Aminihajibashi et al. (2019). The CoV of pupil size was also normalized across participants by z-scoring.

Oslo

The preprocessing of pupil data was automatically conducted by the SMI software and no additional preprocessing was done. However, the data quality was visually inspected prior to analysis and deemed good. The SMI R.E.D reports pupil size in mm, but we again normalized the data by z-scoring so that the two samples could be combined. A coefficient of variation (CoV) was computed in the same way as for the Groningen data.

Task scores

Since the two matrix tasks included a different total number of matrices (Groningen: 16, Oslo: 8), the proportion (as opposed to the number) of correct answers was used for analysis. For the LNS, the number of correct answers was used. No trials were excluded for either of these tasks. For the attentional-control task, trials with too short (less than 200 ms) or too long (more than 2500 ms) reaction times were excluded, because in these trials it is likely that the participants were not paying attention to the task. This resulted in excluding approximately 2% of both distractor-present and distractor-absent trials (of the total number of experimental trials across all participants). Incorrect trials were also excluded. Next, average reaction times on correct trials were computed separately for distractor-present and distractor-absent trials, and the difference between the two was used as the final score on the task.

Control variables

The nominal control variables were coded for analysis. For caffeine, nicotine, and alcohol consumption a code of 0 meant the participant had not consumed/used the substance recently and a code of 1 meant they had. Sex was coded such that 1 = female and 0 = male or other. Handedness was coded such that 1 = right-handed and 0 = left-handed or ambidextrous. Whether the participant wore glasses or contact lenses was coded such that 1 = wears glasses or lenses and 0 = does not wear glasses or lenses. Time of day was rounded to the nearest hour. All data processing was done with scripts written in Python 3 (Rossum & Drake, 2009).

Analysis steps

Further analyses were performed using JASP version (0.16.4). One additional participant (the last one) from the sample collected in Groningen was excluded from the analysis to maintain proper counterbalancing across the order of the three tasks. Thus, the final sample size for Groningen was 102 and for Oslo 122.

Following the data processing steps outlined above we also analyzed data quality (e.g., reliability analyses and outlier detection). Next, we proceeded with the analyses intended to answer the research questions. The first analysis of interest focused on the relationship between baseline pupil size (average size and variability) and fluid intelligence, and was conducted using the combined sample of data from both experiments (N = 224). The second analysis of interest investigated the relationships between baseline pupil size (average size and variability) and WMC and attentional control. This analysis was performed on only the sample collected in Groningen, as WMC and attentional control were not assessed in Oslo. We report Bayesian correlations and Bayesian linear regression results for the main analyses, as they allow for accumulation of evidence for the null (Dienes, 2014).

Furthermore, since two pupil size measurements were taken in Groningen, a control analysis using the second measurement was conducted. This did not alter the results. Therefore, only results from analyses with the first measurement are reported, consistent with prior research favoring pre-task pupil measurements.

DISTRIBUTION OF PUPIL SIZES

The first step in the analysis was to determine the range and inter-individual variability of recorded pupil sizes. We attempted to find a way to convert the arbitrary unit measurements of the Groningen sample into meaningful units but were unable to. Thus, information on range and variability will only be reported for the sample collected in Oslo (N = 122), as these measures were in mm, making comparison with other studies possible. Average pupil size ranged from 3.01 mm to 7.22 mm with a mean of 4.39 mm and a standard deviation of 0.72 (Figure 3). The range, mean and standard deviation shown in Figure 3 were extracted from the raw data. However, for the remainder of this section, average pupil size and CoV of pupil size refer to normalized (i.e., z-scored) measures as described above.

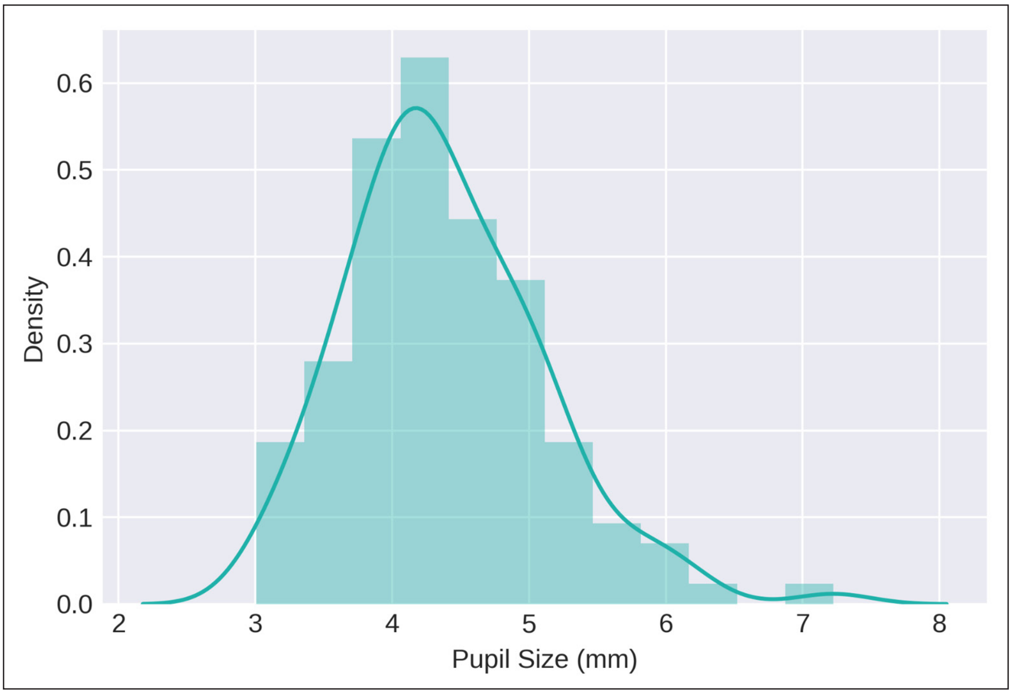

Figure 3 Distribution of resting-state pupil size in mm, as recorded from the sample collected in Oslo (N = 122).

Outlier detection

We checked for outliers on the z-scored pupil sizes (both average and variability) by isolating subjects that had z-scores above 3 or below –3. There were two outliers detected based on average pupil size and three based on CoV. The analyses described below were performed both with and without these outliers, which did not change the results. Hence, we report results including the outliers. This is because we believe that while unusual, these are nevertheless true scores, and not a result of measurement errors.

TASK SCORES

Outlier detection

To check for outliers in tasks scores we used the interquartile range (IQR) method. The IQR is the range between the first (Q1) and third (Q3) quartiles of the data, and outliers are defined as points that fall 1.5 × IQR below Q1 or 1.5 × IQR above Q3 (Vinutha et al., 2018). There were two outliers detected in the attentional-control task, and eight in the LNS task. The matrix task scores did not include any outliers.

The analyses were performed both with and without the outliers included. This generally did not change the results, with one minor exception of the correlation between the LNS and matrix task scores, detailed below. Thus, results reported here include outliers.

Descriptive statistics and reliability

The descriptive statistics for each task are reported in Table 1. The matrix task scores reported in the table are based on the combined sample. The mean in the Groningen subsample was 49% (*SD* = 13%; *Min* = 06%; *Max* = 75%), and in the Oslo subsample 54% (*SD* = 22%; *Min* = 0%; *Max* = 87%).

There is an indication of skewness in the distributions of the LNS and attentional-control task scores. For the LNS this is likely due to many participants scoring relatively high on the task. For the attention control task the source of the skew is likely to be the two participants who were faster on distractor-present than distractor-absent trials, thus ending up with negative final scores on the task.

In addition to the descriptive information reported in Table 1, we also conducted analyses on the reliability (as in internal consistency) of each task. For the matrix tasks and the attentional-control task this was done by means of a split-half reliability analysis. For the LNS a split-half analysis was not possible since items progressively increased in difficulty. So, we opted for Guttman's Lambda-2 (Guttman, 1945), which is a metric of internal consistency that is able to account for hierarchical differences in item difficulty. The results are reported below.

Matrix task

To determine the internal consistency of the matrix tasks, we split the items with an odd-even split, computed the final scores for each half (as the total number of correct responses), and compared the results. This procedure was conducted separately for the two samples. In the Groningen sample some participants completed different numbers of matrices due to the time constraint. If a participant had not completed all 16 matrices and if the last matrix they did complete was oddly numbered, the final response was deleted before splitting the data. This way we made sure that the same number of matrices per participant was included in each subset of the data.

For both samples, there was a significant positive correlation between the scores on each half of the task (Oslo: $r = 0.49$, $p < 0.001$; Groningen: $r = 0.47$, $p < 0.001$).

Letter-number-sequencing task

For the LNS, reliability was determined by way of Guttman's Lambda-2, using the original formula which can be found in Guttman (1945). The analysis indicated good internal consistency ($\lambda_2 = 0.81$).

Attentional control task

For the sake of consistency, we opted for split-half reliability with an odd-even split for the attentional-control task as well. Specifically, after removing practice trials and trials with too short (< 200 ms) or too long (> 2500 ms) reaction times, we split the trials into odd and even numbered trials, computed the final score (as the difference in average RT between distractor-present and distractor-absent trials) and compared the scores for the two halves. There was a significant positive correlation between the scores ($r = 0.29$, $p = 0.002$).

Table 1 Descriptive statistics of task scores.

TASK	M	SD	N	MIN	MAX	SKEW.	KURT.
Matrix	52%	19%	224	0%	88%	–0.27	0.05
LNS	19.07	3.62	102	5	25	–1.49	3.63
Attentional control	113.05	82.02	102	–324.79	299.03	–1.58	7.45

Correlations between measures

Correlations between pupil measures, task scores, and control variables are reported in Table 2. The correlations in the table are based on analyses including outliers. As was the case for the other analyses, also here the inclusion or removal of outliers did not notably affect the results, with the exception that removing the outliers from the LNS task scores increased the correlation between the LNS and the matrix task to $r = 0.17$ ($p = 0.09$). Further, the correlation between the two pupil measures (before and after the tasks) taken in Groningen was $r = 0.72$ ($p < .001$).

Table 2 Zero-order correlations between all measures.

	1	2	3	4	5	6	7	8	9
1. Average pupil size (z)	–								
2. Pupil size CoV (z)	.01	–							
3. Matrix task score	–.01	–.03	–						
4. LNS score	.05	0.11	.09	–					
5. Attentional control score	.11	–.06	.12	–.02	–				
6. Age	–.08	–.05	.01	.00	–.02	–			
7. Sex	.05	**.14***	–.12	–.06	.08	**–.18***	–		
8. Nicotine consumption	–.13	.01	–.01	.03	.03	.08	**–.18***	–	
9. Caffeine consumption	–.02	–.05	**–.18***	.07	–.13	–.06	–.15	**.23***	–

*p < .05.

Baseline Pupil Size and Fluid Intelligence

The Pearson correlation between average pupil size and percentages of correct answers on the matrix task was not statistically significant ($r = –0.01$, $p = 0.83$; Figure 4a). A Bayesian Pearson correlation showed support for the null hypothesis ($BF_{01} = 11.6$).

Similarly, the Pearson correlation between CoV of pupil size and scores on the matrix task was not significant ($r = –0.03$, $p = 0.7$; Figure 4b). A Bayesian Pearson correlation showed support for the null hypothesis ($BF_{01} = 11.1$).

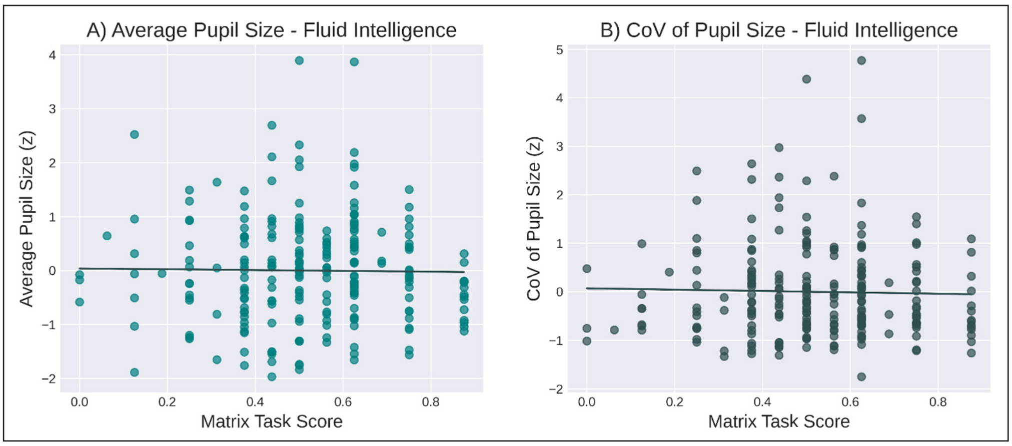

Figure 4 The relationship between baseline pupil size and fluid intelligence (as assessed with a matrix task). **A)** Average pupil size and fluid intelligence. **B)** Variability in pupil size and fluid intelligence.

Following the analysis of the correlations between baseline pupil measures and scores on the matrix task, a Bayesian linear regression was performed to investigate the influence of the measured control variables. First, a regression with average pupil size as the dependent variable was conducted. The included covariates were matrix task score, age, sex, caffeine consumption, and nicotine consumption. The coding of the nominal control variables is explained above. Information on caffeine and nicotine consumption was missing for 11 participants, making the total sample size of this analysis 213. The model that fit the data best included the control variable nicotine but no other covariates (Figure 5). The Bayes factors indicated that the data was 17 times more likely to be observed under this model than the second-best fitting model. However, in the full sample the correlation between average pupil size and nicotine consumption was not significant (Table 2); that is, the extent to which there is evidence for a relationship between average pupil size and nicotine consumption in the full sample is somewhat dependent on which analysis one chooses to look at.

Next, the same analysis was done with the CoV of pupil size as the dependent variable, and the same covariates as above. The model that fit the data best was the null model.

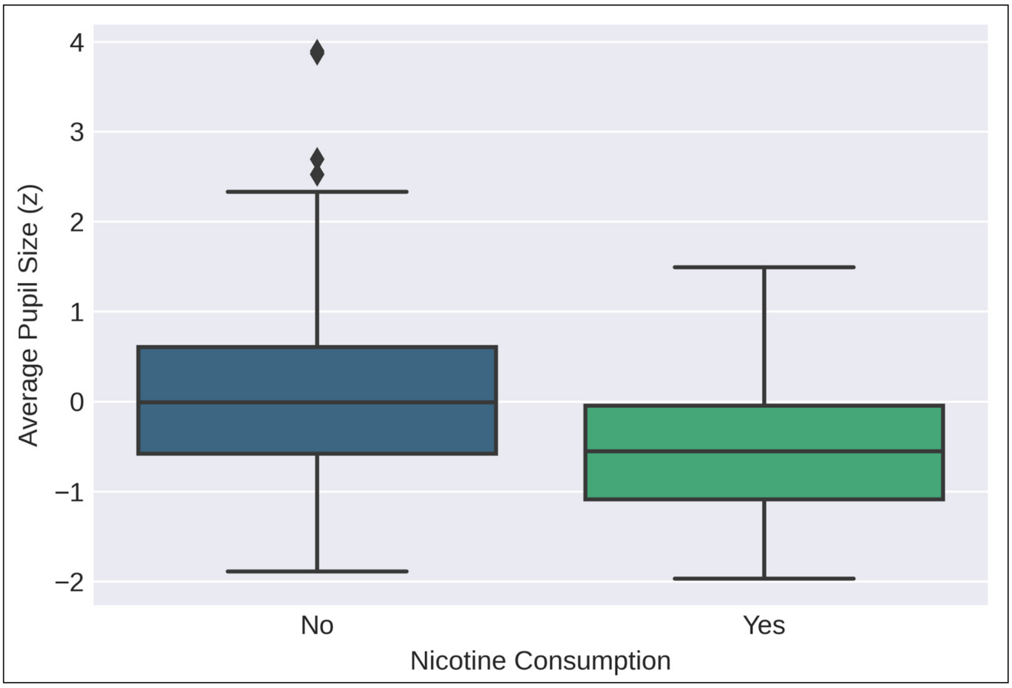

Figure 5 Distribution of pupil sizes across levels of nicotine consumption in the full sample ("no" indicating no nicotine consumption on the day of the experiment (N = 197) and "yes" indicating nicotine consumed on the day of the experiment (N = 16)).

Baseline Pupil Size, Working Memory Capacity and Attentional Control

The second analysis focused on the relationships between baseline pupil size and WMC and attentional control. The sample used for this analysis consisted of 102 participants (the Groningen data set), because only these participants performed the LNS and attentional-control tasks.

The correlations between average pupil size and scores on the LNS and attentional control tasks were not statistically significant ($r = 0.05$, $p = 0.65$ and $r = 0.11$, $p = 0.29$, respectively; Figure 6a; Figure 7a). Bayesian correlations showed support for the null hypothesis for both WMC ($BF_{01} = 7.26$) and attentional control ($BF_{01} = 4.58$).

Similarly, the correlations between the CoV of pupil size and scores on the LNS and attentional-control tasks were not significant ($r = 0.11$, $p = 0.29$ and $r = -0.06$, $p = 0.57$, respectively; Figure 6b; Figure 7b). Bayesian correlations showed support for the null hypothesis for both WMC (BF_{01} = 4.64) and attentional control ($BF_{01} = 6.88$).

To investigate the influence of the measured control variables a Bayesian linear regression analysis was performed. The included covariates were the same as above (with score on the matrix task replaced by scores on the LNS and attentional-control tasks), as well as alcohol consumption, medication use, handedness, sleep, time of day, and whether the participant wore glasses or contact lenses.

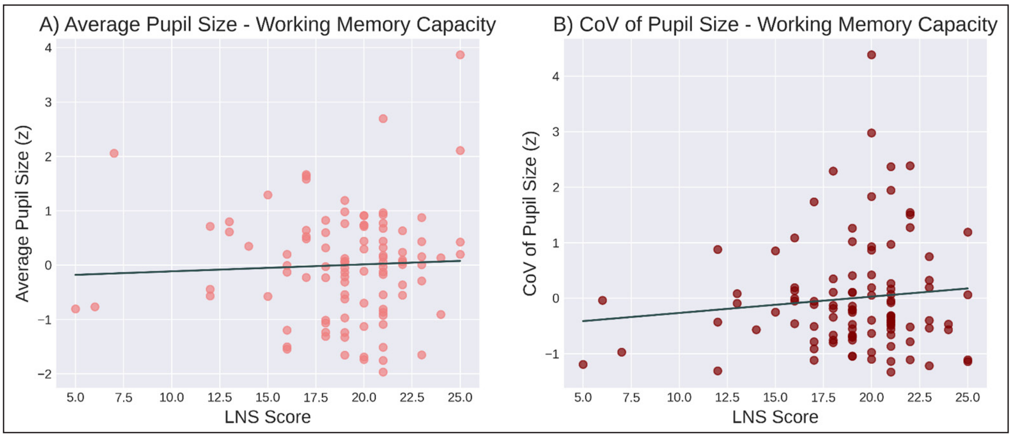

Figure 6 The relationship between baseline pupil size and working memory capacity (as assessed with a Letter-Number-Sequencing task). **A)** Average pupil size and WMC. **B)** Variability in pupil size and WMC.

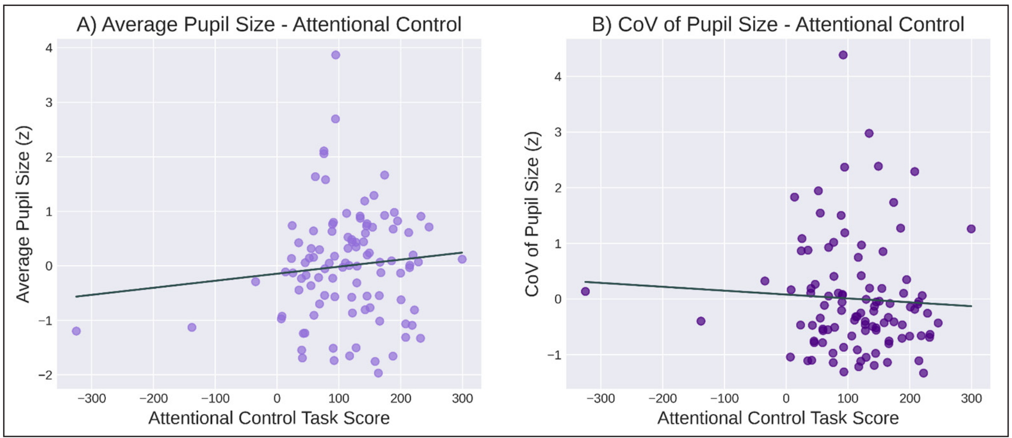

Figure 7 The relationship between baseline pupil size and attentional control (as assessed with an attentional-capture task. **A)** Average pupil size and attentional control. **B)** Variability in pupil size and attentional control.

With average pupil size as the dependent variable, the model that fit the data best included the control variable nicotine consumption but neither of the tasks. Nicotine consumption was significantly related to average pupil size in the Groningen sample (Figure 8; $r = -0.31$, $p = < 0.05$). With CoV of pupil size as the dependent variable, the model that fit the data best was the null model.

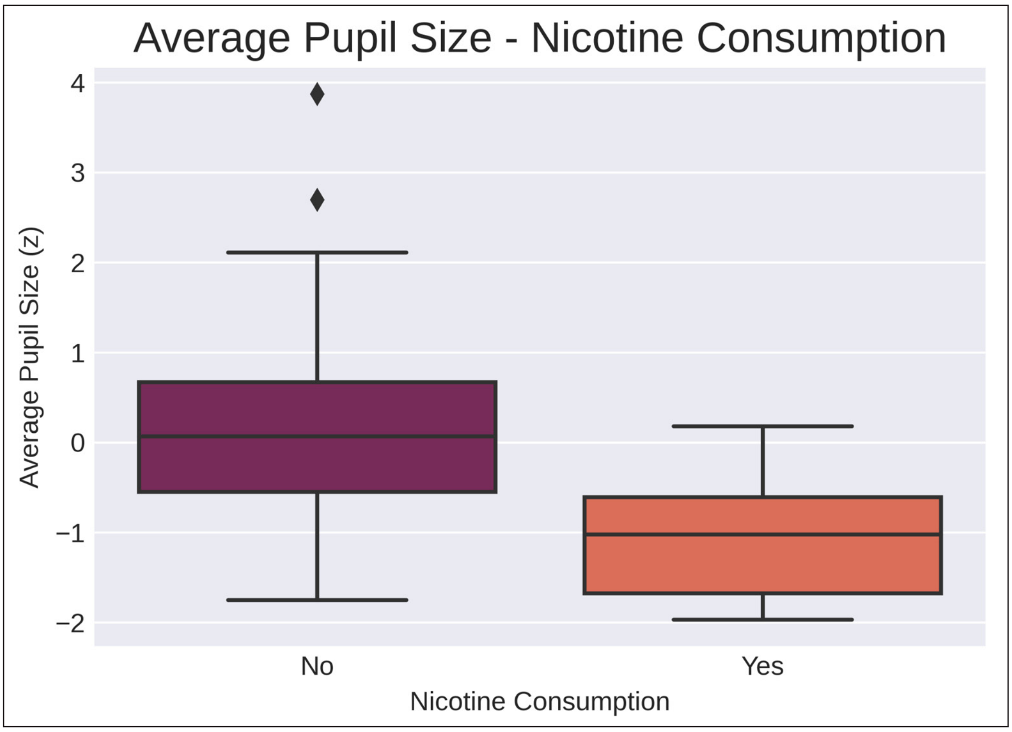

Figure 8 Distribution of pupil sizes across levels of nicotine consumption in the Groningen sample ("no" indicating no nicotine consumption on the day of the experiment (N = 94) and "yes" indicating nicotine consumed on the day of the experiment (N = 8)).

DISCUSSION

Here we have presented the results of two experiments that investigated the relationship between resting state pupil size and cognitive abilities, namely fluid intelligence, working memory capacity and attentional control. We did not find any significant relationships between pupil size (average or variability) and any of the measured constructs. In addition, Bayesian analyses provided support for the absence of these relationships in all analyses. Further, a Bayesian regression analysis indicated that nicotine consumption was a better predictor of average baseline pupil size than any of the task scores (see also Wardhani et al., 2020).

Crucially, we failed to replicate previous findings despite following several of the recommendations given by Tsukahara & Engle (2021b) to ensure sufficient measurement conditions for observing relationships between baseline pupil size and cognitive abilities. Tsukuhara and Engle (2021b) suggested that the relationship between pupil size and fluid intelligence is more robust than that between pupil size and other cognitive abilities, such as WMC (Tsukahara & Engle, 2021b; Unsworth et al., 2021). Therefore, and unlike many previous studies, we measured fluid intelligence in addition to WMC and attentional control, and focused our main analysis on fluid intelligence. However, despite taking the suggestions by Tsukuhara and Engle (2021b) to heart, we did not manage to replicate previous findings (see also Coors et al., 2022; Robison et al., 2022; Robison et al., 2023; Robison & Campbell, 2023).

Another possible problem in previous studies was that pupil size was measured in overly bright settings, leading to constricted pupils and reduced inter-individual variability in pupil size. In both of our experiments, participants fixated on a dark (black or gray) background during the baseline pupil measurement. This ensures a sufficiently large average pupil size, and also sufficient inter-individual variability. Further, participants in Groningen fixated on the eye tracker rather than the screen; this gives a frontal view of the eye, giving the best pupil recording. After data collection was finished we also measured the illuminance at the chinrest where the participants eyes would have been situated. In Groningen we measured approximately 130 lux and in Oslo 42 lux. While these values are brighter than the ideal recommended by Tsukahara and Engle (2021b), this does not seem to have a major detrimental impact on the inter-individual variability in pupil sizes. The mean and standard deviation of pupil sizes recorded in our Oslo sample were 4.39 mm and 0.72, respectively. This falls between the values reported by Tsukahara et al. (2016; mean 5.92 mm, standard deviation 1.09) and Unsworth et al. (2019; mean 3.21, standard deviation 0.49). For the Groningen sample we unfortunately do not have this information since we were not able to convert the arbitrary measurement units into mm. However, as mentioned above, we were able to find relationships between pupil size and one of the control variables (nicotine consumption) and (for the Groningen sample) observed reliable and stable individual differences in pupil size as recorded before and after the tasks; together, this suggests that the pupil measures were sensitive enough to pick up individual differences.

In addition to sufficiently dark background brightness, a notable strength of our measurement conditions was using a chin rest to stabilize the participant's head and to keep the eye-to-monitor distance fixed across participants. In contrast, Tsukahara et al. (2016) did not use a chin rest, which may influence recorded pupil sizes. For instance, highly motivated participants might sit closer to the eye tracker, leading to a larger average pupil size and more inter-individual variability.

Considering the recommendations by Tsukahara and Engle (2021b), there were some limitations in the tasks employed in this study, such as using only a single task to measure each construct. Different measures may tap into different aspects of the underlying construct, making it advisable to use multiple measures and derive the common variance across them for a more precise measurement (Ackerman & Hambrick, 2020). However, it seems unlikely that the specific task used influences results if the underlying correlation is robust. In addition, the WMC and fluid-intelligence tasks employed in our experiments were the same or similar to those used in previous studies (Aminihajibashi et al., 2019; Tsukahara et al., 2016), further supporting the conclusion that the observed results are not a consequence of the tasks used here. Other studies have also failed to replicate the link between baseline pupil size and fluid intelligence despite using multiple tasks (Robison & Campbell, 2023).

We did not observe significant correlations between the three measures as would be expected based on previous research (Unsworth et al., 2014). This is especially notable given that two of the tasks (LNS and Matrix task) had decent internal consistency and split-half reliability, respectively. In part, this may reflect that the tasks do not tap into the exact same cognitive constructs (Kane et al., 2005), although it is still surprising that they did not correlate at all. For the third task, the attentional-control task, the split-half reliability was relatively low (0.29). While singleton-search tasks (and other RT based tasks) are generally robust and often employed in attentional research in experimental settings, there is controversy over their reliability as measures of individual differences (Draheim et al., 2019; Hedge et al., 2018). The reliability of a task places limits on its ability to correlate with other measures (Spearman, 1904), which may explain why the attentional control task scores did not significantly correlate with the matrix task and LNS task scores, as would be expected (Unsworth et al., 2014). Similarly, the reliability of the task limits its ability to correlate with pupil size, potentially explaining our null result. On the other hand, other studies that have measured attentional control with non-RT based tasks (e.g., antisaccade tasks) have also failed to show a relationship with baseline pupil size (Robison & Brewer, 2022; Robison et al., 2022). Future research examining relationships between pupil size and attentional control should consider several ways of measuring attentional control to clarify the discrepancies (see Draheim et al., 2019 and Draheim et al., 2021 for more in depth discussion and recommendations).

Aside from pupil measurement conditions and task used, another concern raised about previous studies is their lack of diversity, as most samples have consisted mainly of North American college students, with relatively few exceptions (Aminihajibashi et al., 2019; Coors et al., 2022; Robison et al., 2022; Robison et al., 2023). Here, we believe we have achieved a more diverse sample through the combination of two populations. Some participants were recruited through social media, resulting in a more ability-diverse subsample (with a mean age of 25.56 years) than the typical undergraduate student sample. Furthermore, the majority of our participants came from various European countries, which could introduce differences in latent variables such as motivation or arousal level in testing situations when compared to the typical North American college student sample (for a more in-depth discussion, see Aminihajibashi et al., 2019).

As previously mentioned, it has been suggested that failures to replicate previous findings linking baseline pupil size to individual differences in cognitive ability have been largely due to improper measurement conditions and other methodological issues (Tsukahara & Engle, 2021b). As such, we have also focused here on discussing the methodological merits and limitations of our study. However, in theoretical terms, our results provide another data point in the discussion on the role of the LC-NE system in individual differences in cognitive ability (Tsukahara & Engle, 2021a; Unsworth and Robison, 2017). Here, we have failed to provide evidence for the theory positing that individual differences in cognitive ability arise from differences in the functional organization of the resting-state brain, manifested as differences in resting-state baseline pupil size (Tsukahara & Engle, 2021a).

In conclusion, we do not find evidence for a relationship between baseline pupil size (average or variability) and cognitive abilities (fluid intelligence, WMC, and attentional control), despite rigorous pupil-size measurement conditions and a sufficient (though modest compared to some other studies) sample size. These results suggest that the relationship between baseline pupil size and cognitive abilities is likely to be weak or non-existent.

APPENDIX

Appendix A Letter-Number-Sequencing items and associated response times.

STRING	RESPONSE TIME
2-B	7s
D-1	7s
4-C	7s
E-5	7s
3-A	7s
C-1	7s
5-C-A	7s
F-E-1	7s
3-2-A	7s
1-G-7	7s
H-9-4	7s

(Contd.)

STRING	RESPONSE TIME
3-Q-7	7s
Z-8-N	7s
M-6-U	7s
P-2-N	7s
V-1-J-5	10s
7-X-4-G	10s
S-9-T-6	10s
8-E-6-F-1	12s
K-4-C-2-S	12s
5-Q-3-H-6	12s
M-4-P-7-R-2	14s
6-N-9-J-2-S	14s
U-6-H-5-F-3	14s
R-7-V-4-Y-8-F	15s
9-X-2-J-3-N-7	15s
M-1-Q-8-R-4-D	15s
6-P-7-S-2-N-9-A	16s
U-1-R-9-X-4-K-3	16s
7-M-2-T-6-F-9-A	16s

DATA ACCESSIBILITY STATEMENT

The data and analysis code are available on the OpenScience Framework (https://osf.io/vf9sb/). We did not pre register either study.

ETHICS AND CONSENT

This research was conducted in accordance with the Declaration of Helsinki on research involving human subjects. All participants provided written informed consent prior to participation. Both studies were approved by the local ethics review board (study approval code Groningen: PSY1920-S-0178; study approval code Oslo: 1439337).

FUNDING INFORMATION

This research was partly funded by the Dutch Research Council (NWO; Grant VI.Vidi.191.045) and partly by the University of Oslo.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Veera Ruuskanen orcid.org/0000-0002-9959-5117
Department of Psychology, University of Groningen, The Netherlands

Thomas Hagen orcid.org/0000-0003-1481-389X
Department of Psychology, University of Oslo, Norway

Thomas Espeseth orcid.org/0000-0001-9076-6365
Department of Psychology, University of Oslo, Norway; Department of Psychology, Oslo New University College, Norway

Sebastiaan Mathôt orcid.org/0000-0003-3379-8584
Department of Psychology, University of Groningen, The Netherlands

REFERENCES

Ackerman, P. L., & **Hambrick, D. Z.** (2020). A primer on assessing intelligence in laboratory studies. *Intelligence*, *80*, 101440. DOI: https://doi.org/10.1016/j.intell.2020.101440

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & **Laeng, B.** (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4), 1. DOI: https://doi.org/10.1167/14.4.1

Aminihajibashi, S., Hagen, T., Foldal, M. D., Laeng, B., & **Espeseth, T.** (2019). Individual differences in resting-state pupil size: Evidence for association between working memory capacity and pupil size variability. *International Journal of Psychophysiology*, *140*, 1–7. DOI: https://doi.org/10.1016/j.ijpsycho.2019.03.007

Aston-Jones, G., & **Cohen, J. D.** (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. DOI: https://doi.org/10.1146/annurev.neuro.28.061604.135709

Aston-Jones, G., Rajkowski, J., & **Cohen, J.** (1999). Role of locus coeruleus in attention and behavioral flexibility. *Biological Psychiatry*, *46*(9), 1309–1320. DOI: https://doi.org/10.1016/S0006-3223(99)00140-7

Baddeley, A. D., & **Hitch, G.** (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation, 8*, 47–89. Academic Press. DOI: https://doi.org/10.1016/S0079-7421(08)60452-1

Birren, J. E., Casperson, R. C., & **Botwinick, J.** (1950). Age Changes in Pupil Size. *Journal of Gerontology*, *5*(3), 216–221. DOI: https://doi.org/10.1093/geronj/5.3.216

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1–22. DOI: https://doi.org/10.1037/h0046743

Chandler, D. J., Gao, W.-J., & **Waterhouse, B. D.** (2014). Heterogeneous organization of the locus coeruleus projections to prefrontal and motor cortices. *Proceedings of the National Academy of Sciences of the United States of America, 111*(18), 6816–6821. DOI: https://doi.org/10.1073/pnas.1320827111

Coors, A., Breteler, M. M. B., & **Ettinger, U.** (2022). Processing speed, but not working memory or global cognition, is associated with pupil diameter during fixation. *Psychophysiology*, *59*(11), e14089. DOI: https://doi.org/10.1111/psyp.14089

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163–191. DOI: https://doi.org/10.1037/0033-2909.104.2.163

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in psychology, 5*, 781. DOI: https://doi.org/10.3389/fpsyg.2014.00781

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. DOI: https://doi.org/10.1037/bul0000192

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & **Engle, R. W.** (2021). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology: General*, *150*(2), 242–275. DOI: https://doi.org/10.1037/xge0000783

Fukuda, K., Vogel, E., Mayr, U., & **Awh, E.** (2010). Quantity, not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, *17*(5), 673–679. DOI: https://doi.org/10.3758/17.5.673

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika, 10*, 255–282. DOI: https://doi.org/10.1007/BF02288892

Hedge, C., Powell, G., & **Sumner, P.** (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior research methods*, *50*, 1166–1186. DOI: https://doi.org/10.3758/s13428-017-0935-1

Heitz, R. P., Schrock, J. C., Payne, T. W., & **Engle, R. W.** (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology*, *45*(1), 119–129. DOI: https://doi.org/10.1111/j.1469-8986.2007.00605.x

Hess, E. H., & **Polt, J. M.** (1960). Pupil size as related to interest value of visual stimuli. *Science*, *132*, 349–350. DOI: https://doi.org/10.1126/science.132.3423.349

Hess, E. H., & **Polt, J. M.** (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(Whole No. 3611), 1190–1192. DOI: https://doi.org/10.1126/science.143.3611.1190

Heydasch, T. (2014). *The Hagen Matrices Test (HMT)*. DOI: https://doi.org/10.13140/RG.2.2.31433.75361

Joshi, S., Li, Y., Kalwani, R. M., & **Gold, J. I.** (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–234. DOI: https://doi.org/10.1016/j.neuron.2015.11.028

Kahneman, D., & **Beatty, J.** (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583–1585. DOI: https://doi.org/10.1126/science.154.3756.1583

Kane, M. J., Hambrick, D. Z., & **Conway, A. R. A.** (2005). Working Memory Capacity and Fluid Intelligence Are Strongly Related Constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 66–71. DOI: https://doi.org/10.1037/0033-2909.131.1.66

Keene, P. A., deBettencourt, M. T., Awh, E., & **Vogel, E. K.** (2022). Pupillometry signatures of sustained attention and working memory. *Attention, Perception, & Psychophysics*, *84*(8), 2472–2482. DOI: https://doi.org/10.3758/s13414-022-02557-5

Laeng, B., Sirois, S., & **Gredebäck, G.** (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. DOI: https://doi.org/10.1177/1745691611427305

Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, *1*(1). DOI: https://doi.org/10.5334/joc.18

Mathôt, S., Schreij, D., & **Theeuwes, J.** (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314–324. DOI: https://doi.org/10.3758/s13428-011-0168-7

Mathôt, S., & **Vilotijević, A.** (2022). Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis. *Behavior Research Methods*. DOI: https://doi.org/10.3758/s13428-022-01957-7

Mielicki, M. K., Koppel, R. H., Valencia, G., & **Wiley, J.** (2018). Measuring working memory capacity with the letter–number sequencing task: Advantages of visual administration. *Applied Cognitive Psychology*, *32*(6), 805–814. DOI: https://doi.org/10.1002/acp.3468

Miller, G. A. (1994). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *101*(2), 343–352. DOI: https://doi.org/10.1037/0033-295X.101.2.343

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & **Balsters, J. H.** (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, *35*(8), 4140–4154. DOI: https://doi.org/10.1002/hbm.22466

Murphy, P. R., Robertson, I. H., Balsters, J. H., & **O'connell, R. G.** (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology*, *48*(11), 1532–1543. DOI: https://doi.org/10.1111/j.1469-8986.2011.01226.x

Raven, J., Raven, J. C., & **Court, J. H.** (2000). *Manual for Raven's progressive matrices and vocabulary scales. Section 3, Standard progressive matrices (including the Parallel and Plus versions)*. OPP Ltd.

Robison, M. K., & **Brewer, G. A.** (2022). Individual differences in working memory capacity, attention control, fluid intelligence, and pupillary measures of arousal. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(9), 1296–1310. DOI: https://doi.org/10.1037/xlm0001125

Robison, M. K., & **Campbell, S.** (2023). Baseline pupil diameter does not correlate with fluid intelligence. *Psychonomic Bulletin & Review*. DOI: https://doi.org/10.3758/s13423-023-02273-7

Robison, M. K., Coyne, J. T., Sibley, C., Brown, N. L., Neilson, B., & **Foroughi, C.** (2022). An examination of relations between baseline pupil measures and cognitive abilities. *Psychophysiology*, *59*(12). DOI: https://doi.org/10.1111/psyp.14124

Robison, M. K., Ralph, K. J., Gondoli, D. M., Torres, A., Campbell, S., Brewer, G. A., & **Gibson, B. S.** (2023). Testing locus coeruleus-norepinephrine accounts of working memory, attention control, and fluid intelligence. *Cognitive, Affective, & Behavioral Neuroscience, 1–45*. DOI: https://doi.org/10.3758/s13415-023-01096-2

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, *15*, 72–101. DOI: https://doi.org/10.2307/1412159

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29. DOI: https://doi.org/10.1037/h0093759

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*, 599–606 DOI: https://doi.org/10.3758/BF03211656

TheEyeTribe. (n.d.). *The eye tribe tracker*. Retrieved from https://theeyetribe.com/

Tryon, W. W. (1975). Pupillometry: A Survey of Sources of Variation. *Psychophysiology*, *12*(1), 90–93. DOI: https://doi.org/10.1111/j.1469-8986.1975.tb03068.x

Tsukahara, J. S., & **Engle, R. W.** (2021a). Fluid intelligence and the locus coeruleus–norepinephrine system. *Proceedings of the National Academy of Sciences*, *118*(46), e2110630118. DOI: https://doi.org/10.1073/pnas.2110630118

Tsukahara, J. S., & **Engle, R. W.** (2021b). Is baseline pupil size related to cognitive ability? Yes (under proper lighting conditions). *Cognition*, *211*. DOI: https://doi.org/10.1016/j.cognition.2021.104643

Tsukahara, J. S., Harrison, T. L., & **Engle, R. W.** (2016). The relationship between baseline pupil size and intelligence. *Cognitive Psychology*, *91*, 109–123. DOI: https://doi.org/10.1016/j.cogpsych.2016.10.001

Unsworth, N., Fukuda, K., Awh, E., & **Vogel, E. K.** (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26. DOI: https://doi.org/10.1016/j.cogpsych.2014.01.003

Unsworth, N., Miller, A. L., & **Robison, M. K.** (2021). Is working memory capacity related to baseline pupil diameter? *Psychonomic Bulletin & Review*, *28*, 228–237. DOI: https://doi.org/10.3758/s13423-020-01817-5

Unsworth, N., & **Robison, M. K.** (2017). The importance of arousal for variation in working memory capacity and attention control: A latent variable pupillometry study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(12), 1962–1987. DOI: https://doi.org/10.1037/xlm0000421

Unsworth, N., Robison, M. K., & **Miller, A. L.** (2019). Individual differences in baseline oculometrics: Examining variation in baseline pupil diameter, spontaneous eye blink rate, and fixation stability. *Cognitive, Affective & Behavioral Neuroscience*, *19*(4), 1074–1093. DOI: https://doi.org/10.3758/s13415-019-00709-z

van der Meer, E., Beyer, R., Horn, J., Foth, M., Bornemann, B., Ries, J., Kramer, J., Warmuth, E., Heekeren, H. R., & **Wartenburger, I.** (2010). Resource allocation and fluid intelligence: Insights from pupillometry. *Psychophysiology*, *47*(1), 158–169. DOI: https://doi.org/10.1111/j.1469-8986.2009.00884.x

Van Rossum, G., & **Drake, F. L.** (2009). Introduction to python 3: python documentation manual part 1. *CreateSpace*.

Vilotijević, A., & **Mathôt, S.** (2023). Functional benefits of cognitively driven pupil-size changes. *Wiley Interdisciplinary Reviews: Cognitive Science*, e1672. DOI: https://doi.org/10.1002/wcs.1672

Vinutha, H. P., Poornima, B., & **Sagar, B. M.** (2018). Detection of Outliers Using Interquartile Range Technique from Intrusion Dataset. In: S. Satapathy, J. Tavares, V. Bhateja & J. Mohanty (Eds.), *Information and Decision Sciences. Advances in Intelligent Systems and Computing*, *701*. Singapore: Springer. DOI: https://doi.org/10.1007/978-981-10-7563-6_53

Wardhani, I. K., Mathôt, S., Boehler, C. N., & **Laeng, B.** (2020). Effects of nicotine on pupil size and performance during multiple-object tracking in non-nicotine users. *International Journal of Psychophysiology*, *158*, 45–55. DOI: https://doi.org/10.1016/j.ijpsycho.2020.09.005

Wechsler, D. (1997). *WAIS-III: Administration and scoring manual: Wechsler Adult Intelligence Scale* (3rd ed). Psychological Corporation. DOI: https://doi.org/10.1037/t49755-000

TO CITE THIS ARTICLE:
Ruuskanen, V., Hagen, T., Espeseth, T., & Mathôt, S. (2024). Baseline Pupil Size Seems Unrelated to Fluid Intelligence, Working Memory Capacity, and Attentional Control. *Journal of Cognition,* 7(1): 41, pp. 1–18. DOI: https://doi.org/10.5334/joc.365

Submitted: 06 October 2023
Accepted: 22 April 2024
Published: 10 May 2024

Journal of Cognition is a peer-reviewed open access journal published by Ubiquity Press.