Multisensory Signals Inhibit Pupillary Light Reflex: Evidence from Pupil Oscillation

Authors: I understand. Please provide the Chinese academic text with tags, LaTeX formulas, and citations, and I will translate it to English following all specified requirements., Yuhui Cheng, Yi Jiang, Xiangyong Yuan, Yi Jiang

Date: 2021-04-27T00:00:00+00:00

Abstract

Multisensory integration, which enhances the stimulus saliency at the early stage of processing hierarchy, is recently shown to produce a larger pupil size than its unisensory constituents. Theoretically, any modulation on pupil size ought to be associated with the sympathetic and parasympathetic pathways that are sensitive to lights. But it remains poorly understood how pupillary light reflex is changed in a multisensory context. The present study evoked an oscillation of pupillary light reflex by periodically changing the luminance of a visual stimulus at 1.25 Hz. It was found that such induced pupil oscillation was substantially attenuated when the bright but not the dark phase of the visual flicker was periodically and synchronously presented with a burst of tones. This inhibition effect persisted when the visual flicker was task-irrelevant and out of attentional focus, but disappeared when the visual flicker was moved from the central field to the periphery. These findings not only offer a comprehensive characterization of the multisensory impact on pupil response to lightness, but also provide valuable clues to the individual contributions of the sympathetic and parasympathetic pathways to multisensory modulation of pupil size.

Full Text

Title Page

Multisensory Signals Inhibit Pupillary Light Reflex: Evidence from Pupil Oscillation

Xiangyong Yuan^{a,b,c,} , *Yuhui Cheng^{a,b,c}* , *Yi Jiang^{a,b,c,d,}*

^{a} State Key Laboratory of Brain and Cognitive Science, CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101, China

^{b} Department of Psychology, University of Chinese Academy of Sciences, Beijing, 100049, China

^{c} Chinese Institute for Brain Research, Beijing, 102206, China

^{d} Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, 230088, China

Short title: Multisensory inhibition of pupillary light reflex

To whom correspondence should be addressed. Email: yuanxy@psych.ac.cn or yijiang@psych.ac.cn

Abstract

Multisensory integration, which enhances stimulus saliency at early processing stages, has recently been shown to produce larger pupil sizes than its unisensory constituents. Theoretically, any modulation of pupil size should be associated with the sympathetic and parasympathetic pathways that are sensitive to light. However, it remains poorly understood how the pupillary light reflex changes in a multisensory context. The present study evoked an oscillation of the pupillary light reflex by periodically changing the luminance of a visual stimulus at 1.25 Hz. We found that such induced pupil oscillation was substantially attenuated when the bright—but not the dark—phase of the visual flicker was periodically and synchronously presented with a burst of tones. This inhibition effect persisted when the visual flicker was task-irrelevant and outside attentional focus, but disappeared when the visual flicker was moved from the central field to the periphery.

These findings not only provide a comprehensive characterization of the multisensory impact on pupil responses to lightness, but also offer valuable clues regarding the individual contributions of the sympathetic and parasympathetic pathways to multisensory modulation of pupil size.

Keywords: multisensory; pupil size; pupillary light reflex; oscillation; stimulus eccentricity; task relevance

Introduction

Combining information from distinct sensory modalities is beneficial for interacting with the environment. For instance, numerous studies have shown that multisensory integration facilitates detection, discrimination, and search (Leo, Bertini, di Pellegrino, & Làdavas, 2008; Noesselt, Bergmann, Hake, Heinze, & Fendrich, 2008; Van der Burg, Olivers, Bronkhorst, & Theeuwes, 2008), amplifies activation in sensory cortical areas (Kayser, Philiastides, & Kayser, 2017; Lewis & Noppeney, 2010; Noesselt et al., 2010; Van der Burg, Talsma, Olivers, Hickey, & Theeuwes, 2011; Werner & Noppeney, 2010, 2011) and subcortical nuclei (most importantly, the superior colliculus; see Stein & Stanford, 2008; Stein, Stanford, & Rowland, 2020). All this evidence reflects an enhancement of stimulus saliency by multisensory integration at an early processing stage. Since pupil size is sensitive to salient stimuli, with larger pupil sizes corresponding to stimuli with higher saliency (e.g., objectively high contrast, or subjectively easy-to-notice) regardless of modality (Liao, Kidani, Yoneya, Kashino, & Furukawa, 2016; Wang, Boehnke, Itti, & Munoz, 2014; Wang & Munoz, 2014), it is assumed that multisensory signals could dilate the pupil to a greater degree than their unisensory constituents.

The breakthrough came from a study on rhesus monkeys, which found that concurrently presented flashes and beeps in the periphery elicited a transient pupil dilation equal to the linear summation of the pupil sizes when they were presented in isolation (Wang et al., 2014). This finding was later replicated in humans by two independent studies, which further indicated in a detection task that larger pupil sizes were associated with faster saccadic or manual responses to audiovisual stimuli (Rigato, Rieger, & Romei, 2016; Wang, Blohm, Huang, Boehnke, & Munoz, 2017). Moreover, it has been shown that the enlarged pupil size when visual stimuli are presented in the central field in combination with auditory stimuli exceeds the linear summation of the pupil sizes obtained in each modality (Rigato et al., 2016, but see Van der Stoep, Van der Smagt, Notaro, Spock, & Naber, 2021). As is well known, pupil size is controlled by two antagonistic pathways: the sympathetic pathway that enlarges the pupil and the parasympathetic pathway that constricts it (Eckstein, Guerra-Carrillo, Miller Singley, & Bunge, 2017; Joshi & Gold, 2020; Larsen & Waters, 2018; Wang & Munoz, 2015). Therefore, pupil dilation induced by multisensory integration may reflect either increased sympathetic activation, decreased parasympathetic activation, or a combination of both (refer to the discussion in Wang et al., 2014 for more details).

Notably, both pathways are sensitive to ambient luminance. Pupil constriction to brightness (the pupillary light reflex) is primarily driven by parasympathetic activation, while pupil dilation to darkness is mainly driven by sympathetic activation. Investigations of how pupillary responses to different light levels are modulated in a multisensory context can provide insightful clues about the individual contributions of these two pathways to such modulation. It has already been shown that the onset latency of pupil dilation evoked by stimulus saliency can be as early as that of the pupillary light reflex, suggesting that the initial component of transient pupil dilation induced by higher visual contrast is likely a result of inhibited parasympathetic activation (Wang & Munoz, 2014). It is thus presumed that multisensory signals, if they enhance stimulus saliency, might specifically inhibit parasympathetic activation very rapidly, which may in turn transiently attenuate the pupillary light reflex. However, this hypothesis that multisensory signals could inhibit the pupillary light reflex has rarely been empirically tested.

To address this issue, the present study, following the pupil frequency tagging method (Naber, Alvarez, & Nakayama, 2013), periodically presented a simple, emotionally neutral stimulus and modulated its luminance at 1.25 Hz to elicit an oscillation in pupil size. In a series of four experiments, we presented a tone periodically at the same frequency as the repeated visual stimulus and manipulated the temporal congruency between the tone pulses and the bright phase of the visual flicker. Using this method, when the tone synchronized with the bright phase, the amplitude of this pupil oscillation could serve as a quantitative measure of the multisensory impact on the pupillary light reflex. In contrast, when the tone synchronized with the dark phase, the oscillatory amplitude quantified the multisensory impact on the dark reflex (or the relaxation of the pupillary light reflex). We examined whether this pupil oscillation was attenuated by tone pulses synchronized with the bright phase (Experiments 1 and 2) and further delineated the respective roles of stimulus eccentricity and task relevance in the multisensory inhibition of the pupillary light reflex (Experiments 3 and 4).

Experiment 1

Experiment 1 examined whether multisensory inputs inhibit the pupillary light reflex. The visual flickering stimulus, which changed its luminance periodically, would induce a dynamic change in pupil size—in other words, an oscillation of pupil size. If multisensory inputs inhibit the light reflex, the pupil oscillation would fluctuate within a smaller range (i.e., a smaller oscillatory amplitude) when the auditory stimuli were temporally congruent with the bright phase of the visual flicker, even though the actual luminance remained constant.

Participants

Sixteen participants were recruited for Experiment 1 (8 females; mean age: 21.9 ± 2.7 years). All participants had normal or corrected-to-normal vision and normal hearing, and were naïve to the purpose of the experiment. They provided informed written consent before the experiment and were compensated for their participation afterward. The study was approved by the institutional review board of the Institute of Psychology, Chinese Academy of Sciences (H18029), and adhered to the tenets of the Declaration of Helsinki.

Stimuli and Apparatus

A pioneering study has revealed that pupil oscillation is evoked by visual stimuli flickering at frequencies below ~3 Hz (Naber et al., 2013). In Experiment 1, a disc presented in the central field (radius: 1.61 degrees of visual angle), which flickered between brightness (22.56 cd/m^2) and darkness (15.15 cd/m^2) at 1.25 Hz, served as the visual stimulus [Figure 1: see original paper]a. The auditory stimulus was a tone (carrier frequency: 700 Hz; sample rate: 44100 Hz) with a duration of 0.4 seconds, played binaurally through headphones (Sennheiser HD 201). The loudness of the tone was set at a comfortable level throughout the experiment (~60 dB(A)) and kept constant for all participants.

The experiment was conducted in a dim, sound-attenuated room. Participants sat comfortably at a viewing distance of approximately 60 cm from the screen (refresh rate: 60 Hz, resolution: 1920 × 1080). The luminance of the gray background was 18.67 cd/m^2. All stimuli were generated by Matlab (The MathWorks Inc.) and presented using Psychtoolbox (Brainard, 1997; Pelli, 1997). Pupil size and eye position of the left eye were recorded using a video-based iView X Hi-Speed system (SMI, Berlin, Germany) at 500 Hz. Participants rested their heads on a chin-rest and were instructed to minimize head movements during the recording period. The recorded pupil size was analyzed and reported in arbitrary units (a.u.) without transformation into actual units (mm), as the relative change in pupil size was our primary interest. Generally, a pupil size of ~33 a.u. corresponded to a pupil size of 5 mm in the present study.

Figure 1 [Figure 1: see original paper]. Stimulus and an exemplar trial. (a) The luminance of the disc modulated at 1.25 Hz. The red arrow points out the oddball dot that participants had to count. (b)(c) The tone is synchronized with the bright phase of the disc in the AVbright condition (AVb), while synchronized with the dark phase of the disc in the AVdark condition (AVd).

Procedures

In each trial, a fixation point (a small dot, diameter: 0.16º) was first presented as a warning signal to inform participants that they should fixate at this position, prepare for the appearance of the visual stimuli, and avoid eye blinks. After a random duration of 1.5-2 seconds, the flickering disc was presented for 10 seconds [Figure 1: see original paper]a. To maintain participants' attention on the disc, they were required to complete an oddball counting task, in which small dots (diameter: 0.27º) flashed for 0.05 seconds at random positions on the disc, and participants counted how many times they saw the oddballs. There were 0-3 oddballs per trial, randomly determined and never presented simultaneously.

The oddball, if presented during the bright phase of the disc, had the same luminance as the dark phase of the disc, and vice versa. After inputting their answers, participants could relax their eyes for a while and then press the SPACE key to initiate the next trial.

There were four conditions in Experiment 1. In the visual-only condition (V-only), the disc was presented silently. In the auditory-only condition (A-only), the tone was periodically presented at 1.25 Hz, but the luminance of the disc remained constant, either bright or dark. The tone was synchronized with the bright phase of the disc in the AVbright condition (AVb), while synchronized with the dark phase of the disc in the AVdark condition (AVd; [Figure 1: see original paper]b and 1c). There were 64 trials total, divided into 4 blocks. In each block, each condition was repeated 4 times. A 5-point standard calibration of eye position was routinely conducted before the first and third blocks, but could be performed before any block if necessary.

Data Analysis

The accuracy of the oddball counting task was calculated as the number of trials with correct answers divided by the total number of trials. The raw pupil diameter in each trial was visually inspected, and trials with more than three blinks or other artifacts were excluded (2.1 trials excluded on average). For the remaining trials, data points where eye position deviated by 3 SDs from the mean, pupil diameter deviated by 3 SDs from the mean, or dropped substantially due to blinks or blink-like artifacts (i.e., the recording system failed to detect the corneal reflex but the pupil diameter still showed a blink-like shrinkage) were linearly interpolated. The artifact-free pupil diameter was then downsampled by averaging data points in every 0.05-second non-overlapping window, and detrended to minimize slow drift. To quantify pupil oscillation, fast Fourier Transform (FFT) was conducted for each trial, with the first second excluded to remove the transient response to stimulus onset (Naber et al., 2013). The amplitude of pupil oscillation was calculated as the modulus of the FFT complex coefficients and averaged across trials for each condition. Finally, the amplitude spectra were normalized by subtracting the amplitude averaged across the four neighboring frequency points (within \$±\$0.156 Hz) from the amplitude at each frequency point.

Statistics

To evaluate whether pupil size oscillated at 1.25 Hz, we performed one-sample t-tests on the normalized amplitude at 1.25 Hz for each condition. A normalized amplitude significantly larger than zero indicates robust pupil oscillation at that frequency. Next, we compared normalized amplitudes between conditions showing significant pupil oscillation using paired-sample t-tests to examine how multisensory signals modulate pupil oscillation. Reported p-values were Bonferroni-corrected for multiple comparisons unless otherwise specified. Additionally, we computed the JZS Bayesian factor (BF_{10}, H_1 versus H_0) using a Matlab toolbox developed by Bart Krekelberg, retrieved from GitHub (https://www.github.com/klabhub/bayesFactor). BF_{10} assesses the relative evidence for H_1 over H_0. A BF_{10} larger than 3 provides substantial evidence for H_1, while a BF_{10} smaller than 1/3 provides substantial evidence for H_0 (Dienes, 2014).

Results and Discussion

Accuracy on the oddball counting task approached 100% in all conditions (V-only: 0.98 ± 0.04; A-only: 0.97 ± 0.06; AVb: 0.99 ± 0.02; AVd: 0.98 ± 0.04), indicating that participants focused their attention on the central flicker during eye recording. As shown in [Figure 2: see original paper]a and 2b, pupil size oscillated during flicker presentation in all conditions except the A-only condition. One-sample t-tests confirmed that the normalized amplitude of pupil oscillation at 1.25 Hz was significantly greater than zero in the V-only, AVb, and AVd conditions (ts > 9, ps < 4e-7, $BF_{10} > 1e+5$), but not in the A-only condition

($t_{15} = 0.002$, $p > 0.9$, $BF_{10} = 0.255$; [Figure 2: see original paper]c and 2d). Therefore, the oscillatory amplitude in the A-only condition was excluded from subsequent comparisons examining audiovisual effects on pupil oscillation. As shown in [Figure 2: see original paper]d, paired-sample t-tests revealed that the strength of pupil oscillation significantly decreased when tones were temporally congruent with the bright phase of the visual stimuli, relative to visual stimuli presented alone (V-only vs. AVb: $t_{15} = 3.032$, $p = 0.025$, $BF_{10} = 6.313$). No other significant effects were found (AVd vs. AVb: $t_{15} = 1.475$, $p = 0.483$, $BF_{10} = 0.632$; V-only vs. AVd: $t_{15} = 0.111$, $p > 0.9$, $BF_{10} = 0.257$).

Experiment 1 demonstrated that pupil oscillation was induced by luminance modulation of the visual stimulus, consistent with previous findings (Naber et al., 2013). More importantly, it indicated that the pupillary light reflex was suppressed in a multisensory context. By contrast, when tones were synchronized with the dark phase of the visual flicker, pupillary responses were not significantly altered. Therefore, the relatively fast pupil frequency tagging method in Experiment 1 specifically captured multisensory inhibition of the pupillary light reflex with virtually no impact on the dark reflex (or relaxation from the pupillary light reflex; refer to the General Discussion for a possible account of this finding). To replicate these results, we conducted Experiment 2. Instead of luminance modulation, we periodically flashed a disc that was either brighter (Experiment 2a) or darker (Experiment 2b) than the background, and played a tone synchronously at the onset of the disc. This method allowed us to induce pupil oscillation as in Experiment 1 and examine whether tones had distinct impacts on the strength of pupil oscillations in Experiments 2a and 2b.

Figure 2. Results of Experiment 1. The baseline-corrected oscillation of pupil size when the disc started flickering from the bright phase (a) or the dark phase (b). The dashed colored lines represent pupil size in the first second of the trial, which is excluded from FFT analysis. (c) The amplitude spectra after FFT. The dashed lines indicate the target frequency of 1.25 Hz. (d) The normalized oscillatory amplitude at 1.25 Hz. Each circle represents the amplitude of pupil oscillation from one participant. The error bar indicates the standard error of the mean. ** indicates $p < 0.01$, uncorrected. AVb represents the AVbright condition; AVd represents the AVdark condition.

Experiment 2

In Experiment 2, the visual stimulus was repeatedly presented against the background, with tone pulses either synchronous with the stimulus or not. If the findings from Experiment 1 were robust, we expected that in Experiment 2a, where the visual stimulus was brighter than the background, pupil oscillation would be suppressed by synchronous tones, whereas in Experiment 2b, there would still be no increased pupil oscillation by synchronous tones when the visual stimulus was darker than the background.

Participants

Thirty-two new participants took part in Experiment 2, with 16 in Experiment 2a (12 females; mean age: 21.8 ± 2.5 years) and 16 in Experiment 2b (10 females; mean age: 21.2 ± 2.5 years).

Stimuli and Apparatus

The luminance of the disc was always 32.40 cd/m^2 in Experiment 2a and 9.20 cd/m^2 in Experiment 2b. The duration of the disc was 0.4 seconds. The tone and all other aspects were identical to Experiment 1.

Procedures

The main procedure of Experiment 2 was the same as Experiment 1, except that in each trial the disc flashed periodically at 1.25 Hz against the background to induce pupil oscillation. There were three conditions: V-only, AVb (in Experiment 2a) or AVd (in Experiment 2b), and AVbackground (AVbkg). In the V-only condition, the disc was presented alone. In the AVb or AVd condition, the tone and disc were presented simultaneously, while in the AVbkg condition, the tone was presented when the disc disappeared. There were 48 trials total, divided into 4 blocks. In each block, each condition was repeated 4 times.

Data Analysis and Statistics

The analysis and statistics were identical to Experiment 1.

Results and Discussion

Regardless of experiment or condition, all participants performed well on the oddball counting task (V-only: 0.94 ± 0.06; AVb: 0.97 ± 0.04; AVbkg: 0.98 ± 0.03 in Experiment 2a, and V-only: 0.95 ± 0.04; AVd: 0.96 ± 0.06; AVbkg: 0.96 ± 0.04 in Experiment 2b). Apparent pupil oscillation was observed in all conditions of Experiment 2 ([Figure 3: see original paper]a and 3b, ts > 7, ps $<$ 4e-5, $BF_{10} >$ 5e+3; pupil oscillation in each condition is shown in Supplementary Fig. 1). The results of Experiment 2a replicated Experiment 1. The amplitude of pupil oscillation decreased when the tone was synchronized with the disc of brighter luminance ([Figure 3: see original paper]a), compared to when the disc was presented alone (V-only vs. AVb: $t_{15} = 3.766$, $p = 0.006$, $BF_{10} = 22.385$) or accompanied by an asynchronous tone (AVbkg vs. AVb: $t_{15} = 3.192$, $p = 0.018$, $BF_{10} = 8.279$; V-only vs. AVbkg: $t_{15} = -0.233$, $p > 0.9$, $BF_{10} = 0.262$). In contrast, no significant amplitude changes in pupil oscillation were found in Experiment 2b where the tone was synchronized with the darker disc (ts < 1, ps > 0.9; V-only vs. AVd: $BF_{10} = 0.337$; AVbkg vs. AVd: $BF_{10} = 0.277$; V-only vs. AVbkg: $BF_{10} = 0.284$; [Figure 3: see original paper]b). Experiment 2 therefore revealed that audiovisual signals attenuated the strength of pupil oscillation evoked by repeated brighter visual stimuli, while they did not increase pupil oscillation when the visual stimuli were darker than the background. As hypothesized, these results support the notion that at relatively fast stimulus repetition speeds (e.g., 1.25 Hz), the pupillary light reflex can be specifically inhibited in a multisensory context.

According to the principle of inverse effectiveness, the strength of cross-modal stimuli should be relatively low for maximal enhancement of multisensory integration (Noesselt et al., 2010; Stein & Stanford, 2008; Stein et al., 2020). One might argue that the failure to reveal enhanced pupil oscillation in Experiment 2b is attributable to relative stimulus strength rather than the speed of the induced pupil oscillation. In response, we reduced the luminance difference between the visual stimulus and the background and examined whether multisensory signals could enhance pupil oscillation. However, in a supplemental experiment using the same analysis protocol, although repeated presentation of a visual stimulus isoluminant with the background induced pupil oscillation at an extremely low magnitude (~0.03 a.u.), we still could not observe increased pupil oscillation (Supplementary Fig. 2).

Furthermore, we noticed that among the four previous studies reporting pupil dilation induced by audiovisual integration, two presented stimuli in the peripheral visual field as they were interested in orienting behaviors (Wang et al., 2017; Wang et al., 2014), while two presented stimuli in the central visual field (Rigato et al., 2016; Van der Stoep et al., 2021). It appears that audiovisual signals can dilate pupil size regardless of where the visual stimulus appears. To further characterize the multisensory modulation of pupil oscillation induced by luminance changes, we conducted Experiment 3 by moving the visual stimulus from the central to the peripheral field to examine the role of visual eccentricity in the observed effect.

Figure 3. Results of Experiments 2-4. The normalized oscillatory amplitude at 1.25 Hz for Experiments 2, 3, and 4, respectively. Each circle represents the amplitude from one participant. The error bar indicates the standard error of the mean. ** indicates $p < 0.01$, * indicates $p < 0.05$, both uncorrected. AVb represents the AVbright condition; AVd represents the AVdark condition; AVbkg represents the AVbackground condition. AVr represents the AVrandom condition.

Experiment 3

Some studies have revealed differential multisensory effects dependent on stimulus eccentricity (Gleiss & Kayser, 2013; Leo et al., 2008; Nidiffer, Stevenson, Fister, Barnett, & Wallace, 2016; van Atteveldt, Peterson, & Schroeder, 2014), but the impact of multisensory integration on pupil size appears to be independent of stimulus eccentricity (Rigato et al., 2016; Van der Stoep et al., 2021; Wang et al., 2017). Experiment 3 evaluated whether audiovisual inhibition of the pupillary light reflex remained when visual stimuli were moved from the central to the peripheral field.

Participants

A new group of 16 participants took part in Experiment 3 (10 females; mean age: 23.3 ± 3.9 years).

Stimuli and Apparatus

In Experiment 3, the visual stimulus was also a disc, but presented in the left or right peripheral visual field (eccentricity: 10.72º from the center of the disc to fixation). The luminance of the disc changed at 1.25 Hz between brightness (47.47 cd/m^2) and darkness (3.03 cd/m^2), as in Experiment 1. The luminance range of the disc was expanded because in our preliminary data, the disc had to flicker across a larger luminance range to induce pupil oscillation with amplitude approaching that in the central field. The auditory stimulus, still presented binaurally through headphones, had its sound level attenuated by 50% in the left or right channel to mimic tones coming from the opposite side. For instance, we would perceive a tone source from the left side if the sound level of the right channel was set somewhat lower than that of the left channel. Although this manipulation could not precisely align the positions of the tones and flickers, minor spatial misalignment may not affect the results according to previous findings (Gleiss & Kayser, 2013; van Atteveldt et al., 2014).

Procedures, Data Analysis, and Statistics

The procedure, analysis, and statistics were identical to Experiment 1.

Results and Discussion

Accuracies on the oddball counting task were 0.97 ± 0.05 in the V-only condition, 0.98 ± 0.03 in the A-only condition, 0.95 ± 0.07 in the AVb condition, and 0.96 ± 0.04 in the AVd condition. As in Experiments 1 and 2, we observed significant pupil oscillation in the three conditions where the flickering disc was presented, with amplitudes at 1.25 Hz significantly greater than zero (ts > 7, ps < 2e-5, BF_{10} > 6e+3), but not in the A-only condition (t_{15} = 1.859, p > 0.3, BF_{10} = 1.024; [Figure 3: see original paper]c). However, paired-sample t-tests failed to reveal any significant differences in pupil oscillation amplitudes across the three conditions (ts < 1, ps > 0.9; V-only vs. AVb: BF_{10} = 0.370; AVd vs. AVb: BF_{10} = 0.322; V-only vs. AVd: BF_{10} = 0.257). The evidence thus tends to support that the pupillary light reflex is not inhibited by audiovisual signals when the visual stimulus is presented in the periphery.

The lack of pupil oscillation inhibition in Experiment 3 can be attributed neither to the relatively weaker amplitude of the evoked pupil oscillation (see [Figure 3: see original paper]d) nor to an absence of audiovisual combination in a repetition paradigm (Noesselt et al., 2007; Talsma & Woldorff, 2005; also see Supplementary Information and Supplementary Fig. 3, where we found that onset pupil size was significantly dilated by audiovisual inputs relative to visual inputs, consistent with Wang et al., 2017). It is most likely that in Experiment 3, despite being fused, multisensory signals failed to inhibit the pupillary light reflex evoked by peripheral visual stimuli. This result contrasts with previous findings focusing on the multisensory impact on pupil orienting responses (Wang et al., 2017; Wang et al., 2014).

Thus far, the visual flicker was always required to be attended because it was task-relevant. Given that several studies have found that even task-irrelevant bimodal signals show signs of fusion relative to unimodal signals (Heeman, Nijboer, Van der Stoep, Theeuwes, & Van der Stigchel, 2016; Krause, Schneider, Engel, & Senkowski, 2012; Mühlberg & Müller, 2020; Matusz et al., 2015), we hypothesized that inhibition of the pupillary light reflex would persist even when visual and auditory stimuli were task-irrelevant and outside attentional focus. We conducted Experiment 4 to test this hypothesis.

Experiment 4

Experiment 4 replaced the oddball counting task with a Rapid Serial Visual Presentation (RSVP) task (following Santangelo & Spence, 2007) and relocated the visual flicker to the surround of the RSVP display so that the visual flicker was now completely task-irrelevant. We examined whether the induced pupil oscillation was still inhibited when tone pulses were temporally congruent with the bright phase of the surround visual flicker, as in Experiment 1.

Participants

Sixteen participants took part in Experiment 4 (9 females; mean age: 22.0 ± 2.3 years).

Stimuli and Apparatus

For the visual stimulus, the disc was replaced by a ring (inner circle radius: 1.34^{o}; outer circle radius: 2.68^{o}), with its luminance flickering between 26.8 cd/m^2 and 34.4 cd/m^2 at a frequency of 1.25 Hz. A stream of letters ($1.61^{o} \times 1.61^{o}$) was rapidly presented at 6 Hz within the inner circle of the ring without blank intervals, so each letter lasted 167 ms [Figure 3: see original paper]d. The letters were randomly selected from the alphabet, with B, F, I, J, L, O, P, Q, W, and Z excluded. Each letter was always different from its neighbors in the stream. Among the letters, some numbers of the same size were embedded, randomly selected from 2, 3, 4, 6, 7, and 9. The auditory stimulus was identical to Experiment 1.

Procedures

In Experiment 4, participants performed an RSVP task. In each trial, they counted how many times numbers appeared (0–3 times) among the rapidly presented stream of letters and were instructed in advance to ignore the flickering

ring outside the letter streams throughout the experiment. The visual inducer of pupil oscillation was therefore outside attentional focus and should be considered task-irrelevant. There were three conditions: V-only, AVbright, and AVrandom. The V-only and AVb conditions were the same as in Experiments 1 and 3, except that a new AVrandom condition (AVr) was used as a control. In this condition, the tone was not played synchronously with the dark phase of the ring, but was randomly played at any time from 0.2 seconds after bright phase onset to 0.2 seconds before dark phase offset. Comparing pupil oscillations from the AVb and AVr conditions could further demonstrate whether changes in the pupillary light reflex were affected by the temporal synchrony between auditory and visual stimuli. Participants completed 48 trials total, divided into 4 blocks, with each condition repeated 16 times.

Data Analysis and Statistics

The analysis and statistics were identical to Experiments 1-3.

Results and Discussion

Performance on the oddball counting task was 0.96 ± 0.05 in the V-only condition, 0.97 ± 0.06 in the AVb condition, and 0.93 ± 0.08 in the AVr condition, indicating that attention was concentrated on the RSVP task. Although task-irrelevant, the visual flicker also induced significant pupil oscillation ([Figure 3: see original paper]d, ts > 5, ps $< 1e-4$, $BF_{10} > 700$). The pupil oscillated at a relatively lower amplitude (approximately 2/3 of the amplitude in Experiments 1 and 2a), probably because the stimuli were unattended (Naber et al., 2013) or eccentrically located. Consistent with Experiments 1 and 2a, the amplitude of pupil oscillation decreased when tones were temporally congruent with the bright phase of the visual stimuli, compared to when visual stimuli were presented alone (V-only vs. AVb: $t_{15} = 2.904$, $p = 0.033$, $BF_{10} = 5.093$) and when audiovisual stimuli were temporally asynchronous (AVr vs. AVb: $t_{15} = 2.898$, $p = 0.033$, $BF_{10} = 5.040$; V-only vs. AVr: $t_{15} = -0.694$, $p > 0.9$, $BF_{10} = 0.316$). These results indicate that the pupillary light reflex can be inhibited in a multisensory context even when the visual inducer is task-irrelevant. They also demonstrate that inhibition of the pupillary light reflex is sensitive to cross-modal temporal relationships.

To further explore whether task relevance modulates this inhibition effect, we calculated an inhibition index (i.e., the difference in oscillatory amplitude between the AVb condition and other conditions, including V-only, AVd, AVbkg, or AVr conditions, with the latter three conditions uniformly represented by AVincongruent, abbreviated as AVinc for convenience) for Experiments 1, 2a, and 4 separately, then compared the inhibition index of Experiment 4 with those from Experiments 1 and 2a using independent-sample t-tests. The results revealed no significant effects [for Experiment 1 vs. 4, ts < 0.8, ps > 0.9, BF_{10} (IndexVonly-AVb) $= 0.384$, BF_{10} (IndexAVinc-AVb) $= 0.410$; for Experiment 2a vs. 4, ts < 0.4, ps > 0.9, BF_{10} (IndexVonly-AVb) $= 0.341$, BF_{10} (IndexAVinc-

AVb) = 0.352]. Taken together, these results tend to suggest that inhibition of the pupillary light reflex in a multisensory context is immune to task irrelevance.

General Discussion

Previous studies have shown that multisensory integration enlarges pupil size (Rigato et al., 2016; Van der Stoep et al., 2021; Wang et al., 2017; Wang et al., 2014). Using a pupil oscillation frequency tagging method (Naber et al., 2013), the present study demonstrated for the first time that pupil oscillation evoked by a visual flicker was attenuated when a sequence of tone pulses was synchronized with the bright phase of the visual flicker, relative to when it was synchronized with the dark phase or when no tones were presented. This implies that multisensory signals can specifically inhibit the pupillary light reflex when luminance alternation occurs at relatively fast speeds (e.g., 1.25 Hz).

As parasympathetic activation constricts pupil size and sympathetic activation dilates it (Eckstein et al., 2017; Joshi & Gold, 2020; Larsen & Waters, 2018; Wang & Munoz, 2015), there are parallel explanations for the previously observed stronger pupil dilation to multisensory signals (Rigato et al., 2016; Van der Stoep et al., 2021; Wang et al., 2017; Wang et al., 2014): inhibited parasympathetic activation, enhanced sympathetic activation, or a combination of both. The currently observed inhibition of the pupillary light reflex is likely caused by inhibition of parasympathetic activation, as the pupillary light reflex is primarily driven by parasympathetic activation (Clarke, Zhang, & Gamlin, 2003; Joshi & Gold, 2020). However, considering that the two pupil-related pathways are complexly interconnected (see Joshi & Gold, 2020, Box 1), inhibition of the pupillary light reflex could equally be explained by increased phasic sympathetic activity, which can dilate the pupil and thereby counteract the pupillary light reflex. Because both unimodal and bimodal stimuli were repeated periodically at the relatively fast rate of 1.25 Hz in our experiments, only multisensory impacts that rapidly changed the trough or peak of the pupil oscillation within the cyclic period (e.g., 400 ms) could alter the amplitude of pupil oscillation (otherwise the trough and peak would be equally changed, leaving the oscillatory amplitude largely unchanged). Parasympathetic activity, which has a very short onset latency for pupil constriction (< ~270 ms with less than ~800 ms to reach its extreme; Clarke, Zhang, & Gamlin, 2003; Wang & Munoz, 2014), is capable of being transiently inhibited within this limited time. By contrast, pupil dilation caused by sympathetic activation (primarily through the locus coeruleus-noradrenergic system) arises slowly, with an onset latency of ~330 ms or more (often with a peak latency exceeding 1 second; Chapman, Oka, Bradshaw, Jacobson, & Donaldson, 1999; Liao, Yoneya, Kidani, Kashino, & Furukawa, 2016; Steiner & Barry, 2011; Wang & Munoz, 2014), and may be too sluggish to be sufficiently enhanced within this cyclic period. Moreover, we would have concurrently observed enhanced pupil oscillation when the tone synchronized with the dark phase if phasic sympathetic activation was enhanced, but this was not the case in Experiments 1 and 2.

One might further argue that this phasic sympathetic activity, although arising slowly, could be gradually enhanced and accumulated during repetition of bimodal inputs, and that the inhibited pupillary light reflex might be confounded by possible pupil dilation caused by this accumulation. We provide further evidence against this possibility. First, although an oddball event can enlarge pupil size, pupil size for a repeated event habituates as its novelty gradually decreases (Liao et al., 2016; Netser, Ohayon, & Gutfreund, 2010; Steiner & Barry, 2011). Based on these results, our experiments could hardly produce a gradual increase in pupil size through periodic presentation of simple, emotionally neutral visual stimuli and pure tones. Second, additional analyses splitting trials into early and late parts were performed to statistically assess whether gradual changes in pupil size during stimulus repetition influenced multisensory inhibition of the pupillary light reflex. Analyses for Experiments 1, 2a, and 4 found almost identical results in early and late parts (with significant inhibition of the pupillary light reflex more frequently observed in the early part), indicating little evidence for gradual pupil dilation and corresponding confounding of our main observation (for detailed analysis, see Supplementary Information).

Although inhibition of parasympathetic activation most likely accounts for our observations, we do not claim that sympathetic activation cannot be enhanced in a multisensory context. Unlike the parasympathetic pathway, which can be transiently inhibited, we propose that the sympathetic pathway may be enhanced by multisensory signals in a slow and sustained manner. This is compatible with previous findings demonstrating that pupil dilation to multisensory signals could, on one hand, be as early as that of the pupillary light reflex (Wang et al., 2017), while on the other hand arise late and persist for a relatively long time (Rigato et al., 2016; Van der Stoep et al., 2021). This assumption can also explain the inconsistency between our observation and a recent study (Van der Stoep et al., 2021), which reported no distinction between phasic pupil responses to light and dark when each trial included only one unimodal or bimodal stimulus but with adequate time to observe pupil changes.

Setting aside possible explanations regarding the underlying pathway, the present study further revealed that multisensory inhibition of the pupillary light reflex was only observed when the visual flicker was located in the central field. This result contrasts with findings that pupil dilation by multisensory signals may be independent of stimulus eccentricity (Rigato et al., 2016; Van der Stoep et al., 2021; Wang et al., 2017; Wang et al., 2014). However, it is not entirely unexpected, as multisensory integration in central and peripheral fields has been proposed to be functionally complementary. Stimuli in the central field may be prioritized for accurate discrimination and recognition regarding their properties and features, whereas stimuli in the periphery may signal potential threats requiring rapid orienting responses, either overt or covert (Chen, Maurer, Lewis, Spence, & Shore, 2017; Gleiss & Kayser, 2013; Leo et al., 2008; Nidiffer et al., 2016; van Atteveldt et al., 2014). It is thus possible that once the visual flicker had already attracted covert attention in Experiment 3, which required fixation at the center, overt orienting responses such as saccades toward the target location would be suppressed thereafter. Given that the superior colliculus (SC) is an important nucleus for both saccade generation (Coe & Munoz, 2017) and multisensory integration (King, 2004; Stein & Stanford, 2008; Stein et al., 2020), suppression of saccades may be accompanied by attenuation of multisensory interaction in SC. This likely leads to no multisensory modulation of the pupillary light reflex in the periphery.

Although dependent on stimulus eccentricity, fusion of multisensory inputs has been proposed to be independent of task relevance. Previous studies have reported that even task-irrelevant cross-modal signals can exert stronger interference on the currently performed task compared to unimodal distractors (Heeman et al., 2016; Krause et al., 2012; Matusz et al., 2015, but see improvement in Mühlberg & Müller, 2020 and no effect in Experiment 4 of Lunn, Sjoblom, Ward, Soto-Faraco, & Forster, 2019). Despite finding no interference with the RSVP task in the present study, the pupillary light reflex induced by visual stimuli that were task-irrelevant and outside attentional focus was inhibited by temporally congruent tone pulses in Experiment 4. This result confirms that multisensory inhibition of the pupillary light reflex may be insensitive to the attentional set defined by task goals, and perhaps controlled by a bottom-up, stimulus-driven mechanism. Moreover, it suggests that changes in pupil size can serve as an effective physiological proxy for task-irrelevant multisensory effects, similar to other indices such as steady-state visual evoked potentials (Krause et al., 2012). Notably, task irrelevance does not necessarily imply immunity to attentional load. The higher RSVP accuracy in Experiment 4 ensures that task-relevant stimuli were fully attended on one hand, but indicates a medium level of attentional load on the other. As several studies have reported that the effect of multisensory integration is attenuated at higher attentional loads (Fairhall & Macaluso, 2009; Morís Fernández, Visser, Ventura-Campos, Ávila, & Soto-Faraco, 2015; Senkowski, Talsma, Herrmann, & Woldorff, 2005; Talsma, Doty, & Woldorff, 2007; Talsma & Woldorff, 2005, but see Santangelo & Spence, 2007; Wahn & König, 2015), it remains to be investigated in future research how the pupillary light reflex in a multisensory context would be affected when attentional load is substantially increased.

Regarding the neural substrate underlying this multisensory influence on the pupillary light reflex, we infer that the most relevant structure is the SC. The SC has been shown to project directly or indirectly to the pretectal olivary nucleus and the Edinger-Westphal nucleus on the parasympathetic pathway (Harting, Huerta, Frankfurter, Strominger, & Royce, 1980; May, 2006; May, Warren, Bohlen, Barnerssoi, & Horn, 2016; Wang & Munoz, 2015). It also receives input from the locus coeruleus and may indirectly influence the sympathetic pathway through the mesencephalic cuneiform nucleus (Joshi & Gold, 2020; Wang & Munoz, 2015). Electrical microstimulation of the intermediate layers of the SC can produce transient pupil dilation, verifying the SC' s ability to modulate pupil size (Wang et al., 2014; Wang, Boehnke, White, & Munoz, 2012). Importantly, the SC, whose deeper layers can integrate multisensory inputs, has repeatedly been shown to be a subcortical hub of multisensory integration (Stein & Stanford, 2008; Stein et al., 2020). Taken together, it is most probable that the SC first combines temporally congruent auditory and visual inputs, then modulates pupil size by suppressing parasympathetic activation (or enhancing sympathetic activation). Cross-modal integration in the SC is also compatible with the observed stimulus eccentricity dependence, as discussed earlier. However, it remains possible that auditory inputs may directly inhibit parasympathetic activity (or increase sympathetic activity) through the locus coeruleus (Joshi, Li, Kalwani, & Gold, 2016). It is difficult to disentangle how multisensory signals interact to affect the pupillary light reflex purely from the physiological data reported here, although the SC might be a key neural candidate involved in this process.

In conclusion, the present study demonstrated that the pupillary light reflex in response to a central visual inducer can be specifically inhibited in a multisensory context regardless of task relevance. This inhibition of the pupillary light reflex not only implies that multisensory signals can modulate pupil-related neural pathways, but also provides another easily measured pupillometric indicator of multisensory interaction independent of explicit responses. Intriguingly, if signals from other modalities are capable of promoting pupil constriction, would an increased pupillary light reflex be specifically observed? This would be considered complementary to the current findings.

Data Accessibility

All data used for statistics and code to generate figures can be found at https://osf.io/npaer/?view_{only}=287f4f90a4304065b4aecf243246f134.

Competing Interests

The authors declare no competing financial interests.

Funding

This study was supported by research grants from the National Natural Science Foundation of China (31830037, 31525011, 31600884), the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences (Y8CX372005), and the Strategic Priority Research Program and the Key Research Program of Frontier Sciences of the Chinese Academy of Sciences (XDB32010300, QYZDB-SSW-SMC030), Beijing Municipal Science & Technology Commission, Shenzhen-Hong Kong Institute of Brain Science, and Fundamental Research Funds for the Central Universities.

Acknowledgments

We appreciate the critical comments of Dr. Ying Wang and Dr. Xilei Zhang on this study.

References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433. doi: 10.1163/156856897X00357

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. doi: https://doi.org/10.1017/S0048577299970373

Chen, Y.-C., Maurer, D., Lewis, T. L., Spence, C., & Shore, D. I. (2017). Central-peripheral differences in audiovisual and visuotactile event perception. *Attention, Perception, & Psychophysics*, 79(8), 2552-2563. doi: 10.3758/s13414-017-1396-4

Clarke, R. J., Zhang, H., & Gamlin, P. D. R. (2003). Characteristics of the Pupillary Light Reflex in the Alert Rhesus Monkey. *Journal of Neurophysiology*, 89(6), 3179-3189. doi: 10.1152/jn.01131.2002

Coe, B. C., & Munoz, D. P. (2017). Mechanisms of saccade suppression revealed in the anti-saccade task. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1718), 20160192. doi: doi:10.1098/rstb.2016.0192

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5(781). doi: 10.3389/fpsyg.2014.00781

Eckstein, M. K., Guerra-Carrillo, B., Miller Singley, A. T., & Bunge, S. A. (2017). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience*, 25, 69-91. doi: 10.1016/j.dcn.2016.11.001

Fairhall, S. L., & Macaluso, E. (2009). Spatial attention can modulate audiovisual integration at multiple cortical and subcortical sites. *European Journal of Neuroscience*, 29(6), 1247-1257. doi: 10.1111/j.1460-9568.2009.06688.x

Gleiss, S., & Kayser, C. (2013). Eccentricity dependent auditory enhancement of visual stimulus detection but not discrimination. *Frontiers in Integrative Neuroscience*, 7, 52. doi: 10.3389/fnint.2013.00052

Harting, J. K., Huerta, M. F., Frankfurter, A. J., Strominger, N. L., & Royce, G. J. (1980). Ascending pathways from the monkey superior colliculus: An autoradiographic analysis. *Journal of Comparative Neurology*, 192(4), 853-882. doi: 10.1002/cne.901920414

Heeman, J., Nijboer, T. C. W., Van der Stoep, N., Theeuwes, J., & Van der Stigchel, S. (2016). Oculomotor interference of bimodal distractors. *Vision Research*, 123, 46-55. doi: 10.1016/j.visres.2016.04.002

Joshi, S., & Gold, J. I. (2020). Pupil Size as a Window on Neural Substrates of Cognition. *Trends in Cognitive Sciences*, 24(6), 466-480. doi: 10.1016/j.tics.2020.03.005

Joshi, S., Li, Y., Kalwani, Rishi M., & Gold, Joshua I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, 89(1), 221-234. doi: 10.1016/j.neuron.2015.11.028

Kayser, S. J., Philiastides, M. G., & Kayser, C. (2017). Sounds facilitate visual motion discrimination via the enhancement of late occipital visual representations. *NeuroImage*, 148, 31-41. doi: 10.1016/j.neuroimage.2017.01.010

King, A. J. (2004). The superior colliculus. *Current Biology*, 14(9), R335-R338. doi: https://doi.org/10.1016/j.cub.2004.04.018

Krause, H., Schneider, T., Engel, A., & Senkowski, D. (2012). Capture of visual attention interferes with multisensory speech processing. *Frontiers in Integrative Neuroscience*, 6(67). doi: 10.3389/fnint.2012.00067

Larsen, R. S., & Waters, J. (2018). Neuromodulatory Correlates of Pupil Dilation. *Frontiers in Neural Circuits*, 12(21). doi: 10.3389/fncir.2018.00021

Leo, F., Bertini, C., di Pellegrino, G., & Làdavas, E. (2008). Multisensory integration for orienting responses in humans requires the activation of the superior colliculus. *Experimental Brain Research*, 186(1), 67-77. doi: 10.1007/s00221-007-1204-9

Lewis, R., & Noppeney, U. (2010). Audiovisual Synchrony Improves Motion Discrimination via Enhanced Connectivity between Early Visual and Auditory Areas. *The Journal of Neuroscience*, 30(37), 12329-12339. doi: 10.1523/jneurosci.5745-09.2010

Liao, H.-I., Kidani, S., Yoneya, M., Kashino, M., & Furukawa, S. (2016). Correspondences among pupillary dilation response, subjective salience of sounds, and loudness. *Psychonomic Bulletin & Review*, 23(2), 412-425. doi: 10.3758/s13423-015-0898-0

Liao, H.-I., Yoneya, M., Kidani, S., Kashino, M., & Furukawa, S. (2016). Human Pupillary Dilation Response to Deviant Auditory Stimuli: Effects of Stimulus Properties and Voluntary Attention. *Frontiers in Neuroscience*, 10(43). doi: 10.3389/fnins.2016.00043

Lunn, J., Sjoblom, A., Ward, J., Soto-Faraco, S., & Forster, S. (2019). Multisensory enhancement of attention depends on whether you are already paying attention. *Cognition*, 187, 38-49. doi: https://doi.org/10.1016/j.cognition.2019.02.008

Mühlberg, S., & Müller, M. M. (2020). Alignment of Continuous Auditory and Visual Distractor Stimuli Is Leading to an Increased Performance. *Frontiers in Psychology*, 11(790). doi: 10.3389/fpsyg.2020.00790

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children’ s limited attention. *Cognition*, 136, 156-165. doi: 10.1016/j.cognition.2014.11.031

May, P. J. (2006). The mammalian superior colliculus: laminar structure and connections. In J. A. Büttner-Ennever (Ed.), *Progress in Brain Research* (Vol. 151, pp. 321-378): Elsevier.

May, P. J., Warren, S., Bohlen, M. O., Barnerssoi, M., & Horn, A. K. E. (2016). A central mesencephalic reticular formation projection to the Edinger–Westphal nuclei. *Brain Structure and Function*, 221(8), 4073-4089. doi: 10.1007/s00429-015-1147-z

Morís Fernández, L., Visser, M., Ventura-Campos, N., Ávila, C., & Soto-Faraco, S. (2015). Top-down attention regulates the neural expression of audiovisual integration. *NeuroImage*, 119, 272-285. doi: 10.1016/j.neuroimage.2015.06.052

Naber, M., Alvarez, G. A., & Nakayama, K. (2013). Tracking the allocation of attention using human pupillary oscillations. *Frontiers in Psychology*, 10.3389/fpsyg.2013.00919

Netser, S., Ohayon, S., & Gutfreund, Y. (2010). Multiple Manifestations of Microstimulation in the Optic Tectum: Eye Movements, Pupil Dilations, and Sensory Priming. *Journal of Neurophysiology*, 104(1), 108-118. doi: 10.1152/jn.01142.2009

Nidiffer, A. R., Stevenson, R. A., Fister, J. K., Barnett, Z. P., & Wallace, M. T. (2016). Interactions between space and effectiveness in human multisensory performance. *Neuropsychologia*, 88, 83-91. doi: 10.1016/j.neuropsychologia.2016.01.031

Noesselt, T., Bergmann, D., Hake, M., Heinze, H.-J., & Fendrich, R. (2008). Sound increases the saliency of visual events. *Brain Research*, 1220, 157-163.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H.-J., & Driver, J. (2007). Audiovisual Temporal Correspondence Modulates Human Multisensory Superior Temporal Sulcus Plus Primary Sensory Cortices. *Journal of Neuroscience*, 27(42), 11431-11441. doi: 10.1523/jneurosci.2252-07.2007

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Heinze, H.-J., & Driver, J. (2010). Sound-Induced Enhancement of Low-Intensity Vision: Multisensory Influences on Human Sensory-Specific Cortices and Thalamic Bodies Relate to Perceptual Enhancement of Visual Detection Sensitivity. *Journal of Neuroscience*, 30(41), 10.1523/jneurosci.4524-09.2010

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437-442. doi: https://doi.org/10.1163/156856897X00366

Rigato, S., Rieger, G., & Romei, V. (2016). Multisensory signalling enhances pupil dilation. *Scientific Reports*, 6, 26188. doi: 10.1038/srep26188

Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology:*

Human Perception and Performance, 33(6), 1311-1321. doi: 10.1037/0096-1523.33.6.1311

Senkowski, D., Talsma, D., Herrmann, C., & Woldorff, M. (2005). Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Experimental Brain Research*, 166(3), 411-426. doi: 10.1007/s00221-005-2381-z

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9(4), 255-266.

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2020). Multisensory Integration and the Society for Neuroscience: Then and Now. *The Journal of Neuroscience*, 40(1), 3-11. doi: 10.1523/jneurosci.0737-19.2019

Steiner, G. Z., & Barry, R. J. (2011). Pupillary responses and event-related potentials as indices of the orienting reflex. *Psychophysiology*, 48(12), 1648-1655. doi: 10.1111/j.1469-8986.2011.01271.x

Talsma, D., Doty, T. J., & Woldorff, M. G. (2007). Selective Attention and Audiovisual Integration: Is Attending to Both Modalities a Prerequisite for Early Integration? *Cerebral Cortex*, 17(3), 679-690. doi: 10.1093/cercor/bhk016

Talsma, D., & Woldorff, M. G. (2005). Selective Attention and Multisensory Integration: Multiple Phases of Effects on the Evoked Brain Activity. *Journal of Cognitive Neuroscience*, 17(7), 1098-1114. doi: doi:10.1162/0898929054475172

van Atteveldt, N. M., Peterson, B. S., & Schroeder, C. E. (2014). Contextual control of audiovisual integration in low-level sensory cortices. *Human Brain Mapping*, 35(5), 2394-2411. doi: 10.1002/hbm.22336

Van der Burg, E., Olivers, C., Bronkhorst, A., & Theeuwes, J. (2008). Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 34(5), 1053-1065.

Van der Burg, E., Talsma, D., Olivers, C. N. L., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *NeuroImage*, 55(3), 1208-1218. doi: 10.1016/j.neuroimage.2010.12.068

Van der Stoep, N., Van der Smagt, M. J., Notaro, C., Spock, Z., & Naber, M. (2021). The additive nature of the human multisensory evoked pupil response. *Scientific Reports*, 11(1), 707. doi: 10.1038/s41598-020-80286-1

Wahn, B., & König, P. (2015). Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration. *Frontiers in Psychology*, 6. doi: 10.3389/fpsyg.2015.01084

Wang, C.-A., Blohm, G., Huang, J., Boehnke, S. E., & Munoz, D. P. (2017). Multisensory integration in orienting behavior: Pupil size, microsaccades, and saccades. *Biological Psychology*, 129(Supplement C), 36-44. doi: 10.1016/j.biopsycho.2017.07.024

Wang, C.-A., Boehnke, S. E., Itti, L., & Munoz, D. P. (2014). Transient Pupil Response Is Modulated by Contrast-Based Saliency. *The Journal of Neuroscience*, 34(2), 408-417. doi: 10.1523/jneurosci.3550-13.2014

Wang, C.-A., Boehnke, S. E., White, B. J., & Munoz, D. P. (2012). Microstimulation of the Monkey Superior Colliculus Induces Pupil Dilation Without Evoking Saccades. *The Journal of Neuroscience*, 32(11), 3629-3636. doi: 10.1523/jneurosci.5512-11.2012

Wang, C.-A., & Munoz, D. P. (2014). Modulation of stimulus contrast on the human pupil orienting response. *European Journal of Neuroscience*, 40(5), 2822-2832. doi: 10.1111/ejn.12641

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

Werner, S., & Noppeney, U. (2010). Distinct Functional Contributions of Primary Sensory and Association Areas to Audiovisual Integration in Object Categorization. *The Journal of Neuroscience*, 30(7), 2662-2675. doi: 10.1523/jneurosci.5091-09.2010

Werner, S., & Noppeney, U. (2011). The Contributions of Transient and Sustained Response Codes to Audiovisual Integration. *Cerebral Cortex*, 21(4), 920-931. doi: 10.1093/cercor/bhq161

Note: Figure translations are in progress. See original paper for figures.