



OPEN

The effects of emotional arousal on pupil size depend on luminance

Jasmine Pan^{1,2✉}, Xuelin Sun^{1,3}, Edison Park¹, Marine Kaufmann¹, Michaela Klimova^{1,2}, Joseph T. McGuire^{1,2} & Sam Ling^{1,2✉}

Pupillometry is widely used to measure arousal states. The primary functional role of the pupil, however, is to respond to the luminance of visual inputs. We previously demonstrated that cognitive effort-related arousal interacted multiplicatively with luminance, with the strongest pupillary effects of arousal occurring at low-to-mid luminances ($< 37 \text{ cd/m}^2$), implying a narrow range of conditions ideal for assessing cognitive arousal-driven pupillary differences. Does this generalize to other forms of arousal? To answer this, we assessed luminance-driven pupillary response functions while manipulating emotional arousal, using well-established visual and auditory stimulus sets. At the group level, emotional arousal interacted with the pupillary light response differently from cognitive arousal: the effects occurred primarily at much lower luminances ($< 20 \text{ cd/m}^2$). Analyses at the individual-participant level revealed qualitatively distinct patterns of modulation, with a sizable number of individuals displaying no arousal response to the visual or auditory stimuli, regardless of luminance. Together, our results suggest that effects of arousal on pupil size are not monolithic: different forms of arousal exert different patterns of effects. More practically, our findings suggest that lower luminances create better conditions for measuring pupil-linked arousal, and when selecting ambient luminance levels, consideration of the arousal manipulation and individual differences is critical.

Our pupils are often considered a “window into the mind”^{1,2}: pupillary responses are believed to give researchers access to a person’s arousal state, affect, and cognitive processing. Indeed, numerous studies over the past several decades have linked pupillary responses to changes in behavioral and cortical arousal states^{3–5}, cognitive load and effort⁶, attention^{7,8}, affect^{9,10}, decision-making^{11,12} and more. As a result, pupillometry has become a popular physiological marker of internal state in psychology and neuroscience research, with pupils typically dilating with increased levels of arousal. However, in addition to arousal, the pupils also respond to a variety of different factors – with luminance being one of the primary factors dictating pupil size.

Our pupils dilate in the dark, a response known as the pupillary dark reflex, and constrict in the light, a response known as the pupillary light reflex. Since our pupils are responsive to a combination of environmental factors such as luminance and endogenous factors such as arousal, studies that employ pupillometry to measure arousal often keep the luminance consistent within a study in order to avoid visually-driven confounds. However, different luminances are often employed across studies (e.g. Nassar et al.: screen luminance at $0.457 \pm 0.010 \text{ cd/m}^2$; Niikuni et al.: screen luminance at 7.8 cd/m^2 ^{21,14}). In doing so, researchers have tacitly assumed similar increases in pupil size across all luminance levels.

Recent findings exploring the impact of arousal on pupillary light responses to ambient light intensity changes have begun to challenge the assumption of additivity across luminances, highlighting the crucial role luminance plays in influencing arousal-based pupil responses^{15–19}. In most instances, studies report larger arousal modulations on pupillary responses under *low* luminance conditions compared to *high* luminance conditions^{15–17}. However, some studies have indicated either an absence of task- and arousal-based pupillary differences across luminances, or larger modulations at higher luminances^{19–22}. Overall, gaps persist in our understanding of the relationship and interaction between luminance and arousal on pupillary responses—in particular, all of the aforementioned studies only examined 2 or 3 luminance levels, with highly variable luminance levels and arousal manipulations between studies, making it challenging to reconcile the mixed findings.

Pan et al.¹⁸ addressed some of the discrepancies in previous studies by mapping out pupil size across ten different luminances (0.9 to 233.37 cd/m^2), referred to as the pupillary light function (PLF), and examining its interaction with *cognitive arousal*, which we define as arousal tied with cognitive factors such as effort and task difficulty. The examination of cognitive arousal is built on the idea that cognitive effort and arousal are interrelated, such that arousal is a core concept and that effort and load can drive arousal. The pupillary light function

¹Psychological & Brain Sciences, Boston University, Boston, MA, USA. ²Center for Systems Neuroscience, Boston University, Boston, MA, USA. ³John Hopkins University, Baltimore, MD, USA. ✉email: jasmineep@bu.edu; samling@bu.edu

(PLF) is a mapping of pupil size across a range of luminances, which encompasses both the pupillary light and pupillary dark reflex. The PLF exhibits a nonlinear profile, with pupil sizes typically ranging between 2 and 8 mm (as shown in Fig. 1). At the group level, cognitive arousal, manipulated by varying the difficulty of arithmetic problems, multiplicatively interacted with luminance, with the largest arousal-driven pupillary effects happening mainly at low to mid luminances ($< 37 \text{ cd/m}^2$), and with little-to-no effect of arousal at higher luminances (see Fig. 1). At the individual subject level, there were qualitatively distinct modulatory effects of cognitive arousal on luminance-driven pupillary responses, with individuals showing the largest modulation by arousal at different luminance levels. However, regardless of individual differences, the authors found that overall, high luminances were the least modulated by arousal. In other words, if pupillometry were used as a measure of cognitive arousal concurrently with a high-luminance background, a null result might be obtained merely because of physiological constraints of constricted pupils at high luminances. If the same hypothetical study were conducted using a low ambient luminance level, modulation of pupil size by arousal might be more readily observable. In summary, these findings suggest that there is a narrow range of luminance levels more suited for assessing cognitive arousal-driven pupillary differences.

While the aforementioned study focused on cognitive arousal, the degree to which the findings generalize to other forms of arousal (e.g. physical arousal, affective arousal) remains unknown. Here, we investigate the generalizability of the findings from Pan et al.¹⁸ to arousal driven by affective processing. It has been suggested that cognitive arousal, linked with cognitive load and effort, and emotional arousal, tied with affective processing, modulate pupillary responses by differentially engaging the sympathetic and parasympathetic nervous systems. Specifically, pupil dilation with cognitive arousal is suggested to be mainly mediated through the inhibition of the parasympathetic pathway^{19,23}, whereas emotional arousal is mediated through the sympathetic pathway^{17,24}.

Pupillary changes to luminance are primarily mediated through a balance between the parasympathetic and sympathetic systems: the pupillary light reflex is mainly mediated through the parasympathetic pathway, whereas the pupillary dark reflex is mediated through the sympathetic system²⁵. Studies investigating emotional arousal-based pupillary responses often report larger pupil dilation in dark conditions compared to bright ones, suggesting a predominant role of the sympathetic pathways in mediating these responses^{9,17,24}. Conversely, research investigating cognitive processing-related pupil dilation often report greater pupil dilation in brighter conditions, suggesting the primary role of parasympathetic inhibition in mediating these responses. Pharmacological blocking of the parasympathetic sphincter muscle has also been shown in diminishing cognitive arousal's dilation effects^{17,24}. These differences suggest distinct underlying mechanisms for various arousal-induced pupil responses¹⁷.

Nonetheless, cognitive and emotional arousal likely co-occur for a given task. For instance, challenging cognitive tasks can evoke emotional responses, while emotional stimuli often require cognitive processing. Therefore, while both sympathetic-activation and parasympathetic-inhibition likely contribute to different arousal-induced pupil responses, the precise balance and interactions between these systems across various tasks in modulating pupil size remain unclear. Additionally, the extent to which different arousal manipulations differentially impact pupillary responses across a wide range of luminances is unknown.

In principle, there are at least four distinct ways the PLF might be modulated by arousal (see Fig. 1). We could observe (1) an overall increase in pupil size across all luminance levels (*Additive shift*), (2) an increase in pupil size only at lower luminances (*Maximum response gain*), (3) an increase in pupil size only at mid-luminances (*Inflection point gain*), (4) an increase in pupil size only at higher luminances (*Baseline shift*), or some combination of these patterns. If different types of arousal lead to similar modulation of the PLF, we would expect the effect of emotional arousal to resemble the pattern previously observed for cognitive arousal¹⁸, which involved shifts in both the maximum response and the inflection point, or more specifically, at low and mid-luminances (see Fig. 1).

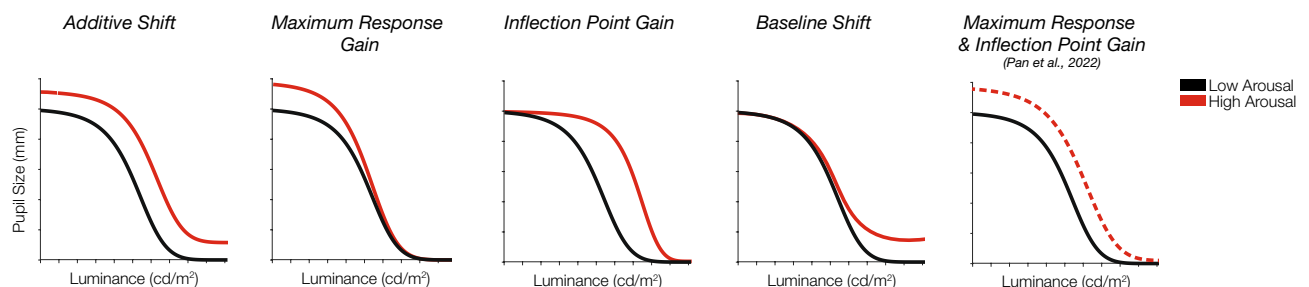


Fig. 1. Possible modulation by arousal on the pupillary light function. The black curve depicts a hypothetical pupillary light response under a low-arousal condition, while the red curves represent the different possible modulations of the pupil response under a higher arousal condition. Different modulations include: (1) An additive shift, which indicates an overall increase in pupil size with increased arousal across all luminance levels. (2) A maximum response gain, in which arousal modulates pupil response at the low luminances. (3) An inflection point gain, or a horizontal shift of the curve, in which arousal's modulation of the pupil response occurs at mid-, but not high- and low-luminances. (4) A baseline shift, wherein arousal affects pupil response only at higher luminances. Pan et al.¹⁸ observed a combination of maximum response gain and inflection point gain modulation by cognitive arousal on the pupillary light response, in which cognitive arousal's effect occurred at mainly at low- and mid-luminances.

To investigate how emotional arousal interacts with luminance, we used affective stimuli from well-established databases, such as the International Affective Digitized Sounds (IADS²⁶), expanded version of the International Affective Digitized Sounds (IADS-E²⁷), and the International Affective Pictures System (IAPS^{10,24}). We selected these stimuli sets for three reasons: (1) the stimuli have been systematically studied and shown to be reliable; (2) all of the stimuli have normative ratings (i.e. using the Self-Assessment Manikin using a nine-point rating scale) along various dimensions of affect, such as arousal, dominance, and valence; and (3) they are extensively used in experiments investigating emotion, arousal, and more. Studies of pupillary responses to affective auditory stimuli using IADS have often found increased pupil size when listening to negative and positive auditory stimuli – categories with higher mean arousal ratings – than when listening to neutral stimuli^{9,17}. This indicates that pupil responses are sensitive to emotional stimulation and affective processing. Likewise, in experiments employing affective visual stimuli using IAPS, negative and positive images result in larger pupil dilation than neutral images^{24,28,29}.

The aims of this study are twofold: (1) to examine whether emotional arousal mimics cognitive arousal in modulating the pupillary light function; and (2) to investigate whether and how affective stimuli of different sensory modalities, visual and auditory, may influence the modulation in distinct ways. In Experiment 1, participants listened to and rated IADS and IADS-E sound clips of low and high emotional arousal while the display cycled through multiple luminances, allowing us to measure the PLF. In Experiment 2A, participants viewed and rated grayscale IAPS visual stimuli under high and low emotional arousal conditions while we measured the PLF. In a control experiment, Experiment 2B, participants viewed the same IAPS visual stimuli as in Experiment 2A, but in color in order to better preserve image content. Overall, we found the largest modulatory effects by emotional arousal occurred mainly at low luminances, regardless of modality. However, this modulatory pattern was contingent on stimulus factors, with specific manipulations to the stimuli playing a role in altering the strength of modulation. Furthermore, at the individual level, we observed qualitatively distinct patterns of modulation, with a substantial proportion of individuals not displaying any pupil-driven arousal response to the stimuli.

Method

Observers

In Experiment 1, thirty-six observers (mean age: 24.08 ± 5.06 ; age range: 19–48; 22 females) participated in this one-session study. Eight observers were excluded from the analysis due to eye-data loss and/or failure to maintain fixation, resulting in a final sample size of twenty-eight observers. In Experiment 2A, thirty-six observers (mean age: 24.71 ± 5.69 ; age range: 19–48; 20 Females) participated in the experiment, with eight observers excluded from the analysis for the same reasons as Expt. 1. Thus, our final sample size in Expt. 2A also consisted of twenty-eight observers. In Experiment 2B, thirty-three observers (mean age: 24.367 ± 4.82 ; age range: 18–35; 19 Females; for detailed description of participants see Supplementary S1) participated in the experiment. Three observers were excluded, resulting in a final sample size of thirty observers. The sample size is consistent with prior pupillometry experiments^{9,17,24,29–34}. Fifteen observers participated in two out of three of the experiments, and three observers participated in all three experiments (see Supplementary S2, S3). All observers had normal or corrected-to-normal vision, provided written informed consent and were compensated for their participation with the exception of the authors. The study was approved by the Boston University Institutional Review Board and conducted in accordance with relevant guidelines and regulations.

Apparatus

The stimuli were created and presented using MATLAB (2017b) in conjunction with Psychophysics Toolbox^{35–37} and displayed on a gamma-corrected, linearized Display + LCD monitor (Cambridge Ltd, 1440×1080 px, 100 Hz). Observers sat and viewed the stimuli from a viewing distance of 116 cm, with their head position stabilized with a chin and forehead rest. The room had no other illumination other than the display screen. Auditory stimuli were presented through Sennheiser HD 202 II headphones, with the volume set to a comfortable level for each observer. The left eye was recorded using Eyelink 1000 Plus Desktop Mount (SR Research) sampled at 1000 Hz, measured using centroid mode, and calibrated using a standard 5-point calibration. Skin conductance was recorded using a Biopac MP160 system. Electrodermal activity (EDA) was collected using Biopac's EDA disposable electrodes (TSD203) and conductive isotonic gel (GEL101). Two electrodes, with gel applied, were placed on the observer's index and middle finger. The physiological measurements were attached to the observer's hand of choice, which was not used for button responses.

Experiment 1: affective sounds (IADS)

Affective auditory stimuli

Affective auditory stimuli were selected from the International Affective Digitized Sounds (IADS; ²⁶) and International Affective Digitized Sounds expanded (IADS-E²⁷) database. Within the IADS and IADS-E databases, each sound has standardized ratings on three dimensions, pleasure, arousal and dominance, derived from the Self-Assessment Manikin affective rating system on a scale from 1 to 9, with 1 being the lowest rating on a given dimension and 9 being the highest^{38,39}. A total of 180 sounds were selected, with 90 sounds classified as negative and 90 sounds as neutral. The high emotional arousal condition consists of negative sound clips with a valence rating range from 1.28 to 2.83 (mean: 2.26; std: 0.40), and an arousal rating range from 4.09 to 8.46 (mean: 6.66; std: 0.85). The low emotional arousal condition consists of neutral sound clips with a valence rating range from 4 to 6.5 (mean: 5.30; std: 0.69), and an arousal rating range from 2.48 to 4.40 (mean: 3.73; std: 0.44). From the IADS stimuli set, a total of 37 sounds (26 negative, and 11 neutral) were selected, and from the IADS-E stimuli set, a total of 143 sounds were selected (64 negative, 79 neutral). Erotic sound clips were excluded (see Supplementary S4 for list of all stimuli used). All sound clips were pre-processed with a Gaussian envelope ($\sigma = 1$) to

soften the attack and decay of the clips, and normalized in volume to prevent potential confounding effects of surprise and loudness on pupil responses.

Procedure

Observers underwent a short practice block prior to the main experiment. During the main experiment, observers maintained fixation on a fixation cross (diameter: 0.1° ; luminance: 15.57 cd/m^2) at the center of the screen to minimize distortions and changes in pupil size due to eye movements. In each 10.5-min block of the task, ten different luminances (0.92 to 233.37 cd/m^2 , log-spaced: $0.92, 1.69, 3.13, 5.80, 10.74, 19.88, 36.81, 68.12, 126.09, 233.37 \text{ cd/m}^2$) were presented on the screen for 6 s each, with the luminance order randomized. Simultaneously, while viewing the luminances, observers listened to the IADS stimuli and provided subjective arousal ratings to each sound on a scale from 1 (low arousal) to 5 (high arousal). Each trial consisted of one sound clip presented for 6 s, during which the observer could respond at any time or within the 1 s response window that followed the sound. Within a given luminance level, 9 auditory stimuli of the same valence (neutral, negative) were presented, with the valence order randomized across luminances within a block, resulting in a total of 90 auditory stimuli and trials per block. Observers completed a total of 4 blocks, leading to a total of 360 trials total: 180 negative trials and 180 neutral trials, with each unique sound appearing twice in a randomized order during the entire experiment (see Fig. 2).

Experiment 2A: grayscale affective images (IAPS)

Affective visual stimuli

Affective visual stimuli were drawn from the International Affective Pictures System IAPS^{10,24} database. A total of 280 visual stimuli were chosen, consisting of 140 negative and 140 neutral images. The negative stimuli set (high emotional arousal condition) had images with a standardized valence rating ranging from 1.31 to 2.95 (mean: 2.19 std: 0.43), and a standardized arousal rating ranging from 5.11 to 7.35 (mean: 6.17 std: 0.58). The neutral stimuli (low emotional arousal condition) set had images with a standardized valence rating ranging between 4.5 and 5.99 (mean: 5.20 std: 0.38), and an arousal rating ranging between 1.72 to 6.79 (mean: 3.65 std: 0.98). We sampled images from all different categories, excluding erotic images (see Supplementary S4 for list of all stimuli used).

Twenty-eight vertically oriented images were cropped in order to fit the same dimensions of the images of the horizontal orientation (1024×768 pixels). The vertically oriented images were all manually cropped to ensure that the primary content of the image was still visible and contained within the new dimensions. Additionally, seventeen pictures that had a surrounding border around the image were also cropped to exclude the border. In order to control for potential confounds in pupillary responses due to differences in luminance and/or spatial frequency content, the images were converted to grayscale and standardized for luminance, ensuring consistent mean luminance (117.26 cd/m^2 , SD: 21.99 cd/m^2) and spatial frequency across the entire stimuli set using the SHINE toolbox⁴⁰.

Procedure

Prior to the main experiment, observers participated in a short practice block. During the main experiment, observers maintained fixation on a fixation cross (diameter: 0.1° ; luminance: 15.57 cd/m^2) throughout each 10.5-min task block. Each block contained 18-s trials. Here, each trial began with a luminance adaptation period (10 s) to allow the pupil to adjust to the luminance being tested. During the adaptation period, one of ten luminances, order-randomized, was presented full-screen. However, at the center of the screen surrounding the fixation cross there was a small gray box (3.8° width \times 2.9° height; luminance: 117.26 cd/m^2) of the same mean luminance as the stimuli. This gray box contained the area where the stimulus would be presented. Following the adaptation period, the affective image stimulus ($3.8^\circ \times 2.9^\circ$) was presented in the gray box for 6 s, with the surrounding luminance unchanged from the adaptation period. The stimulus valence was randomized trial-to-trial. During stimulus presentation, observers were instructed to view the image without providing any key responses. A 2-s response period followed (full-screen luminance: 116.67 cd/m^2 with fixation cross), in which observers indicated their subjective arousal rating on a scale from 1 to 5. Observers were also instructed to click any other key other than the keys 1–5 if the image was confusing or difficult to discern. Each block consisted of 35 trials, and a total of 8 blocks were completed by observers, resulting in a total of 280 trials: 140 negative trials and 140 neutral trials, with each unique image, order-randomized, appearing only once during the entire experiment (see Fig. 3).

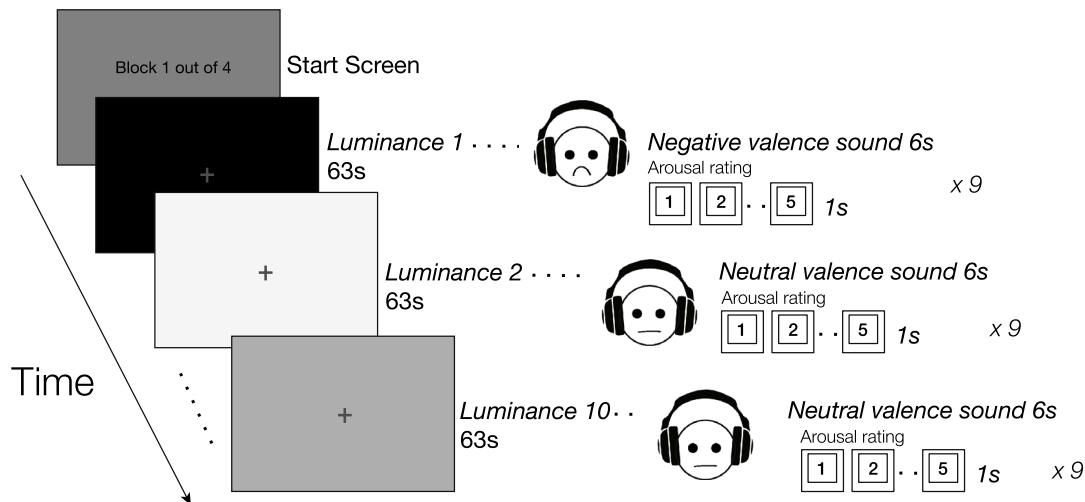
Experiment 2B: color affective images (IAPS)

Affective visual stimuli

Similar to Experiment 2A, the images for this experiment were chosen from the IAPS database, but the stimulus set consisted of a subset of the images used in Experiment 2A. Here a total of 200 images (100 negative, 100 neutral) were selected. The negative stimuli set consisted of images with a standardized valence rating ranging from 1.4 to 2.95 (mean: 2.14 std: 0.40), and a standardized arousal rating ranging from 5.19 to 7.35 (mean: 6.21 std: 0.55). The neutral stimuli set consisted of images with a standardized valence rating ranging from 4.52 to 5.99 (mean: 5.22 std: 0.36), and an arousal rating ranging from 1.72 to 6.79 (mean: 3.48 std: 0.85). Images from all different categories were sampled, excluding erotic images (see Appendix for list of all stimuli used).

In order to address the possibility that the grayscale, luminance and spatial frequency-matched images may have been difficult to discern in Experiment 2A (e.g. blood, tomatoes; see Results, Supplementary S5), the images in this control experiment remained in full color. The mean luminance (mean luminance: 91.61 cd/m^2 ; std: 22.90 cd/m^2) of the colored images was matched across the entire stimuli set using the SHINE_color toolbox⁴¹ to prevent any confounding effect of luminance on pupillary responses to the affective images.

A. Experiment 1: Affective Sound Stimuli (IADS/IADS-E)



B. Experiment 2A & 2B: Affective Picture Stimuli (IAPS)

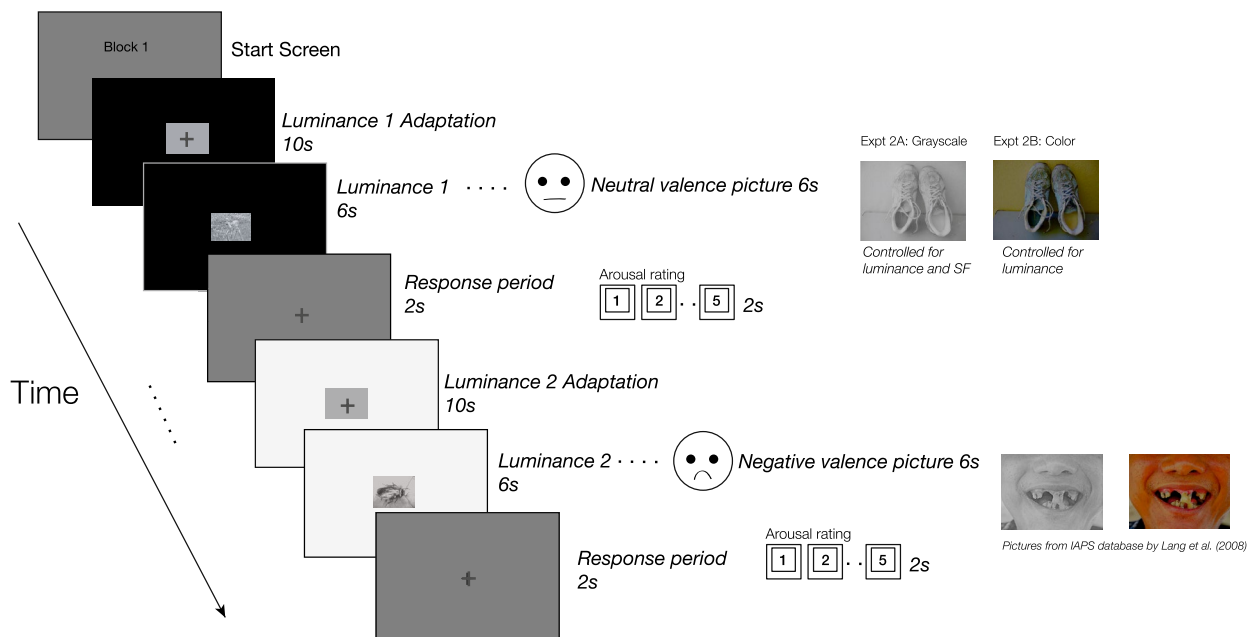


Fig. 2. Experimental design. (A) Example sequence for Experiment 1 using affective sounds. Participants viewed 10 different luminances presented pseudorandomly for 63 s each. While viewing the luminances, participants simultaneously listened to auditory stimuli of Negative and Neutral valence, while rating their subjective arousal for each sound on the keyboard on a scale from 1 to 5. (B) Example sequence for Experiment 2A & 2B using affective pictures. Every trial consisted of three parts: (1) A baseline luminance adaptation period was presented for 10 s to allow the pupil to adjust to the current luminance; (2) After the adaptation period, the image was presented at fixation for 6 s while the background luminance was held constant from the adaptation period; (3) Following the image presentation, the image and the luminance were cleared from the screen, which indicated the start of the 2 s response period. The 10 different luminances were presented in pseudorandom order. While viewing the luminances, participants simultaneously viewed images of Negative and Neutral valence presented centrally and entered their subjective arousal rating for each image on the keyboard on a scale from 1 to 5. In Experiment 2A, the images were presented in grayscale, luminance and spatial frequency-matched. In Experiment 2B, in which the images were presented in color, luminance-matched.

Experiment 1: IADS Affective Sounds

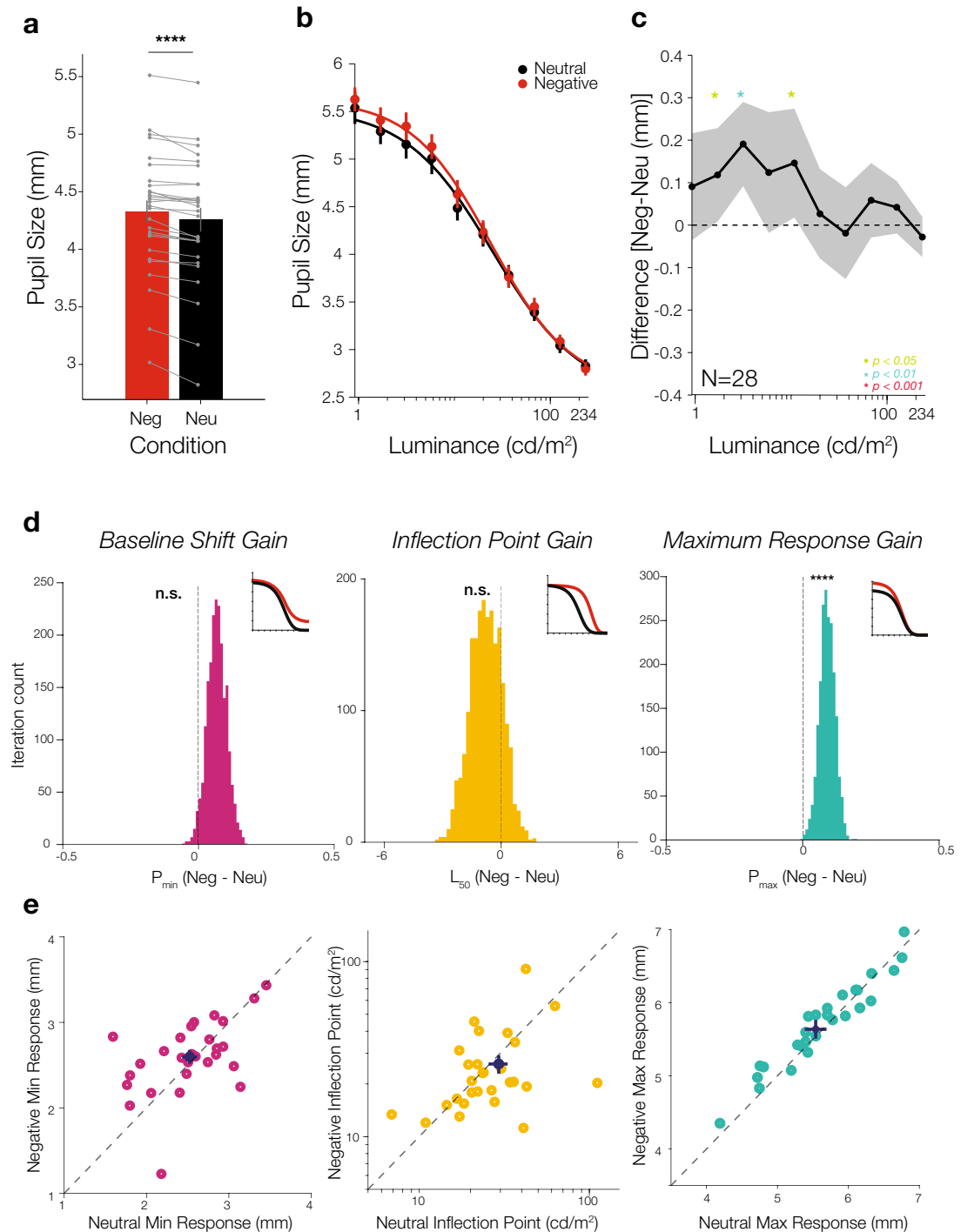


Fig. 3. Group results for IADS. **(a)** Mean pupil size in the Negative and Neutral condition, collapsing across luminance. Each point on the graph represents an individual subject. Overall larger pupil size was observed in the Negative condition compared to Neutral condition. **(b)** Group-level mean pupillary light reflex functions for the Negative and Neutral condition, fit with the modified decrement Naka-Rushton function. Error bars represent SEM. **(c)** The group-average difference in pupil size between the Negative and Neutral conditions was tested separately at each of the ten luminance levels. The shaded gray area represents the 95% confidence interval. **(d)** Bootstrapped distributions (2000 iterations) of the difference between Negative and Neutral parameter estimates for P_{\min} , L_{50} and P_{\max} . The dotted line indicates no difference in the Negative and Neutral parameter estimates. **(e)** Scatterplot of individual observer parameter estimates comparing Neutral versus Negative for P_{\min} (minimum pupil response), L_{50} (inflection point), and P_{\max} (maximum pupil response). The purple dot indicates the group-averaged parameter estimates, with the error bars indicating the SEM. The dotted line indicates the unity line of no difference in parameter estimates between the Negative and Neutral conditions.

Procedure

The procedure for this experiment was mostly the same as Experiment 2A, with a few differences: (1) the size of the gray box (mean luminance: 91.61 cd/m²) in the adaptation period and affective image stimuli was slightly increased in size (5.2° width × 3.9° height), to make it easier for observers to discern the image; (2) the experiment was shorter, with observers completing a total of 25 trials in each of eight 7.5-min blocks, leading to a total of 200 trials. There were 100 negative trials and 100 neutral trials, with each unique image appearing only once during the entire experiment.

Analysis

Eye data preprocessing

In order to map out the pupillary light response function, the raw pupil data were converted into absolute units of millimeters using Hayes & Petrov's formula⁴². For the conversion, pupil size was converted from Eyelink's arbitrary diameter units to area units, and the distance parameter between the eye-tracker and eye was approximately 520 mm. Blinks were interpolated using cubic-spline interpolation⁴³. Time points with, and 1 s prior to, abnormal pupil sizes (< 1 mm or > 9 mm) were treated as signal loss and excluded from analysis. Furthermore, time points where the horizontal or vertical eye positions exceeded 2.5° in Expt. 1 and 3° in Expt. 2A and 2B from the observer's mean x- and y- center of fixation were also excluded from analysis. Standard deviation was computed for the x- and y- eye positions to ensure observers were maintaining fixation. Observers that failed to maintain fixation or had excessive eye movements, with an x- and y- eye position standard deviation greater than 2 degrees and/or having greater than 10% of the data excluded from analysis, were excluded from further analysis. In Experiments 1 and 2A, eight participants were excluded from each experiment. In Experiment 2B, three participants were excluded. In Expt. 1, the remaining 28 subjects had a mean x- and y-center of fixation of − 0.74 and − 1.56 degrees, with an average of 1.39% of data excluded from analysis across participants. In Expt. 2A, the 28 subjects had a mean x- and y-center of fixation of − 0.17 and − 0.11 degrees, with an average of 2.30% of data excluded from analysis across participants. In Expt. 2B, the 30 subjects had a mean x- and y-center of fixation of 0.18 and − 0.19 degrees, with an average of 1.99% of data excluded from analysis across participants.

In all experiments, the mean pupil diameter in millimeters was computed for each trial by averaging the pupil trace within the 2 to 6 s time window after stimulus onset. This time window excludes the initial constriction response, in which the pupil initially constricts in response to changes in in foveal vision^{17,44}. The mean pupil diameter across each luminance and arousal condition was then computed. Due to recalibration of the eye for each run and potential differences in head placement within and between runs, pupil size measurements and conversions may not be exact⁴².

Skin conductance analysis

Skin conductance response (SCR) is a traditional measure of arousal, as it tracks changes in the sympathetic nervous system. Changes in skin conductance have been shown to covary with changes in pupil size with emotional stimuli, supporting the notion that pupillary changes during emotional arousal reflect sympathetic activity⁴⁵. Here, we recorded skin conductance as an additional measure of arousal. The recordings were analyzed offline using AcqKnowledge (version 5.0) to compute phasic SCRs. SCR was calculated from the original electrodermal activity (EDA) time course using the AcqKnowledge function, "Derive phasic EDA from Tonic", using a 0.05 Hz high pass filter, a baseline estimation of 0.25, SCR threshold of 0.03 and a 10% rejection of maximum⁴⁶. The skin conductance data were visually inspected and trials that had movement artifacts were removed. The peak amplitude of the phasic response 2 to 8 s after stimulus onset was then computed per trial.

Subjects who did not show a skin conductance response or had more than 20% of trials with signal loss were excluded from our SCR analysis. In Experiment 1, we conducted SCR analysis on 22 out of 28 subjects. In Experiment 2A, we conducted analyses on 26 out of 28 subjects, and in Experiment 2B, we conducted analyses on 23 out of 30 subjects.

Quantifying the PLF

The pupillary light reflex function (Fig. 1) typically exhibits a nonlinear decreasing sigmoidal profile⁴⁷. In order to quantify the PLF and test whether emotional arousal interacts multiplicatively with luminance similarly to cognitive arousal, the PLFs for each condition (Negative, Neutral) were fit with a hyperbolic ratio function. For our experiment, we modified the Naka-Rushton contrast response function⁴⁸, such that pupil size is assumed to decrease monotonically with luminance. In this decrement Naka-Rushton function (Eq. 1):

$$P(L) = (M) \frac{L^n}{L^n + L_{50}^n} + P_{max},$$

where

$$M = (P_{min} - P_{max}). \quad (1)$$

$P(L)$ is the pupil response as a function of luminance; L is the luminance level; P_{max} is the maximum pupil response; L_{50} is the luminance at half the maximum pupil response; and n is the slope, or the nonlinearity in the gain of response to the input intensity. M represents the negative magnitude of the PLF, which is derived by subtracting P_{max} from P_{min} , or the minimum, baseline pupil response. Variants of this hyperbolic function have been used in prior studies examining pupillary or neural responses as a function of luminance^{18,49}.

This function allowed us to quantify the modulatory effects of arousal on the PLF—significant differences between conditions for any of these parameters would indicate different modulatory signatures of emotional

arousal on the PLF. Our focus centered on four primary modulatory effects: maximum response gain, inflection point gain, baseline and additive shift. Maximum response gain modulation is indicated by a shift in the parameter P_{max} , signifying a multiplicative increase of the maximum pupil size, highlighting that arousal's effects are primarily observed at the lowest luminances. Inflection point gain is marked by a shift in the parameter $L50$, causing a horizontal shift of the PLF, wherein arousal's modulation of pupil size is prominent at mid-luminance levels. A baseline shift is recognized by a shift in the parameter P_{min} , reflecting an increase in pupil size due to arousal at high luminances. Lastly, an additive shift is characterized by arousal modulating both parameters P_{max} and P_{min} , resulting in an overall increase in pupil size across all luminance levels.

The pupillary light reflex curves were fit individually for the Negative and Neutral conditions using MATLAB's *fmincon* function by optimizing the parameter estimates using nonlinear least squares. There were four free parameters: P_{max} , P_{min} , $L50$, and n . $L50$ was constrained between the luminance levels tested, 0.92 to 233.37 cd/m². P_{max} was constrained to the greatest mean pupil size observed across all luminances for each observer for individual-subject fits. P_{max} represents the maximum pupil response in the context of our experiment, as non-saturating PLFs may lead to fits of unrealistically large P_{max} values.

Bootstrapping analyses were used as a measure of reliability of the parameter estimates. We sampled with replacement the data for each observer to obtain a PLF for each condition (Negative, Neutral) on each of 2000 iterations (Expt 1: 18 data points per condition for each luminance level; Expt 2A: 14 data points; Expt 2B: 10 data points). We then fit the resampled PLF with the modified Naka-Rushton function (Eq. 1) in order to obtain the parameter estimates. The parameter estimates of interest, P_{max} , P_{min} , and $L50$ were then subtracted between the Negative and Neutral conditions to obtain a difference score, which served to quantify an individual subject's modulation of the PLF by arousal. The bootstrapping procedure yielded a distribution of difference scores for each of the parameter estimates, in which p-values were computed as a proportion of the difference scores that fell below or above zero. We then doubled the smaller value to obtain a two-sided p-value. At the group level, we computed a group-averaged PLF by averaging the data points across observers for each of the 2000 iterations and fit each of the group-averaged PLFs with the modified decrement Naka-Rushton function. Again, we then generated a distribution of difference scores for each parameter estimate, and computed p-values in an identical manner.

Results

Experiment 1: affective sounds (IADS)

Overall pupil diameter was larger in the Negative condition (mean: 4.35 mm, SEM = 0.11) than in the Neutral condition (mean: 4.27 mm, SEM = 0.11) when collapsed across luminance levels (paired *t*-test $t(27) = 6.21$, 95% CI [0.05 0.10], $p < 0.001$, Cohen's $d = 1.17$; Fig. 3a), indicating a large effect of emotional arousal on pupillary responses, consistent with prior work^{9,17}. Average subjective arousal rating (on a scale from 1 to 5, with 5 as highest arousal) in the Negative condition (mean: 3.33; SEM = 0.04) was also higher than in the Neutral condition (mean: 2.10; SEM = 0.04; paired *t*-test $t(27) = 10.16$, 95% CI [0.97 1.47], $p < 0.001$, Cohen's $d = 1.93$). We found that the subjective arousal ratings provided by our participants for the IADS stimuli were similar to those reported by participants in Bradley & Lang²⁶ and Yang et al.²⁷ ($r = 0.87$, $p < 0.001$; Supplementary S6). However, across luminances, only lower luminances exhibited larger pupil dilation in the Negative condition compared to the Neutral condition (luminances: 1.69, 3.13, 10.74 cd/m²; $p = 0.04$, $p < 0.001$, $p = 0.03$; Cohen's $d = 0.42$, $d = 0.75$, $d = 0.44$; Fig. 3b,c), with little-to-no modulation of pupil size by arousal in the mid- to higher-luminances levels. Since the modulation by emotional arousal across luminances was not uniform, we then quantified the PLF and its interaction with arousal. This approach allowed us to test whether the relationship between luminance and arousal is additive, as has been assumed in previous research, or whether a multiplicative interaction exists, similar to that observed with cognitive arousal.

To do so, we fit each PLF (Negative, Neutral) at the individual and group level with a modified decrement Naka-Rushton function (see Method). The fits well-characterized individual subjects' PLFs (group-averaged mean $R^2 = 0.96$, 95% CI [0.94 0.98]). The primary modulatory effects we tested were maximum response gain, inflection point gain, baseline and additive shift. A maximum response gain modulation is denoted by a shift in the parameter P_{max} , which indicates that arousal leads to a multiplicative increase in maximum pupil size at the lowest luminances. An inflection point gain is denoted by a shift in the parameter $L50$, which causes a horizontal shift of the PLF, such that arousal's modulation of pupil size occurs predominantly at mid-luminance levels. A baseline shift is denoted by a shift in the parameter P_{min} , in which an increase in pupil size by arousal occurs at high luminances. Lastly, an additive shift is characterized by arousal modulating both the parameters P_{max} and P_{min} , leading to an overall increase in pupil size across all luminance levels.

Bootstrapping analysis on the parameter estimates revealed that the group results were best characterized by a maximum response gain modulation, in which arousal modulated the P_{max} parameter (Negative P_{max} bootstrapped mean = 5.63 mm, 95% CI [5.59 5.66]; Neutral P_{max} mean = 5.54 mm, 95% CI [5.51 5.57]; bootstrap test, 95% CI [0.04 0.14], $p < 0.001$; Fig. 3d), affecting pupil size only at lower luminances. There was no difference in P_{min} (Negative P_{min} bootstrapped mean = 2.60 mm, 95% CI [2.56 2.64]; Neutral P_{min} mean = 2.53 mm, 95% CI [2.49 2.58]; bootstrap test, 95% CI = [0.01 0.13]), $p = 0.06$ or $L50$ (Negative $L50$ bootstrapped mean = 24.05 cd/m², 95% CI [23.30 24.85]; Neutral $L50$ mean = 24.82 cd/m², 95% CI [23.74 25.91]; bootstrap test, 95% CI = [-2.16 0.54], $p = 0.36$) between the Negative and Neutral PLF, indicating a lack of a baseline shift or inflection point gain, which would have corresponded to modulation of pupil size at mid- and high luminances. Scatterplots of the parameter estimates across individuals in Fig. 3e corroborate the group-level modulation; the majority of the participants exhibited a larger maximum pupil response (P_{max}) in the Negative condition compared to Neutral, whereas there was considerable heterogeneity across participants in the minimum pupil response (P_{min}) and inflection point ($L50$) parameters between the two arousal conditions. In summary, we found that, despite

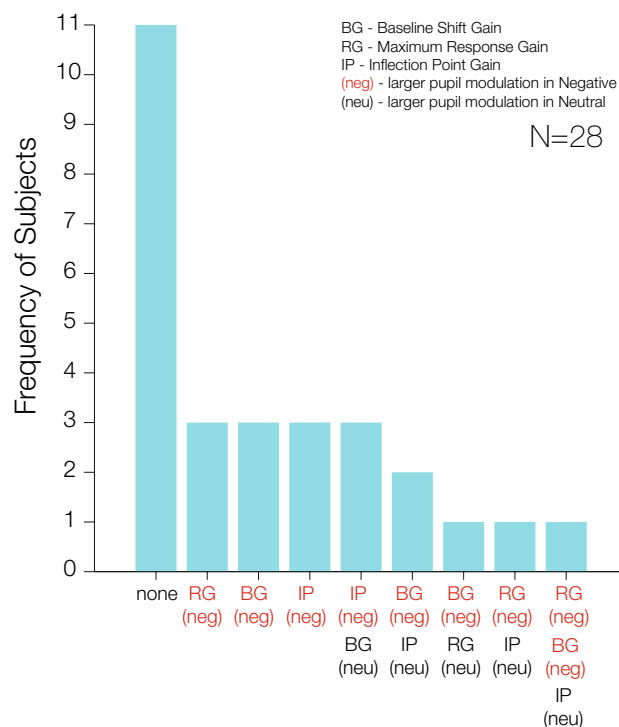
individual differences across observers, modulation of pupil size by emotional arousal mainly occurred at low luminance, with little-to-no modulation at mid- or higher- luminances.

Because we observed substantial variability in parameter estimates across individuals, we then sought to identify the model that best captured the modulatory effect of emotional arousal on the PLF for each observer. To do so, we conducted a bootstrapping analysis to ensure the accuracy and reliability of the parameter estimates within an individual. We resampled with replacement the PLF for each condition (Negative, Neutral) for 2000 iterations, and fit each PLF with the modified decrement Naka-Rushton function to obtain parameter estimates for each iteration. In doing so, we end up with 2000 parameter estimates for each of the conditions (Negative, Neutral), for which we then generated a distribution of difference scores between the parameter estimates. The bootstrapping analysis identified different winning models across participants: the largest group of participants (11 out of the 28 observers) displayed no modulation (Fig. 4). While the majority of participants (17 out of 28) had modulatory effects of arousal on the PLFs, the qualitative nature of the modulation was highly heterogeneous across observers, with individuals showing a maximum response gain modulation, inflection point modulation, an additive shift, or a combination of different modulations. The 11 observers who showed no difference in pupillary response between the Negative and Neutral condition can be considered “non-responders” to the IADS stimuli: the auditory affective sounds did not seem to evoke an arousal response in these individuals. Figure 4 displays exemplar subjects: certain observers had a maximum response gain modulation, some had an inflection point gain modulation, and others exhibited no modulation at all.

Next, we assessed possible factors that could explain inter-individual variation in the magnitude of emotion-driven pupil responses. First we examined the correlation across participants between individual’s negative-minus-neutral difference in pupil size and subjective arousal rating. We found no correlation between pupil size and subjective arousal ratings ($r = 0.04$, $p = 0.83$; see Supplementary S7). However, skin conductance response (SCR) related to pupil size ($r = 0.46$, $p = 0.03$; see Supplementary S8), consistent with prior work linking skin conductance and pupil dilation to the same underlying sympathetic nervous pathway. More specifically, subjects who exhibited no difference in pupillary response between Negative and Neutral sounds also displayed no skin conductance response between the two conditions, thus highlighting heterogeneity in responsiveness to the

Experiment 1: IADS

a. Winning model across subjects



b. Example subjects

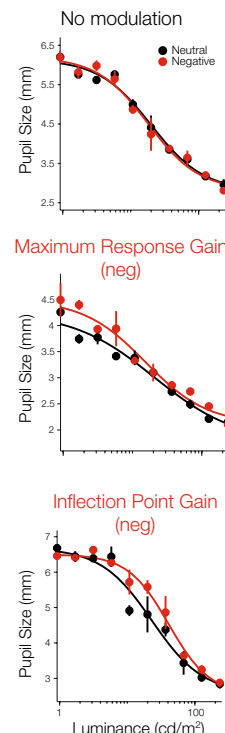


Fig. 4. Heterogeneity in modulation by emotional arousal of the PLF in experiment 1 using IADS. **(a)** Histogram of the winning model that best captures the modulation by emotional arousal on an individual’s PLF. Red text indicates larger pupil modulation in the Negative compared to Neutral condition. Black text indicates either no modulation, or a larger pupil modulation in the neutral compared to negative condition. **(b)** Three example subjects exhibiting different modulation of the PLF by emotional arousal.

IADS stimuli set, suggesting that the sounds did not evoke a detectable physiological arousal response in certain individuals. The absence of an arousal response could be due to either the sounds being ineffective or failure to recognize the sounds for a given individual. However, when examining the group-level data, we did observe an overall higher skin conductance response in the Negative condition (mean: 0.37 μ S, SEM: 0.06) than the Neutral condition (mean: 0.30 μ S, SEM: 0.05; paired *t*-test(21) = 2.59, 95% CI [0.01 0.12], *p* = 0.017, Cohen's *d* = 0.92). Skin conductance exhibited high variability across luminances; nevertheless, there are increases in skin conductance observed at both low and high luminances in the Negative condition compared to Neutral (see Supplementary S9). This finding is inconsistent with the possibility that arousal levels vary across luminance levels, thereby disfavoring the notion that the observed pupillary modulation at lower luminances is primarily attributed to increased arousal modulation at those lower luminance levels.

Experiment 2A: grayscale affective images (IAPS)

Replicating previous work examining pupil size response to IAPS images, we observed that overall pupil size, collapsed across luminance, was larger with a small effect size in the Negative condition (mean: 3.61 mm, SEM: 0.10) than in the Neutral condition (mean: 3.58 mm, SEM: 0.10; paired *t*-test *t*(27) = 2.24, 95% CI [0.002 0.07], *p* = 0.03, Cohen's *d* = 0.42; Fig. 5a). Average subjective arousal ratings were also higher in the Negative condition (mean: 3.53, SEM: 0.1691) than in the Neutral condition (mean: 1.78, SEM: 0.09; paired *t*-test *t*(27) = 12.31, 95% CI [1.43 2.00], *p* < 0.0001, Cohen's *d* = 2.17; Supplementary S7), with our subject pool's subjective arousal to the images reflecting the arousal ratings by the subject pool in Bradley & Lang¹⁰ (*r* = 0.86, *p* < 0.0001; Supplementary S6). When broken down by luminance, we observed a small difference in pupil size between the Neutral and Negative PLFs, with significant pupil differences at two luminances of 68.12 and 126.09 cd/m² (*p* = 0.03, *p* = 0.04 Cohen's *d* = 0.43, *d* = 0.40; Fig. 5b,c).

However, on the group level, there was no detectable modulation of the PLF by emotional arousal: there was no difference between the Negative and Neutral conditions for the *P_{min}* (Negative *P_{min}* bootstrapped mean = 2.34 mm, 95% CI [2.29, 2.39]; Neutral *P_{min}* mean = 2.26 mm, 95% CI [2.21, 2.31]; bootstrap test, 95% CI = [0.00, 0.15], *p* = 0.08), *L50* (Negative *L50* bootstrapped mean = 25.82 cd/m², 95% CI [23.98, 27.83]; Neutral *L50* mean = 27.06 cd/m², 95% CI [24.92, 29.28]; bootstrap test, 95% CI = [-4.04, 1.51], *p* = 0.45), or *P_{max}* (Negative *P_{max}* bootstrapped mean = 4.52 mm, 95% CI [4.49, 4.56]; Neutral *P_{max}* mean = 4.50 mm, 95% CI [4.47, 4.53]; bootstrap test, 95% CI [-0.02, 0.07], *p* = 0.37) parameters (Fig. 5d). Similar to Experiment 1, the fits well-characterized individual subjects' PLFs (group-averaged mean *R*² = 0.979, 95% CI [0.973 0.985]). While emotional arousal evoked by grayscale IAPS images do modulate pupil size, the observed effects are relatively small and cannot be solely attributed to any specific PLF parameter or modulatory pattern of effect. Scatterplots of the parameter estimates in Fig. 5e further reveal the lack of a modulatory effect driven by grayscale IAPS images, as the observers displayed considerable heterogeneity across all parameter estimates tested.

Due to variability in parameter estimates across observers, we then fit the modified Naka-Rushton function to individual observer's PLFs, assessing which model (e.g. maximum response gain) best characterized the modulation by IAPS-driven emotional arousal within individuals. As in Experiment 1, we conducted a bootstrapping analysis on individual subject's PLFs and the parameter estimates, in order to test for parameter estimate reliability. The test also captured the direction of the modulation – whether pupil size was greater in the Negative or Neutral condition at low, mid and high luminances. We found that most subjects (19 out of 28) exhibited no modulation between the Negative and Neutral IAPS grayscale images, which is reflected in our group results. Unlike Experiment 1, three of the subjects displayed the opposite pattern of effect to the group pattern – that is, having an overall larger pupillary response to the Neutral images compared to Negative images (Fig. 6).

In order to disentangle these patterns of effects, we assessed different possible factors that could play a role in influencing an individual's distinct modulatory effects. We found no correlation between pupil size and subjective arousal ratings (*r* = 0.14, *p* = 0.48; Supplementary S7). Interestingly, however, we found that the difference in skin conductance response between the Neutral and Negative conditions correlated to the difference in pupil size (*r* = 0.58, *p* = 0.001; Supplementary S8). When examining the group-level data, we did not observe a significant difference in skin conductance response between the Negative condition (mean: 0.21 μ S, SEM: 0.05) and the Neutral condition (mean: 0.20 μ S, SEM: 0.05; paired *t*-test *t*(25) = -1.09, 95% CI [-0.01 0.04], *p* = 0.29, Cohen's *d* = 0.21), due to the variability in skin conductance response across subjects, likely stemming from individual differences in polarity of responses to either the Negative or Neutral images. Additionally, there was no discernible difference in skin conductance response between the Negative and Neutral conditions across different luminance levels (Supplementary S9). This finding further corroborates the presence of high variability in reactivity to the IAPS images, evident in both pupil size and skin conductance. While some individuals exhibited an increased arousal response to the Negative images, others showed no discernible arousal response to the Negative images, and a smaller group of individuals demonstrated an increased arousal response to the Neutral images. We conducted Experiment 2B to address several factors that might have accounted for this heterogeneity.

Experiment 2B: color affective images (IAPS)

In Experiment 2A, we used grayscale IAPS images to manipulate emotional arousal while controlling for the effects of luminance, color, and spatial frequency. However, our results showed that the stimuli had little-to-no effect in modulating pupil size across luminances. To further investigate the potential factors that might have led to our observed effects, we conducted Experiment 2B. We hypothesized that presenting the images in grayscale might have reduced the impact of arousal-provoking content (such as blood and body mutilation). The small size and reduced clarity of the images might also have reduced the strength or consistency of participants' affective response. Our analysis showed that on average, participants found 6.17% (range: 0 to 19.29%) of the images ambiguous or hard to discern in Experiment 2A (see Supplementary S5 for examples). To rule out the possibility

Experiment 2A: IAPS Affective Pictures (grayscale)

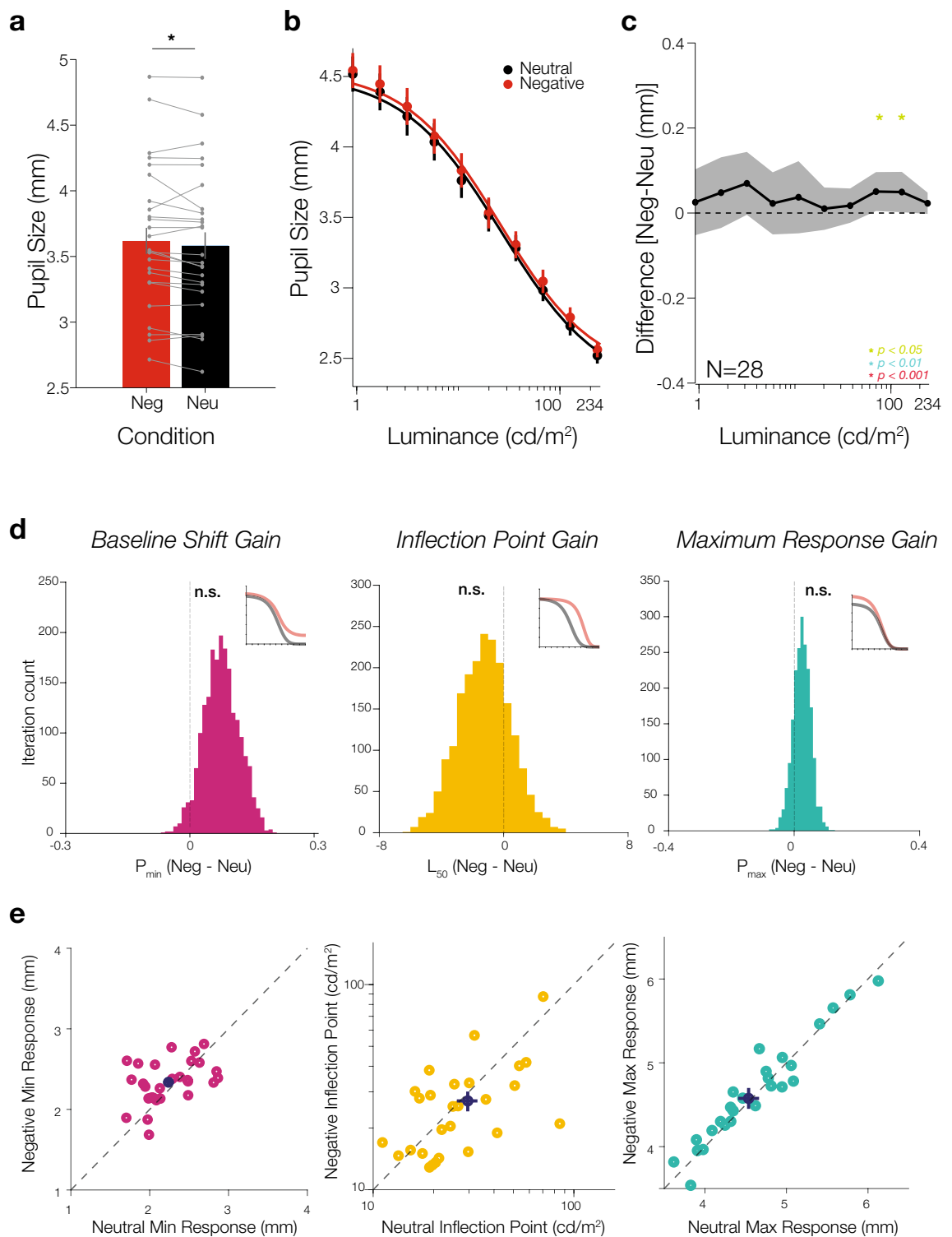


Fig. 5. Group results for grayscale IAPS. **(a)** Mean pupil size in the Negative and Neutral condition using IAPS grayscale images, collapsing across luminance. Each point on the graph represents an individual subject. Overall, larger pupil size was observed in the Negative condition than the Neutral condition. **(b)** Group-level mean pupillary light reflex functions for the Negative and Neutral condition, fit with the modified decrement Naka-Rushton function. Error bars represent SEM. **(c)** Tests at each luminance level for a significant difference between pupil size in the Negative and Neutral conditions. The shaded gray area represents the 95% confidence interval. **(d)** Bootstrapped distributions (2000 iterations) of the difference between Negative and Neutral parameter estimates for P_{min} , L_{50} and P_{max} . The dotted line indicates no difference in the Negative and Neutral parameter estimates. **(e)** Scatterplot of individual observer parameter estimates comparing Neutral versus Negative for P_{min} (minimum pupil response), L_{50} (inflection point), and P_{max} (maximum pupil response). The purple dot indicates the group-averaged parameter estimates, with the error bars indicating the SEM. The dotted line indicates the unity line of no difference in parameter estimates between the Negative and Neutral condition.

Experiment 2A: IAPS (grayscale)

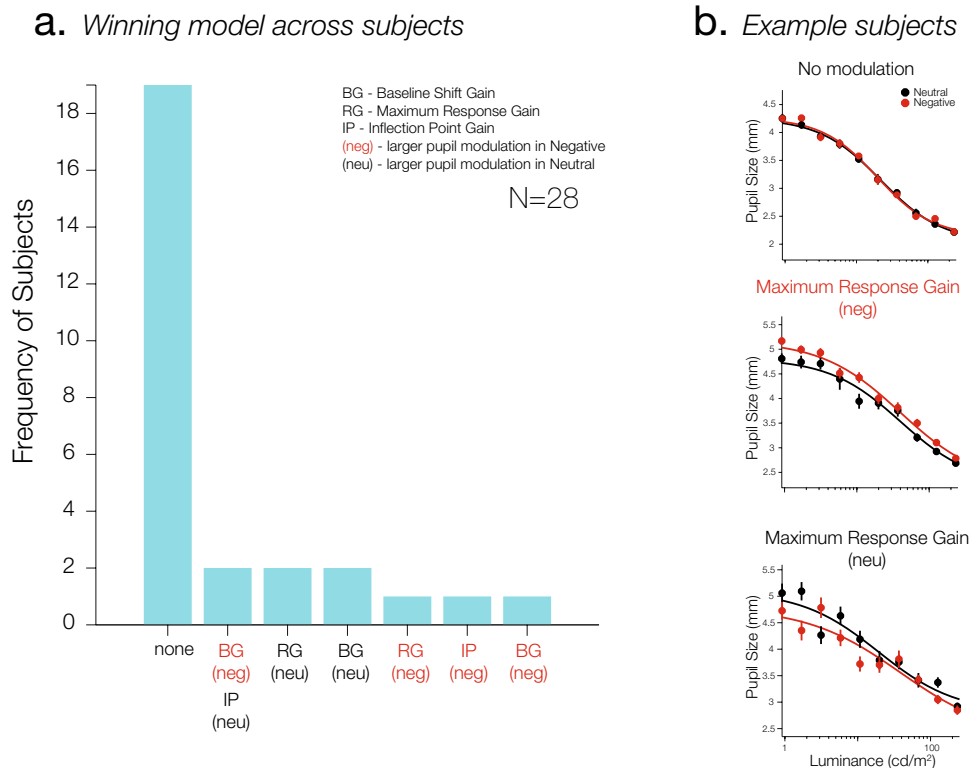


Fig. 6. Heterogeneity in modulation by arousal on the PLF in Expt 2A using IAPS in grayscale. **(a)** Histogram of the winning model that best captures the modulation by emotional arousal on an individual's PLF. Red text indicates larger pupil modulation in the Negative compared to Neutral condition. Black text indicates either no modulation, or a larger pupil modulation in the Neutral compared to Negative condition. **(b)** Three example subjects exhibiting different modulation by arousal on the PLF. Overall, majority of the subjects show no emotional arousal modulation on pupil size.

that our results were driven by confusing images, we re-analyzed the individual subjects' and group PLFs and excluded the trials marked as ambiguous. We found little-to-no discernable changes in the overall results, except for some minor pupil differences.

We then conducted a follow-up study in which we made several changes to our arousal manipulation. First, we kept the IAPS images in color but controlled for luminance content across the images. Second, we increased the size of the stimuli presented on the screen. Third, we shortened the experiment by using a subset of the images from Experiment 2A. Similar to Experiment 2A and prior studies, we found that pupil responses to the colored IAPS images were larger in the Negative condition (mean: 3.71 mm, SEM: 0.12) compared to the Neutral condition (mean: 3.65 mm, SEM: 0.12) with a medium effect when collapsed across luminance (paired t -test $t(29) = 3.43$, 95% CI [0.02 0.09], $p = 0.002$, Cohen's $d = 0.63$; Fig. 7). This was also reflected in the observers' subjective arousal rating, with higher rating in the Negative condition (mean: 3.75 out of 5, SEM: 0.08) than in the Neutral condition (mean: 1.65 out of 5, SEM: 0.11; paired t -test $t(29) = 19.85$, 95% CI [1.88 2.31], $p < 0.0001$). Furthermore, the subjective arousal ratings of our subject pool corresponded with the arousal ratings of the original Lang et al. (2008) dataset ($r = 0.63$, $p < 0.0001$; Supplementary S6). In contrast to Experiment 2A, we observed differences in pupil size between the Neutral and Negative conditions at four individual luminances (luminances: 1.69, 5.80, 19.88, 68.12 cd/m²; $p = 0.03$, $p = 0.01$, $p = 0.04$, $p = 0.001$; Cohen's $d = 0.42$, $d = 0.48$, $d = 0.39$, $d = 0.65$ Fig. 7b,c). To test the relationship between luminance and arousal in influencing pupillary responses, we quantified the modulation of the PLF by arousal by conducting bootstrapping and model-fitting analysis on the group PLF, with the fits capturing the PLFs well (group-averaged mean = 0.96, 95% CI [0.949 0.977]), to test for the three main modulations of interest: maximum response gain, inflection point gain, and baseline shift.

Bootstrapping analysis revealed that the modulation of the PLF by arousal was best captured by a maximum response gain modulation, indicating that the greatest effects of arousal on pupil size occurred at low luminances, similar to what was observed in Experiment 1 using auditory IADS stimuli. There was a difference between the Negative and Neutral conditions for the P_{max} parameter (Negative P_{max} bootstrapped mean = 4.49 mm, 95% CI [4.47 4.51]; Neutral P_{max} mean = 4.42 mm, 95% CI [4.40 4.44]; bootstrap test, 95% CI [0.04 0.10], $p < 0.0001$) but no difference for the L_{50} parameter (Negative L_{50} bootstrapped mean = 44.48 cd/m², 95% CI [40.39 49.25]; Neutral L_{50} mean = 40.40 cd/m², 95% CI [36.45 44.82]; bootstrap test, 95% CI [-2.36 10.30], $p = 0.27$) or the P_{min}

Experiment 2B: IAPS Affective Pictures (color)

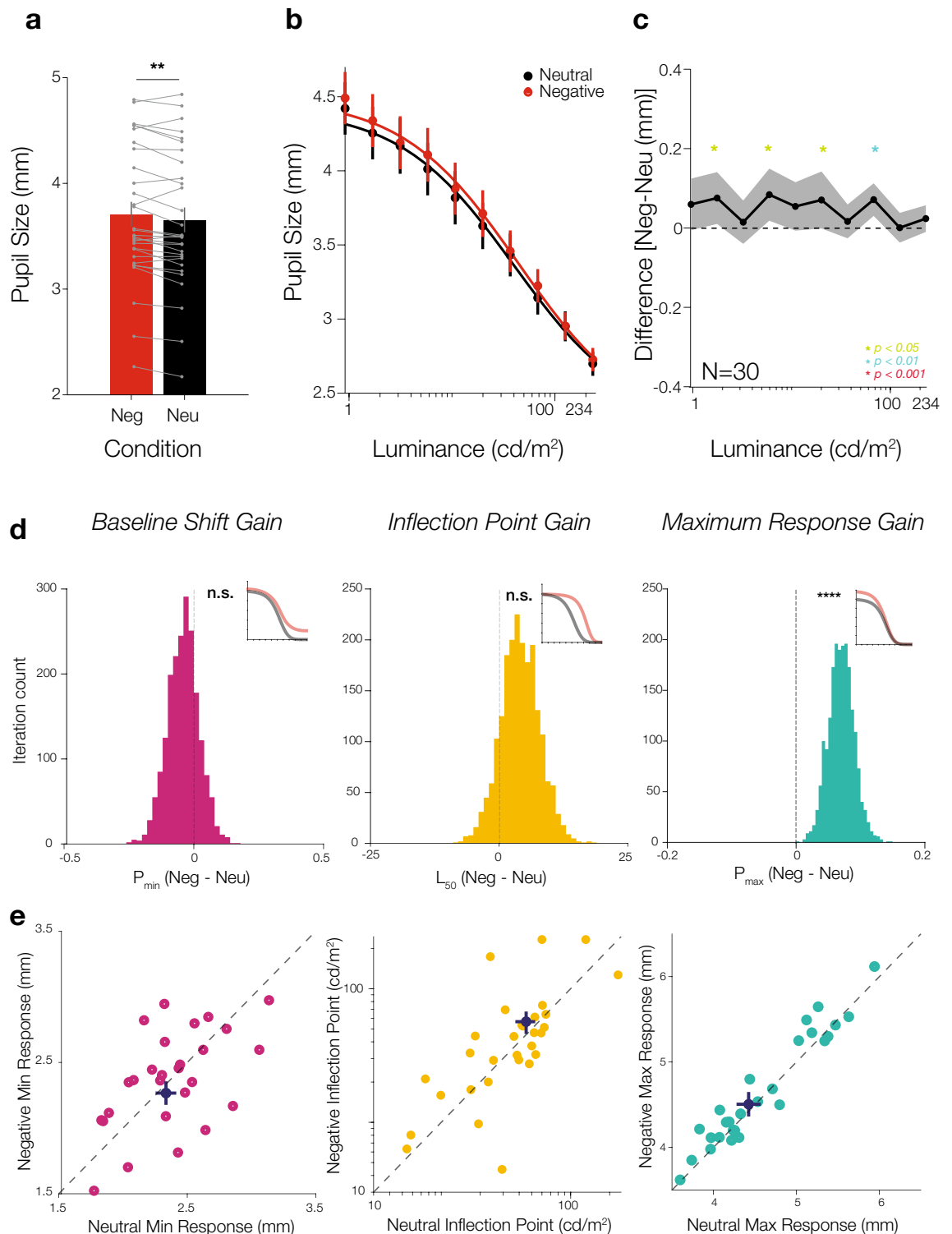


Fig. 7. Group results for IAPS colored. **(a)** Mean pupil size in the negative and neutral condition, collapsing across luminance. Each point on the graph represents an individual subject. Overall, larger pupil size was observed in the Negative condition than the Neutral condition. **(b)** Group-level mean pupillary light reflex functions for the Negative and Neutral condition, fit with the modified decrement Naka-Rushton function. Error bars represent SEM. **(c)** Tests at each luminance level for a significant difference between pupil size in the Negative and Neutral conditions. The shaded gray area represents the 95% confidence interval. **(d)** Bootstrapped distributions (2000 iterations) of the difference between Negative and Neutral parameter estimates for P_{\min} , L_{50} and P_{\max} . The dotted line indicates no difference in the Negative and Neutral parameter estimates. **(e)** Scatterplot of individual observer parameter estimates comparing Neutral versus Negative for P_{\min} (minimum pupil response), L_{50} (inflection point), and P_{\max} (maximum pupil response). The purple dot indicates the group-averaged parameter estimates, with the error bars indicating the SEM. The dotted line indicates the unity line of no difference in parameter estimates between the Negative and Neutral condition.

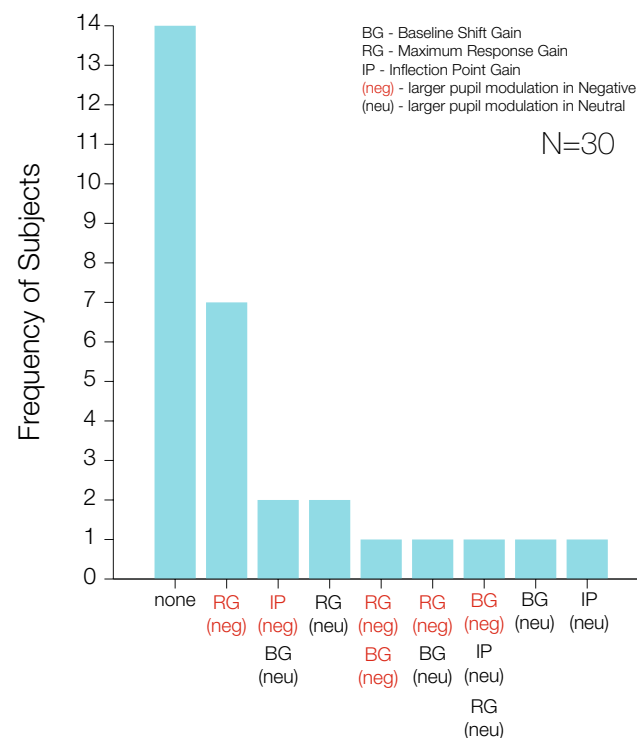
parameter (Negative P_{min} bootstrapped mean = 2.26 mm, 95% CI [2.18 2.32]; Neutral P_{min} mean = 2.30 mm, 95% CI [2.23 2.36]; bootstrap test, 95% CI [-0.14 0.06], $p = 0.48$; Fig. 7d). Across observers, there was a consistent increase in P_{max} , or maximum pupil size at the low luminances, in the Negative condition compared to Neutral. Meanwhile, there was considerable variability across individuals in the $L50$ and P_{min} parameters, indicating little-to-no systematic pupil modulation at mid- and high-luminances (Fig. 7e). The similarity in modulation between Experiment 2B, using IAPS colored images, and Experiment 1, using IADS sounds, demonstrates that when emotional arousal is effectively manipulated, it evokes similar modulatory patterns on the PLF, regardless of the sensory modality of the stimuli used.

We then assessed what form of modulation by emotional arousal best captured individual subjects' PLFs in the Negative and Neutral conditions. To do this, we conducted a bootstrapping analysis on each subject's PLFs and parameter estimates. Our findings revealed that 14 out of 30 subjects showed no significant parameter changes between the Negative and Neutral colored IAPS images, while the next largest group of subjects (7) showed a maximum response gain modulation with greater pupil size in the Negative condition at low luminances, consistent with our group results. The remaining subjects displayed a variety of different modulations (see Fig. 8). Interestingly, similar to Experiment 2A using grayscale images, but dissimilar to Experiment 1 using sounds, a subset (5) of individuals displayed greater overall pupil response in the Neutral condition compared to the Negative condition. Due to the heterogeneity of parameter estimates, we further examined whether subjective arousal ratings could explain individual differences in modulatory effects, but found no correlation between pupil size and subject arousal ratings ($r = -0.21$, $p = 0.26$; Supplementary S7).

Skin conductance responses did not differ significantly between the Neutral and Negative conditions at the group level (paired t -test $t(22) = -0.09$, 95% CI [-0.03 0.03], $p = 0.93$) due to the heterogeneity in skin conductance response across subjects. However, the heterogeneity was driven in part by a positive correlation between individual differences in Negative-minus-Neutral pupillary response and skin conductance response ($r = 0.52$, $p = 0.01$; Supplementary S8). This suggests that subjects who showed a larger pupillary response to the Neutral condition compared to the Negative condition also displayed higher skin conductance response to the Neutral

Experiment 2B: IAPS (color)

a. Winning model across subjects



b. Example subjects

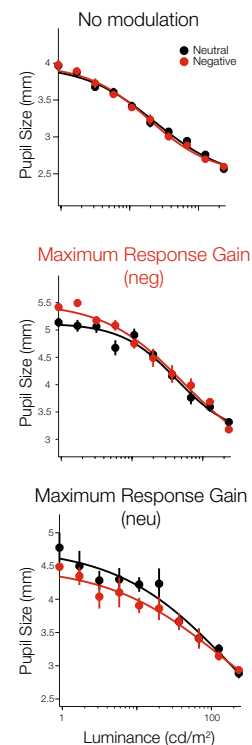


Fig. 8. Heterogeneity in modulation by arousal on the PLF in Expt 2B using IAPS in color. **(a)** Histogram of the winning model that best captures the modulation by arousal on an individual's PLF. Red text indicates larger pupil modulation in the Negative compared to Neutral condition. Black text indicates either no modulation, or a larger pupil modulation in the Neutral compared to Negative condition. **(b)** Three example subjects exhibiting different modulation by arousal on the PLF. Overall, almost half of our subjects show no modulation by emotional arousal on pupil size, the next largest group of subjects show a response gain modulation, or modulation at low luminances, consistent with our group results.

condition, and vice versa. In addition, we did not observe any significant differences in skin conductance response between the Negative and Neutral conditions across different luminance levels (Supplementary S9), disfavoring the possibility that the observed pupil modulation at lower luminances was attributed to an increased arousal modulation at those levels.

In sum, in Experiment 2A, using grayscale IAPS images and controlling for color, luminance, and spatial frequency, our manipulation was ineffective in evoking an emotional arousal response compared to using colored IAPS images in Experiment 2B, where we controlled only for luminance. Our findings suggest that when the arousal manipulation is effective, the interaction of emotional arousal with luminance is similar across auditory and visual sensory modalities. We found that emotional arousal multiplicatively interacted with luminance in influencing pupil responses, leading to greater effects at lower luminances.

Discussion

We investigated the interaction between emotional arousal and luminance in modulating pupillary responses. To do so, we mapped out the pupillary light function using a wide range of luminances under high and low emotional arousal states, which were manipulated using negative (high arousal) and neutral valence (low arousal) stimuli of different sensory modalities sourced from the International Affective Picture System (IAPS), International Affective Digitized Sounds (IADS) and expanded version of the International Affective Digitized Sounds (IADS-E). Our findings revealed that emotional arousal, manipulated using both auditory and visual sensory modalities, interacted with luminance in modulating pupil size differently from cognitive arousal: the effects occurred primarily at much lower luminances ($< 20 \text{ cd/m}^2$). Moreover, we observed qualitatively distinct individual differences in the modulation of the pupillary light reflex by emotional arousal, with some participants showing the greatest modulation at low-luminances, mid-luminances, different combinations of luminances, or no modulation across all luminances.

Our results suggest that differences in arousal manipulation play a critical role in influencing the magnitude and modulation of pupil responses. Notably, emotional arousal-based pupil modulation using auditory IADS stimuli showed less variability, with fewer non-responders, compared to using visual IAPS stimuli. Furthermore, when using IAPS stimuli, presenting the images in luminance-matched grayscale, as done in previous studies^{24,28,29}, evoked a smaller and more variable modulatory effect compared to presenting the image in luminance-matched color. In this study, we also matched for spatial frequency in addition to luminance under the grayscale condition, which may have further contributed to the observed variability. In comparison, cognitive arousal induced by auditory arithmetic problems of Easy (low arousal) and Hard difficulty (high arousal) effectively evoked larger pupil size in the high arousal condition for all participants¹⁸. On the group level, pupil size was larger in the high cognitive arousal condition across *all* luminances but with the largest effects occurring at low and mid-luminances. Meanwhile for *emotional arousal*, at the group level, pupil size was only larger at lower luminances under the high arousal condition. These findings indicate that arousal-based pupillary responses are highly sensitive to different task and stimuli manipulations.

Moreover, our findings align with previous studies that explored the interaction between arousal and luminance on pupillary responses, albeit only testing 2 or 3 luminance levels^{15–19}. The majority of these studies report larger arousal modulation on pupillary responses under low luminance conditions compared to high luminance conditions^{15–17}. However, the *specific luminance levels* and *arousal manipulation* tested were highly variable between these studies. For instance, Chering et al.¹⁷ observed more pronounced pupil effects by emotional arousal (using emotional sounds) in their low luminance condition at 2 cd/m^2 , compared to the high luminance condition at 16 cd/m^2 . Our study aligns with these findings, revealing greater and significant pupil effects to emotional sounds at luminances at and below 10.74 cd/m^2 , reinforcing the observations made by Chering et al.

Examining cognitive arousal's interaction with luminance, Peysakhovich et al.¹⁵ measured pupillary responses across black, gray, and white backgrounds, with luminance levels of 3.4, 24, and 54.8 cd/m^2 , respectively, while observers performed a short-term memory task. In a similar vein, Pflieger et al.¹⁶ assessed pupillary responses during a cognitive load recall task under three light conditions, using 1, 2, or 3 lamps, as well as three stimulus brightness conditions, low, medium, and high brightness, where the exact luminance measurements were not reported. In all, whether induced by cognitive arousal or emotional arousal, the effects of arousal on pupil size were more pronounced at the delineated low luminance condition compared to the high luminance condition. The pattern of effects by arousal at the luminances employed across the aforementioned studies align with our findings examining emotional arousal's effect, and Pan et al.'s¹⁸ findings examining cognitive arousal's effect, across a range of luminances^{15–17}.

Other studies have indicated either an absence of arousal-based pupillary differences across the selected luminances, or a larger modulation at the higher luminances tested^{19–22}. For example, Peysakhovich et al.²⁰ investigated the effects of tonic and phasic pupil responses to luminance and cognitive arousal, utilizing an n-back task and a math task. Under the two luminance conditions tested (low: $\sim 11 \text{ cd/m}^2$; high: $\sim 28 \text{ cd/m}^2$), they found that luminance interacted with arousal modulation for tonic pupillary responses, but not phasic pupil responses. The lack of effect may be due to the luminance levels chosen: Pan et al.¹⁸ found the most consistent and significant modulation by cognitive arousal on pupil responses to be between the luminances of 5 and 37 cd/m^2 across observers, and Peysakhovich and colleague's selected luminances of 11 and 28 cd/m^2 falls within this optimal range. Conversely, Steinhauer et al.¹⁹ reported greater pupillary modulation by cognitive load in their high-luminance light condition at 0.59 cd/m^2 compared to their low-luminance dark condition measured at around 0.03 cd/m^2 . While Pan et al.¹⁸ and the current study did not explicitly measure responses at those luminance levels (0.03 and 0.59 cd/m^2), the high luminance condition employed by Steinhauer et al.¹⁹ is much lower than the low luminance levels used in the other studies discussed. Future research could explore fine-scale

differences in arousal-based pupillary responses at even wider ranges of luminances and using other forms of arousal (e.g. reward-driven arousal).

In the present study, we manipulated emotional arousal by comparing *negative* versus *neutral* valence stimuli to create, respectively, our *high* and *low* arousal conditions. Our goal was to maximize the emotional arousal effect by using Negative versus Neutral stimuli. Instead of determining our low and high arousal conditions based on standardized arousal ratings provided by the databases, we chose to group stimuli by valence in order to prevent potential interactions of valence within our high arousal condition that could influence pupillary responses, since stimuli rated at the higher end of the standardized arousal scale can consist of both negative and positive valence stimuli. Previous research suggests that positive and negative stimuli, both rated higher in arousal than neutral, tend to elicit similar responses, with only minor gender-based differences in response to positive stimuli^{9,24,28}. However, we cannot rule out the possibility of a different modulatory pattern due to slight differences in response to positive and negative stimuli and potential interactions.

An important aspect that is often overlooked in group-averaged data is the heterogeneity in the effectiveness of the affective stimuli across participants. While the arousal ratings of our subject pool are highly consistent with the ratings from the original IADS and IAPS dataset, subjective arousal rating, across the group, was not a strong indicator of an individual's pupil and skin conductance response, both of which are commonly used measures of arousal. Specifically, we observed that a sizable group of individuals did not display a physiological arousal response, measured by pupil and skin conductance response, to the emotional stimuli across all luminances. In the case of IAPS visual stimuli, a few subjects even displayed the opposite pattern of effects, showing higher arousal responses to our Neutral stimuli compared to Negative stimuli. Individual differences in response to emotional stimuli, such as IAPS and IADS stimuli, have been reported in prior studies where differences in behavior, reactivity or brain activation to specific sets of emotional stimuli have been linked to differences that include but is not limited to individual traits such as disgust tolerance and social loneliness, as well as differences in hedonic capacity and depressed mood^{50–52}. In our study, we used a wide range of emotional stimuli from different categories (e.g. animals, objects, humans, etc.) within both Negative and Neutral valence. Due to the number and range of emotional images and sounds we employed (i.e. 90–140 unique stimuli for each arousal condition), there could be various individual difference factors and traits that contribute to the responsiveness of individuals in our subject pool, ranging from highly responsive to not responsive to the emotional stimuli set.

Despite the presence of individual differences and manipulation-based effects, our findings suggest that overall emotional arousal modulates pupil responses more strongly at lower luminances (< 20 cd/m²). We predict that this pattern of modulatory effect would also apply to other stimuli or manipulations of emotional arousal, as long as the manipulation is effective. However, due to the variability of results depending on the type of arousal manipulation, we do not discount the possibility of different modulatory effects by arousal when using various emotional arousal manipulations and stimuli. This is different from Pan et al.¹⁸, which found that cognitive arousal had the largest effects on pupil size at low-to-mid luminances (< 37 cd/m²), but with mid-luminances (~ 5 to 37 cd/m²) consistently modulated across all participants despite individual variability. There are similarities in the modulatory effects of emotional and cognitive arousal on the pupil light reflex, wherein the modulatory effects are larger at lower luminances and smallest at high luminances; however, mid-luminances are only modulated by cognitive arousal. Future studies can investigate whether these similarities and differences are due to overlapping and distinct underlying mechanisms and pathways associated with different forms of arousal and their interactions with the pupillary system.

In sum, while different forms of arousal exhibit different patterns of effects on pupil size across luminances, our findings, consistent with previous studies, suggest that high luminances are least ideal for pupillometry studies that aim to measure arousal. This may be due to a physiological constraint of constricted pupils at high luminances, in which the luminance-constricted pupil may restrict or suppress arousal-driven dilation of the pupils. However, more work is needed to understand the mechanisms and biological basis underlying pupillary responses at higher luminances under different arousal states. These results have implications for real-world eye-tracking studies that aim to measure pupil responses in environments with typically higher luminances (e.g. 1000 cd/m² on a cloudy day), which may make it challenging to detect significant arousal-based pupillary differences. Further complicating matters is the finding that there are qualitatively distinct individual differences in the modulatory effect of emotional arousal on the pupillary light reflex. Given the different effects of cognitive and emotional arousal on pupil modulation, it is possible that other forms of arousal (e.g. pain, surprise) may yield yet other modulatory patterns – a worthwhile question to pursue in future work.

Data availability

The datasets generated during and/or analysed during the current study are available in the OSF repository, <https://osf.io/xnjbe/>.

Received: 26 January 2024; Accepted: 22 August 2024

Published online: 19 September 2024

References

1. Kahneman, D., Beatty, J. & Pollack, I. Perceptual deficit during a mental task. *Science* **157**(3785), 218–219 (1967).
2. Cohen Hoffing, R. A. et al. Dissociable mappings of tonic and phasic pupillary features onto cognitive processes involved in mental arithmetic. *PLoS One* **15**(3), e0230517 (2020).
3. Reimer, J. et al. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat. Commun.* **7**, 13289 (2016).
4. McGinley, M. J. et al. Waking state: Rapid variations modulate neural and behavioral responses. *Neuron* **87**(6), 1143–1161 (2015).

5. Vinck, M. *et al.* Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron* **86**(3), 740–754 (2015).
6. Kahneman, D. & Beatty, J. Pupil diameter and load on memory. *Science* **154**(3756), 1583–1585 (1966).
7. Binda, P. & Murray, S. O. Spatial attention increases the pupillary response to light changes. *J. Vis.* **15**(2), 1 (2015).
8. Unsworth, N. & Robison, M. K. Pupillary correlates of lapses of sustained attention. *Cogn. Affect. Behav. Neurosci.* **16**(4), 601–615 (2016).
9. Partala, T. & Surakka, V. Pupil size variation as an indication of affective processing. *Int. J. Hum. Comput. Stud.* **59**, 185–198 (2003).
10. Lang, P.J., Bradley, M.M. and Cuthbert, B.N. *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-8 (University of Florida, 2008).
11. Zeelenberg, M. Anticipated regret, expected feedback and behavioral decision making. *J. Behav. Decis. Mak.* **12**(2), 93–106 (1999).
12. Wichary, S., Mata, R. & Rieskamp, J. Probabilistic inferences under emotional stress: How arousal affects decision processes. *J. Behav. Decis. Mak.* **29**(5), 525–538 (2016).
13. Niikuni, K. *et al.* Pupil dilation reflects emotional arousal via poetic language. *Percept. Mot. Skills* **129**(6), 1691–1708 (2022).
14. Nassar, M. R. *et al.* Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat. Neurosci.* **15**(7), 1040–1046 (2012).
15. Peysakhovich, V. *et al.* Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort. *Int. J. Psychophysiol.* **97**(1), 30–37 (2015).
16. Pflieger, B. *et al.* A model relating pupil diameter to mental workload and lighting conditions. In *Proc. of the 2016 CHI Conference on Human Factors in Computing Systems*, p. 5776–5788 (Association for Computing Machinery, 2016).
17. Cherno, Y. G. *et al.* Background luminance effects on pupil size associated with emotion and saccade preparation. *Sci. Rep.* **10**(1), 15718 (2020).
18. Pan, J. *et al.* Arousal-based pupil modulation is dictated by luminance. *Sci. Rep.* **12**(1), 1390 (2022).
19. Steinhauer, S. R. *et al.* Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *Int. J. Psychophysiol.* **52**(1), 77–86 (2004).
20. Peysakhovich, V., Vachon, F. & Dehais, F. The impact of luminance on tonic and phasic pupillary responses to sustained cognitive load. *Int. J. Psychophysiol.* **112**, 40–45 (2017).
21. Bradshaw, J. L. Brightness of the dominant field, and pupillary reflexes in retinal rivalry. *Br. J. Psychol.* **60**(3), 351–356 (1969).
22. Reilly, J. *et al.* The human task-evoked pupillary response function is linear: Implications for baseline response scaling in pupillometry. *Behav. Res. Methods* **51**(2), 865–878 (2019).
23. Steinhauer, S. R., Condray, R. & Kasperek, A. Cognitive modulation of midbrain function: task-induced reduction of the pupillary light reflex. *Int. J. Psychophysiol.* **39**(1), 21–30 (2000).
24. Bradley, M. M. *et al.* The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* **45**(4), 602–607 (2008).
25. McDougal, D. H. & Gamlin, P. D. Autonomic control of the eye. *Compr. Physiol.* **5**(1), 439–473 (2015).
26. Bradley, M.M. and Lang, P.J. *The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual*. Technical report B-3 (University of Florida, 2007).
27. Yang, W. *et al.* Affective auditory stimulus database: An expanded version of the international affective digitized sounds (IADS-E). *Behav. Res. Methods* **50**(4), 1415–1429 (2018).
28. Bradley, M. M., Sapigao, R. G. & Lang, P. J. Sympathetic ANS modulation of pupil diameter in emotional scene perception: Effects of hedonic content, brightness, and contrast. *Psychophysiology* **54**(10), 1419–1435 (2017).
29. Nakakoga, S. *et al.* Pupillary response reflects attentional modulation to sound after emotional arousal. *Sci. Rep.* **11**(1), 17264 (2021).
30. Nakakoga, S. *et al.* Asymmetrical characteristics of emotional responses to pictures and sounds: Evidence from pupillometry. *PLoS One* **15**(4), e0230775 (2020).
31. Vasquez-Rosati, A. *et al.* Pupillary response to negative emotional stimuli is differentially affected in meditation practitioners. *Front. Hum. Neurosci.* **11**, 209 (2017).
32. Snowden, R. J. *et al.* The pupil's response to affective pictures: Role of image duration, habituation, and viewing mode. *Psychophysiology* **53**(8), 1217–1223 (2016).
33. Ferrari, V. *et al.* Novelty and emotion: Pupillary and cortical responses during viewing of natural scenes. *Biol. Psychol.* **113**, 75–82 (2016).
34. van Steenbergen, H., Band, G. P. & Hommel, B. Threat but not arousal narrows attention: Evidence from pupil dilation and saccade control. *Front. Psychol.* **2**, 281 (2011).
35. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**(4), 433–436 (1997).
36. Pelli, D. G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat. Vis.* **10**(4), 437–442 (1997).
37. Kleiner, M., Brainard, D. H. & Pelli, D. What's new in psychtoolbox-3?. *Perception* **36**, 1–16 (2007).
38. Bradley, M. M. & Lang, P. J. Measuring emotion: The self-assessment Manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **25**(1), 49–59 (1994).
39. Lang, P. J. *Behavioral Treatment and Bio-Behavioral Assessment: Computer Applications* 119–137 (Ablex: Norwood, 1980).
40. Willenbockel, V. *et al.* Controlling low-level image properties: The SHINE toolbox. *Behav. Res. Methods* **42**(3), 671–684 (2010).
41. Dal Ben, R. SHINE_color: Controlling low-level properties of colorful images. *MethodsX* **11**, 102377 (2021).
42. Hayes, T. R. & Petrov, A. A. Mapping and correcting the influence of gaze position on pupil size measurements. *Behav. Res. Methods* **48**(2), 510–527 (2016).
43. Mathôt, S., *A simple way to reconstruct pupil size during eye blinks* (2013).
44. Barbur, J. *Learning from the Pupil: Studies of Basic Mechanisms and Clinical Applications* Vol. 1, 641–656 (The MIT Press, 2004).
45. Wang, C. A. *et al.* Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. *Front. Neurol.* **9**, 1029 (2018).
46. Braithwaite, J. J. *et al.* Guide for Analysing Electrodermal Activity & Skin Conductance Responses for Psychological Experiments. *CTIT tech. rep. seri.* (2013).
47. Watson, A. B. & Yellott, J. I. A unified formula for light-adapted pupil size. *J. Vis.* **12**(10), 12 (2012).
48. Naka, K. I. & Rushton, W. A. S-potentials from luminosity units in the retina of fish (Cyprinidae). *J. Physiol.* **185**(3), 587–599 (1966).
49. Clarke, R. J., Zhang, H. & Gamlin, P. D. R. Characteristics of the pupillary light reflex in the alert rhesus monkey. *J. Neurophysiol.* **89**(6), 3179–3189 (2003).
50. Mataix-Cols, D. *et al.* Individual differences in disgust sensitivity modulate neural responses to aversive/disgusting stimuli. *Eur. J. Neurosci.* **27**(11), 3050–3058 (2008).
51. Cacioppo, J. T. *et al.* In the eye of the beholder: Individual differences in perceived social isolation predict regional brain activation to social stimuli. *J. Cogn. Neurosci.* **21**(1), 83–92 (2009).
52. Saxena, A. *et al.* Individual differences in hedonic capacity, depressed mood, and affective states predict emotional reactivity. *Motiv. Emot.* **41**(3), 419–429 (2017).

Acknowledgements

We thank the members of the Ling Lab for providing helpful feedback and comments on this work.

Author contributions

J.P., X.S., M.K., J.T.M. and S.L. conceived and designed the experiments. Testing and data collection were performed by J.P., X.S., M.K., and E.P. Data analysis was conducted by J.P. and X.S. under the supervision of S.L. J.P., J.T.M. and S.L. wrote the manuscript and approved the final version of the manuscript for submission.

Funding

The funding was supported by National Institutes of Health, F31EY033650, EY028163.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-70895-5>.

Correspondence and requests for materials should be addressed to J.P. or S.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024