ORIGINAL ARTICLE

Processing speed, but not working memory or global cognition, is associated with pupil diameter during fixation

Annabell Coors[1] | Monique M. B. Breteler[1,2] | Ulrich Ettinger[3]

[1]Population Health Sciences, German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany

[2]Institute for Medical Biometry, Informatics and Epidemiology, Faculty of Medicine, University of Bonn, Bonn, Germany

[3]Department of Psychology, University of Bonn, Bonn, Germany

Correspondence
Monique M. B. Breteler, Population Health Sciences, German Center for Neurodegenerative Diseases (DZNE), Venusberg-Campus 1/99, 53127 Bonn, Germany.
Email: monique.breteler@dzne.de

Abstract

Mean pupil size during fixation has been suggested to reflect interindividual differences in working memory and fluid intelligence. However, due to small samples with limited age range (17–35 years) and suboptimal light conditions in previous studies, these associations are still controversial and it is unclear whether they are observed at older ages. Therefore, we assessed whether interindividual differences in cognitive performance are reflected in pupil diameter during fixation and whether these associations are age-dependent. We analyzed pupillometry and cognition data of 4560 individuals aged 30–95 years of the community-based Rhineland Study. Pupillometry data were extracted from a one-minute fixation task. The cognitive test battery included tests of oculomotor control, working memory, episodic verbal memory, processing speed, executive function, and crystallized intelligence. For data analysis, we used multivariable regression models. Working memory and global cognition were not associated with pupil diameter during fixation. Better processing speed performance was associated with larger pupil diameter during fixation. Associations between cognition and pupil diameter during fixation hardly varied with age, but pupil diameter during fixation declined linearly with age (adjusted decline: 0.33 mm per 10 years of age). There were no significant sex differences in pupil size. We conclude that interindividual differences in mean pupil diameter during fixation may partly reflect interindividual differences in the speed of processing and response generation. We could not confirm that interindividual differences in working memory and fluid intelligence are reflected in pupil size during fixation; however, our sample differed in age range from previous studies.

KEYWORDS
age, cognition, cohort studies, intelligence, pupillometry

1 | INTRODUCTION

The pupil dilates and constricts in response to changes in lighting, known as the pupillary light reflex (Lowenstein & Loewenfeld, 1950). Non-luminance-driven pupil dilations have been linked to activity in superior colliculus and locus coeruleus (Joshi & Gold, 2020). The locus coeruleus is one of the first sites showing Alzheimer's disease pathology (Mather & Harley, 2016) and is the main source of norepinephrine in the central nervous system (Joshi & Gold, 2020). Norepinephrine is a major neuromodulator that mediates arousal (Joshi & Gold, 2020) and has been linked to cognitive performance (Wang et al., 2013). Specifically, pupil dilations have been associated with arousal (Joshi & Gold, 2020) and mental effort (van der Wel & van Steenbergen, 2018). However, it is controversial whether interindividual differences in cognitive abilities are reflected in pupil diameter during fixation (Tsukahara et al., 2016; Tsukahara & Engle, 2021; Unsworth et al., 2020). Evidence of a relationship between interindividual differences in cognitive abilities and pupil diameter during fixation would further support the importance of the locus coeruleus-norepinephrine system in cognitive performance (Joshi & Gold, 2020).

Some studies have reported that individuals with higher working memory capacity (Heitz et al., 2008; Tsukahara et al., 2016; Tsukahara & Engle, 2021) and higher fluid intelligence (Bornemann et al., 2010; Tsukahara et al., 2016; Tsukahara & Engle, 2021; van Der Meer et al., 2010) have larger pupil diameter during a passive baseline condition than individuals with lower working memory capacity and fluid intelligence. However, a recent study and meta-analysis found no association between working memory capacity and baseline pupil diameter (Unsworth et al., 2020). Only three of the 26 studies included in the meta-analysis reported effect sizes that were statistically significantly different from zero for the association between working memory capacity and mean pupil size during fixation, with one study reporting a negative association (Sibley et al., 2018) and two studies reporting a positive association (Tsukahara et al., 2016; Tsukahara & Engle, 2021). Since the variability in effect sizes in the meta-analysis was large, the authors performed a moderator analysis and concluded that the primary moderator for heterogeneity across studies was the laboratory in which the study was conducted (laboratory of Tsukahara versus other) (Unsworth et al., 2020).

In response, Tsukahara and Engle (2021) argued that one major reason for negative findings from other laboratories may have been sub-optimal experimental light conditions. They argued that because of the pupillary light reflex (Lowenstein & Loewenfeld, 1950), bright experimental light conditions reduce the variance in pupil diameter, which in turn reduces the chances to detect statistically significant associations with cognitive performance across individuals (Tsukahara & Engle, 2021). They then systematically varied the light conditions in their own experiments and found that a reduction of the mean-variance in pupil diameter to a similar level as in the study by Unsworth et al. (2020) completely eliminated the association between working memory capacity and pupil diameter and decreased the association with fluid intelligence (Tsukahara & Engle, 2021).

However, Unsworth et al. (2021) reanalyzed the data of Tsukahara and Engle (2021) and found that controlling for the additional confounder race also eliminated the association between fluid intelligence and baseline pupil diameter. This finding is in line with previous studies that reported differences in baseline pupil size between races (Tsukahara et al., 2016). In response to the re-analysis of their data, Tsukahara et al. (2021) combined the data of their different samples ($N = 831$), controlled for age and race and reported that there is an association between fluid intelligence and pupil diameter.

Given these controversies and given that the meta-analysis was only based on 4356 persons in total with more than half of the studies originating from two laboratories, the association between cognitive performance and pupil diameter during fixation needs further investigation in a large, independent sample, and under comparable experimental conditions as in the study by Tsukahara et al. (2016). According to the findings reported in the study by Tsukahara et al. (2016) and their follow-up studies (Tsukahara et al., 2021; Tsukahara & Engle, 2021), the chances of detecting associations between cognition and mean pupil size during fixation can be maximized by measuring each cognitive domain with multiple tests, and by ensuring that the experimental room is dark enough to allow for large variation in pupil size across participants (standard deviation across participants in Tsukahara et al. (2016): 1.1 mm). Further, it should be controlled for interindividual differences in genetic background and age as these represent important confounders (Unsworth et al., 2021).

Linear age-related decline in pupil size and more restricted range in pupil dilation have also been reported in older adults as a result of a degenerating dilation muscle in the iris (Van Gerven et al., 2004). If age also reduced interindividual variation in pupil size, then the chances to detect statistically significant associations between cognitive performance and pupil diameter during fixation across individuals might be extremely low in older age (Tsukahara & Engle, 2021). As all studies included in the meta-analysis by Unsworth et al. (2020) and the additional studies cited before were conducted in individuals aged 17 to 35 years, it remains an open question whether associations between cognitive performance and pupil diameter during fixation are age-dependent.

We therefore investigated, in a large population-based cohort under experimental conditions deemed optimal (Tsukahara et al., 2016; Tsukahara & Engle, 2021), whether interindividual differences in cognitive performance (including global cognition and working memory performance) are reflected in pupil diameter during fixation and whether these associations are age-dependent.

2 | METHOD

2.1 | Participants

We used data from the first 5000 participants of the Rhineland Study, a community-based cohort study in Bonn, Germany. Inclusion criteria are living in one of two geographically defined areas in Bonn and being at least 30 years or older. Participation is only possible upon invitation and there are no financial incentives. The only exclusion criterion is not having sufficient comprehension of the German language to provide written informed consent. Eligibility is irrespective of health status. The ethics committee of the Medical Faculty of the University of Bonn approved the study, which was carried out in accordance with the recommendations of the International Council for Harmonisation Good Clinical Practice standards (ICH-GCP).

Participants underwent 8 h of examinations, including a one-hour cognitive and eye movement test battery. The eye movement tests included fixation, smooth pursuit, prosaccade, and antisaccade tasks (Coors et al., 2021). Pupillometry data were taken from the fixation task (see below) and were available in 4568 participants out of the first 5000 participants. Missing values ($N = 432$) were primarily due to technical issues during data acquisition and post-processing (71.1%), exclusion after visual inspection of data (14.8%), contraindications (8.1%), non-compliance (5.1%), refusal (0.7%), or timeout (0.2%). An additional 8 participants did not have pupil data available for all fixation positions and were therefore excluded. This left 4560 participants for our analysis.

2.2 | Pupillometry data

We recorded eye movements using video-based infrared oculography (EyeLink 1000 and EyeLink 1000 Plus; SR Research Ltd) at 1000 Hz. We extracted the mean pupil diameter during a one-minute fixation task. In the fixation task, participants had to fixate the target at the center ($x = 0°$, $y = 0°$), the left ($x = −9.63°$, $y = 0°$), the right ($x = 9.63°$, $y = 0°$), the top ($x = 0°$, $y = 9.63°$) and the bottom ($x = 0°$, $y = −9.63°$). The order within which the target was presented in these positions was randomized across

participants but eccentric locations were always followed by the central location. The central position thus had to be fixated four times in total. The target appeared at each eccentric location for 10 s and at the central location for 5 s each time. The target was a white (RGB 255, 255, 255) circle 0.35° in diameter presented on black (0, 0, 0) background. A PCE-172 light meter (PCE Instruments, Meschede, Germany) was used to measure luminance. During the examination, the overall room luminance value was about 1 Lux (light meter placed on the table at a distance of 70 cm from the monitor screen, faced up) and the screen luminance value was about 4 Lux (light meter placed at a distance of about 1 cm from the fixation point displayed on the monitor screen, faced toward the screen). A detailed description of oculomotor data acquisition can be found in Coors et al. (2021).

Pupil diameter values typically vary between 1.5 and 8 mm (McDougal & Gamlin, 2008) with more extreme values possibly representing measurement artifacts. Based on a visual inspection of a plot of pupil diameter values against age, we decided to consider all participants with pupil diameter values above 9 mm as outliers ($N = 6$, Figure 1). We performed all analyses including and excluding outlier values and compared the results.

Pupil diameter values were highly correlated across fixation positions (Pearson $r \geq 0.97$), suggesting high internal consistency of measurement. Therefore, we calculated the mean pupil diameter across all positions and used this value for subsequent analyses.

2.3 | Cognitive tests

The cognitive test battery measured working memory, episodic verbal memory, processing speed, executive function, and crystallized intelligence. We assessed working memory with the orally performed Digit Span forward (sequence length 3 to 9) and backward (sequence length 2 to 9) tasks and the touchpad-based Corsi block-tapping test (Corsi) forward and backward (sequence lengths 2 to 9), adapted from the PEBL battery (Mueller & Piper, 2014). In both tasks, the sequence was increased until the participant made two errors within one sequence. The length of the last successfully completed sequence was then taken as a measure of the working memory span. Episodic verbal memory was assessed with the Auditory Verbal Learning and Memory test (AVLT) with a list length of 15 words (immediate recall: sum of correctly recalled nouns in the first five trials out of 75 words; delayed recall: number of correctly recalled words out of 15 words after a time delay of 20–30 min) (Boenniger et al., 2021). Results on a numbers-only trail-making test (TMT-A: time to connect 24 digits that were randomly distributed on a computer screen in ascending order) and prosaccade latency (time needed to initiate a saccade after
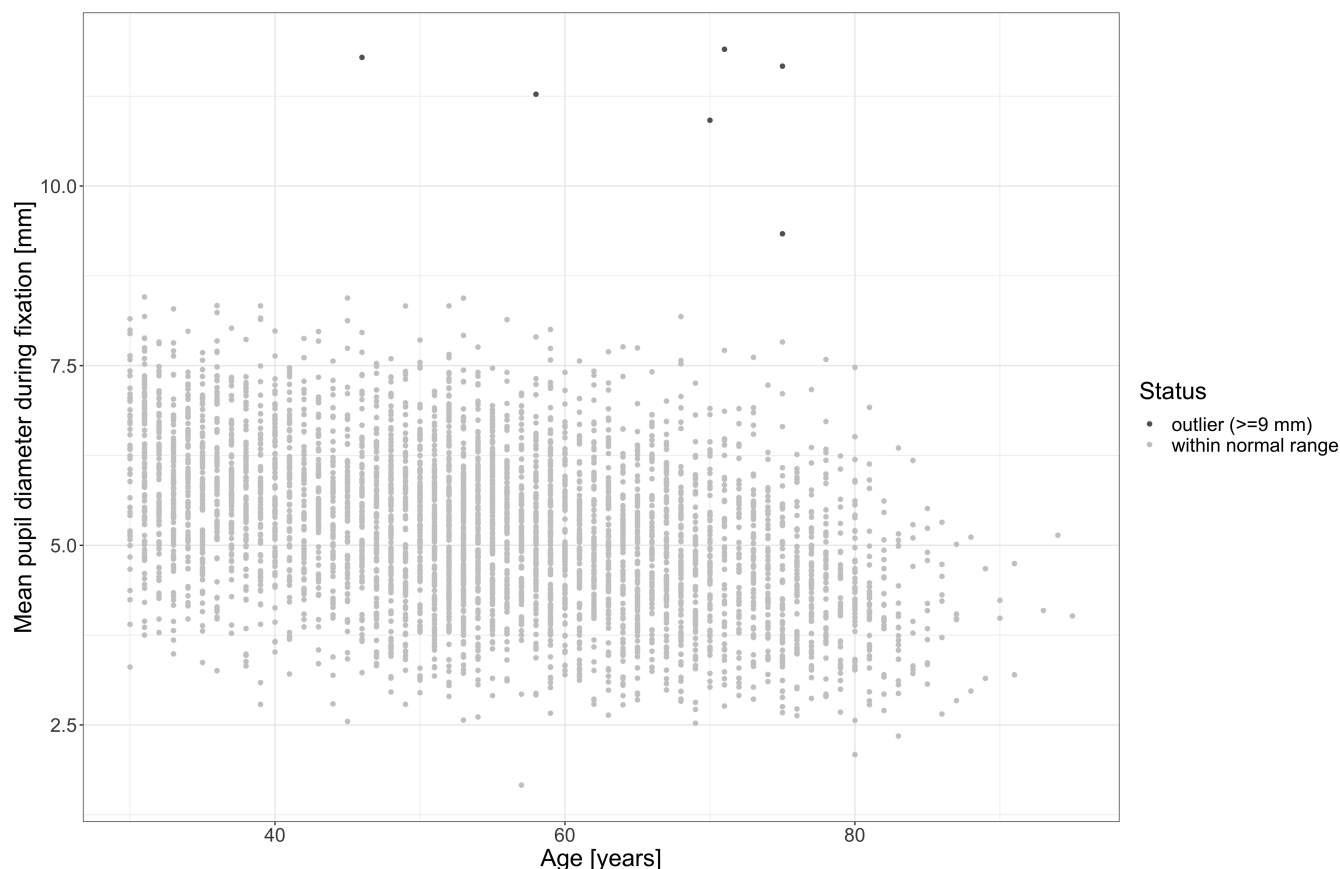
FIGURE 1 Mean pupil diameter during fixation across age – Outlier determination

the target has stepped to one side) were combined to assess processing speed. We assessed executive function with a 60s categorical word fluency task (number of uniquely named animals), a number-and-letters trail-making test (TMT-B: time to connect 12 digits and 12 letters in ascending order and in an alternating fashion [1-A-2-B]), and antisaccade error rate (percentage of trials in which the initial saccade was made toward the target instead of the opposite direction). Crystallized intelligence was measured with the 37-item Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B), which is a vocabulary test in which participants select an existing German word among four non-words in each of 37 trials (Lehrl, 2005).

We calculated domain scores based on averaged z-scores for the separate test scores in that domain. The assignment of cognitive tests to domain scores was based on previous literature. Finally, we averaged the domain scores for working memory, episodic verbal memory, processing speed, and executive function to obtain a global cognition score. Individuals whose native language was not German or those with severe cognitive impairment as a result of traumatic brain injury or other non-aging-related diseases did not contribute to the mean and standard deviation statistics used in the standardization process. However, we computed

z-scores for those participants based on the mean and standard deviation of the remaining sample, except for the crystallized intelligence score where German as a native language is a test requirement.

2.4 | Genetic ancestry

We used genome-wide single-nucleotide polymorphisms (SNP) arrays (Illumina Omni-2.5 exome array) to determine heterogeneity in genetic background. Genotype data were processed with GenomeStudio (version 2.0.5), and quality-controlled using PLINK (version 1.9) (Purcell et al., 2007). We calculated six principal components using EIGENSTRAT (version 16000) (Price et al., 2006) to represent differences in genetic background.

2.5 | Statistical analyses

To quantify associations between cognitive scores and pupil diameter during fixation, we calculated a separate regression model for each cognitive score that included the cognitive score as independent variable, pupil diameter during fixation as dependent variable, and age, sex,

best-corrected visual acuity, educational level, and six principal components to correct for population stratification as additional variables to adjust for. In addition, we added native language as additional variable to all models except for the two models that included prosaccade latency and antisaccade error rate as independent variables. To evaluate whether associations between cognitive performance and pupil diameter during fixation vary with age, we ran an additional model for each cognitive predictor that included an age*cognition interaction term in addition to the covariates mentioned above. In addition, we performed age-stratified analysis for the strata 30- to 49-year-olds, 50- to 60-year-olds, and 61+-year-olds. These age cut-offs were chosen to yield approximately equal stratum sizes, and to reflect age ranges with little or no age-related cognitive decline (30–49 years), an intermediate age range (50–60 years), and an age range where some participants might start showing age-related cognitive decline (≥61 years).

We then quantified how pupil diameter during fixation differs across age and between men and women using a linear regression model that included age and sex as independent variables, pupil diameter during fixation as dependent variable, and best-corrected visual acuity, educational level, and population stratification as covariates.

To make our results on the associations between cognitive performance and baseline pupil diameter comparable to previous studies, we additionally calculated zero-order Pearson correlations for the whole sample and for the three aforementioned age strata. For the correlation analyses, we used participants without missing values in any of the cognitive scores ($N = 3702$).

Statistical analyses were carried out in RStudio (version 1.3.959, R-base version 4.0.3). Missing covariate data in the linear regression models (<1% missing values for visual acuity, education and native language, and 16.2% missing values for population stratification) were imputed using predictive mean matching (Hmisc package, 10 bootstrap replicates). We present standardized point estimates of association with 95% confidence intervals. We did not correct for multiple testing as we conducted an exploratory analysis on the associations between cognitive and eye movement measures and mean pupil size during fixation (Althouse, 2016). Thus, only additional confirmatory studies can rule out the possibility of false discoveries (Althouse, 2016).

3 | RESULTS

3.1 | Study sample

Table 1 gives descriptive characteristics of the study sample. Our sample was on average highly educated and had

TABLE 1 Sample characteristics

Number of participants, N (%)	4560 (100)
30–49 years	1656 (36.3)
50–60 years	1346 (29.5)
61+ years	1558 (34.2)
Age, M (SD) in years	54.7 (13.9)
Sex, N (%) women	2604 (57.1)
Education level, N	4519
High	2377 (52.6)
Middle	2052 (45.4)
Low	90 (2.0)
Best-corrected visual acuity, N	4535
High (≥0.8)	3931 (86.7)
Middle (0.32–0.63)	571 (12.6)
Low (<0.32)	33 (0.7)
Mean pupil size during fixation, M (SD)	5.2 (1.1)
30- to 49-year-olds	5.7 (1.0)
50- to 60-year-olds	5.2 (1.0)
61+-year-olds	4.7 (1.0)
Digit span forward [number of digits], M (SD), max = 9	6.3 (1.2)
Digit span backward [number of digits], M (SD), max = 9	4.8 (1.2)
Corsi forward [number of blocks], M (SD), max = 9	4.9 (1.1)
Corsi backward [number of blocks], M (SD), max = 9	4.8 (1.0)
AVLT - immediate recall [sum of recalled words over recall 1 to 5], M (SD), max = 75	51.3 (10.0)
AVLT - delayed recall [number of recalled words], M (SD), max = 15	10.3 (3.3)
Trail-making test A [completion time in s], median [IQR]	32.9 (26.8, 41.7)
Mean prosaccade latency [ms], median [IQR]	185.8 (170.6, 205.5)
Trail-making test B [completion time in s], median [IQR]	43.5 (33.5, 60.3)
Word fluency [number of unique words], M (SD)	26.0 (7.0)
Antisaccade error rate [%], M (SD)	31.9 (23.9)
MWT-B [sum of correctly recognized words], M (SD), max = 37	30.5 (3.4)

Note: Education level was determined using the International Standard Classification of Education 2011 (ISCED) and was coded as low (lower secondary education or below), middle (upper secondary education to undergraduate university level), and high (postgraduate university study). Assessment of best-corrected visual acuity was based on visual scores from the right eye and was measured using an automated refractometer (Ark-1 s, NIDEK CO., Tokyo, Japan). Categorization of the visual acuity values was based on the guidelines of the International Council of Ophthalmology.

Abbreviations: AVLT, Auditory Verbal Learning and Memory test; M, mean; MWT-B = Mehrfachwahl-Wortschatz-Intelligenztest; N, number of participants; SD, standard deviation.

TABLE 2 Pearson correlations between different cognitive domain and test scores

	Global cognition score	Working memory	Episodic verbal memory	Processing speed	Executive function	Crystallized intelligence	Digit span forward
Working memory	0.728						
Episodic verbal memory	0.791	0.375					
Processing speed	0.729	0.396	0.413				
Executive function	0.784	0.521	0.498	0.411			
Crystallized intelligence	0.014	0.056	0.044	−0.149	0.096		
Digit span forward	0.415	0.674	0.184	0.175	0.287	0.123	
Digit span backward	0.511	0.718	0.288	0.223	0.377	0.127	0.424
Corsi forward	0.537	0.681	0.271	0.346	0.388	−0.075	0.220
Corsi backward	0.535	0.667	0.286	0.344	0.376	−0.026	0.191
AVLT-immediate recall	0.779	0.397	0.954	0.410	0.503	0.054	0.221
AVLT-delayed recall	0.732	0.321	0.956	0.380	0.449	0.030	0.132
Trail-making test A	−0.655	−0.402	−0.392	−0.709	−0.502	0.034	−0.188
Prosaccade latency	−0.436	−0.176	−0.229	−0.801	−0.110	0.170	−0.053
Trail-making test B	−0.647	−0.461	−0.413	−0.461	−0.658	−0.004	−0.225
Word fluency	0.505	0.270	0.360	0.220	0.705	0.194	0.178
Antisaccade error rate	−0.447	−0.322	−0.253	−0.122	−0.710	−0.017	−0.164

Note: The cognitive test scores and the corresponding domain score as well as the correlations between them are shown in the same color. Crystallized intelligence was only measured with one single test. The global cognition score is a composite measure consisting of the domain scores for working memory, episodic verbal memory, processing speed, and executive function.

high best-corrected visual acuity. The mean pupil diameter during fixation was 5.2 mm with a standard deviation of 1.1 mm. In each age group, the standard deviation in pupil diameter was 1.0 mm.

3.2 | Associations between cognition and pupil diameter

Correlations between the cognitive test and domain scores are shown in Table 2. In Table 3, we present the regression results for the associations between cognitive performance and mean pupil size during fixation for the cognitive domain scores and in Table 4 for the individual cognitive test scores. Across the sample, better performance in the domain score for processing speed (Table 3) and its subtest prosaccade latency (Table 4) were associated with larger pupil diameter during fixation.

When we included an age*cognition-interaction term in the models, it did not become statistically significant for any model ($p \geq .071$). In the age-stratified analyses, the processing speed domain score was significantly associated with pupil size during fixation in the 50- to 60-year-olds and in the 61+-year-olds. Moreover, prosaccade latency was significantly associated with mean pupil diameter during fixation in all three strata (Table 5). Digit span forward performance was significantly associated with mean pupil size

during fixation in the 30–49 years old (p value $= .048$), but not in any of the other age strata (Table 5). Visualizations of the associations between cognitive domain scores and mean pupil diameter during fixation can be found in Figure 2.

Across the sample and before adjusting for covariates, some cognitive scores were correlated with baseline pupil diameter ($-0.25 \leq$ Pearson $r \leq 0.31$) (Table 6). However, when we age-stratified the correlation analyses, the correlations were largely diminished and very small ($r < 0.14$) for all cognitive scores except for the domain score for processing speed ($0.09 \leq r \leq 0.20$) and prosaccade latency ($-0.18 \leq r \leq -0.09$). For both cognitive scores, the strongest correlation emerged in the 61+-year-olds (Table 6).

Our sensitivity analysis revealed that the exclusion of the six participants with pupil diameter values above 9 mm did not materially change any of the results.

3.3 | Associations between age, sex, and pupil diameter

Pupil diameter during fixation declined linearly with increasing age (average adjusted decrease 0.33 [95% CI: −0.35 to −0.30] mm per 10 years of age). There were no sex differences in mean pupil diameter during fixation (adjusted difference between men and women −0.03 [95% CI: −0.09 to 0.03] mm).

COORS ET AL.

Digit span backward	Corsi forward	Corsi backward	AVLT-immediate recall	AVLT-delayed recall	Trail-making test A	Prosaccade latency	Trail-making test B	Word fluency
0.269								
0.263	0.386							
0.316	0.270	0.280						
0.236	0.247	0.267	0.825					
−0.265	−0.327	−0.324	−0.391	−0.358				
−0.055	−0.186	−0.191	−0.224	−0.212	0.215			
−0.311	−0.374	−0.355	−0.408	−0.381	0.481	0.221		
0.228	0.166	0.166	0.367	0.321	−0.253	−0.074	−0.246	
−0.211	−0.258	−0.251	−0.258	−0.225	0.287	−0.090	0.281	−0.205

TABLE 3 Associations between cognitive domain scores and mean pupil diameter during fixation

Cognitive predictor	β (95% CI)	p-value	Variance explained (%)
Global cognition score	0.060 (−0.004, 0.124)	.065	16.3
Working memory	0.020 (−0.025, 0.065)	.377	16.3
Episodic verbal memory	0.006 (−0.029, 0.041)	.755	16.5
Processing speed	**0.125 (0.080, 0.169)**	**<.001**	**17.2**
Executive function	−0.020 (−0.064, 0.023)	.360	16.6
Crystallized intelligence	0.014 (−0.017, 0.044)	.389	16.6

Note: The table displays the change in standard deviations in mean pupil diameter during fixation per one standard deviation increase in the cognitive domain score (β and 95% CI). Regression coefficients were obtained from the following multivariable linear regression model: Mean pupil diameter during fixation $\sim \beta_0 +$ cognitive domain score$*\beta_1 +$ age $+$ sex $+$ population stratification $+$ education $+$ visual acuity $+$ native language $+$ residual error. Variance explained refers to the adjusted variance explained by the total model including the cognitive predictor. The adjusted variance explained in the model without cognitive predictor was 16.8%. In bold are those associations with a p value below .05.

Abbreviations: 95% CI, 95% confidence interval; β, standardized regression coefficient.

4 | DISCUSSION

Of all cognitive measures, only processing speed, and prosaccade latency were associated with pupil diameter during fixation across the sample. Associations between cognitive performance and pupil diameter during fixation did not significantly vary with age, yet age-stratified analyses showed larger effect sizes for processing speed and prosaccade latency at higher age. Mean pupil diameter during fixation declined about 0.33 mm per 10 years of age. Men and women did not differ in mean pupil size during fixation.

We could not confirm that working memory capacity and fluid intelligence are associated with mean pupil size during fixation (Tsukahara et al., 2016; Tsukahara & Engle, 2021; van Der Meer et al., 2010). Our finding for working memory capacity is in line with the result of the meta-analysis by Unsworth et al. (2020). However, as experimental conditions may play a major role, we compared our experimental conditions with the ones in the study by Tsukahara et al. (2016), in which associations were found, and present the comparison in Table 7. Based on this comparison we can rule out that our null findings

TABLE 4 Associations between cognitive test scores and mean pupil diameter during fixation

Cognitive predictor	β (95% CI)	p-value	Variance explained (%)
Digit span forward	0.023 (−0.005, 0.051)	.108	16.4
Digit span backward	0.007 (−0.022, 0.035)	.642	16.3
Corsi forward	0.007 (−0.024, 0.038)	.660	15.8
Corsi backward	−0.001 (−0.031, 0.029)	.948	16.0
AVLT-immediate recall	0.002 (−0.032, 0.035)	.920	16.6
AVLT-delayed recall	0.008 (−0.024, 0.040)	.643	16.6
Trail-making test A	0.002 (−0.030, 0.034)	.911	16.6
Prosaccade latency	**−0.099 (−0.129, −0.069)**	**<.001**	**17.5**
Trail-making test B	0.008 (−0.024, 0.041)	.616	16.6
Word fluency	−0.028 (−0.058, 0.001)	.062	16.5
Antisaccade error rate	−0.009 (−0.038, 0.020)	.541	16.4

Note: The table displays the change in standard deviations in mean pupil diameter during fixation per one standard deviation increase in the cognitive test score (β and 95% CI). Regression coefficients were obtained from the following multivariable linear regression model: Mean pupil diameter during fixation ~ β_0 + cognitive test score*β_1 + age + sex + population stratification + education + visual acuity + (native language) + residual error. Native language was not included in the models with prosaccade latency and antisaccade error rate as predictors. Variance explained refers to the adjusted variance explained by the total model including the cognitive predictor. The adjusted variance explained in the model without cognitive predictor was 16.7% for the model excluding native language and 16.8% for the model including native language as covariate. In bold are those associations with a *p* value below .05.

Abbreviations: 95% CI, 95% confidence interval; AVLT, Auditory Verbal Learning and Memory test; β, standardized regression coefficient.
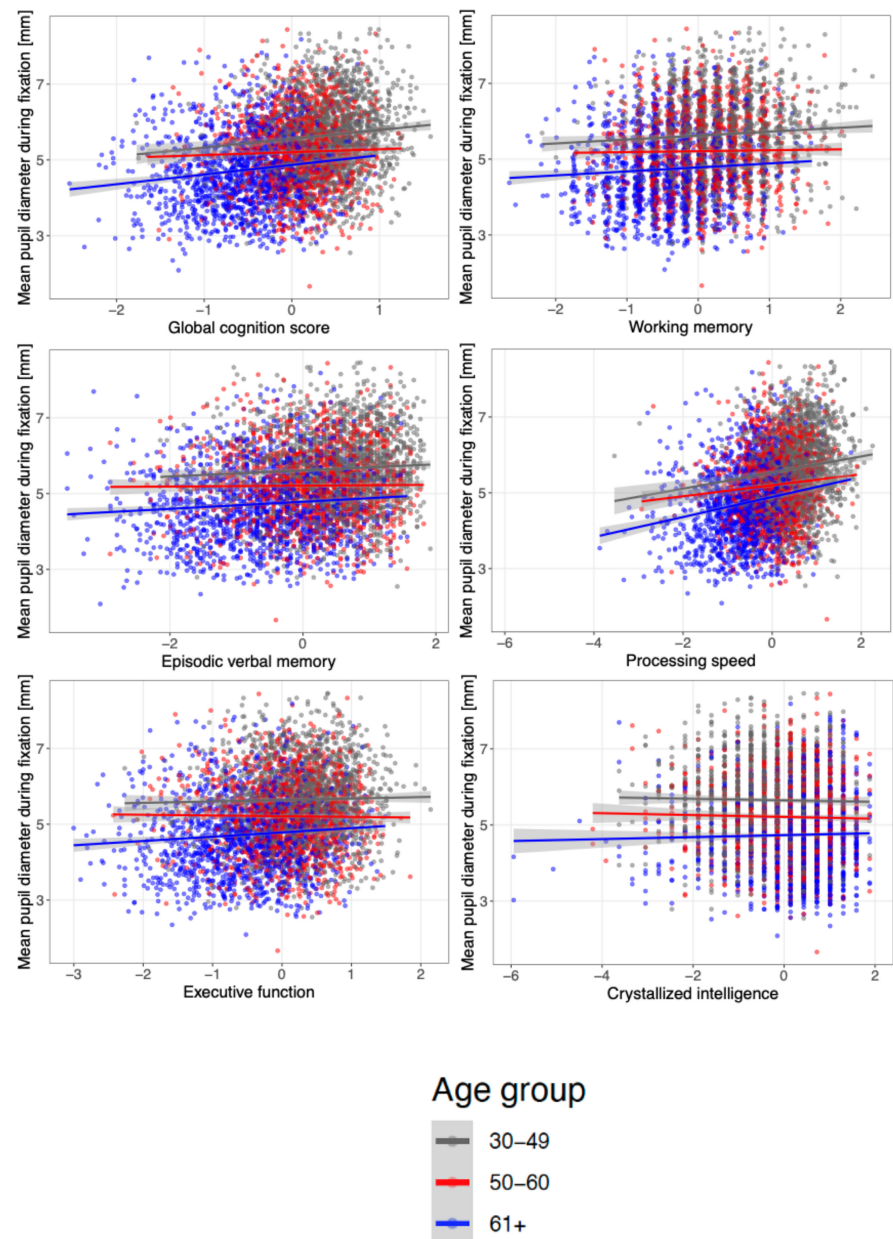
TABLE 5 Associations between cognitive test scores and mean pupil diameter during fixation in age groups 30–49, 50–60, and 61+ years

Cognitive predictor	β (95% CI) in the 30- to 49-year-olds	β (95% CI) in the 50- to 60-year-olds	β (95% CI) in the 61+-year-olds
Global cognition score	0.068 (−0.050, 0.186)	0.075 (−0.049, 0.199)	0.057 (−0.044, 0.159)
Working memory	0.032 (−0.043, 0.106)	0.022 (−0.063, 0.107)	0.017 (−0.059, 0.093)
Episodic verbal memory	0.022 (−0.045, 0.088)	0.013 (−0.053, 0.079)	−0.014 (−0.069, 0.040)
Processing speed	0.076 (−0.006, 0.159)	**0.129 (0.042, 0.217)**	**0.150 (0.083, 0.216)**
Executive function	−0.016 (−0.098, 0.067)	−0.021 (−0.105, 0.064)	−0.010 (−0.078, 0.058)
Crystallized intelligence	0.033 (−0.020, 0.086)	−0.012 (−0.073, 0.048)	0.020 (−0.030, 0.070)
Digit span forward	**0.047 (0.000, 0.094)**	0.003 (−0.049, 0.054)	0.018 (−0.031, 0.066)
Digit span backward	0.006 (−0.039, 0.051)	0.010 (−0.042, 0.061)	0.011 (−0.042, 0.063)
Corsi forward	0.002 (−0.050, 0.053)	0.022 (−0.036, 0.079)	0.004 (−0.047, 0.056)
Corsi backward	−0.004 (−0.054, 0.045)	0.002 (−0.057, 0.060)	0.006 (−0.045, 0.056)
AVLT-immediate recall	0.012 (−0.050, 0.073)	0.007 (−0.056, 0.070)	−0.013 (−0.066, 0.041)
AVLT-delayed recall	0.025 (−0.037, 0.086)	0.015 (−0.045, 0.075)	−0.013 (−0.062, 0.036)
Trail-making test A	0.017 (−0.055, 0.089)	−0.017 (−0.102, 0.067)	0.002 (−0.037, 0.042)
Prosaccade latency	**−0.066 (−0.128, −0.005)**	**−0.098 (−0.158, −0.038)**	**−0.112 (−0.154, −0.069)**
Trail-making test B	0.007 (−0.099, 0.114)	−0.016 (−0.090, 0.059)	0.017 (−0.022, 0.056)
Word fluency	−0.033 (−0.082, 0.016)	−0.041 (−0.097, 0.014)	−0.004 (−0.056, 0.047)
Antisaccade error rate	−0.043 (−0.099, 0.013)	0.002 (−0.053, 0.056)	−0.008 (−0.051, 0.035)

Note: The table displays the change in standard deviations in mean pupil diameter during fixation per one standard deviation increase in the cognitive test or domain score (β and 95% CI) for three different age strata. Regression coefficients were obtained from the following multivariable linear regression model: Mean pupil diameter during fixation ~ β_0 + cognitive score*β_1 + age + sex + population stratification + education + visual acuity + (native language) + residual error. Native language was not included in the models with prosaccade latency and antisaccade error rate as predictors. In bold are those associations with a *p* value below .05.

Abbreviations: 95% CI, 95% confidence interval; AVLT, Auditory Verbal Learning and Memory test; β, standardized regression coefficient.

FIGURE 2 Associations between cognitive domain scores and mean pupil diameter during fixation. For each cognitive domain, there is a single scatterplot with the domain score (z-score) on the x-axis and the mean pupil diameter during fixation [in mm] on the y-axis. Each data point represents one participant. Since the data points of some participants overlap, some data points have a darker color than others. Data of three age groups are depicted in different colors, gray for the 30–49 years old, red for the 50–60 years old, and blue for the 61+-years old. For each age group there exists one superimposed linear function for the zero-order correlation between the cognitive domain score and mean pupil diameter during fixation

were due to differences in pupil size measurement leading to limited variation in pupil size as the standard deviation in pupil diameter across all participants was 1.1 mm, as in the study by Tsukahara et al. (2016), and more than double of the standard variation in pupil diameter reported in Unsworth et al. (2020). Tsukahara and Engle (2021) also highlighted that each domain should be measured with several different tasks to avoid that task-specific ability that are unrelated to the domain strongly influence the domain score. Since we measured working memory with four test scores from two tasks and built the global cognition score based on 11 outcomes from eight tasks, we also took this into account. Further, our regression models were also adjusted for interindividual differences in age and genetic background (see discussion in Tsukahara et al., 2021; Unsworth et al., 2021). It should be noted,

however, that our participants deviate considerably in age range from the study by Tsukahara et al. (2016).

Age had a strong effect on pupil size, with the average pupil diameter during fixation decreasing by 0.33 mm per 10 years of increase in age. For comparison, the pupil dilates rarely more than 0.5 mm in response to non-luminance-driven changes (e.g., changes in mental effort) (Beatty & Lucero-Wagoner, 2000). Small correlations ($-0.32 \leq r \leq -0.17$) between age and pupil diameter have been reported in previous studies with more restricted age ranges (Tsukahara et al., 2016; Unsworth et al., 2020). Further, our correlation analyses showed that age largely accounted for correlations between cognitive performance and mean pupil diameter during fixation as the zero-order correlations were largely diminished in the age-stratified analysis compared to the analysis across the whole sample.

PSYCHOPHYSIOLOGY SPR

Cognitive score	Whole sample	30- to 49-year-olds	50- to 60-year-olds	61+-year-olds
Global cognition score	**0.29**	**0.10**	0.03	**0.14**
Working memory	**0.18**	**0.06**	0.02	**0.07**
Episodic verbal memory	**0.20**	**0.06**	0.01	**0.09**
Processing speed	**0.31**	**0.12**	**0.09**	**0.20**
Executive function	**0.18**	0.02	−0.01	**0.09**
Crystallized intelligence	**−0.11**	−0.02	−0.02	0.03
Digit span forward	**0.09**	**0.06**	0.00	**0.05**
Digit span backward	**0.09**	0.03	0.00	0.04
Corsi forward	**0.18**	0.04	0.03	**0.05**
Corsi backward	**0.14**	0.04	0.01	**0.05**
AVLT- immediate recall	**0.19**	0.05	0.00	**0.09**
AVLT-delayed recall	**0.19**	**0.06**	0.01	**0.08**
TMT-A	**−0.21**	−0.03	−0.02	**−0.07**
Prosaccade latency	**−0.25**	**−0.10**	**−0.09**	**−0.18**
TMT-B	**−0.20**	**−0.06**	−0.01	**−0.07**
Word fluency	**0.04**	−0.03	−0.04	**0.06**
Antisaccade error rate	**−0.12**	**−0.05**	0.00	−0.05

TABLE 6 Zero-order correlations between cognitive scores and mean pupil diameter during fixation across the sample and in different age strata

Note: The table displays Pearson correlations between cognitive scores and mean pupil diameter during fixation for the whole sample and for three different age strata. In bold are those zero-order correlations with a *p* value below .05.

However, we found that the age*cognition interaction term was not significant in any model, indicating that the associations between cognitive performance and mean pupil size during fixation were quite consistent across age. Additionally, the standard deviation in mean pupil size was highly comparable in all age groups (Table 1), which indicates that interindividual variation in pupil size remained consistent across age groups. This rules out the possibility that the associations between cognitive performance and baseline pupil diameter vary with age due to an age-related decrease in interindividual variation in pupil size (Tsukahara & Engle, 2021).

Nevertheless, we additionally performed age-stratified analyses. In the age-stratified analysis, better digit span forward performance was significantly associated with larger mean pupil diameter during fixation in the 30- to 49-year-olds (*p* value = .048). However, the effect size was small (ß = .0471) and the confidence interval was very close to zero (95% CI: 0.0003 to 0.0939). Moreover, we did not observe a consistent pattern across the age strata. Further, the three other working memory test scores and the domain score for working memory were not associated with mean pupil diameter during fixation in this age group. Thus, the association between digit span forward performance and pupil size in this age group should be interpreted with caution.

Investigating associations not only of working memory and global cognitive scores with mean pupil size during fixation but also of additional cognitive measures, we found that higher values in the processing speed domain score were associated with larger mean pupil diameters during fixation across the sample and most strongly in the 50- to 60-year-olds and the 61+-year-olds. In line with this, higher prosaccade latency, which indicates worse performance, was associated with lower mean pupil diameter during fixation across the sample and in all three age strata, with the strength of the effect increasing with increasing age. Since the processing speed domain score only consists of TMT-A and prosaccade latency, the association between processing speed and pupil diameter may have been mainly driven by the oculomotor outcome. Higher pre-saccadic pupil dilations have been found to precede lower saccade latencies (Jainta et al., 2011; Wang et al., 2015). In addition, there is evidence for a link between the pupil control circuit and the superior colliculus (Wang & Munoz, 2015). Thus, pupil size may reflect preparatory neural activity related to saccade generation (Jainta et al., 2011; Wang et al., 2015). This suggests that individuals with larger pupil diameter during fixation may generally have higher levels of preparatory neural activity and, therefore, on average also lower saccade latencies.

Despite known sex differences in the locus coeruleus–norepinephrine system (Bangasser et al., 2016), which has been linked to pupil size (Joshi & Gold, 2020), and sex differences in cognitive performance (Herrera et al., 2019), we did not find sex differences in pupil diameter during fixation. This may be due to the very simple nature of the fixation task as sex differences in arousal are more

TABLE 7 Method comparison with the study by Tsukahara et al. (2016)

Experimental condition	Study by Tsukahara et al. (2016)	Our study	Comparable conditions?
Sample characteristics	• Age range: 17 to 35 years • Analysis based on 337 participants • Data come in part from college students • No information on sex distribution	• Age range: 30 to 95 years with a mean age of 54.7 years (SD: 13.9) • Analysis based on 4560 participants • Data come from a population-based cohort study; women and individuals with high educational level are slightly over-represented	• Age range not comparable but sensitivity analyses conducted (age-stratified analyses; no interaction effects between age and cognition on mean pupil size during fixation) • Over-representation of women in our sample not detrimental as we found no sex effect on mean pupil size during fixation
Measurement of cognitive domains	• Working memory: z-score averaging of the cognitive outcomes from three complex span tasks • Fluid intelligence: z-score averaging of the outcomes from three fluid intelligence tasks (Raven advanced progressive matrices, Letter sets, Number series)	• Working memory: z-score averaging of four test scores derived from two cognitive tasks (digit span and Corsi forward and backward) • Fluid intelligence: z-score averaging of 11 cognitive test scores derived from eight tasks measuring four different domains (processing speed, working memory, episodic verbal memory, executive function)	Yes
Pupil size measurement	• Average pupil size during fixation on a gray fixation point on black background for 30 s	• Average pupil size during fixation on a white fixation point on black background that appears at five different positions during a one-minute fixation task	Yes
Experimental light conditions that lead to sufficient variation in pupil size	• SD in pupil diameter across all participants: 1.1 mm	• SD in pupil diameter across all participants: 1.1 mm • SD in pupil diameter in the three separate age groups: 1.0 mm	Yes
Statistical models	• Self-reported ethnicity (three categories: Caucasian, African-American, other) was included in the regression models • All models were controlled for age, nicotine use in the past 10 h, and intake of medication in the past 24 h that may affect attention and memory (all variables were self-reported)	• Six principal components included in the regression models to control for differences in genetic background • All models were controlled for age, sex, best-corrected visual acuity, educational level, and native language (except for the two models with an eye movement predictor)	• In both studies, the models were controlled for the confounders age and race • Some confounders were not present in our study and vice versa

Abbreviations: SD, standard deviation.

pronounced under stress (Bangasser et al., 2011, 2016). In addition, the mean pupil size during fixation is not a pure marker of activity in the locus coeruleus-norepinephrine system but has also been linked to activity in other brain areas such as the superior colliculus (Joshi & Gold, 2020), for which there is no clear evidence of sex differences.

Some strengths and weaknesses of our study should be addressed. Our study tested the associations between cognitive performance and pupil diameter during fixation in a large sample under comparable experimental conditions as in the study by Tsukahara et al. (2016) except for a much wider age range. The wide age range allowed us to extend the findings from the literature to an older and broader age range and to examine whether any associations were age-dependent. However, the wide age range limits a direct comparison of our findings with previous findings (Table 7). We tried to address this limitation by calculating age-stratified analyses. Still, even the participants in our youngest age group (30- to 49-year-olds) were older than the participants in previous studies that mainly included individuals aged 17 to 35 years (Unsworth et al., 2020). We thus cannot entirely exclude on the basis of our data that there may be a relationship between working memory capacity and fluid intelligence with pupil size in adolescents and younger adults. However, given the absence of such a relation across the entire older age range, as well as the absence of any indication that associations between cognitive performance and pupil size strongly depended on age, we consider this unlikely. Another limitation of our study is that there was only one run per cognitive test and that we were not able to validate the internal consistencies of our cognitive measures. Further, the variance explained did not largely differ between the models including different cognitive predictors (Tables 3 and 4). This suggests that processing speed and prosaccade latency explained little variance in mean pupil size during fixation and that most of the variance in pupil outcome could be explained by the covariates included in the regression models (age, sex, best-corrected visual acuity, educational level, race, and native language).

Future research could help to clarify how far performance in non-oculomotor processing speed tasks is associated with pupil diameter during fixation. Further, the assumption that pupil diameter and processing speed are associated because of differences in the level of neural activation could also be investigated by testing whether resting-state network activity partly mediates the association between processing speed and pupil diameter during fixation.

5 | CONCLUSIONS

Working memory and global cognition were not associated with pupil diameter during fixation, which cannot be explained by limited interindividual variation in pupil diameter during fixation or a small number of tests per cognitive domain. Processing speed and prosaccade latency were associated with pupil diameter, suggesting that differences in pupil diameter may inform about differences in levels of preparatory neural activity for saccades. These associations were consistent across a large age range. Pupil diameter declined linearly with age but did not differ between men and women.

AUTHOR CONTRIBUTIONS

Annabell Coors: Conceptualization; data curation; formal analysis; investigation; methodology; validation; visualization; writing – original draft; writing – review and editing. **Monique M. B. Breteler:** Conceptualization; funding acquisition; methodology; project administration; resources; supervision; validation; writing – review and editing. **Ulrich Ettinger:** Conceptualization; methodology; supervision; writing – review and editing.

ORCID
Annabell Coors https://orcid.org/0000-0001-9000-6390
Monique M. B. Breteler https://orcid.org/0000-0002-0626-9305
Ulrich Ettinger https://orcid.org/0000-0002-0160-0281

REFERENCES
Althouse, A. D. (2016). Adjust for multiple comparisons? It's not that simple. *Annals of Thoracic Surgery*, *101*(5), 1644–1645. https://doi.org/10.1016/j.athoracsur.2015.11.024

Bangasser, D. A., Eck, S. R., Telenson, A. M., & Salvatore, M. (2011). Sex differences in stress regulation of arousal and cognition. *Physiology & Behavior*, *103*(3–4), 342–351. https://doi.org/10.1016/j.physbeh.2011.02.037

Bangasser, D. A., Wiersielis, K. R., & Khantsis, S. (2016). Sex differences in the locus coeruleus-norepinephrine system and its regulation by stress. *Brain Research*, *1641*, 177–188. https://doi.org/10.1016/j.brainres.2015.11.021

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142–162). Cambridge University Press.

Boenniger, M. M., Staerk, C., Coors, A., Huijbers, W., Ettinger, U., & Breteler, M. M. B. (2021). Ten parallel German versions of Rey's auditory verbal learning test: Age and sex effects in 4,000 adults of the Rhineland study. *Journal of Clinical and Experimental Neuropsychology*, *43*(6), 637–653. https://doi.org/10.1080/13803395.2021.1984398

Bornemann, B., Foth, M., Horn, J., Ries, J., Warmuth, E., Wartenburger, I., & van der Meer, E. (2010). Mathematical cognition: Individual differences in resource allocation. *ZDM - International Journal on Mathematics Education*, *42*(6), 555–567. https://doi.org/10.1007/s11858-010-0253-x

Coors, A., Merten, N., Ward, D. D., Schmid, M., Breteler, M. M. B., & Ettinger, U. (2021). Strong age but weak sex effects in eye movement performance in the general adult population: Evidence from the Rhineland study. *Vision Research*, *178*, 124–133. https://doi.org/10.1016/j.visres.2020.10.004

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. https://doi.org/10.1111/j.1469-8986.2007.00605.x

Herrera, A. Y., Wang, J., & Mather, M. (2019). The gist and details of sex differences in cognition and the brain: How parallels in sex differences across domains are shaped by the locus coeruleus and catecholamine systems. *Progress in Neurobiology*, *176*, 120–133. https://doi.org/10.1016/j.pneurobio.2018.05.005

Jainta, S., Vernet, V., Yang, Y., & Kapoula, Z. (2011). The pupil reflects motor preparation for saccades - even before the eye starts to move. *Frontiers in Human Neuroscience*, *5*, 1–10. https://doi.org/10.3389/fnhum.2011.00097

Joshi, S., & Gold, J. I. (2020). Pupil size as a window on neural substrates of cognition. *Trends in Cognitive Sciences*, *24*(6), 466–480. https://doi.org/10.1016/j.tics.2020.03.005

Lehrl, S. (2005). *Mehrfachwahl-Wortschatz-Intelligenztest MWT-B*. Spitta Verlag.

Lowenstein, O., & Loewenfeld, I. E. (1950). Role of sympathetic and parasympathetic systems in reflex dilatation of the pupil. *Archives of Neurology and Psychiatry*, *64*(3), 313–340. https://doi.org/10.1001/archneurpsyc.1950.02310270002001

Mather, M., & Harley, C. W. (2016). The locus coeruleus: Essential for maintaining cognitive function and the aging brain. *Trends in Cognitive Sciences*, *20*(3), 214–226. https://doi.org/10.1016/j.tics.2016.01.001

McDougal, D. H., & Gamlin, P. D. R. (2008). Pupillary control pathways. In R. H. Masland, T. D. Albright, P. Dallos, D. Oertel, S. Firestein, G. K. Beauchamp, M. C. Bushnell, A. I. Basbaum, J. H. Kaas, & E. P. Gardner (Eds.), *The senses: A comprehensive reference* (pp. 521–536). Academic Press.

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, *222*, 250–259. https://doi.org/10.1016/j.jneumeth.2013.10.024

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Purcell, S. M., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Sibley, C., Foroughi, C., Brown, N., & Coyne, J. T. (2018). Low cost eye tracking: Ready for individual differences research? *Proceedings of the Human Factors and Ergonomics Society*, *2*, 741–745. https://doi.org/10.1177/1541931218621168

Tsukahara, J. S., Draheim, C., & Engle, R. W. (2021). Baseline pupil size is related to fluid intelligence: A reply to Unsworth. *Cognition*, *215*, 104826. https://doi.org/10.1016/j.cognition.2021.104826

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

Tsukahara, J. S., & Engle, R. W. (2021). Is baseline pupil size related to cognitive ability? Yes (under proper lighting conditions). *Cognition*, *211*, 104643. 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. https://doi.org/10.1016/j.cogpsych.2016.10.001

Unsworth, N., Miller, A. L., & Robison, M. K. (2020). Is working memory capacity related to baseline pupil diameter? *Psychonomic Bulletin and Review*, *28*, 228–237. https://doi.org/10.3758/s13423-020-01817-5

Unsworth, N., Miller, A. L., & Robison, M. K. (2021). No consistent correlation between baseline pupil diameter and cognitive abilities after controlling for confounds—A comment on Tsukahara and Engle (2021). *Cognition*, *215*, 104825. https://doi.org/10.1016/j.cognition.2021.104825

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. https://doi.org/10.1111/j.1469-8986.2009.00884.x

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin and Review*, *25*(6), 2005–2015. https://doi.org/10.3758/s13423-018-1432-y

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167–174. https://doi.org/10.1111/j.1469-8986.2003.00148.x

Wang, C. A., Brien, D. C., & Munoz, D. P. (2015). Pupil size reveals preparatory processes in the generation of pro-saccades and anti-saccades. *European Journal of Neuroscience*, *41*(8), 1102–1110. https://doi.org/10.1111/ejn.12883

Wang, C. A., & Munoz, D. P. (2015). A circuit for pupil orienting responses: Implications for cognitive modulation of pupil size. *Current Opinion in Neurobiology*, *33*, 134–140. https://doi.org/10.1016/j.conb.2015.03.018

Wang, L. Y., Murphy, R. R., Hanscom, B., Li, G., Millard, S. P., Petrie, E. C., Galasko, D. R., Sikkema, C., Raskind, M. A., Wilkinson, C. W., & Peskind, E. R. (2013). Cerebrospinal fluid norepinephrine and cognition in subjects across the adult age span. *Neurobiology of Aging*, *34*(10), 2287–2292. https://doi.org/10.1016/j.neurobiolaging.2013.04.007