Naval Research Laboratory
Washington, DC 20375-5320

NRL/5580/MR—2022/3

Final Report: Pupillary Dynamics More than Meets the Eye

JOSEPH COYNE

CIARA SIBLEY

Information Management and Decision Architectures Branch
Information Technology Division

March 15, 2022

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
15-03-2022

2. REPORT TYPE
NRL Memorandum Report

3. DATES COVERED *(From - To)*
01-01-2018 – 12-31-2021

4. TITLE AND SUBTITLE
Final Report: Pupillary Dynamics More than Meets the Eye

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Joseph Coyne and Ciara Sibley

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
1L45

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/5580/MR--2022/3

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375-5320

10. SPONSOR / MONITOR'S ACRONYM(S)
NRL

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report details a three year effort to expand our understanding of pupillary dynamics and improve the methods by which cognitive science researchers measure pupil diameter. Researchers established in the early 1960s that pupil diameter increases during cognitive effort, and that the size of this increase is proportional to the magnitude of that effort. More recent research has suggested that pupil size at rest is significantly correlated with cognitive ability. However, after conducting several studies investigating resting pupil size and cognitive ability, with a combined sample of over 500 active duty personnel, we were unable to replicate this finding; calling this work into question. Additionally, these results were obtained after we developed a new technique to improve the accuracy of remote eye tracking systems by computing a known reference for each participant, using a fiducial marker and inter-pupillary distance. The lack of ground truth has been a long-standing issue for remote eye tracking systems and this research effort provided a solution. Furthermore, although we found no link between pupil size and ability we did find that the degree of variability is correlated with an individual's cognitive ability. This report discusses future applications of pupil diameter measurement for the DoD.

15. SUBJECT TERMS
Pupil diameter, Eye tracking, Physiological measurement, Cognitive ability, Individual differences

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	U	9	Joseph Coyne

19b. TELEPHONE NUMBER *(include area code)*
(703) 346-9002

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This page intentionally left blank.

Final Report: Pupillary Dynamics Closeout

Introduction and Objective

Since the 1960s pupil size has been used as an indicator of cognitive effort and engagement across a broad range of cognitive functions, to include decision making, attention control, memory retrieval, and analogical reasoning. Recent research tied pupillary fluctuations to activity in the locus coeruleus – norepinephrine system (LC-NE), which is believed to play a role in the control of attentional resources, with even brief spikes of activation to a stimulus being immediately captured in the pupil. The LC is the source of norepinephrine (NE) projections throughout the brain, which mobilizes the brain and body for action. NE is one of the primary neurotransmitters involved in our "flight or fight" response, increasing overall arousal and focusing attention. Eye tracking can serve as an unobtrusive and cost effective means to capture activation in this important neural system. The recent availability of low-cost eye tracking hardware (e.g., the Gazepoint GP3 costs $695) provides the opportunity to exploit this association outside the laboratory and provide streaming information about individuals' fluctuating level of processing load and attention, while engaged with real world systems.

Research over the last 60 years on the pupillary system has been hyper-focused on identifying peak mental exertion in response to very brief (e.g., 1-3 sec) stimuli exposure, and have all but ignored the variability present within the pupillary response over time, and across individuals. Studies generally extract one point of the pupillary signature, ignoring the rest of the signal, and treat all individuals as the same, aggregating data across individuals and grouping by cognitive demand of the stimuli. Research is limited on the variability that exists within and across individual pupil signals, although this information is important for application and exploitation purposes. Additionally, there is a paucity of pupillometry research during tasks that require sustained attention (e.g., monitoring a sensor feed), the cognitive equivalent of running a marathon.

The United States military increasingly relies on automation to provide new capabilities, enhance warfighter effectiveness, and reduce manpower requirements. Nonetheless, the human still plays a central role in monitoring and directing automated systems, and all the benefits that come from automation also come at a cost to human performance as humans are relegated to a supervisory role. Nearly 100 years of literature has consistently shown that most humans perform poorly at monitoring tasks, even after short durations of time (e.g., vigilance decrement), and as reliance on automation continues to increase, human performance will further suffer. As such, new methods are needed to 1) identify which individuals have the ability to perform well in a highly automated environment, and 2) capture when an individual is in a suboptimal state. Pupil diameter has the ability to support both of these objectives.

Researchers recently revealed that individuals with greater fluid intelligence have significantly larger (~1 mm) resting pupil sizes Tsukahara, Harrison, & Engle, (2016), theorizing that these differences are indicative of different levels of baseline activation in the brain. This suggests that pupil size may have some practical application in understanding differences across individuals.

Manuscript approved March 11, 2022.

This could have implications in the area of selection and classification, which is a critical element of manning in the DoD.

Further, pupillometry is an ideal measure to reveal fluctuating cognitive states within an individual, over time. Tasks such as air traffic control or unmanned vehicle operation require individuals to focus their attention over long periods of time. Pupil diameter and eye tracking can provide insight into how attention and brain activation are fluctuating during that time. This can be invaluable in detecting when an individual is in a sub optimal state.

Our overall objective is to capitalize on recent scientific discoveries linking brain activation and pupillary dynamics, and improve human cognitive performance prediction by augmenting existing models with pupil data. Our goal is to better understand how pupil diameter can be used to both measure differences between individuals as well as fluctuations of cognitive state within individuals. As part of this objective we will be focusing on new ways of analyzing pupil data so as to make comparisons across individuals and tasks.

Technical Approach

This work was designed to exploit recent research findings that fluctuations in pupil diameter capture activation of the Locus coeruleus-Norepinephrine (LC-NE) system, and that differences in cognitive ability may be related to baseline LC-NE activation, which is measurable via resting pupil diameter. Additionally, this work aimed to explore how task based variations in pupillary response can be used to capture the increased resources necessary to complete cognitively demanding tasks.

The first emphasis of our approach was improving the way in which pupil data can be analyzed and used to understand individual differences in cognitive ability. One of the unique challenges of using pupil diameter data to understand individual differences is that most eye tracking companies point to the limitations in the data. A known issue with all remote eye-tracking systems (regardless of cost) is that they do not provide a reference point of known size against which to calibrate pupil size measurements from the device. As such, manufacturers warn that "Because video-based eye trackers don't report an actual ground-truth, physical pupil diameter (most commonly a model-generated value either in millimeters or arbitrary units) it's advised to use only relative change measures. For example, percentage dilation." (Tobii Pro, 2021). Given that the latest research shows that baseline pupil size (reported in mm) –measured with remote trackers without the use of a chin rest – is correlated with fluid intelligence, we needed to understand and ensure our pupil size measurements are accurate.

Data collected initially for this effort revealed that pupil diameter collected via low cost systems temporally changed with task demands, however the link between pupil size and cognitive ability was either not present or in the opposite direction. As such much of our initial approach was directed at improving the accuracy of pupil diameter measurements. This involved numerous experiments aimed at quantifying the accuracy of raw size measurements and comparing different techniques (e.g., use of a fiducial marker, and the use of interpupillary distance measured via corneal reflection pupillometer) for correcting those values and reducing measurement noise and error. Our findings from this effort have been incorporated into the Gazepoint GP3's algorithmic computation of pupil diameter in mm. Although their new measurement reduces noise generated due to head movements, our research still suggests that either a fiducial marker, or a measurement of interpupillary distance should be used if the raw size of the pupil is of interest, as research in individual differences suggests that it is.

Our new correction techniques allowed us to compare absolute changes in pupil diameter across participants, and test the hypothesis that baseline pupil size is correlated with cognitive ability. However, our initial testing using a fiducial marker with student naval aviators (SNA) failed to replicate previous research and did not identify any relationship between resting pupil size and cognitive ability. The second emphasis of our approach was figuring out how to collect a large amount of human subject data from a diverse set of individuals, i.e., people with variable levels of cognitive abilities. This challenge was made more difficult as a terrorist event at the flight school at NAS Pensacola prevented data for over 2 months and was quickly followed by the Covid 19 pandemic. In spite of data collection being halted for 15 months, we resumed in 2021 with a very diverse sample. We were able to couple our research in pupillary response with our new research in investigating the use of attention control as a predictor or military success. During the Pandemic we looked at standardized scores on the ASVAB for different enlisted schools and were able to work with several schools (Air Traffic Control, Aviation Ordnance, and Aviation Maintenance) and gain access to their students. The immediate benefit to the school houses was the possibility of improving their current selection tests, however this also afforded us the opportunity to gather pupil data from all of their participants.

Over the course of 11 months, we collected eye tracking and cognitive performance data from over 1100 sailors and marines. This provided us with an immense amount of high quality pupil data to mine and explore for correlates of cognitive performance. One large analysis effort involved replicating the association between baseline pupil size and cognitive ability. The scope of this replication effort and the diversity of the sample call into question the previous research, which was published in a number of popular science outlets. This failed replication uncovered additional concerns with this line of research: one issue we uncovered is instability within resting pupil size that suggests baseline pupil size may have issues with reliability. This was discovered by having participants record pupil data at multiple computer stations within the same room and during the same data collection session. This instability was neither attributable to fatigue nor lighting conditions at the computer station. However, we also uncovered some initial evidence that links variation in baseline pupil activity with cognitive performance. This may suggest that individuals with more variable pupil size have less attention control, an important cognitive ability which is theorized to be linked to performance in a broad array of cognitively demanding tasks.

Technical Progress

FY 19

- Found a small but negative correlation between baseline pupil size and cognitive ability (working memory capacity) in a group of 90 student naval aviators at Naval Air Station (NAS) Pensacola. These results contradicted results obtained by other researchers at GA Tech who used a larger more diverse sample, and a more expensive eye tracking system. This null result shifted our focus on improving analysis techniques of pupil data towards addressing the issue of how a remote eye tracking system can measure pupil size in mm without having a known reference. This is a requirement for comparing absolute pupil size differences between individuals. This issue is a problem for all remote eye tracking systems, not just low cost ones. Tobii manufactures a number of more expensive remote eye tracking systems and note that pupil diameter measurements are outputted in millimeters, however they should be treated as an arbitrary unit and recommends normalizing changes in pupil size within an individual versus discussing absolute differences in millimeters.

- Identified that inter pupillary distance (IPD), the distance between an individual's eyes could be used as a known reference when measuring an individual's pupil size. IPD can be measured using a corneal reflection pupilometer (CRP), a device used by optometrists when prescribing glasses. An individual's IPD has a small amount of variation based upon the distance at which the eyes are fixating on an object, however the distance while looking at a computer display will have very little variation within an individual.
- Demonstrated that use of IPD could be used to compute pupil size in mm using pixel measurements of pupil size and position within the eye trackers camera, and significantly reduce the noise in pupil size generated by head movement compared to the systems algorithmic computation of mm.
- Demonstrated that normative IPD data (based upon population averages for males and females) could be used to reduce pupil size measurement error within a low cost eye tracking system due to noise from head movements.
- Revealed pupil measurement problems with low cost eye tracking manufacturer, Gazepoint, and worked with them on developing and validating a new alternative correction technique using a fiducial marker worn on a pair of lens-less glasses.
- Demonstrated that low cost eye tracking could be used to measure within task learning by extracting phasic and tonic pupillary response over time within a group. These results improved and extended previous research by using a new pre-processing technique for converting pixel size to millimeters, and by accounting for tonic activation.

FY 20

- Demonstrated that pupil size measurements computed from CRP measured IPD and fiducial marker were nearly identical with an average difference of 0.08mm. This average difference between values computed by both techniques and those computed by the system's algorithmic estimation of pupil size was .36mm. This is considerable since the magnitude of pupillary response to a cognitive stimulus is typically less than .5mm.
- Found that a hybrid technique where an individual's IPD is measured using a fiducial marker could overcome limitations of both correction techniques. Measurement using the CRP is not conducive to group data collection, however the use of a fiducial marker resulted on average a loss of approximately 10% of the eye tracking data collected. The hybrid technique has a mean difference of only .03 mm and results in no data loss.
- NRLs use of IPD as a correction technique was adopted by the low cost eye tracking company Gazepoint and a software update significantly improved the quality of their algorithmic estimation of mm.
- NRL purchased and installed new higher resolution, low cost eye tracking systems (150 Hz vs 60 Hz) with fixed mounts at NAS Pensacola's Operational Psychology Department in an effort to further improve signal quality and reduce data loss due to head movements.
- Due to an act of Terrorism at NAS Pensacola, data collection with student Naval aviators was suspended from November 2019 to February 2020. All in person data collection at

NAS Pensacola was further suspended due to Covid 19 beginning in March 2020 and continuing into FY21.

- Met with other military schools, air traffic control, aviation maintenance and aviation ordinance to discuss the importance of our research to the Navy and use of their students in our research once in person data collection was possible.
- Demonstrated large individual variability exists in pupil data during a memory task, between individuals and even within individuals on a trial by trial basis. The aggregate group pupil data collected demonstrated pupil diameter was sensitive to changes in task difficulty and pupil.

FY 21

- Due to Covid 19, in person data collection did not resume until February 2021 and once data collection resumed the number of people we obtained data from simultaneously was cut in half to ensure adequate social distancing.
- Collected baseline pupil size data and cognitive ability data on a large sample (over 500 sailors and marines) with a broad range of cognitive abilities and failed to find a link between baseline pupil size and cognitive ability. These latest data seriously call into question scientific results linking pupil size and cognitive ability that received much attention in the popular press in 2021.
- Demonstrated the baseline resting pupil size within individuals can have large fluctuations within many individuals (> 2 mm) further calling into question the use of baseline pupil size as a reliable measurement of cognitive ability. The measurements were all collected on the same day with minimal cognitive testing (only 15 minutes) so fatigue should not have been a factor in the variable pupil sizes. Furthermore, luminance was recorded at each station, using a light meter, and the small variation in luminance at each testing station did not account for the pupil variability.
- Found evidence that stability in baseline pupil size is significantly correlated with cognitive ability. We hypothesize there may be a link to pupil stability and attention control which may be used to account for individual differences. Large variations in pupil size during baseline may be associated with cognitive states such as mind wandering.

Naval / Marine Corps Needs:

As the world's largest employer, the optimal function of the DoD depends on understanding individuals' cognitive abilities. Pupil diameter provides an unobtrusive and inexpensive means of gathering real time data on brain activation that can be used for a host of DoD relevant issues, ranging from selection and understanding an individual's capability to perform a certain job, to assessing a sailor's cognitive state in real time.

Dual Use:

Research in pupillary dynamics extends beyond just the military. Our research on improving pupil diameter assessment has already been incorporated by several universities and has been adopted by the only manufacturer of low cost eye tracking systems that are designed for research. Pupil diameter provides a means of inexpensively assessing an individual's cognitive state. Outside of understanding cognition and individual differences it can be used to assess learning, monitor cognitive state (e.g., fatigue), or even be used to help diagnose traumatic brain injuries (TBI).

Transition Plan:

NRL's Pupillary dynamics effort helped establish a new capability to collect accurate pupil data that can be reliably compared across individuals. The collection of baseline pupil size and in-task pupil size will continue as part of NRL's research on improving selection testing for Naval Aviators as well as our research focused on selecting enlisted service members. NRL's ability to gather large quantities of pupil data from a healthy and cognitively diverse sample is unique. NRL is planning to work with the Defense Health Agency to gather baseline pupillary response data on a large population of healthy sailors to aid them in their efforts to detect TBI. The task DHA is using looks at attention is short, uses the same low cost eye tracking systems, and fits within our cognitive assessment battery. NRL will be able to leverage our unique pupil data gathering capability to benefit both the DoD and the broader healthcare community.

With respect to our ongoing selection test development, pupil size can help understand how motivation and an individual's cognitive state can impact test performance over an extended battery of cognitive ability tasks. The ability to detect fatigue within a testing session can improve how we analyze different service member's data and a test's effectiveness. Further, this research in individual differences will be used to continue to expand our knowledge and demonstrate how pupillary dynamics can be used to benefit the fleet.

Publications from Effort:

Flynn, S. R., Quartuccio, J. S., Sibley, C., & Coyne, J. T. (2018). Variation in Pupil Diameter by Day and Time of Day. In *International Conference on Engineering Psychology and Cognitive Ergonomics* (pp. 296-305). Springer, Cham.

Sibley, C., Foroughi, C., Brown, N., & Coyne, J. T. (2018). Low Cost Eye Tracking: Ready for Individual Differences Research?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 741-745). Sage CA: Los Angeles, CA: SAGE Publications.

Coyne, J.T., Moclaire, C., Brown, N., Foroughi, C., & Sibley, C. (2019) Normative Interpupillary Distance Data Reduces Pupil Size Noise in a Remote Eye Tracking System. In Proceedings of the 20th International Symposium on Aviation Psychology. Dayton, OH.

Coyne, J. T., Brown, N., Foroughi, C. K., & Sibley, C. M. (2019). Improving Pupil Diameter Measurement Accuracy in a Remote Eye Tracking System. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 63, No. 1, pp. 49-53). Sage CA: Los Angeles, CA: SAGE Publications.

Sibley, C., Foroughi, C. K., Moclaire, C. M., King, K. M., Brown, N. L., & Coyne, J. T. (2019). Additional Evidence for Pupil Size as a Measure of Within-Task Learning. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 63, No. 1, pp. 54-58). Sage CA: Los Angeles, CA: SAGE Publications.

Sibley, C., Foroughi, C. K., Brown, N. L., Phillips, H., Drollinger, S., Eagle, M., & Coyne, J. T. (2020). More than Means: Characterizing Individual Differences in Pupillary Dilations. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 64, No. 1, pp. 57-61). Sage CA: Los Angeles, CA: SAGE Publications.

Sibley, C., Foroughi, C., Brown, N., Drollinger, S., Phillips, H., & Coyne, J. (2020). Augmenting Traditional Performance Analyses with Eye Tracking Metrics. In *International Conference on Applied Human Factors and Ergonomics* (pp. 118-125). Springer, Cham.

Robison, M., Coyne, J., Sibley, C., Brown, N.L., Neilson, B., & Foroughi, C. (*Submitted, 2022*). An Examination of relations between baseline pupil measures and cognitive ability