Received: 6 February 2023 | Revised: 26 April 2023 | Accepted: 29 April 2023

DOI: 10.1111/psyp.14350

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

Eye blinking, musical processing, and subjective states—A methods account

Elke B. Lange[1] | Lauren K. Fink[1,2,3]

[1]Department of Music, Max Planck Institute for Empirical Aesthetics, Frankfurt/M, Germany

[2]Max Planck NYU Center for Language, Music, & Emotion, Frankfurt/M, Germany

[3]Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Ontario, Canada

Correspondence
Elke B. Lange, MPIEA, Grüneburgweg 14, 60322 Frankfurt, Germany.
Email: elke.lange@ae.mpg.de

Abstract

Affective sciences often make use of self-reports to assess subjective states. Seeking a more implicit measure for states and emotions, our study explored spontaneous eye blinking during music listening. However, blinking is understudied in the context of research on subjective states. Therefore, a second goal was to explore different ways of analyzing blink activity recorded from infra-red eye trackers, using two additional data sets from earlier studies differing in blinking and viewing instructions. We first replicate the effect of increased blink rates during music listening in comparison with silence and show that the effect is not related to changes in self-reported valence, arousal, or to specific musical features. Interestingly, but in contrast, felt absorption reduced participants' blinking. The instruction to inhibit blinking did not change results. From a methodological perspective, we make suggestions about how to define blinks from data loss periods recorded by eye trackers and report a data-driven outlier rejection procedure and its efficiency for subject-mean analyses, as well as trial-based analyses. We ran a variety of mixed effects models that differed in how trials without blinking were treated. The main results largely converged across accounts. The broad consistency of results across different experiments, outlier treatments, and statistical models demonstrates the reliability of the reported effects. As recordings of data loss periods come for free when interested in eye movements or pupillometry, we encourage researchers to pay attention to blink activity and contribute to the further understanding of the relation between blinking, subjective states, and cognitive processing.

KEYWORDS

absorption, emotion, eye blink, eye tracking, mixed effects models, music, outlier analysis

1 | INTRODUCTION

Music can induce powerful changes in one's mental and physical state (e.g., alterations in attention, emotion, stress level, movement energy, etc.), and has therefore been used to induce such states (see Vjästfjäll, 2002; Warrenburg, 2020, for reviews), often evaluated by self-reports. We seek more implicit measures of music-induced states via measurement of eye-blinking activity by video-based eye tracking. This measure is less obtrusive than other prominent measures in emotion research, such as electromyography, electrocardiography, skin conductivity, and respiration changes (e.g., Kim & André, 2008). Further, recent webcam eye tracking (e.g., Saxena et al., 2022) methods enable new, exciting perspectives for blink research in large-scale applications. However, blinking is understudied in the field of emotion research, and, for video-based recordings, there is no agreement regarding (1) blink definition, (2) outlier detection, and (3) the necessity, or not, of rejecting outlier participants. Our study closes this gap by suggesting a data-driven approach for defining blinks and outliers, and by illustrating the consequences of outlier rejection.

Generally, one reason to blink is to keep intact the thin tear film that protects the cornea (Sweeney et al., 2013). Physiological properties of the eye affect blink frequency (e.g., dry eyes, Nakamori et al., 1997). Also, several psychological processes are associated with blink activity, some related to changes of states and demands, others to individual differences in pathophysiology. For example, visual demands decrease blink activity (Recarte et al., 2008; Veltman & Gaillard, 1996), but arousal (De Jong & Merckelbach, 1990) and fatigue, or time-on-task, increase it (Fukuda et al., 2005; Stern et al., 1994). Likewise, blink rates increase when attention shifts from outward (e.g., stimulus perception) to inward (e.g., thoughts; Annerer-Walcher et al., 2018; Nakano et al., 2013; Schäfer & Fachner, 2014; Smilek et al., 2010). An increase of blinking marks the chunking of information (e.g., the end of speaking units during conversation; Sacks et al., 1974). Furthermore, dopamine-related differences in the nervous system affect blink frequency (Jongkees & Colzato, 2016).

Regarding music listening, previous studies suggest that emotions are related to blink activity. For example, emotional states induced by music have consequences for reflexive blink control shown by the startle reflex (Roy et al., 2009). Also, a few studies exist that relate spontaneous blink activity with musical processing, though blinking was not in the main focus of those studies and results diverge. The general presence of music, compared to silence, increases blink activity (Hammerschmidt & Wöllner, 2018; Schäfer & Fachner, 2014), but the presence of tones and words does not (Liu et al., 2020), or not reliably (see group A and B for conditions 1 and 3 in table 1 of Huber et al., 2022), indicating that musical processing affects blink activity differently from other auditory stimuli. The subjective feeling of being absorbed by music (Lange et al., 2017) decreases blink rates. Blink rates did not change for self-selected, favorite music in comparison with experimenter-selected, control music (Schäfer & Fachner, 2014), pointing to the fact that absorption is a specific subjective state related to blink activity, whereas preferences are not. In one study, self-selected, chill-inducing music did not effect blink rate in comparison with control music (Laeng et al., 2016), but in another study it reduced blink rates (Laeng et al., 2021).

Comparison of the above studies is difficult for several reasons. First, experimental settings differed. In some studies, participants were instructed to inhibit blinking or keep their eyes open because of the interest in recording measures of saccadic or pupil activity (Laeng et al., 2016; Lange et al., 2017; Lange et al., 2020), whereas, in others, spontaneous blink rate was directly investigated (e.g., Huber et al., 2022; Liu et al., 2020). Further, the studies differed by the additional presence of visual stimuli (Hammerschmidt & Wöllner, 2018; Huber et al., 2022; Laeng et al., 2021; Schäfer & Fachner, 2014) or not (Laeng et al., 2016; Lange et al., 2017; Liu et al., 2020), and by instructions to centrally fixate (Lange et al., 2017; Liu et al., 2020) or fixate somewhere on a scale (Laeng et al., 2021) versus free viewing (Hammerschmidt & Wöllner, 2018; Huber et al., 2022; Laeng et al., 2016; Schäfer & Fachner, 2014). Some studies did not include a measurement in silence (Laeng et al., 2016; Laeng et al., 2021). The task to inhibit blinking (Lange et al., 2017) can be understood as a dual task. Likewise, processing visual information might alter blink activity. Two music studies included a rather low number of trials and/or participants (Hammerschmidt & Wöllner, 2018; Schäfer & Fachner, 2014). Therefore, we aimed to replicate the effect of music listening on blink rates by presenting a broad range of musical styles, with a reliable measurement of blink activity (i.e., spontaneous blinking, no visual stimulation), and with overall increased experimental power. In addition to the present new study, we also report reanalyses from two earlier studies (Lange et al., 2017), in which the same stimuli were presented, but blinking had to be inhibited in the context of fixation. We explore, post hoc, how these different settings might have contributed to diverging results reported in the literature (Table S1).

Second, with respect to blink analyses, there is no agreed-upon practice of how to define blinks, particularly when measured by video-based, infra-red eye trackers. Eye trackers record pupil size and shape, which is used to infer eye position. During blinking, the eye lid covers the pupil, resulting in data loss. Such data loss periods have been used to define blinks. Very short periods might simply be measurement noise, for example, caused by eye lashes covering the pupil. Long periods do not reflect natural blinking, but phases of micro-sleep (Rodriguez et al., 2018), or instances when participants' gaze is not directed to the recording camera, or other recording failure. A broad range of lower and upper limits have been reported in the literature, with lower limits ranging between 50 and 300 ms (50 ms: Brych et al., 2020, 2021; Caffier et al., 2003; Wang et al., 2011; 100 ms: Aarts et al., 2012; Rodriguez et al., 2018; 200 ms: Jongkees & Colzato, 2016; 300 ms: Nomura et al., 2015) and upper limits between 200 and 1000 ms (200 ms: Naicker et al., 2016; 500 ms: Aarts et al., 2012; Brych et al., 2020; Caffier et al., 2003; Jongkees & Colzato, 2016; Nakano et al., 2013; Rodriguez et al., 2018; Wang et al., 2011; 1000 ms: Brych et al., 2021; Nomura et al., 2015). Note that defining blinks by data loss periods is a simplification, as those do not cover the start and end of the eye-lid movement (see Caffier et al., 2003). Therefore, the sophisticated blink detection algorithms take those times into account, for example, by capturing the movement of the gaze (e.g., Geng et al., 2008) or a change in the pupil diameter based on the closing lid (e.g., Van Ordern et al., 2000). To know the exact time points of the start and end is particularly important when data loss has to be interpolated in order to analyze the continuous pupil signal (Hershman et al., 2018; Mathôt, 2013; Piquado et al., 2010). Sometimes results of such an algorithm are reported together with the note that the experimenter has controlled every single blink detection (e.g., Nakano et al., 2013), indicating that there is no agreement about the reliability of these algorithms by different researchers. But a recently published algorithm reports being as precise as human judges (Hershman et al., 2018). Note that when measuring blinks by electrooculography (EOG) or application of a magnetic search coil on the eye lid, the precision of the recorded blink movement is higher than in case of video recordings, and elaborate suggestions on how to define blink parameters have been published (e.g., Caffier et al., 2003; Cruz et al., 2011; Kleifges et al., 2017; Liu et al., 2017; VanderWerf et al., 2003). We focus here on video-based eye tracking, for which blink information can easily be extracted without being the main interest of the researcher.

Third, there is also no agreed-upon practice about how to treat outlier trials and participants. Spontaneous blink activity varies tremendously. Ranges of 2.8–48 blinks per minute (Doughty & Naase, 2006) and 5.3–27.4 blinks per minute (Jongkees & Colzato, 2016: healthy human controls) have been reported. Investigations into spontaneous eye-blink activity are complicated by the fact that person-level characteristics interact with mental states and task demands, which change dynamically, sometimes causing blinking effects to go in different directions, cancel each other out, or interact with each other, ultimately making it difficult to relate physiological measures to unique states or processes. Thus, the endeavor to relate blink activity to music-induced changes in states can be a challenge.

In light of these difficulties, it is highly important to understand how to define and treat outliers. Several approaches exist, differing in whether and on which level data are excluded from further analyses. One way is to relate blink activity of several treatments to a baseline, either by subtraction (De Jong & Merckelbach, 1990; Naicker et al., 2016), or proportion (Huber et al., 2022). This handling can be interpreted as normalization and all data are then included. Another option is to trim the means by excluding a certain percentage of data at the left and right side of the distribution (e.g., Huber et al., 2022). For example, trimming a distribution of all participants' mean blink rates (e.g., all individual mean blink rates) excludes the means of high- or low-blinking participants. Trimming relates to 10% or 20% of each side of the distribution (Mair & Wilcox, 2020), excluding up to 40% of the data. Furthermore, an arbitrary cutoff has been chosen for high-blinking participants, that is, individuals with blink rates of 50 per minute (Brych et al., 2021), 55 per minute (Brych et al., 2020), or 1 per second (Nakano, 2015; Nakano et al., 2009, 2013), or for high- and low-blinking participants falling outside $M \pm 1$ *SD* (Nakano & Kitazawa, 2010). Such cutoffs define which participants are excluded from the data set. Often, reports do not include information about the issue of outlier participants, indicating that high- and low-blinking participants were included in the data set (e.g., Aarts et al., 2012; Annerer-Walcher et al., 2018; Fukuda, 2001; Fukuda et al., 2005; Nomura et al., 2015; Recarte et al., 2008; Smilek et al., 2010; Veltman & Gaillard, 1996). These discrepancies raise the question of whether outlier rejection is an important tool to apply to blink data analyses.

The purpose of the current study, then, is manifold: We explore whether blink activity is related to music listening, focusing on three perspectives: (i) an effect of music presence; (ii) a relation between subjective states during music listening and blink activity; (iii) a relation between acoustic features and blink activity. Furthermore, we are interested in the methods of defining and analyzing blinks. We propose a definition of blinks based on the individual distributions of blink durations and suggest data-driven accounts to define outlier trials and participants. Importantly, we investigate the consequences of outlier rejection by comparing results including all trials and participants or excluding the outliers from the data set. We discuss different accounts of how to treat blink data using mixed effects models. Finally, we compare blink activity between the silence and music condition to establish the silence condition as proper control for future studies.

2 | METHOD

We build on both the current study as well as the published eye-movement data from our lab (Lange et al., 2017). In all data sets, we used music as auditory stimuli, no visual stimulation, and manipulated blink and fixation instructions. In one experiment, we asked participants to inhibit blinking (IB) while fixating a central target (IB-fix; Experiment 1 in Lange et al., 2017) and in the other experiment to inhibit blinking while gazing freely around the empty screen (IB-free; Experiment 2 in Lange et al., 2017). The exact instruction for blink inhibition in both studies (IB-fix, IB-free) was less commanding but more indirect by explaining that our goal is to record the pupil and adding: "It is beneficial if you avoid blinking and squinting your eyes as much as possible." The current study measured spontaneous blinking (SB), without mentioning any information on blinking. The corresponding passage was omitted in the instructions. We decided on instructing participants to fixate a central target (SB-fix), as the IB-free experiment showed unusual, drifting eye movements. For an overview of the different data sets, see Table 1. In the analyses and conclusion, we focus on the current spontaneous blinking (SB-fix) experiment and complement these results with the results of the other data sets. All data and code required to recreate the reported analyses are publicly available at https://osf.io/65nyh/. A variety of further analyses are reported in the Supporting Information.

2.1 | Participants and sample size

Twenty-nine volunteers participated in this experiment, with a mean age of 23 years ($SD=3$, range: 18–29). Five were male, 24 were students (eight from psychology, two from music/musicology, 14 from other disciplines). Twenty participants had corrected-to-normal and nine had normal vision. Participants showed a broad range of musical taste, evaluated on 14 musical styles. Participants showed a medium level of musical sophistication, $M=74$, $SD=15$, range: 53–112 (Gold-MSI, Müllensiefen et al., 2014).

Participants gave written informed consent for the experiment, which occurred in two sessions, each with a mean duration of 71 min. They were compensated by an honorarium of 10 Euro per hour. The experimental procedures were approved by the Ethics Council of the Max Planck Society. The sample size was chosen according to the two prior studies ($N=30$, $N=35$, Lange et al., 2017). For IB-fix, in which participants fixated a central target, similar to the current study, the post hoc calculation of effect size and power resulted in dz = 0.75 and $1-\beta=.99$ for the t test between blink rates during silence versus music. Fitting this post hoc effect size into an a priori power analyses ($\alpha=.05$, $\beta=.20$), resulted in a sample estimate of 13. Given the fact that we had no clear assumption about the effect size for SB-fix, if

TABLE 1 Overview of experiments: current study SB-fix; earlier studies IB-fix, IB-free.

	SB-fix	IB-fix	IB-free
Blinks	Spontaneous	Inhibition	Inhibition
Viewing instruction	Fixate (central target)	Fixate (central target)	Free viewing (no central target)
Trials: music & silence = total	59 & 7 = 66	56 & 14 = 70	56 & 14 = 70
Subjects numbers	29	30	35
Total number of trials	1914	2100	2450
Missing trials	82	8	6
Remaining trials	1832	2092	2444
Trials with zero blinking	105	458	370
All recorded data loss periods	26,595	7991	14,292
Defined blinks within all music or silent trials	24,644	7795	13,749
Trial outliers	11 from 8 partic.	24 from 20 partic.	27 from 18 partic.
Number of High (H) or low (L) blinking participant	H: 4, L: 2	H: 3, L: 8	H: 3, L: 10
Remaining number of trials, subjects and blinks (% of data missing or rejected)	Trial $n=1497$ (−21.79%) Subject $n=23$ (−20.69%) Blink $n=16{,}037$ (−34.93%)	Trial $n=1312$ (−37.19%) Subject $n=19$ (−36.67%) Blink $n=4958$ (−36.40%)	Trial $n=1525$ (−37.76%) Subject $n=22$ (−37.14%) Blink $n=8781$ (−36.13%)

anything, we should have a well-powered experiment with 29 participants.

2.2 | Apparatus

Participants were placed in a sound attenuated, dimly lit booth, equipped with an eye tracker (EyeLink 1000, SR Research, 500 Hz sample rate, binocular recording), monitor (resolution 1920 × 1080, refresh rate 144 Hz), standard computer mouse and keyboard, and Neumann KH 120 A G loud speakers. The height of the chair and table was adjusted to the participants' needs. Participants' heads were supported by a chin and head rest. The distance between monitor and chin rest was 66 cm. The experiment ran on a Windows PC. Stimulus presentation and data collection were programmed in PsychoPy 1.82.01 (Peirce, 2007).

2.3 | Audio materials

2.3.1 | Stimuli

The 59 musical stimuli were taken from prior studies (56 from Lange et al., 2017; additional three from Lange & Frieler, 2018). They were excerpts from a broad selection of instrumental music from 14 different styles (blues, country, electronica, folk, hiphop, classical, jazz, metal, pop, reggae, rock, soul, German Volksmusik, world music). The loudness of all stimuli was normalized. For a complete list of stimuli and adjustments, please see appendix of Lange and Frieler (2018) with stimulus No 60 being excluded in the present study due to a programming error in the experimental procedure. Participants adjusted the loudness to their personal level of comfort at the beginning of the first session (range of −21 to −67, modus of −30 LUFS). Musical stimuli were 43–61 s long. The seven silent trials without music had each a duration of 40 s. Note that this duration is on the lower end of the durations of music trials and differed from IB-fix and IB-free, in which we recorded silent trials of 60 s.

2.3.2 | Acoustic properties

In an earlier study (Lange & Frieler, 2018), we analyzed the musical features of the stimuli used in the current study. From the 85 extracted features, we selected 10 features based on their interpretability and rather low correlations (see Lange & Frieler, 2018): mean root mean square of amplitude (loudness), mean pitch (brightness), standard deviation of attack time (variability of articulation), mean pulse clarity (beat strength), mean key clarity (chromaticism), mean mode (modality), standard deviation of mode (changes in modality), mean low energy (loudness contrasts), mean regularity (homogeneity of spectral peaks, polyphony), mean spectral novelty (spectral/musical contrasts). We added a measure for tempo, tapped by two professional percussionists, and extracted via peak frequency from Tomic and Janata's (2008) Beyond-the-Beat Model, but proceeded with the peak frequency from the linear oscillator model, as this value is easily reproducible.

2.4 | Procedure

In the beginning of each session, participants filled in the questionnaires on musical sophistication and taste, split and serial order balanced across the two sessions. The musical part consisted of 66 trials (59 with music, 7 in silence), 33 trials per session. Three silence trials were spread equally spaced across each session, one was randomly added. The serial order of the 59 music excerpts was randomized for each participant. The musical part started with a 9-pt calibration and validation procedure of the eye tracker. Calibration was repeated on every fifth trial or whenever participants failed to fixate properly. Each trial started self-paced with the presentation of a dot (radius = 0.15° visual angle) centrally on the screen. Participants had to fixate this dot. If they did so within 800 ms, the musical excerpt or silence period started, otherwise the calibration procedure was repeated. During the trial, participants' task was to continue fixation, and in trials with music to immerse themselves into the music. They were informed that they might be absorbed in some music, and in other music not. After each listening episode, participants evaluated their musical absorption, mind wandering, felt groove, felt valence, felt arousal, and liking of the music, on a 5-pt rating scale from *fully disagree* to *fully agree*. For valence and arousal, the 5-pt Self-Assessment Manikins were used, spanning positive–negative valence and high-low arousal (Bradley & Lang, 1994), recoded for analyses to range from negative to positive and low to high arousal. In silent trials, participants evaluated valence and arousal, too. To ensure spontaneous blinking, blinking was not mentioned at all in the instructions.

2.5 | Data treatment

We recorded the ocular data from the beginning of the fixation check until the end of the musical excerpt or the silence trial. From the total of 1914 trials, 82 trial recordings failed for technical reasons, resulting in a total of 1832 trials.

2.5.1 | Blink identification

As reported above, there is a lack of agreement on durations for valid blinks, for the upper as well as lower limits. Hence, we guided our definition by the individual distributions of the durations. We set up the following multi-step procedure for blink identification: First, data loss periods were defined by missing pupil data from either eye, abrupt fluctuations of the pupil area (>20 units per sample), and gaze positions out of the screen. Data loss periods less than 20 samples apart were merged into one period. Then, blinks were defined from the data loss periods based on the following criteria: (i) include binocular data loss periods only; (ii) durations <50 ms and >2000 ms were excluded as rough duration outliers; then, (iii) data loss periods longer than the individual $M_{\text{subject}} + 3$ *SD* were excluded; finally, (iv) blinks during the initial fixation period (800 ms before music started) were excluded as were blinks starting before music onset.

Figure 1, upper panel, shows the distributions of binocular data loss periods with rough outliers excluded (criteria i and ii applied) from four example participants (subject IDs 1, 4, 7, and 14), with the individual duration $M_{\text{subject}} + 3$ *SD* added as a red vertical line (see Figure S1 for all participants). Distributions of individual durations were normal to right-skewed, and the red lines captured the properties and ranges of the individual distributions. They hence appeared to be most suitable to differentiate valid blinks from micro-sleep or other reasons for long data loss. Note that in our data most cutoffs were in the bin 450–500 ms (Figure 1b, upper row), which is what several authors have applied in the literature (500 ms cutoff: Aarts et al., 2012; Caffier et al., 2003; Jongkees & Colzato, 2016; Nakano et al., 2013; Rodriguez et al., 2018; Wang et al., 2011). However, Figure 1b (upper row) shows how strongly individuals differ in their personal $M_{\text{subject}} + 3$ *SD*. This is convincing evidence to adjust the upper cutoff individually instead of applying an arbitrary value.

In the data set of 1832 recorded trials, 26,595 data loss periods were detected by the above criteria (i); 1487 by (ii), and additionally 389 by (iii), resulting in rejection of 7.05% of binocular data loss periods. An additional 75 blinks fit criteria (iv). That is, we recorded 24,644 valid blinks in total. There was at least one blink in 1727 of the 1832 recorded trials.

Figure 2 shows histograms of blink counts across trials split for different blink durations, discussed in the literature. The differences in counts between trials with music (blue bars) and in silence (red bars) are partially due to the difference in the durations of the trials. The measure of blink rate adjusts for these differences in length and will be reported later. Overall, most blinks were in the duration range 100–300 ms. For these blinks, the least within and between session effects occur, whereas the other ranges showed strong differences. Note that for these histograms outlier trials and participants were not excluded. Most of the within and between session effects were attenuated, after excluding outlier trials and participants (see Figure S2). It seems, then, that for studies interested in splitting blinks by durations, outlier rejection is of importance. Such differences could otherwise be erroneously

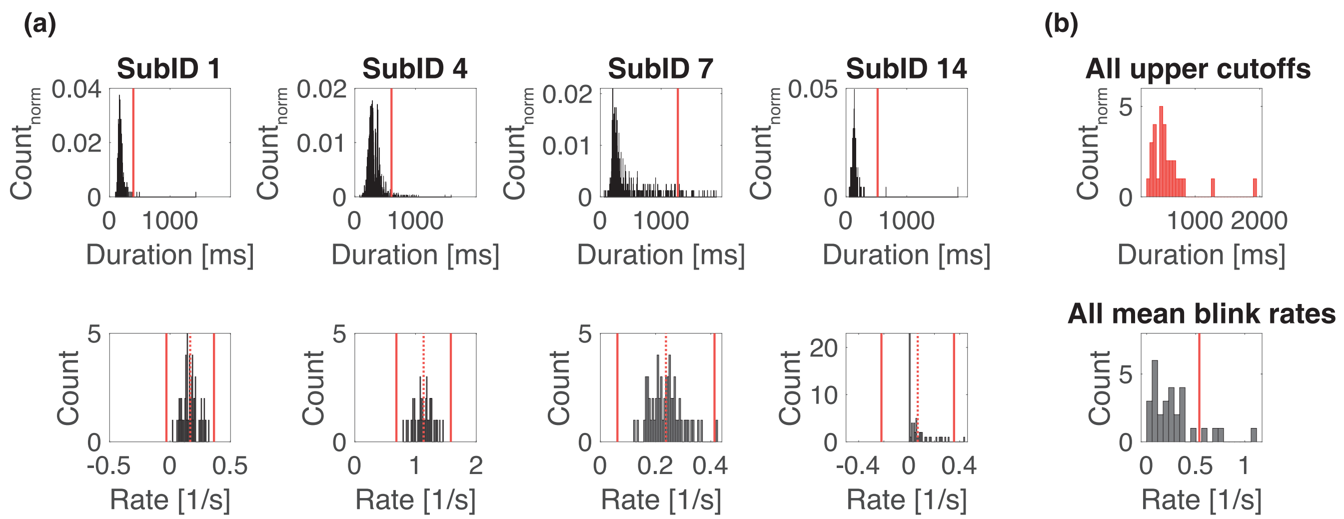

FIGURE 1 Examples for the blink and outlier definitions. (a) Individual distributions from four example participants (SubID 1, 4, 7, 14), (b) Distributions from the results of all participants. (a), upper row: individual normalized distributions of blink durations; the red line depicts $M_{\text{subject}} + 3$ *SD*; (a), lower row: the distributions of blink rates from the same four participants, with $M_{\text{subject}} \pm 3$ *SD* marked by red vertical lines. (b), upper part: distribution of upper cutoffs from all participants ($M_{\text{subject}} + 3$ *SD*), cases match the individual upper cutoffs marked in red in (a), upper row; (b), lower part: the individual mean blink rates per trial from all participants. Four high-blinking participants have rates to the right of the red line at $M_{\text{group}} + 1$ *SD* and were defined as outliers.

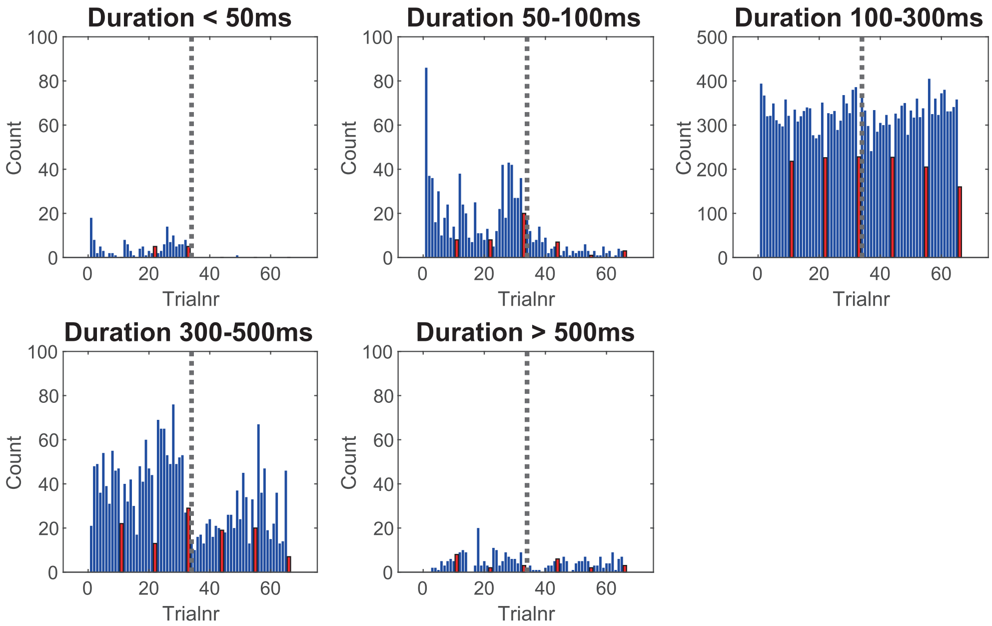

FIGURE 2 Histogram of number of blinks by duration, based on the serial position of trials within two sessions of the SB-fix experiment (trial and participant outliers included, see Figure S2 for outliers excluded). Blue bars depict trials with music, red trials in silence. The dashed line splits trials from the first and the second session.

interpreted as systematic differences in behavior (e.g., learning processes across time), or as different amounts of measurement noise between sessions (e.g., failure to optimize recordings in one session).

2.5.2 | Outlier rejection of trials and participants

We related the definition of trial and participant outliers to the measure of interest: blink rate (count per second). We visually inspected trial outliers by plotting the distributions of mean blink rates per trial for each participant (see Figures S3–S5). Blink rates were rather normally distributed. Then outlier trials were defined by applying $M_{subject} \pm 3$ *SD* as cutoffs. Figure 1a, lower panel, shows the blink rates from the four sample participants. Three of them had rather normally distributed blink rates, but they differed in their mean, e.g., the subject with ID 4 had high blink rates. One sample distribution of participant with ID 14 was heavily skewed, with a mode of zero blinking. Because there might be a systematic difference between music and silence conditions, we defined outliers of trials separately for music and silent trials. Only 11 trials from eight participants were defined as outliers and excluded.

Next, we defined outlier participants. The outlier criterion for high-blinking participants was based on Nakano and Kitazawa (2010): $M_{group} + 1$ *SD*. We re-calculated the individual means across all conditions after excluding the outlier trials. The subplot in the lower panel of Figure 1b identified four high-blinking participants as outliers, falling above the cutoff. For the low-blinking participants, we needed a different criterion, because participants such as ID 14 were not identified by the criterion $M_{group} - 1$ *SD* (Nakano & Kitazawa, 2010). We decided to differentiate between participants with a mode of zero blinks and a heavily skewed distribution, and low-blinking participants with a mean of slightly above zero but a more normal distribution. The two groups can be differentiated by $M_{subject} - 1$ *SD* ≤ 0 for the first, and $M_{subject} - 1$ *SD* > 0 for the second group. We treated only the first as outlier participants. Four participants (mean blink rates: 1.12, 0.78, 0.59, 0.72 blinks per second) exceeded the upper cutoff, and two participants (mean blink rates: 0.06, 0.01 blinks per second) fit the lower criterion. In the literature, a mean blink rate of 60 per minute (i.e., 1 per second) has been applied to define high-blinking participants (Nakano, 2015; Nakano et al., 2009, 2013), a criterion which only fits one of our high-blinking participants. No participant met the criterion of $M_{subject} < M_{group} - 1$ *SD* (Nakano & Kitazawa, 2010).

When outlier trials and participants were rejected, 21.16% of the total number of 1914 trials were discarded: 11 were excluded as outlier trials, 264 as trials from high-blinking subjects, 130 as trials from low-blinking subjects. Recordings of additional 12 trials were missing. This resulted in 1497 remaining trials for the analyses and 16,037 blinks within these trials.

2.5.3 | Analyses

We report *t* and *F* statistics, analyzed with MATLAB (R2019b), and Cohen's *d* with computeCohen_d (Bettinardi, 2021). The mixed effects models were analyzed in R (R Core Team, 2012), with the packages lme4 (Bates et al., 2015), GLMMadaptive (Rizopoulos, 2022) and MuMIn (Bartoń, 2009). We also applied rmcorr (Bakdash & Marusich, 2017).

3 | RESULTS

3.1 | Blinking during music vs. silence

3.1.1 | Does blink rate increase during music vs silence?

Figure 3, upper row, shows the normalized histograms of blink rates. In the IB studies, the distributions were skewed with most blink rates near zero, as can be expected when participants were instructed to inhibit blinking. The distribution for SB-fix is much broader with a greater number of high blink rates, and with slightly more high and slightly less low blink rates for music trials (blue lines) in comparison with silence (red lines), which points to higher mean blink rates for music than silent trials. Accordingly, most participants showed a proportional increase in blink rates from silence to music trials (Figure 3, second row), whereas about 1/4 in IB-fix and 1/3 in IB-free showed a decrease (negative difference score). The increase in mean blink rate by music was significant in SB-fix and IB-fix (Figure 3, third row), but not for IB-free (see Table 2 for statistics; note a tendency of $p < .10$ for IB-free when outliers were rejected). Thus, the effect of increased blink rate during music showed irrespective of blinking instructions (spontaneous versus inhibit) but seemed affected by viewing instructions (fixation versus free viewing). Results of the t tests were comparable for data sets including or rejecting outliers, but in the analyses with outlier rejection, Cohen's d's increased (see also Table S2 for between-experiment

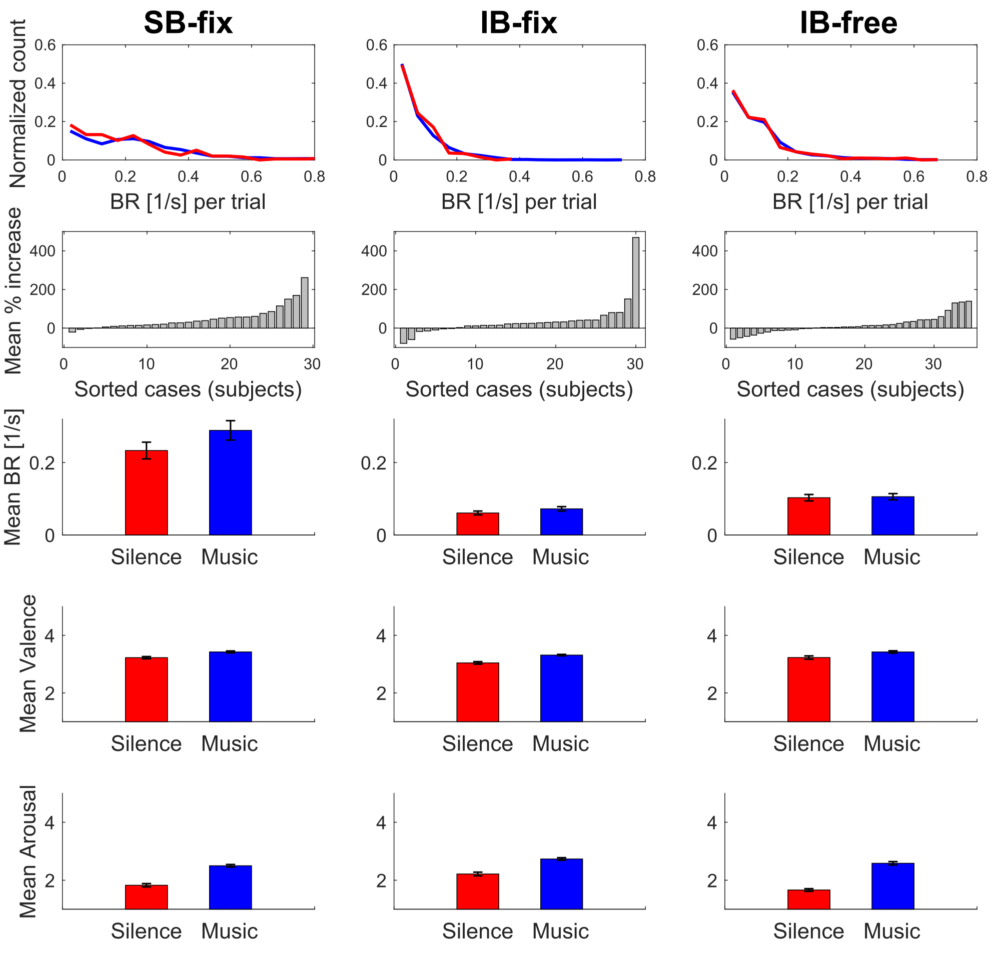

FIGURE 3 Comparison of the silence and music condition (outliers included, see Figure S6 for outliers excluded) in three data sets SB-fix, IB-fix, IB-free (see Table 1 for more information). Blue color represents the music and red the silence condition. Top row: distributions of blink rates for all trials and participants. Second row: proportional subject-based increase in blink rates by music, taking silence as baseline. Third to fifth row: subject-based mean blink rates (BR), valence, and arousal ratings. Error bars show the standard errors of the means.

TABLE 2 *T*-statistics and Cohen's *d* of the comparisons of participant's mean blink rates in two conditions (silence or music).

	Outlier trials and participants included				Outlier trials and participants excluded			
	df	*t*	*p*	*d*	df	*t*	*p*	*d*
SB-fix	28	4.35	<.001	0.81	22	4.82	<.001	1.00
IB-fix	29	3.50	.002	0.64	18	3.16	.005	0.73
IB-free	34	0.60	.551	0.10	21	1.74	.096	0.36

Note: The left side of the table corresponds to the data shown in Figure 3 (see Figure S6 for outliers excluded).

TABLE 3 Pearson's correlation coefficient and regression parameters for the relation between silent and music mean blink rates.

	Outliers included				Outliers excluded			
	Slope	Intercept	df	*r*	Slope	Intercept	df	*r*
SB-fix	1.13	0.03	27	.97	1.09	0.03	21	.94
IB-fix	1.09	0.01	28	.96	0.94	0.01	17	.89
IB-free	0.92	0.01	33	.95	0.80	0.03	21	.94

Note: All correlations resulted in $p < .0001$.

comparisons with the same consequence of outlier rejection for effect sizes). Stronger effect sizes can be expected, as outlier rejection makes the data more homogenous, reducing the standard error at the expense of losing data at the low and high end of the distributions (Figure S6).

3.1.2 | How comparable is blink activity during music listening vs. silence?

Mean blink rates for silent and music trials

The individual blink rates for the two conditions correlated highly in all three data sets (see Table 3 for statistical results, Figure 4, first row, and Figure S7 excluding outliers). Linear regression showed an increase in intercept and change in slope. That is, music does not add blinking in an unsystematic or complex (e.g., polynomial) way, but has a linear effect (increase in intercept), as well as a proportional one (slope differs from 1).

Blink rates across trials

To examine continuous blink dynamics, we created binary blink time series by coding blink onset time as 1 and all other values as 0 (i.e., blink duration was not encoded) in bins of 1 s (as in Goldstein et al., 1992). We smoothed the data for plotting, by a moving window of 10 s. Figure 4, second row, depicts rates locked to the beginning of the trial, and the third row locked to the end. There are four observations: (i) The 95% confidence intervals were highly overlapping. In comparison with music trials, the silent trials had a much larger confidence interval and were less smoothed due to the lower number of trials. (ii) Mostly, rates were parallel across time for music and silent trials. (iii) Two specific effects emerged: A music onset response with higher blink rate when music started than in silence (Figure 4, second row). This effect occurs in SB-fix and IB-free, but not in IB-fix. There was also a very small tendency for an increase in blink rates in IB-fix and IB-free in the end of music trials (Figure 4, third row). Note that we do not add statistics for these tendencies, as the choice of window for the comparisons would be decidedly post hoc. Also note that we stopped collecting data at the end of the trial and an effect of the end of the music might be more visible, when the recordings had continued after trial offset.

Temporal structure: Burstiness versus periodicity

Figure 4 (row two and three) indicates that blink rates are not modulated very much across time. We now ask more specifically about the regularity of blink events, by analyzing the distribution of blink inter-onset intervals (IOIs). One useful measure is the calculation of the *burstiness parameter B* (Goh & Barabási, 2008),

$$B = \frac{\sigma - \mu}{\sigma + \mu}$$

which takes into account the standard deviation and mean of the distribution of blink IOIs. If events are periodically executed, the distribution of IOIs will be rather normal. In contrast, if events are chunked in bursts with longer time in between the bursts, the distribution will be heavily right skewed. B varies between $B = -1$, a periodic or anti-bursty pattern of events, and $B = 1$, a bursty signal, with $B = 0$ denoting a random signal. The parameter estimations depend

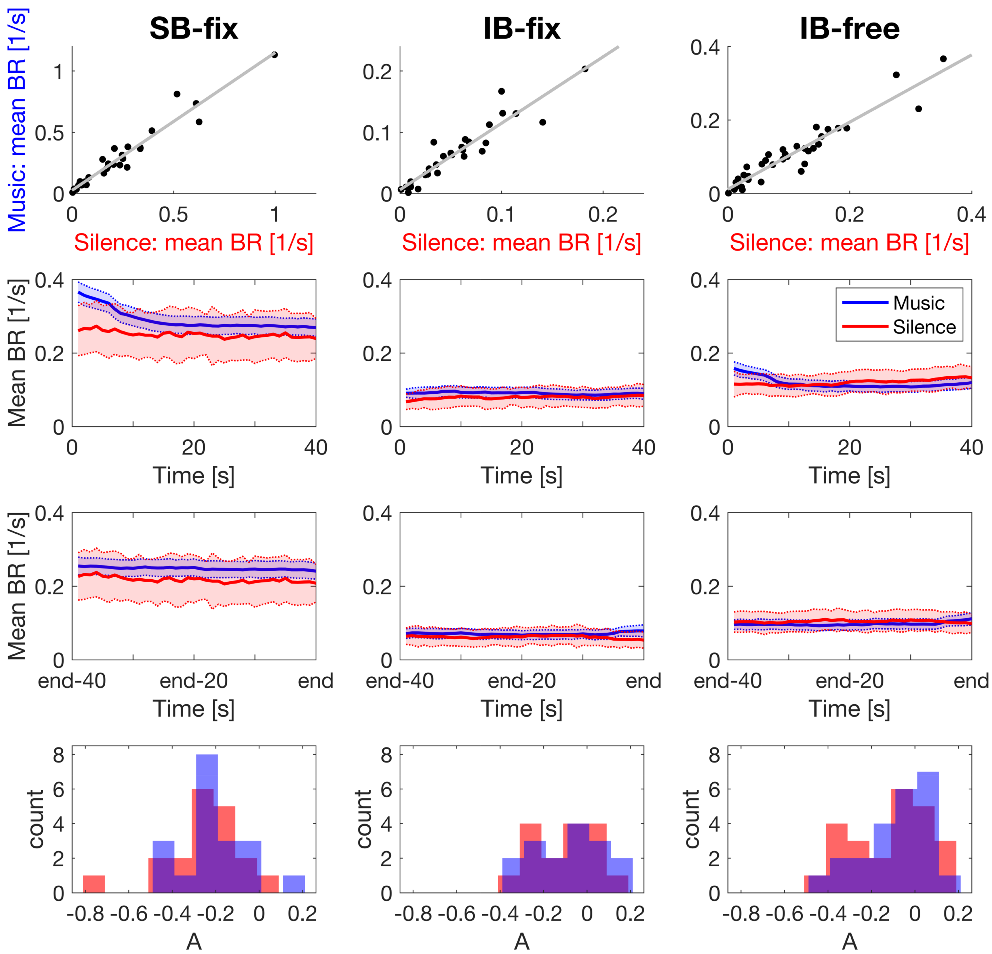

FIGURE 4 Comparing blink rates between the (red) silent and (blue) music conditions (outliers included, see Figure S7 excluding outliers) in the three data sets SB-fix, IB-fix, IB-free (see Table 1 for more information). First row: A linear regression line is added to the scatter plot; individual blink rates show a linear relation between conditions (see Table 3 for statistics). Second and third row: mean blink rates across time, locked to the beginning (second row) or the end (third row) of trials. The shaded areas show 95% confidence intervals for each 1 s bin across time. Last row: histograms of the *burstiness parameter A*, that can range from −1 (perfectly periodic) to 0 (random) to +1 (perfectly bursty).

on the number of events in the time series; therefore, an updated equation has been suggested by Kim & Jo, 2016 (see equation no. 23) and named *burstiness parameter A* to avoid confusion. For series longer than 100 units, the two estimates of B and A converge (Xu et al., 2020). We report the *burstiness parameter A*.

Because blinks are sparse events within a trial, we combined all IOIs within each condition for each participant. For the calculation of the *burstiness parameters A*, we included participants, when they had at least a sum of $n_{ioi} = 30$ IOIs in each condition, to ensure that we had a reasonable number of IOIs for the analyses. Thereby, participants with a low blink count were excluded.

In all three data sets, the majority of individuals' As were negative, which means that the signal was more periodic than bursty (Figure 4, last row). For inhibited blinking, the distributions shifted slightly to the right, becoming more random ($A = 0$). Distributions of A overlapped between the silent and music condition. We combined the data of all three sets to gain statistical power, resulting in $n_{subjects} = 59$ for the estimates of A for each the silent and music condition. The *burstiness parameter* differed between music and silence, $t(58) = -2.17$, $p = .034$, and was statistically different from zero (random) in the silent condition, $t(58) = -6.47$, $p < .001$, as well as the music condition, $t(58) = -5.48$, $p < .001$. In other words, blink IOIs were periodic and not bursty or random; they showed greater periodicity during silence than music.

We repeated these analyses, excluding outliers. The combined data included $n_{subjects} = 45$, and the t test showed no difference in the *burstiness parameter A* between conditions, $t(44) = -1.09$, $p = .280$. Inspection of the histograms showed that the distributions now overlapped much more (see Figure S7), particularly at the right part of the distribution, with lower numbers of cases of burstiness around zero

TABLE 4 R syntax for different types of models fitting self-rated absorption to predict blink activity.

	Model	Treatment of zero cases
(i)	glmer(br ~ abs+(1 + abs|subF) + (1 + abs|stimF), data = df, family = Gamma(link = "log"))	NaN
(ii)	glmer(br ~ abs+(1 + abs|subF) + (1 + abs|stimF), data = df, family = Gamma(link = "log"))	Set to 10^{-3}
(iii)	mixed_model(bno1 ~ abs, random = ~ 1|subF, data = df, family = zi.negative.binomial(), zi_fixed = ~abs, zi_random = ~1|subF))	Included
(iv)	glmer.nb(bno ~ abs+(1 + abs|subF) + (1 + abs|stimF), data = df)	Included

Note: Example models for absorption (abs) as predictor. Models differ by how they treat zero cases, by the dependent variable being blink rate (br) or number of blinks (bno), and the distribution (gamma for continuous distributions, negative binomial for count data). For information on the syntax of the zero-inflation model (iii) see Rizopoulos (2022). Random effects: subF = subjects (factorized), stimF = stimulus (factorized).

in the music condition. We tentatively conclude that the high-blinking participants—that had not been excluded when taking all data into account—might have more noisy blink data, eventually increasing the randomness in the signal. This hypothesis needs further testing. But the analyses show again that outlier rejection results in a more homogeneous data set and might be beneficial, when looking into the dynamics of the blink signal—unless one is interested in individual differences between participants.

Several other interesting tools exist to understand the dynamics within time series, such as detrended fluctuation analysis, cross-recurrence quantification analysis, or inter-subject correlation. Unfortunately, for blink data, the events of interest are rather sparse. Therefore, these analyses are not suitable in our current setting. For further discussion on this issue, please see Supporting Information, Section 8.

3.2 | Blinking and subjective states

3.2.1 | Do subjective states predict blink activity during music listening?

In addition to measuring self-rated states of absorption, mind wandering, and being in groove, we also measured participants' emotional experience via valence and arousal ratings, and their reported liking of the music. The goal was to use subjective ratings to predict spontaneous blink activity. However, ratings were correlated, some highly (absorption and valence: $r(1654) = .52$, absorption and liking: $r(1654) = .65$, both $p < .001$), some low or medium (absorption and arousal: $r(1654) = -.11$, absorption and groove: $r(1654) = .31$, both $p < .001$; repeated measures correlations). For the mixed effects modeling account, correlated predictors are not applicable as calculated weights cannot be interpreted. Hence, we did not create one model with all subjective ratings as additive predictors, but fit separate models for absorption, valence, arousal, and liking. MW and groove did not correlate ($r(1654) = -.03$, $p = .31$), and were therefore fit both in one model.

As there were trials without blinking, our data set included blink rates of zero. There are several ways to treat data with zero cases (see Table 4 for R code): (i) setting the zero cases of blink rates to NaN and fitting the remaining data with glmer and gamma family; (ii) setting the zero cases of blink rates to an arbitrarily small constant (e.g., 0.001) and fitting all data with glmer and gamma family; (iii) keeping the zero cases as blink counts and not rate, and fitting count data with the choice of using zero-inflated models and negative binomial distributions; or (iv) fitting the count data in a more simple way with glmer (without a separate zero-inflated part) and negative binomial distributions. However, the four approaches are not equally suitable.

Results of all four approaches are reported in Table S4. We report here the last option (iv), which we regard as straight forward and the optimal way to treat the data. The rational is the following: First, it can be argued that measuring nothing cannot be simply translated into the digit zero but is an invalid trial and needs to be set to NaN. But setting zero cases to NaN as in (i) trims the data distributions in a systematic way. In addition, zero blinking can be regarded as special case of low blinking, which might be of importance for the research question. In fact, as absorption might affect blinking by reducing its rate, exclusion of zero blinking is not sensible. Second, setting zero cases to a constant as in (ii) is suitable but less elegant. Third, our data showed inflation of zero cases above what can be expected (e.g., by a Poisson distribution). We then fitted zero-inflated models as in (iii). The advantage of these

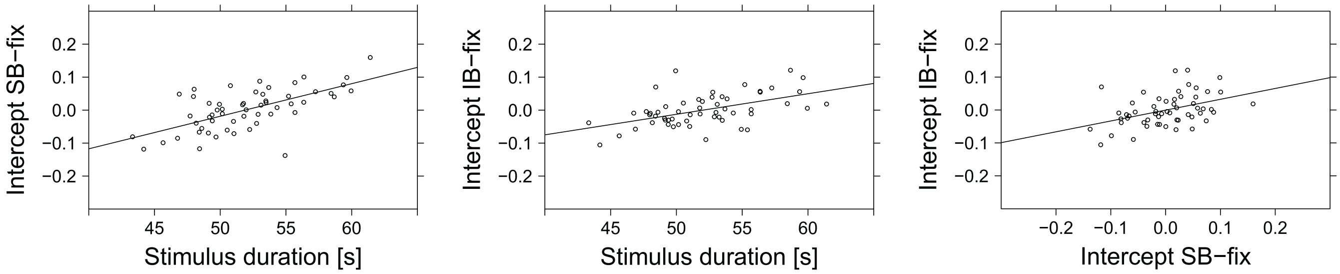

FIGURE 5 Scatter plots relating stimulus duration and the fitted stimulus-specific intercepts in the mixed effects models with absorption as fixed effect, in SB-fix and IB-fix experiments, outliers included (for more information on the data sets IB-fix and SB-fix see Table 1).

models is that the complete distribution can be split by one distribution for the zero part, and a second one for the remaining part. There are two obvious choices for the applied distributions: Poisson and negative binomial. As distributions were overdispersed, the negative binomial distribution was more suitable. In fact, when comparing models with different distributions, the negative binomial resulted in better fits (e.g., lower BIC). However, it seemed that these models were too complex for our data.[1] Hence, we finally decided to fit the data with mixed effects models on blink count data and a negative binomial distribution, without modeling a separate zero part as in (iv).

One potential problem of analyzing blink counts (as in iii and iv) is that counts do not take the stimulus duration into account. The musical excerpts ranged between 43 and 61 s, that is the longest stimulus was 50% longer than the shortest one. However, one advantage of mixed effects models is that the stimulus-specific variance can be captured. The question is, whether we captured a systematic increase of counts by stimulus duration. Figure 5 depicts the correlations between stimulus duration and the stimulus-based random intercept of the models on absorption, outliers included. The related correlations were $rho = .65$, $p < .001$ for SB-fix, and $rho = .49$, $p < .001$ for IB-fix. In addition, random intercepts correlated between experiments, $rho = .43$, $p < .001$. This is strong evidence in favor of the assumption that systematic differences in blink counts, due to differences in stimulus durations, were taken into account by the blink count models. We conclude, then, that applying models based on count data provides the best framework to fit our data. Note that in the models for IB-free, the random intercept did not contribute to the best model fit, that is, no stimulus-specific variance was captured by the model.

We analyzed the models from account (iv) by an iterative model comparison process of nested models (see Baayen et al., 2008), starting with the full model and deleting non-significant effects until the best model with the least number of meaningful parameters was found. Table 5 shows the resulting final models (see Table S4 for results of the other modeling accounts), including either all cases, or trial and subject outliers excluded. The trade-off between the two data sets is one between having more data but potentially more noise and having less data but more reliable measures. Interestingly, having outliers included or not did not matter very much for the resulting significance of the fixed effect, but sometimes increased the complexity of the structure of the best fitting model (IB-fix, IB-free).

Results showed that absorption and liking predicted blinking. More specifically, the fixed effect of absorption was significant in all but one comparison, for which the effect was marginal ($p < .10$, IB-fix, outlier included). The effect of liking was absent in IB-fix, and only marginal in SB-fix and IB-free, when outliers were rejected ($p < .10$). Overall, SB-fix resulted in the most complex models (e.g., including random intercepts and slopes). It seems obvious that complex relations between blinking and subjective states are best measured with spontaneous blinking. Blink inhibition alters the systematic variance that can be captured, reducing the possibility to include stimulus-specific random effects in the models.

Within experiment, the model structure for absorption and liking in SB-fix and IB-free were the same, enabling a direct comparison of the effect sizes for the fixed effects. Effect sizes were larger for absorption than liking and increased when outliers were excluded than included. The latter indicates, that excluding outliers is beneficial for analyses. But as mixed effects models capture systematic individual differences, such as high and low blinking,

[1]Models of highest complexity did not fit (random slopes and intercepts for both parts of the model); nested iterative testing resulted in error messages likely due to overfitting. Error messages disappeared, when setting iter_EM = 0 (to only use the quasi Newton part of the optimization). Then the simplest model fits best, having no fixed and random effects for the zero part. Also, with more complex models and not setting iter_EM to zero, the fixed effect of the zero-inflated model (e.g., for absorption) was not significant, whereas the one for the non-zero part was. Taken together, this indicated that mixed models without a specific zero part were sufficient for our data.

TABLE 5 Resulting best model fits for subjective ratings predicting blink activity (count), using either the complete data set or excluding trial and subject outliers for SB-fix, IB-fix, and IB-free.

Data	Model glmer.nb()	Predictor's z value	Effect size
SB-fix, outlier included, 1639 observations			
	bno ~ abs +(1 + abs|subF) + (1|stimF)	**$z = -2.34$, $p = .019$***	**.0023**
	bno ~ val + (1 + val|subF) + (1|stimF)	$z = -0.44$, $p = .662$	
	bno ~ arous +(1|subF) + (1|stimF)	$z = -0.14$, $p = .891$	
	bno ~ MW + groove +(1 + MW|subF) + (1|stimF)	MW: $z = -0.39$, $p = .694$ Groove: $z = -1.53$, $p = .125$	
	bno ~ like +(1 + like|subF) + (1|stimF)	**$z = -2.20$, $p = .028$***	**.0011**
SB-fix, outlier excluded, 1339 observations			
	bno ~ abs +(1 + abs|subF) + (1|stimF)	**$z = -2.38$, $p = .018$***	**.0026**
	bno ~ val + (1 + val|subF) + (1|stimF)	$z = -0.99$, $p = .322$	
	bno ~ arous +(1|subF) + (1|stimF)	$z = -1.59$, $p = .111$	
	bno ~ MW + groove +(1 + MW|subF) + (1|stimF)	MW: $z = -1.38$, $p = .169$ Groove: z = −0.81, $p = .421$	
	bno ~ like +(1 + like|subF) + (1|stimF)	**$z = -1.91$, $p = .056$**	**.0017**
IB-fix, outlier included, 1672 observations			
	bno ~ abs +(1 + abs|subF) + (1|stimF)	**$z = -1.84$, $p = .066$**	**.0022**
	bno ~ val + (1|subF)	z = 0.94, $p = .35$	
	bno ~ arous +(1|subF)	$z = 0.42$, $p = .966$	
	bno ~ like +(1|subF)	$z = -0.25$, $p = .803$	
IB-fix, outlier excluded, 1048 observations			
	bno ~ abs +(1|subF)	**$z = -2.34$, $p = .019$***	**.0037**
	bno ~ val + (1|subF)	$z = 0.63$, $p = .532$	
	bno ~ arous +(1|subF)	$z = -0.87$, $p = .382$	
	bno ~ like +(1|subF)	$z = -0.12$, $p = .903$	
IB-free, outlier included, 1954 observations			
	bno ~ abs +(1|subF)	**$z = -3.83$, $p < .001$***	**.0026**
	bno ~ val + (1|subF) + (1|stimF)	$z = -2.20$, $p = .028$*	.0009
	bno ~ arous +(1 + arous|subF) + (1|stimF)	$z = 1.42$, $p = .157$	
	bno ~ like +(1|subF)	**$z = -2.48$, $p = .013$***	**.0010**
IB-free, outlier excluded, 1217 observations			
	bno ~ abs +(1|subF)	**$z = -2.19$, $p = .029$***	**.0027**
	bno ~ val + (1|subF)	$z = -0.98$, $p = .329$	
	bno ~ arous +(1 + arous|subF)	$z = 0.47$, $p = .640$	
	bno ~ like +(1|subF)	**$z = -1.78$, $p = .075$**	**.0015**

Note: Re-occuring significant effects and tendencies are printed in bold, significant fixed effects are denoted by *. We used the glmer.nb() function from the lme4 package. Effect sizes (estimated by r.squaredGLMM from the MuMIn package) for fixed effects are reported, when they were significant or marginal ($p < .10$). bno = number of blinks, abs = absorption, val = valence, arous = arousal, MW = mind wandering, like = liking; MW and groove were evaluated in SB-fix only; measures of abs, val, arous, like, MW, and groove were based on self-reports. Random effects: subF = subjects (factorized), stimF = stimulus (factorized).

mixed effects models can handle outliers rather well. Statistical results do not differ much between data sets including or excluding outliers, but, for IB-fix and IB-free experiments, including outliers increased the model complexity in several cases.

3.2.2 | Did differences in blink instructions between experiments affect felt absorption?

We reported that higher absorption was related to lower blink activity in SB-fix and IB-fix. Did then, the instruction to inhibit blinking increase absorption? If so, felt absorption should differ between experiments with spontaneous and inhibited blink instructions. However, felt absorption was on a similar level for all three experiments (SB-fix: $M = 3.49$, $SD = 0.53$, IB-fix: $M = 3.46$, $SD = 0.50$, IB-free: $M = 3.71$, $SD = 0.47$), confirmed by a one-factor between-subject ANOVA: $F(2) = 2.56$, $p = .083$, $\eta_p^2 = .053$ (outlier excluded: $F(2) = 0.71$, $p = .50$, $\eta_p^2 = .023$). There is no indication that the instructions confounded felt absorption.

3.2.3 | Do changes of self-reported valence and arousal by music presence predict changes in blink rate?

Differently from the evaluations of absorption and liking, we have collected valence and arousal ratings not only during music but also in silence. In all data sets, participants reported more positive valence during music listening than in silence, all t's ≤ 2.98, p's $\leq .0053$, d's ≥ 0.5042 and felt more aroused during music than silent trials, all t's ≤ 4.69, p's $< .0001$, d's ≥ 0.8563 (Figure 3; see Table S3 for complete statistics). That is, music had an effect on participants' felt emotional state. We can then explore whether changes in valence or arousal levels by music were related to the increase of blink rates. To do so, we calculated difference scores between measures with music and in silence of subject-based mean valence and arousal levels, as well as of blink rates. Neither the difference in valence, $rho = .02$, $p = .86$, nor in arousal, $rho = -.03$, $p = .77$, correlated with the difference in blink rates. Correlations remained non-significant when outliers were excluded (both p's $> .32$). Hence, there is no indication, that the change of blink rates by music presence was related to changes in valence or arousal states.

3.3 | Blinking and acoustic features

When adding music to silence, the added auditory stimuli contain information that has to be processed to some extent. One way to describe an auditory stimulus is by its acoustic properties, which we extracted from the digital files in an automated way. We next explore whether musical features relate to blink rates during listening to music. We applied generalized linear mixed effects models predicting blink counts by the 11 extracted musical features (see Section 2.3.2), with participants and stimuli as random effects, and, if possible, the musical feature as subject-specific slope. As some features were correlated, we set up separate models, one for each feature. However, none of the fixed effects were significant, all $|z|$'s < 1.2, p's $> .25$ for the complete data set, and all $|z|$'s < 1.4, p's $> .16$ when outliers were rejected (see Table S5 for detailed statistics). That is, even though blink counts increased with music present, in comparison with silence, no particular musical feature seems to consistently drive the number of blinks during music listening. However, one might argue that there are features not captured by our selection. Then, blink rates should be music-specific, that is the mean blink rates for the 56 stimuli should correlate between experiments (SB-fix, IB-fix, IB-free). But this was not the case, all $|rho|$'s $< .18$, p's $> .18$ for the complete data set, all $|rho|$'s $< .23$, p's $> .08$ when outliers were rejected (see Table S6 for detailed statistics). Hence, there is no indication that blink rates were driven by the acoustic features of the stimuli.

4 | GENERAL DISCUSSION

We measured spontaneous blink activity by video-based eye tracking during music listening and in silence. Our aims were to (i) replicate and further understand the blink rate increase for music presence vs. silence, (ii) explore the relationship between blinking and music-induced subjective states, (iii) assess the association between blinking and musical features, and (iv) provide other researchers with methodological advice for analyzing blink data. More specifically, we defined blinks and outliers by individual distributions, and conducted a multitude of analyses to illustrate the consequences of outlier rejection.

We replicated the effect of increased spontaneous blinking by music. This effect of music presence has been reported earlier (Hammerschmidt & Wöllner, 2018; Schäfer & Fachner, 2014), but did not occur using different auditory stimuli like simple auditory tones and words (Liu et al., 2020). However, studies differed in many further aspects, for example, visual stimuli on the screen (Hammerschmidt & Wöllner, 2018; Schäfer & Fachner, 2014) or not (Liu et al., 2020), instructions to centrally fixate (Liu et al., 2020; SB-fix, IB-fix) or not (Hammerschmidt & Wöllner, 2018; Schäfer & Fachner, 2014; IB-free). Comparisons thus have the potential to shed light on the question of why earlier findings diverged. If visual stimulation is the key for an effect of music presence, we should not have found an effect of music listening, but we did. If free viewing instead of central fixation was the key, we should have found an effect in IB-free but not in IB-fix and SB-fix, but the pattern was—if anything—the other way around. The comparisons of blink rates between conditions were significant for SB-fix and IB-fix, and not significant for IB-free including outliers, but marginal ($p < .10$) when outliers were excluded. These results perhaps imply the importance of

PSYCHOPHYSIOLOGY SPR

visual stimuli or fixation targets in eliciting a reliable effect of music presence. This argument is further supported by two studies comparing the effect of chill-inducing versus control music (Laeng et al., 2016; Laeng et al., 2021): Blink rates decreased by chill-inducing music when there was a visual scale on the screen (2021), but not when the screen was blank (2016). Importantly, music in comparison with silence had a rather consistent effect on blink activity in our data. Then, the question is what makes music special?

We were not able to find a reasonable answer to this question. We tested two lines of argumentation: Music has the power to induce emotional states and/or consists of specific acoustic properties which might alter blink activity. Notably, the effect of absorption into music on blink rates showed in all three data sets (in IB-fix the effect was reliable only when outliers were excluded). The effect of absorption mirrors the study showing decreased blink rates by chill-inducing in comparison with control music (Laeng et al., 2021; but see Laeng et al., 2016). However, absorption into music was related to decreased blink rates, showing the opposite effect from increased blinking by music presence. On the other hand, participants felt more positive and more aroused during music listening in comparison with silence, but the changes were not related to the increase of blink rates by music. That is, the increase in blink rates during music presence cannot be explained by the state of absorption induced by music, or changes in self-reported valence and arousal levels.

We then extracted 11 acoustic features that describe the mean characteristics of our musical excerpts (Lange & Frieler, 2018). The features spanned attributes such as brightness, chromaticism, beat strength, or tempo. None of the features were related to blink rate, showing that there is no simple relation between acoustics and blink rate. Moreover, comparisons across the three data sets showed no music-specific blink rates, as they did not correlate. Note that, while the diversity of our stimulus set offers many advantages, it is also possible that because our material comes from music of so many different styles, there may not be specific features that consistently signal certain types of events (e.g., phrase boundaries). In the future, others may find correlations if analyzing within a specific style.

The final aim of our investigation was to report and explore analysis methods for blink data. There are several issues that are worth noting. First, as no agreement exists across different researchers about how to define blinks by their duration (e.g., difference in the chosen cutoffs), we suggest to apply a data-driven approach to define blinks from binocular data loss periods. We excluded rough outliers (e.g., <50 ms and >2000 ms) and individualized the upper duration limit by $M_{\text{subject}} + 3\ SD$ (Figure 1a, upper row). The distribution of upper cutoffs peaked around 450–500 ms, corresponding to what has been applied in the literature in several studies (Aarts et al., 2012; Caffier et al., 2003; Jongkees & Colzato, 2016; Nakano et al., 2013; Rodriguez et al., 2018; Wang et al., 2011). But the broad range of individual differences in our data (Figure 1b, upper row) speaks to the individualized account. Note that our duration-based definition of blinks was applied to the detected, binocular data loss periods. Such data loss periods can be defined by custom algorithms (which are agnostic with respect to the type of eye tracker used) or by proprietary software offered by individual companies that produce the eye tracker being used (e.g., EyeLink DataViewer, Tobii Pro Lab, SMI BeGazeTM, etc.). These companies usually already define the output of their algorithms as blinks. Regardless of the chosen approach, we encourage researchers to report their blink definition algorithm and any relevant parameters that need to be set (e.g., including those applied to proprietary tools) in enough detail that others can replicate and conduct future meta-analyses.

Second, we suggest to identify outlier trials or participants by the distributions of the measure of interest, that is, mean blink rates. For outlier trials, the individual $M_{\text{subject}} \pm 3\ SD$ was a suitable cutoff. To define outlier participants, we related their mean blink rate to the group mean. High-blinking participants can be detected by the criterion $M_{\text{subject}} > M_{\text{group}} + 1\ SD$. In our data set, low-blinking subjects were not detected by $M_{\text{subject}} < M_{\text{group}} - 1\ SD$ (as in Nakano & Kitazawa, 2010), but rather by $M_{\text{subject}} - 1\ SD \leq 0$.

Third, we evaluated the consequences of removing trial and participant outliers for statistical results. The question is, whether less but cleaned data are preferable to more but noisier data. Some studies are silent about outliers (e.g., Aarts et al., 2012; Annerer-Walcher et al., 2018; Fukuda, 2001; Nomura et al., 2015; Recarte et al., 2008; Smilek et al., 2010; Veltman & Gaillard, 1996), indicating that outliers were not defined and rejected; other studies excluded them (Nakano et al., 2009). In our three studies, about 21%–38% of the data were defined as outliers, which is a big portion of the collected data to be rejected. Statistical tests on mean blink rates showed more conclusive results and had better power, when outliers were removed (e.g., Cohen's d in Table 2, Table S2). Similarly, in trial-based statistics such as mixed effects models, estimated effect sizes were slightly stronger when outliers were removed (Table 5). However, mixed effects models can take into account stimulus- and participant-specific variance (e.g., high- or low-blinking participants), and statistical results converged broadly for the data sets including and excluding outliers. Including outliers increased the number of observations and allowed for models of higher complexity in some data sets (IB-fix, IB-free). As a rule of thumb, 1600 observations have been suggested to have sufficient power in linear mixed effects models (Brysbaert & Stevens, 2018), which was reached only when outliers were included in our data.

Note that in the context of open science it has been suggested to explore the compatibility of results using different data sets derived from recorded measures (e.g., Steegen et al., 2016). Comparing analyses including or excluding outliers is an important contribution to such kind of multiverse analysis and might show diverging results questioning final conclusions (Verkoeijen et al., 2018). However, in our case we found highly converging results including or excluding outliers: increased blink rates by music (Table 2) and musical absorption predicting blink count (Table 5). In addition, for the latter finding we provide results from different modeling accounts in the Supporting Information (Table S4). Results converge largely, speaking strongly for the reliability of the effect of absorption. We noted a few cases where outliers seemed to matter. When plotting blink durations across experimental sessions (Figure 2, and Figure S2), histograms were clearly more homogeneous for the two sessions for durations <100 and >300 ms, when participant outliers were excluded, indicating that outlier blink durations are related to individual differences in mean blink rates. In addition, the periodicity in the dynamic blink signal seems to differ for high-blinking participants.

Fourth, blinks are rather rare events. As a result, trials with zero blinks occur. The instruction to inhibit blinking further promotes mean blink rates of zero. We evaluated how to treat such data with zero blink rates by different types of modeling accounts. Setting zero trials to NaN changes the blink rate distributions in a systematic and not desirable way. Setting zero to a small constant is possible but less elegant. Using models built for many zero cases, such as zero-inflated models, did not show significant fixed effects for the zero part of the model, indicating that these models were too complex. Therefore, we settled on applying simple mixed effects models to fit count data. For the spontaneous blink data, we arrived at an interesting observation. For these data, it was possible to fit random intercepts for the stimuli. The random intercept correlated with stimulus durations, and random intercepts for different models correlated with each other. That is, there was a systematic difference of the number of blinks due to stimulus duration, and this systematic difference was taken into account by the model structure.

The fact of many zero cases also has consequences for analyzing blink events as time series, for example, coding a blink onset as 1 in a vector representing the trial duration. The attentive researcher will notice that high correlations between such time series are then driven by the zero-samples and not the measure of interest (blinks). However, dynamic effects are interesting, because they can reveal differences between conditions, without driving trial means apart (e.g., Huber et al., 2023). We think future research exploring similarity in blink timing across participants would be highly interesting but would require much longer stimuli, and likely more participants or trials than in our studies, for appropriate statistical power. For further discussion regarding the analysis of blink time series of our data, please see Supporting Information, Section 8.

Fifth, we observed a rather constant blink rate across the time of a trial with the exceptions of the start and end of music listening. At the beginning, the higher blink rate leveled out at about 10–15 s after music onset, at the end, blink rate started to increase at about 5 s before the end of the musical pieces. The musical excerpts included composed endings, making the end of the stimulus foreseeable. Overall, the deviations are small, and the question might arise, whether the effect of music presence on mean blink rates can be attributed to those deviations at the beginning and end. However, this is highly unlikely, as IB-free showed a clear effect of beginning, but no clear effect of music presence. But the deviations in the beginning and end for the music condition might have inflated the effect of music presence on mean blink rate. Final conclusions on the importance of dynamic changes for an overall effect on music presence require solid replication, including the same number of trials for the music and silent condition, as well as matched lengths of the stimulus durations.

Sixth, for future studies, our findings clearly suggest that the silence condition is a proper control condition to understand blink activity during music listening (see also Huber et al., 2022). Mean blink rates between music and silence are correlated, showing a systematic relation (Figure 4, first row). The temporal structure of blinking is periodic for both conditions (Figure 4, last row). In time series analyses, one potential control is to randomly scramble the series. But simply randomizing the time series would wash out the demonstrated periodicity, which would clearly not be suitable.

Finally, we want to point out that when interested in blink activity, measuring spontaneous blinking is most suitable. This argument is promoted by the modeling results, showing the models of highest complexity (including random slopes and intercepts) fit for spontaneous but not inhibited blinking. However, there are certain conditions that require inhibition of blinking, for example, when recording EEG, blinking introduces artifacts. For such a setting it is reassuring that statistical results matched in our study broadly between inhibit and spontaneous blink conditions, at least for the models for absorption.

Nevertheless, our data indicate that instructions affect blink activity. The instruction to inhibit blinking decreased blink rates, and the additional instruction to fixate—that is to inhibit saccadic activity—decreased blink rates further (Table S2). For reflexive saccades, it has been shown that the instruction to inhibit physiological activity diminished the startle eye-blink modulation (Verschuere et al., 2007). In another study, event-related potentials (ERPs) were compared between recordings differing by blink instructions (Ochoa & Polich, 2000). The instruction "not to blink" changed the amplitude and latency of the P300 (a positivity about 300 ms after event onset) in comparison with no instruction, but there was no effect on earlier ERP components. This differential finding indicates that rather higher level than sensory processes were affected by blink inhibition. Related to this issue, it has been shown that visual demands (e.g., visual search) decrease blink rates (Recarte et al., 2008). That is, our instructions to inhibit blinking and saccades likely reduced blink rates by further adding higher-level processes or demands to the listening task. It is important to note that researchers have to be sensitive for potential effects of instructions, and a systematic investigation on how to best reduce blinking without increasing cognitive demands would be highly desirable for researchers interested in saccade and pupil tracking.

To sum, we think that analyzing blink activity with video-based eye trackers can be useful to explore emotional processes, and a variety of other questions. For example, blink activity is part of face-to-face communication (Hömke et al., 2018) and plays a role in social communication, even across species (Tada et al., 2013). Studies on blinking during music performance in concert settings might uncover hidden connectivity between performers and audience. In addition, eye-blink activity is related to concurrent simple motor processes, such as voluntary tapping (Cong Khac et al., 2010), or facial motor activity during speaking (Brych et al., 2021; but see Brych et al., 2020). As music can induce the urge to move (Madison, 2006; Stupacher et al., 2016), it seems obvious for music to affect blink activity. We measured this urge to move by asking for groove in SB-fix. Even though our results did not support this conclusion, further studies measuring groove in a more elaborated way, and over longer time periods, might be needed to shed light on this issue. Importantly, recording blinks is a non-intrusive measure, which can be covertly recorded via camera-based methods. Then, blink activity is less influenced by participants' intentions than more subjective measures, such as evaluations, or more obtrusive measures. Hence, it seems highly desirable to increase the amount of blink research. For researchers mainly interested in saccadic activity or fixations, eye-blinks as a dependent variable come for free (but at the expense of a disturbed pupil signal) and might reveal interesting findings. Therefore, we highly encourage researchers to analyze blink data and contribute to a better understanding of blink activity.

AUTHOR CONTRIBUTIONS

Elke B. Lange: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; resources; software; supervision; visualization; writing – original draft; writing – review and editing. **Lauren K Fink:** Conceptualization; methodology; writing – review and editing.

ACKNOWLEDGEMENTS

We thank Klaus Frieler, Methods' Specialist at the MPIEA, for very fruitful discussions on data analyses. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in osf at https://osf.io/65nyh.

ORCID

Elke B. Lange https://orcid.org/0000-0002-8928-3718

REFERENCES

Aarts, H., Bijleveld, E., Custers, R., Dogge, M., Deelder, M., Schutter, D., & van Haren, N. E. M. (2012). Positive priming and intentional binding: Eye-blink rate predicts reward information effects on the sense of agency. *Social Neuroscience*, *7*, 105–112. https://doi.org/10.1080/17470919.2011.590602

Annerer-Walcher, S., Körner, C., & Benedek, M. (2018). Eye behavior does not adapt to expected visual distraction during internally directed cognition. *PLoS One*, *13*(9), e0204963. https://doi.org/10.1371/journal.pone.0204963

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, *8*, 456. https://doi.org/10.3389/fpsyg.2017.00456

Bartoń, K. (2009). MuMIn: Multi-model inference (R package version 0.12.2).

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bettinardi, R.G. (2021). computeCohen_d(x1, x2, varargin). MATLAB Central File Exchange. Retrieved September 29, 2021, from https://www.mathworks.com/matlabcentral/fileexchange/62957-computecohen_d-x1-x2-varargin

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*, 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Brych, M., Händel, B. F., Riechelmann, E., Pieczykolan, A., & Huestegge, L. (2020). Effect of vocal demands on pupil dilation. *Psychophysiology*, *58*, e13729. https://doi.org/10.1111/psyp.13729

Brych, M., Murali, S., & Händel, B. (2021). How the motor aspect of speaking influences the blink rate. *PLoS One*, *16*(10), e0258322. https://doi.org/10.1371/journal.pone.0258322

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 9. https://doi.org/10.5334/joc.10

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, *89*, 319–325. https://doi.org/10.1007/s00421-003-0807-5

Cong Khac, D., Sharikadze, M., Staude, G., Deubel, H., & Wolf, W. (2010). Spontaneous eye blinks are entrained by finger tapping. *Human Movement Science*, *29*(1), 1–18. https://doi.org/10.1016/j.humov.2009.08.003

Cruz, A. A. V., Garcia, D. M., Pinto, C. T., & Cechetti, S. P. (2011). The ocular surface, *9*, 29–41. https://doi.org/10.1016/S1542-0124(11)70007-6

De Jong, P. J., & Merckelbach, H. (1990). Eyeblink frequency, rehearsal activity, and sympathetic arousal. *International Journal of Neuroscience*, *51*, 89–94. https://doi.org/10.3109/00207459009000513

Doughty, M. J., & Naase, T. (2006). Further analysis of the human spontaneous eye blink rate by a cluster analysis-based approach to categorize individual with 'normal' versus 'frequent' eye blink activity. *Eye & Contact Lens*, *32*, 294–299. https://doi.org/10.1097/01.icl.0000224359.32709.4d

Fukuda, K. (2001). Eye blinks: New indices for the detection of deception. *International Journal of Psychophysiology*, *40*, 239–245. https://doi.org/10.1016/S0167-8760(00)00192-6

Fukuda, K., Stern, J. A., Brown, T. B., & Russo, M. B. (2005). Cognitions, blinks, eye-movements, and pupillary movements during performance of a running memory task. *Aviation, Space & Environmental Medicine*, *76*(7, Suppl), C75–C85.

Geng, J. J., Ruff, C. C., & Driver, J. (2008). Saccades to a remembered location elicit spatially specific activation in human retinotopic visual cortex. *Journal of Cognitive Neuroscience*, *21*, 230–245. https://doi.org/10.1162/jocn.2008.21025

Goh, K.-I., & Barabási, A.-L. (2008). Burstiness and memory in complex systems. *Europhysics Letters*, *81*, 48002. https://doi.org/10.1209/0295-5075/81/48002

Goldstein, R., Bauer, L. O., & Stern, J. A. (1992). Effect of task difficulty and interstimulus interval on blink parameters. *International Journal of Psychophysiology*, *13*, 111–117. https://doi.org/10.1016/0167-8760(92)90050-L

Hammerschmidt, D., & Wöllner, C. (2018). The impact of music and stretched time on pupillary responses and eye movements in slow-motion film scenes. *Journal of Eye Movement Research*, *11*(2), 10. https://doi.org/10.16910/jemr.11.2.10

Hershman, R., Henik, A., & Cohen, N. (2018). A novel blink detection method based on pupillometry noise. *Behavior Research Methods*, *50*, 107–114. https://doi.org/10.3758/s13428-017-1008-1

Hömke, P., Holler, J., & Levinson, S. C. (2018). Eye blinks are perceived as communicative signals in human face-to-face interaction. *PLoS One*, *13*(12), e0208030. https://doi.org/10.1371/journal.pone.0208030

Huber, S. E., Martini, M., & Sachse, P. (2022). Patterns of eye blinks are modulated by auditory input in humans. *Cognition*, *221*, 104982. https://doi.org/10.1016/j.cognition.2021.104982

Huber, S. E., Martini, M., & Sachse, P. (2023). Task-synchronized eye blink modulation neither requires visual stimulation nor active motor response and is modulated by task predictability. *International Journal of Psychophysiology*, *187*, 1–10. https://doi.org/10.1016/j.ijpsycho.2023.01.011

Jongkees, B., & Colzato, L. S. (2016). Spontaneous eye blink rate as predictor of dopamine-related cognitive function – A review. *Neuroscience and Biobehavioral Reviews*, *71*, 58–82. https://doi.org/10.1016/j.neubiorev.2016.08.020

Kim, E.-K., & Jo, H.-H. (2016). Measuring burstiness for finite event sequences. *Physical Review E*, *94*, 032311. https://doi.org/10.1103/PhysRevE.94.032311

Kim, J., & André, E. (2008). Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*, 2067–2083. https://doi.org/10.1109/TPAMI.2008.26

Kleifges, K., Bigdely-Shamlo, N., Kerick, S. E., & Robbins, K. A. (2017). BLINKER: Automated extraction of ocular indices from EEG enabling large-scale analysis. *Frontiers in Neuroscience*, *11*, 12. https://doi.org/10.3389/fnins.2017.00012

Laeng, B., Eidet, L. M., Sulutvedt, U., & Panksepp, J. (2016). Music chills: The eye pupil as a mirror to music's soul. *Consciousness and Cognition*, *44*, 161–178. https://doi.org/10.1016/j.concog.2016.07.009

Lange, E. B., & Frieler, K. (2018). Challenges and opportunities of predicting musical emotions with perceptual and automatized features. *Music Perception*, *36*, 217–242. https://doi.org/10.1525/MP.2018.36.2.217

Laeng, B., Garvija, L., Løseth, G., Eikemo, M., Ernst, G., & Leknes, S. (2021). 'Defrosting' music chills with naltrexone: The role of endogenous opioids for the intensity of musical pleasure. *Consciousness and Cognition*, *90*, 103105.

Lange, E. B., Zweck, F., & Sinn, P. (2017). Microsaccade-rate indicates absorption by music listening. *Consciousness and Cognition*, *55*, 59–78. https://doi.org/10.1016/j.concog.2017.07.009

Liu, C. C., Ghosh Hajra, S., Cheung, T. P. L., Song, X., & D'Arcy, R. C. N. (2017). Spontaneous blinks activate the precuneus: Characterizing blink-related oscillations using magnetoencephalography. *Frontiers in Human Neuroscience*, *11*, 489. https://doi.org/10.3389/fnhum.2017.00489

Liu, C. C., Ghosh Hajra, S., Pawlowski, G., Fickling, S. D., Song, X., & D'Arcy, R. C. N. (2020). Differential neural processing of spontaneous blinking under visual and auditory sensory environments: An EEG investigation of blink-related oscillations. *NeuroImage*, *218*, 116879. https://doi.org/10.1016/j.neuroimage.2020.116879

Madison, G. (2006). Experiencing groove induced by music: Consistency and phenomenology. *Music Perception*, *24*, 201–208. https://doi.org/10.1525/mp.2006.24.2.201

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, *52*, 464–488. https://doi.org/10.3758/s13428-019-01246-w

Mathôt, S. (2013). A simple way to reconstruct pupil size during eye blinks (version 1). Figshare https://doi.org/10.6084/m9.figshare.688001

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS One*, *9*(2), e89642. https://doi.org/10.1371/journal.pone.0089642

Naicker, P., Anoopkumar-Dukie, S., Grant, G. D., Neumann, D. L., & Kavanagh, J. J. (2016). Central cholinergic pathway involvement in the regulation of pupil diameter, blink rate and cognitive function. *Neuroscience*, *334*, 180–190. https://doi.org/10.1016/j.neuroscience.2016.08.009

Nakamori, N., Odawara, M., Nakajima, T., Mizutani, T., & Tsubota, K. (1997). Blinking is controlled primarily by ocular surface conditions. *American Journal of Ophthalmology*, *124*, 24–30. https://doi.org/10.1016/s0002-9394(14)71639-3

Nakano, T. (2015). Blink-related dynamic switching between internal and external orienting networks while viewing videos. *Neuroscience Research*, *96*, 54–58. https://doi.org/10.1016/j.neures.2015.02.010

Nakano, T., Kato, M., Morito, Y., Itoi, S., & Kitazawa, S. (2013). Blink-related momentary activation of the default mode network while viewing videos. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 702–706. https://doi.org/10.1073/pnas.1214804110

Nakano, T., & Kitazawa, S. (2010). Eyeblink entrainment at breakpoints of speech. *Experimental Brain Research*, *205*, 577–581. https://doi.org/10.1007/s00221-010-2387-z

Nakano, T., Yamamoto, Y., Kitajo, K., Takahashi, T., & Kitazawa, S. (2009). Synchronization of spontaneous eyeblinks while viewing video stories. *Proceedings of the Royal Society B*, *276*, 3635–3644. https://doi.org/10.1098/rspb.2009.0828

Nomura, R., Hino, K., Shimazu, M., Liang, Y., & Okada, T. (2015). Emotionally excited eyeblink-rate variability predicts an experience of transportation into the narrative world. *Frontiers in Psychology*, *6*, 447. https://doi.org/10.3389/fpsyg.2015.00447

Ochoa, C. J., & Polich, J. (2000). P300 and blink instructions. *Clinical Neurophysiology*, *111*, 93–98. https://doi.org/10.1016/S1388-2457(99)00209-6

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, *47*, 560–569. https://doi.org/10.1111/j.1469-8986.2009.00947.x

R Core Team. (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Recarte, M. Á., Pérez, E., Conchillo, Á., & Nunes, L. M. (2008). Mental workload and visual impairmen: Differences between pupil, blink, and subjective rating. *The Spanish Journal of Psychology*, *11*, 374–385. https://doi.org/10.1017/S1138741600004406

Rizopoulos, D. (2022). GLMMadaptive: Generalized linear mixed models using adaptive gaussian quadrature. https://drizopoulos.github.io/GLMMadaptive/, https://github.com/drizopoulos/GLMMadaptive

Rodriguez, J. D., Lane, K. J., OuslerIII, G. W., Angjeli, E., Smith, L. M., & Abelson, M. B. (2018). Blink: Characteristics, controls, and relation to dry eyes. *Current Eye Research*, *43*, 52–66. https://doi.org/10.1080/02713683.2017.1381270

Roy, M., Mailhot, J.-P., Gosselin, N., Paquette, S., & Peretz, I. (2009). Modulation of the startle reflex by pleasant and unpleasant music. *International Journal of Psychophysiology*, *71*, 37–42. https://doi.org/10.1016/j.ijpsycho.2008.07.010

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn taking for conversation. *Language*, *50*, 696–735. https://doi.org/10.2307/412243

Saxena, S., Fink, L. K., & Lange, E. B. (2022). Deep learning models for webcam eye-tracking in online experiments. Manuscript Preprint. https://osf.io/42qej/

Schäfer, T., & Fachner, J. (2014). Listening to music reduces eye movements. *Attention, Perception, & Psychophysics*, *77*, 551–559. https://doi.org/10.3758/s13414-014-0777-1

Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010). Out of mind, out of sight: Eye blinking as indicator and embodiment of mind wandering. *Psychological Science*, *21*, 786–789. https://doi.org/10.1177/0956797610368063

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Stern, J. A., Boyer, D., & Schroeder, D. J. (1994). *Blink rate as a measure of fatigue: A review*. Civil Aerospace Medical Institute Tech Report Number DOT/FAA/AM-94/17.

Stupacher, J., Hove, M. J., & Janata, P. (2016). Audio features underlying perceived groove and sensorimotor synchronization in music. *Music Perception*, *33*, 571–589. https://doi.org/10.1525/mp.2016.33.5.571

Sweeney, D. F., Millar, T. J., & Ryju, S. R. (2013). Tear film stability: A review. *Experimental Eye Research*, *117*, 28–38. https://doi.org/10.1016/j.exer.2013.08.010

Tada, H., Omori, Y., Hirokawa, K., Ohira, H., & Tomonaga, M. (2013). Eye-blink behaviors in 71 species of primates. *PLoS One*, *8*(5), e66018. https://doi.org/10.1371/journal.pone.0066018

Tomic, S. T., & Janata, P. (2008). Beyond the beat: Modeling metric structure in music and performance. *The Journal of the Acoustical Society of America*, *124*, 4024–4041. https://doi.org/10.1121/1.3006382

Van Ordern, K. F., Jung, T.-P., & Makeig, S. (2000). Combined eye activity measures accurately estimate changes in sustained visual task performance. *Biological Psychology*, *52*, 221–240. https://doi.org/10.1016/S0301-0511(99)00043-5

VanderWerf, F., Brassinga, P., Reits, D., Aramideh, M., & Ongboer de Visser, B. (2003). Eyelid movements: Behavioral studies of blinking in humans under different stimulus conditions. *Journal of Neurophysiology*, *89*, 2784–2796. https://doi.org/10.1152/jn.00557.2002

Veltman, J. A., & Gaillard, A. W. K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology*, *42*, 323–342. https://doi.org/10.1016/0301-0511(95)05165-1

Verkoeijen, P. J. L., Polak, M. G., & Bouwmeester, S. (2018). A practical illustration of methods to deal with potential outliers: A multiverse outlier analysis of study 3 from Brummelman, Thomaes, Orobio de Castro, Overbeek, and Bushman (2014) Collabra. *Collabra: Psychology*, *4*(1), 30. https://doi.org/10.1525/collabra.118

Verschuere, B., Crombez, G., Koster, E. H. W., Van Bockstaele, B., & De Clerq, A. (2007). Startling secrets: Startle eye blink modulation by concealed crime information. *Biological Psychology*, *76*, 52–60. https://doi.org/10.1016/j.biopsycho.2007.06.001

Vjästfjäll, D. (2002). Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae*, *5*(1_suppl), 173–211. https://doi.org/10.1177/10298649020050S107

Wang, Y., Toor, S. S., Gautam, R., & Henson, D. B. (2011). Blink frequency and duration during perimetry and their relationship to test–retest threshold variability. *Investigative Ophthalmology*

& Visual Science, *52*, 4546–4550. https://doi.org/10.1167/iovs.10-6553

Warrenburg, L. A. (2020). Choosing the right tune: A review of music stimuli used in emotion research. *Music Perception*, *37*, 240–258. https://doi.org/10.1525/mp.2020.37.3.240

Xu, T. L., de Barbaro, K., Abney, D. H., & Cox, R. F. A. (2020). Finding structure in time: Visualizing and analyzing behavioral time series. *Frontiers in Psychology*, *11*, 1457. https://doi.org/10.3389/fpsyg.2020.01457

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supporting information S1. The Supporting Information includes results of additional analyses, Tables S1–S8, and Figures S1–S8.

How to cite this article: Lange, E. B., & Fink, L. K. (2023). Eye blinking, musical processing, and subjective states—A methods account. *Psychophysiology*, *00*, e14350. https://doi.org/10.1111/psyp.14350