

Correspondences among pupillary dilation response, subjective salience of sounds, and loudness

Hsin-I Liao¹ · Shunsuke Kidani¹ · Makoto Yoneya¹ · Makio Kashino^{1,2} · Shigeto Furukawa¹

Published online: 11 July 2015
© Psychonomic Society, Inc. 2015

Abstract A pupillary dilation response is known to be evoked by salient deviant or contrast auditory stimuli, but so far a direct link between it and subjective salience has been lacking. In two experiments, participants listened to various environmental sounds while their pupillary responses were recorded. In separate sessions, participants performed subjective pairwise-comparison tasks on the sounds with respect to their salience, loudness, vigorousness, preference, beauty, annoyance, and hardness. The pairwise-comparison data were converted to ratings on the Thurstone scale. The results showed a close link between subjective judgments of salience and loudness. The pupil dilated in response to the sound presentations, regardless of sound type. Most importantly, this pupillary dilation response to an auditory stimulus positively correlated with the subjective salience, as well as the loudness, of the sounds (Exp. 1). When the loudnesses of the sounds were identical, the pupil responses to each sound were similar and were not correlated with the subjective judgments of salience or loudness (Exp. 2). This finding was further confirmed by analyses based on individual stimulus pairs and participants. In Experiment 3, when salience and loudness were manipulated by systematically changing the sound pressure level and acoustic characteristics, the pupillary dilation

response reflected the changes in both manipulated factors. A regression analysis showed a nearly perfect linear correlation between the pupillary dilation response and loudness. The overall results suggest that the pupillary dilation response reflects the subjective salience of sounds, which is defined, or is heavily influenced, by loudness.

Keywords Salience · Loudness · Audition · Pupil · Rating · Autonomic nervous system · Endocrine system

Pupillary responses are known to reflect not only ambient light (the pupillary light reflex; Ellis, 1981) but also mental states, such as emotional arousal (Bradley, Miccoli, Escrig, & Lang, 2008; Bradshaw, 1967), and cognitive functions, including memory (Goldinger & Papesh, 2012; Naber, Frassle, Rutishauser, & Einhäuser, 2013), decision making (Einhäuser, Koch, & Carter, 2010; Einhäuser, Stout, Koch, & Carter, 2008; Lavin, Martin, & Jubal, 2014; Preuschoff, 't Hart, & Einhäuser, 2011), and attention (Eldar, Cohen, & Niv, 2013; Gabay, Pertzov, & Henik, 2011; Privitera, Renninger, Carney, Klein, & Aguilar, 2010). Previous studies have suggested that the cognitive-evoked pupillary responses are due to modulation of the locus coeruleus–norepinephrine (LC-NE) system. The LC-NE system receives direct inputs from the anterior cingulate and orbitofrontal cortices to achieve optimal performance in a task, and the effect of modulation is reflected in the pupillary responses (Aston-Jones & Cohen, 2005; Sara, 2009). Related studies have also shown that neural activation in the superior colliculus evokes a pupillary response to salient audiovisual stimuli (Netser, Ohayon, & Gutfreund, 2010; Wang, Boehnke, Itti, & Munoz, 2014; Wang, Boehnke, White, & Munoz, 2012). Taken together, these findings indicate that pupillary responses are mediated by the endocrine system, modulated by neural activations in the midbrain, and

✉ Hsin-I Liao
liao.hsini@lab.ntt.co.jp

¹ Human Information Science Laboratory, NTT Communication Science Laboratories, NTT Corporation, Kanagawa, Japan

² Department of Information Processing, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Tokyo, Japan

that they reflect cognitive functions related to attention and salient stimulus detection.

Indeed, it has been shown that pupillary dilation responses can be evoked by salient auditory or visual stimuli of different species. Bala and Takahashi (2000) showed that the pupillary dilation response in the barn owl could be evoked by deviant auditory stimuli, which are characterized by their frequency and location differences from background sounds. Wang and colleagues (Wang et al., 2014; Wang & Munoz, 2014) showed that transient pupillary dilation responses in monkeys and humans could be evoked by salient audiovisual stimuli, which are characterized by contrast differences in the visual stimuli and intensity differences in the auditory stimuli. Our previous study (Liao, Yoneya, Kidani, Kashino, & Furukawa, 2014) showed that in humans, the pupil dilates in response to deviant auditory oddballs embedded in an auditory stream, which are characterized by spectrum (pure-tone vs. noise) and frequency (pure-tone) differences.

These studies suggest that the pupillary dilation response signals salient auditory stimulus detection, objectively characterized by contrast in the stimulus characteristics (e.g., spectrum or intensity) or by context (e.g., the deviant oddballs). However, whether there is a direct link between pupillary dilation responses and subjectively defined psychological salience remains unclear. It has been established that saccadic allocation is a valid indicator of visual salience detection (Itti & Koch, 2001), whereas, until now, a physiological marker has been lacking for auditory salience detection. In the present study, we tried to determine whether the pupillary dilation response reflects subjective salience, defined as being “easy to notice,” and can thus be used as a physiological index of auditory salience.

Experiment 1

We used various environmental sounds to investigate the relationship between pupillary response and the subjective salience of sounds. The sounds included synthesized sounds, such as tone, noise, chirps, and beeps, and environmental sounds, such as a dog barking, a bird singing, a phone ringing, nails scratching, human crying, and human laughing. Participants listened to the sounds passively while their pupillary responses were recorded. Later on, they made subjective salience judgments of the sounds. We examined whether pupillary responses to sounds can be predicted by subjective salience judgments of them.

Method

Participants Eight adults (four males, four females; age 21–43, median = 33 years old) participated in the present experiment, but one was excluded because he was unable to

participate in the pairwise-comparison experiment conducted on another day. None had been diagnosed with hearing disorders, and all had normal or corrected-to-normal vision. The sample size was predefined as a fixed number of participants being recruited. No replacement was applied. The present study was approved by the NTT Communication Science Laboratories Ethical Committee. All participants gave informed written consent before the experiment and received payment for their participation.

Materials Separate configurations were set up for the pupillary response recording and pairwise-comparison task. In the pupillary-response recording experiment, all of the auditory and visual stimuli were controlled by a personal computer (Dell OptiPlex 755). The auditory stimuli were generated digitally by the personal computer, transformed by an audio interface (Roland OCTA-CAPTURE), sent through a headphone amplifier (Grace design m903), and presented through headphones (Sennheiser HD 595). The visual stimuli were presented on an 18.1-in. monitor (Eizo FlexScan L685Ex) with a frame rate of 60 Hz and resolution of $1,024 \times 768$ pixels. Throughout the whole pupillary-response recording session, a dark gray central fixation point ($0.25 \times 0.25^\circ$, 0.33 cd/m^2) was presented against a light gray background (12.9 cd/m^2). The participants’ pupillary responses were recorded binocularly with an infrared eyetracker camera (Eyelink 1000 Desktop Mount, SR Research Ltd.). The camera was positioned below the monitor, and the sampling rate of the recording was 1000 Hz.

In the pairwise-comparison experiment, all auditory and visual stimuli were controlled by another personal computer (Dell OptiPlex 980). The auditory stimuli were sent through a digital-to-analog converter (Tucker-Davis Technologies, Inc. System III RM1) connected to the headphone amplifier (Grace design m903) and presented on headphones (Sennheiser HD 595). The visual stimuli were presented on a 19-in. monitor (EizoFlexScan L761T-C). A panel shown on the screen enabled the participant to manipulate the auditory stimuli, display the questionnaire, and indicate the chosen responses. The panel had four buttons arranged in a square. Each button was 2.8 cm (W) \times 1.4 cm (H) and were separated from each other by 7.3 cm horizontally and 4.3 cm vertically. The top two buttons were used to play the sound pairs and were, respectively, labeled “play sound A” and “play sound B.” The bottom buttons were used to indicate the choice: “sound A” or “sound B.” The following question was shown in the center of the panel: “Which sound is more salient?”, and “A” and “B” were used to represent the abstract indices of the paired sounds, to avoid any semantic influence from the real sound names.

Ten auditory stimuli were used in both setups (Fig. 1). Four were generated by the personal computer: bursts of a 1000-Hz tone, white noise, chirps, and beeps. Five were selected from a database (Audio pro sound effects

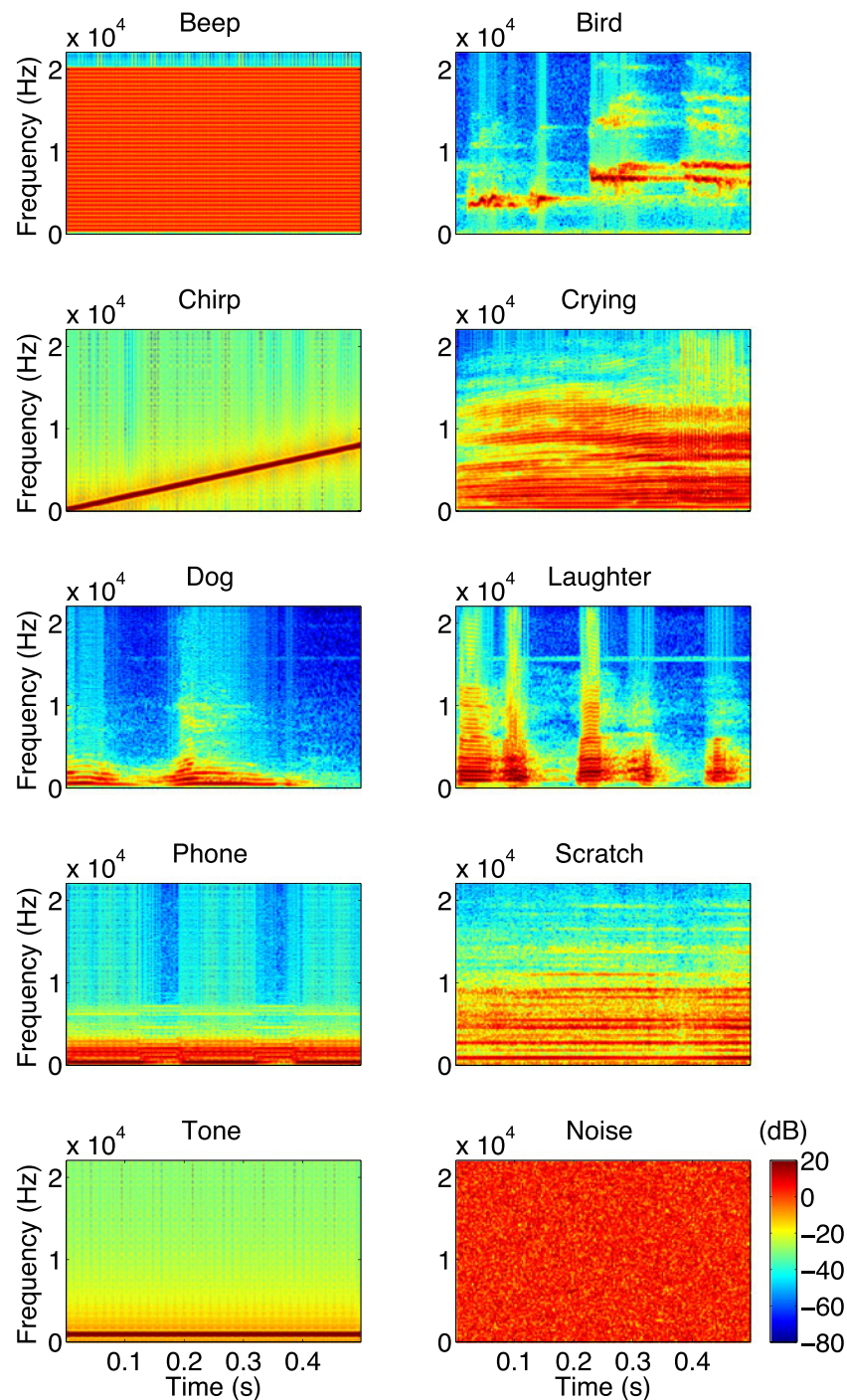


Fig. 1 Spectrograms of the auditory stimuli. The colors represent the spectrum levels in decibels relative to the maximum value of that stimulus, as is indicated by the color bar.

by Yannick Chevalier): a dog barking, a bird singing, a phone ringing, a human crying, and a human laughing. One was a self-recorded sound of nails scratching over a blackboard. The interstimulus interval was 10 s, and all sounds were presented diotically for 500 ms and had an A-weighted sound pressure level of 65 dB.

These ten sounds may be classified into several categories in terms various attributes, such as their

acoustical properties, emotional valence, and source. We used only ten sounds to represent the sounds in different categories. It should be noted, however, that it was not our intention to compare pupillary responses among the stimulus categories. Our major interest was the relationship between the behavioral and pupillary responses induced by the sounds, not that between stimulus categories and pupillary responses.

Procedure In the pupillary-response recording experiment, participants sat in front of the monitor at a viewing distance of 80 cm in a dark chamber, with their head fixed on a chinrest. They were not involved in any task, and just listened to an auditory sequence while their pupillary responses were recorded. The auditory sequence consisted of the ten types of auditory stimuli, in which each stimulus was presented ten times. These 100 sounds were presented in random order. The interstimulus interval was 10 s. During the recording, participants were instructed to fixate on the fixation point to avoid gaze-position-dependent pupil size artifacts. The standard five-point calibration procedure of the EyeLink system was conducted prior to the auditory sequence, and a 30-s interval was interposed between the calibration and the auditory sequence presentation. This was done to ensure that the pupillary response was not influenced by the flash calibration procedure, and to allow it to adapt to the constant illuminant environment. The whole pupillary-response recording session took about 20 min.

After the pupillary-response recording, in seven separate sessions, participants were requested to perform subjective pairwise-comparison tasks on salience, loudness, vigorousness, preference, beauty, annoyance,¹ and hardness. The instructions were given in Japanese as follows: 目立つ (noticeable, remarkable, standing out), 大きい (loud), 迫力がある (power, punch), 好き (liking), 綺麗 (beautiful), 不快 (unpleasant, discomfort), and 固い (hard, tough), respectively. The order of the sessions was randomly assigned. Each session consisted of 45 trials, in which all of the possible combinations of pairs chosen from the ten test sounds were used. The paired sounds were assigned to “sound A” and “sound B” in a random order. In each trial, participants were asked to choose the stimulus that sounded stronger regarding the aspect being tested—for example, more salient in the salience session. Participants were allowed to listen to the paired sounds as long as they wanted until they had made a decision. They were able to replay the sounds and to indicate their decisions by clicking the corresponding buttons using the computer’s mouse. After the response button had been clicked, the next trial started. The pupillary response recording was performed separately from the pairwise-comparison tasks to avoid any confounding pupillary response owing to the task involvement.

Results

The pupil diameter data acquired during blinks or breaks in fixation were treated as missing. Only the data of the right eye were analyzed, since pupil responses in both eyes have been confirmed to be consensual (Purves et al., 2004). The EyeLink

system output pupil diameter data in arbitrary units [au] that were not comparable across conditions. To normalize the pupillary responses to sounds across conditions, the raw pupil diameter data were subtracted from the mean of the data during the 1-s period before each trial. The subtractive value differed across all trials for all participants and all sound types. The pupil diameter data during the reference period (i.e., 1 s before the sound onset) had a symmetrical long-tailed distribution and did not fit the normal distribution (Kolmogorov–Smirnov goodness-of-fit test, $p < .001$). To avoid undesirable contributions from outliers, we calculated the median, instead of the mean, to represent the change in the pupil diameter over time for individual participants. The data were then averaged with a moving window of 250 ms to remove high-frequency noise.

The means of the pupil diameters across participants are shown in Fig. 2A. It has been shown that the pupillary response to sounds consists of multiple components that differ in latency (e.g., Bala & Takahashi, 2000; Wang et al., 2012; Wang & Munoz, 2014). In order to examine how the pupil responds to sounds and changes over time, and whether these responses interact with sound type, the data were subjected to a two-way repeated measures analysis of variance (ANOVA) with Sound Type (tone, noise, chirp, beep, dog, bird, phone, crying, laughter, and scratch) and Time Delay (centered at 0.125, 0.375, 0.625, 0.875, 1.125, 1.375, 1.625, 1.875, 2.125, 2.375, 2.625, 2.875, 3.125, 3.375, 3.625, and 3.875 s) as the within-subject factors. The results showed a main effect of time delay [$F(15, 90) = 4.64, p < .001, \eta_p^2 = .44$], but not a main effect of sound type [$F(9, 54) = 0.91, p = .52, \eta_p^2 = .13$] or an interaction [$F(135, 810) = 1.07, p = .30, \eta_p^2 = .15$]. A trend analysis showed that both the linear and quadratic trends were significant [$F_s(1, 6) = 7.27$ and $8.89, p_s = .04$ and $.03, \eta_p^2_s = .55$ and $.60$, respectively]. Overall, the results indicated that the pupil diameter increased gradually as the sound was presented, regardless of the sound type. This pupillary dilation response to sound was most effective around 1–2 s after the sound presentation.

To evaluate whether sound affects blinks, we performed an analysis of blink rates as a function of time from stimulus onset (Fig. 2B). The data were subjected to a two-way repeated measures ANOVA with Sound Type and Time Delay as the within-subject factors. The results showed that none of the effect was significant ($p_s > .05$). The results not only indicated that participants did fixate on the central fixation points throughout the experiment, even though no task was required, but also suggested that the effect of the pupil dilation response to sounds was not related to blinks.

Two indices were derived in order to examine the relationship between the pupillary dilation response to sounds and the subjective judgments about them. First, the pupil diameter data collected for a period of 1–2 s after the sound presentation were averaged in order to

¹ Annoyance corresponds to “noisiness” in the terminology used in the field of acoustics (Berglund, Berglund, & Lindvall, 1975). The judgment was based on the quality of sound, independent of the context of a given situation.

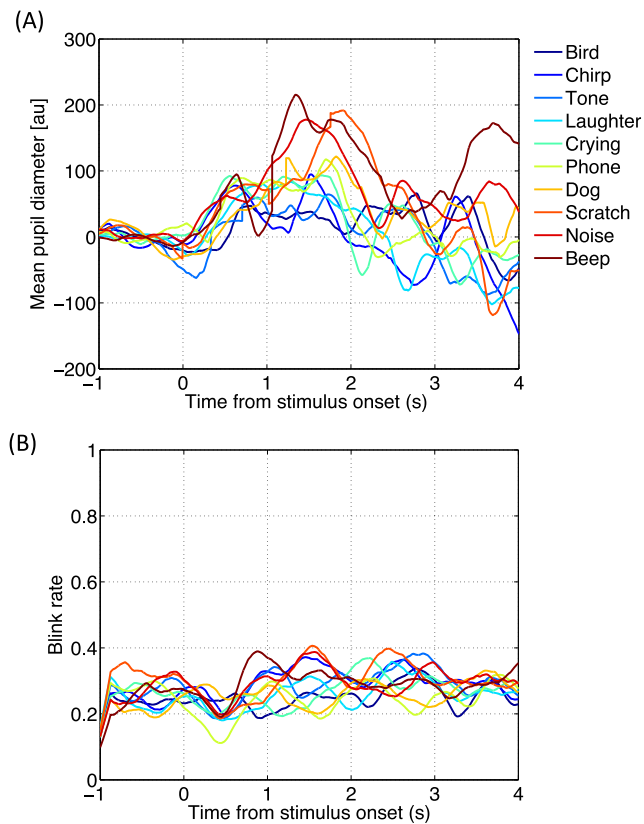


Fig. 2 Mean pupil diameter (A) and blink rate (B) changes as a function of the time relative to sound onset in Experiment 1.

represent the pupillary dilation response to sound. Second, the pairwise-comparison data were converted to the Thurstone scale. For each aspect of the sounds, the pairwise-comparison data were counted, and matrices of the probabilities of all pairs of sounds to be chosen were calculated. The probabilities were converted to the inverse of the standard normal cumulative distribution. For each sound, the inverses were averaged across its comparisons with all of the other sounds to represent the rating numbers of the Thurstone scale. A positive value indicated more salient, louder, more vigorous, preferable, more beautiful, more annoying, or harder. Sounds that evoked larger pupillary dilation responses were rated as being more salient, as well as louder, more vigorous, and more annoying, but less preferred and less likely to be judged as beautiful (see Fig. 3). Table 1 lists the Pearson correlation coefficients, with hypothesis-testing results among all of the variables.

In addition to the group analysis, we conducted an analysis based on individual participants to examine whether the correlations between the pupillary dilation responses and subjective ratings held true individually. This analysis was necessary for the following reasons. First, although there was significant correlation between the pupillary dilation response and the salience scale, the analysis was based on the data averaged across participants. It can be argued that the pupillary dilation

response just reflects the acoustical characteristics inherent to the stimuli, which might determine the subjective salience to the “ideal” listener, but not the individual participant’s judgment per se. Second, in the ANOVA wherein the participant factor was treated as random noise, the pupil size in fact did not significantly differ across stimulus types. However, the effect of the sound-induced pupillary dilation response may still have corresponded to the subjective ratings of salience for individual participants. In other words, different participants might have different feeling about the sounds, which might be reflected in the individual participant’s pupillary responses. The individual-based analysis would therefore help us to clarify whether we can predict people’s subjective judgments by examining their pupillary responses.

For each participant and each stimulus pair in the pairwise-comparison task, we compared the sizes of the pupillary dilation responses (i.e., the mean pupil diameter over the period of 1–2 s) to the corresponding stimuli. If the pupillary dilation response reflected a subjective judgment of the paired stimuli, we would expect to observe a tendency for the pupillary dilation response to be consistently larger (or smaller) for the chosen than for the unchosen stimulus. In other words, the odds ratio of the larger pupillary dilation response for the chosen stimulus was expected to be larger (or smaller) than 1—that is, the point of no relation between the pupil diameter and the choice. The results are shown in Table 2. As expected, there were more trials in which the participants’ pupil diameters were larger for the chosen *salient* stimuli; namely, all of the odds ratios were larger than 1 (sign test $\chi^2 = 7, p = .008$). The same tendency was found for the *loudness* judgments ($\chi^2 = 7, p = .008$), but not for the other judgments ($\chi^2 < 2, ps > .250$). Thus, the pupillary dilation response predicted the salience and loudness judgments of the paired sounds.

Discussion

The pupil dilated to auditory stimulus presentation, regardless of sound type. The pupillary dilation response to sound was most effective during 1–2 s after stimulus onset (sound presentation duration was 500 ms). Most importantly, this pupillary dilation response not only correlated with the subjective salience of the sounds, but also with their subjective loudness, a finding that is backed up by the effect of sound (i.e., the correlation on the Thurstone scale) and by the individual pairwise analysis. Together with the finding that salience and loudness are highly correlated, the overall results suggest that the salience and loudness judgments may share a common mechanism related to pupillary responses. Although salience was also correlated with all of the other psychological judgments on the Thurstone scale, the prediction of a pupillary dilation response to the subjective judgment did not hold true for all the other

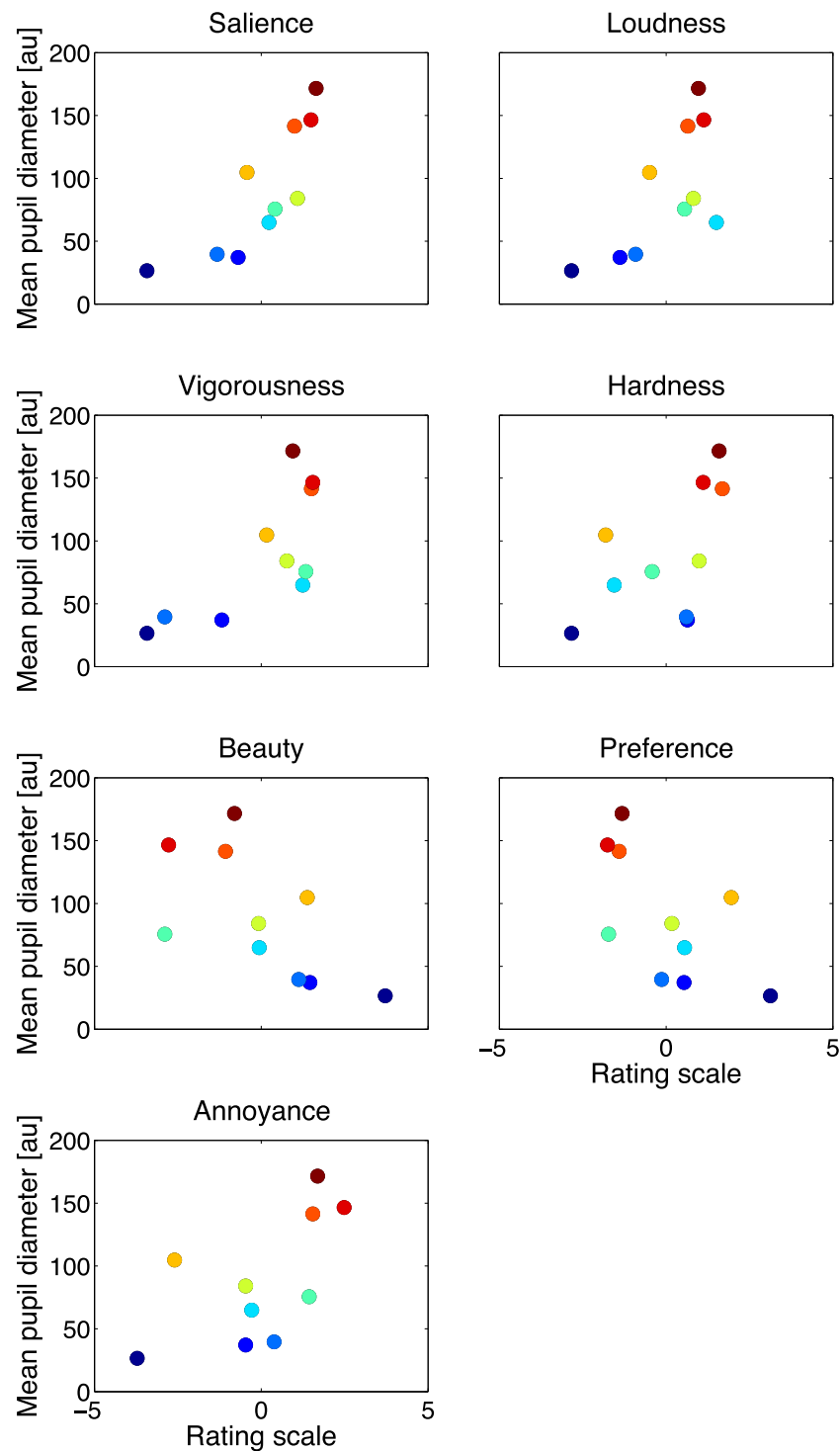


Fig. 3 Scatter plots of mean pupil diameters, averaged across the period 1–2 s after stimulus onset, as a function of the Thurstone rating scale derived from the pairwise comparison judgments in Experiment 1. Each

dot represents each sound. The colors are labeled in correspondence to Fig. 2.

psychological judgments in the individual-based analysis. The correlations may come from different sources, which could derive from the cognitive evaluation or emotional aesthetic pleasure of the shared acoustic features or timbre elements.

Experiment 2

It remains unclear whether pupillary dilation responses reflect salience or loudness, or both. Furthermore, it is possible that salience is more-or-less equivalent to loudness in the current

Table 1 Pearson's correlation coefficients with hypothesis testing among all of the ratings in the Thurstone scale and mean pupil diameter in Experiment 1

	Salience	Loudness	Vigorousness	Preference	Beauty	Annoyance	Hardness
Loudness	.92**						
Vigorousness	.90**	.91**					
Preference	-.81**	-.72*	-.67*				
Beauty	-.84**	-.82**	-.83**	.92**			
Annoyance	.79*	.72*	.63†	-.99**	-.88**		
Hardness	.72*	.50	.42	-.80*	-.57	.81**	
Pupil diameter	.81**	.67*	.73*	-.62†	-.63*	.61†	.57

* Statistical significance of Pearson's coefficient at $p < .05$. ** Statistical significance of Pearson's coefficient at $p < .01$. † Marginally significant at $p = .06$

context, in which salience is defined as being “easy be notice” and may inherently covary with loudness. In Experiment 2, we examined how the pupil responds to sounds when loudness is dissociated from salience, and how the relationship between salience and loudness would be subjectively judged.

We used the same ten sounds as in Experiment 1, except that we adjusted the sound pressure level of each sound so that all of the sounds would have the same loudness on the basis of the loudness model developed by Glasberg and Moore (2002). If the loudness manipulation did not change the salience judgments, the effect of pupillary dilation response on salience could be dissociated. However, if participants instead interpreted salience to be more or less equivalent to loudness, there would still be a correlation between subjective judgments of salience and loudness. Pupillary responses were recorded in order to examine how the pupillary dilation response is related to sounds with equal loudness.

Method

Participants A group of 12 adults (eight females, four males; age 22–43, median = 38 years old) participated in the present experiment. In a hearing test from 125 to 8000 Hz, all participants showed normal hearing at the test frequencies of 125, 250, 500, 100, 2000, and 4000 Hz (threshold below 25 dB). For the 8000-Hz test tone, two participants showed hearing

loss (one's threshold was 30 dB in both ears, and the other's threshold exceeded 60 dB in the right ear). All of the other participants showed normal hearing of the 8000-Hz test tone.

Materials and procedure The stimuli, apparatus, and procedure were the same as in Experiment 1, except for the following. The loudness as a function of the stimulus duration of the ten sounds was estimated by analyzing the waveform of the stimulus on the basis of the loudness model developed by Glasberg and Moore (2002). Since the stimuli were complex sounds, we chose the maximum of the long-term loudness to represent the overall loudness impression of each sound. We then adjusted the sound pressure level of each sound so that the estimated maximum of the long-term loudness was 65 phons.

Results

The correlations among all of the pairs of subjective judgments on the Thurstone scale are shown in Table 3. Despite the equivalent loudnesses of the sounds, the subjective judgments of salience still significantly correlated with the subjective judgments of loudness, as well as with those of vigorousness. Salience was *positively* correlated with beauty, whereas a negative correlation had been found between salience and

Table 2 Odds ratios of larger mean pupil diameter for the chosen stimulus in Experiment 1

	Salience	Loudness	Vigorousness	Preference	Beauty	Annoyance	Hardness
CK	1.25	1.05	0.50	1.50	1.65	0.61	0.50
MO	1.50	2.46	2.00	0.88	0.80	1.50	1.50
HM	1.25	1.65	0.88	0.61	0.88	1.65	1.81
YY	2.00	2.46	2.21	0.29	0.55	2.00	1.37
MY	1.14	1.25	1.50	0.45	0.41	3.09	1.25
TK	1.81	1.81	1.37	1.65	0.73	0.61	0.73
YT	1.81	1.50	1.81	0.61	0.41	1.37	1.37

Table 3 Pearson's correlation coefficients with hypothesis testing among all of the ratings in the Thurstone scale and mean pupil diameter in Experiment 2

	Salience	Loudness	Vigorousness	Preference	Beauty	Annoyance	Hardness
Loudness	.83**						
Vigorousness	.82**		.94	**			
Preference	.14	-.12	-.31				
Beauty	.73*	.47	.40	.53			
Annoyance	.28	.49	.64*	-.84**	-.34		
Hardness	.22	.33	.43	-.79*	-.13	.76*	
Pupil diameter	.02	-.21	-.25	.57	.24	-.53	-.68*

* Statistical significance of Pearson's coefficient at $p < .05$. ** Statistical significance of Pearson's coefficient at $p < .01$.

beauty in Experiment 1. In addition, there were in general fewer significant correlations among all comparisons, probably due to less discriminability among the sounds on the Thurstone scale than in Experiment 1. Indeed, the mean of the relative distances on the Thurstone scale among all sounds was smaller in Experiment 2 than in Experiment 1 for all subjective judgments [$t_s > 2$, $p_s < .02$], except for the judgments of salience [$t(44) = 1.46$, $p = .08$] and hardness [$t(44) = 0.47$, $p = .32$].

The means of pupil diameter and blink rate as a function of time are shown in Fig. 4. The data were subjected to a two-way repeated measures ANOVA with Sound Type and Time Delay as within-subject factors, as in Experiment 1. For pupil diameter changes, the main effect of time delay was significant [$F(15, 165) = 3.71$, $p < .001$, $\eta_p^2 = .25$], but not the main effect of sound type [$F(9, 99) = 0.46$, $p = .90$, $\eta_p^2 = .04$] or the interaction [$F(135, 1485) = 0.67$, $p > .99$, $\eta_p^2 = .06$]. A trend analysis showed that both the quadratic and cubic trends were significant [$F_s(1, 11) = 9.36$ and 7.64 , $p_s = .01$ and $.02$, $\eta_p^2_s = .46$ and $.41$, respectively], but not the linear trend [$F(1, 11) = 1.45$, $p = .25$, $\eta_p^2 = .12$]. The overall results for pupillary responses were similar to those in Experiment 1: The pupil dilated to sounds around 1–2 s after stimulus onset, regardless of sound type.

For the blink rate, the main effect of time delay and the interaction [$F(15, 165) = 6.74$, $p < .001$, $\eta_p^2 = .38$, and $F(135, 1485) = 1.44$, $p = .001$, $\eta_p^2 = .12$, respectively] were both significant, but the main effect of sound type was not [$F(9, 99) = 0.93$, $p = .51$, $\eta_p^2 = .08$]. The results indicate that the blink rate changes over time relative to stimulus onset, and that the fluctuation depends on the sound type.

To examine the relationship between pupillary dilation responses and the subjective judgments of sounds, we performed a correlation analysis based on the Thurstone scale and an individual-based analysis, as in Experiment 1. The results of the correlation between pupillary dilation response and subjective judgment on the Thurstone scale are shown in Fig. 5. Pupillary dilation responses did not correlate with salience or loudness, or with the other subjective judgments, except for a negative correlation with hardness. The finding

was the same as in the individual-based analysis (see Table 4): Participants showed a consistent tendency in the hardness judgment ($\chi^2 = 12$, $p < .001$), but not in the other judgments ($\chi^2_s < 3$, $p_s > .08$).

Discussion

With equivalence of the loudness of sounds on the basis of the loudness model (Glasberg & Moore, 2002), the subjective

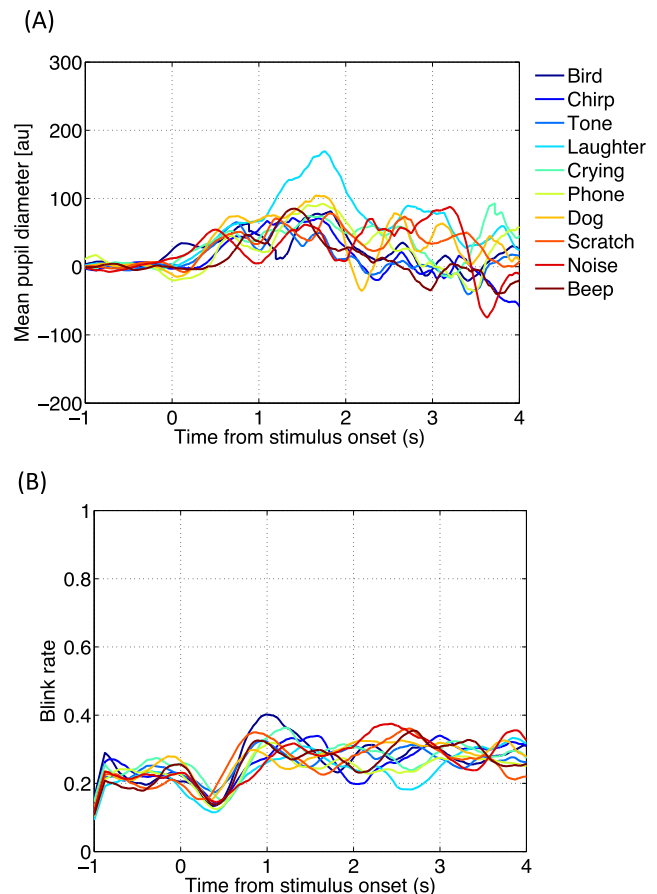


Fig. 4 Mean pupil diameter (A) and blink rate (B) changes as a function of the time relative to sound onset in Experiment 2.

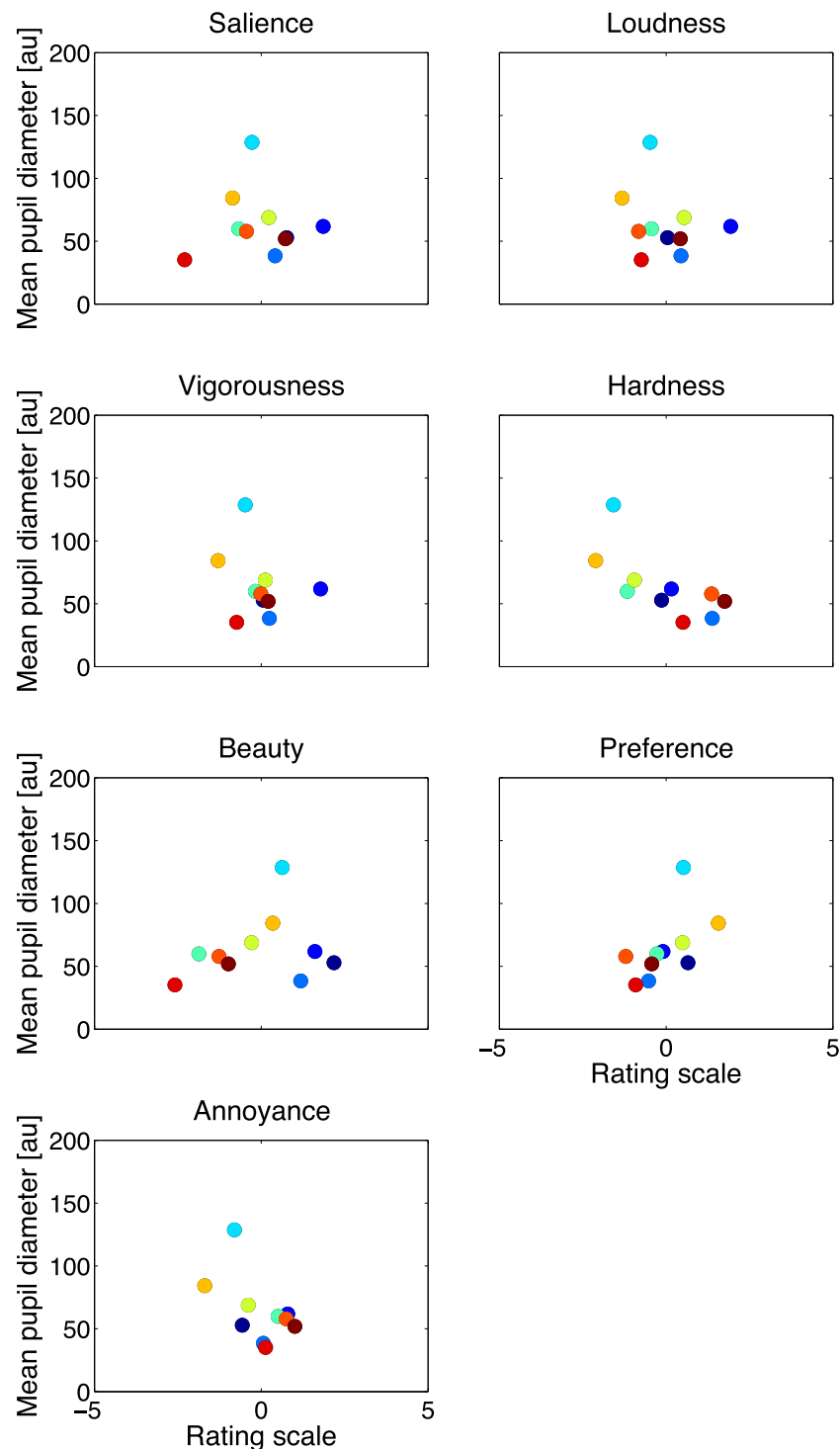


Fig. 5 Scatter plots of mean pupil diameters, averaged across the period 1–2 s after stimulus onset as a function of the Thurstone rating scale derived from the pairwise comparison judgments in Experiment 2. Each

dot represents each sound. The colors are labeled in correspondence to Fig. 4.

judgment of salience still positively correlated with the subjective judgment of loudness, as well as with that of vigorousness. This result suggests that salience is more or less equivalent to loudness or vigorousness in the current context. Salience is defined as being “easy to notice,” and it is thus

inherent that loudness or vigorousness (having power) is the major factor determining salience, and the reverse may also be true.

The pupil responded to each equivalent-loudness sound similarly, and the responses did not correlate with salience or

Table 4 Odds ratios of larger mean pupil diameter for the chosen stimulus in Experiment 2

	Saliency	Loudness	Vigorousness	Preference	Beauty	Annoyance	Hardness
IT	1.81	1.64	2.00	0.88	0.96	1.05	0.88
MM	2.00	1.64	1.65	0.96	1.25	1.50	0.96
MM	1.05	1.14	0.96	1.05	1.81	1.14	0.80
WY	0.55	0.61	0.96	0.67	0.55	1.37	0.80
SM	0.67	0.96	0.96	0.36	0.29	2.00	0.80
SK	1.25	1.50	1.37	0.73	1.65	1.14	0.96
TS	2.00	1.25	1.25	0.96	1.05	0.88	0.73
TK	0.96	0.50	0.88	1.25	1.37	0.61	0.73
NN	1.14	1.25	0.80	0.67	0.88	1.25	0.67
IK	2.75	2.46	2.21	1.64	1.81	0.41	0.96
OR	0.55	0.50	0.50	0.96	0.67	1.50	0.80
YA	0.73	0.50	0.96	0.29	0.36	4.00	0.67

with the loudness rating. Together with the finding in Experiment 1 that the pupil dilated more to the sounds evaluated as being more salient and yet louder, the overall results provide converging evidence of the correspondence between the pupillary dilation responses and subjective judgments of the saliency, as well as the loudness, of sounds.

Two inconsistent and yet intriguing results were found, as compared to Experiment 1. First, subjective judgments of saliency were positively correlated with beauty, but they had been negatively correlated with beauty in Experiment 1. Second, the pupillary dilation response was negatively correlated with subjective judgments of hardness, but no such correlation was found in Experiment 1. The inconsistency could be due to an interaction between loudness and sound perception regarding the psychological acoustic evaluation. Alternatively, it could be due to a determinant role of loudness in acoustic evaluation. In other words, loudness could be one of the major features for psychological acoustic evaluation. When variant loudness is embedded in sounds, as in Experiment 1, the evaluation of a sound, such as its beauty, is heavily influenced by loudness. In contrast, when the loudnesses of different sounds are equivalent, and loudness therefore no longer provides one of the major features of sounds, other acoustic features, such as hardness, emerge to influence the evaluation. The argument above of the major role of loudness in acoustic features is supported by the results that sounds were less discriminable in general when their loudness was the same.

Experiment 3

In the previous two experiments, we found that the pupillary dilation responses reflected the subjective saliency, as well as the loudness, of sounds. In Experiment 1, when the sounds were presented with the sound pressure level constant, the

pupillary dilation response was positively correlated with the subjective salient judgments of sounds and with loudness. In Experiment 2, when the sounds were equally loud, the pupillary dilation responses were similar to each other and did not correlate with saliency or loudness. In both experiments, saliency and loudness were highly correlated. The overall results suggest that the subjective judgment of saliency is more or less equivalent to loudness, and that the pupillary dilation response reflects both.

In Experiment 3, we examined the effect of saliency on the pupillary dilation response by manipulating loudness. We manipulated the sound pressure level (45, 55, 65, 75, and 85 dB) and acoustic characteristics (1000-Hz pure tones and white noise) to adjust the loudness. The larger the sound pressure level was, the louder the sound became. According to the loudness model (Glasberg & Moore, 2002) and the results of Experiment 1, white noise is louder than a 1000-Hz pure tone when the sound pressure level is constant. It was thus expected that the pupillary dilation response would be greater for sounds with larger sound pressure levels, and greater for white noise than for the 1000-Hz pure tone.

Method

Participants Another group of eight adults (four females, four males; age 21–42, median = 37.5 years old) participated in the present experiment.

Materials and procedure The stimuli, apparatus, and procedure were the same as in Experiment 1, except for the following. First, only bursts of a 1000-Hz tone and white noise were used. Second, the A-weighted sound pressure levels of the tone and the noise were varied to be 45, 55, 65, 75, and 85 dB. Third, only the procedure with pupillary response recording was conducted; no subjective judgments were required after the pupillary response recording.

Results

The means of pupil diameter and blink rate as a function of time are shown in Figs. 6A and 6B, respectively. The data were subjected to a three-way repeated measures ANOVA with Sound Type (tone, noise), Intensity (45, 55, 65, 75, and 85), and Time Delay (centered at 0.125, 0.375, 0.625, 0.875, 1.125, 1.375, 1.625, 1.875, 2.125, 2.375, 2.625, 2.875, 3.125, 3.375, 3.625, and 3.875 s, as in the previous two experiments) as within-subject factors. For pupil diameter changes, the main effects of sound type, intensity, and time delay were all significant [$F(1, 7) = 5.34, p = .05, \eta_p^2 = .43$; $F(4, 28) = 4.18, p = .01, \eta_p^2 = .37$; and $F(15, 105) = 5.91, p < .001, \eta_p^2 = .46$, respectively], as was the two-way interaction of Sound Type \times Time Delay [$F(15, 105) = 1.91, p = .03, \eta_p^2 = .22$]. The two-way Intensity \times Time Delay interaction was marginally significant [$F(60, 420) = 1.34, p = .056, \eta_p^2 = .16$]. All of the other interactions were not significant ($F_s < 0.7, p_s > .6$). A trend analysis of time delay showed that both the quadratic and cubic trends were significant [$F_s(1, 7) = 20.76$ and $27.21, p_s = .003$ and $.001, \eta_p^2_s = .75$ and $.80$, respectively], but not the linear trend [$F(1, 7) = 0.60, p = .46, \eta_p^2 = .08$]. The general profile of the pupillary responses to sounds showed a consistent pattern, as had been found in the previous experiments—that is, the pupils dilated to sounds around 1–2 s after stimulus onset. Unlike in the previous experiments, the sound type and intensity showed significant effects.

For the blink rate, a three-way repeated measures ANOVA showed a main effect of time delay [$F(15, 105) = 4.88, p < .001, \eta_p^2 = .41$], but not other main effects or interactions ($F_s < 3, p_s > .1$). A trend analysis on time delay showed significant linear and quadratic trends [$F(1, 7) = 7.08, p = .03, \eta_p^2 = .50$, and $F(1, 7) = 14.40, p = .007, \eta_p^2 = .67$, respectively] and the marginal significance of the cubic trend [$F(1, 7) = 5.07, p = .059, \eta_p^2 = .42$]. The overall results indicate that the blink rate changes over time relative to stimulus onset, regardless of sound type or intensity. The effects of sound type and intensity on the pupillary dilation response cannot be explained by the blinks.

To examine the effects of sound type and intensity, the pupillary dilation responses to sound—that is, the average of the pupil diameter data collected during 1–2 s after the sound presentation—were subjected to a two-way repeated measures ANOVA with Sound Type and Intensity as within-subject factors. The results are shown in Fig. 6C. Both the main effects of sound type and intensity were significant [$F(1, 7) = 6.93, p = .03, \eta_p^2 = .50$, and $F(4, 28) = 6.73, p = .001, \eta_p^2 = .49$, respectively], but the two-way interaction was not [$F(4, 28) = 0.68, p = .61, \eta_p^2 = .09$]. A trend analysis on intensity showed a strong linear relationship between sound intensity and the pupillary dilation response [$F(1, 7) = 20.36, p = .003, \eta_p^2 = .74$].

To examine the effect of loudness on the pupillary dilation responses across stimulus types and experiments, we

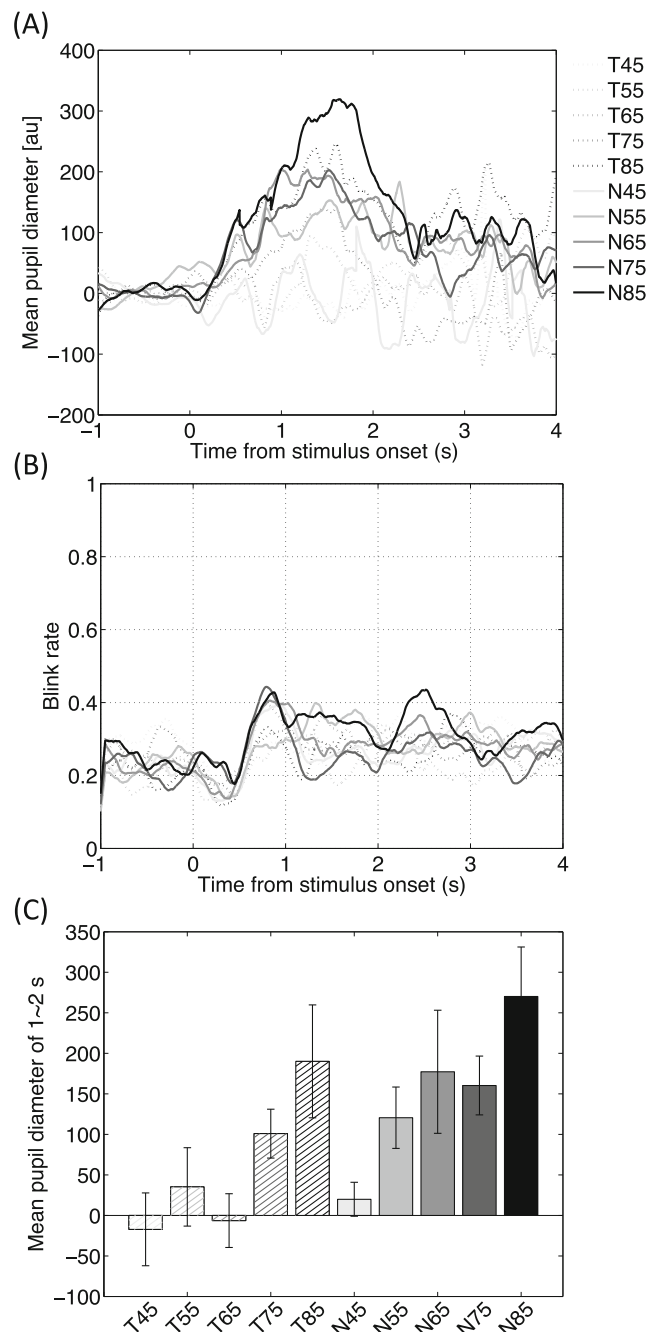


Fig. 6 Mean pupil diameter (A) and blink rate (B) changes as a function of the time relative to sound onset in Experiment 3. (C) Mean pupil diameters averaged across the period 1–2 s after stimulus onset as a function of the sound type. T represents “tone,” and N represents “noise.” The numbers following “T” and “N” represent the sound pressure level.

conducted a regression analysis to predict the pupillary dilation responses by loudness, estimated by the loudness model of Glasberg and Moore (2002). The results are shown in Fig. 7. The regression coefficients were 4.87 and 5.50 for the environmental stimuli used in Experiment 1 and for the tone and noise stimuli with various intensities used in Experiment 3, respectively. The regression coefficients did

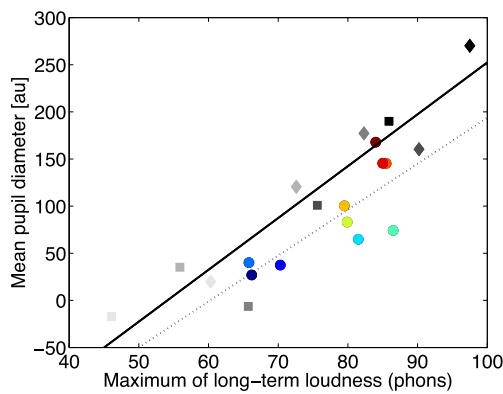


Fig. 7 Scatter plots of mean pupil diameters, averaged across the period 1–2 s after stimulus onset (i.e., the pupillary dilation response) as a function of loudness. The color dots represent the sounds used in Experiment 1, and the colors are labeled in correspondence to Fig. 2. The squares and diamonds represent the tones and noises, respectively, with different sound pressure levels in Experiment 3. The darker the symbol, the greater the sound pressure level. The dashed and solid lines represent the regression lines of loudness to the pupillary dilation responses to the sounds used in Experiments 1 and 3, respectively.

not differ from each other [$t(16) = 0.39, p = .35$]. This suggests that the pupillary dilation response can be predicted by the loudness, regardless of whether the auditory stimuli are various environmental sounds or synthesized sounds with different intensities.

Discussion

The pupillary dilation response was greater for the sounds with stronger intensity and greater for bursts of white noise than for 1000-Hz pure tones. The amplitude of the pupillary dilation response was linearly predicted well by loudness estimated on the basis of the loudness model of Glasberg and Moore (2002). The results indicate that loudness, which depends on the physical power (sound pressure level) and acoustic characteristics of stimuli, affects the pupillary dilation response.

General discussion

In three experiments, we found that the pupil dilated in response to auditory stimulus presentations without task involvement. The dilation peaked around 1–2 s after stimulus onset, regardless of stimulus type. Although the peak of the pupillary dilation responses may differ depending on the duration of stimuli or on species (see Bala & Takahashi, 2000; Wang et al., 2014; Wang & Munoz, 2014), our finding of the pupillary dilation responses to auditory stimuli were consistent with previous studies showing that pupillary dilation is a response to auditory deviant oddballs (Bala & Takahashi, 2000) and contrast-based salient audiovisual stimuli (Wang et al., 2014; Wang & Munoz, 2014). In contrast, when visual stimuli are presented for a long time and a cognitive task—for example,

memory—is required, the pupil contracts, rather than dilates, in response to the visual presentation (Naber et al., 2013; Naber & Nakayama, 2013). These opposing directions of pupillary responses suggest the existence of multiple perceptual- and/or task-related mechanisms.

Importantly, the pupillary dilation response to an auditory stimulus was found to be positively correlated with the subjective salience, as well as the loudness, of the sounds (Exp. 1). When the loudness was the same, the pupil responded to each sound similarly, and the response did not correlate with the subjective judgment of salience or loudness (Exp. 2). The findings were backed up by an analysis based on the Thurstone scale derived from pooled participants, as well as by a pairwise analysis on the individual participants. Note that in our procedure, the pupillary response recording was conducted separately from the subjective pairwise comparison; therefore, the correlations between the pupillary dilation responses and salience, as well as loudness, were not mediated by the task demands.

In the present study, salience was defined as being “easy to notice,” and it was expected to be associated with loudness and vigorousness. The argument was backed up by the high correlation between the subjective judgments of salience and loudness/vigorousness, regardless of the match between the sound pressure level (Exp. 1) or loudness (Exp. 2) among the sounds. No such tight correlation was found between salience and any other subjective judgment. Together with the results that the pupillary dilation response was highly correlated with salience and loudness, but not with vigorousness, the overall results suggest that pupillary dilation reflects the subjective salience of sounds, which is defined, or is heavily influenced, by loudness.

Accordingly, we examined the effect of salience on the pupillary dilation responses by manipulating loudness (Exp. 3). The results confirmed our hypothesis that both the physical sound intensity and acoustic characteristics that contribute to loudness induce pupillary dilation responses. A regression analysis showed a strong linear correlation between the pupillary dilation response and loudness. The conclusion also holds true for the various environmental sounds with equal sound pressure levels (the stimuli used in Exp. 1). Although the aim of the study was to investigate the effects of subjective salience on pupillary responses, the results demonstrated that salience is indicative of, or is heavily influenced by, loudness.

The underlying mechanism of the correspondence between the pupillary dilation response and salient yet loud sounds may be due to an autonomic response to the auditory stimulation that shares a common mechanism with salience detection, arousal, or maybe even pain. For example, auditory stimuli that have spectral differences activate different auditory channels, and thus have different impacts on the superior colliculus (King & Carlile, 1994; Wise & Irvine, 1983). On the basis of the finding that neural activation in the superior

colliculus evokes a pupillary response (Netser et al., 2010; Wang et al., 2012), it is reasonable to suppose that different stimulations of the superior colliculus caused by different auditory spectra would result in different pupillary responses. The finding that the superior colliculus is involved in the neural circuits of visual salience detection (Comoli et al., 2003) and receives multisensory inputs (Meredith & Stein, 1986) suggests that it could also be involved in auditory salience detection. The co-occurrence of the involvements of the superior colliculus in pupillary responses to auditory stimulation and in salience detection may indicate a close link between pupillary responses and auditory salience. Alternatively, it could be that the loudness induces arousal or even pain, which in turn influences pupil dilation. In any case, the results of the individual analysis indicate that the pupillary dilation response and physical properties (e.g., spectrum) are not linked merely in a straightforward stimulus–response reflex. Instead, the pupillary dilation response reflects an individual’s subjective feelings—that is, of salience as well as of loudness—which arises from an interaction between the listeners and the sound properties.

The present study suggests that pupillary responses make it possible to probe auditory salience. Previous studies have shown that pupillary dilation responses are evoked by high-contrast visual and high-intensity auditory stimuli that are presumably salient (Wang et al., 2014; Wang & Munoz, 2014; cf. Mathôt, Siebold, Donk, & Vitu, 2015), but they did not establish a link between pupillary dilation response and psychological salience. In the present study, we showed that the pupillary dilation response is nicely correlated with the subjective feeling of the salience of sounds. Moreover, in the present study we used various sounds, including synthesized sounds and sounds recorded in natural environments. This extended the ecologically valid circumstances in which the pupillary dilation response can be used as a physiological index of auditory salience.

Author note This study is supported in part by Japan MIC SCOPE (Grant No. 121803022).

References

- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus–norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, 28, 403–450. doi:10.1146/annurev.neuro.28.061604.135709
- Bala, A., & Takahashi, T. (2000). Pupillary dilation response as an indicator of auditory discrimination in the barn owl. *Journal of Comparative Physiology A*, 186, 425–434.
- Berglund, B., Berglund, U., & Lindvall, T. (1975). Scaling loudness, noisiness and annoyance of aircraft noise. *Journal of the Acoustical Society of America*, 57, 930–934.
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45, 602–607. doi:10.1111/j.1469-8986.2008.00654.x
- Bradshaw, J. (1967). Pupil size as a measure of arousal during information processing. *Nature*, 216, 515–516.
- Comoli, E., Coizet, V., Boyes, J., Bolam, J. P., Canteras, N. S., Quirk, R. H., ... & Redgrave, P. (2003). A direct projection from superior colliculus to substantia nigra for detecting salient visual events. *Nature Neuroscience*, 6, 974–980.
- Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, 105, 1704–1709. doi:10.1073/pnas.0707727105
- Einhäuser, W., Koch, C., & Carter, O. L. (2010). Pupil dilation betrays the timing of decisions. *Frontiers in Human Neuroscience*, 4, 18. doi:10.3389/fnhum.2010.00018
- Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature Neuroscience*, 16, 1146–1153. doi:10.1038/nn.3428
- Ellis, C. J. (1981). The pupillary light reflex in normal subjects. *British Journal of Ophthalmology*, 65, 754–759.
- Gabay, S., Pertzov, Y., & Henik, A. (2011). Orienting of attention, pupil size, and the norepinephrine system. *Attention, Perception, & Psychophysics*, 73, 123–129. doi:10.3758/s13414-010-0015-4
- Glasberg, B. R., & Moore, B. C. J. (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society*, 50, 331–342.
- Goldinger, S. D., & Pappas, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science*, 21, 90–95.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2, 194–203. doi:10.1038/35058500
- King, A. J., & Carlile, S. (1994). Responses of neurons in the ferret superior colliculus to the spatial location of tonal stimuli. *Hearing Research*, 81, 137–149.
- Lavin, C., Martin, R. A., & Jubal, E. R. (2014). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience*, 7, 218. doi:10.3389/fnbeh.2013.00218
- Liao, H. I., Yoneya, M., Kidani, S., Kashino, M., & Furukawa, S. (2014). *Human pupil dilation responses to auditory stimulations: Effects of stimulus property, context, probability, and voluntary attention*. San Diego: Poster presented at the Association for Research in Otolaryngology (ARO) 37th Annual MidWinter Meeting.
- Mathôt, S., Siebold, A., Donk, M., & Vitu, F. (2015). Large pupils predict goal-driven eye movements. *Journal of Experimental Psychology: General*, 144, 513–521. doi:10.1037/a0039168
- Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *Journal of Neurophysiology*, 56, 640–662.
- Naber, M., & Nakayama, K. (2013). Pupil responses to high-level image content. *Journal of Vision*, 13(6), 7. doi:10.1167/13.6.7
- Naber, M., Frassle, S., Rutishauser, U., & Einhäuser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of Vision*, 13(2), 11. doi:10.1167/13.2.11
- 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, 108–118.
- Preuschhoff, K., Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline’s role in decision making. *Frontiers in Neuroscience*, 5, 115. doi:10.3389/fnins.2011.00115
- Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., & Aguilar, M. (2010). Pupil dilation during visual target detection. *Journal of Vision*, 10(10), 3. doi:10.1167/10.10.3

- Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A., McNamara, J. O., & Williams, S. M. (Eds.). (2004). *Neuroscience* (3rd ed.). Sunderland: Sinauer.
- Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*, 211–223.
- Wang, C. A., & Munoz, D. P. (2014). Modulation of stimulus contrast on the human pupil orienting response. *European Journal of Neuroscience*, *40*, 2822–2832. doi:[10.1111/ejn.12641](https://doi.org/10.1111/ejn.12641)
- 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. *Journal of Neuroscience*, *32*, 3629–3636. doi:[10.1523/JNEUROSCI.5512-11.2012](https://doi.org/10.1523/JNEUROSCI.5512-11.2012)
- Wang, C. A., Boehnke, S. E., Itti, L., & Munoz, D. P. (2014). Transient pupil response is modulated by contrast-based saliency. *Journal of Neuroscience*, *34*, 408–417. doi:[10.1523/JNEUROSCI.3550-13.2014](https://doi.org/10.1523/JNEUROSCI.3550-13.2014)
- Wise, L. Z., & Irvine, D. R. (1983). Auditory response properties of neurons in deep layers of cat superior colliculus. *Journal of Neurophysiology*, *49*, 674–685.