



Article

# Cognitive Load Changes during Music Listening and its Implication in Earcon Design in Public Environments: An fNIRS Study

Eunju Jeong <sup>1,2,\*</sup>, Hokyung Ryu <sup>1,3</sup>, Geonsang Jo <sup>1</sup> and Jaehyeok Kim <sup>4</sup>

<sup>1</sup> Department of Arts and Technology, Hanyang University, Seoul 04763, Korea; hryu@hanyang.ac.kr (H.R.); azureluna@hanyang.ac.kr (G.J.)

<sup>2</sup> Division of Industrial Information Studies, Hanyang University, Seoul 04763, Korea

<sup>3</sup> Graduate School of Technology and Innovation Management, Hanyang University, Seoul 04763, Korea

<sup>4</sup> Department of Industrial Engineering, Hanyang University, Seoul 04763, Korea; jhyeok.kim@letinar.com

\* Correspondence: ejeong@hanyang.ac.kr; Tel.: +82-2-2220-2260

Received: 21 August 2018; Accepted: 17 September 2018; Published: 21 September 2018



**Abstract:** A key for earcon design in public environments is to incorporate an individual's perceived level of cognitive load for better communication. This study aimed to examine the cognitive load changes required to perform a melodic contour identification task (CIT). While healthy college students ( $N = 16$ ) were presented with five CITs, behavioral (reaction time and accuracy) and cerebral hemodynamic responses were measured using functional near-infrared spectroscopy. Our behavioral findings showed a gradual increase in cognitive load from CIT1 to CIT3 followed by an abrupt increase between CIT4 (i.e., listening to two concurrent melodic contours in an alternating manner and identifying the direction of the target contour,  $p < 0.001$ ) and CIT5 (i.e., listening to two concurrent melodic contours in a divided manner and identifying the directions of both contours,  $p < 0.001$ ). Cerebral hemodynamic responses showed a congruent trend with behavioral findings. Specific to the frontopolar area (Brodmann's area 10), oxygenated hemoglobin increased significantly between CIT4 and CIT5 ( $p < 0.05$ ) while the level of deoxygenated hemoglobin decreased. Altogether, the findings indicate that the cognitive threshold for young adults (CIT5) and appropriate tuning of the relationship between timbre and pitch contour can lower the perceived cognitive load and, thus, can be an effective design strategy for earcon in a public environment.

**Keywords:** melody perception; contour identification; auditory attention; cognitive load; near-infrared spectroscopy; earcon design

## 1. Introduction

Earcon has been defined as “abstract, synthetic tones that can be used in structured combinations to create auditory messages” [1,2] and can represent almost any type of event or interaction in public environments [3]. Earcon takes advantage of the auditory modality, which is omnidirectionally perceivable and reaches across distances, thus making it preferable to other sensory modalities, for example, in situations where visual information is unavailable or not fully available [4]. When individuals are faced with a high visual cognitive load or the inaccessibility of visual information, auditory earcons can deliver information [5–7], thus facilitating almost immediate decision-making and action [8]. An assumed disadvantage of non-verbal and non-iconic auditory earcon is that their meaning might be vague and require more time to learn [9].

In a real-world auditory context, multi-layered sound streams are concomitant (e.g., voices, natural sounds, and electronic noises). Especially in public environments, where high levels of background speech or noises are present, communication challenges are common [10–12]. Earcon can

be the best option for successful communication and decision-making [13,14], and be preferable even over speech-based auditory display [15,16], which requires a longer duration to distinguish the source from environmental sounds or to completely understand the sentence [17,18].

It is a natural phenomenon that we attend to certain features of sounds and segregate meaningful information from complex auditory scenes [19,20]. In this process, we are compelled to use a combined strategy of attention, which embraces a broad range of subtypes, such as stimulus orientation, vigilance, selective attention, and executive control [21–25]. As the complexity of auditory environments increases, individuals tend to utilize more attentional resources to effectively process task-relevant against task-irrelevant information (e.g., the cocktail party effect [26–28]), resulting in increased mental effort. In this sense, the cognitive load theory is applicable to auditory attention in public environments [19,20].

The load theory of attention [29–32] proposes that the level and type of information provided in a task can determine the degree of mental effort. More specifically, stimulus complexity (i.e., perceptual demand) and task difficulty (i.e., cognitive control load) are believed to generate and impose different amounts of cognitive load [30–33]. Perceptual demands are perceived cognitive loads caused by the stimulus complexity, which further emphasizes the “bottom-up” aspects of information processing, while cognitive control demands are associated with the “top-down” aspects of cognitive control, such as cognitive flexibility, working memory, and executive functioning [33–35].

Many studies have suggested that cognitive load is the defining factor for designing an efficient and effective earcon [36–38]. The existing approach to earcon design, specific to multiple and concurrent earcons in auditory display, focuses on the issues around masking and interfering, providing solutions that utilize spatially disparate sound locations [39–41]. It is the assumption that the allocation of information to the auditory field of view can increase the efficiency in attentional resource management and thus contribute to reducing cognitive loads [42–44]. Without spatial separation, the multiple sound sources will have a greater tendency to fuse together, making them more difficult to understand [45].

Various acoustic features are also of great importance in controlling cognitive loads [16,46]. The core features include amplitude (loudness), frequency (pitch), and waveform (timbre). Loudness is a dominant characteristic of auditory perception and provides a strong affordance to assign a considerable amount of attentional resource [47]. An auditory alarm system, for example, recommends a sound level of at least 75 dB and allows decibel increases in a very emergent situation. However, sounds that are more than 140 dB or in the case of a sudden increase of more than 30 dB can cause emotional displeasure and sudden fright [48,49]. Loudness control seems less effective and informative, so applying other perceptual features such as pitch and timbre have been recognized as potentially useful.

Using a combination of timbre and pitch is more promising in earcon design. Recently, researchers have reported the interaction effects between two perceptual features. In particular, pitch and timbre showed a symmetric relationship, so modulation in the spectral feature of timbres yields changes in pitch recognition [50]. Li et al. [51] reported enhancement in pitch identification when provided with timbre information. When more perceptually distinguishable timbres were used, there was a significant improvement in performance of the identification task [52]. Collectively, the findings suggested that auditory events and their constituent acoustic features (i.e., bottom-up process) can influence cognitive load changes by increasing the chance to establish prediction or anticipation [53]; thus, designing the features are of importance.

In fact, earcons are presented in a simple music-like pattern [3] and are experienced as melody. Given that concurrent earcon patterns are very similar to music where multi-layered streams (e.g., two melodies, one melody and accompaniment) convey aesthetic and meaningful information, we can integrate neurocognitive evidence indicating the attention system in the brain involving melody perception. Previous studies have revealed that melody perception leads to voluntary and involuntary activation of the brain [54–59]. Especially in multi-voice music-listening, in which diverse musical streams are presented concurrently, neural activation seems to be differentiated by the type of musical texture and the manner of listening (e.g., holistic, selective, or divided) [47–50]. A group of studies employing various types of musical tasks (e.g., error detection and target tone identification) have also

reported common but specific neural activation depending on the given stimulus and task [60–62]. Such findings indicated that listening to music in various contexts (e.g., polyphonic music) can utilize and activate multiple attention systems in the brain, which are associated with the level of cognitive load. Although the existing literature on music and neuroscience indicates that various features may provide a hint toward the reduction of cognitive load (i.e., automatic or involuntary processing), how much cognitive load the other features of sound, such as pitch contour and timbre or their combination, would impose is still open to question.

Several studies have employed earcons presented simultaneously and examined their effect on cognitive loads. Concurrent spatialized sounds, for example, were used in a mobile-based system in Nomadic Radio [63]. Gaver et al. [64,65] used concurrent presentation of auditory information to determine the status and monitor the vital processes of plants. More recently, McGookin et al. [2] emphasized the importance of stimulus and task characteristics in designing earcon. The authors examined the effect of the characteristics on task performance and found that the amount of auditory information presented within the same time window and the acoustic properties (e.g., timbre) can together influence overall task performance. Although the effectiveness of concurrent earcon has been evaluated by assessing behavioral performance, the physiological aspects of cognitive load and its relationship with behavioral aspects has rarely been examined so far.

In the present study, we focused on cognitive loads associated with the subtypes of attention during a melodic contour identification task (CIT). In our previous studies [66,67], we developed a method of music-based attention assessment to examine the attentional function of individuals with a moderate-to-severe level of traumatic brain injury (TBI). Findings from our first study indicated that melodic contour identification was feasible both for healthy adults and adults with TBI, and had a high reliability (split-half coefficient = 0.836, Cronbach's  $\alpha$  = 0.940) [67]. In the second study, Jeong [66] investigated construct validity using factor analysis, yielding four factors of attention. The overall findings imply that the melodic CIT can distinguish different types and levels of auditory attention and, thus, is a valid and reliable task.

Here, we measured hemodynamic changes in the frontopolar area (Brodmann's area 10 (BA10)) to examine cognitive loads associated with CITs. This area receives information from the primary auditory cortex and relays the same to other regions of the prefrontal cortex (PFC). Laguë-Beauvais et al. [68] reported that this area receives both verbal and non-verbal information from the superior temporal gyrus. This area generally modulates attention-related function and becomes more active as the task complexity increases. Plakke and Romanski [69] indicated that BA10 manages audition-related task performances that require heavy cognitive processing. Previous studies have not measured the cognitive load imposed during different types of auditory attention, which would be important in public earcon design. For this purpose, the present study used functional near-infrared spectroscopy (fNIRS), a recent advancement in brain imaging technology, to measure cognitive load changes in the prefrontal regions [68,70–72].

Designing a nonverbal auditory earcon using both pitch and timbre requires the understanding of how such musical components impose cognitive loads. The present study, thus, aimed to determine the cognitive loads required to perform CITs, using pitch contours in conjunction with timbre in a simulated real-world auditory environment. With a multi-faceted approach to cognitive load, we assessed behavioral data along with hemodynamic changes using fNIRS. The neurophysiological response complements the sensitivity of the behavioral data and is a good indicator for decision-making in relation to perceived sensory information [35,73].

## 2. Materials and Methods

### 2.1. Participants

Sixteen college student volunteers (10 men and 6 women), who were not majoring in music, were recruited from a university in Seoul, Republic of Korea. We intentionally controlled the level of the

participants' general education and musical experience. All volunteers were freshmen or sophomores, limiting the years of education to 13 or 14 years. None of the participants were professionally trained in music (less than 1 year of professional music training), nor did they have a neurological medical history or any sensory impairment. The mean age of the participants was 23.5 years (standard deviation = 1.7 years).

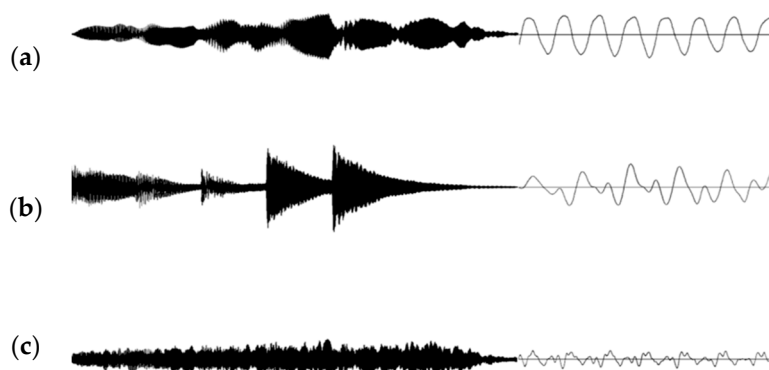
## 2.2. Music Stimuli

As shown in Figure 1, melodic contours are a series of tones moving in different directions (i.e., ascending, descending, and stationary). Two different types of contour were combined consecutively to yield six types of test items (i.e., ascending and descending, ascending and stationary, stationary and ascending, stationary and descending, descending and ascending, descending and stationary). The presentation time of each test item was 5250 ms, including two contours (2250 ms for each) and inter-contour interval (750 ms). Figure 1 shows examples of the pitch contours used in the study.

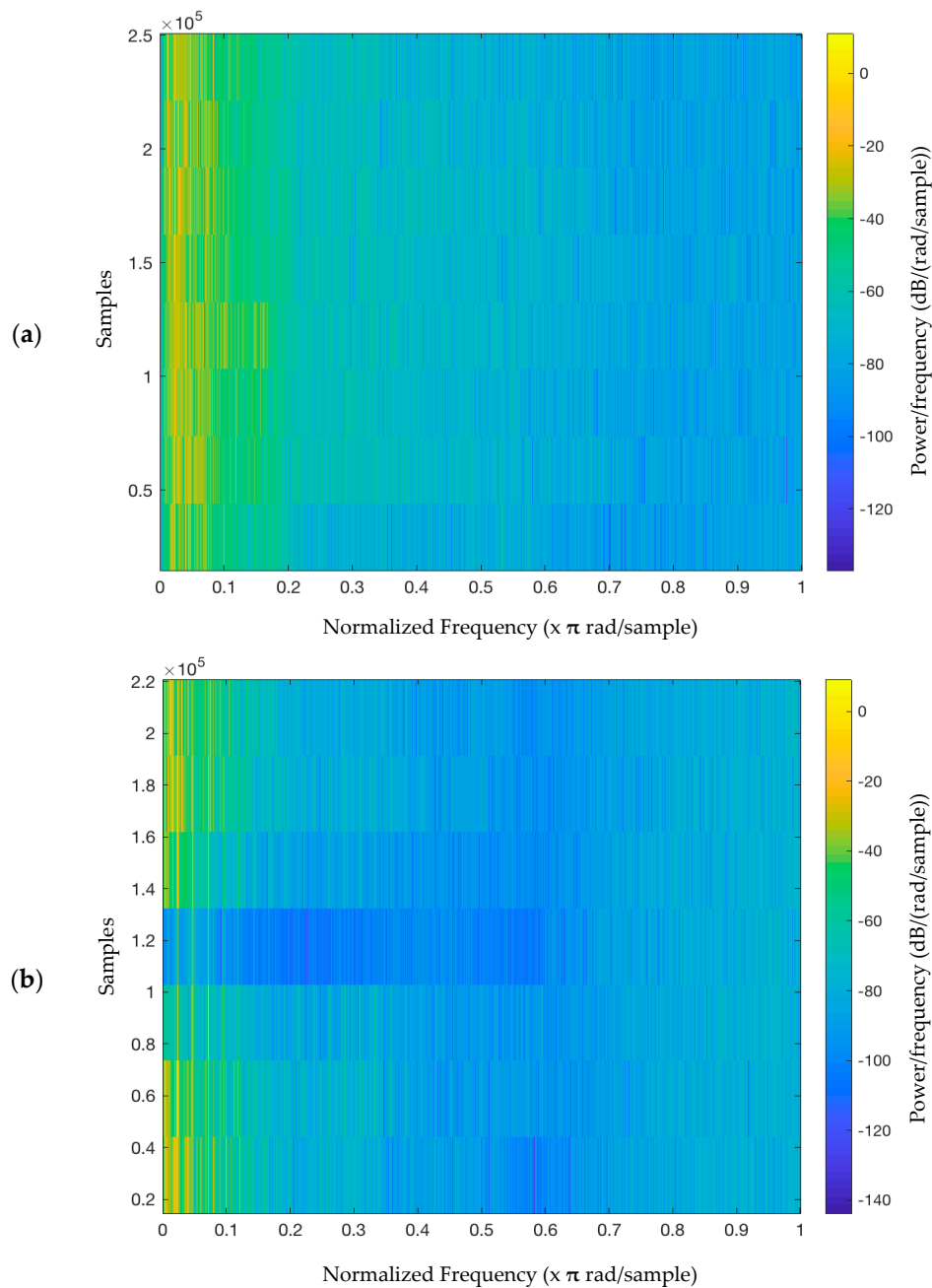


**Figure 1.** Samples of melodic contours (adopted from [66,67]). Item (a) includes ascending and stationary contours and Item (b) includes ascending and descending contours.

The timbres of the three instruments have different spectral and temporal complexities. The flute has a relatively simple spectrum with little attack; the piano has a more complex spectrum with a sharp attack; the strings have a complex spectrum with a soft attack (Figure 2). We selected the instruments following a previous study that classified various musical instruments based on the spectral features of timbre, such as the harmonic structure, inharmonicity, and harmonic energy skewness [74]. Figure 3 shows a spectrogram of target contours presented combined with environmental noise (Figure 3a), and with a target-like distractor (Figure 3b) using a short-time Fourier transform. Signal-to-noise ratio (SNR) was 7.2385 for the target contour presented by the flute against environmental noise and 8.0835 dB for the target contours with target-like contour, for example.



**Figure 2.** Waveforms of instrument timbres. The instruments are (a) flute, (b) piano, and (c) strings.



**Figure 3.** Spectrogram of target contours presented with (a) environmental noise and (b) a target-like distractor (b).

The six types of test items were modulated in five different keys (G# to C major) and presented with three instrument timbres, yielding a total of 90 test items. In each CIT, participants were randomly presented with 18 out of 90 items and asked to identify the directions of contours. The melodic contours were generated by a musical instrument digital interface (MIDI) synthesizer (YAMAHA DGX 230, Hamamatsu, Japan) with a digital audio workstation (Logic Pro X, Apple Inc., Cupertino, CA, USA). The experimental test was developed as a computerized version, using Visual Studio (Microsoft, Washington, DC, USA).

### 2.3. Contour Identification Task

The computerized version of the CIT was designed to measure different types of auditory attention and the associated cognitive load changes (Table 1). The task stimuli and structures were adopted

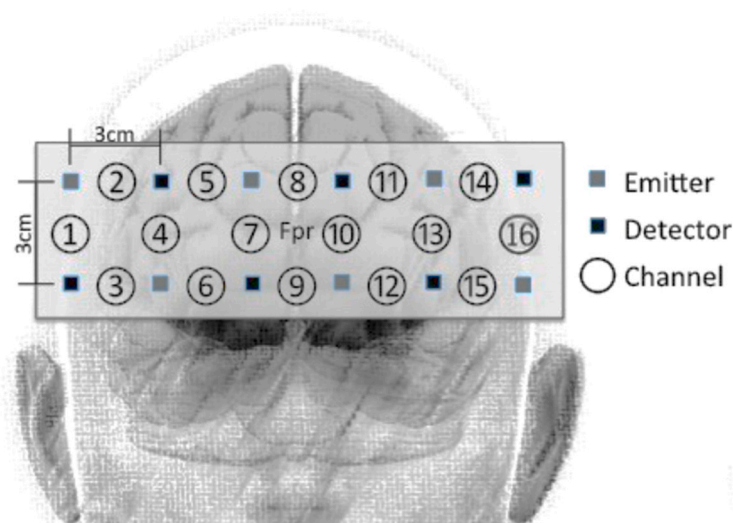
from previous studies [66,67,75] and modified for the current purpose. In CIT1, two consecutive contour directions were presented as a target contour without distraction (i.e., focused identification). Ten types of environmental sounds, including traffic, raining, twittering, ticktack, bustling, laughing, gabbling, applause, crying, and jeering sounds were randomly presented against target contours in the second task (i.e., CIT2, selective identification against environmental noise). Both CIT3 and CIT4 simultaneously presented two melodic contour streams. For CIT3, the participants were asked to attend to a target musical stimulus while ignoring the target-like distraction (i.e., selective identification against a more competing distractor). Target-like distractors had the same or different direction of contour and were played using different musical instruments. The CIT4 was more complex since the task involved an alternating focus between two melodic contours. Participants were asked to intentionally shift and re-focus their attention between the two auditory stimuli. Lastly, in CIT5, participants were asked to divide their attentional focus over both melodic contours and to completely identify all four contour directions.

**Table 1.** Structure of the contour identification tasks (CITs).

CIT	Target	Distractor	Given Task	Cognitive Load
1	Melodic contour	None	Focus	Low
2	Melodic contour	Environmental sounds	Focus	↓
3	Melodic contour	Target-like contours	Select	
4	Melodic contour	Target-like contours	Shift	
5	Melodic contour	Target-like contours	Divide	High

2.4. fNIRS Data Acquisition and Pre-Processing

In this study, we used fNIRS (16-channel Spectratech OEG-16, Yokohama, Japan) to measure hemodynamic changes, which is assumed to indicate the cognitive loads involved in each CIT [76–78]. The center of the measurement unit was placed between the frontopolar areas (Fp1 and Fp2), according to the international 10–20 system (Figure 4). The task-related hemodynamic changes were recorded through 16 channels, with a sampling rate of 0.65 s. The hemodynamic changes that were measured included oxygenated hemoglobin (HbO<sub>2</sub>), deoxygenated hemoglobin (HHb), and the sum of oxygenated and deoxygenated hemoglobin (HbT).



**Figure 4.** Placement of emitters, detectors, and channels for Spectratech OEG-16. Fpr refers to the frontopolar region.

The fNIRS data were collected and converted into concentration changes of hemoglobin using the modified Beer-Lambert law. In general, raw fNIRS data are affected by other physiological signals,

such as the heart rate, breathing, and eye-blinking; therefore, a zero-phase low- and high-pass filter with cut-off frequencies of 0.01–0.09 Hz was applied, using MATLAB (The Mathworks Korea LLC, Gangnam-gu, Korea), to pre-process the raw NIRS signals [79–81]. The values were standardized by subtracting the mean values of HbO<sub>2</sub>, HHb, and HbT obtained during the first 20 s baseline period, which is obtained prior to any type of stimulus presentation, from the means of each of the five CITs. This was performed since each individual had different baseline values of oxygen metabolism [81].

## 2.5. Procedure

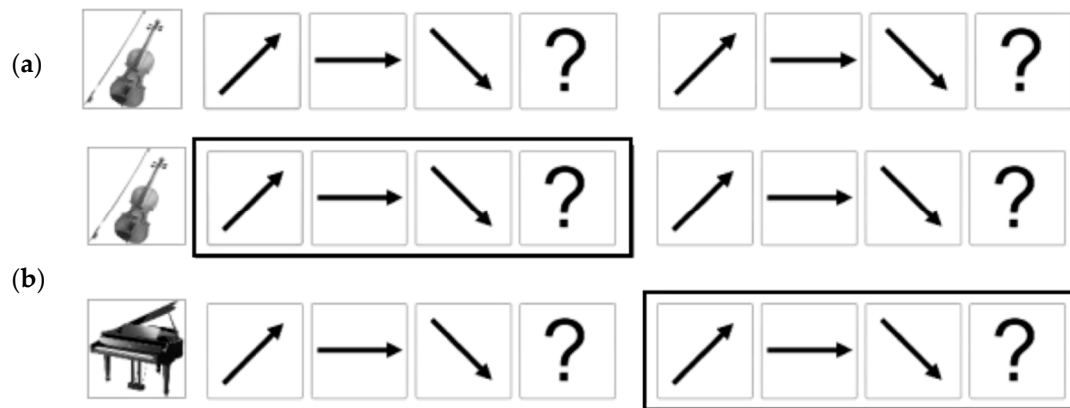
The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Institutional Review Board of Hanyang University (HYI-141273). The experiment was announced to college students registering for the course, “An Introduction to Cognitive Psychology”. All recruitment processes were performed electronically, and 16 out of 20 volunteers participated in this study. Three were excluded because of their professional experience in music training. One left-handed volunteer was also excluded. Informed consent was obtained from all individual participants included in the study. Once participants agreed to participate voluntarily and provided full permission for the publication, they filled out a nine-item demographic questionnaire (age, sex, academic major, previous musical experience, etc.).

A band-type NIRS containing an array of 12 probes (i.e., emitters and detectors) was attached to each participant’s forehead. The probes were connected to the main board of the NIRS, which was connected to a computer. Auditory stimuli were delivered diotically via headphones at a constant volume, and visual cues specifying the target musical stimulus were presented on a monitor. A 20 s baseline was recorded prior to stimulus presentation, before and after each of the five CITs, while the participants fixed their eyes on the center of the monitor.

The experiment started with an initial familiarization session. All participants took part in a brief stimulus-familiarization phase. In a practice session, participants were asked to identify the directions of the contours until they could correctly identify more than 80% of the directions. In a main session, each of the five CITs started with a brief instruction in terms of the task characteristics given in each CIT and how to respond to test items. Participants were also instructed to identify the directions of the target contours by clicking the arrow corresponding to the contour direction, as accurately and immediately as possible.

In terms of a visual cue, CIT1 and CIT2 had no visual cues; however, in CIT3, a picture of an instrument that plays target contours was presented prior to presenting the item to indicate which contour the participants selectively listened to. In CIT4, outlined boxes were additionally used to hint at which contours the participants selectively listened to and shifted from one to another instrument (see Figure 5). For example, the first outlined box appeared in the upper or lower line with the first set of contours, and the second box appeared with the second set of contours. In CIT5, the two boxes appeared in a random manner after the test item presentation, so the participants were asked to attend and hold information from four contour directions, but identified two of them (i.e., the contours of which the outlined boxes appeared).

The participants were also informed that their behavioral and hemodynamic responses were being recorded throughout the experiment. Reaction time was the sum of overall times between post-stimulus to the first arrow selection and between the first arrow selection and the second arrow selection. In each CIT, a total of 18 test items were presented (a blocked design) and the order of CIT was randomized across participants. When participants identified directions of all contours correctly, one point was assigned, so participants could obtain a maximum of 18 points for each CIT. The main experimental session required approximately 30 min to complete and was performed in a sound-proof room to control for other noises. The ambient light and temperature remained constant throughout the experimental sessions.



**Figure 5.** Examples of answer pages. (a) is given in CIT3 and (b) is given in CIT4. The boxes were presented prior to stimulus presentation (CIT4), while they appeared after the stimulus presentation (CIT5).

### 2.6. Statistical Analysis

We used a repeated measures design for the statistical analysis. The independent variables were the CITs (CIT1, CIT2, CIT3, CIT4, and CIT5) for the behavioral analysis and the sessions (pre-baseline, CIT1, CIT2, CIT3, CIT4, CIT5, and post-baseline) for the fNIRS analysis. The dependent variables were behavioral responses (performance accuracy and reaction time) and hemodynamic responses (HbO<sub>2</sub> and HHb). For behavioral analysis, the rate corrected score (RCS) was estimated using  $c/\Sigma RT$ , in which *c* refers to correct responses and *RT* refers to response time [82,83]. We selected the following NIRS channels representing the frontopolar area (BA10, BA11), including Channels 7, 8, 9, and 10. All statistical analyses were performed using the repeated measures analysis of variance (ANOVA) in SPSS version 20 (SPSS Inc., Chicago, IL, USA).

## 3. Results

### 3.1. Accuracy and Response Time

The behavioral responses for the five CITs are shown in Table 2. The mean accuracy was almost perfect for CIT1 (97%), followed by CIT2, with environmental noise (96%); CIT3, with a target-like melodic distractor (92%); and CIT4, with two alternating targets (89%). As expected, accuracy was the lowest (67%) for the divided identification of two concurrent target contours (CIT5). A similar trend was found for reaction times across the CITs. A gradual increase was observed until CIT3 (ranging from 2896 to 2999 ms), followed by a large increase at CIT4 (3907 ms). The time required to complete CIT5 was the longest (7825 ms).

**Table 2.** Behavioral responses across the CITs (*N* = 16).

CITs	Characteristics	Accuracy		Reaction Time (ms)	
		Mean	SD	Mean	SD
1	Focused identification task	0.97	0.11	2895	842
2	Focused identification task against noise	0.96	0.11	2489	785
3	Selective identification task	0.92	0.11	2998	1443
4	Alternating identification task	0.89	0.12	3906	902
5	Divided identification task	0.67	0.17	7825	2292

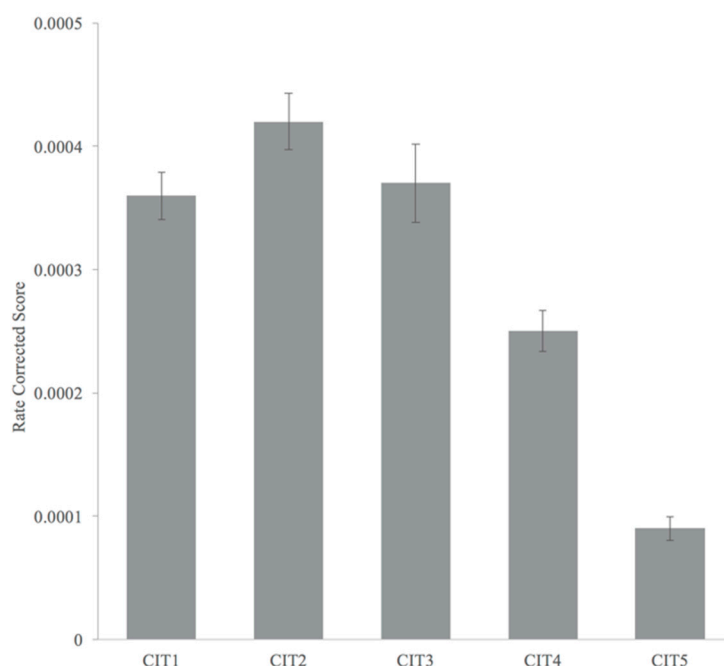
CIT: contour identification task; SD: standard deviation.

For further analysis, we performed a non-parametric Friedman test since performance accuracy and reaction time showed non-normality in distribution (Shapiro–Wilk test,  $p < 0.05$ ). Results of accuracy rendered a chi-square value of 52.80 ( $p < 0.001$ ). Nemenyi post-hoc analyses showed a significant difference between CIT1 and CIT4 ( $p < 0.05$ ), CIT1 and CIT5 ( $p < 0.001$ ), CIT2 and CIT5 ( $p < 0.001$ ),



CIT3 and CIT5 ( $p < 0.05$ ), and CIT4 and CIT5 ( $p < 0.05$ ). Results of reaction time rendered a chi-square value of 47.25 ( $p < 0.001$ ). Nemenyi post-hoc analyses showed a significant difference between CIT1 and CIT5 ( $p < 0.001$ ), CIT2 and CIT4 ( $p < 0.01$ ), CIT2 and CIT5 ( $p < 0.001$ ), and CIT3 and CIT5 ( $p < 0.001$ ).

In order to clarify the trend in behavioral responses, RCSs were calculated. We confirmed normality of the data using Shapiro–Wilk test ( $p > 0.05$ ). A repeated measures ANOVA showed a significant main effect of CIT on the RCS ( $F_{(4,60)} = 61.189$ ,  $p < 0.001$ ), indicating that performance worsened as the CIT became more difficult. Pairwise post-hoc analyses with Bonferroni correction revealed that the differences between CIT1 and CIT2, CIT2 and CIT3, and CIT1 and CIT3 were not statistically significant ( $p > 0.05$ ). All remaining pairwise tests revealed significant differences, including those between CIT3 and CIT4 ( $p < 0.001$ ) and between CIT4 and CIT5 ( $p < 0.001$ ). These findings indicated that the five CITs could be organized into three groups. Briefly, CIT1, CIT2, and CIT3 tended to be grouped together, while CIT4 and CIT5 showed clear distinctions in terms of difficulty (Figure 6).



**Figure 6.** Changes in rate corrected scores across contour identification tasks (CITs).

Altogether, our behavioral findings showed a clear distinction among the CITs. Accuracy decreased and reaction time increased as the CITs became more complicated (CIT4 and CIT5). More importantly, the behavioral responses implied that CIT4 is sensitive to the cognitive threshold of healthy young adults. Furthermore, CIT5 is the most challenging task in terms of general attentional capability.

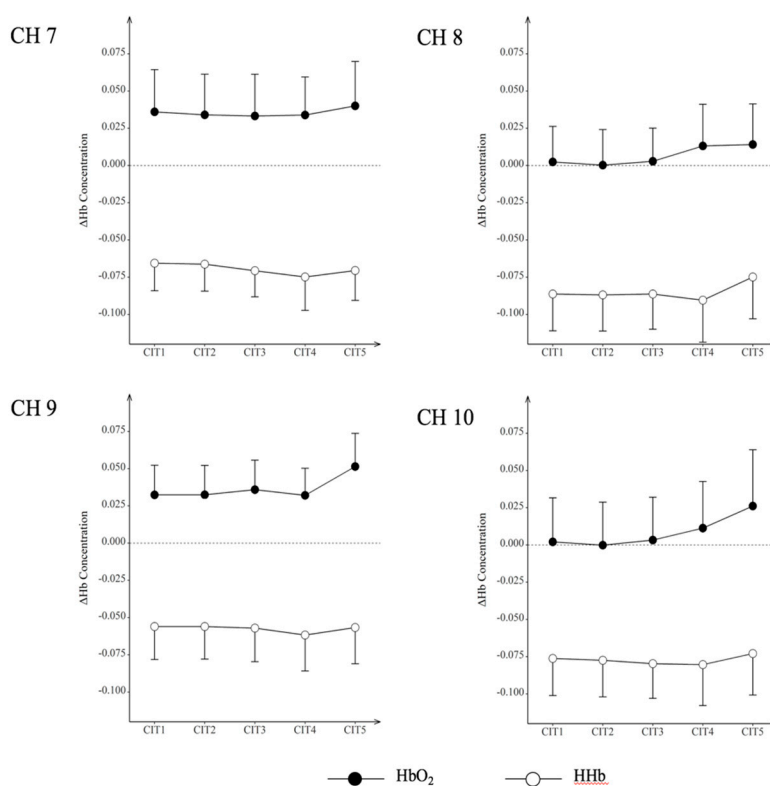
### 3.2. Hemodynamic Responses

We analyzed hemodynamic responses using HbO<sub>2</sub> and HHb across the CITs obtained from the frontopolar areas (BA10 and BA11). Table 3 and Figure 7 presents the overall hemodynamic changes across these channels, i.e., HbO<sub>2</sub> and HHb at Channel 7 (CH7), CH8, CH9, and CH10. Overall, HbO<sub>2</sub> increased and HHb decreased while the level of CITs increased. Subsequently, we performed a repeated measures ANOVA to identify the channels that may indicate cognitive loads across CITs. The findings showed a significant main effect of CIT on HbO<sub>2</sub> ( $F_{(4,60)} = 3.204$ ,  $p < 0.05$ ) in CH9. Post-hoc pairwise comparisons revealed that HbO<sub>2</sub> increased significantly between CIT1 and CIT5, CIT2 and CIT5, and CIT4 and CIT5 ( $p < 0.05$ , respectively).

**Table 3.** Descriptive statistics across CITs.

CIT	Oxy/Deoxygenation	CH7		CH8		CH9		CH10	
		M	SD	M	SD	M	SD	M	SD
1	HbO <sub>2</sub>	0.036	0.028	0.002	0.024	0.032	0.020	0.002	0.030
	HHb	−0.066	0.019	−0.086	0.025	−0.056	0.022	−0.076	0.025
2	HbO <sub>2</sub>	0.034	0.027	0.000	0.024	0.032	0.02	0.000	0.029
	HHb	−0.066	0.018	−0.087	0.024	−0.056	0.022	−0.077	0.025
3	HbO <sub>2</sub>	0.033	0.028	0.003	0.022	0.036	0.02	0.003	0.029
	HHb	−0.071	0.018	−0.086	0.024	−0.057	0.023	−0.080	0.023
4	HbO <sub>2</sub>	0.034	0.026	0.013	0.028	0.032	0.018	0.011	0.031
	HHb	−0.075	0.022	−0.090	0.028	−0.062	0.024	−0.080	0.028
5	HbO <sub>2</sub>	0.04	0.03	0.014	0.027	0.051	0.022	0.026	0.038
	HHb	−0.071	0.020	−0.075	0.028	−0.057	0.024	−0.073	0.028

CIT: contour identification task; CH: channel; M: mean; SD: standard deviation; HbO<sub>2</sub>: oxygenated hemoglobin; HHb: deoxygenated hemoglobin.



**Figure 7.** Changes in HbO<sub>2</sub> and HHb across contour identification tasks. HbO<sub>2</sub>: oxygenated hemoglobin; HHb: deoxygenated hemoglobin.

#### 4. Discussion

The main purpose of this study was to examine how various components and textures of music can impose different levels of cognitive loads, as evidenced by behavioral and hemodynamic responses. In our behavioral findings, accuracy and reaction times together indicated that the CITs can impose different levels of cognitive load. A gradual increase in RCS from CIT1 to CIT3 was followed by an abrupt increase between CIT3 and CIT4, and again between CIT4 and CIT5. The hemodynamic findings showed a significant increase between CIT1 and CIT5, CIT2 and CIT5, and CIT4 and CIT5. Collectively, the findings indicate that the perceived cognitive load can differ depending on the type of properties of the auditory stimuli and the manner of organizing the stimuli. Individuals can process two simultaneous auditory information provided in a selective attention task (i.e., CIT3) without an increase in cognitive load, while significant increases in cognitive load were observed in processing

two auditory stimuli provided in a shifting or divided attention task (i.e., CIT4 and CIT5). At CH9, hemodynamic changes between CIT3 and CIT4 were non-significant; however, those at CH7, CH8, and C10 showed an overall increasing tendency, suggesting overall increases in cognitive load in the prefrontal regions.

#### 4.1. Lessons from Behavioral and Hemodynamic Findings

Our findings revealed that healthy young adults performed tasks involving single-target melody perception (CIT1 and CIT2) almost perfectly. Additional cognitive loads between CIT1 and CIT2 were not indicated despite environmental noise presented against target contours (note that accuracy was similar and that reaction time even showed non-significant decreases in CIT2). A similar tendency was observed in CIT3. The findings indicate that melodic contours without a distractor or with a non-target-like distractor can be autonomously processed by healthy young adults. These results are consistent with the lack of effort required to process melodic components, as evidenced in a mismatch negativity study and its magnetic counterpart [62,84,85], as well as in an event-related potentials study [57]. Moreover, melody perception against melody-like distractors (CIT3) did not increase cognitive loads, indicating that intact selective attention (inhibitory attention control over irrelevant stimuli) is characteristic of healthy young adults [86–88].

In more challenging tasks, such as CIT4, in which participants were asked to shift their attentional focus from one timbre to another (note that CIT3 did not require such a switch), performance deteriorated significantly, as evidenced by both reaction time and accuracy. The present findings are similar to those reported by previous studies, which demonstrated that shifting attention to a particular location or timbre increases cognitive load and perceived difficulty of the task [89]. When participants were presented with CIT5, in which melodic contour directions must be identified simultaneously, their task performance became much worse; CIT5 appeared to require participants to listen to the directions of melodic contours either in the manner of rapidly alternating or continuously applied attentional focus between multiple streams of CIT5 [60,90,91]. The ability to perform this has been reported to be the most difficult to achieve in the auditory modality, and both typical and clinical populations showed the worst performance in a similar type of auditory attention task [2,45,67,92,93].

In hemodynamic analysis, the current findings showed a significant increase in HbO<sub>2</sub> and a non-significant but continuous decrease in HHb across CITs. Increases in HbO<sub>2</sub> were prominent between CIT1 and CIT5, CIT2 and CIT5, and CIT4 and CIT5, indicating increases in cognitive load during CIT5 (i.e., divided attention task for CIT5). Decreases in HHb were accompanied with cognitive load changes associated with HbO<sub>2</sub>. In general, both indicators (i.e., increased HbO<sub>2</sub> and decreased HHb) accompany each other, indicating an increase in cognitive load [94,95]. The increase in hemodynamic responses were obvious in the frontopolar area (BA10). This area is known for its role in modulating attention performance in the auditory modality. Laguë-Beauvais et al. [68] reported that the BA10 receives both verbal and non-verbal information from the superior temporal gyrus and becomes more active as the task complexity increases. Plakke and Romanski [69] also indicated that the BA10 manages auditory tasks that require heavy cognitive processing. Significant increases in HbO<sub>2</sub> in the frontopolar areas, thus, indicated an increased involvement of these areas in performing the given auditory task with increased cognitive loads.

The current findings can be interpreted as follows: (1) together, the characteristics of auditory stimuli and tasks can generate different levels of cognitive loads, and (2) cognitive load changes in the auditory modality can be measured by the hemodynamic activation in the frontopolar areas. Further, they indicate that individuals (typically younger adults) easily perceive multiple melodic contours that are presented simultaneously and process them in a goal-directed manner (i.e., CIT3) without any perceived increase in the cognitive load, both behaviorally and hemodynamically. The performance was better than we expected; however, there certainly was a limited capacity of allocating cognitive resources involving auditory information in our real-world auditory surroundings, in which multi-layered sounds exist. Behavioral and hemodynamic responses to CIT5 are considered

the potential markers of the highest cognitive capability in healthy young adults. Such cognitive threshold can further provide a guideline for earcon design in a public environment.

#### 4.2. Lessons from Additional Analyses on Directional Congruence and Timbre Similarity

What drew our attention here is the cognitive load involved in the performance of CIT3 and CIT4, in which two melodic contours are presented simultaneously. As mentioned earlier, our auditory environments embrace multi-layered sound streams. This is not exceptional in public environments, in which various types of products and services are in operation. For example, during sound design in the context of human–robot interaction, there are multi-layered sound streams, including sounds of machinery platforms, monitoring, alarming, and feedback sounds [96]. Since each sound aims to deliver certain information, it is very important to consider the priority of the information, the individuals' cognitive ability, and contexts. Our findings clearly indicate that CIT5 is not comparable to the other CITs. Hence, for earcon design for a public use, CIT3 and CIT4 were further examined to identify the features of evacuation earcon design that might trigger a perception advantage.

In CIT3 and CIT4, the target and distractor have either congruent or incongruent pitch contour direction. That is, two pitch contours moving in the same direction are considered congruent, and those moving in different directions are called incongruent. In a similar vein, pitch contours with similar timbres are considered similar. The timbre of the string and flute are similar in terms of temporal and spectral dimensions, whereas the piano is dissimilar to both the flute and string [85–87]. We, thus, performed an additional analysis to examine stimulus- and task-specific effects on behavioral performance. The items of CIT3 and CIT4 were recorded in relation to direction congruence and timbre similarity between target and target-like contours (Table 4).

**Table 4.** Accuracy based on timbre similarity and direction congruence in CIT3 and CIT4.

Timbre Similarity	Direction Congruence	Number of Items	Mean	SD
Similar	Congruent	27	1.00	0.00
	Incongruent	176	0.84	0.37
Dissimilar	Congruent	69	0.97	0.17
	Incongruent	340	0.95	0.21

CIT, contour identification task; SD, standard deviation.

A two-way ANOVA revealed a significant effect of direction congruence ( $F_{(1,611)} = 7.79, p < 0.01$ ) but not a statistically significant difference based on timbre similarity ( $F_{(1,611)} = 1.71, p > 0.05$ ). Interestingly, there was a significant interaction between direction congruence and timbre similarity ( $F_{(1,611)} = 4.93, p < 0.05$ ), which suggests that direction identification of incongruent contours is affected by timbre similarity. When the directions between the target and the distractor were congruent, accuracy was maintained (mean, 0.97 vs. 1.00) irrespective of timbre similarity. However, when their directions were incongruent, accuracy was much lower when the timbre of the target and the distractor was similar (mean, 0.84 vs. 0.95). Taken together, the results indicate that both the pitch contour and timbre can be supportively used for earcon design.

#### 4.3. Implications for Auditory Earcon Design in Public Environments

In situations in public environments, specifically when each of the various streams of earcon needs to be clearly perceived and understood by individuals, it is important to organize the acoustic features of music based on the psychology of cognitive load. When the timbre of earcon is clearly different acoustically as well as perceptually, the contour directions are not necessarily incongruent. That is, the pitch contour itself is not of a higher value if the distracting music stimuli have different pitch contours and similar timbres. However, when the contours are different, the timbres must be

dissimilar to each other. Other than that, there is a very high chance that information from different earcons will be integrated, forming a single percept [45,91].

Timbre refers to the quality that makes one particular musical sound different from another [97–99] even when they have the same pitch. Schröter et al. [100] found that two concurrent pitch contours presented with similar timbres are likely to generate an integrated single percept, indicating auditory redundancy gains. In a similar vein, Galvin, Fu, and Oba [101] stated that direction identification is significantly affected by the level of competition in terms of timbre—the existence of masker sound lowers identification performance, and the performance is even worse when the timbre of the masker becomes stronger. This implies that earcons are perceived and understood without any additional increases in cognitive load and thus enhance effective and efficient decision-making and action when two concurrently present earcons have congruent/incongruent directional information and dissimilar timbres.

This study has several limitations, including the small sample size ( $N = 16$ ) and the skewed female-to-male ratio. Some previous fNIRS studies have reported sex-related differences [102–104]. This issue remains controversial, with different findings depending on the areas of behavioral function. For example, the difference is more obvious when emotional aspects are the focus of the study, including emotional stress [104] and cooperation [102]. In this study, we found no sex-related differences in any of the behavioral or hemodynamic responses. However, in future studies, we will directly address this issue by controlling the sample size, sex ratio, and age (note that our study only included participants who were in their early to mid-20s). Additionally, in the present study, we selected synthesized instrument timbres from default MIDI instruments based on an approach of music psychology. Given that timbre is a multidimensional attribute that deals with spectral and temporal features [105], brightness and sharpness—which correspond to the harmonic spectral centroid and envelope attack time—for example, should have been considered in stimulus and task design. In future studies, the multidimensionality of timbre needs to be carefully designed, and the effect of such features on cognitive load changes needs to be investigated.

Another limitation of this study is that neurophysiological responses were measured using fNIRS. Although fNIRS is portable and relatively tolerant to motion, unlike other brain imaging methods [106,107], it measures hemodynamic responses only at the cortical level, while techniques such as functional magnetic resonance imaging and magnetoencephalography, which can measure responses at the sub-cortical level or determine the influence of emotion (e.g., panic and fear), are important in fire evacuation scenarios. Further, fNIRS only examines the frontopolar activation; therefore, the relationships between the PFC and other brain regions have not been included in the current study. Thus, we must be cautious before generalizing that the cognitive thresholds observed in CIT4 and CIT5 are only dependent on the role of the PFC.

## 5. Conclusions

Providing timely, efficient information is critical, and auditory signals are predominantly used to achieve this because they allow simultaneous processing in some contexts. The present study employed a reductionist approach in relation to the basic processing of complex auditory signals. We examined cognitive load changes associated with different types of CITs, which can simulate various types of attention required in a real-world auditory environment. Our findings show that melodic CITs can impose different levels of cognitive loads. Moreover, melodic CITs could be a potentially scalable marker to measure cognitive threshold in terms of auditory perception. The cognitive threshold for young adults can shift, and attention can be divided. Further, additional analysis (i.e., analysis of contour congruence and timbre similarity) revealed that appropriate tuning of the relationship between timbre and pitch contour can lower the perceived cognitive load and, thus, can be an effective design strategy for earcon in a public environment, even with concurrent streams without cognitive overload. As a concluding remark, we hope that, by applying the music-based tasks proposed in this article to other conditions and contexts, researchers can provide the insights required to understand the

relationship between music and cognition and, thus, to design a more effective and esthetic nonverbal auditory earcon that is applicable to public auditory environments.

**Author Contributions:** Conceptualization: E.J.; Methodology: J.K.; Formal Analysis: E.J. and G.J.; Visualization: G.S.; Investigation: E.J.; Writing—Original Draft Preparation: E.J. and H.R.; Writing—Review & Editing: E.J., H.R., J.K. and G.J.

**Funding:** This research was funded by a National Research Foundation of Korea grant funded by the Korean government (NRF-2018R1C1B5044305).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Brewster, S.A. Providing a structured method for integrating non-speech audio into human-computer interfaces. *Hum.-Comput. Interact.* **1994**, *277*.
2. McGookin, D.K.; Brewster, S.A. Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *ACM Trans. Appl. Percept.* **2004**, *1*, 130–155. [[CrossRef](#)]
3. Blattner, M.M.; Sumikawa, D.A.; Greenberg, R.M. Earcons and icons: Their structure and common design principles. *Hum.-Comput. Interact.* **1989**, *4*, 11–44. [[CrossRef](#)]
4. Sanders, M.S.; McCormick, E.J. *Human Factors in Engineering and Design*; McGraw-Hill: New York, NY, USA, 1998; ISBN 0-07-112826-3.
5. Kirschner, P.A. Cognitive load theory implications of cognitive load theory design-learning. *Learn. Instr.* **2002**, *12*, 1–10. [[CrossRef](#)]
6. Oviatt, S. Human-centered design meets cognitive load theory. In Proceedings of the 14th Annual ACM International Conference on Multimedia-MULTIMEDIA '06, Santa Barbara, CA, USA, 23–27 October 2006; p. 871.
7. Königshulte, A. Sound as effective design feature in multimedia learning—benefits and drawbacks from a cognitive load theory perspective. In Proceedings of the 12th International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA), Greater Dublin, Ireland, 24–26 October 2015; pp. 75–83.
8. Hermann, T.; Hunt, A.; Neuhoff, J.G. *The Sonification Handbook*; Logos Publishing House: Berlin, Germany, 2011; ISBN 9783832528195.
9. Wogalter, M.S.; Conzola, V.C.; Smith-Jackson, T.L. Research-based guidelines for warning design and evaluation. *Appl. Ergon.* **2002**, *33*, 219–230. [[CrossRef](#)]
10. Edworthy, J. The design and implementation of non-verbal auditory warnings. *Appl. Ergon.* **1994**, *25*, 202–210. [[CrossRef](#)]
11. Edworthy, J. Designing effective alarm sounds. *Biomed. Instrum. Technol.* **2011**, *45*, 290–294. [[CrossRef](#)] [[PubMed](#)]
12. Fritz, J.B.; Elhilali, M.; David, S.V.; Shamma, S.A. Auditory attention-focusing the searchlight on sound. *Curr. Opin. Neurobiol.* **2007**, *17*, 437–455. [[CrossRef](#)] [[PubMed](#)]
13. Stevens, C.; Brennan, D.; Petocz, A.; Howell, C. Designing informative warning signals: Effects of indicator type, modality, and task demand on recognition speed and accuracy. *Adv. Cogn. Psychol.* **2009**, *5*, 84–90. [[CrossRef](#)] [[PubMed](#)]
14. Edworthy, J.; Hellier, E. Alarms and human behaviour: Implications for medical alarms. *Br. J. Anaesth.* **2006**, *97*, 12–17. [[CrossRef](#)] [[PubMed](#)]
15. Peres, S.C.; Best, V.; Brock, D.; Shinn-Cunningham, B.; Frauenberger, C.; Hermann, T.; Neuhoff, J.G.; Nickerson, L.V.; Stockman, T. Auditory Interfaces. In *HCI Beyond the GUI. Design for Haptic, Speech, Olfactory and Other Nontraditional Interfaces*; Kortum, P., Ed.; Morgan Kaufman: Burlington, MA, USA, 2008; pp. 147–195, ISBN 978-0-12-374017-5.
16. Fagerlön, J. Informative auditory warning signals: A review of published material within the HCI and Auditory Display communities. In Proceedings of the 39th Nordic Ergonomics Society Conference, Lysekil, Sweden, 1–3 October 2007; pp. 1–3.
17. Graham, R. Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics* **1999**, *42*, 1233–1248. [[CrossRef](#)] [[PubMed](#)]
18. Nilsson, D.; Frantzich, H. Design of Voice Alarms—The benefit of mentioning fire and the use of a synthetic voice. In *Pedestrian and Evacuation Dynamics*; Springer: Berlin, Heidelberg, 2008; pp. 135–144, ISBN 9783642045035.

19. Pressnitzer, D.; Sayles, M.; Micheyl, C.; Winter, I.M. Perceptual Organization of Sound Begins in the Auditory Periphery. *Curr. Biol.* **2008**, *18*, 1124–1128. [[CrossRef](#)] [[PubMed](#)]
20. Alain, C.; Arnott, S.R.; Picton, T.W. Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *J. Exp. Psychol. Hum. Percept. Perform.* **2001**, *27*, 1072. [[CrossRef](#)] [[PubMed](#)]
21. Mirsky, A.F.; Anthony, B.J.; Duncan, C.C.; Ahearn, M.B.; Kellam, S.G. Analysis of the elements of attention: A neuropsychological approach. *Neuropsychol. Rev.* **1991**, *2*, 109–145. [[CrossRef](#)] [[PubMed](#)]
22. Ocasio, W. Attention to attention. *Organ. Sci.* **2011**, *22*, 1286–1296. [[CrossRef](#)]
23. Posner, M.I.; Petersen, S.E. The attention system of the human brain. *Annu. Rev. Neurosci.* **1990**, *13*, 25–42. [[CrossRef](#)] [[PubMed](#)]
24. Petersen, S.E.; Posner, M.I. The attention system of the human brain: 20 Years After. *Annu. Rev. Neurosci.* **2012**, *35*, 73–89. [[CrossRef](#)] [[PubMed](#)]
25. Cohen, R.A. *The Neuropsychology of Attention*; Springer: Boston, MA, USA, 2014; ISBN 9780387726397.
26. Broadbent, D.E. The role of auditory localization in attention and memory span. *J. Exp. Psychol.* **1954**, *47*, 191–196. [[CrossRef](#)] [[PubMed](#)]
27. Kim, S. The cocktail party effect. *Am. Acad. Neurol.* **2013**, *9*, 13. [[CrossRef](#)]
28. Bronkhorst, A.W. The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Atten. Percept. Psychophys.* **2015**, *77*, 1465–1487. [[CrossRef](#)] [[PubMed](#)]
29. Lavie, N. Distracted and confused? Selective attention under load. *Trends Cogn. Sci.* **2005**, *9*, 75–82. [[CrossRef](#)] [[PubMed](#)]
30. Lavie, N.; De Fockert, J. The role of working memory in attentional capture. *Psychon. Bull. Rev.* **2005**, *12*, 669–674. [[CrossRef](#)] [[PubMed](#)]
31. Lavie, N. Attention, distraction, and cognitive control under load. *Curr. Dir. Psychol. Sci.* **2010**, *19*, 143–148. [[CrossRef](#)]
32. Ayres, P.; Paas, F. Cognitive load theory: New directions and challenges. *Appl. Cogn. Psychol.* **2012**, *26*, 827–832. [[CrossRef](#)]
33. Botvinick, M.M.; Cohen, J.D.; Carter, C.S. Conflict monitoring and anterior cingulate cortex: An update. *Trends Cogn. Sci.* **2004**, *8*, 539–546. [[CrossRef](#)] [[PubMed](#)]
34. Gruber, O.; Goschke, T. Executive control emerging from dynamic interactions between brain systems mediating language, working memory and attentional processes. *Acta Psychol. (Amst.)* **2004**, *115*, 105–121. [[CrossRef](#)] [[PubMed](#)]
35. Miller, E.K.; Cohen, J.D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **2001**, *24*, 167–202. [[CrossRef](#)] [[PubMed](#)]
36. Holmqvist, K.; Nyström, M.; Andersson, R.; Dewhurst, R.; Jarodzka, H.; Van de Weijer, J. *Eye Tracking: A Comprehensive Guide to Methods and Measures*; Oxford University Press: Oxford, UK, 2011; ISBN 9780198738596.
37. Oviatt, S.; Coulston, R.; Lunsford, R. When do we interact multimodally? Cognitive load and multimodal communication patterns. In Proceedings of the 6th Annual ACM International Conference on Multimodal Interfaces, State College, PA, USA, 13–15 October 2004; p. 129.
38. Paletta, L.; Wagner, V.; Kallus, W.; Schrom-Feiertag, H.; Schwarz, M.; Pszeida, M.; Ladstätter, S.; Matyus, T. Human factor modeling from wearable sensed data for evacuation based simulation scenarios. In Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics (AHFE 2014), Krakow, Poland, 19–23 July 2014.
39. Barreto, A.B.; Jacko, J.A.; Hugh, P. Impact of spatial auditory feedback on the efficiency of iconic human-computer interfaces under conditions of visual impairment. *Comput. Human Behav.* **2007**, *23*, 1211–1231. [[CrossRef](#)]
40. Marston, J.R.; Loomis, J.M.; Klatzky, R.L.; Golledge, R.G.; Smith, E.L. Evaluation of spatial displays for navigation without sight. *ACM Trans. Appl. Percept.* **2006**, *3*, 110–124. [[CrossRef](#)]
41. Singh, D. Spatial auditory based interaction model for driver assistance system. *World Appl. Sci. J.* **2012**, *20*, 560–564. [[CrossRef](#)]
42. Roginska, A.; Childs, E.; Johnson, M.K. Monitoring real-time data: A sonification approach. In Proceedings of the 12th International Conference on Auditory Display (ICAD2006), London, UK, 20–23 June 2006.
43. Levinson, J. *The Oxford Handbook of Aesthetics*; Oxford University Press: Oxford, UK, 2009; ISBN 9780191577239.

44. Choi, I.; Wang, L.; Bharadwaj, H.; Shinn-Cunningham, B. Individual differences in attentional modulation of cortical responses correlate with selective attention performance. *Hear. Res.* **2014**, *314*, 10–19. [[CrossRef](#)] [[PubMed](#)]
45. Bigand, E.; McAdams, S.; Forêt, S. Divided attention in music. *Int. J. Psychol.* **2000**, *35*, 270–278. [[CrossRef](#)]
46. Liljedahl, M.; Fagerlön, J. Methods for sound design: A review and implications for research and practice. In Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound, Piteå, Sweden, 15–17 September 2010; ACM: New York, NY, USA; pp. 1–8.
47. Hodges, D.; Sebald, D.C. *Music in the Human Experience: An Introduction to Music Psychology*; Routledge: Abingdon-on-Thames, UK, 2010; ISBN 0203834976.
48. National Fire Protection Association. *Standard on Disaster/Emergency Management and business Continuity Programs*; National Fire Protection Association: Quincy, MA, USA, 2013; ISBN 978-145590648-2.
49. Bukowski, R.; Moore, W.D. *Fire Alarm Signaling Systems*; National Fire Protection Association: Quincy, MA, USA, 1994; ISBN 978-0877653998.
50. Allen, E.J.; Oxenham, A.J. Symmetric interactions and interference between pitch and timbre. *J. Acoust. Soc. Am.* **2014**, *135*, 1371–1379. [[CrossRef](#)] [[PubMed](#)]
51. Li, D.; Cowan, N.; Sauls, J.S. Estimating working memory capacity for lists of nonverbal sounds. *Atten. Percept. Psychophys.* **2012**, *75*, 145–160. [[CrossRef](#)] [[PubMed](#)]
52. Golubock, J.L.; Janata, P. Keeping timbre in mind: Working memory for complex sounds that can't be verbalized. *J. Exp. Psychol. Hum. Percept. Perform.* **2013**, *39*, 399–412. [[CrossRef](#)] [[PubMed](#)]
53. Pearce, M.T.; Wiggins, G.A. Auditory Expectation: The information dynamics of music perception and cognition. *Top. Cogn. Sci.* **2012**, *4*, 625–652. [[CrossRef](#)] [[PubMed](#)]
54. Deutsch, D. Grouping Mechanisms in Music. In *The Psychology of Music*; Academic Press: San Diego, CA, USA, 2013; pp. 183–248. ISBN 9780123814609.
55. Deutsch, D. The Processing of Pitch Combinations. In *The Psychology of Music*; Academic Press: San Diego, CA, USA, 2013; pp. 249–325. ISBN 9780123814609.
56. Crawley, E.J.; Acker-Mills, B.E.; Pastore, R.E.; Weil, S. Change detection in multi-voice music: The role of musical structure, musical training, and task demands. *J. Exp. Psychol. Hum. Percept. Perform.* **2002**, *28*, 367. [[CrossRef](#)] [[PubMed](#)]
57. Demorest, S.M.; Osterhout, L. ERP responses to cross-cultural melodic expectancy violations. *Ann. N. Y. Acad. Sci.* **2012**, *1252*, 152–157. [[CrossRef](#)] [[PubMed](#)]
58. Macken, W.J.; Tremblay, S.; Houghton, R.J.; Nicholls, A.P.; Jones, D.M. Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *J. Exp. Psychol. Hum. Percept. Perform.* **2003**, *29*, 43. [[CrossRef](#)] [[PubMed](#)]
59. Snyder, J.S.; Alain, C.; Picton, T.W. Effects of attention on neuroelectric correlates of auditory stream segregation. *J. Cogn. Neurosci.* **2006**, *18*, 1–13. [[CrossRef](#)] [[PubMed](#)]
60. Janata, P.; Tillmann, B.; Bharucha, J.J. Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cogn. Affect. Behav. Neurosci.* **2002**, *2*, 121–140. [[CrossRef](#)] [[PubMed](#)]
61. Lee, Y.-S.; Janata, P.; Frost, C.; Hanke, M.; Granger, R. Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI. *Neuroimage* **2011**, *57*, 293–300. [[CrossRef](#)] [[PubMed](#)]
62. Fujioka, T.; Trainor, L.J.; Ross, B.; Kakigi, R.; Pantev, C. Automatic encoding of polyphonic melodies in musicians and nonmusicians. *J. Cogn. Neurosci.* **2005**, *17*, 1578–1592. [[CrossRef](#)] [[PubMed](#)]
63. Sawhney, N.; Schmandt, C. Nomadic radio: Speech and audio interaction for contextual messaging in nomadic environments. *ACM Trans. Comput. Interact.* **2000**, *7*, 353–383. [[CrossRef](#)]
64. Gaver, W.W.; Smith, R.B.; O'Shea, T. Effective sounds in complex systems: The Arkola simulation. *Proc. CHI 1991*, *1991*, 85–90. [[CrossRef](#)]
65. Gaver, W.W. Auditory interfaces. In *Handbook of Human-Computer Interaction (Second Edition)*; Helander, M.G., Landauer, T.K., Prabhu, P.V., Eds.; Elsevier: Amsterdam, The Netherlands, 1997; pp. 1003–1041, ISBN 978-0-444-81862-1.
66. Jeong, E. Psychometric validation of a music-based attention assessment: Revised for patients with traumatic brain injury. *J. Music Ther.* **2013**, *50*, 66–92. [[CrossRef](#)] [[PubMed](#)]
67. Jeong, E.; Lesiuk, T.L. Development and preliminary evaluation of a music-based attention assessment for patients with traumatic brain injury. *J. Music Ther.* **2011**, *48*, 551–572. [[CrossRef](#)] [[PubMed](#)]



68. Laguë-Beauvais, M.; Brunet, J.; Gagnon, L.; Lesage, F.; Bherer, L. A fNIRS investigation of switching and inhibition during the modified Stroop task in younger and older adults. *Neuroimage* **2013**, *64*, 485–495. [[CrossRef](#)] [[PubMed](#)]
69. Plakke, B.; Romanski, L.M. Auditory connections and functions of prefrontal cortex. *Front. Neurosci.* **2014**, *8*, 199. [[CrossRef](#)] [[PubMed](#)]
70. McKendrick, R.; Ayaz, H.; Olmstead, R.; Parasuraman, R. Enhancing dual-task performance with verbal and spatial working memory training: Continuous monitoring of cerebral hemodynamics with NIRS. *Neuroimage* **2014**, *85*, 1014–1026. [[CrossRef](#)] [[PubMed](#)]
71. Ogawa, Y.; Kotani, K.; Jimbo, Y. Relationship between working memory performance and neural activation measured using near-infrared spectroscopy. *Brain Behav.* **2014**, *4*, 544–551. [[CrossRef](#)] [[PubMed](#)]
72. Fishburn, F.A.; Norr, M.E.; Medvedev, A.V.; Vaidya, C.J. Sensitivity of fNIRS to cognitive state and load. *Front. Hum. Neurosci.* **2014**, *8*, 76. [[CrossRef](#)] [[PubMed](#)]
73. Rustichini, A. Decision-making and neuroeconomics. In *Encyclopedia of Neuroscience*; Squire, L.R., Ed.; Elsevier: New York, NY, USA, 2010; pp. 323–328. ISBN 9780080450469.
74. Agostini, G.; Longari, M.; Pollastri, E. Musical instrument timbres classification with spectral features. *EURASIP J. Appl. Signal Process.* **2003**, *2003*, 5–14. [[CrossRef](#)]
75. Jeong, E.; Ryu, H. Melodic contour identification reflects the cognitive threshold of aging. *Front. Aging Neurosci.* **2016**, *8*. [[CrossRef](#)] [[PubMed](#)]
76. Peck, E.M.M.; Yuksel, B.F.; Ottley, A.; Jacob, R.J.K.; Chang, R. Using fNIRS brain sensing to evaluate information visualization interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 27 April–2 May 2013; ACM: New York, NY, USA; pp. 473–482.
77. Yasumura, A.; Inagaki, M.; Hiraki, K. Relationship between neural activity and executive function: An NIRS study. *ISRN Neurosci.* **2014**, *2014*. [[CrossRef](#)] [[PubMed](#)]
78. Akgul Sankur, B.; Akin, A.C.B. Spectral analysis of event-related hemodynamic responses in functional near infrared spectroscopy. *J. Comput. Neurosci.* **2005**, *18*, 67–83. [[CrossRef](#)] [[PubMed](#)]
79. Bauernfeind, G.; Scherer, R.; Pfurtscheller, G.; Neuper, C. Single-trial classification of antagonistic oxyhemoglobin responses during mental arithmetic. *Med. Biol. Eng. Comput.* **2011**, *49*, 979–984. [[CrossRef](#)] [[PubMed](#)]
80. Morren, G.; Wolf, M.; Lemmerling, P.; Wolf, U.; Choi, J.H.; Gratton, E.; De Lathauwer, L.; Van Huffel, S. Detection of fast neuronal signals in the motor cortex from functional near infrared spectroscopy measurements using independent component analysis. *Med. Biol. Eng. Comput.* **2004**, *42*, 92–99. [[CrossRef](#)] [[PubMed](#)]
81. Herff, C.; Heger, D.; Putze, F.; Hennrich, J.; Fortmann, O.; Schultz, T. Classification of mental tasks in the prefrontal cortex using fNIRS. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2013**, *2013*, 2160–2163. [[CrossRef](#)] [[PubMed](#)]
82. Woltz, D.J.; Was, C.A. Availability of related long-term memory during and after attention focus in working memory. *Mem. Cognit.* **2006**, *34*, 668–684. [[CrossRef](#)] [[PubMed](#)]
83. Vandierendonck, A. A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behav. Res. Methods* **2017**, *49*, 653–673. [[CrossRef](#)] [[PubMed](#)]
84. Trainor, L.J.; Marie, C.; Bruce, I.C.; Bidelman, G.M. Explaining the high voice superiority effect in polyphonic music: Evidence from cortical evoked potentials and peripheral auditory models. *Hear. Res.* **2014**, *308*, 60–70. [[CrossRef](#)] [[PubMed](#)]
85. Trainor, L.J.; McDonald, K.L.; Alain, C. Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *J. Cogn. Neurosci.* **2002**, *14*, 430–442. [[CrossRef](#)] [[PubMed](#)]
86. Hugenschmidt, C.E.; Peiffer, A.M.; McCoy, T.P.; Hayasaka, S.; Laurienti, P.J. Preservation of crossmodal selective attention in healthy aging. *Exp. Brain Res.* **2009**, *198*, 273–285. [[CrossRef](#)] [[PubMed](#)]
87. Couperus, J.W. Perceptual Load Influences Selective Attention across Development. *Dev. Psychol.* **2011**, *47*, 1431–1439. [[CrossRef](#)] [[PubMed](#)]
88. Quigley, C.; Müller, M.M. Feature-selective attention in healthy old age: A selective decline in selective attention? *J. Neurosci.* **2014**, *34*, 2471–2476. [[CrossRef](#)] [[PubMed](#)]
89. Hill, K.T.; Miller, L.M. Auditory attentional control and selection during cocktail party listening. *Cereb. Cortex* **2010**, *20*, 583–590. [[CrossRef](#)] [[PubMed](#)]

90. Gallun, F.J.; Mason, C.R.; Kidd, G. Task-dependent costs in processing two simultaneous auditory stimuli. *Percept. Psychophys.* **2007**, *69*, 757–771. [[CrossRef](#)] [[PubMed](#)]
91. Shinn-Cunningham, B.G.; Ihlefeld, A. Selective and divided attention: extracting information from simultaneous sound sources. In Proceedings of the 10th Meeting of the International Conference on Auditory Display, Sydney, Australia, 6–9 July 2004; pp. 1–8.
92. Azouvi, P.; Couillet, J.; Leclercq, M.; Martin, Y.; Asloun, S.; Rousseaux, M. Divided attention and mental effort after severe traumatic brain injury. *Neuropsychologia* **2004**, *42*, 1260–1268. [[CrossRef](#)] [[PubMed](#)]
93. Blanchet, S.; Paradis-Giroux, A.-A.; Pépin, M.; McKerral, M. Impact of divided attention during verbal learning in young adults following mild traumatic brain injury. *Brain Inj.* **2009**, *23*, 111–122. [[CrossRef](#)] [[PubMed](#)]
94. Cui, X.; Bray, S.; Reiss, A.L. Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage* **2010**, *49*, 3039–3046. [[CrossRef](#)] [[PubMed](#)]
95. Matsukawa, K.; Ishii, K.; Liang, N.; Endo, K.; Ohtani, R.; Nakamoto, T.; Wakasugi, R.; Kadowaki, A.; Komine, H. Increased oxygenation of the cerebral prefrontal cortex prior to the onset of voluntary exercise in humans. *J. Appl. Physiol.* **2015**. [[CrossRef](#)] [[PubMed](#)]
96. Jeong, E.; Kwon, G.H.; So, J. Exploring the taxonomy and associative link between emotion and function for robot sound design. In Proceedings of the 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence, Jeju, Korea, 28 June–1 July 2017.
97. Krumhansl, C.L. Why is musical timbre so hard to understand. *Struct. Percept. Electroacoust. Sound Music* **1989**, *9*, 43–53.
98. Krumhansl, C.L. *Cognitive Foundations of Musical Pitch*; Oxford University Press: Oxford, UK, 2010; ISBN 978-0195148367.
99. Lakatos, S. A common perceptual space for harmonic and percussive timbres. *Percept. Psychophys.* **2000**, *62*, 1426–1439. [[CrossRef](#)] [[PubMed](#)]
100. Schröter, H.; Ulrich, R.; Miller, J. Effects of redundant auditory stimuli on reaction time. *Psychon. Bull. Rev.* **2007**, *14*, 39–44. [[CrossRef](#)] [[PubMed](#)]
101. Galvin, J.J., III; Fu, Q.-J.; Oba, S.I. Effect of a competing instrument on melodic contour identification by cochlear implant users. *J. Acoust. Soc. Am.* **2009**, *125*, EL98–EL103. [[CrossRef](#)] [[PubMed](#)]
102. Baker, J.M.; Liu, N.; Cui, X.; Vrticka, P.; Saggat, M.; Hosseini, S.M.H.; Reiss, A.L. Sex differences in neural and behavioral signatures of cooperation revealed by fNIRS hyperscanning. *Sci. Rep.* **2016**, *6*, 26492. [[CrossRef](#)] [[PubMed](#)]
103. Li, T.; Luo, Q.; Gong, H. Gender-specific hemodynamics in prefrontal cortex during a verbal working memory task by near-infrared spectroscopy. *Behav. Brain Res.* **2010**, *209*, 148–153. [[CrossRef](#)] [[PubMed](#)]
104. Yang, H.; Zhou, Z.; Liu, Y.; Ruan, Z.; Gong, H.; Luo, Q.; Lu, Z. Gender difference in hemodynamic responses of prefrontal area to emotional stress by near-infrared spectroscopy. *Behav. Brain Res.* **2007**, *178*, 172–176. [[CrossRef](#)] [[PubMed](#)]
105. Fastl, H.; Zwicker, E. *Psychoacoustics: Facts and Models*; Springer: Berlin/Heidelberg, Germany, 2007; ISBN 978-3-540-23159-2.
106. Moriguchi, Y.; Hiraki, K. Prefrontal cortex and executive function in young children: A review of NIRS studies. *Front. Hum. Neurosci.* **2013**, *7*, 867. [[CrossRef](#)] [[PubMed](#)]
107. Sato, H.; Yahata, N.; Funane, T.; Takizawa, R.; Katura, T.; Atsumori, H.; Nishimura, Y.; Kinoshita, A.; Kiguchi, M.; Koizumi, H. A NIRS-fMRI investigation of prefrontal cortex activity during a working memory task. *Neuroimage* **2013**, *83*, 158–173. [[CrossRef](#)] [[PubMed](#)]

