Enhancing tonic arousal improves voluntary but not involuntary attention in humans

Aurélie Grandjean[*,1], Anne Mathieu[*,1], Sophie Chen[1], Andreas Widmann[2,3], Nicole Wetzel[3,4], Aurélie Bidet-Caulet[1,5]

* Equally contributing authors

[1] Aix Marseille Univ, Inserm, INS, Inst Neurosci Syst, Marseille, France

[2] Leipzig University, Wilhelm Wundt Institute for Psychology, Leipzig, Germany

[3] Leibniz Institute for Neurobiology, Magdeburg, Germany

[4] University of Applied Sciences Magdeburg-Stendal

[5] Université Claude Bernard Lyon 1, INSERM, CNRS, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, F-69500, Bron, France

Abstract

Arousal and attention are pivotal brain functions for optimizing performance. Kahneman's attention model (1973) theorizes a key interplay between attention and arousal, yet this relationship remains poorly understood. We investigated this interaction in 16 healthy young adults performing an auditory attention task that simultaneously assessed phasic arousal, voluntary attention and involuntary attention. Tonic arousal was modulated by low or high arousing music, as measured using skin conductance, pupil size, and heart rate. Pupil dilation responses to distracting sounds highlight an intricate interplay between tonic and phasic arousal. Importantly, increasing tonic arousal does not influence involuntary attention, whereas it does improve voluntary attention, as shown by shorter and less variable reaction times and larger electroencephalographic brain responses to task-relevant targets. This study provides experimental evidence in humans that tonic arousal can influence the attentional balance by improving voluntary attention in a transient and sustained manner, rather than by impacting involuntary attention.

Funding

This work was supported by the French National Research Agency [ANR-19-FRAL-0007-01 & ANR-16-CONV-0002 (ILCB)] and the German Research Foundation (DFG WE5026/4-1).

Acknowledgements

We thank Victor Collombat for support in piloting the experiment and Benjamin Morillon for advice on the framing of the manuscript.

1. INTRODUCTION

Crucial for the selection and prioritization of sensory information, attention enables individuals to navigate their complex surroundings efficiently. It is a multifaceted cognitive process, manifesting in two primary forms: voluntary and involuntary attention. The former facilitates focused engagement with specific tasks, encompassing processes such as voluntary orienting, anticipation, and sustained attention. In contrast, involuntary attention is elicited by salient, unexpected stimuli outside the intended focus. The intricate balance between these two aspects is essential for an optimal attentional control. Over the last 50 years, several theoretical and physiological models of attention have been proposed[1–6], many of them emphasizing the interplay between attention and arousal. Arousal is generally described as the state of physiological reactivity of a subject[1,4,7–9], and more recently as a general state of excitation of the cortex, modulated by the activity of the Locus Coeruleus – Norepinephrine (LC-NE) system[10,11].

Since Yerkes and Dodson (1908)[12] pioneer works in mice, an inverse U-shape relationship between arousal and performance has been proposed, with moderate levels of arousal promoting maximal performance, while lower and upper levels would lead to poorer performance. This inverse U-shape relationship was later associated to visual voluntary attention by Easterbrook (1959)[13] in a review of behavioral findings: at a moderate level of arousal, attention would be focused on relevant information while irrelevant one could be ignored. In hypo- or hyper-arousal states, relevant information would become indistinguishable from irrelevant information. Importantly, the optimal arousal level would be higher for simple than for complex tasks. In 1973, Kahneman[4] proposed a comprehensive capacity model of attention encompassing both transient (i.e., phasic) and sustained (i.e., tonic) arousal. This model suggests that available attentional capacities and arousal can influence each other to achieve optimal performance. This model also considers that the orienting response to novel stimuli encompasses both involuntary allocation of attention to the novel and a transient increase in arousal (see also [8] for a similar proposal). Later on, based on neuropsychological observations, Mesulam (1981)[5] proposed a strong influence of the reticular system underlying arousal on both sensory and higher-level brain regions in the frontal and parietal cortices involved in attention, a model regularly refined over the following decades[2,3,6].

These hypotheses were confirmed by experimental evidence, mostly in rodents and monkeys, showing widespread projections from the LC-NE system to the cerebral cortex, the cerebellum, and subcortical structures (thalamus, amygdala)[15]. Conversely, attention control has been shown to be supported by brain networks including frontal and parietal regions[3,6], with descending connections to the locus coeruleus[2]. Animal studies have demonstrated that neurons in the LC present two different modes of activity: a tonic

mode, that defines the baseline activity of the system and changes relatively slowly, and a phasic mode during which neurons fire at a high frequency (10–15 spikes per second) in response to a salient stimulus[16]. Tonic firing of LC neurons is relatively low when the organism is at rest, medium when engaged in a task requiring the filtering of irrelevant information, and high when the exploration of the environment requires to stay attentive to potentially unexpected occurring events (for a review see [10]). Therefore, experimental findings confirm that the LC-NE system presents both phasic and tonic activities which can modulate or be modulated by signal from cortical regions supporting attention. However, the interplay between attention and arousal remains poorly understood, in particular in humans, and is still at the heart of recent influential theories of task performance and stimuli processing[10,14,15].

Involuntary allocation of attention to novel stimuli can be behaviorally measured by a cost in performance (longer reaction times and reduced accuracy), named distraction effect[17]. However, an increasing number of studies have reported improved performance when novel sounds are presented[18,19], particularly in the context of highly arousing emotional stimuli[20]. This facilitation effect has been attributed to an increase in phasic arousal triggered by the distracting sound[21,22], in line with attention models of the orienting response[4,8] and animal studies showing an increase in LC-NE activity in response to salient stimuli[23–26]. At the cerebral level, novel unexpected sounds elicit a distinctive sequence of event-related potentials (ERPs): the N1/Mismatch Negativity around 100ms, followed by a P3 complex around 300ms (also called novelty-P3 or P3a)[8,27] . The P3 complex is commonly thought to reflect the involuntary orienting of attention towards distracting stimuli[27,28]. Moreover, this P3 complex is dissociable in two phases: a fronto-central early phase peaking around 235ms called early-P3 and a fronto-parietal late phase peaking around 320ms called late-P3[29]. Recently, the amplitude of the early-P3 was found strongly correlated to the arousal rating of distracting sounds[22,30], suggesting that the early-P3 would reflect the arousal component of the orienting response. In the last few years, an increasing number of studies used pupillometry recordings to investigate attentional processes and showed that novel oddball sounds cause a transient dilation of the pupil higher for arousing emotional stimuli, presumably due to the co-activation of the sympathetic nervous and the LC-NE systems[30].

Several experimental studies have investigated the link between tonic arousal and voluntary attention. The influential Aston-Jones theory of the LC-NE function in cognition relies mainly on animal studies. They were able to record LC neuron activity using electrophysiology in monkeys, and correlate levels of tonic and phasic activity with performance during visual attention tasks[31–33]. They confirmed the Yerkes and Dodson inverse U-shape relationship with moderate levels of tonic arousal promoting strong phasic activity in the LC-NE and maximal performance; and lower and upper levels resulting in small or no phasic

LC-NE activity and poorer performance. In the human literature, for technical reasons, evidence is much rarer and mainly in the visual modality. Tonic arousal modulations are achieved using a wide range of methodologies: pharmacological manipulations[34], stress induction[35], meditation[36], isometric exercise[37] or background music presentation[38–40]. In addition, some behavioral studies have used pre-task music to induce tonic arousal modulation, with varying type and duration of music extracts[41,42]. Because arousal level in human cannot be assessed directly by measuring the activity of neurons within the LC, most studies have relied on subjective ratings of the subject arousal state[43] or on indirect physiological markers such as heart rate or salivatory cortisol levels[35]. The obtained behavioral results are quite mixed, with for example relaxing (i.e. low-arousing) music being beneficial to the performance of a Stroop test[40]; while relaxing music impairs performance in a visuo-spatial flanker task[38]. Very few studies have investigated the link between tonic arousal, phasic arousal and attention at the brain level[35,37]. Interestingly, Mather et al. (2020)[37] conducted a fMRI study in which isometric exercises were used to modulate tonic arousal shortly before an auditory oddball task. Handgrip was found to decrease tonic pupil size during the subsequent task while increasing phasic pupil responses to oddball sounds, increasing activation of the frontoparietal network and speeding oddball detection, suggesting that reducing tonic arousal increases phasic arousal and activation of the brain network of voluntary attention.

Therefore, the interplay between arousal and attention remains largely unexplored, in particular in humans. To our knowledge, the effect of tonic or phasic arousal levels on the balance between voluntary and involuntary attention has not been directly addressed. In this study, we propose to unravel the complex relationship between voluntary attention, involuntary attention, tonic and phasic arousal. For this purpose, we used the Competitive Attention Task (CAT) which provides measures of voluntary and involuntary attention processes and their interplay[21]. In the CAT, voluntary attention orienting is modulated by modifying the information provided by a cue preceding a target sound to detect. Involuntary attention and phasic arousal are triggered by rare unexpected task-irrelevant sounds (distracting sounds) presented at random times between the cue and the target (**Fig. 1**). Phasic arousal is further modulated by varying the arousing content of the distracting sounds. Low and moderate levels of tonic arousal are induced by presenting calm or dynamic short music tracks, respectively, to the participants before each block of the CAT. The level of tonic arousal is indirectly assessed using peripheral measures of autonomic nervous system activity such as the heart rate, skin conductance and pupil size[44,45]. The pupil dilation response (PDR) to the different events of the CAT is used as an indirect measure of phasic arousal[11,46,47], while EEG measures are used to assess brain processes involved in voluntary and involuntary attention[21]. According to theorical models of attention and previous experimental findings, we expect that increasing

tonic arousal from a low to a moderate level to improve voluntary attention, to reduce involuntary attention triggered by distracting sounds and to increase the phasic arousal response to these same distracting sounds. We also speculate that reaching a moderate level of tonic arousal would impact the attentional balance by favoring voluntary over involuntary attention.

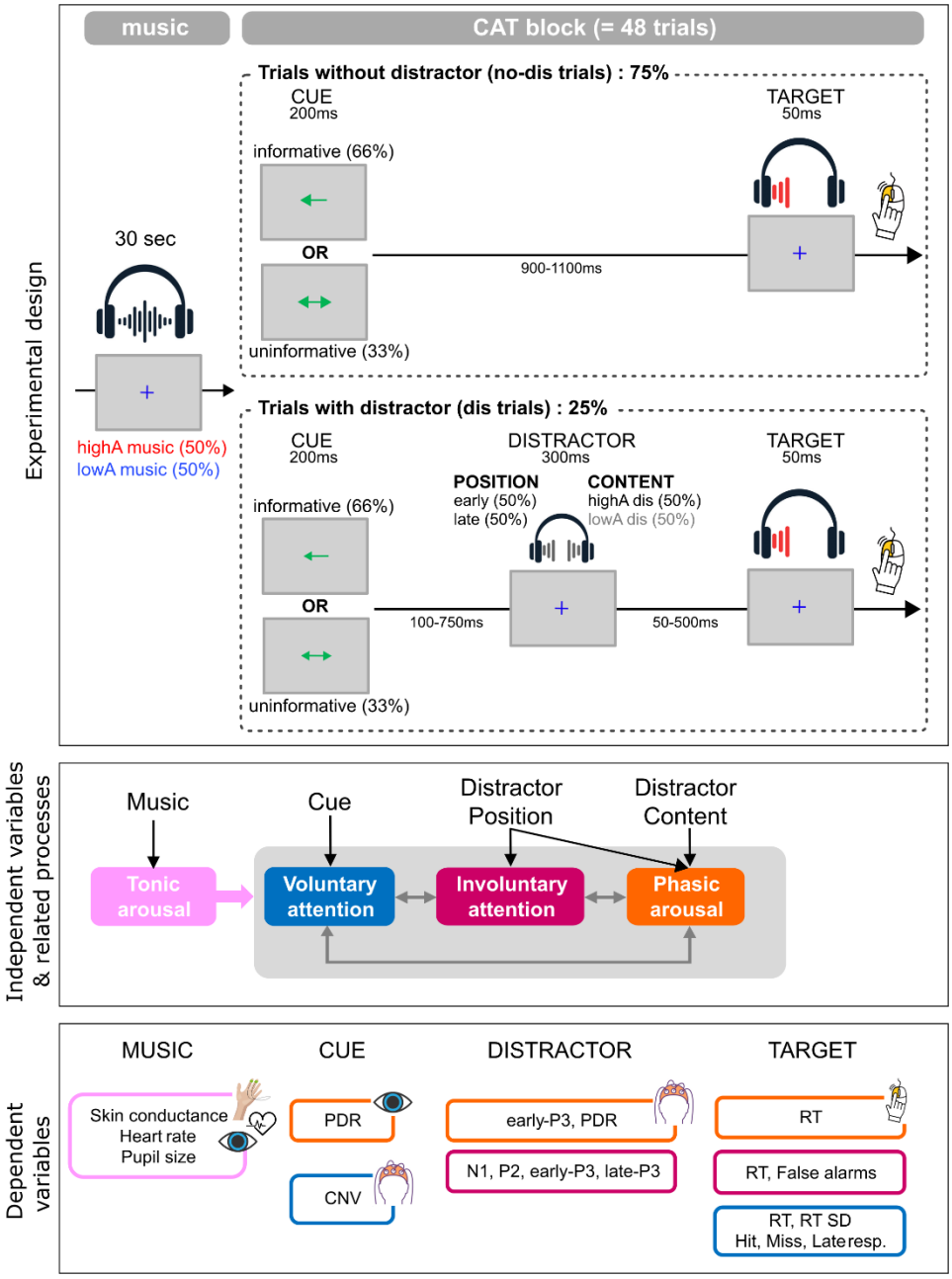

Fig 1. Protocol. **a.** Experimental design. Each block of the Competitive Attention Task (CAT; 48 trials) is preceded by 30 sec of music (**high arousing**: **highA music**, or **low arousing**: **lowA music**). In trials without distractor (no-dis, *top*), a cue (**informative** or **uninformative**) is first presented on the screen, and after a delay (between 900 and 1100ms) a target sound is played in the left or right ear. In trials with distractor (*bottom*), a distractor sound is presented binaurally in the delay between the cue and the target sound. Distractors presented between 250 and 500ms before target onset

cue are called **early distractors**. Distractors presented between 50 and 250ms before target onset are called **late distractors**. These distracting sounds can be either **high arousing** (**highA dis**), or **low arousing** (**lowA dis).** Subjects were asked to press the mouse button as fast as possible when they hear the target sound. **b.** Independent variables and related brain processes. High and low arousing **music** were used to modulate the level of *tonic arousal*. Informative compared to uninformative **cues** increase *voluntary attention* to the task. Distractors trigger both *involuntary attention* and an increase in *phasic arousal* that can be dissociated according to the **distractor position**. The arousing **distractor content** allows further modulation of *phasic arousal*. This design allows investigation of the interaction between involuntary attention, voluntary attention and phasic arousal, as well as the impact of tonic arousal on these interactions. c. Dependent variables. The impact of the independent variables was investigated on different physiological or behavioral measures locked to the different events of the paradigm and associated with distinct brain processes (see color correspondence with panel b). CNV: contingent negative variation, PDR: pupil dilation response, RT: reaction time, RT SD: reaction time standard deviation.

3. RESULTS

The impact of the tonic arousal level, induced by presenting low arousing (lowA) or high arousing (highA) music before each block of the Competitive Attention Task (CAT, Fig. 1), was explored on peripheral measures (skin conductance, pupil size and heart rate), as well as on performance, event-related pupil dilation response (PDR) and event-related potential (ERP) during the CAT. We used Bayesian Wilcoxon tests when only the effect of music was tested. We used linear mixed models (LMM) or generalized linear mixed models (GLM) to estimate the effect of the independent variables (Music, Cue, Distractor Position, Distractor Content). The model characteristics are presented in Table 1 and the full results in Supplementary Results).

3.1. Effect of music on peripheral indices of arousal

To validate the modulation of the tonic arousal level, we explored the effect of the music type on peripheral measures (Fig. 2 & 3). We found moderate to strong evidence for an effect of music on skin conductance (W = 113.0, BF_{10} = 5.150), pupil size (W = 31.0, BF_{10} = 4.104) and heart rate (W = 118.0, BF_{10} = 18.538). These measures were found larger during the listening of high arousing music compared to low arousing music (mean ± SEM: change of skin conductance, highA: -1.0 ± 2.4 nS & lowA: -7.2 ± 2.0 nS; change of pupil size, highA: 0.024 ± 0.063 mm & lowA: -0.042 ± 0.056 mm; heart rate, highA: 70.21 ± 2.28 bpm & lowA: 69.52 ± 2.27 bpm). As expected, all physiological measures are enhanced during listening to high arousing music compared to low arousing music, validating the tonic arousal modulation.

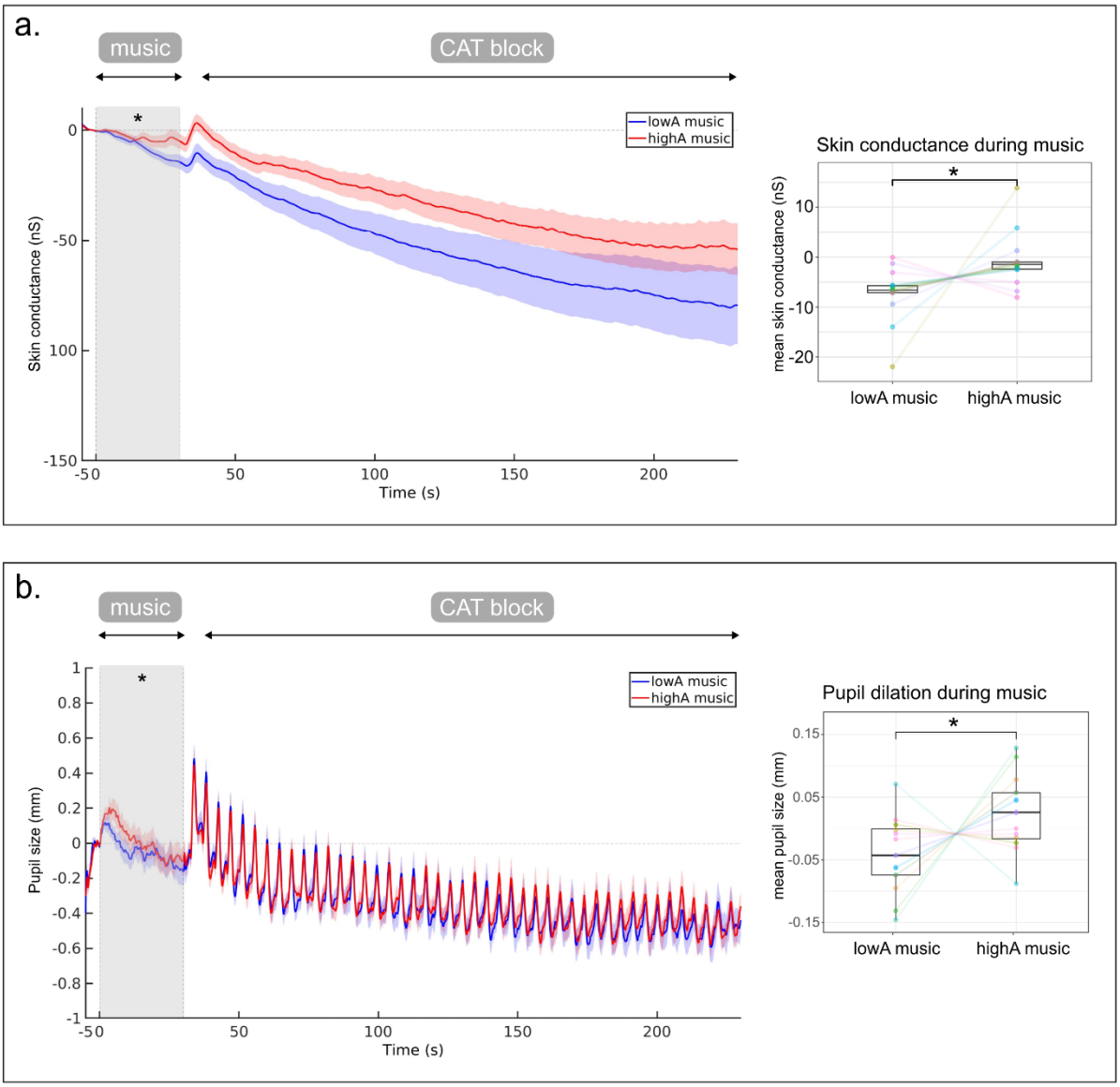

Fig 2. Physiological measures as a function of music type. **a.** Skin conductance during music presentation and the following CAT block (*left*) as a function of music type (blue: lowA music, red: highA music). Mean skin conductance during the music as a function of music type (*right*). **b**. Pupil size during music presentation and the following CAT block (*left*) as a function of music type (blue: lowA music, red: highA music). Mean pupil size during the music as a function of music type (*right*). On the time course plots, shadowed areas represent standard errors of the mean. Boxplots were plotted using normalized individual data (= raw data - subject mean + group mean). Within each boxplot, the horizontal line represents the group median, the lower and upper hinges correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (IQR = inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge. Superimposed to each boxplot, the colored dots represent individual means. * $p < 0.05$. lowA: low arousing, higA: high arousing.

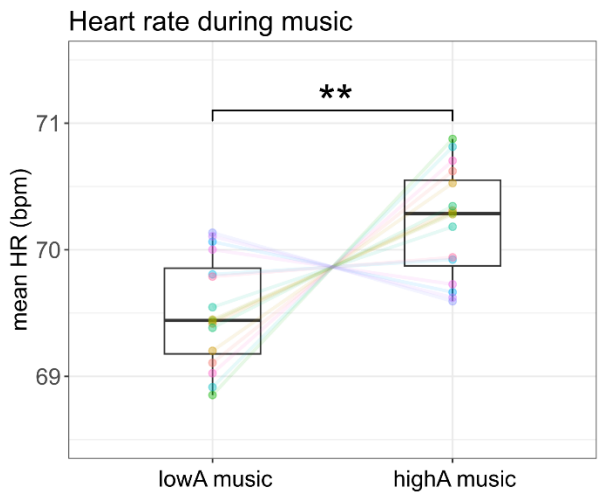

Fig 3. Mean heart rate during the music as a function of music type. Boxplots were plotted using normalized individual data (= raw data - subject mean + group mean). Within each boxplot, the horizontal line represents the group median, the lower and upper hinges correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (IQR = inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge. Superimposed to each boxplot, the colored dots represent individual means. ** $p < 0.01$. lowA: low arousing, higA: high arousing.

3.2. Behavioral measures during the Competitive Attention Task

3.2.1. Response Type Rates

Participants performed well the task, with mean rates of hits, missed responses, false alarms and late responses of 91.92% (SEM = 1.67%), 0.62% (SEM = 0.11%), 7.43% (SEM = 1.6%) and 0.03% (SEM = 0.02%), respectively. We found weak to moderate evidence against an effect of music on the rate of hits (W = 55.0, BF_{10} = 0.267), misses (W = 67.0, BF_{10} = 0.300), false alarms (W = 55.5, BF_{10} = 0.273) and late responses (W = 1.0, BF_{10} = 0.361). Response rates are not modulated by tonic arousal.

	Dependent variable	Time-window (ms)	trials	Elec	Model specification	conditional R^2	marginal R^2	Significant main effects	Significant interactions
single-trial analysis	**RT+**		all		logRT ~ DisPosition3 * Cue * Music + (DisPosition3 + Cue + Music | Participant)	0,171	0,031	DisPosition3, Cue	DisPosition3 x Cue Cue x Music DisPosition3 x Cue x Music
	RT+		dis		logRT ~ DisPosition2 * DisContent * Cue * Music + (DisPosition2 + DisContent + Cue + Music | Participant	0,221	0,05	DisPosition2, Cue	DisPosition2 x Cue x Music
trials averaged per subject	mean **PDR cue**-locked	500-1000	no-dis		mean_amplitude ~ Cue * Music + (Cue + Music | Participant)	0,887	0,018	Cue	*none*
	mean **PDR cue**-locked	1000-1500	no-dis		mean_amplitude ~ Cue * Music + (Cue + Music | Participant)	0,959	0,019	Cue	*none*
	mean **PDR cue**-locked	1500-2000	no-dis		mean_amplitude ~ Cue * Music + (Cue + Music | Participant)	0,978	0,006	Cue	*none*
	mean **PDR cue**-locked	2000-2500	no-dis		mean_amplitude ~ Cue * Music + (Cue + Music | Participant)	0,934	0,003	*none*	*none*
	mean **PDR dis**-locked	300-800	dis		mean_amplitude ~ DisPosition2 * DisContent * Cue * Music + (DisPosition2 + DisContent + Music | Participant)	0,282	0,075	DisPosition2, DisContent	Cue x Music
	mean **PDR dis**-locked	950-1450 (early-dis) 1100-1800 (late-dis)	dis		mean_amplitude ~ DisPosition2 * DisContent * Cue * Music + (DisPosition2 + DisContent + Music | Participant)	0,494	0,074	DisPosition2, DisContent	*none*
	mean **ERP cue**-locked	650-850	no-dis	Cz	mean_amplitude ~ Cue * Music + (Cue + Music | Participant)	0,888	0,006	*none*	*none*
	mean **ERP cue**-locked	850-1050	no-dis	Cz	mean_amplitude ~ Cue * Music + (Cue + Music | Participant)	0,943	0,006	*none*	*none*
	mean **ERP target**-locked	110-150	no-dis	FCz	mean_amplitude ~ Cue * Music + (Cue + Music | Participant)	0,967	0,003	*none*	*none*
	mean **ERP target**-locked	190-230	no-dis	CPz	mean_amplitude ~ Cue * Music + (Cue | Participant)	0,977	0,004	Music	*none*
	mean **ERP target**-locked	250-500	no-dis	CPz	mean_amplitude ~ Cue * Music + (Cue | Participant)	0,957	0,011	Cue	*none*
	mean **ERP dis**-locked	90-130	dis	FCz	mean_amplitude _amp ~ Music * Cue * DisPosition2 * DisContent + (Music*Cue + Cue*DisPosition2 + DisPosition2*DisContent + Cue*DisContent + Music*DisPosition2 + Music*DisContent|Participant)	0,801	0,024	DisPosition2	DisContent x DisPosition2 x Music
	mean **ERP dis**-locked	160-200	dis	Cz	mean_amplitude ~ Music * Cue * DisPosition2 * DisContent + (Music*Cue + Cue*DisPosition2 + DisPosition2*DisContent + Cue*DisContent + Music*DisPosition2 + Music*DisContent | Participant)	0,881	0,028	DisPosition2, DisContent	*none*
	mean **ERP dis**-locked	220-260	dis	FCz	mean_amplitude ~ Music * Cue * DisPosition2 * DisContent + (Cue + DisPosition2 + DisContent | Participant)	0,761	0,049	DisPosition2, DisContent	Music x DisPosition2 x DisContent
	mean **ERP dis**-locked	320-360	dis	Fz (& CPz)	mean_amplitude ~ Music * Cue * DisPosition2 * DisContent + (Music + Cue + DisPosition2 + DisContent | Participant)	0,535 (CPz: 0,516)	0,075 (CPz : 0,094)	DisPosition2 (CPz: DisContent)	*none*

Table 1. Details of the statistical models for each analysis. For each dependent variable, are provided details of the mixed model (fixed within-subject factors and random factors), information on model performance (marginal and conditional R^2 when available) and results of the statistical analyses (significant main effect and significant interactions). Factors levels: Music (2 levels: highA music, lowA music), Cue (2 levels: informative, uninformative), Dis Position2 (2 levels: early, late), Dis Position3 (3 levels: no-dis, early, late), and Dis Content (2 levels: highA dis, lowA dis). RT+: positive RT.

3.2.2. Reaction Times

Reaction times (RT) were fitted to a linear mixed model with Music, Cue and DisPosition3. Consistent with previous works[21,48,49], RT were significantly modulated by the type of cue (χ^2 (1) = 45.07, $p < 0.001$) and the position of the Distractor (χ^2 (2) = 63.97, $p < 0.001$). RT were shorter in the informative cue condition (mean = 235.6 ms, SEM = 2.7) compared to the uninformative cue condition (mean = 249.4 ms, SEM = 3.4). RT were shorter in the early dis condition (mean = 222.6 ms, SEM = 4.2) compared to the late dis condition (mean = 250.8 ms, SEM = 4.0) and to the no-dis condition (mean = 241.7 ms, SEM = 2.9).

Importantly, significant interactions were also revealed: DisPosition3 x Cue (χ^2 (2) = 8.53, $p < 0.05$), Cue x Music (χ^2 (2) = 4.08, $p < 0.05$) and DisPosition3 x Cue x Music (χ^2 (2) = 8.02, $p < 0.05$). Post-Hoc HSD tests were conducted to analyze this triple interaction (Fig. 4 a & b). In the highA music condition, the cue effect was significant in the no-dis and early dis conditions and close to significance in late dis; while in the lowA music condition, the cue effect was only significant in the no-dis and late dis conditions. The cue effect on RT is larger in trials with an early distractor after listening to high arousing music, suggesting a larger effect of voluntary attention under higher tonic arousal.

Then, in order to test the effect of the arousing content of distractor in trials with a distractor, RT were fitted to a linear mixed model with Music, Cue, DisPosition2 and DisContent. RT were significantly modulated by the cue (χ^2 (1) = 8.61, $p < 0.01$) and the position of the distractor (χ^2 (1) = 49.98, $p < 0.01$). No significant effect of the DisContent was observed. The test also revealed a significant Music x Cue x DisPosition2 interaction (χ^2 (1) = 5, $p < 0.05$). As shown in the previous analysis, the cue effect was significant in the early dis condition only after highA music. No significant effect of the arousal content of the distracting sound is observed on RT.

Distracting sounds trigger two effects on RT: the facilitation phasic arousal effect (median RT in no-dis minus median RT in early dis) and the detrimental involuntary attention effect (median RT in late dis minus median RT in early dis). To specifically test the impact of the music on these effects, Bayesian repeated measures analysis of variance were performed with the factors music and cue. We found that the null model is the best model, with moderate evidence for no effect of music on the phasic arousal (no-dis – early dis) effect ($BF_{inclusion}$ = 0.292) and on the involuntary attention (late dis – early dis) effect ($BF_{inclusion}$ = 0.282). There is no effect of tonic arousal on the behavioral measure of involuntary attention and phasic arousal.

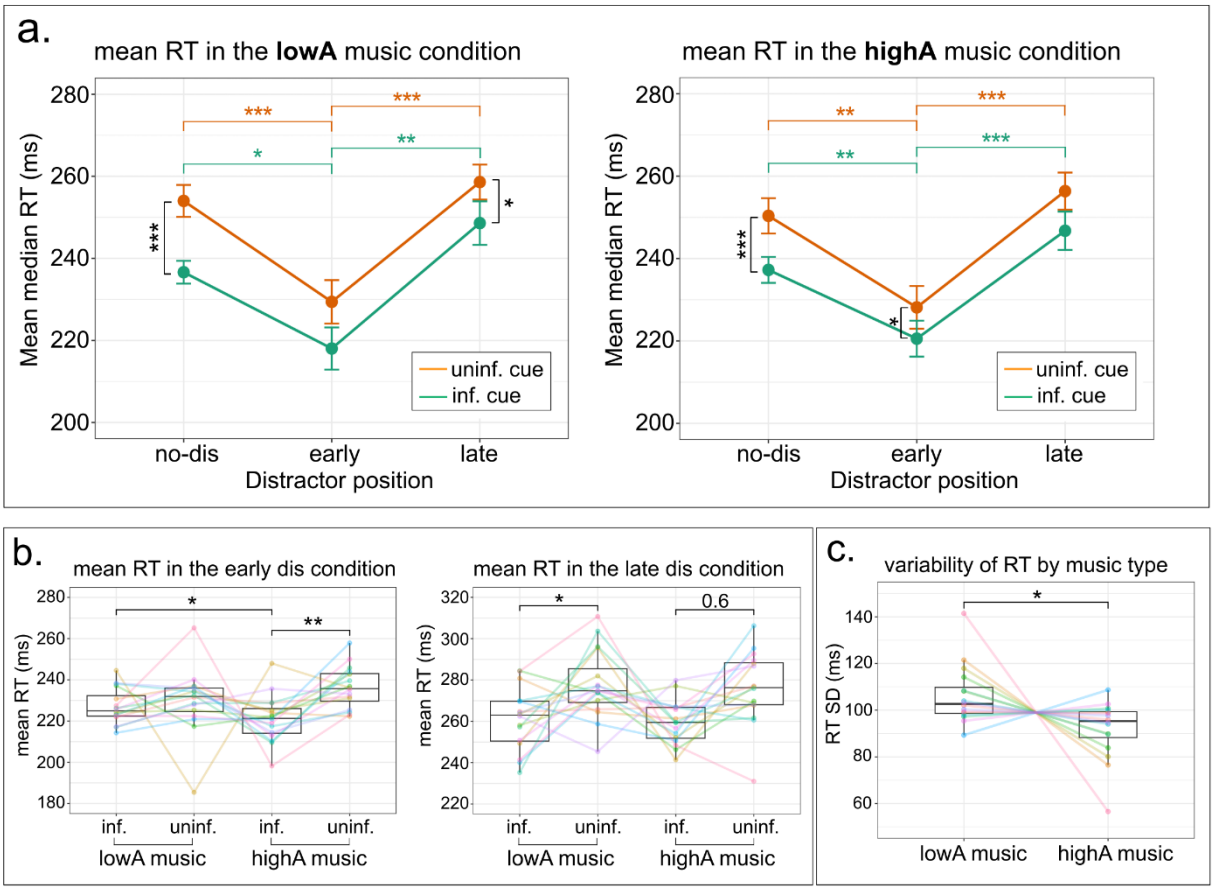

Fig 4. Behavioral results. **a**. Mean RT by distractor position and cue type, in lowA music condition (*left*) and highA music condition (*right*). Each point corresponds to values averaged across subjects. Within-subject standard errors of the mean are plotted around group means. **b**. Mean RT by music and cue types, in trials with an early (*left*) or late (*right*) distractor. **c.** RT variability (SD: standard deviation) by music type. Boxplots were plotted using normalized individual data (= raw data - subject mean + group mean). Within each boxplot, the horizontal line represents the group median, the lower and upper hinges correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (IQR = inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge. Superimposed to each boxplot, the colored dots represent individual means. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. lowA: low arousing, higA: high arousing, inf: informative, uninf: uninformative, dis: distractor.

To investigate sustained attention, we explored the effect of the music type on the variability of RT (standard deviation of RT) in trials with no distractor. We found moderate evidence for an effect of music ($W = 108$, $BF_{10} = 3.9$) with a reduced RT variability in the highA music condition (mean = 91.8ms, SEM = 8.9ms) compared to the lowA music condition (mean = 106.2ms, SEM = 10.2ms) as shown in Fig. 4c. The variability of RT is reduced after listening to high arousing music, suggesting improved sustained attention under higher tonic arousal

3.3. Physiological measures during the Competitive Attention Task

3.3.1. PDR to the cue

To test the impact of music and cue types on phasic arousal to task-relevant stimuli, mean amplitudes of the PDR to the cue in trials with no distracting sound were analyzed in four time-windows following cue onset (500 – 1000 ms; 1000 – 1500 ms; 1500 – 2000 ms; 2000 – 2500 ms). No significant effect of music was found. A significant effect of the cue on the PDR amplitude was found in time-windows one to three from 500 to 2000 ms post-cue (χ^2 (1) = 6.32, $p < 0.05$; χ^2 (1) = 19.30, $p < 0.001$; χ^2 (1) = 13.40, $p < 0.001$, respectively). In these time-windows, the mean amplitude of the PDR following a cue was larger after an informative than after an uninformative cue (Fig. 5). No significant effect of the cue or the music was found in the last time-window (from 2000 to 2500 ms post-cue). In the absence of distracting sound, the PDR to the cue is larger after an informative cue than after an uninformative cue from 500 to 2000 ms, suggesting increase in phasic arousal with enhanced voluntary attention.

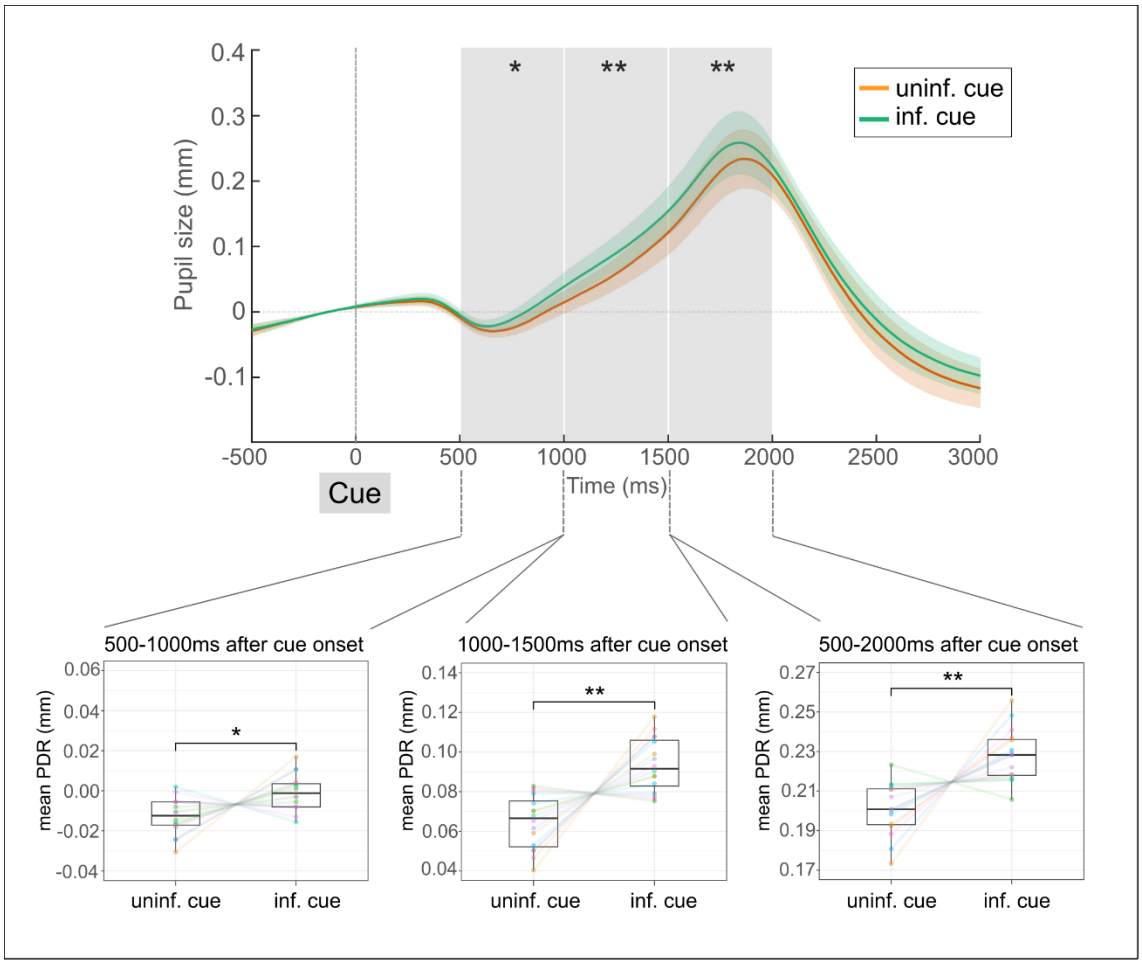

Fig 5. Pupil Dilation Response (PDR) to the presentation of the cue, as a function of cue type. Grey areas correspond to time-windows where the cue effect (informative vs. uninformative) was significant. Shadowed areas represent standard errors of the mean. Mean values for the significant time-windows are presented below the PDR time-course. Boxplots were plotted using normalized individual data (= raw data - subject mean + group mean). Within each boxplot, the horizontal line represents the group median, the lower and upper hinges correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (IQR = inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge. Superimposed to each boxplot, the colored dots represent individual means. * $p < 0.05$, ** $p < 0.01$. inf: informative, uninf: uninformative.

3.3.2. PDR to the distractor

To test the impact of music on phasic arousal to salient task-irrelevant stimuli, mean amplitudes of the PDR to the distractor were analyzed in two time-windows following a distractor: 300 – 800 ms and 950 – 1450 ms for early dis, and 300 – 800 ms and 1100 – 1800 ms for late dis.

During the first time-window, a significant main effect of the DisPosition2 (χ^2 (1) = 5.66, $p < 0.05$) and DisContent (χ^2 (1) = 5.35, $p < 0.05$) was found on PDR amplitude. Specifically, the PDR was larger after an early (mean = 0.092 mm, SEM = 0.005 mm) than after a late distractor (mean = 0.068 mm, SEM = 0.005 mm), and larger after a high arousing (mean = 0.09 mm, SEM = 0.004 mm) than after a low arousing distractor (mean = 0.07 mm, SEM = 0.004 mm). Additionally, a significant interaction between cue and music (χ^2 (1) = 3.93, $p < 0.05$) was observed (Fig. 6). Post-hoc HSD analysis revealed that the PDR amplitude was larger for informative compared to uninformative trials, in the high arousing music condition only ($p = 0.08$).

During the second time-window a significant main effect of the DisPosition2 (χ^2 (1) = 14.67, $p < 0.001$) and DisContent (χ^2 (1) = 6.12, $p < 0.05$) was found on PDR amplitude. Specifically, PDR was larger after an early (mean = 0.194 mm, SEM = 0.005 mm) than after a late distractor (mean = 0.163 mm, SEM = 0.005 mm), and larger after a high arousing (mean = 0.204 mm, SEM = 0.006 mm) than after a low arousing distractor (mean = 0.158 mm, SEM = 0.006 mm).

As already shown[30], the PDR is larger after a high arousing rather than a low arousing distractor. Importantly, the cue effect on the PDR to the distractor is larger under higher tonic arousal during the early time-window.

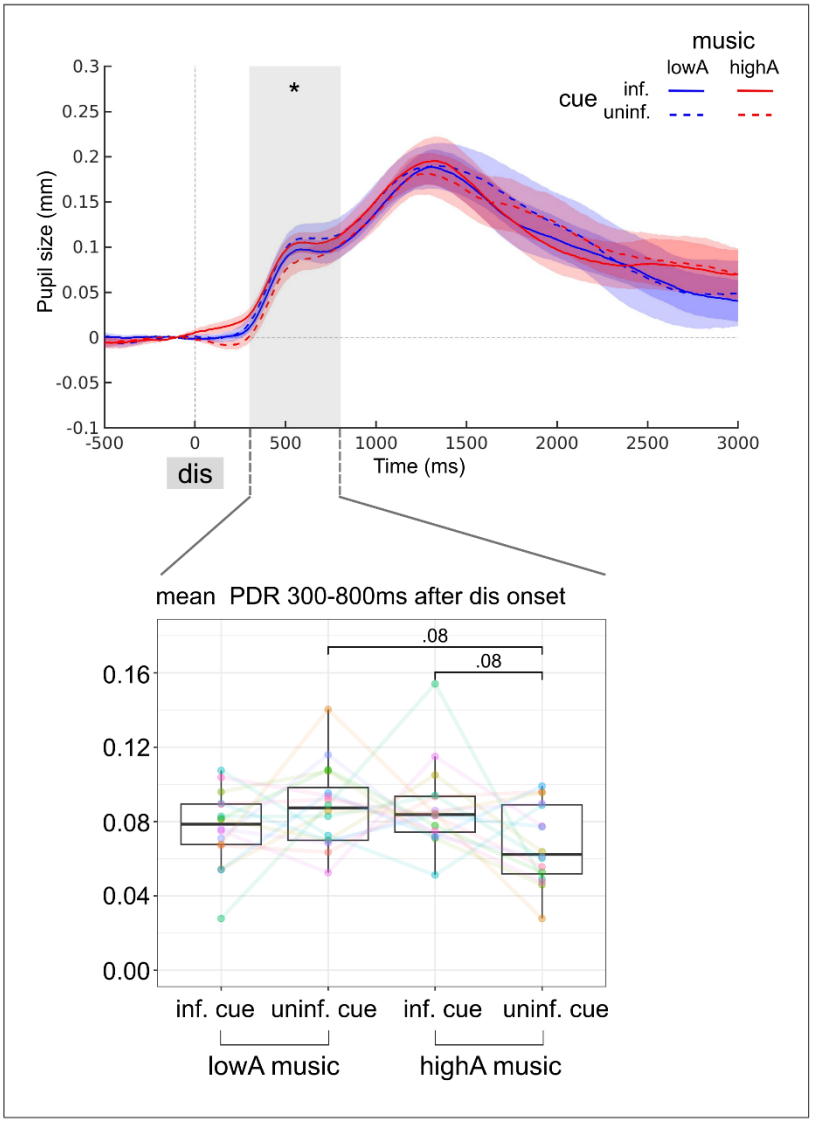

Fig 6. Pupil Dilation Response (PDR) to the presentation of the distractor sound, as a function of music and cue type. The grey area corresponds to the time window where the Music x Cue interaction was significant. Shadowed areas represent standard errors of the mean. Mean values for the significant time-windows are presented below the PDR time-course. Boxplots were plotted using normalized individual data (= raw data - subject mean + group mean). Within each boxplot, the horizontal line represents the group median, the lower and upper hinges correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (IQR = inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge. Superimposed to each boxplot, the colored dots represent individual means. * $p < 0.05$. lowA: low arousing, higA: high arousing, inf: informative, uninf: uninformative.

3.3.3. ERP to the cue

To investigate the impact of music on brain correlates of voluntary attention orienting, we analyzed the Contingent Negative Variation (CNV) elicited between the cue and the target, in trials with no distracting sounds. Cue-related potential were analyzed on **Cz** in two time-windows following a cue: 650 – 850 ms and 850 – 1050 ms. In both time-windows, no significant effect was found. The CNV is not significantly impacted by the type of music.

3.3.4. ERP to the distractor

To investigate the impact of music on distractor processing, we investigated the successive ERPs to the distracting sounds: a fronto-central N1 (~110ms), a central P2 (~180ms), a fronto-central early-P3 (~240ms) and a late-P3 with frontal and parietal components (~340 ms).

The mean amplitude of N1 response to distractors was analyzed on **FCz** in the **90-130ms** time-window following a distractor. A significant effect of the DisPosition2 (χ^2 (1) = 7.74, $p < 0.01$) showed a larger N1 amplitude in response to an early (mean = -5.380 µV, SEM = 0.160 µV) compared to a late distractor (mean = -4.506 µV, SEM = 0.160 µV). A significant DisContent x DisPosition2 x Music interaction (χ^2 (1) = 6.51, $p < 0.05$) was also found. However, post-hoc analysis revealed no significant difference regarding this interaction ($p > 0.1$).

For the P2 response, the distractor-related potential was analyzed on **Cz** in the **160-200ms** time-window following a distractor. A significant effect of DisPosition2 (χ^2 (1) = 8.50, $p < 0.01$) showed larger P2 amplitude in response to a late distractor (mean = 3.737 µV, SEM = 0.178 µV) compared to an early one (mean = 2.775 µV, SEM = 0.178 µV). Finally, a significant effect of the DisContent (χ^2 (1) = 14.19, $p < 0.001$) indicated a larger P2 amplitude in response to a highly arousing distractor (mean = 3.763 µV, SEM = 0.236 µV) than to a low arousing one (mean = 2.748 µV, SEM = 0.236 µV).

The mean amplitude of early-P3 response to distractors was analyzed on **FCz** in the **220-260ms** time-window following a distractor. The early-P3 amplitude was found modulated by DisPosition2 (χ^2 (1) = 5.04, $p < 0.05$) and DisContent (χ^2 (1) = 26.87, $p < 0.001$). Early-P3 amplitude was larger in response to a late distractor (mean = 5.867 µV, SEM = 0.214 µV) compared to an early one (mean = 4.775 µV, SEM = 0.214 µV), and larger in response to a highly arousing distractor (mean = 6.434 µV, SEM = 0.178 µV) compared to a low arousing distractor (mean = 4.207 µV, SEM = 0.178 µV). Finally, a significant DisContent x DisPosition2 x Music interaction (χ^2 (1) = 5.47, $p < 0.05$) was found. Early-P3 amplitude was larger in response to a highly arousing distractor compared to a low arousing distractor in all conditions, except in

trials with an early distractor after listening to high arousing music (Fig. 7). The impact of the distractor arousing content on the early-P3 to distractor is modulated by tonic arousal.

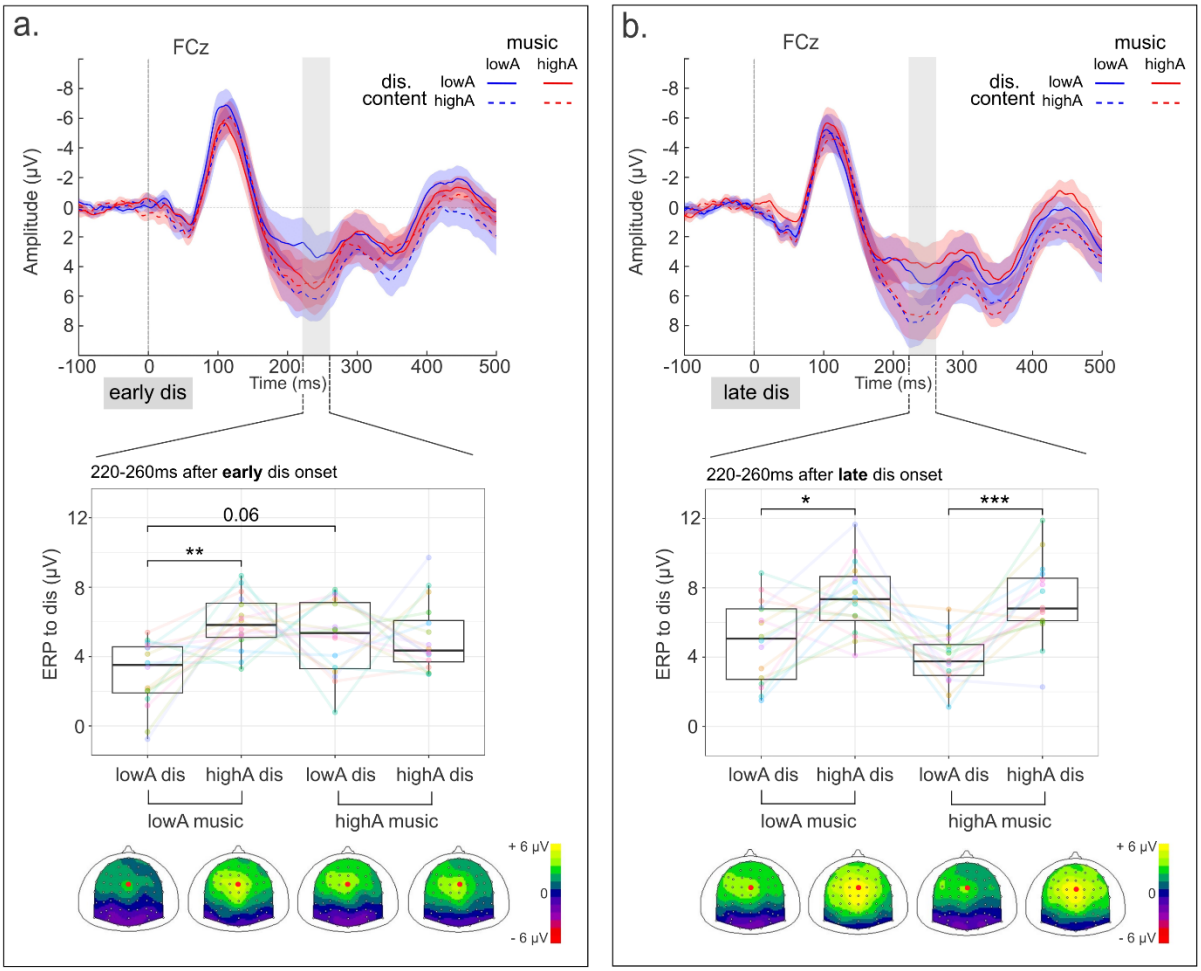

Fig 7. Event-Related Potentials (ERPs) on FCz electrode to the presentation of early (panel **a.**) or late distractor (panel **b.**) as a function of music type (lowA music vs. highA music) and distractor content (lowA dis vs. highA dis). Grey areas correspond to the time-windows where the Music x Distractor position x Distractor content was significant. Mean values for the significant time-windows are presented below the ERP time-course, with corresponding scalp topographies (top view). On the scalp topographies, the red dot indicates the FCz electrode. Shadowed areas represent standard errors of the mean. Mean values for the significant time-windows are presented below the ERP time-course. Boxplots were plotted using normalized individual data (= raw data - subject mean + group mean). Within each boxplot, the horizontal line represents the group median, the lower and upper hinges correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (IQR = inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge. Superimposed to each boxplot, the colored dots represent individual means. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. lowA: low arousing, higA: high arousing, inf: informative, uninf: uninformative, dis: distractor.

For the late-P3 response, the distractor-related potential was analyzed on **Fz** & **CPz** in the 320-360ms time-window following a distractor. On **Fz**, a significant effect of DisPosition2 (χ^2 (1) = 20.07, $p < 0.001$) showed larger late-P3 amplitude after a late distractor (mean = 5.945 µV, SEM = 0.227 µV) than after an early distractor (mean = 3.237 µV, SEM = 0.227 µV). On **CPz**, a significant effect of DisPosition2 (χ^2 (1) = 15.76, $p < 0.001$) showed larger late-P3 amplitude after a late distractor (mean = 4.952 µV, SEM = 0.239 µV) than after an early distractor (mean = 3.181 µV, SEM = 0.239 µV). A significant effect of DisContent (χ^2 (1) = 16.01, $p < 0.001$) showed larger late-P3 amplitude after a high arousing distractor (mean = 5.301 µV, SEM = 0.303 µV) than after a low arousing distractor (mean = 2.832 µV, SEM = 0.303 µV).

3.3.5. ERP to the target

To test the impact of music and cue types on target processing, we analyzed the successive fronto-central N1 (~130ms) and parietal target-P2/P3 (200-500ms), in trials with no distracting sound.

The mean amplitude of N1 response to targets was analyzed on **FCz** in the **110-150ms** time-window following a target. No significant effect was found.

The mean amplitude of target-P2/P3 response was analyzed on **CPz** in the **190-230ms** time-window following a target. A significant effect of music (χ^2 (1) = 9.17, $p < 0.01$) indicated a larger target-P2/P3 amplitude in the high arousing music condition compared to the low arousing music condition (Fig. 8).

The mean amplitude of the target-P3 response was also analyzed on **CPz** in the **250-500ms** time-window following a target. A significant effect of cue (χ^2 (1) = 6.83, $p < 0.01$) indicated a larger target-P3 amplitude after an uninformative compared to an informative cue.

The processing of task-relevant stimuli around 200 ms is enhanced under higher arousal.

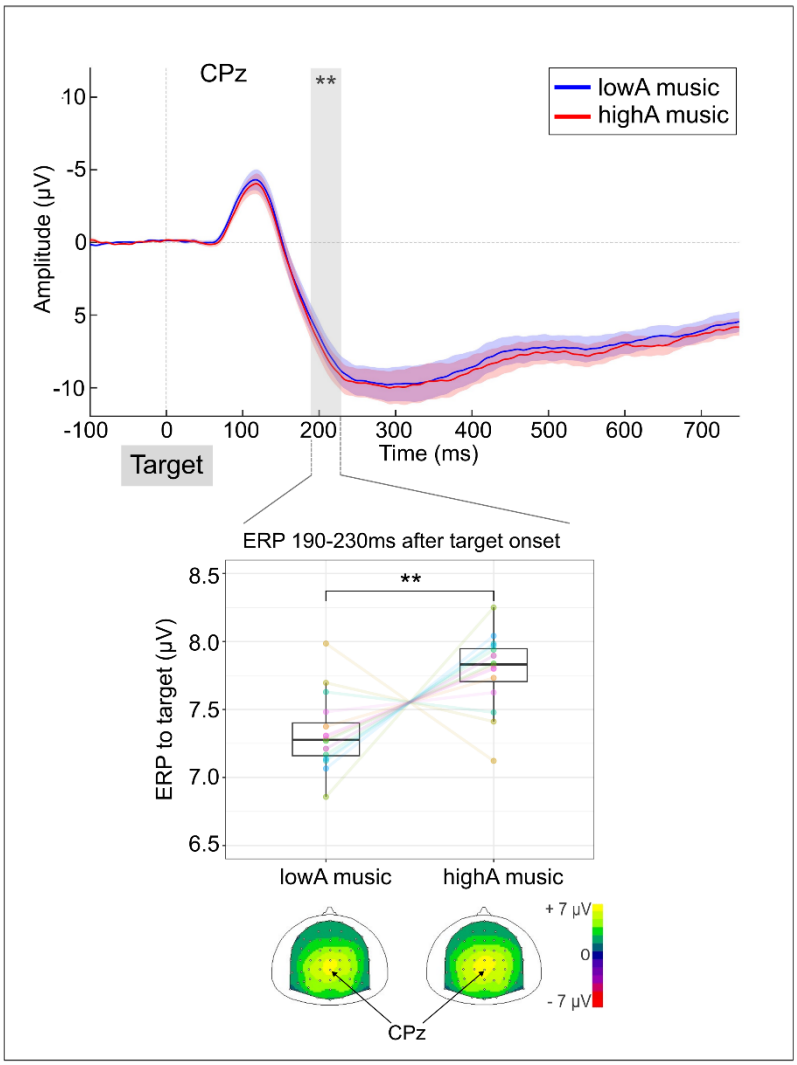

Fig 8. Event-Related Potentials (ERP) on CPz electrode to the presentation of the target sound, as a function of music type. Grey areas correspond to the time-window where the main effect of music was significant. Mean values for the significant time-window are presented below the ERP time-course, with corresponding scalp topographies (top view). On the scalp topographies, the red dot indicates the CPz electrode. Shadowed areas represent standard errors of the mean. Mean values for the significant time-windows are presented below the ERP time-course. Boxplots were plotted using normalized individual data (= raw data - subject mean + group mean). Within each boxplot, the horizontal line represents the group median, the lower and upper hinges correspond to the first and third quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (IQR = inter-quartile range, or distance between the first and third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge. Superimposed to each boxplot, the colored dots represent individual means. ** $p < 0.01$. lowA: low arousing, highA: high arousing.

4. DISCUSSION

The aim of the present study was to investigate the interplay between arousal and the voluntary/involuntary attentional balance in the auditory modality.

Indirect physiological measures of the LC-NE system (skin conductance, pupil dilation, heart rate) were enhanced during high compared to low arousing music excerpts, indicating a successful modulation of tonic arousal from a low to a moderate level. Voluntary attention was measured by comparing informative and uninformative cues (i.e. cue effect), as done in previous studies using the same paradigm[21,49]. Consistently, the cue effect was observed at the behavioral level, with shorter reaction times in the informative cue condition[21,50,51], and at the physiological level, with a smaller P3 in response to target in the informative cue condition[21,52]. Reaction time variability, calculated using the standard deviation of mean reaction times, was used as an index of sustained voluntary attention[53–55]. Involuntary attention was triggered by unexpected task-irrelevant salient sounds (distractors). At the behavioral level, involuntary orienting of attention can be indirectly assessed by comparing reaction times in the presence of early and late distractors[22]. At the physiological level, event-related potentials (ERPs) to distractors are used as a marker of involuntary attention[57]. However, these ERPs encompass both an involuntary orienting of attention and a phasic arousal responses triggered by distractors[22]. Phasic arousal can be assessed at the behavioral level by measuring the impact of early distractors on reaction times, and at the cerebral level by analyzing the effect of the arousing content on the early-P3 to distractor[22]. Moreover, the pupil dilation response (PDR) to sounds, can be used as an indirect measure of phasic arousal[11,46,47]. In the present study, an enhanced PDR in response to distractors with high compared to low arousing content showed an effective modulation of phasic arousal[30,58].

Firstly, we demonstrate a relationship between voluntary attention and phasic arousal: increasing voluntary attention can lead to heightened phasic arousal responses, as evidenced by the larger pupil dilation response (PDR) observed during target expectation after informative cues (i.e., under enhanced voluntary attention). This result is consistent with the few studies investigating the effect of voluntary attention on PDR using visual cueing paradigms[59,60] (but see [61]). In agreement with Kahneman theoretical model of attention (see Fig. 9), the present result experimentally confirms that voluntary attention can modulate the arousal level in a transient manner in order to adapt to task demand[4]. This anticipatory increase in arousal, indirectly reflecting an increase in LC-NE activity, might participate in facilitating the processing of the expected target sound, as observed with a smaller P3 to the target and shorter reaction times after informative cues.

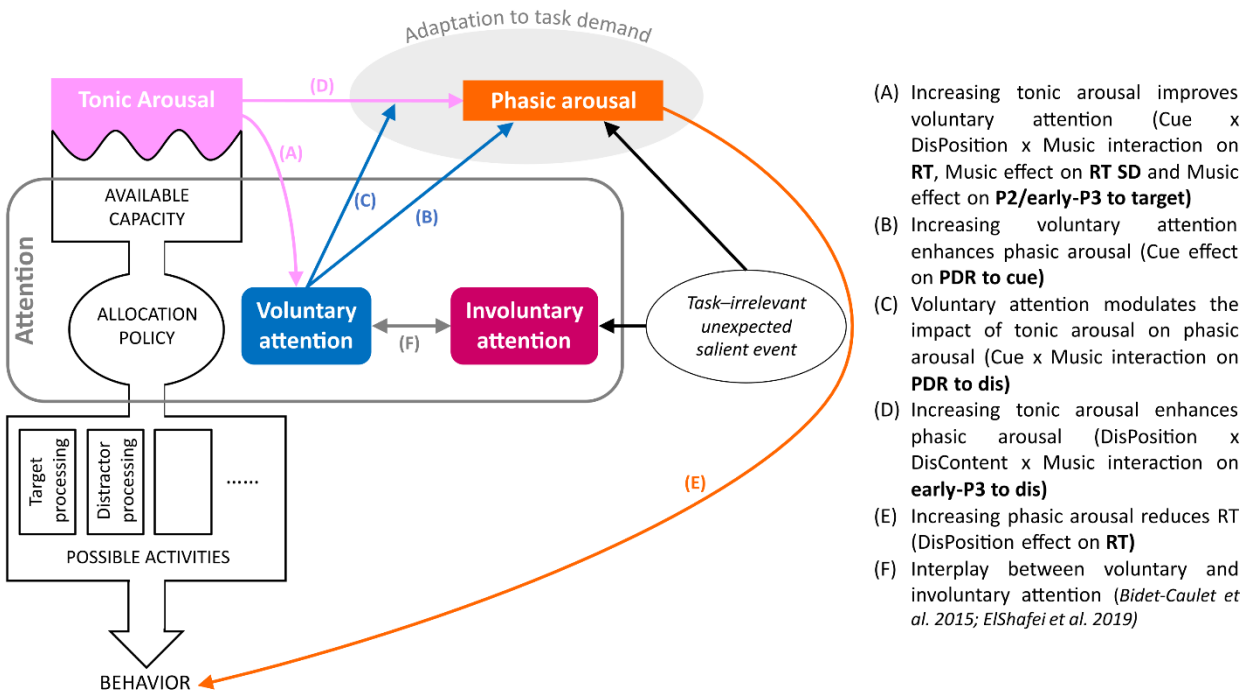

Fig 9. Schematic illustration of arousal and attention interactions (after Kahneman, 1973). This model reflects a general framework for viewing how arousal and attention processes could interact. Overall, the level of tonic arousal determines the amount of available capacity for attention, and attention allocates these resources to different possible activities resulting in an adapted behavior. More specifically, in this study, we find that tonic arousal influences voluntary attention (A), which can in turn impact phasic arousal responses directly (B) or in interaction with tonic arousal (C). The level of tonic arousal also influences the amplitude of phasic arousal responses (D), which can have an impact on behavior (E). The interaction between tonic arousal, phasic arousal and voluntary attention allows the overall arousal level to be adapted to the task demands. Task-irrelevant unexpected salient events (distractors) trigger both an increase in phasic arousal and involuntary attention processes. A balance between involuntary and voluntary attention (F) controls the allocation of available capacity to the possible brain activities, task-relevant ones (e.g. target processing) or task-irrelevant ones (e.g. distractor processing).

Moreover, our study delves into the intricate relationship between tonic and phasic arousal. In human, very few studies have investigated this relationship and found a reduced phasic arousal response under higher tonic arousal[35,37,46,63]. In the present study, tonic arousal interacts in a more complex way with the PDR and ERP indirect measures of phasic arousal.

First, tonic arousal modulated the effect of the cue on the PDR to the distractor. The PDR elicited by a distractor was reduced after an uninformative cue in the high tonic arousal condition. This suggests a larger effect of voluntary attention (cue effect) on phasic arousal (indexed by the PDR amplitude) when the level of tonic arousal increases.

Second, tonic arousal interacted with the impact of the arousing content of distracting sounds at the cerebral level, more precisely at the latency of the early-P3 to distractor sounds (220-260 post sound onset) and at fronto-central sites. In most conditions, the early-P3 was larger in response to high compared to low arousing distractors, in agreement with previous results[22,30]. However, in response to early distractors under high tonic arousal, no difference in amplitude of the early-P3 in response to low and high arousing distractors was found, with both amplitudes reaching a maximum (ceiling effect). The early-P3 has been associated for several years to the involuntary orienting of attention towards unexpected salient stimuli (for review see [64]). Recently, its amplitude at fronto-central sites has been found strongly impacted by the arousal content of distracting sounds, with increasing amplitude as arousal content increases[22,30]. This arousal content effect on the early-P3 amplitude is quite likely to be associated to the arousal component of the orienting response[4,8]. Thus, in specific cases (here early presentation of the distractor), enhancing tonic arousal could increase phasic arousal in response to task-irrelevant sounds up to a maximum, resulting in a modulation of the cortical activity around 250 ms. This result of a small phasic response under low tonic arousal and a larger phasic response under moderate tonic arousal is consistent with in the inverse-U relationship between tonic and phasic arousal observed in animals[23,32,33,62]. These PDR and ERP results highlight a complex interaction between tonic and phasic arousal that could be mediated or modulated by voluntary attention and follow the inverse-U relationship in specific cases (Fig. 9 B & C).

Notably, our findings reveal how tonic arousal, modulated through music listening, impacts voluntary and involuntary attention. To begin with, increasing tonic arousal was found to improve voluntary attention at the behavioral and cerebral levels (Fig. 9A). At the behavioral level, a greater cue effect was found under high tonic arousal in trials with an early distractor. This result suggests that the effect of voluntary orienting of auditory attention is increased under high tonic arousal, with more anticipation enabling a faster response to the target sound. In the visual modality, inconsistent results were found with improved[41] or unchanged[43,65] attention orienting with increasing tonic arousal. Moreover, the decreased variability in reaction times under high arousing music conditions suggests improved sustained attention capacities. To our knowledge, very few studies (and only in the visual modality) investigated the effect of tonic arousal on RT variability, with inconsistent results: reduced variability[66] or no effect[39]. These discrepancies might be due to the diverse methods employed to modulate tonic arousal, including factors such as the type of music or the duration of induction. Additionally, at the brain level, the target-P2/P3 response at parietal electrodes was enhanced in the high arousing music condition. The P3b or P300 response elicited by task-relevant stimuli is commonly observed on parietal electrodes[67], and its amplitude is deemed to reflect the

relevance of the stimuli and the attentional resources allocated to the task[68,69]. Therefore, an enhanced target-P2/P3 associated with an increase of tonic arousal indicates an increased processing of the target sound and suggests allocation of more attentional resources to this task-relevant stimulus. Those consistent results at the behavioral and cerebral levels show that several processes of voluntary attention, namely attention orientation and sustained attention, are improved with increasing moderate levels of tonic arousal. This aligns with previous theoretical[4,10,12,13] and experimental works in animals[23,33,62] (for review see [15]), highlighting that an intermediate level of tonic arousal is associated with optimal performances.

Finally, the level of tonic arousal does not seem to have any direct or indirect impact on involuntary attention as assessed with Bayesian statistics. To our knowledge, this is the first study to investigate the effect of tonic arousal on involuntary attention in Human. Measuring the impact of unexpected task-irrelevant sounds is not straightforward. Here, we capitalize on a series of 10 experiments showing the suitability of the CAT paradigm to explore auditory involuntary attention at the behavioral and cerebral levels[21,22,48,49,70–75]. Neither the behavioral distraction effect (cost in reaction times observed after late distractors compared with early ones), nor the ERP to the distractors were modulated by the level of tonic arousal, independently of the arousing content. This novel result shows that the moderate levels of tonic arousal used in the present study do not influence involuntary orienting of attention.

While the interplay between attention and arousal has been at the heart of physiological works in animals and theoretical reflections, there are very few experimental results in humans. Here, we provide experimental evidence in humans that (1) heightened voluntary attention is linked to intensified phasic arousal responses, (2) tonic arousal influences phasic arousal in interaction with voluntary attention and following the inverse-U relationship in specific cases, (3) increasing tonic arousal is associated with improved voluntary attention, (4) moderate levels of tonic arousal do not influence involuntary orienting of attention. These experimental results align nicely with Kahneman's theoretical model, affirming “a mutual relation between attention and arousal” with voluntary attention being optimized at intermediate levels of tonic arousal and transiently influencing arousal levels to adapt to task demands[4]. Crucially, these results demonstrate at behavioral and brain levels that the tonic arousal level can influence the attentional balance by improving voluntary attention in a transient and sustained manner, rather than by impacting involuntary attention. Music-induced increase in tonic arousal could be a promising clinical perspective for improving voluntary attention deficit disorders.

2. MATERIAL AND METHODS

2.1. Participants

Twenty-two participants took part in the experiment. The data of six participants had to be discarded because of technical issues or difficulties performing the task. Then, the analysis was performed on data from sixteen participants (8 female, 4 left-handed and 12 right-handed, mean age ± standard deviation (SD): 24.3 ± 2.5 years). All participants were free from neurological or psychiatric disorder, did not take any medication that could impact brain functioning, and had normal hearing and normal or corrected-to-normal vision. The study was approved by a national ethical committee, and subjects gave written informed consent, according to the Declaration of Helsinki, and they were paid for their participation. To limit inter-individual differences in tonic arousal, only participants drinking no more than two cups of coffee (or equivalent in caffeinated drinks) and one glass of alcohol per day were included.

2.2. Stimuli and Task

The CAT consisted of twenty-four blocks (separated in two sessions of twelve blocks each, one to six days apart, at the same period of the day). Each block started with a music track played during 30s followed by a 2s pause and then by 48 trials of the CAT (**Fig. 1**). To modulate tonic arousal, a low or high arousing musical excerpt (*lowA* or *highA* music) was presented during 30s at the beginning of each block (see Supplementary Materials for more details).

Each trial of the CAT started with a visual cue (200ms duration), corresponding to a green arrow pointing either to the left, to the right or to both sides, centrally presented on a grey background screen. The visual cue was followed by a delay randomly chosen between 900 and 1100ms. Then, a target sound (50ms duration), corresponding to a monaural harmonic sound (fundamental frequency: 200 Hz, 5 harmonics; 5ms rise-time, 5ms fall-time) was presented in earphones. In 25% of the CAT trials, a binaural distracting sound (300ms duration) was played during the delay between the cue and the target. In the absence of the visual cue, a blue fixation cross was presented at the center of the screen.

The cue and target categories were manipulated in the same proportion for trials with and without distracting sound. In 66.6% of trials, the cue was informative (pointing left and the target sound was played in the left ear (33.3%), or pointing right and the target sound was played in the right ear (33.3%)). In the last 33.3 % of the trials, the cue was uninformative, pointing in both directions while the target sound was played in the left (16.7 %) or right (16.7 %) ear.

A total of 24 different sounds were used as distracting sounds (clock-alarm, doorbell, laughter, gunfire, etc.). They were chosen from a pool of sounds from Bidet-Caulet et al. (2015)[21] and Widmann et al. (2018)[30] in order to modulate phasic arousal: 12 low arousing and emotionally positive distracting sounds (*lowA dis*) and 12 high arousing and emotionally negative distracting sounds (*highA dis*) (see Supplementary Materials for more details). These distracting sounds were binaurally presented, their onset equiprobably fell in two different time periods before the target onset: in the 600-800ms range (*early dis*), or in the 350-550ms range (*late dis*). To compare behavioral and physiological responses to acoustically matched sounds, the same distracting sounds were played for each combination of music type (lowA and highA music), cue category (informative left, informative right and uninformative) and distractor condition (early or late dis). Each distracting sound was played 12 times during the whole experiment, but no more than twice during each single block to limit habituation.

Participants were instructed to perform a detection task by pressing a mouse button as fast as possible when they heard the target sound, and to ignore the distracting sounds. They were asked to focus their attention to the cued side in the case of informative cues. Moreover, they were informed that informative cues were 100 % predictive and that a distracting sound could be sometimes played. Participants were also instructed to blink naturally during the task but to keep their eyes fixated on the cross and to minimize eye movements while performing the task. Participants had between 3000 and 3200ms to answer after target sounds, each CAT trial lasted therefore around 4350ms on average, leading to block duration of ≈4 min and EEG session of ≈1h (breaks included).

Participants performed a total of twenty-four blocks (48 trials each). Within blocks, CAT trials were randomized differently for each subject to limit sequence effects. In each subject, across both sessions, the same twelve blocks were presented in the low and in the high arousing music conditions. In each session, high and low arousal blocks were alternating, with one session starting with a low arousing music and the other one with a high arousing music. The order of sessions was balanced across participants.

During the whole experiment, for each music condition, 144 no-dis, 12 early lowA dis, 12 early highA dis, 12 late lowA dis, 12 late highA dis were presented in each cue category (informative left, informative right & uninformative).

2.3. Procedure

For each session, participants were seated in a comfortable armchair at a 1 m distance from the screen. All stimuli were delivered using Presentation® software (Version 23.0, Neurobehavioral Systems, Inc.,

Berkeley, CA, www.neurobs.com). Sounds were delivered through earphones (Etymotic, ER4SR). First, the auditory threshold was determined for the target sound, in each ear, for each participant using the Bekesy tracking method[76]. The target sounds were monaurally presented at 10 dB SL (between 23 and 31 dBA across subjects), while distracting sounds were binaurally played at 40 dB SL (between 69 and 77 dBA across subjects), low arousing music at 40 dB SL (between 54 and 62 dBA across subjects), high arousing music at 45 dB SL (between 69 and 77 dBA across subjects) above the target sound thresholds. Second, participants were trained with a short sequence of the task. Finally, for each session, the electroencephalogram (EEG), the electrooculogram (EOG), the electrocardiogram (ECG), the skin conductance and the pupil size were recorded while participants performed 12 blocks. Each block started with a five-point eye-tracker calibration and validation procedure.

2.4. Data recordings

2.4.1. Physiological recordings

EEG data were recorded from 64 electrodes using the ActiveTwo system (BioSemi, the Netherlands). Eight additional Flat-Type active electrodes were used for vertical EOG (left supraorbital and infraorbital ridge locations), ECG (left shoulder and stomach), and skin conductance (medial phalanx of index and middle fingers of the non-dominant hand) recordings, and two electrodes were placed on earlobes for offline referencing. Data were amplified (−3 dB at ∼204 Hz low-pass, DC coupled), digitized (1024 Hz), and stored for offline analysis. EEG data were re-referenced offline to the average potential of the two earlobe electrodes.

2.4.2. Pupillometry

The pupil size of the left eye was recorded during each entire block and monitored using an eye-tracking system (Eyelink Portable Duo, SR Research Ltd., Mississauga, Ontario, Canada). Participants were placed at a 45 cm distance from the eye-tracker (remote mode, sampling rate of 500 Hz). At the beginning of each block, a five-point calibration followed by a five-point validation procedure was performed.

2.5. Data analysis

Data analysis was performed using the software package for electrophysiological analysis (ELAN Pack[77]) developed at the Lyon Neuroscience Research Center (elan.lyon@inserm.fr) and custom MATLAB programs.

2.5.1. Behavior

For each participant, the shortest reaction time (RT) limit for correct response (RT lower limit) was set to 150 ms after the target, from previous studies using a similar paradigm in adults[48,49,70]. The longest RT for a correct response (RT upper limit) was calculated from all RT > 0 ms using the John Tukey's method of leveraging the Interquartile Range (3rd Quartile + 1.5 x Interquartile Range). A button press before the RT lower limit was considered as a false alarm (FA). A trial with a button press after the RT upper limit was considered as a late response. A trial with no button press before the next cue onset was considered as a missed trial. A trial with no FA and with a button press between the RT lower limit and upper limit was counted as a correct trial. For statistical analysis of reaction times (RTs), all RT > 0 ms were considered. Additionally, the standard deviation of positive reaction time (SD RT) in trials with no distractor was calculated for each subject.

2.5.2. Skin conductance

Skin conductance data were low-pass filtered offline at 10Hz. For each block, a Z-score normalization to the baseline (period of 2500ms before the onset of the music) was applied.

2.5.3. Heart rate

ECG data were filtered offline with a 1 Hz high-pass filter and a 40 Hz low-pass filter. During the 30s music presentation, R-peaks were automatically detected, interbeat interval (IBI) were derived from the signal, and artifacts and ectopic beats were hand corrected using the QRSTool software[78]. Then, the heart rate was extracted using the CMetX software[78] (www.psychofizz.org). Analyses were conducted on mean heart rate values.

2.5.4. Pupillometry

The eye tracker pupil size digital counts were converted to millimeters using an artificial eye. Full blinks were detected and marked by the eye tracker. Partial blinks were detected from the smoothed velocity time as pupil size changes exceeding a velocity adjusted for each subject around 15 mm/s. Data were removed in a 50ms pre-blink and a 100ms post-blink interval and a 50ms pre-partial blink and 150ms post-partial blink interval. Subsequently, segments with blinks or missing data were interpolated with linear interpolation. Trials with more than 50% of interpolated data were removed from further analysis.

Three distinct analyses were conducted with pupillometry data: the pupil size changes during the music presentation, the pupil dilation responses (PDR) to the cue, and the PDR to the distractor.

To investigate the effect of music presentation, a baseline subtraction was applied (period of 455ms before the onset of the music) and the mean amplitude was computed during the 30s music presentation.

To analyze the PDR to the cue, a baseline subtraction was applied using the 250ms period before stimulus onset and the mean PDR was computed in 500ms time-windows from 500 to 2500ms after cue onset. To analyze the PDR to the distractor, for each distractor onset time-range, surrogate dis-locked PDR were created in the no-dis trials and subtracted from the actual dis-locked responses. The obtained PDR was thus clear of cue-related activity. Additionally, a baseline subtraction was applied using the 250ms period before stimulus onset and the mean PDR was computed in time-windows after distractor onset: 300 – 800 ms and 950 – 1450 ms post early dis, and 300 – 800 ms and 1100 – 1800 ms post late dis.

2.5.4. EEG

EEG data were filtered offline with a 0.1 Hz high-pass filter and a 40 Hz lowpass filter. Eye-related activities were detected using independent component analysis (ICA) and were selectively removed via the inverse ICA transformation[79] (EEGLAB toolbox). Only 1 or 2 ICs were removed in each participant. In 10 participants, the flat or excessively noisy signals at 1 to 6 electrodes were replaced by their values interpolated from the remaining electrodes using spherical spline interpolation[80]. Trials including false alarms or undetected target, and trials contaminated with excessive muscular activity were excluded from further analysis.

ERPs were averaged for each stimulus event: cue-related potentials (cueRPs) were averaged locked to cue onset, target-related potentials (targetRPs) were averaged locked to target onset, and distractor-related potentials (disRPs) were averaged locked to distractor onset. Different baseline corrections were applied according to the investigated processes. ERP scalp topographies were computed using spherical spline interpolation[80].

To investigate the deployment of top-down attention mechanisms in the absence of distracting sound (no-dis), cueRPs were baseline corrected to the mean amplitude of the -100 to 0ms period before cue onset, and targetRPs were corrected to the mean amplitude of the -100 to 0ms period before target onset. To analyze ERPs to distracting sound, for each distractor onset time-range, surrogate disRPs were created in the no-dis trials and subtracted from the actual disRPs. Then, disRPs were baseline corrected to the mean amplitude of the -100 to 0mms period before distractor onset. The obtained disRPs were thus clear of cue-related activity.

For the analysis of EEG and pupil data, we considered only trials that were not rejected in both EEG and pupil data analysis. On average across participants and music conditions, the number of considered trials for analysis was 205 no-dis (SEM = 2.5), 17 early lowA dis (SEM = 0.4), 16 early highA dis (SEM = 0.4), 17 late lowA dis (SEM = 0.3), and 17 late highA dis (SEM = 0.4).

2.6. Statistical analysis

When only two paired conditions were compared, Bayesian non-parametric Wilcoxon signed-rank test were conducted using JASP software (JASP—A Fresh Way to Do Statistics, 2021; Version 0.14.1). In contrast to Frequentist statistics, Bayesian analyses allow one to assess the credibility of both the alternative and null hypotheses. We reported Bayes factor (BF_{10}) as a measure of evidence in favor of the null hypothesis (BF_{10} 0.33–1, 0.1–0.33, 0.01–0.1, and lower than 0.01: weak, moderate, strong, and decisive evidence, respectively) and in favor of the alternative hypothesis (value of 1–3, 3–10, 10–100, and more than 100: weak, moderate, strong, and decisive evidence, respectively[81]).

When the analysis included more than one factor, we used linear mixed models (LMM) or generalized linear mixed models (GLM) using R (R Core Team, 2021, version 4.1.2). We estimated the effect of the manipulated task parameters (Music, Cue, Distractor Position, Distractor Content) on measured variables (reactions times, pupil dilation responses, event-related potentials; see Table 1 for a list of the model characteristics and Supplementary Methods for more details on model selection and analysis). We considered the results of all analyses significant at $p < 0.05$.

2.6.1. Peripheral indices of arousal

Non-parametric statistics were used to test the effect of music on skin conductance, pupil size and heart rate. A paired Wilcoxon signed-rank test was used on the skin conductance mean amplitude, mean pupil size and mean heart rate during the presentation of the music (0-30 sec) to test the tonic arousal modulation.

2.6.2. Behavior

Response Type Rates

The effect of music on the rate of Hits, Miss, False Alarms, and Late Responses was investigated using Bayesian non-parametric paired Wilcoxon signed-rank tests.

Reaction Times

Raw RT were log-transformed at the single-trial scale to be better fitted to a linear model with Gaussian family distribution. In a first model, logRT were fitted to a model with the following three within-subject fixed-factors: Music (2 levels: highA-music, lowA-music), Cue (2 levels: informative, uninformative) and Dis Position3 (3 levels: no-dis, early, late). In a second model, to analyse the effect of the distractor content, logRT in trials with a distractor were fitted to a model with the following four within-subject fixed-factors: Music (2 levels: highA-music, lowA-music), Cue (2 levels: informative, uninformative), Dis Position (2 levels: early, late), Dis Content (2 levels: highA-dis, lowA-dis).

Moreover, we performed a Bayesian repeated measures analysis of variance (rmANOVA), with two within-subject factors: Music (2 levels: highA-music, lowA-music), Cue (2 levels: informative, uninformative), on two specific measures of the distractor effect: the distractor occurrence (median RT in no-dis minus median RT in early dis) and the distractor position (median RT in late dis minus median RT in early dis). Based on previous results[21,22], these differences can be respectively considered as good approximations of the facilitation phasic arousal effect and the detrimental involuntary attention effect triggered by distracting sounds. We report the (BF_{10} for each model, as well as the $BF_{inclusion}$ for each factor and interaction, which compares models that contain the effect to equivalent models stripped of the effect and was considered as a relative measure of evidence supporting the inclusion of a factor.

Additionally, we investigated the effect of music type on the **variability of reaction times** using a Bayesian non-parametric signed-rank paired Wilxocon test.

2.6.3. Pupil dilation response (PDR)

For statistical analysis of the PDR to the cue (no-dis trials), the mean pupil size per subject in the 500-1000ms, 1000-1500ms, 1500-2000ms and 2000-2500ms time windows following a cue was fitted to a linear model with Gaussian family distribution model with two within-subject fixed-factors: Music (2 levels: highA-music, lowA-music) and Cue (2 levels: informative, uninformative).

For statistical analysis of the PDR to the distractor, mean pupil size was calculated separately for early and late distractors: in the 300-800ms and 950-1450ms time-windows following an early distractor, and in the 300-800ms and 1100-1600ms time-windows following a late distractor. Then, mean amplitude for the first and second time-windows were fitted to a linear model with Gaussian family distribution model. with four within-subject fixed-factors: Music (2 levels: highA-music, lowA-music), Cue (2 levels: informative, uninformative), Dis Position2 (2 levels: early, late), Dis Content (2 levels: highA-dis, lowA-dis).

2.6.4. Event-related potential (ERP)

Statistical analyses were performed on specific electrode and time-windows for each main ERP component (N1, P2, early-P3 and late-P3). The electrodes and time-windows were chosen based on previous studies[21,22].

For statistical analysis of cue-related potentials (no-dis trials), the mean amplitude per subject of the Contingent Negative Variation (CNV) on Cz electrode in the 650-850 ms and 850-1050 ms time-windows following a cue was fitted to a linear model with Gaussian family distribution model with two within-subject fixed-factors: Music (2 levels: highA-music, lowA-music) and Cue (2 levels: informative, uninformative).

For statistical analysis of distractor-related potentials, the mean amplitude per subject of the N1 (90-130ms at FCz), of the P2 (160-200ms at Cz), of the early-P3 (220-260ms at FCz) and of the late-P3 (320-360ms at Fz and CPz) following a distractor was fitted to a linear model with Gaussian family distribution model with three within-subject fixed-factors: Music (2 levels: highA-music, lowA-music), Cue (2 levels: informative, uninformative), Dis Position2 (2 levels: early, late) and Dis Content (2 levels: highA-dis, lowA-dis).

For statistical analysis of target-related potentials (no-dis trials), the mean amplitude per subject of the N1 (110-150 ms at FCz), of the target-P2/P3 (190-230 ms at CPz) and of the target-P3 (250-500 ms at CPz) following a target was fitted to a linear model with Gaussian family distribution model with two within-subject fixed-factors: Music (2 levels: highA-music, lowA-music) and Cue (2 levels: informative, uninformative).

References

1. Broadbent, D. E. *Decision and Stress*. xiv, 522 (Academic Press, Oxford, England, 1971).
2. Corbetta, M., Patel, G. & Shulman, G. L. The Reorienting System of the Human Brain: From Environment to Theory of Mind. *Neuron* **58**, 306–324 (2008).
3. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).
4. Kahneman, D. *Attention and Effort*. (Prentice-Hall, Englewood Cliffs, N.J, 1973).
5. Mesulam, M. M. A cortical network for directed attention and unilateral neglect. *Ann. Neurol.* **10**, 309–325 (1981).
6. Posner, M. I. & Petersen, S. E. The attention system of the human brain. *Annu. Rev. Neurosci.* **13**, 25–42 (1990).
7. Coull, J. T. Neural correlates of attention and arousal: insights from electrophysiology, functional neuroimaging and psychopharmacology. *Prog. Neurobiol.* **55**, 343–361 (1998).
8. Näätänen, R. *Attention and Brain Function*. xx, 494 (Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, US, 1992).
9. Sturm, W. & Willmes, K. On the functional neuroanatomy of intrinsic and phasic alertness. *NeuroImage* **14**, S76-84 (2001).
10. Aston-Jones, G. & Cohen, J. D. AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal Performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
11. Nieuwenhuis, S., de Geus, E. J. & Aston-Jones, G. The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology* **48**, 162–175 (2011).
12. Yerkes, R. M. & Dodson, J. D. The Relation of Strength of Stimulus to Rapidity of Habit Formation. *J. Comp. Neurol. Psychol.* **18**, 459–482 (1908).

13. Easterbrook, J. A. The effect of emotion on cue utilization and the organization of behavior. *Psychol. Rev.* **66**, 183–201 (1959).

14. Mather, M., Clewett, D., Sakaki, M. & Harley, C. W. Norepinephrine ignites local hot spots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behav. Brain Sci.* **39**, e200 (2016).

15. Berridge, C. W. & Waterhouse, B. D. The locus coeruleus–noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain Res. Rev.* **42**, 33–84 (2003).

16. Vazey, E. M., Moorman, D. E. & Aston-Jones, G. Phasic locus coeruleus activity regulates cortical encoding of salience information. *Proc. Natl. Acad. Sci.* **115**, E9439–E9448 (2018).

17. Wetzel, N. & Schröger, E. On the development of auditory distraction: A review. *PsyCh J.* **3**, 72–91 (2014).

18. SanMiguel, I., Linden, D. & Escera, C. Attention capture by novel sounds: Distraction versus facilitation. *Eur. J. Cogn. Psychol.* **22**, 481–515 (2010).

19. Wetzel, N., Widmann, A. & Schröger, E. Distraction and facilitation—two faces of the same coin? *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 664–674 (2012).

20. Max, C., Widmann, A., Kotz, S. A., Schröger, E. & Wetzel, N. Distraction by emotional sounds: Disentangling arousal benefits and orienting costs. *Emotion* **15**, 428–437 (2015).

21. Bidet-Caulet, A., Bottemanne, L., Fonteneau, C., Giard, M.-H. & Bertrand, O. Brain Dynamics of Distractibility: Interaction Between Top-Down and Bottom-Up Mechanisms of Auditory Attention. *Brain Topogr.* **28**, 423–436 (2015).

22. Masson, R. & Bidet-Caulet, A. Fronto-central P3a to distracting sounds: An index of their arousing properties. *NeuroImage* **185**, 164–180 (2019).

23. Aston-Jones, G., Chiang, C. & Alexinsky, T. Chapter 35 - Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance. in *Progress in Brain Research* (eds. Barnes, C. D. & Pompeiano, O.) vol. 88 501–520 (Elsevier, 1991).
24. Devilbiss, D. M. Consequences of tuning network function by tonic and phasic locus coeruleus output and stress: Regulating detection and discrimination of peripheral stimuli. *Brain Res.* **1709**, 16–27 (2019).
25. Grant, S. J., Aston-Jones, G. & Redmond, D. E. Responses of primate locus coeruleus neurons to simple and complex sensory stimuli. *Brain Res. Bull.* **21**, 401–410 (1988).
26. Rasmussen, K., Morilak, D. A. & Jacobs, B. L. Single unit activity of locus coeruleus neurons in the freely moving cat: I. During naturalistic behaviors and in response to simple and complex stimuli. *Brain Res.* **371**, 324–334 (1986).
27. Escera, C., Alho, K., Winkler, I. & Näätänen, R. Neural Mechanisms of Involuntary Attention to Acoustic Novelty and Change. *J. Cogn. Neurosci.* **10**, 590–604 (1998).
28. Polich, J. & Criado, J. R. Neuropsychology and neuropharmacology of P3a and P3b. *Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol.* **60**, 172–185 (2006).
29. Yago, E., Escera, C., Alho, K., Giard, M.-H. & Serra-Grabulosa, J. M. Spatiotemporal dynamics of the auditory novelty-P3 event-related brain potential. *Cogn. Brain Res.* **16**, 383–390 (2003).
30. Widmann, A., Schröger, E. & Wetzel, N. Emotion lies in the eye of the listener: Emotional arousal to novel sounds is reflected in the sympathetic contribution to the pupil dilation response and the P3. *Biol. Psychol.* **133**, 10–17 (2018).
31. Aston-Jones, G., Chiang, C. & Alexinsky, T. Discharge of noradrenergic locus coeruleus neurons in behaving rats and monkeys suggests a role in vigilance. *Prog. Brain Res.* **88**, 501–520 (1991).
32. Aston-Jones, G., Rajkowski, J. & Cohen, J. Role of locus coeruleus in attention and behavioral flexibility. *Biol. Psychiatry* **46**, 1309–1320 (1999).

33. Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J. & Aston-Jones, G. The role of locus coeruleus in the regulation of cognitive performance. *Science* **283**, 549–554 (1999).

34. Coull, J. T., Jones, M. E. P., Egan, T. D., Frith, C. D. & Maze, M. Attentional effects of noradrenaline vary with arousal level: selective activation of thalamic pulvinar in humans. *NeuroImage* **22**, 315–322 (2004).

35. Rojas-Thomas, F. *et al.* Impact of acute psychosocial stress on attentional control in humans. A study of evoked potentials and pupillary response. *Neurobiol. Stress* **25**, 100551 (2023).

36. Kozhevnikov, M., Irene Strasser, A. V., McDougal, E., Dhond, R. & Samuel, G. Beyond mindfulness: Arousal-driven modulation of attentional control during arousal-based practices. *Curr. Res. Neurobiol.* **3**, 100053 (2022).

37. Mather, M. *et al.* Isometric exercise facilitates attention to salient events in women via the noradrenergic system. *NeuroImage* **210**, 116560 (2020).

38. Cloutier, A., Fernandez, N. B., Houde-Archambault, C. & Gosselin, N. Effect of Background Music on Attentional Control in Older and Young Adults. *Front. Psychol.* **11**, 557225 (2020).

39. Kiss, L. & Linnell, K. J. The effect of preferred background music on task-focus in sustained attention. *Psychol. Res.* **85**, 2313–2325 (2021).

40. Nadon, É., Tillmann, B., Saj, A. & Gosselin, N. The Emotional Effect of Background Music on Selective Attention of Adults. *Front. Psychol.* **12**, 729037 (2021).

41. Dovorany, N., Brannick, S., Johnson, N., Ratiu, I. & LaCroix, A. N. Happy and sad music acutely modulate different types of attention in older adults. *Front. Psychol.* **14**, 1029773 (2023).

42. Quan, Y., Ma, W., Li, H. & Thompson, W. F. Slow tempo music preserves attentional efficiency in young children. *Atten. Percept. Psychophys.* **85**, 978–984 (2023).

43. McConnell, M. M. & Shore, D. I. Upbeat and happy: arousal as an important factor in studying attention. *Cogn. Emot.* **25**, 1184–1195 (2011).

44. Bellato, A., Arora, I., Kochhar, P., Hollis, C. & Groom, M. J. Indices of Heart Rate Variability and Performance During a Response-Conflict Task Are Differently Associated With ADHD and Autism. *J. Atten. Disord.* **26**, 434–446 (2022).

45. Wass, S. V., de Barbaro, K. & Clackson, K. Tonic and phasic co-variation of peripheral arousal indices in infants. *Biol. Psychol.* **111**, 26–39 (2015).

46. Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M. & Cohen, J. D. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cogn. Affect. Behav. Neurosci.* **10**, 252–269 (2010).

47. Joshi, S., Li, Y., Kalwani, R. & Gold, J. I. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron* **89**, 221–234 (2016).

48. Hoyer, R. S., Elshafei, H., Hemmerlin, J., Bouet, R. & Bidet-Caulet, A. Why Are Children So Distractible? Development of Attention and Motor Control From Childhood to Adulthood. *Child Dev.* **92**, e716–e737 (2021).

49. Hoyer, R. S. *et al.* When do we become more distractible? Progressive evolution of different components of distractibility from early to late adulthood. *J. Exp. Psychol. Gen.* **152**, 3403–3417 (2023).

50. Posner, M. I. Orienting of attention. *Q. J. Exp. Psychol.* **32**, 3–25 (1980).

51. Posner, M. Attentional Networks and Consciousness. *Front. Psychol.* **3**, (2012).

52. Abramov, D. M. *et al.* Visuospatial information processing load and the ratio between parietal cue and target P3 amplitudes in the Attentional Network Test. *Neurosci. Lett.* **647**, 91–96 (2017).

53. Manly, T., Davison, B., Heutink, J., Galloway, M. & Robertson, I. H. Not enough time or not enough attention? Speed, error and self-maintained control in the Sustained Attention to Response Test (SART). *Clin. Neuropsychol. Assess. Int. J. Res. Clin. Pract.* **3**, 167–177 (2000).

54. Antonini, T. N., Narad, M. E., Langberg, J. M. & Epstein, J. N. Behavioral correlates of reaction time variability in children with and without ADHD. *Neuropsychology* **27**, 201–209 (2013).

55. Esterman, M., Noonan, S. K., Rosenberg, M. & Degutis, J. In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cereb. Cortex N. Y. N 1991* **23**, 2712–2723 (2013).

56. Flehmig, H. C., Steinborn, M., Langner, R., Scholz, A. & Westhoff, K. Assessing intraindividual variability in sustained attention: Reliability, relation to speed and accuracy, and practice effects. *Psychol. Sci.* **49**, 132–149 (2007).

57. Escera, C., Alho, K., Schröger, E. & Winkler, I. W. Involuntary Attention and Distractibility as Evaluated with Event-Related Brain Potentials. *Audiol. Neurotol.* **5**, 151–166 (2000).

58. Bonmassar, C., Widmann, A. & Wetzel, N. The impact of novelty and emotion on attention-related neuronal and pupil responses in children. *Dev. Cogn. Neurosci.* **42**, 100766 (2020).

59. Geva, R., Zivan, M., Warsha, A. & Olchik, D. Alerting, orienting or executive attention networks: differential patters of pupil dilations. *Front. Behav. Neurosci.* **7**, (2013).

60. Dragone, A. *et al.* Expectancy modulates pupil size during endogenous orienting of spatial attention. *Cortex* **102**, 57–66 (2018).

61. Aminihajibashi, S., Hagen, T., Laeng, B. & Espeseth, T. Pupillary and behavioral markers of alerting and orienting: An individual difference approach. *Brain Cogn.* **143**, 105597 (2020).

62. Rajkowski, J., Kubiak, P. & Aston-Jones, G. Locus coeruleus activity in monkey: phasic and tonic changes are associated with altered vigilance. *Brain Res. Bull.* **35**, 607–616 (1994).

63. Murphy, P. R., Robertson, I. H., Balsters, J. H. & O'connell, R. G. Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. *Psychophysiology* **48**, 1532–1543 (2011).

64. Escera, C. & Corral, M. J. Role of Mismatch Negativity and Novelty-P3 in Involuntary Auditory Attention. *J. Psychophysiol.* **21**, 251–264 (2007).

65. Rowe, G., Hirsh, J. B. & Anderson, A. K. Positive affect increases the breadth of attentional selection. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 383–388 (2007).

66. O'Connell, R. G. *et al.* Self-Alert Training: volitional modulation of autonomic arousal improves sustained attention. *Neuropsychologia* **46**, 1379–1390 (2008).

67. Rugg, M. D. & Coles, M. G. H. *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*. xii, 220 (Oxford University Press, New York, NY, US, 1995).

68. Herrmann, C. S. & Knight, R. T. Mechanisms of human attention: event-related potentials and oscillations. *Neurosci. Biobehav. Rev.* **25**, 465–476 (2001).

69. Kok, A. On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* **38**, 557–577 (2001).

70. Hoyer, R. S., Pakulak, E., Bidet-Caulet, A. & Karns, C. M. Relationships among age, socioeconomic status, and distractibility in preschoolers as assessed by the Competitive Attention Test. *J. Exp. Child Psychol.* **227**, 105584 (2023).

71. ElShafei, H. A., Fornoni, L., Masson, R., Bertrand, O. & Bidet-Caulet, A. Age-related modulations of alpha and gamma brain activities underlying anticipation and distraction. *PLOS ONE* **15**, e0229334 (2020).

72. ElShafei, H. A., Fornoni, L., Masson, R., Bertrand, O. & Bidet-Caulet, A. What's in Your Gamma? Activation of the Ventral Fronto-Parietal Attentional Network in Response to Distracting Sounds. *Cereb. Cortex* **30**, 696–707 (2020).

73. ElShafei, H. A., Bouet, R., Bertrand, O. & Bidet-Caulet, A. Two Sides of the Same Coin: Distinct Sub-Bands in the α Rhythm Reflect Facilitation and Suppression Mechanisms during Auditory Anticipatory Attention. *eNeuro* **5**, (2018).

74. ElShafei, H. A. *et al.* Age-related differences in bottom-up and top-down attention: Insights from EEG and MEG. *Eur. J. Neurosci.* **55**, 1215–1231 (2022).

75. Ruby, P. *et al.* High dream recall frequency is associated with an increase of both bottom-up and top-down attentional processes. *Cereb. Cortex N. Y. N 1991* **32**, 3752–3762 (2022).

76. Haupt, G. T. Audiometry and Sound Calibration Utilities. (2003).

77. Aguera, P.-E., Jerbi, K., Caclin, A. & Bertrand, O. ELAN: a software package for analysis and visualization of MEG, EEG, and LFP signals. *Comput. Intell. Neurosci.* **2011**, 5:1-5:11 (2011).

78. Allen, J. J. B., Chambers, A. S. & Towers, D. N. The many metrics of cardiac chronotropy: a pragmatic primer and a brief comparison of metrics. *Biol. Psychol.* **74**, 243–262 (2007).

79. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).

80. Perrin, F., Pernier, J., Bertrand, O. & Echallier, J. F. Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* **72**, 184–187 (1989).

81. Lee, M. D. & Wagenmakers, E.-J. *Bayesian Cognitive Modeling: A Practical Course*. (Cambridge University Press, Cambridge, 2014). doi:10.1017/CBO9781139087759.