Investigating Pupil Dilation in Decision Research

Joseph Tao-yi Wang*[1] and Wei James Chen[2]

National Taiwan University[1] and The Chinese University of Hong Kong[2]

Pupil dilation (PD) or the measure of pupil diameter/size has been used in the decision research starting from the 1960s (Hess and Polt, 1960, 1964; Kahneman and Beatty, 1966; Hicks, Reaney, and Hill, 1967). Since then, pupil dilation has been viewed as an index of arousal and cognitive load (Beatty, 1982). In more recent years, the finding that the pupil diameter is highly correlated with the Locus Coeruleus (LC) neuron's firing activity (Rajkowski et al., 1994) gave raised to a neural foundation of what pupil dilation indicates, i.e. pupil dilation is linked with neural gain and the exploration and exploitation trade-off (Aston-Jones and Cohen, 2005).

Regarding pupil dilation as an emotional reaction, Hess and Polt (1960) reported pupillary dilation responses to what they call "emotionally toned or interesting visual stimuli". Other early studies reporting pupillary responses to specifically sexual arousal include Hicks, Reaney, and Hill (1967), and Bull and Shead (1979). More recently, Aboyoun and Dabbs (1998) also reported pupillary responses to arousal. Oka et al. (2000) reported pupillary responses when pain was administered to either fingertips and ear lobes.

Regarding task-evoked pupillary responses to cognitive load, Hess and Polt (1964) first reported differential pupillary dilation responses while mentally calculating the product of two one digit or two digit numbers. Kahneman and Beatty (1966) showed how more difficult memory tasks (memorizing numbers with more digits vs. less digits) induced larger pupillary response. Later studies establish pupillary dilation responses as an indicator of cognitive load in various tasks, including signal detection and letter matching. More recently, Bailey and Iqbal (2008) utilized pupillary responses to provide a steady stream of workload data to study its changes within a single task.

For single responses, pupil dilation occurs and peaks several seconds after stimuli were presented, with timing depending on the type of stimuli. For example, Hess (1972) reported that such dilation would occur two to seven seconds after emotional stimuli were presented and suggested faster dilation for stronger stimuli. Chapman et al. (1999) found that pupil dilation responses to pain began at 330 ms and peaked at 1250 ms after stimulus onset. Peak dilation increased significantly as pain intensity increased. Using different sounds as stimuli (baby crying, laugh or regular office noise), Partala and Surakka (2003) showed that there was first seemingly no response for about 400 ms, and then a steep increase in pupil size peaking at two to three seconds after stimulus onset. When performing a cognitively demanding task, pupils dilate in responses to mental workload, and peak at about one to two seconds after the onset of demand (Beatty, 1982). Pupils constrict after the task is completed, either gradually (Kahneman and Beatty, 1966; Hess, 1972) or instantly (Bernhardt et al, 1996).

Recently, Locus Coeruleus (LC) neuron's firing activity is found to be highly correlated with pupil diameter in primates (Rajkowski et al., 1994).[i] LC is a nucleus located in the pons, part of the brainstem. It has wide projections throughout the entire brain, and it is the primary source of norepinephrine (NE), which is an important neurotransmitter and neuromodulator that, among other functions, increases arousal and focuses attention. Together, the LC and the areas of the brain that are affected by NE are jointly called the LC-NE system. Since there is a positive relationship between NE and pupillary response, pupil dilation can be viewed as an indirect but non-invasive measure of NE.

In fact, Aston-Jones and Cohen (2005) proposed an integrated model, the Adaptive Gain Theory (AGT), of the LC-NE system. In particular, AGT predicts that the LC phasic mode (pupil dilation) is associated with exploitation behavior that optimizes current task performance, while the LC tonic mode (pupil constriction or baseline pupil) is associated with worse performance (in the current task) but broader attention that induces exploration behavior (of other rewarding tasks). Several researchers have demonstrated this effect in different contexts, including ambiguous perceptual identification (Einhäuser, Stout, Koch, and Carter, 2008; Hupe, Lamirel, and Lorenceau, 2009), perceptual discrimination (Gabay, Pertzov, and Henik, 2011; Gilzenrat, Nieuwenhuis, Jepma, and Cohen, 2010), and gambling (Jepma and Nieuwenhuis, 2011). The link to breadth of attention has been shown in several studies including visual and semantic learning (Eldar et al., 2013), ambiguous letter identification (Eldar et al., 2016) and binary food-choice (Chen and Krajbich, 2017b).

This chapter is designed for researchers who are interested in using pupil diameter as a data input in their research but have little or no experience before. Therefore, the authors will talk about the modern method of measuring pupil diameter (section II), the common environmental control (luminance, visual angle, etc.) and experimental design that need to be taken care of in pupil dilation research (section III), post-experiment data analysis method (section IV), and the difficulties in interpreting pupil dilation results along with some solutions (section V). We will end this chapter with some examples of actual pupil dilation research (section VI).

II. Modern Method of Measuring Pupil Diameter

There are several different methods to measure pupil dilation. See Duchowski, (2007) for a complete review of all the different techniques. Among these methods, one of the most reliable and non-invasive method utilizes high speed video cameras. In particular, video-based eye-trackers put cameras and infrared illuminators in front of subject's eyes, and videotape subject's pupil. This is typically performed by either placing cameras in front of the computer screen subjects are viewing (desk-mount), or by placing cameras on a head restraint similar to a bicycle helmet so they are located right in front of each eye (head-mount). Since images of the pupil are recorded, the eye-tracker is able to measure pupil diameter by either counting the number of pixels of the (dark color) pupillary area or fitting an ellipse on the pupil image and calculating the length of the major axis. Note that when using video cameras, pupil diameter is usually reported in relative terms, since actual image size depends on camera position. This is in contrast to traditional methods that focus on absolute measure of pupil dilation.

As discussed in Klingner et al. (2008), the pixel-counting method is subject to the pupil foreshortening error (PFE), namely that the pupil diameter varies with the gaze direction, while the ellipse-fitting method is not. However, a more recent paper (Hayes and Petrov, 2015) explicitly models the geometric relationship between gaze angle and the change in observed pupil diameter. They proposed a simple method that uses the gaze data to correct the PFE, which can reduce it by 82.5%, and a slightly more complicated method can reduce it by 97.5%. Therefore, these two pupil measuring methods should be treated equally now.

III. Experimental Design

Experimental design for pupil dilation research differ from the usual eye-tracking experiments for several reasons.

First of all, pupil dilation experiments generally need longer inter-trial interval (ITI, up to 2000 ms longer) and longer trial time (at least 1000 ms longer). This is because the cognitive effect on the pupil dilation has a delay about 1000 ms and most research define pupil dilation as a percentage relative to a pre-stimuli baseline from 1000 to 2000 ms (Beatty and Lucero-Wagoner, 2000). For example, in a repeated binary-choice experiment, if we only want to obtain the gaze data, the inter-trial interval can be as short as possible (as long as it is sufficient for subjects to fixate at the center) and each trial can be ended once subjects made their decision. However, if we want to further obtain the pupil diameter data, a longer inter-trial interval will be needed to establish the baseline, and after subjects made their decision, we still want to record the pupil diameter for an additional 1000 to 2000 ms.

Second, pupil dilation is heavily affected by the luminance level on the screen. This is typically handled by controlling luminance during the course of the experiment. For example, Beatty (1982) contains instructions on producing stimulus slides with comparable brightness and contrast, and most recent work (Einhäuser et al., 2008; Einhäuser, 2010; Eldar et al., 2016; Chen and Krajbich, 2017a) use isoluminant stimuli. This is more easily done with computerized display, but still remains a challenge in many naturally-occurring settings, such as flying an airplane at night or viewing online search results that contain images of different colors. For example, Dehais, Causse, and Pastor (2008) embedded an eye-tracker in a real aircraft to record pupillary responses (as well as fixations) when pilots performed different flying routines, but eventually decided to analyze pupil dilation for only a subset of the pilots to make luminance conditions comparable. Similar to the luminance problem, pupil dilation is also affected by the ambient light and noise. Hence, a common practice to handle this problem is to use a dim/dark and quiet room when running the experiment (Einhäuser et al., 2008; Chen and Krajbich, 2017a).

Lastly, as mentioned in section II, pupil diameter can be affected by the PFE. Besides the method we mentioned in section II (Hayes and Petrov, 2015), the common practice to handle this problem is to design the experiment so that subjects always gaze at the center position on the monitor. However, this is obviously not a universal solution.

IV. Preprocessing and Data Analysis

Besides the experimental design, the preprocessing and data analysis process for pupil diameter data is also different from other eye-tracking data.

Preprocessing

Video-based eye-trackers report raw data (usually called the "sample report") for each instance of observation. Each record consists of a time stamp, the pupil diameter of the eye(s), the X-Y location of the left and right eye (or just one eye under monocular mode), and possible "messages" that are sent to the eye-tracker. These messages are commonly used as syncing timestamp between the data recorded by the eye-tracker (eye data) and by the experimental program (behavioral data).

First of all, blinks are very common in the pupil data. These are usually recognize by the manufacturer's own data analysis software or by the algorithm designed by the researcher. The most common practice is to do a linear interpolation to remove the blinks, but other method such as cubic-spline fit are also available (Mathôt, 2013).

Second, the pupil data measured by the eye-tracker has a high-frequency noise (Duchowski, 2007). Since pupils dilate and constrict at low frequency, it is a common practice to smooth the data with a low-pass filter. Klingner et al. (2008) provides a detailed method of how to determine the cutoff frequency for this filter.

Lastly, some other corrections may be required depending on the experimental design. For example, if the experimental design does not restrict subjects to always gaze at some point on the monitor and the stimuli are not isoluminant, then the correction for the PFE and luminance level will be needed. Note that the correction for luminance is usually done by regressing out the effect of different luminance levels. It is not a very precise method. We recommend using isoluminant stimuli for any pupil dilation experiment.

Data Analysis

The first step of data analysis is to establish a baseline for each trial of observation. This is usually defined as the average pupil diameter of the 1000 (up to 2000) ms time interval before the trial onset. Each trial's pupil data after the trial onset is then calculated as a percentage change compared to the baseline in that trial.

There are some common static measurements that are being reported regarding the trial pupil data: mean pupil dilation, peak pupil dilation and latency to peak (Beatty and Lucero-Wagoner, 2000). As for the dynamic measurement, the common practice is to compare the time course of pupil dilation between different experimental treatments/conditions (Hayes and Petrov, 2016; Chen and Krajbich, 2017a).

V. Interpretations of Pupillary Responses

Since there are different causes that could all trigger pupillary responses, care must be taken to identify the exact cause that activated the response to avoid "reverse inference," occurs when interpreting functional magnetic resonance imaging (fMRI) data. Here we discuss some of these issues and possible solutions.

The main challenge regarding interpretation is to isolate the exact cause of the pupil dilation. For studies conducted to observe pupillary responses to specific tasks, this is done by designing control trials that is identical to the treatment trials except for only one particular factor of interest. For example, to identify the pupillary responses to (positive) affect, Partala and Surakka (2003) compared mean pupil size of trials where subjects heard baby laughing (positive affect) with that of trials where subjects heard regular office noise (neutral). Oliveira et al. (2009) asked subjects to investigate three different Google search results sequentially, and determine which was most relevant. They found pupillary responses to the relevance of search results when the results were each shown for exactly five seconds. However, they also noted that their analysis was made possible by separating the search process into several stages, which is typically not the case in actual web searches.

In more naturally-occurring settings, pupillary responses could be compared with other subjective or psychophysical measures, such as skin conductance, heart rate variability, subjective ratings, etc. By combining the results from various measures, researchers are more likely to identify the unique cause that could explain all of them. For example, Kahneman et al. (1969) combined pupillary responses with skin conductance and heart rate to find similar responses in all three measures during information intake and processing, as well as peak responses depending on task difficult. Building on Just and Carpenter (1993)'s comparison of pupillary dilation when processing object-relative sentences ("The reporter that the senator attacked admitted the error") and subject-relative sentences ("The reporter that attacked the senator admitted the error"), Just et al. (1996) performed the same syntactic tasks under functional magnetic resonance imaging (fMRI) to identify the Wernicke's area and the Broca's area as related to more complex language processing.

In some studies, pupil dilation is used to inform interpretation and analysis of other measures, such as fMRI data. In other studies, pupil dilation is aided by other measurements to provide more accurate interpretation. For example, Siegle, Steinhauer, Stenger, Konecky, and Carter (2003) compared the time course of pupil dilation with that of the fMRI signal in the middle frontal gyrus during a digit sorting task to suggest that activity in that area indexed the working memory subtask of digit sorting. In contrast, Kang et al. (2009) find a pupillary response before seeing the answer to an interesting question, which could be attributed to either arousal due to anticipation of the answer, or pain due to the impatience waiting for the answer, but was interpreted as "anticipation of an interesting answer" because the same self-reported curiosity is also shown by fMRI to activate the ventral striatum, a region involved in anticipated reward. Another piece of evidence is the time series of pupil size, which gradually increase during the countdown and peaks immediately after the answer is displayed. This favors the anticipation story since impatience should decrease rather than increase as the remaining waiting time decreases.

In addition to combining measurements, interpretation of pupil dilation can also be aided by relevant theory that explains behavior. In particular, Wang et al. (2010) studied strategic information transmission, which consists of an informed sender (e.g., stock analyst) who sends a (possibly deceptive) message to an uninformed receiver who will then take action (e.g., invest in the stock). They find pupil dilation proportional to the size of the lie (how much subjects inflate the true state), which could be due to simply guilt, or a more complicated process (cognitive difficulty) involved in lying. Wang et al. (2010) eventually affirmed that pupil dilation is evoked by cognitive difficulty because individual differences in subject behavior is explained by a level-k model that generates heterogeneous types each performing different steps of thinking.

Note that the level-k model provides a particular decomposition of the decision-making process which could be empirically tested (with either behavioral or psychophysical data). In fact, researchers in computer science have been performing similar decomposition of tasks using well-known modeling techniques (Card et al., 1983; John, 1995; John and Kieras, 1996). These models of task execution also played important roles in explaining pupillary responses in human-computer interaction. For example, Iqbal et al. (2004) recorded pupillary responses when subjects were performing reading comprehension, mathematical reasoning, product searching, and email classification. Lin et al. (2007) and Lin et al. (2008) observed pupillary responses when subjects played a video game. Both found little difference in pupil dilation for simple and difficult tasks when averaged across entire trials, but Iqbal et al. (2004) did find significant difference in "cognitive subtasks" (but not in "motor subtasks") when they decomposed the original task into several lower-level subtasks.

VI. Applications in Decision-making

In behavioral decision making, Hochman and Yechiam (2011) report pupil dilation when subjects chose between a risky and safe lottery. They find asymmetric responses for gains (vs. losses) in pupil dilation, even when subjects did not exhibit loss aversion behaviorally. They interpret this as supporting the loss signals risk (LSR) hypothesis that a loss signals a threat in the environment and induces arousal and alertness to all outcomes. Yechiam and Hochman (2013) later propose an attentional model to explain this.

In addition, research fields such as cognitive psychology and neuroscience have used pupil diameter as a measure of attention. For example, researchers in e-commerce have attempted to utilize eye-tracking to build automated systems that recognize consumer's attention regarding certain items and respond by offering particular products. In particular, Bee, Prendinger et al. (2006) built an "AutoSelect" system that selects one necktie out of a pair of ties presented on screen and asks subjects to verify if this is what they prefer. This system was mainly based on gaze bias studied by Shimojo et al. (2003), and achieved 81% accuracy rate in selecting subject's preferred choice.

We now turn to recent examples of using pupillary response as an indicator of NE and LC-NE system activation during decision making.

To begin with, Einhäuser et al. (2010) let subjects choose one of the ten seconds in a time interval to push a button, and received a reward if the chosen interval matches a random draw. The authors found maximum pupil dilation occurring one to two seconds after the button press, and take this as supporting evidence that the neurotransmitter NE released by the LC played an important role in reaching a cognitive decision. To rule out alternative causes of pupil dilation, they conducted a series of follow-up experiments: First, they modify the task so subjects instead chose one out of five numbers between 1 and 10. When first showing these numbers in ascending order, the timing of maximum pupil dilation corresponded to the eventually chosen number. Since decisions were made at the end, the increase in pupil dilation during viewing cannot be attributed to button press. To rule out arousal due to reward, the authors also ran the same experiment without payment, yielding similar results. They even ran an instructed-pick version to rule out anticipation since the instruction to pick a particular number was only shown on the spot. This is a good example of careful experimental design for pupil dilation studies.

Secondly, Preuschoff et al. (2011) conducted a gambling task where two numbers were randomly drawn from 1 to 10 and subjects were rewarded for correctly predicting which draw is larger. By showing the two drawn numbers sequentially, the authors induce changes in risk (variance of reward), reward prediction error (actual minus expected reward, RPE) and risk prediction error (actual minus expected size of RPE) as information arrives.

They find pupil dilation corresponding to risk prediction error that measures surprise (instead of expected reward) and propose a risk prediction error hypotheses of NE. Since Preuschoff et al. (2011), much research has shown that pupil dilation signals surprise (or large prediction error) in the Iowa gambling task (Lavín, San Martín, and Rosales Jubal, 2014), while making intertemporal choices (Lempert, Glimcher, and Phelps, 2015), or playing two-person guessing games (Chen and Krajbich, 2017a).

In Nassar et al. (2012), the authors generate a random sequence from a normal distribution with known variance but unknown mean and asked subjects to guess the next number after viewing the current one. Subjects were also told that the unknown mean would change from time to time, but did not know exactly when. Intuitively, the larger the current prediction error, the more likely there was a shift; the more observations one has (after a shift), the more certain one is about the mean. Indeed, a reduced Bayesian model predicts people would estimate and respond to the likelihood the mean has shifted (change-point probability), and the relative uncertainty about the underlying true mean after a shift. The authors found that pupil change during outcome-viewing is positively correlated with change-point probability, linking in to LC phasic mode, except when the prediction was an exact hit (i.e. no change) which also triggered pupil dilation. Also, the average pupil diameter is at the highest level right after the change point (with the largest relative uncertainty), linking it to LC tonic mode.

Finally, Hayes and Petrov (2016) let subjects play a sequence of tasks taken from Raven's Advanced Progressive Matrices (APM) under the think-aloud verbal protocols (speak out what they were thinking, see Chapters 12-14). In these tasks, subjects saw a 3x3 matrix containing eight items with various characteristics and were asked to fill in the last item that fits the pattern. Voice-recording data were used to classify subjects' thinking into pattern recognition and pattern application stages. Subjects' pupils dilate only in the pattern recognition stage, but constrict in the pattern application stage. In fact, those with higher Raven's APM scores have a higher percentage change in pupil diameter (PCPD) averaged across the pattern recognition stage.[ii]

These findings indicate that pupil dilation can be very informative in many different domains, however, its interpretation should be taken carefully since there are multiple causes of pupil dilation. Nonetheless, the LC-NE system and its AGT provide a neural foundation to pupillary response, and hence, has the potential to provide a unifying framework to understand pupil dilation. After all, NE directly increases arousal, and cognitive demanding tasks could trigger the LC phasic mode to induce better performance, both increasing pupil dilation. Future work in this direction would facilitate better interpretation of pupillary responses, and more applications of pupil dilation would surely follow.

References

Aboyoun, D. C., & Dabbs, J. N. (1998). The Hess pupil dilation findings: Sex or novelty? *Social Behavior and Personality*, *26*(4), 415–419.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.*, *28*, 403–450.

Bailey, B. P., & Iqbal, S. T. (2008). Understanding Changes in Mental Workload During Execution of Goal-directed Tasks and Its Application for Interruption Management. *ACM Trans. Comput.-Hum. Interact.*, *14*(4), 21:1–21:28.

Beatty, J. (1982). Phasic Not Tonic Pupillary Responses Vary With Auditory Vigilance Performance. *Psychophysiology*, *19*(2), 167–172.

Beatty, J., & Lucero-Wagoner, B. (2000). The Pupillary System. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd edition, pp. 142–162). Cambridge: Cambridge University Press.

Bee, N., Prendinger, H., Nakasone, A., André, E., & Ishizuka, M. (2006). AutoSelect: What You Want Is What You Get: Real-Time Processing of Visual Attention and Affect. In *Perception and Interactive Technologies* (pp. 40–52). Springer, Berlin, Heidelberg.

Bernhardt, P. C., Dabbs, J. M., & Riad, J. K. (1996). Pupillometry system for use in social psychology. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 61–66.

Bull, R., & Shead, G. (1979). Pupil-Dilation, Sex of Stimulus, and Age and Sex of Observer. *Perceptual and Motor Skills*, *49*(1), 27–30.

Card, S. K., Newell, A., & Moran, T. P. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc.

Chapman, C. R., Oka, S., Bradshaw, D. H., Jacobson, R. C., & Donaldson, G. W. (1999). Phasic pupil dilation response to noxious stimulation in normal volunteers: Relationship to brain evoked potentials and pain report. *Psychophysiology*, *36*(1), 44–52.

Chen, W. J., & Krajbich, I. (2017a). Computational modeling of epiphany learning. *Proceedings of the National Academy of Sciences*, *114*(18), 4637–4642.

Chen, W. J., & Krajbich, I. (2017b). Pupil Dilation and Attention in Value-based Choice. *Working Paper*.

Costa, V. D., & Rudebeck, P. H. (2016). More than Meets the Eye: the Relationship between Pupil Size and Locus Coeruleus Activity. *Neuron*, *89*(1), 8–10.

Crawford, V. P., & Sobel, J. (1982). Strategic Information Transmission. *Econometrica*, *50*(6), 1431–1451.

Dehais, F., Causse, M., & Pastor, J. (2008). Embedded eye tracker in a real aircraft: new perspectives on pilot/aircraft interaction monitoring. Presented at the 3rd International Conference on Research in Air Transportation (ICRAT 2008).

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice* (Vol. 373). Springer Science & Business Media.

Einhäuser, W., Koch, C., & Carter, O. L. (2010). Pupil dilation betrays the timing of decisions. *Frontiers in Human Neuroscience*, *4*(18).

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, *105*(5), 1704–1709.

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature Neuroscience*, *16*(8), 1146–1153.

Eldar, E., Niv, Y., & Cohen, J. D. (2016). Do You See the Forest or the Tree? Neural Gain and Breadth Versus Focus in Perceptual Processing. *Psychological Science*, *27*(12), 1632–1643.

Gabay, S., Pertzov, Y., & Henik, A. (2011). Orienting of attention, pupil size, and the norepinephrine system. *Attention, Perception, & Psychophysics*, *73*(1), 123–129.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252–269.

Hayes, T. R., & Petrov, A. A. (2015). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods*, 1–18.

Hayes, T. R., & Petrov, A. A. (2016). Pupil Diameter Tracks the Exploration–Exploitation Trade-off during Analogical Reasoning and Explains Individual Differences in Fluid Intelligence. *Journal of Cognitive Neuroscience*, *28*(2), 308–318.

Hess, E. H. (1972). Pupillometrics. In N. S. Greenfield & R. A. Sternbach (Eds.), *Handbook of Psychophysiology* (pp. 491–531). Holt, Rinehart & Winston.

Hess, E. H., & Polt, J. M. (1960). Pupil Size as Related to Interest Value of Visual Stimuli. *Science*, *132*(3423), 349–350.

Hess, E. H., & Polt, J. M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science*, *143*(3611), 1190–1192.

Hicks, R. A., Reaney, T., & Hill, L. (1967). Effects of Pupil Size and Facial Angle on Preference for Photographs of a Young Woman. *Perceptual and Motor Skills*, *24*(2), 388-.

Hochman, G. & Yechiam, E. (2011). Loss aversion in the eye and in the heart: The autonomic nervous system's responses to losses. *Journal of Behavioral Decision Making*, 24: 140-156.

Hupe, J. M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision*, *9*(7), 10–10.

Iqbal, S. T., Zheng, X. S., & Bailey, B. P. (2004). Task-evoked pupillary response to mental workload in human-computer interaction. In *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1477–1480). ACM.

Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration–exploitation trade-off: evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*(7), 1587–1596.

John, B. (1995). Why GOMS? *Interactions*, *2*(4), 80–89.

John, B. E., & Kieras, D. E. (1996). The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *3*(4), 320–351.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–234.

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.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science*, *274*, 114–116.

Kahneman, D., & Beatty, J. (1966). Pupil Diameter and Load on Memory. *Science*, *154*(3756), 1583–1585.

Kahneman, D., Tursky, B., Shapiro, D., & Crider, A. (1969). Pupillary, heart rate, and skin resistance changes during a mental task. *Journal of Experimental Psychology*, *79*, 164–167.

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological Science*, *20*(8), 963–973.

Klingner, J., Kumar, R., & Hanrahan, P. (2008). Measuring the task-evoked pupillary response with a remote eye tracker (pp. 69–72). Presented at the Proceedings of the 2008 symposium on Eye tracking research & applications, ACM.

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27.

Lavín, C., San Martín, R., & Rosales Jubal, E. (2014). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience*, *7*.

Lempert, K. M., Glimcher, P. W., & Phelps, E. A. (2015). Emotional arousal and discount rate in intertemporal choice are reference dependent. *Journal of Experimental Psychology: General*, *144*(2), 366–373.

Lin, T., Imamiya, A., Hu, W., & Omata, M. (2007). Combined user physical, physiological and subjective measures for assessing user cost. In *Universal access in ambient intelligence environments* (pp. 304–316). Springer.

Lin, T., Imamiya, A., & Mao, X. (2008). Using multiple data sources to get closer insights into user cost and task performance. *Interacting with Computers*, *20*(3), 364–374.

Mathôt, S. (2013). *A simple way to reconstruct pupil size during eye blinks*. Available.

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus: Pupil Diameter and Locus Coeruleus Activity. *Human Brain Mapping*, *35*(8), 4140–4154.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*(7), 1040–1046.

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*(2), 97–105.

Oliveira, F. T., Aula, A., & Russell, D. M. (2009). Discriminating the relevance of web search results with measures of pupil size. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2209–2212). ACM.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, *59*(1), 185–198.

Preuschoff, K. (2011). Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, *5*.

Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1994). Locus coeruleus activity in monkey: Phasic and tonic changes are associated with altered vigilance. *Brain Research Bulletin*, *35*(5–6), 607–616.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322.

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., & Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *NeuroImage*, *20*(1), 114–124.

Wang, J. T., Spezio, M., & Camerer, C. F. (2010). Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games. *The American Economic Review*, *100*(3), 984–1007.

Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin*, 139(2), 497-518.

[i] Although there is direct (single neuron) evidence only on primates, research with humans (Costa and Rudebeck, 2016; Joshi et al., 2016; Laeng et al., 2012; Murphy et al., 2014) has shown this relationship using fMRI data.

[ii] Hayes and Petrov (2016) refer to the pattern recognition stage as "exploration" and the pattern application stage as "exploitation," which may be confusing: This indicates pupil dilation during "exploration," instead of "exploitation" predicted by AGT, while they claim their results are consistent with Jepma and Nieuwenhuis (2011) which support AGT. Interestingly, they report mean PCPD averaged across segments of the same stage in a trial, instead of baseline pupil diameter used in Jepma and Nieuwenhuis (2011) to represent the LC tonic mode.