Maastricht University

Memory load and the cognitive pupillary response in aging

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Download date: 06 Feb. 2023

Memory load and the cognitive pupillary response in aging

PASCAL W. M. VAN GERVEN,[a] FRED PAAS,[b] JEROEN J. G. VAN MERRIËNBOER,[b]
AND HENK G. SCHMIDT[c]

[a]Maastricht University, Faculty of Psychology, Department of Neurocognition, Maastricht, The Netherlands
[b]Educational Technology Expertise Center (OTEC), Open University of the Netherlands, Heerlen, The Netherlands
[c]Erasmus University, Department of Social Sciences, Institute of Psychology, Rotterdam, The Netherlands

Abstract

The effect of memory load on the cognitive pupillary response among 16 young adults and 16 older adults was investigated. Mean pupil dilation and reaction time were measured during a Sternberg memory-search task, which involved six levels of memory load. A classic interaction pattern was obtained in which the reaction times of the elderly participants increased more as a function of memory load than the reaction times of the young participants. In the encoding phase of the experiment, mean dilation increased with memory load. No age differences were observed here. In the search phase of the experiment, however, mean pupil dilation was considerably greater in the young than in the elderly participants. Moreover, mean dilation of the older participants was not sensitive to memory load, whereas mean dilation increased as a function of memory load in the young participants. The results suggest that the usefulness of the pupillary response as a correlate of subtle fluctuations in memory load diminishes with old age.

Descriptors: Cognitive pupillary response, Memory load, Memory search, Aging

Numerous attempts have been made to find reliable and sensitive physiological correlates of cognitive load in various "real life" situations, such as studying instructional material and flying an airplane (e.g., Paas, Van Merriënboer, & Adam, 1994; Wierwille, Rahimi, & Casali, 1985). In addition to the widely used subjective scaling techniques (e.g., Nygren, 1991; Paas & Van Merriënboer, 1994; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2002), there is still a need for these more "objective" measures (e.g., Paas, Tuovinen, Tabbers, & Van Gerven, 2003). One of the most promising physiological measures, both in terms of sensitivity and reliability, is the cognitive pupillary response. The aim of the present study is to test whether the pupillary response, regarding its sensitivity to physiological change, retains its quality as a measure of cognitive load into old age.

Pupillometry had its bloom in the 1960s and early 1970s. Although its history actually begins a little earlier, pupillometric research aimed at processes, which nowadays would be attributed to working memory, was initiated by Hess and Polt (1964). They found pupil diameter to become larger in people who were engaged in solving multiplication problems of increasing difficulty. Kahneman and Beatty (1966) were the first to relate pupil size to memory load. Participants were verbally presented with an increasing number of digits, which had to be reproduced after a short pause. It was found that pupil size increases during encoding and decreases during recall. During the entire procedure, however, pupil diameter was larger when more digits were involved. Peavler (1974) investigated what happens to the size of the pupil when memory is overloaded by an excessive number of digits. Similar to earlier research, he found the pupil to widen as the number of digits increased. However, when the number of digits exceeded nine, the diameter of the pupil stabilized, which in his view indicates a "momentary suspension of processing effort" (p. 559) due to the perceived risk of overload.

It is peculiar that after this successful initial period (see Janisse, 1977, for a review), pupillometry subsequently seemed to pass into oblivion. Strong attempts to inject new life into it were made by Ahern and Beatty (1979) and Beatty (1982). Only throughout the last decade, however, has pupillometry as an instrument for measuring cognitive load seemed to regain a great deal of its initial appeal. This period marks a number of new directions in such areas as physiology and pharmacology (Matthews, Middleton, Gilmartin, & Bullimore, 1991), ergonomics (Backs & Walrath, 1992), computational modeling

(Hoeks & Levelt, 1993), language processing (Hyönä, Tommola, & Alaja, 1995; Just & Carpenter, 1993), visual perception (Verney, Granholm, & Dionisio, 2001), and abnormal psychology (Granholm, Morris, Sarkin, Asarnow, & Jeste, 1997). There is also a renewed interest in pupil size as an indicator of resource limits and memory overload. Granholm, Asarnow, Sarkin, and Dykes (1996), for instance, found a progressive increase of pupil size with memory load, which stabilized at resource limits (i.e., nine digits). Contrary to the results of Peavler (1974), however, pupil size declined after the processing demands exceeded the available resources.

It seems that the new boost of research has consolidated the value of pupillometry as a psychometric instrument. That is, within the limits of cognitive functioning, it has proven to be a reliable and accurate measure of mental activity. However, little is known about its value for measuring mental load in elderly people. It is widely accepted, though, that pupil size substantially decreases with old age, a phenomenon commonly referred to as *senile miosis*. Much of the predominantly medical literature on senile miosis (e.g., Bitsios, Prettyman, & Szabadi, 1996; Korczyn, Laor, & Nemet, 1976; Loewenfeld, 1979; Sekuler & Owsley, 1983; Winn, Whitaker, Elliott, & Phillips, 1994) focuses on its biological basis and its consequences for peripheral functions, such as the pupillary light reflex. The core message is that senile miosis results from a degenerating dilation muscle (the dilator pupillae) in the iris, which causes a linear decline in pupil size with age. It is interesting to note that the dilator pupillae is controlled by sympathetic pathways from the central nervous system (in contrast to the constriction muscle—the sphincter pupillae—which is controlled by the parasympathetic system). Because the sympathetic nervous system is associated with activation, there is reason to believe that task-evoked dilation of the older pupil—if detectable—is an attenuated correlate of mental activity. However, none of the consulted literature mentions an effect of senile miosis on the cognitive pupillary response.

The same holds for the more psychologically oriented literature. There are hardly any age-related studies aimed at pupillary responses that can be attributed to cognitive activity. One of the few comes from Morris, Granholm, Sarkin, and Jeste (1997), who related pupil dilation to memory performance in both healthy and schizophrenic participants ranging in age from 26 to 75 years. They presented their participants with series of 4, 6, 8, and 10 digits and measured the pupillary response at last-digit presentation and at digit retrieval. The participants were required to recall as many digits as possible. Because Morris et al. were primarily interested in performance differences between the schizophrenic participants and the controls at high memory-load levels, they only report data of the 8- and 10-digit conditions. Differences between these conditions were not statistically tested, however. Moreover, the reported means are based on the whole age range. Both at last-digit presentation and at digit retrieval, regression analyses revealed a decline of the pupillary response with increasing age.

One of the most recent pupillometric studies comparing young and older adults comes from Kim, Beversdorf, and Heilman (2000). These researchers were interested in pupillary responses induced by novel stimuli. They required both young (mean age = 30 years) and older participants (mean age = 60 years) to attend to a digit presented in the center of a screen. At certain moments, a new digit appeared somewhere in the periphery of the screen. The subsequent arousal response was recorded after the participant's gaze had shifted to the new stimulus. The data revealed a typical response pattern, comprising three phases: a brief constriction of the pupil, followed by a dilation—the "arousal response"—and a second constriction or "habituation response." It was found that the interval between the arousal response and the habituation response is longer in the elderly than in the young participants. Remarkably, however, no significant age-related difference was found in the magnitude of the dilation response.

Apparently, there is no research that has shown an effect of different levels of "normal" memory load on pupil size in elderly persons. What we do know is that the pupil in older people dilates as a result of excessive memory load and stimulus novelty. We also know that, generally, the older pupil shows less dilation than the young pupil as a result of the physiological decline of the dilator pupillae. The question remains whether in spite of this decline, subtle variations in memory load, that is, within the classical limits of working memory (Miller, 1956), result in a detectable pupillary response.

An interesting task in this respect is Sternberg's (1966, 1967) memory-search task. This task addresses both the storage and processing characteristics of working memory. Sequences of one to six digits have to be memorized. After memorizing a sequence or memory set, a series of single-digit probes is presented. For every probe, the participant has to decide whether it belongs to the memory set (positive probe) or not (negative probe). Response latencies typically increase as the memory set contains more digits. Sternberg did not find effects of the serial position of a positive probe in the memory set, which suggests that the search process is exhaustive, rather than self-terminating (i.e., the search process does not stop as soon as the probe is found). This implies comparable levels of mental activity in response to positive and negative probes. The Sternberg task entails two different, subsequently evolving working-memory activities: encoding and searching. Encoding requires more effort as the memory set becomes larger. Thus, higher levels of pupil dilation can be expected for larger memory sets (e.g., Granholm et al., 1996). In the same light, searching larger memory sets is more effortful and can thus be expected to lead to more pupil dilation.

Beatty and Lucero-Wagoner (2000) describe three possible task-evoked pupillary responses (TEPRs), which can be treated similarly to event-related potentials: (1) *mean dilation* during a critical poststimulus time interval; (2) *peak dilation*, which is the maximum dilation during this time interval; and (3) *peak latency*, which is the time elapsed between probe onset and the appearance of the peak dilation. The magnitude of these TEPRs typically rises as the burden on working memory becomes more intense. Mean dilation is the most commonly used TEPR and was adopted in the present study.

On the basis of the literature discussed, it can be expected that the elderly will show less dilation than the young. Regarding Morris et al.'s (1997) study, in which working memory was loaded to the limit and the young still exhibited more pupil dilation than the elderly, it can furthermore be expected that the elderly's pupil dilation will not exceed that of the young, even though the elderly will probably experience higher levels of cognitive load.

Modern theories on age-related cognitive slowing (e.g., Salthouse, 2000) predict that the reaction times of the elderly will be much longer than those of the young. Furthermore, there is the generally found *complexity effect*, which implies an interaction between age and task complexity in that performance differences between young and old people become larger when

the complexity of the task increases (e.g., Craik & Byrd, 1982; Perfect & Maylor, 2000). The complexity effect is likely to occur in the response latencies resulting from the Sternberg task. The question arises whether the complexity effect will also be reflected by the pupillary responses. That is, will the elderly show relatively more dilation in the higher memory sets? Based on the senile-miosis literature, the effect will at least show a different pattern. That is, the young will display a larger overall pupillary response than the elderly, but in the greater memory sets (i.e., higher memory load), the pupillary responses of the two age groups are likely to converge.

Method

Participants

Participants were 16 health-science, medicine, and psychology students of Maastricht University (1 man and 15 women; mean age = 20.4 years, *SD* = 2.2) and 16 elderly adults (11 men and 5 women; mean age = 68.6 years, *SD* = 4.7), who were recruited via advertisements in local newspapers. All participants were in good health and had normal or corrected-to-normal vision. None of them took medication or had ever suffered an eye disorder that could influence their pupillary response. They received €5 (about $5) for their participation.

Memory-Search Task

The memory-search task (Sternberg, 1966, 1967) comprises two phases. In the encoding phase, participants are required to memorize a certain number of randomly selected digits (ranging from 0 to 9), none of which occurs more than once. The number of digits in this so-called *memory set* varies from one to six. In the search phase, the participant is presented with a sequence of randomly selected single-digit probes. In each trial, the participant has to decide as quickly as possible whether a probe belongs to the memory set (i.e., positive probe) or not (i.e., negative probe).

Apparatus

Pupil size was measured with the SMI (SensoMotoric Instruments, Berlin, Germany) Remote Eyetracking Device (RED) with an angular resolution of less than 0.5°. The device consisted of an infrared source, which was aimed at the participant's eye at a distance of about 90 cm, and an infrared camera. Rays from the infrared source were reflected by the retina and picked up by the camera, which generated a high-resolution image of the pupil on a computer screen. The participant was seated in an adjustable chair and his or her head was stabilized by a chin and forehead rest. The ambient luminance was 516 lx. At a rate of 50 Hz, that is, every 20 ms, both the horizontal and vertical pupil diameters were measured in screen pixels, which were transformed into millimeters. Stimuli were presented by an IBM-compatible computer on a 21-in. monitor, which was placed at a distance of about 110 cm from the participant. The size of the stimulus digits on the screen was approximately 2.5 × 1.5 cm. The software controlling the stimulus presentation was programmed in ERTS (Experimental Run Time System; Beringer, 1996). The participant responded by pressing one of two response buttons. Reaction times were measured in milliseconds.

Design and Procedure

Age group (levels young and old) was the independent between-groups variable and memory load (levels 1 to 6) was the independent within-groups variable. The function of the response buttons was counterbalanced. That is, half of the participants were instructed to press the left button if a negative probe was presented and the right button if a positive probe was presented; the other half of the participants received opposite instructions. The order of memory-set presentation was counterbalanced by presenting one half of the participants with an increasing number of digits (1 to 6), whereas the other half was presented with a decreasing number of digits (6 to 1). The number of positive and negative probes was balanced as well.

The experiment was performed in individual sessions. After calibrating the eyetracking device, the participant was presented with a concise onscreen instruction concerning the memory-search task. Then there was a practice session with a two-digit memory set and eight probes. Subsequently, the first experimental memory set was presented. The digits of a memory set were presented simultaneously. After memorizing a set, the participant had to push a button to continue. Thus, the presentation time of the memory set was variable, dependent on the time needed by the participant. The participant was then presented with a sequence of 24 random probes. A 1000-Hz (45 dB) warning tone was presented 1,000 ms before each probe onset. Probes were presented for 2,000 ms. The intertrial interval was 5,000 ms, which was considered sufficient for the pupil to revert to baseline level. The participant did not receive feedback about the correctness of a response.

Data Processing

Eye blinks were filtered out on the basis of sudden drops in vertical pupil diameter (i.e., the direction in which the eyelids move). The criterion for detecting a drop was the ratio between the vertical (d_v) and horizontal diameter (d_h). For a perfectly circular pupil, the ratio d_v/d_h lies around 1. During a blink or semi-blink, however, this ratio quickly drops toward zero. All data with a ratio differing more than one standard deviation from the mean were replaced by linear interpolation. Subsequently, pupil area (A) was calculated as the area of an ellipse: $A = \frac{1}{4}\pi d_v d_h$. In the encoding phase, baseline pupil area was determined by the average pupil size during a period of 500 ms preceding the presentation of the memory set. In the search phase, baseline pupil area was determined for each trial by the pupil's average size during a period of 500 ms preceding the warning tone. Pupil dilation or constriction was calculated by subtracting this baseline area from the pupil area during the response period, which was a period of 2,000 ms following probe onset.

Mean pupil dilation was calculated for both the encoding phase and the search phase. For the search phase, dilation-by-time plots (see Figure 3 below) reveal that the dilation curves pertaining to the different memory sets start to diverge around 650 ms after probe onset, which roughly corresponds to the average cognitive pupillary response latency found in earlier research (e.g., Beatty, 1982). Therefore, mean dilation was determined as the average dilation during a time window of 650 to 2,000 ms after probe onset. Data were averaged across trials. Values corresponding to an incorrect response (i.e., an incorrect button press) were replaced by the mean score of the participant within a memory set. In this way, 3.8% of the young and 2.5% of the old participants' data were substituted.

Results

A 2 (age group) × 6 (memory load) analysis of variance, with age group as the between-groups independent variable and memory load as the within-groups independent variable, was performed on baseline pupil size, average dilation, and reaction time. A Greenhouse–Geisser correction was applied if sphericity was violated. The corrected degrees of freedom and significance levels are given for the within-groups effects if applicable. Polynomial tests of within-groups contrasts were conducted to evaluate the characteristics of the within-groups effects.

Special care was taken with respect to the presumed relation between baseline pupil size and poststimulus pupil dilation. There is quite some controversy about this so-called *law of initial value* (LIV). The LIV states that "the higher the initial value [of a physiological variable or function], the smaller the response to function-raising, the larger the response to function-depressing stimuli" (Wilder, 1967, p. viii). In the present study, this means that the higher the baseline pupil size (i.e., the initial value), the lower the pupillary response to the probe.

The debate around the LIV has focused on two main issues. First of all, the "law" is challenged with respect to its direction: There seems to be no consensus about whether a higher initial value leads to a higher or a lower poststimulus value. Jin (1992), for instance, argues that the law should be reformulated as: "(a) within the middle range of the initial state, the higher the initial value, the greater the organism's reactivity, and (b) a tendency to reversed responses may occur when the initial value reaches its upper limit" (p. 182). Considering the data in this study, the direction of the law is inconclusive. In the encoding phase of the experiment, the overall correlation between baseline pupil size and dilation was slightly negative, $r = -.15$, $p < .05$, whereas in the search phase the correlation between baseline and dilation was slightly positive, $r = .26$, $p < .01$.

A second source of debate is the variety of competing statistical techniques for dealing with the LIV (e.g., Jennings & Stine, 2000, p. 886). A commonly used technique is an analysis of covariance, in which baseline is the covariate. An ANCOVA is problematic, however, because variation of the covariate is intrinsic to the independent variable (Miller & Chapman, 2001). That is, variation in baseline pupil size is intrinsic to variation in memory load. Instead of an ANCOVA, a relatively simple procedure was applied. Following Jin's (1992) revised version of the LIV, the dilation values were converted to proportions of the corresponding baseline values. In this way, dilation values pertaining to higher baseline levels were attenuated. For comparison reasons, separate ANOVAs were performed on the uncorrected and the proportional data.

Encoding Phase

Means and standard deviations of baseline pupil size can be found in Table 1. Baseline pupil size was significantly greater in the young than in the elderly participants, $F(1,30) = 14.64$, $MSE = 159.82$, $p = .001$. Furthermore, a significant main effect of memory load was found, $F(3.2,94.7) = 4.40$, $MSE = 5.23$, $p < .01$. Baseline was more sensitive to memory load in the young than in the older participants, which is reflected by a significant interaction between age group and memory load, $F(3.2,94.7) = 4.38$, $p < .01$. Tests of within-groups contrasts reveal that baseline follows a quadratic pattern, $F(1,30) = 10.96$, $MSE = 3.00$, $p < .01$, setting off high in the lower levels of memory load, decreasing in the intermediate levels, and rising again in the higher levels of memory load. The relatively high baseline in the lower levels of memory load can be explained as an initial arousal at the beginning of the experiment. A significant interaction, $F(1,30) = 11.35$, $p < .01$, shows that the quadratic pattern is stronger in the young than in the elderly participants, as can be seen in Table 1.

Remarkably, pupil dilation during encoding of the memory set was equal for the age groups, $F(1,30) = 0.92$, $MSE = 4.17$, $p > .05$. There was a main effect of memory load, $F(3.5,106.0) = 2.63$, $MSE = 3.30$, $p < .05$, indicating an increase of pupil dilation. Tests of within-groups contrasts revealed a linear pattern, $F(1,30) = 4.01$, $MSE = 2.77$, $p = .05$. This result is in line with Figure 1, which shows a progressive increase of pupil dilation with memory load. The only aberration in this pattern is a sudden drop in pupil dilation at a memory load of six digits. In the older participants, pupil dilation seems to stabilize at a memory load of five digits. The absence of an interaction, $F(3.5,106.0) = 0.15$, $p > .05$, suggests that the effect of memory load on dilation does not differ between age groups.

Descriptive statistics of dilation as a proportion of baseline are given in Table 2. An analysis of variance reveals the same pattern of results as was found for the dilation data. Again, there is no main effect of age group, $F(1,30) = 0.70$, $MSE = 0.02$, $p > .05$. There is also a main effect of memory load, $F(5,150) = 3.11$, $MSE = 0.01$, $p < .05$. Finally, there is no significant interaction, $F(5,150) = 0.33$, $p > .05$. Inspection of Table 2 reveals that the means of the proportional data roughly follow the same pattern as the dilation data in Figure 1. In the elderly, however, the sudden drop in dilation seems to already have set off at a memory load of five digits.

Search Phase

The search phase is the essential part of the Sternberg task. Here, the reaction times should reveal a typical complexity effect. Figure 2 reveals a pattern of results, which is in line with this classical complexity effect. Age group and memory load seem to

Table 1. *Means and Standard Deviations (in Parentheses) of Baseline Pupil Size in Square Millimeters*

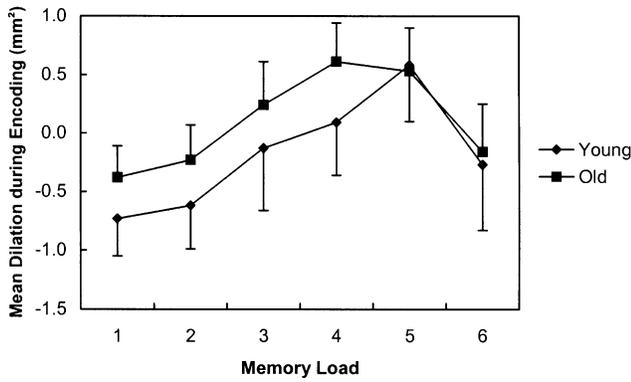|	Memory load					
	1	2	3	4	5	6
Encoding phase						
Young	18.43 (7.01)	17.19 (6.59)	16.49 (6.78)	17.83 (6.17)	16.39 (6.64)	19.49 (8.01)
Old	9.94 (3.64)	10.17 (3.45)	10.34 (3.22)	10.74 (3.03)	11.52 (3.58)	11.22 (3.21)
Search phase						
Young	14.23 (5.79)	13.35 (5.27)	13.05 (5.39)	13.52 (5.41)	13.44 (5.21)	14.53 (5.67)
Old	8.45 (2.74)	8.37 (2.75)	8.50 (2.39)	8.90 (2.41)	8.98 (2.28)	9.42 (2.45)

Figure 1. Mean pupil dilation (in square millimeters) during encoding as a function of memory load. Values are change scores relative to baseline. Bars indicate one standard error of the mean.
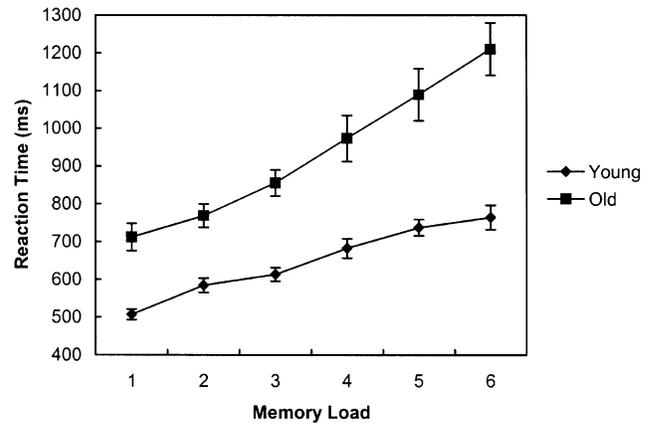

Figure 2. Reaction time (in milliseconds) as a function of memory set. Bars indicate one standard error of the mean.

interact in that the elderly show a greater increase in reaction time than the young. This is confirmed by the ANOVA. First of all, there was a main effect of age group, $F(1,30) = 35.49$, $MSE = 111,197.97$, $p < .001$, where the elderly show greater reaction times than the young. In addition, there was a main effect of memory load, $F(2.5,76.2) = 70.24$, $MSE = 18,663.54$, $p < .001$. Tests of within-groups contrasts reveal a linear increase in reaction time, $F(1,30) = 124.65$, $MSE = 26,571.23$, $p < .001$. Finally, there was a significant interaction, $F(2.5,76.2) = 8.25$, $p < .001$, which establishes the complexity effect. Tests of within-groups contrasts reveal that this interaction had a linear character, $F(1,30) = 13.29$, $p = .001$.

Similar to the encoding phase, baseline pupil area of the young participants is significantly greater than that of the older participants in the search phase, $F(1,30) = 11.21$, $MSE = 103.49$, $p < .01$. There is also a main effect of memory load, $F(3.4,102.4) = 5.96$, $MSE = 1.44$, $p = .001$. However, unlike the baseline data in the encoding phase, the baseline data in the search phase do not reveal a significant interaction, $F(3.4,102.4) = 2.00$, $p > .05$. Tests of within-groups contrasts yield a quadratic development of baseline pupil size with memory load, $F(1,30) = 14.44$, $MSE = 1.14$, $p = .001$. Again, especially the young participants show relatively high baselines in the lower levels of memory load (see Table 1), which is confirmed by a significant interaction, $F(1,30) = 5.85$, $p < .05$.

Figure 3 displays the development of pupil dilation as a function of time after probe onset. Each curve represents a different level of memory load. In both age groups, the curves start to diverge around 650 ms. Divergence seems to be greater in the young (left panel) than in the older participants (right panel). In the young participants, a clear arrangement of the curves

is established. This arrangement is less clear in the elderly participants. In Figure 4, it appears that mean dilation in the young develops differently from that in the elderly. First of all, pupil dilation is greater in the young. Furthermore, Figure 4 suggests a progressively increasing mean pupil dilation with memory load in the young, whereas the elderly's mean dilation more or less stays at the same level. These patterns are supported by the analysis of variance. There was indeed a main effect of age group, $F(1,30) = 4.36$, $MSE = 6.58$, $p < .05$. There was no main effect of memory load, however, $F(2.6,78.4) = 0.71$, $MSE = 0.56$, $p > .05$. Interestingly, there is a trend toward an interaction, $F(2.6,78.4) = 2.34$, $p = .09$. This trend is confirmed by the tests of within-groups contrasts, $F(1,30) = 3.48$, $MSE = 0.78$, $p = .07$, which suggest that the interaction is linear.

An analysis of the proportional data more or less reveals the same results. The effect of age group has faded, however, $F(1,30) = 2.82$, $MSE = 0.03$, $p = .10$. There is still no main effect of memory load, $F(2.8,83.7) = 0.11$, $MSE = 0.004$, $p > .05$. But again, there is a trend toward an interaction, $F(2.8,83.7) = 2.40$, $p = .08$. This interaction appears to be linear, $F(1,30) = 4.01$, $MSE = 0.006$, $p = .05$, as is revealed by the tests of within-groups contrasts.

Discussion

In this study, the cognitive pupillary response of both young and older adults was measured under conditions of normal working-memory load. The central question was whether the cognitive pupillary response of elderly people, which is supposedly affected by physiological decline, can differentiate between moderate

Table 2. *Means and Standard Deviations (in Parentheses) of Dilation as a Proportion of Baseline Pupil Size*

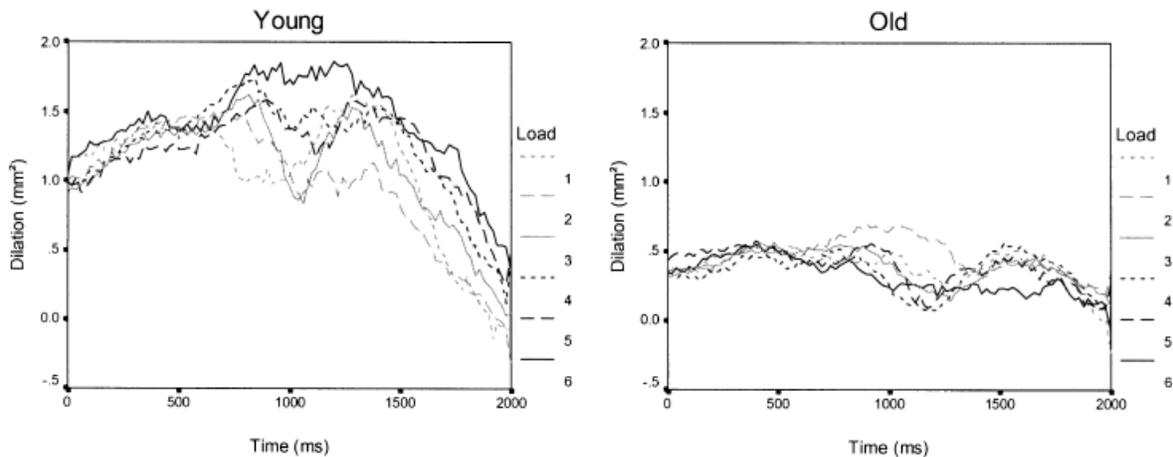|	Memory load					
	1	2	3	4	5	6
Encoding phase						
Young	− .04 (.08)	− .03 (.09)	.01 (.12)	.01 (.13)	.03 (.10)	− .01 (.09)
Old	− .03 (.09)	− .02 (.13)	.02 (.11)	.07 (.12)	.05 (.12)	.00 (.14)
Search phase						
Young	.07 (.07)	.07 (.08)	.09 (.08)	.09 (.09)	.10 (.12)	.10 (.11)
Old	.05 (.04)	.06 (.05)	.04 (.05)	.04 (.08)	.04 (.09)	.02 (.11)

Figure 3. Mean pupil dilation (in square millimeters) as a function of time (in milliseconds after probe onset) in the search phase. Values are change scores relative to baseline. The left panel shows the mean values of the 16 young participants; the right panel shows the mean values of the 16 elderly participants.

levels of memory load. A memory-search task (Sternberg, 1966, 1967), involving six levels of memory load, was used to evoke the cognitive pupillary response. The task comprised two phases: an encoding phase and a search phase. In the encoding phase, the young showed more variation in baseline pupil size, but there were no age differences in pupil dilation. In both age groups, pupil dilation progressively increased with the first five levels of memory load. These findings are not in line with the hypothesis that the young show more dilation than the elderly and that dilation values of the age groups converge at the higher levels of memory load.

In the search phase, the young participants did show significantly more pupil dilation than their elderly counterparts. Regarding the reaction times, however, the elderly experienced higher levels of cognitive load than the young. Thus, it is impossible to compare the age groups in terms of absolute levels of cognitive load on the basis of differences in pupillary response. Remarkably, there was a trend toward a linear interaction, which indicated that the pupillary responses of the age groups did not converge with increasing memory load, as was hypothesized, but rather diverged (Figure 4). That is, pupil dilation in the elderly participants more or less stayed at the
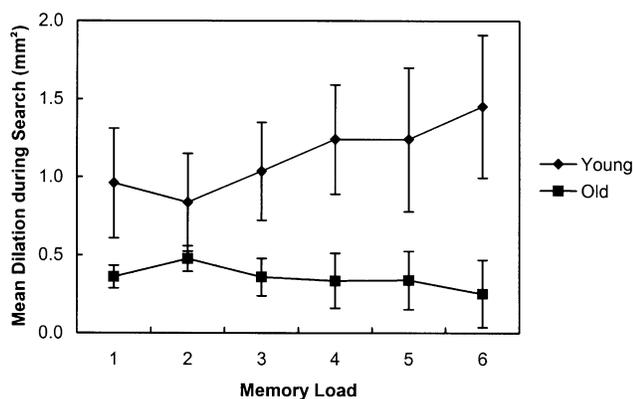
same level, whereas pupil dilation in the young increased with memory load.

How can the absence of an increasing pupillary response in the older participants be explained? It is conceivable that the elderly pupil already reached its maximum size in the encoding phase. This explanation is flawed, however, because the average pupil size of the elderly participants in the search phase ($M = 9.12\,\text{mm}^2$) was smaller than in the encoding phase ($M = 10.76\,\text{mm}^2$). An alternative explanation is that the elderly already reached the limits of their cognitive resources in the encoding phase. This would imply, however, that the elderly performed poor in the search phase, especially under high memory load. This was not the case. The elderly's performance in terms of the percentage correct responses (97.5%) was comparable to the young's performance (96.2%). Moreover, resource limits in elderly people are usually observed at much higher levels of memory load (e.g., Granholm et al., 1996). It is true that the response times revealed a complexity effect in that the elderly were more affected by an increase of memory load than the young, but this is probably due to a different speed–accuracy trade-off resulting from a relatively high impact of "normal" load, not to capacity constraints.

The most likely explanation for the absence of an increasing pupillary response in the older participants is a relatively low sensitivity of the pupil. A low sensitivity may manifest itself under circumstances where cognitive load fluctuates only faintly. Essentially, the Sternberg task entails two types of cognitive load. In the encoding phase, there is the load inflicted by the memorization process. In the search phase, there is the load imposed by the search process. The size of the memory set may have less impact on the level of cognitive load during the search process than during the encoding process. This is plausible if the search process is considered as a type of recall. Kahneman and Beatty (1966), for instance, found a lower pupil dilation during recall in young adults.

Nevertheless, the elderly participants did show pupillary responses in the search phase, although the magnitude of these responses was more or less the same for the whole range of memory load. Possibly, these responses reflected basic arousal levels, which were induced by the mere appearance of the probes (cf. Kim et al., 2000). Although it is disputable whether such

Figure 4. Mean pupil dilation (in square millimeters) during memory search as a function of memory load. Values are change scores relative to baseline. Bars indicate one standard error of the mean.

responses have a cognitive origin, it is still peculiar that different levels of memory load did not elicit different levels of arousal.

It is remarkable that in the encoding phase, both age groups show a decline in mean pupil dilation at a memory load of six digits. It is unlikely that this decline is due to memory overload, which requires much larger memory sets (e.g., Granholm et al., 1996). Regarding the nature of the Sternberg task, it is more likely that the decline marks a strategic switch. A memory set of six randomly selected, unique digits means that there are four possible negative probes (i.e., the remaining digits in the range of 0 to 9). Instead of memorizing the positive set of six digits, participants could have focused on the negative set of only four digits. This may have reduced memory load and thus mean pupil dilation.

With respect to the sample, there was an obvious imbalance of gender and educational background. In the young sample, women were overrepresented; in the elderly sample, men were overrepresented. The scarce psychological literature on possible sex differences in cognitive pupillary response focuses on emotional stimuli. No meaningful influences of gender are reported, however (e.g., Aboyoun & Dabbs, 1998; Oka, Chapman, & Jacobson, 2000). Furthermore, no literature was found that addresses a possible relation between the cognitive pupillary response and educational background or intelligence.

The ultimate test of pupillometry is its usefulness as a "real-time" method for tracing cognitive load in a wide variety of tasks (see Beatty, 1982). The question is whether pupillometry as a measure of cognitive load bears enough advantages in comparison to "simpler" methods, such as secondary-task performance. First and foremost, pupillometry offers a more or less continuous monitor of cognitive load. Furthermore, pupil dilation can be regarded as one of the most sensitive psychophysiological correlates of cognitive load. This study has shown, however, that the sensitivity of pupillometry seems to encounter its limitations in the elderly, who do not show a cognitive pupillary response under all circumstances. That is, the older pupil seems to be less sensitive to subtle variations of cognitive load. Whether this reduced sensitivity is the result of physiological decay (i.e., senile miosis) remains an open question.

In conclusion, pupillometry seems only moderately suitable for measuring "normal" levels of cognitive load in older people. In spite of the dilation effects found for both age groups in the encoding phase, no cognitive pupillary response could be detected in the elderly participants during the search phase of the Sternberg task. This may be explained by a combination of subtle fluctuations of memory load and a reduced sensitivity of the older pupil. Furthermore, age-related differences in the cognitive sensitivity of the pupil make it impossible to compare age groups, because different levels of pupil dilation do not match different levels of actually experienced cognitive load. Regarding these inadequacies, great care should be taken in pupillometric research with older adults.

REFERENCES

Aboyoun, D. C., & Dabbs, J. M., Jr. (1998). The Hess pupil dilation findings: Sex or novelty? *Social Behavior and Personality*, *16*, 415–419.

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science*, *205*, 1289–1292.

Backs, R. W., & Walrath, L. C. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. *Applied Ergonomics*, *23*, 243–254.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*, 276–292.

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, UK: Cambridge University Press.

Beringer, J. (1996). *Experimental Run Time System (ERTS)*. [*Computer software*]. Frankfurt, Germany: BeriSoft.

Bitsios, P., Prettyman, R., & Szabadi, E. (1996). Changes in autonomic function with age: A study of pupillary kinetics in healthy young and old people. *Age and Ageing*, *25*, 432–438.

Craik, F. I. M., & Byrd, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 191–211). New York: Plenum Press.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33*, 457–461.

Granholm, E., Morris, S. K., Sarkin, A. J., Asarnow, R. F., & Jeste, D. V. (1997). Pupillary responses index overload of working memory resources in schizophrenia. *Journal of Abnormal Psychology*, *106*, 458–467.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*, 1190–1192.

Hoeks, B., & Levelt, W. J. M. (1993). Pupillary dilation as a measure of attention: A quantitative system analysis. *Behavior Research Methods, Instruments, & Computers*, *25*, 16–26.

Hyönä, J., Tommola, J., & Alaja, A. M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *Quarterly Journal of Experimental Psychology*, *48A*, 598–612.

Janisse, M. P. (1977). *Pupillometry: The psychology of the pupillary response*. Washington, DC: Hemisphere.

Jennings, J. R., & Stine, L. A. (2000). Salient method, design, and analysis concerns. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 870–899). Cambridge, UK: Cambridge University Press.

Jin, P. (1992). Towards a reconceptualization of the law of initial value. *Psychological Bulletin*, *111*, 176–184.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology*, *47*, 310–339.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*, 1583–1585.

Kim, M., Beversdorf, D. Q., & Heilman, K. M. (2000). Arousal response with aging: Pupillographic study. *Journal of the International Neuropsychological Society*, *6*, 348–350.

Korczyn, A. D., Laor, N., & Nemet, P. (1976). Sympathetic pupillary tone in old age. *Archives of Ophthalmology*, *94*, 1905–1906.

Loewenfeld, I. E. (1979). Pupillary changes related to age. In H. S. Thompson, R. Daroff, L. Frisen, J. S. Glaser, & M. D. Sanders (Eds.), *Topics in neuro-ophthalmology* (pp. 124–150). Baltimore: Williams & Wilkins.

Matthews, G., Middleton, W., Gilmartin, B., & Bullimore, M. A. (1991). Pupillary diameter and cognitive load. *Journal of Psychophysiology*, *5*, 265–271.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*, 40–48.

Morris, S. K., Granholm, E., Sarkin, A. J., & Jeste, D. V. (1997). Effects of schizophrenia and aging on pupillographic measures of working memory. *Schizophrenia Research*, *27*, 119–128.

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors*, *33*, 17–33.

Oka, S., Chapman, C. R., & Jacobson, R. C. (2000). Phasic pupil dilation response to noxious stimulation: Effects of conduction distance, sex, and age. *Journal of Psychophysiology*, *14*, 97–105.

Paas, F., Tuovinen, J. E., Tabbers, H. K., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*, 63–71.

Paas, F., & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, *86*, 122–133.

Paas, F., Van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, *79*, 419–430.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology*, *11*, 559–566.

Perfect, T. J., & Maylor, E. A. (2000). Rejecting the dull hypothesis: The relation between method and theory in cognitive aging research. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 1–18). Oxford, UK: Oxford University Press.

Salthouse, T. A. (2000). Steps toward the explanation of adult age differences in cognition. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 19–49). Oxford, UK: Oxford University Press.

Sekuler, R., & Owsley, C. (1983). Visual manifestations of biological aging. *Experimental Aging Research*, *9*, 253–255.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654.

Sternberg, S. (1967). Retrieval of contextual information from memory. *Psychonomic Science*, *8*, 55–56.

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2002). Cognitive load theory and aging: Effects of worked examples on training efficiency. *Learning and Instruction*, *12*, 87–105.

Verney, P., Granholm, E., & Dionisio, D. P. (2001). Pupillary responses and processing resources on the visual backward masking task. *Psychophysiology*, *38*, 76–83.

Wierwille, W. W., Rahimi, M., & Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. *Human Factors*, *27*, 489–502.

Wilder, J. (1967). *Stimulus and response: The law of initial value*. Bristol, UK: Wright.

Winn, B., Whitaker, D., Elliott, D. B., & Phillips, N. J. (1994). Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology & Visual Science*, *35*, 1132–1137.