GUIDANCE AND CAPTURE OF ATTENTION BY (IRRELEVANT) AUDITORY EVENTS

Ananya Mandal

Graduate School of Systemic Neurosciences
LMU Munich

Dissertation der
Graduate School of Systemic Neurosciences
Ludwig-Maximilians-Universität München

15 July 2024

Supervisor

PD Dr. Heinrich R. Liesefeld

Forschungsmethoden und Kognitive Psychologie

Universität Bremen

First Reviewer:	PD Dr. Heinrich R. Liesefeld
Second Reviewer:	Prof. Dr. Zhuanghua Shi

Date of Submission:	15.07.2024
Date of Defense:	14.03.2025

Summary

As sensory beings, we are often inundated with a multitude of stimuli, each one vying for our attention all at once. Navigating this requires us to continuously discern and prioritize the most significant stimuli while simultaneously filtering out the irrelevant ones – and is therefore a significant challenge. This ongoing task of selective attention, though complex, is an essential part of our daily lives. Fortunately, our attentional system is aided by a powerful preattentional guidance mechanism that helps prioritize locations/objects for attentional allocation depending on various bottom-up (stimulus driven factors, e.g., stimulus salience) and top-down (user-driven factors e.g., task goals) guidance factors. Likely as a side effect of this, if an irrelevant stimulus (which is presented concurrent to a target) is highly salient, our attention can sometimes be misguided, causing it to be misallocated towards the irrelevant stimulus instead of the target (attentional capture) – thereby, causing distraction.

Distraction by salient events is a widespread phenomenon. In everyday life, an unexpected loud ringtone or a sudden flicker of light can often interfere with our current behavioral goals and cause distraction. While such interference has been extensively studied using visual distractors, auditory distractors also present a similar challenge. This thesis therefore aims to explore auditory distractions in search scenarios, building on the extensive literature on visual distraction. To this end, two broad scenarios of auditory distraction are investigated: (*a*) cross-modal auditory distraction during visual search, covered in Chapters 2.1 and 2.2, and (*b*) unimodal auditory distraction, i.e., auditory distraction during auditory search, discussed in Chapter 2.3.

Chapter 2.1 investigates the impact of an irrelevant auditory stimulus on the performance of a visual search task. To our surprise, there was no effect of the auditory distractor on the performance of the visual search task observed through a series of experiments. The visual search performance remained robust against auditory distraction even when the auditory distractor was presented as a rare oddball stimulus or

had a temporal advantage. Only when the auditory modality was made globally relevant (while maintaining its irrelevance to the visual search task) by introducing a simple additional auditory task, significant cross-modal distraction was observed. This indicates that our attentional system can effectively block auditory stimuli from interfering with visual search, provided the auditory modality remains completely irrelevant.

Chapter 2.2 extends Chapter 2.1's findings by exploring the effect of auditory distractors on visual search tasks at varying levels of difficulty. Consistent with Study 1, irrelevant auditory distractors did not affect performance across any difficulty level. Consequently, as observed in Chapter 2.1, the introduction of a secondary auditory task (which made the auditory modality globally relevant) resulted in significant interference from auditory distractors.

Chapter 2.3 explores auditory distraction during auditory search, focusing on the dynamics of spatial auditory attention using the N2ac component as an ERP marker of an attention allocation towards an auditory stimulus. The study demonstrates that a salient auditory distractor causes attentional capture, that is, attention being misallocated towards the distractor instead of the target. Conversely, a less salient distractor does not capture attention but still delays response times. This research provides comprehensive ERP evidence for the spatio-temporal dynamics of attentional capture in the auditory modality, paralleling similar findings in the visual modality. Additionally, it distinguishes between two types of auditory distraction: those with and without attentional capture.

Table of Contents

1. General Introduction

In an era marked by the omnipresence of digital technologies and the consequent surge in information flow – distraction seems to be an increasingly pervasive phenomenon. Urban life, in particular, is characterized by a constant bombardment of stimuli, be they from advertisements, traffic signs, or city noises. Further exacerbating this situation is the ubiquity of smartphones, social media, and notifications – to the extent that some might call this the *"age of distraction"* (e.g., Crawford, 2015). For instance, while writing this very thesis, I've been interrupted more times than I'd like to admit, by notifications the likes of *"World's laziest cat refuses to move for vacuum"* – a scenario which I am sure many of my peers will sympathize with. Whether I have been successful in maintaining my focus on writing (by ignoring these distractions) is a different story. Nevertheless, the ubiquitous presence of distraction in our surrounding makes the scientific inquiry into distraction relevant, now more than ever!

Of course, the example above is quite complex and influenced by several external factors: such as, the time of day it is (late evening writing sessions might seem more prone to distraction than early morning ones), the length of time since the last break, how important the notification is, etc. While it is important to study these more *naturalistic* scenarios, it becomes challenging to control for multiple factors influencing the same phenomenon. Therefore, it is often easier to begin investigations with as few confounding variables as possible, starting with simpler situations and gradually building up to more complex ones. In fact, scientists across fields of study employ this reductionist approach (P. W. Anderson, 1972; Nagel, 1961) – breaking down complex phenomena into simpler components to understand their fundamental principles. For instance, only when the ionic mechanisms underlying action potentials in neurons were uncovered by Hodgkin & Huxley (1952) could we understand signal transmission in neurons; similarly, Maxwell's equations (Maxwell, 1865) unifying electricity and magnetism paved the way for the development of technologies such as radio and telecommunications. The study of distraction has similarly been carried out following this reductionist approach – focusing to first understand how attention and distraction works for simple stimuli before delving into more complex, naturalistic settings.

By no means, is the study of distraction a new phenomenon. Ancient philosophers, including Aristotle, pondered on the nature of attention and its limitations, laying the groundwork for future exploration (see Fiecconi, 2021 for a discussion on this). In the 19th century, the advent of experimental psychology saw pioneers like Wilhelm Wundt and William James delving into the mechanisms of attention and distraction. James's influential work, "The Principles of Psychology," highlighted the fluctuating nature of human attention and its susceptibility to external stimuli (James, 1890). Later, in the 1950s Broadbent's filter theory provided a systematic model for understanding selective attention (Broadbent, 1958). However, we still have much to learn about attention and distraction, and distraction remains an active subject of inquiry in the neural and psychological sciences (for example, see Gaspelin et al., 2023; Lavie, 2005, 2010; Wetzel & Schröger, 2014; Wöstmann et al., 2022).

Therefore, in this thesis, I attempt to contribute (although comparable to a mere drop in the ocean) to our understanding of distraction – focusing particularly on auditory distraction. The studies conducted here explore auditory distraction informed through the extensive literature on visual distraction. Through three studies, I explore two types of auditory distraction: (a) cross-modal auditory distraction, where auditory distractors are presented concurrent to a visual search target (Chapters 2.1 and 2.2), and (b) unimodal auditory distraction, where auditory distractors are presented alongside an auditory search target (Chapter 2.3).

1.1 Selective Attention

Our senses are constantly hit with a flood of stimuli, all vying for attention at the same time. It is impossible to process all this information simultaneously, so we only focus on a select few stimuli at any given moment. The cognitive process involved in selecting and focusing on the most important subset of stimuli for enhanced processing and extraction of information – while filtering out the rest, is known as selective attention (Wolfe, 2021). This ability is crucial for everyday functioning. It helps us manage daily activities without sensory overload by prioritizing the stimuli most relevant for immediate processing.

Selective attention has been extensively studied within the visual domain; however, it is not just limited to visual stimuli. Rather, it operates across various sensory modalities, allowing individuals to concentrate on specific inputs – be they visual, auditory, tactile, or olfactory – while simultaneously ignoring irrelevant data. Take, for instance, a driver navigating heavy traffic. They employ visual selective attention to remain focused on the road and relevant signals, ignoring distractions such as roadside billboards. Simultaneously, their auditory selective attention monitors critical sounds like sirens from ambulances, while tuning out less relevant noises like conversations from pedestrians, or music from surrounding vehicles. Likewise, tactile selective attention allows them to notice a slight tap on the shoulder from a passenger, while ignoring less important tactile sensations, like the feel of the seatbelt against their shoulder.

1.1.1 Attention in search

Searching is a frequent activity in our daily lives, even though we do not necessarily realize it. For example, consider the simple task of trying to find a specific green highlighter on your office desk – this activity forms an instance of visual search. Typically, searching involves actively looking for specific information, items, or locations; and attention plays a crucial role in the search process. In fact, attention is the key mechanism that helps identify the search target, even when it is fully visible among a field of non-targets (Wolfe, 2010). Therefore, unsurprisingly search tasks are heavily employed in studying attention in the laboratory. In these paradigms, participants generally look for a target stimulus presented among a set of non-target stimuli, which are either presented simultaneously (as in the *additional-singleton paradigm*), or sequentially (as in the *rapid serial visual presentation RSVP* paradigm). These tasks often replicate the information overload found in real-world scenarios and are particularly useful for studying how attention filters out irrelevant information by focusing on relevant details; thereby, minimizing interference during the search process (Luck & Vecera, 2002).

Visual search paradigms are the most common form of search task employed to study selective attention (Müller & Krummenacher, 2006). Typically, visual-search displays comprise of a target stimulus presented amongst several irrelevant stimuli (Liesefeld et al., 2024). Participants search for the target and their performances – mainly reaction times (RTs) and accuracies are recorded. Visual search can be characterized as efficient or inefficient based on search slopes. Efficient searches occur when the search times remains constant regardless of the number of stimuli in the display (set size). This typically happens when the target is highly salient and pops-out from the non-targets, so adding more non-targets does not cause additional interference. Conversely, in inefficient searches, the time required to find the target increases as the number of stimuli in the display grows.

1.1.2 Guided search: How does 'selection' in selective attention work?

As Wolfe (2021) points out, the *selection* process in attention is rarely random. Instead, attention is believed to be *guided* by various sources of information toward a select subset of items, thereby preventing sensory overload during the search process (Wolfe & Horowitz, 2017). In other words, in a search scenario, several *preattentive* features guide the allocation of attention toward specific items or locations, that are then selected. This forms the basic premise of the Guided Search model (Wolfe, 1994, 2021) – currently the most widely accepted and successful framework for understanding visual search.

There are two broad sources of guidance: bottom-up (salience-driven) and top-down (goal- or experience-driven). Bottom-up or stimulus driven features are considered to be a potent form of guidance, where stimulus salience (defined by local feature contrast) drives attentional priority – for example, a red apple stands out amongst many green apples, and thereby has a higher priority for attentional allocation. Top-down guidance, on the other hand, is any form of guidance which is user-driven – such as, task goals, prior experiences, etc. Beyond task goals, reward/value (B. A. Anderson et al., 2011), prior selection history (Awh et al., 2012), and scene structure (Draschkow & Võ, 2017) can also guide attention and can be considered as forms of top-down guidance (although these are sometimes characterized as distinct forms of guidance beyond top-down and bottom-up guidance, see Wolfe, 2021; Wolfe & Horowitz, 2017). Together, these various sources of information are integrated to guide attention in an informed manner during visual search.

1.1.3 Priority maps: How are the various forms of guidance integrated for attention allocation?

The information from the various sources of guidance, as described in the previous section, are integrated into an attentional *priority map* (Duncan & Humphreys, 1989; Fecteau & Munoz, 2006; Liesefeld & Müller, 2021; Wolfe, 1994, 2021) in a weighted manner. This priority map – a conceptual (Wolfe, 1994, 2021), computational (Itti & Koch, 2001), and neural (Li, 2002) framework – represents the spatial distribution of attentional priority in a visual scene. The activations within this map indicate the 'priority' of each location for attention. Attention is guided in this priority map according to a winner-take-all mechanism (Koch & Ullman, 1985) towards its highest peak of activation – i.e., attention is allocated to the most active location of the map (Wolfe, 2021).

In its classical conception, the priority map was thought to emerge from a weighted integration of both bottom-up (stimulus-driven) properties and top-down (user-driven) influences within a visual scene. The bottom-up component is often represented by a *salience map* (Itti & Koch, 2001), which highlights the most salient or noticeable elements of the scene. This salience map is then further shaped by top-down factors such as task goals, motivation, and other user-driven considerations, creating an internal priority map.

Some models, such as the Dimension Weighting Account (DWA; Found & Müller, 1996; Liesefeld & Müller, 2019), propose a hierarchy of maps known as *dimensional maps* (Liesefeld et al., 2024; Liesefeld, Liesefeld, Pollmann, et al., 2019). The dimensional maps are considered intermediate levels where location-based salience information in each dimension (e.g., orientation, or color) are represented. These saliency signals are then integrated in the master priority map as a weighted sum of the activations from the dimensional maps. While the activations of the individual dimension maps are influenced by bottom-up salience, the weights with which the dimension maps are integrated into the priority map is thought to be influenced by top-down guidance information, such as task goals and experience (Liesefeld & Müller, 2021). Ultimately, these diverse sources of information (be they from salience maps, or dimensional maps) are effectively combined in the integrated priority map (the master map) to guide attention within the visual field.

1.2 Attentional guidance vs. misguidance (attentional capture)

Selecting and attending to the most important stimuli amongst multiple simultaneous stimuli presents a significant computational challenge (Bronkhorst, 2000; Tsotsos, 1990) – such that we cannot process all the information presented to us at a given time. Consequently, due to the limitations of our information processing capacity, crucial information is given priority for attentional processing. The preattentive priority map represents this prioritization of stimuli for attentional allocation and is influenced by several guidance factors (see Chapter 1.1.2). Ultimately, selective attention is *guided* towards specific locations or stimuli effectively solving the problem of "*Where/What to attend next?*".

While efficient in most scenarios, attention is sometimes *misguided* towards a salient-but-irrelevant stimulus – which is not the intended target. This misallocation of attention is known as *attentional capture*. The impact of attentional capture is evident through increased reaction times and/or error rates, leading to a noticeable decline in search performance. Attentional capture primarily occurs due to heightened priority activations triggered by bottom-up salience signals in the priority map. Bottom-up salience causes the salient-but-irrelevant (distractor) stimuli to 'pop-out' of the display – thereby, misguiding (or, capturing) attention towards itself. However, this type of (mis-)guidance is rather short lived and transient (but see Constant & Liesefeld, 2023) in the dynamic and changing priority map (Donk & van Zoest, 2008), which quickly adjusts and reorients attention back to the correct target instead of getting stuck at the high salience distractor.

1.2.1 Irrelevant stimuli: Distractor interference vs. search facilitation

The relevance of a stimulus is often described by its contribution to achieving the main task goal. Therefore, an irrelevant stimulus is one whose processing does not contribute in achieving this goal (see Liesefeld et al., 2024). That is, an irrelevant stimulus in a search display is everything but the target.

The irrelevant stimulus may or may not have any effect on the performance of the search task. When the (unique) irrelevant stimulus (salient or otherwise) can potentially decrease the performance of the task with its presence (by slowing down search speed, decreasing accuracy, or both), it is known as the distractor (Liesefeld et al., 2024). Thus, a distractor is any irrelevant stimuli which has the *potential* to lure away attention from the target causing a decrease in performance; or in other words interfering with the search for the target.

An irrelevant stimulus can also result in a reduction of search time for a target when presented with a statistical regularity (to form a context) compared to when presented without such regularities. For example, in the *contextual cueing paradigm* (Chun & Jiang, 1998) the irrelevant stimulus in the displays (also sometimes called as the distractors/non-targets in the literature) guides attention towards the target, resulting in response to target faster (i.e., reduced interference) when the displays are repeated, compared to novel displays. Finally, an irrelevant stimulus can have no effect on the search performance – for example, in Fig. 1c the vertical bars (non-targets) have no effect on the search performance, i.e., finding the tilted bar. No matter how many more identical vertical bars we add to the display, it would not have effect on the performance for finding the target, since the target will always 'pop-out' of the display. Sometimes, even the (unique/salient) distractor might not have any effect on the search performance in certain scenarios. For example, when a distractor is presented with 100% probability in each trial of a detection task (where the response is to determine whether the target is absent or present) the distractor does not cause any interference (Chan & Hayward, 2009; Kumada, 1999; Zehetleitner et al., 2009). However, this is not the case for a compound task, where a participant responds on a certain feature of the target (e.g., in Fig. 1a to determine if the bar inside the green diamond is horizontal or vertical; Theeuwes, 1992). Similarly, a distractor presented from another modality, like an irrelevant auditory stimulus, also shows no effect on the performance of a visual search task (see Chapter 2.1 and 2.2; Mandal et al., 2024a).

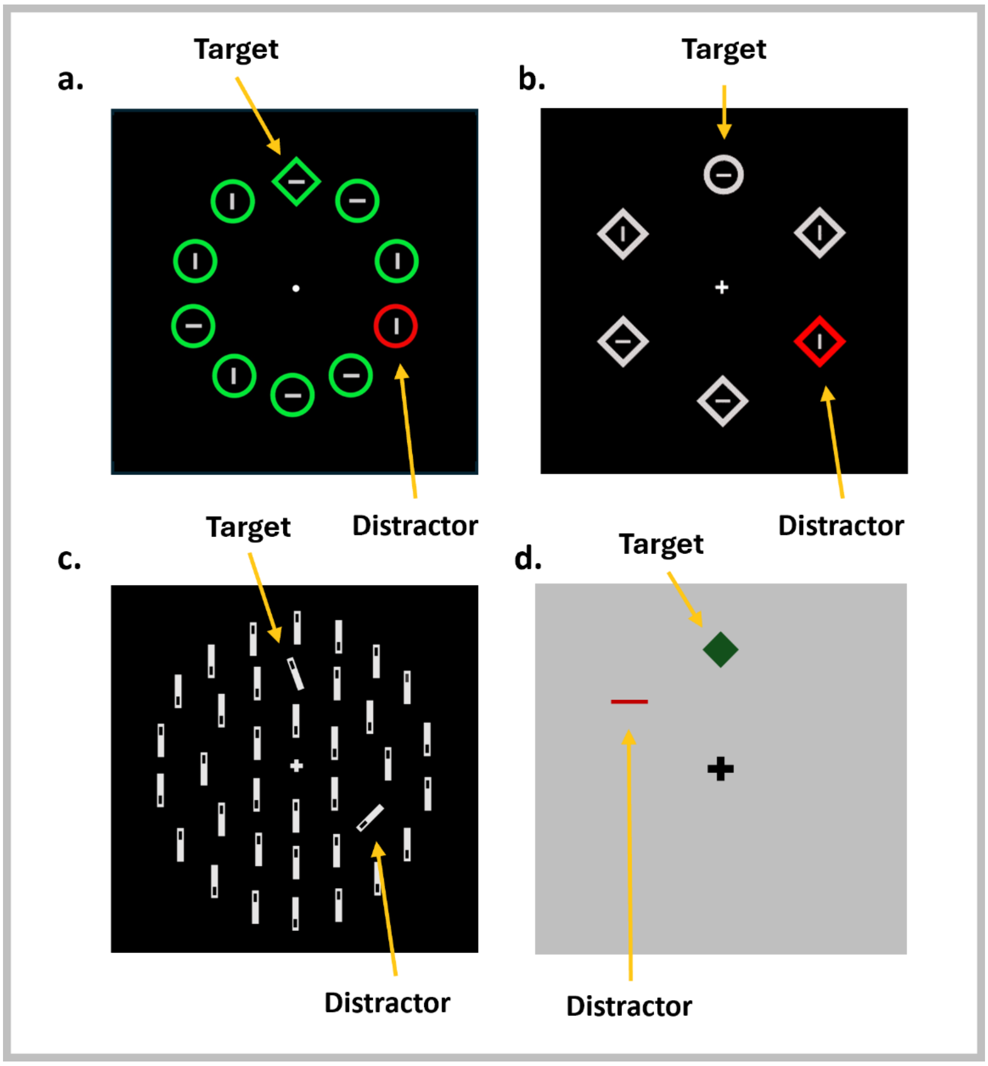


Figure 1. Illustration of four variants of the additional-singleton paradigm. In each variant, participants perform a visual search task to identify a target (which is also a singleton i.e., differing in a target-defining feature from the standard non targets, see Liesefeld et al., 2024 for a definition of the term) presented amongst several non-targets. In a selected percentage of the trials, an additional singleton (differing in a non-target defining feature, may also be salient) – also known as the distractor is presented. The performance of the search task is decreased in the presence of the distractor compared to its absence. Using this same basic principle many different search displays can be constructed as illustrated above. **(a).** Reconstructed from display as used in McDonald et al. (2013), **(b).** Reconstructed from display as used in Kiss et al. (2012), and **(c).** Reconstructed from display as used in Liesefeld et al. (2017). **(d).** Reconstructed from display as used in Hickey et al. (2009).

1.2.2 Distraction: attentional capture vs. filtering costs

Distraction has been extensively investigated in the visual modality using the additional-singleton paradigm (Theeuwes, 1991, 1992). In this paradigm, participants are asked to locate a distinct target among several irrelevant non-targets (for example, identifying a green diamond target among multiple green circular non-targets). In some trials, an irrelevant stimulus is made particularly salient (such as a red circle, serving as the singleton distractor, see examples of some variants of the paradigm in Figure 1), which subsequently delays the response to the target (Chelazzi et al., 2019; Liesefeld & Müller, 2019; Theeuwes, 2010). This impairment of the search performance due to the irrelevant singleton is referred to as *distraction*.

When targets and distractors are presented concurrently in a search display, there is competition for attention allocation between them. Depending on whether the target or the distractor wins the competition for the attention allocation – distraction can be categorized as two specific types: attentional capture, or filtering cost. Attentional capture occurs when the distractor wins the competition for attention, and as a result attention is misallocated towards the distractor instead of the target. This type of scenario often arises when the distractor's relative salience is higher than that of the target's, leading to a higher activation in the priority map. In this case, there are two peaks of activation at the priority map, however, the peak for the distractor is higher than that of the target. Thus, attention is allocated towards the distractor – which happens to be the highest peak in this case. Filtering cost (also known as *non-spatial filtering costs*; Becker, 2007), on the other hand, is the scenario when the distractor fails to win the competition, yet it delays attention towards the target. This scenario generally occurs when the priority of the distractor is lower relative to that of the target – either due to its lower relative salience, or due to other factors induced by the task, e.g., distractor spatial statistical learning (Liesefeld & Müller, 2021). Here again, there are two peaks of activation at the priority map, however, now the peak for the distractor is relatively lower than that of the target, and the attention (although, still delayed) is allocated to the target without going towards the distractor.

Both these types of distraction typically result in delayed response when the distractor is present compared to its absence. Although, attentional capture typically produces a larger behavioral cost (given that in this case, attention is first allocated towards the distractor and only later reallocated towards the target) compared to filtering-costs (where the delay is caused due to resolution of the competition between the two peaks in the priority map); it is not possible to distinguish between these two types of distraction simply by looking at behavioral performances. In the visual modality, the N2pc component (an ERP component indexing the allocation of visual-spatial attention; Constant et al., 2023; Eimer, 1996, 2014) has been instrumental in uncovering the differences in distraction caused by attentional capture and filtering cost (Liesefeld et al., 2017, 2022; Liesefeld & Müller, 2019).

In theory, a specific sequence of events orchestrates the spatio-temporal dynamics involved in the events of attentional capture and filtering-costs. The concept of attentional capture refers to a process where attention is initially drawn to a salient distractor, which is then redirected towards the intended target. This process of shifting attention suggests that the distractor's presence delays the attention allocation towards the target compared to scenarios without distractors. The expected spatio-temporal patterns forecast a specific N2pc pattern (see Figure 2a) for attentional capture: (a) the detection of a distractor N2pc indicating the shift of attention to the distractor, (b) a delay in the emergence of the target N2pc when a distractor is present compared to when it is absent, and (c) a comparative delay in the target N2pc relative to that of the distractor N2pc (Liesefeld & Müller, 2019). The sequence of events governing filtering-costs (see Figure 2b) is quite similar to those of attentional capture, with the notable difference being the absence of a distractor N2pc (as noted in Point (*a*) above). Moreover, the delay observed in the target N2pc when a distractor is present versus its absence (Point (*b*)) is considerably less pronounced than in cases of attentional capture.

Distraction has primarily been explored through the visual modality using visual search paradigms, but it is equally relevant in the auditory modality. By employing an auditory search task—modeled after the additional-singleton task used in visual searches (also refer to the task employed in Hickey et al., 2009; see Fig. 1d) – we have recently shown, for the first time, that both forms of distraction (namely, attentional capture and filtering cost) also occur in the auditory modality (see Chapter 2.3; Mandal et al., 2024b).

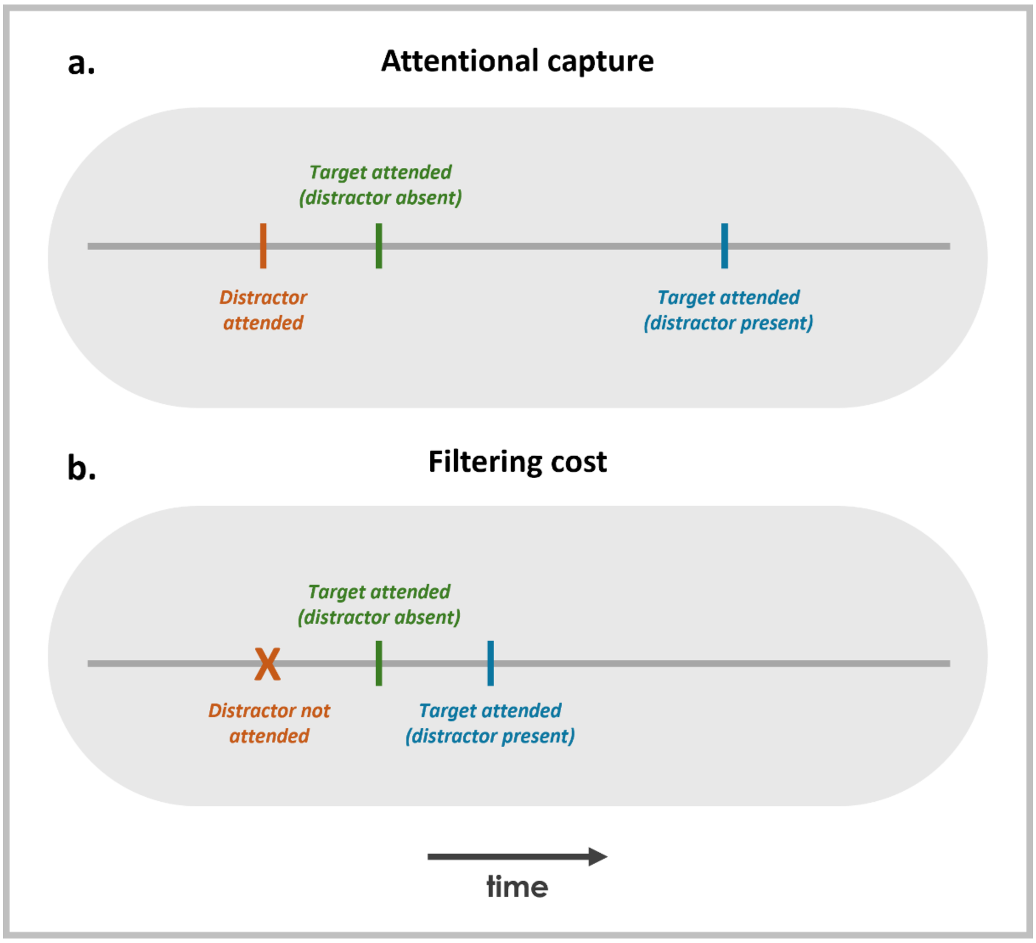


Figure 2. Series of events defining attentional capture and filtering-costs. (a). The temporal dynamics governing attentional capture. Most notably, here we first observe an early attentional allocation towards the distractor. Attention is then reallocated to the target after a delay. There is also a significant delay of attention allocation to the target in the presence of a distractor compared to its absence. **(b).** The temporal dynamics governing filtering-cost. Here, unlike attentional capture here the distractor is not attended. There is still a delay in attentional allocation towards the target in the presence of distractor compared to its absence. However, this delay is of a magnitude lower than that observed for attentional capture.

1.3 Attention in the auditory and visual modality.

Just as in the visual modality, attention in the auditory modality also involves selectively focusing on specific stimuli while filtering out irrelevant information. Although similar in their purpose (i.e., filtering out irrelevant information to focus on the most relevant stimuli), visual and auditory attention differ significantly in how they process and prioritize sensory information. These differences are the reflection of the unique strengths and constraints of each modality. Vision and audition are typically known to be spatial and temporal senses, respectively (Kubovy, 1988). Thus, visual attention is predominantly spatial, enabling us to focus on specific location within the visual field accurately and effortlessly. This spatial aspect of visual attention is thought to be highly developed due to the retinotopically structured organization of visual information on the retina and its subsequent cortical processing (Carrasco, 2011). In contrast, auditory attention is primarily temporal, excelling in processing information over time, such as distinguishing between different sounds in a conversation or tracking changes in auditory stimuli (Shinn-Cunningham, 2008). Auditory attention can also operate effectively in environments where visual attention is limited, such as in darkness or when the visual field is obstructed. Additionally, while visual attention benefits from clear and stable spatial cues, auditory attention often relies on temporal cues and can manage overlapping sounds through mechanisms like stream segregation (Bregman, 1990). These differences underscore the complementary roles of auditory and visual attention in navigating complex environments.

Although the auditory modality is comparatively weaker than the visual modality in terms of spatial acuity and localization accuracy, it is still capable of sound localization. It achieves this through indirect cues (Blauert, 1996; van der Heijden et al., 2019) such as (a) interaural level difference (ILD), which is the difference in sound intensity reaching each ear, making a sound from the left side appear louder in the left ear than in the right; (b) interaural time difference (ITD), which is the difference in the time it takes for sound to reach both ears, with a sound from the left reaching the left ear faster than the right; and (c) spectral cues, which are primarily used to determine the vertical elevation of the sound. Thus, despite its relative weakness, the auditory modality localizes sounds and can effectively allocate spatial attention to the location of auditory stimuli (see Chapter 2.3; Mandal et al., 2024b).

1.3.1 Visual and Auditory search

Search tasks are vital for studying attention (as highlighted in Chapter 1.1.1). While predominantly used in visual studies, these tasks are also crucial for examining auditory attention. Given the constraints of spatial auditory attention (see Chapter 1.3), auditory search tasks often involve identifying a target sound within a rapid sequence of non-target sounds. This paradigm is known as rapid serial auditory presentation (RSAP). For instance, Dalton & Lavie (2004) utilized this paradigm by presenting four sounds in rapid succession, where the goal was to detect the target sound among homogeneous non-targets. Occasionally, a salient deviant distractor was included in the sequence, which caused interference in terms of delaying responses to the target. This auditory task bears some resemblance to the visual additional singleton paradigm, where participants search for a target among non-targets and experience disruption from a deviant singleton. However, a key difference lies in the presentation of stimuli: in auditory search tasks, sounds are presented sequentially, unlike the concurrent presentation in visual tasks.

Handling multiple concurrent sounds presents a significant challenge in the auditory modality as the auditory system finds it difficult to spatially separate more than a few concurrent sounds, particularly when these sounds are abstract and similar (Bregman, 1990). This issue is also pertinent with simultaneous speech, which often has very similar spectral profiles, making it tough to differentiate and isolate specific targets. Thus, developing an auditory paradigm analogous to visual search tasks – where a target is embedded among multiple non-target stimuli – has proven to be quite difficult. Consequently, studying auditory attention in the presence of concurrent auditory stimuli has been limited by this factor. The few studies which have used concurrent auditory stimuli from different locations have limited the number of stimuli presented together, while additionally making sure that the target sounds are somehow detectable

– either by presenting the target at an amplitude higher than that of the non-target (Addleman & Jiang, 2019), or spectrally sparse speech signals (e.g., Ihlefeld & Shinn-Cunningham, 2008).

1.3.2 Can theories of visual attention be extended to auditory attention?

In daily life, people are constantly affected by a combination of sensory inputs from different modalities. Yet, most studies on attentional competition between two stimuli have primarily focused on visual stimuli. Experimental evidence suggests that different sensory modalities can interact to influence attentional allocation. It is then likely that a stimulus from auditory modality can affect the visual priority map, for instance, an irrelevant sound can draw visual attention towards itself (Driver & Spence, 1998; Hillyard et al., 2016; Spence & Driver, 1997). Another example of interaction between auditory and visual modalities is the "pip-and-pop" effect (Van der Burg et al., 2008), where a diotic sound (which is often referred to as non-spatial or spatially unspecific) makes an otherwise inefficient (and hence difficult) visual search task more efficient (by drawing attention to the visual target's location, even though the sound in itself is non-spatial). Such improvements in search efficiency, often measured by the slope of the set-size effect, are a key indicator of the involvement of the priority map.

Similar to how visual attentional guidance is depicted in a priority map, auditory attention can also be represented in a priority map. While the foundational concept of the priority map stays consistent, auditory priority maps also incorporate frequency and temporal dimensions (Kaya & Elhilali, 2012; Kayser et al., 2005) along with spatial characteristics (Golob et al., 2017). This emphasizes the vital role of spectrotemporal properties in auditory attention. Additionally, beyond modality-specific priority maps, such as those for visual or auditory inputs, hierarchical models like the modality-weighting account (Nasemann et al., 2023; Töllner et al., 2009) suggest the existence of a master supramodal priority map. This model, an extension of the dimension-weighting account (see Chapter 1.1.3), proposes that priority activations from all sensory modalities are integrated into this single overarching priority map.

In summary, the presence of auditory and supramodal priority maps indicates that, similar to how an irrelevant visual distraction can activate the visual priority map, an irrelevant auditory distraction can not only impact attention within the auditory priority map but can also influence the visual priority map. This points to the possibility of adapting existing theories of visual attention to include not only cross-modal attentional interactions between auditory and visual stimuli (see Chapter 2.1 and 2.2), but also potentially apply these theories to the auditory modality (see Chapter 2.3).

2. Cumulative Thesis

The following sections contains three original studies which are peer-reviewed, and published articles (Chapter 2.1, Chapter 2.2, and Chapter 2.3).

2.1 The surprising robustness of visual search against concurrent auditory distraction

This manuscript has been published as:

Mandal, A., Liesefeld, A. M., & Liesefeld, H. R. (2024). The surprising robustness of visual search against concurrent auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, 50(1), 99–118. https://doi.org/10.1037/xhp0001168

Author contributions

Ananya Mandal: conceptualization (equal), data curation, formal analysis, investigation, methodology (equal), project administration, software, visualization, writing – original draft (equal), and writing – review and editing (equal).

Anna M. Liesefeld: conceptualization (supporting), methodology (supporting), supervision (supporting), writing–original draft (supporting), and writing–review and editing (supporting).

Heinrich R. Liesefeld: conceptualization (equal), formal analysis (supporting), methodology (equal), project administration (supporting), funding acquisition, resources, supervision, validation, visualization (supporting), writing – original draft (equal), and writing – review and editing (equal).

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: Human Perception and Performance

ISSN: 0096-1523

2024, Vol. 50, No. 1, 99–118
https://doi.org/10.1037/xhp0001168

The Surprising Robustness of Visual Search Against Concurrent Auditory Distraction

Ananya Mandal[1, 2], Anna M. Liesefeld[1], and Heinrich R. Liesefeld[2, 3]
[1] General and Experimental Psychology, Ludwig-Maximilians-Universität Munich
[2] Graduate School for Systemic Neurosciences, LMU Munich
[3] Department of Psychology, Universität Bremen

People often complain about distraction by irrelevant sounds that reportedly hamper performance on concurrent visual tasks demanding the allocation of focused attention toward relevant stimuli, such as processing street signs during driving. To study this everyday issue experimentally, we devised a cross-modal distraction paradigm, inspired by a standard visual-distraction paradigm (additional-singleton paradigm) that is highly sensitive to measure interference on the allocation of attention. In a visual-search pop-out task, participants reported whether a salient target (a tilted bar) was present or absent, while a completely irrelevant, but salient auditory distractor accompanied some trials. To our surprise, the results revealed no notable distraction on visual-search performance (controlled for speed-accuracy tradeoffs). Reliable auditory distraction failed to occur even when the distractor was a (highly salient) auditory oddball or was additionally presented with a temporal advantage of 300 ms. However, when the auditory modality was made relevant globally while maintaining its irrelevance to the visual-search task, we finally observed the expected interference effect.

Public Significance Statement

Distraction by irrelevant sounds while looking for relevant visual information (visual search) is a common experience, but little is known regarding the conditions under which such distraction occurs (or does not occur). The present study demonstrates that visual search is robust to distraction by irrelevant auditory stimuli unless observers have to additionally monitor information from the auditory modality for a secondary task, indicating that our attentional system can block auditory stimuli from interfering with visual search as long as the auditory modality is completely irrelevant. By inducing auditory distraction on visual search, our experimental design opens new avenues for research on the nature and handling of auditory interference that links more closely with the vast scientific literature on visual distraction.

Keywords: visual search, auditory distraction, additional-singleton paradigm, cross-modal distraction, auditory oddball

Supplemental materials: https://doi.org/10.1037/xhp0001168.supp

Distraction by salient events is a ubiquitous phenomenon: in daily life, a loud sound or bright flash that is completely irrelevant to behavioral goals is often perceived as distracting. More specifically, people feel that allocating spatial attention to a task-relevant stimulus can be hampered by a concurrent irrelevant event. For example, a loud sound from the back of the car might cause the driver to unnoticingly pass by a street sign. Such interference with concurrent attention allocations has been extensively studied during visual search, mainly using the additional-singleton paradigm (Theeuwes, 1991, 1992). Typically, participants search for a pop-out target in an array of irrelevant nontargets (e.g., a green target diamond among several green nontarget circles). Visual pop-out search is a laboratory equivalent to common everyday activities involving selection among multiple concurrently presented stimuli such as watching

Ananya Mandal https://orcid.org/0000-0002-2627-0733
Anna M. Liesefeld https://orcid.org/0000-0002-6716-5251
Heinrich R. Liesefeld https://orcid.org/0000-0003-4551-8607

This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) via the RTG 2175 “Perception in Context and its Neural Basis.” We thank Johannes M. Arend for his support in acoustic and audio-related matters and Jan Philipp Röer for the suggestion to use an auditory oddball to increase distractor saliency in Experiments 3 and 4. The authors have no conflicts of interest to disclose.

Ananya Mandal served as lead for data curation, formal analysis, investigation, project administration, software, and visualization. Anna M. Liesefeld served in a supporting role for conceptualization, methodology, supervision, writing–original draft, and writing–review and editing. Heinrich R. Liesefeld served as lead for funding acquisition, resources, supervision, and validation and served in a supporting role for formal analysis, project administration, and visualization. Ananya Mandal and Heinrich R. Liesefeld contributed equally to conceptualization, writing–original draft, writing–review and editing, and methodology.

Correspondence concerning this article should be addressed to Ananya Mandal, General and Experimental Psychology, Ludwig-Maximilians-Universität Munich, 80802 Munich, Germany. Email: Ananya.Mandal@psy.lmu.de

out for street signs, your friend at the main station, or berries to pick (Wolfe, 2021). In some trials of the additional-singleton paradigm, one irrelevant stimulus is rendered salient (e.g., a red circle—the singleton distractor) and therefore delays responding to the target (Chelazzi et al., 2019; Liesefeld & Müller, 2019; Theeuwes, 2010).

The additional-singleton paradigm has proven to be highly sensitive to measure interference by visual distractors. For example, it has been shown that a salient visual distractor causes interference even if participants are almost certain that it will occur on a given trial (Moher et al., 2011) and if it is presented on the majority of trials (e.g., Liesefeld et al., 2022). Interference is even caused by distractors presented outside of the visual-search array (Forster & Lavie, 2008a, 2008b) or at task-irrelevant locations at the same eccentricity as the target (Di Caro et al., 2019). Thus, top-down influences such as the observer's goals and knowledge or expectations and experiences (often referred to as selection history) cannot fully overrule (even if often attenuate; e.g., Liesefeld & Müller, 2019; Theeuwes et al., 2022) pure bottom-up distractor salience.

To provide some context for readers outside the visual-search community, we briefly outline Guided Search (Wolfe, 1994, 2021), the currently most popular and successful framework to understand visual search, because we will introduce the present study and interpret its results from this perspective. The core idea of Guided Search is that search is, well, guided by a spatial mental representation coding for the potential relevance of stimuli. This representation is most commonly referred to as a priority map (Awh et al., 2012; Bisley & Mirpour, 2019; Fecteau & Munoz, 2006; Liesefeld et al., 2023; Luck et al., 2021). Guidance means that spatial attention does not visit each individual stimulus, but only those that are represented by a high activation on the priority map. Thus, based on a preprocessing of (almost) all incoming sensory information in parallel, the priority map is a powerful tool to inform the allocation of spatial attention to only a selected set of stimuli for more in-depth processing. Thus, the priority-map concept explains the stunning ability to efficiently select potentially relevant stimuli for an attention allocation without first attending to each individual stimulus. In typical pop-out search tasks, the target produces a high peak on the map so that it is reliably the first stimulus selected for further processing (henceforth: attended). Phenomenologically, such a target "pops out" of the search display. As the target determines the response, response times are unaffected by the number of concurrently presented stimuli in the search display (set size). This is typically shown by confirming that the linear function relating set size to response times has a zero slope (i.e., the search slope is flat; but see Liesefeld & Müller, 2020; Townsend, 1972).

The core bottom-up influence on the priority map (and thus on the allocation of spatial attention) is salience. Much like "attention" itself (see Chun et al., 2011), the term "salience" has many different meanings in everyday language and in various research communities. The conception of salience we adopt here (in line with the visual-search community; Liesefeld et al., 2023) is that of local feature contrast (Liesefeld et al., 2016; Nothdurft, 1993): a stimulus is salient if one or more of certain of its features differs sufficiently from the respective feature(s) of the stimuli surrounding it (for a comprehensive list of relevant features for the visual domain, see Wolfe & Horowitz, 2017). Therefore, a red tomato among green tomatoes is salient and easily found whereas the same red tomato among other red tomatoes is not. A stimulus presented in isolation is salient by virtue of differing in multiple of these features from the background (e.g., Constant & Liesefeld, 2021, Experiment 3).

From multiple theoretical perspectives, it is crucial that the search target is salient if the goal is to examine distraction and we designed all experiments of the current study as pop-out searches accordingly. Following our proposal of two qualitatively different search modes (Liesefeld & Müller, 2020), we have argued that the priority map is used only if it is useful, namely when the target is salient enough to guide attention. Otherwise, observers will engage in a piecemeal processing of the search display that is not guided by priority (clump scanning; Liesefeld & Müller, 2023; Liesefeld et al., 2021). Similarly, Bacon and Egeth (1994; see also Leber & Egeth, 2006) have proposed that only when the target is salient (they usually refer to it being a singleton, but see Liesefeld et al., 2023, for a differentiation of salience and singleton status) do observers look for salience signals (in the common community lingo: they have an attentional set for salience or they employ singleton-detection mode). Otherwise, observers look for specific target features (feature-search mode) and therefore a distractor with different features—even if highly salient—will not interfere. While rejecting the notion of search modes, Theeuwes (2023) argues that if the target is not salient, observers will shrink their attentional window to increase the signal-to-noise ratio. As a side effect, it is unlikely that a distractor falls within a small attentional window, so that the distractor cannot usually exert its potential to interfere (see also Theeuwes, 1992, 2004). Perceptual load theory makes a similar prediction from another perspective: distractors are processed only under low perceptual load (Lavie, 2005). Notably, searching for a unique target feature among homogenous nontargets (as in the standard additional-singleton paradigm employed here) is usually classified as low in perceptual load, whereas inefficient search tasks (Wolfe, 1998; such as conjunction search, Treisman & Gelade, 1980, or search among heterogenous nontargets, Duncan & Humphreys, 1989) are classified as high in perceptual load (Lavie, 2005). High-perceptual-load search tasks have been shown to be unaffected by distractors (Forster & Lavie, 2008a, 2008b). Thus, from these various theoretical perspectives, employing a salient target is crucial to examine distraction in visual search (but see Stilwell & Gaspelin, 2021, for an alternative view).

According to Guided-Search-based theories of distractor handling (e.g., Liesefeld & Müller, 2019, 2021), a salient distractor can cause interference either because it wins the competition against the salient target and is consequently attended first (attentional capture; note that this term is used differently in the auditory-distraction community as detailed in the General Discussion) or because it renders it harder for the target to win the competition, effectively delaying an attention allocation toward the target ([nonspatial] filtering costs, Folk & Remington, 2008; Kahneman et al., 1983; Wykowska & Schubö, 2011). From the perspective of attentional-engagement theory (Desimone & Duncan, 1995; Duncan & Humphreys, 1989), concurrently presented distractor and target in the additional-singleton paradigm enter a competition for attention that is (ideally) biased in favor of the target (Kerzel & Huynh Cong, 2022; Leonard, 2021; Wolfe, 2021; Zehetleitner et al., 2013). It is this competition for focused attention between concurrently presented stimuli that we are interested in here.

Most research on the competition for focused attention between two salient objects focuses on the visual modality. In daily life, however, people are constantly influenced by inputs from multiple modalities and inputs from different modalities strongly interact, including distraction across sensory modalities as in the introductory street sign example above. In fact, experimental studies have demonstrated cross-modal interactions on attention allocations, where an irrelevant sound draws visual attention (Driver & Spence, 1998; Hillyard et al., 2016; Spence & Driver, 1997) or eases search for an otherwise hard-to-find visual target flickering in sync with the sound (the "pip-and-pop" effect; Van der Burg et al., 2008). In fact, according to a recently proposed distinction between peripheral and central sensation, the clearly peripheral sense of audition (along with peripheral vision and in contrast to foveal vision) should be especially relevant for guiding attention (Zhaoping, 2023). However, to our knowledge, interference by auditory distractors on concurrent visual pop-out search (i.e., comparable to the classical additional-singleton paradigm) has never been examined under controlled laboratory conditions. To fill this gap and establish a link between the two rich scientific literatures on visual and auditory distraction, here we investigate the effect of a salient, task-irrelevant auditory distractor on the performance of visual pop-out search using a cross-modal distraction paradigm inspired by the additional-singleton paradigm (Figure 1).

Previous reports of interference by salient auditory distractors do not require the attentional selection among concurrently presented stimuli. For example, Dalton and Lavie (2004) presented four sounds in rapid succession (rapid serial auditory presentation). The task was to identify the target sound defined by frequency, intensity, or duration among a group of homogeneous nontarget sounds. On some trials, the sequence contained another deviating sound and this salient distractor indeed caused interference in terms of delaying target responses. Dalton and Spence (2007) extended this paradigm to study cross-modal distraction and found similar results for an auditory distractor during rapid serial visual presentation. Another line of evidence for cross-modal orienting of attention comes from a body of literature that uses the cross-modal cueing paradigm: based on the purely visual spatial-cueing paradigm by Folk et al. (1992), participants were presented with a task-irrelevant audio-visual spatial cue prior to performing a visual search task (Matusz & Eimer, 2011). Attentional capture by this cue is inferred from a boost in response speed when the audio-visual cue is presented at the target location (valid cue) compared to when only the visual cue is presented. Unlike the additional-singleton paradigm, in both rapid serial presentation and spatial cueing, target and distractor are not in direct competition, because they are presented at different time points.

From a guided-search perspective is notable that in most auditory-distraction studies, including the present one, the potentially distracting sound was presented via headphones without any spatialization (diotic presentation, Best et al., 2020). It is unclear where on the priority map such a sound is located. If listeners are asked to locate a diotic sound, they most often respond that it is inside their head. As the concept of a priority map has emerged from research on visual attention, this map is typically pictured as a single visual plane within observers' field of view. This is misleading because visuospatial attention can be allocated to different depth planes (Nakayama & Silverman, 1986) and is not limited only to the oculomotor range (Hanning et al., 2019). It therefore does appear possible that a location inside the head is represented on the priority map. Alternatively, a diotic sound is often described as "nonspatial" or "spatially nonspecific," potentially implying that it is everywhere and nowhere at the same time. In any case, it is well established that diotic sounds can affect priority computations in visual search. Van der Burg et al. (2008) employed an inefficient search task with independently flickering stimuli, the target was found efficiently when a diotic sound synchronized with the target flicker was played (pip-and-pop effect). That an inefficient search task, that is, a task that becomes harder as the number of search stimuli increases (positive search slope), becomes efficient, that is the number of search stimuli plays no or only a minor role for performance (flat search slope) is usually interpreted as the target obtaining a boost in priority (by the diotic sound). Thus, a diotic sound can affect the priority map that is relevant for finding a visual-search target.

Even more extensive evidence that diotic auditory distractors can potentially interfere with the processing of visually presented stimuli comes from serial-recall working memory tasks (for a review, see Hughes, 2014). In these tasks, participants are to remember a series of visually presented stimuli (typically verbal material such as words, letters, or digits) for a later ordered recall and the presentation of task-irrelevant sounds during the encoding phase results in reduced recall performance (irrelevant-sound effect). Hence, our expectation that the auditory distractor would cause interference in terms of performance costs in the visual-search task was reinforced in light of previous evidence that auditory distractors can cause interference on other visual tasks and that, more generally, auditory and visual attention interact (see Spence, 2010, for a review).

Despite this theoretical possibility and the related empirical evidence, it remains unknown whether a diotic auditory distractor does indeed enter the competition for focused visual attention against a concurrently presented visual-search target and therefore causes interference. Can people completely ignore a salient diotic sound during search for a salient visual target stimulus, or does an irrelevant-but-salient diotic sound hamper the allocation of spatial attention to a concurrently presented relevant visual stimulus?

Experiment 1

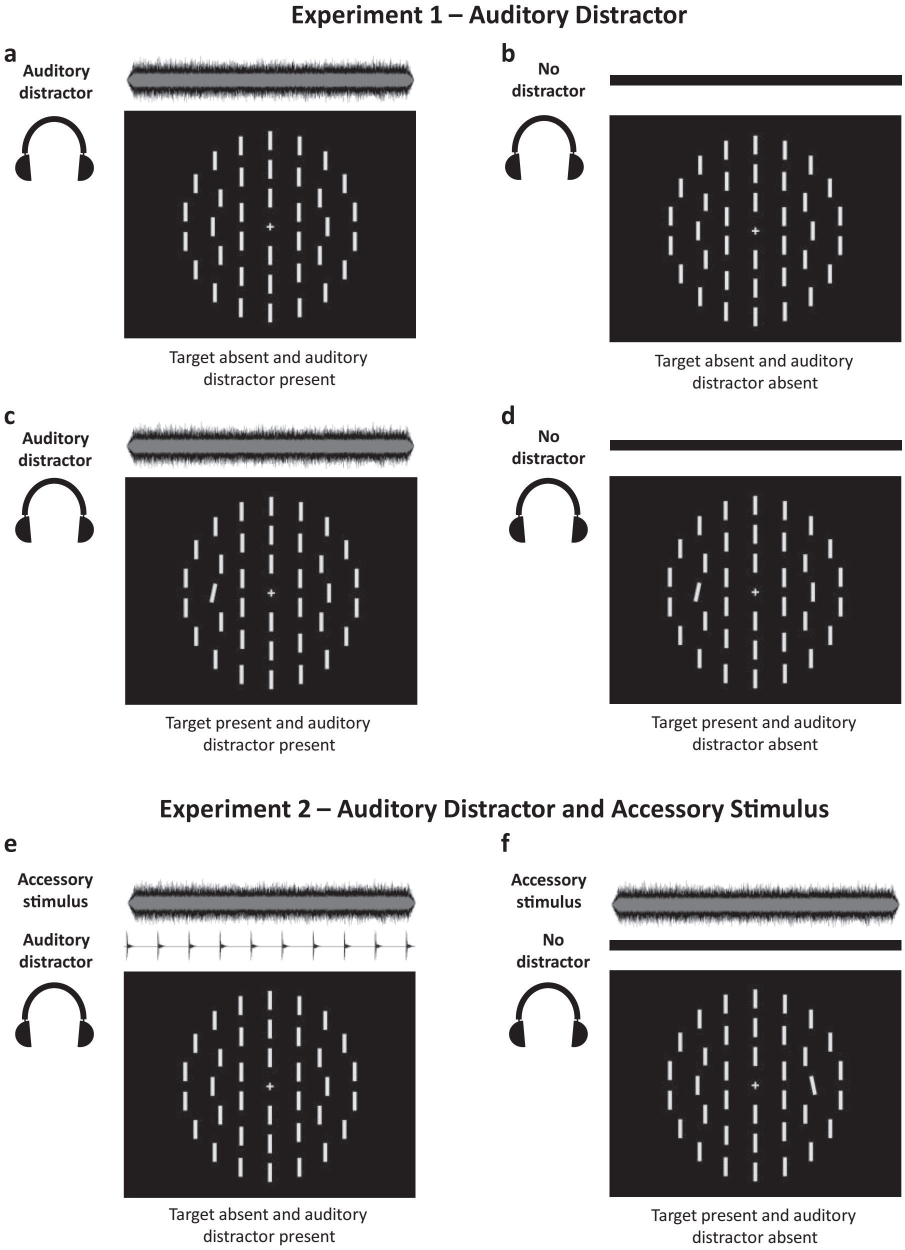
Method

Participants

To determine the appropriate sample size, we used Liesefeld et al. (2019), after which the present study was modeled, as our starting point. We base our sample-size planning on Liesefeld et al. (2019) because this study uses a pop-out search task to investigate distractor interference between same-dimension and different-dimension distractors. The current study provides an extension to this by studying different-modality distractors in the same visual search scenario. Liesefeld et al.'s study consisted of two experiments, with 12 participants each. In particular, their participants also searched for a 12° tilted target bar presented among vertical nontarget bars arranged in concentric circles. Effect sizes (on reaction times, RTs) were lower for the (also more comparable) different-dimension distractors and we thus base our power analyses on these. Based on the effect sizes observed by Liesefeld et al. (2019) in the different-dimension distractor condition for the luminance-target group ($d_z = 1.70$) and the orientation-target group ($d_z = 1.44$), we determined that to achieve a statistical power of $1 - \beta = .95$ at $\alpha = .05$, two-tailed, the required sample size is $n = 7$ and $n = 9$, respectively (calculated using G*Power Version 3.1.9.7; Faul et al., 2009). We further consulted another study (Liesefeld & Müller, 2021) that also employed a comparable design—but with a relatively large sample size (47 participants in Experiment 2)—providing for more stable effect-size estimates. Based on the effect size observed for the different-

Figure 1

Illustration of Search Displays in Experiments 1 (a, b, c, d) and 2 (e, f)



Note. Participants were asked to report whether the target singleton (a 12°-tilted bar) was present or absent among a dense array of nontargets (vertical bars). An additional salient auditory singleton was presented in some trials. (a–d) All four possible conditions for Experiment 1. (e, f) Two example conditions for Experiment 2. Please visit www.doi.org/10.6084/m9.figshare.24174018, for example, trials from each of the experiments. Note that the pink-noise auditory distractor used in Experiment 1 served as the accessory stimulus in Experiment 2, where a metronome tick took on the role of the auditory distractor.

dimensional distractor condition in Experiment 2 ($d_z = 1.01$) and the same parameters as above, we calculated the required sample size to be $n = 15$. To be on the safe side, we base the sample size for all the experiments in the current study on this latter estimate. For counterbalancing responses among participants, we increased that number to $n = 16$ in all experiments of the present study.

Thus, 16 participants took part in Experiment 1 (median age: 28 years; range: 19–44 years; eight female). In all experiments, participants reported normal hearing and normal or corrected-to-normal vision. Written informed consent was obtained from all the participants, and they received course credit or were paid for their participation. No participant had to be excluded from the analysis (all average error rates below 15%). Procedures were approved by the ethics committee of the Department of Psychology and Pedagogics at Ludwig-Maximilians-Universität Munich.

Constraints on Generality Statement

Our samples were mostly composed of European, South Asian, and East Asian young adults residing in Munich, Germany. Given that our task is not tied to any origin-specific behavior (such as reading direction), we do not think that the geographic origin of our participants would influence the present results. However, our results might not be generalizable to all people, especially not to those with severe attentional deficits. Moreover, the results might apply only to tasks where it is a reasonable strategy to look for salience signals (see Liesefeld et al., 2021; Liesefeld & Müller, 2023). Additionally, the age of the participants can influence the impact of auditory oddballs (as employed in Experiments 4 and 5) on the performance of visual tasks (e.g., Leiva et al., 2015).

Stimuli and Design

Stimulus presentation and response collection were controlled using PsychoPy (Peirce et al., 2019). Search displays were presented on a monitor (screen resolution: 1,680 × 1,050 pixels, refresh rate: 60 Hz) at a viewing distance of 70 cm. Search arrays consisted of 36 grey bars (1.35° × 0.25° in size) presented on a black background (see Figure 1). The bars were arranged in three concentric rings (radii of 2.1°, 4.2°, and 6.3°) around a central fixation cross (0.5° × 0.5°). Half of the trials consisted of vertically oriented bars only (target-absent condition). On the other half of the trials, one target bar was tilted 12° to the right[1] (target-present condition).

Along with the presentation of the bars, an irrelevant auditory stimulus (distractor) was randomly presented in half of the trials (distractor-present condition), with no distractor presented in the remaining half (distractor-absent condition). The distractor was a pink noise (high pass filtered at 200 Hz) that was presented binaurally (diotic) using headphones (Sennheiser HD 600, Sennheiser, Wedemark, Germany) at 70–72 dB(A) sound pressure level (SPL), measured at the eardrum, using the miniDSP Earphone Audio Response System device (miniDSP, Hong Kong, China). Pink noise is considered to be relatively potent at causing interference (Marsh et al., 2020, p. 354). The study thus consisted of a 2 (target present vs. target absent) × 2 (distractor present vs. distractor absent) design. The experiment was divided into 15 blocks, with 96 trials in each block. Data from the first block were excluded from all analyses to eliminate any learning effect. Thus, results are based on 1,344 trials in total, with 336 trials in each of the four conditions.

Procedure

The search displays remained on the screen until participants made their response, or until the response deadline of 4,000 ms. For trials with auditory distractors, the sound was played from display onset to display offset, so that it was present for the same duration as a visual distractor in the standard additional-singleton paradigm would. The task was to respond whether the 12°-tilted target bar was present or absent in each trial. In line with the standard additional-singleton paradigm, participants were instructed to try and ignore the distractors that were presented. The occurrence of the distractors was temporally unpredictable in all the tasks, because of the (a) randomization of trials and (b) variable intertrial interval. Responses were given on a keyboard using the left and right index fingers (counterbalanced across participants) and triggered display offset. If the trial response was incorrect or delayed, the central fixation cross turned red or blue, respectively, for 1,000 ms. For correct trials, no feedback was provided, and the program directly proceeded with the intertrial-interval preceding the next trial (i.e., without an extra 1,000-ms delay). Trials were separated by an interstimulus interval of variable duration between 800 and 1,200 ms, during which only the fixation cross was presented on the screen. Participants were instructed to respond accurately and as quickly as possible.

Analyses

Below we report analyses on median correct RTs and error rates. As we observed telltale signs of condition-specific speed-accuracy tradeoffs (SAT) in these performance measures (Heitz, 2014), we additionally report, and eventually base most major conclusions on, a measure that combines RTs and the percentage of correct responses (often also referred to as "percent correct" [PC] or "accuracies"; i.e., 1—error rates) into a combined performance measure termed balanced integration score (BIS) that has been shown to be largely unaffected by SATs (Liesefeld et al., 2015; Liesefeld & Janczyk, 2019, 2023). This measure first standardizes (*z*-transforms) the raw performance measures by subtracting their respective overall mean and dividing by the respective standard deviation across the relevant design cells (here: 16 participants × 2 distractor presence = 32 median RTs and PCs for most of the analyses presented here). These standardized performance measures are combined into one score per design cell ($\text{BIS} = z_{\text{Acc}} - z_{\text{RT}}$; see, e.g., Equation (4) in Liesefeld & Janczyk, 2019) so that negative BIS values indicate lesser than average performance and positive BIS values indicate greater than average performance in the respective design cell. By design, this measure gives equal weight to both performance measures (Appendix B in Liesefeld & Janczyk, 2019) and extensive simulations have shown that it controls for variations in SAT while retaining "true" effects such as differences in actual performance.

As we were interested in whether an auditory distractor interferes with visual search, the reported analyses focus on the effect of distractor presence. Pairwise comparisons between distractor-absent and distractor-present conditions were performed using Wilcoxon signed-rank test (two-tailed comparisons). The nonparametric Wilcoxon signed-rank test was used instead of the Student t test since some of our data violated the assumption of normality that is a prerequisite for this parametric test (see Supplement II in the online supplemental materials for the results of the test for normality, and the results of Student t tests for the comparisons which do not violate the assumption of normality). Effect size is reported using rank-biserial correlation (r), where $r = 1$ refers to perfect rank ordering, that is, an $r = 1$ means that the effect was present and in the same direction for each individual participant. Consequently, we also report the mean (M) and median values (Mdn) per condition with

[1] Note that the target in all subsequent experiments was tilted 12° to the left. Tilt direction does not affect target salience (or task difficulty) in our displays, since the local feature contrast of the target remains the same (Liesefeld et al., 2016).

the respective median absolute deviation (MAD: a measure of dispersion). Results graphs show the means (of individual participants' medians) and within-participant confidence intervals (Cousineau & O'Brien, 2014). Additionally, we report Bayes factors (BFs), for the Bayesian Wilcoxon signed-rank test (result based on data augmentation algorithm with five chains of 10,000 iterations), quantifying evidence for the alternative over the null hypothesis (BF_{10}). BFs were calculated using the standard Jeffreys-Zellner-Siow Cauchy prior with a scale factor of $r = \sqrt{2}/2$. To classify the strengths of evidence through the BFs, we used Jeffreys's criterion (van Doorn et al., 2021)—which states that for the alternative hypothesis, BFs between 1 and 3 are weak evidence, BFs between 3 and 10 are moderate evidence, and BFs greater than 10 are strong evidence. Following these criteria, $1/3\ (=0.33) < BF_{10} < 1$ is considered weak evidence for the null hypothesis, $1/10\ (=0.10) < BF_{10} < 1/3\ (=0.33)$ is considered moderate evidence for the null hypothesis, and $BF_{10} < 1/10\ (=0.10)$ is considered strong evidence for the null hypothesis. All analyses were performed using JASP Version 0.16.4 (JASP Team, 2021) and using custom scripts in Matlab.

Transparency and Openness

The anonymized raw data, the analysis script in Matlab and the analysis files from JASP are available on Open Science Framework (http://www.doi.org/10.17605/OSF.IO/C7K3J). The data for all the experiments were collected between August 2021 and June 2022 in Munich, Germany.

Results

Correct responses were significantly faster (see Figure 2a) on distractor-present (M =564 ms, $Mdn \pm \text{MAD} = 530 \pm 37$ ms) compared to distractor-absent trials ($M = 587$ ms, $Mdn \pm \text{MAD} = 545 \pm 39$ ms), $W = 5.00$, $p < .001$, $r = -.926$, $BF_{10} = 430.65$. Error rates showed no significant difference between distractor-present ($M = 3.59\%$, $Mdn \pm \text{MAD} = 2.31 \pm 0.82\%$) and distractor-absent trials ($M = 2.84\%$, $Mdn \pm \text{MAD} = 2.08 \pm 1.41\%$), $W = 93.00$, $p = .065$, $r = .550$, $BF_{10} = 1.56$. Although nonsignificant, the direction of the distractor effect on error rates is intriguing and might potentially explain the counterintuitive finding that distractor presence speeds rather than slows responses. Distractor presence seems to concurrently decrease RTs and increase error rates, indicating an effect on the SAT rather than an effect on performance (Heitz, 2014). Discounting this interpretation based on the nonsignificant test on error rates would be an instance of accepting the null hypothesis, which is not supported by the inconclusive BF.

Rather than glossing over the error pattern, we combined speed and accuracy to control for SATs by calculating the BIS introduced in the analysis section above. Comparing BIS between distractor-present ($M = -0.04$, $Mdn \pm \text{MAD} = 0.42 \pm 0.44$) and distractor-absent trials ($M = 0.04$, $Mdn \pm \text{MAD} = 0.30 \pm 0.41$), revealed moderate evidence for the absence of a distractor effect (Figure 2b), $W = 74.00$, $p = .782$, $r = .088$, $BF_{10} = 0.26$.

Discussion

Based on the extensive literature on visual distraction and on cross-modal interactions, we had expected visual-search performance to be impeded when an auditory distractor was presented with a visual-search display (Chelazzi et al., 2019; Liesefeld & Müller, 2019; Theeuwes, 2010). To our surprise, we observed a significant decrease rather than the expected increase in RTs. Error rates increased slightly on distractor-present trials, but this effect did not reach statistical significance. This would typically be interpreted as evidence that the auditory "distractor" facilitated search performance. However, ignoring the effect on error rates because it is nonsignificant would be an instance of accepting the null hypothesis which is not supported by frequentist statistics (p values) and the respective BF was highly inconclusive. In fact, controlling for condition-specific SAT using a combined performance measure (BIS) indicated that the auditory distractor had no effect on overall visual-search performance. Similar results are sometimes discussed as "criterion shifts" (Spence & Driver, 1997).

In the broader cognitive-science community, a salient task-irrelevant stimulus presented in a task-irrelevant sensory modality (such as our auditory distractor) with close temporal proximity to an imperative stimulus (our search display or the target within it) is known as accessory stimulus. These accessory stimuli are typically completely uninformative and task-irrelevant and are known to induce speeded responses to stimuli in the target sensory modality—often without a significant change in error rates (e.g., Bernstein et al., 1969; Hackley & Valle-Inclán, 1999; Jepma et al., 2009)[2] compared to when the stimulus in the target sensory modality is presented alone. As SAT effects induced by accessory stimuli—as in the current experiment—have been reported as well (Low et al., 1996; Posner et al., 1973), we feel that the effect of distractor-presence observed in Experiment 1 is best interpreted as an accessory-stimulus effect (see the General Discussion section for more detail).

Experiment 2

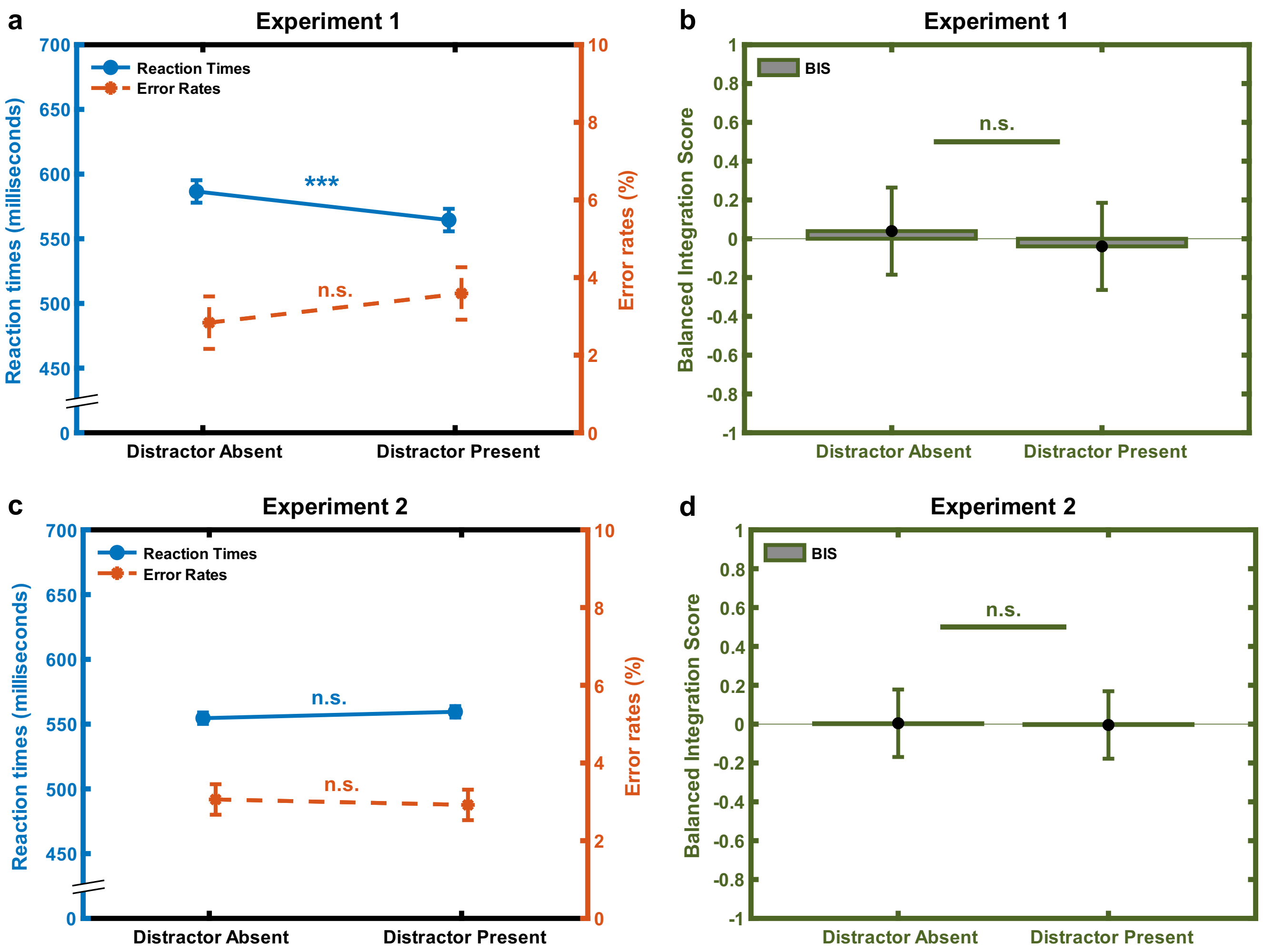
In Experiment 1, the auditory distractor sped up responses, which might be interpreted as the auditory "distractor" acting as an accessory stimulus (Hackley & Valle-Inclán, 1999; Jepma et al., 2009) as detailed above. It appears possible that this speed-up had overshadowed any interference effect, which is typically observed in RTs. To eliminate this potential difference between distractor-present and distractor-absent trials, we played the pink-noise sound (the potential accessory stimulus) on every trial. Another salient and clearly distinguishable sound (metronome tick) served as a distractor in half of the trials. If the auditory distractor in Experiment 1 had acted as an accessory stimulus, then the observed effect on the SAT (the accessory-stimulus effect) should occur on all trials in Experiment 2, so that the comparison between distractor-present trials (two sounds) and distractor-absent trials (one sound) should now reveal an interference effect (if existent).

Method

Sixteen participants participated in Experiment 2 (median age: 24 years; range: 19–43 years; 13 female). Experiment 2 was identical to Experiment 1, except that an auditory stimulus (pink noise at 75 dB(A)—same stimulus as in Experiment 1, just a little louder) was present in all the trials and served as the accessory stimulus.

[2] We thank Daniel Bratzke and Markus Janczyk for making us aware of this literature.

Figure 2
Results for Experiment 1 (a, b) and Experiment 2 (c, d)



Note. (a, c) Mean (average across participants) of median RTs and error rates. (b, d) BIS, a combined performance measure. Error bars represent 95% within-participant confidence intervals. RTs = reaction times; BIS = balanced integration score. See the online article for the color version of this figure.

An additional auditory stimulus (metronome tick—generated using Audacity, with tempo of 300 beats/min at 75 dB(A)) was presented along with the visual-search display in half of the trials (distractor-present condition; see Figure 1e) and, thus, served as the auditory distractor in Experiment 2.

Results

This design produced no significant change in RTs (Figure 2c) for distractor-present trials ($M = 559$ ms, $Mdn \pm$ MAD = 553 ± 57 ms) compared to distractor-absent trials ($M = 555$ ms, $Mdn \pm$ MAD = 546 ± 52 ms), $W = 99.00$, $p = .117$, $r = .46$, $BF_{10} = 0.99$. Even though the evidence for the absence of an effect in terms of the BF is indecisive, the direction of the effect (slower RTs on distractor-present trials) indicates the successful elimination of the accessory-stimulus effect by the auditory distractor, as observed in Experiment 1. Error rates also showed no significant difference when the distractor was present ($M = 2.92\%$, $Mdn \pm$ MAD = 2.23 ± 1.41%) compared to when it was absent ($M = 3.06\%$, $Mdn \pm$ MAD = 2.38 ± 1.79%), $W = 61.00$, $p = .744$, $r = -.103$, $BF_{10} = 0.28$. Additionally, the comparison of BIS for distractor-present ($M = -0.004$, $Mdn \pm$ MAD = 0.09 ± 0.842) and distractor-absent trials ($M = 0.004$, $Mdn \pm$ MAD = 0.13 ± 0.78) showed no significant change in performance (Figure 2d) because of the distractor, $W = 60.00$, $p = .706$, $r = -.12$, $BF_{10} = 0.27$. Rather, the BF again provides moderate evidence for the absence of a distractor effect.

Discussion

The result pattern of Experiment 2 indicates that presenting a sound on each trial successfully eliminated the confound with the accessory-stimulus effect. That is, presenting an auditory distractor no longer results in an effect on participants' SAT. Even without this confound, we do not observe any sign of interference by the auditory distractor.

Experiment 3

One reason for not observing interference by an auditory distractor in Experiments 1 and 2 might be that neither of the employed distractors was salient enough. We chose the distractor stimuli so that they appeared salient to us, and we played them relatively loud (70–75 dB(A)—similar to the sound level of a classical vacuum cleaner measured in the free field at 1 m distance), but maybe this was not yet enough to render them more salient than the visual target stimulus. While even visual stimuli less salient than the target are known to produce interference effects (Gaspar & McDonald, 2014; Zehetleitner et al., 2013), the potentially low salience remains a valid critique of our findings.

A common approach to render an auditory stimulus salient is to present it as an oddball within a stream of homogenous nontargets. It has been shown that such auditory oddball stimuli are processed against the observers' will, even if they are engaged in taxing visual tasks (Escera et al., 1998; Parmentier, 2016). Thus, if visual search can in principle be hampered by auditory distraction, an auditory oddball should produce this effect. Furthermore, visual distractors interfere more strongly when presented only rarely, likely because participants have less incentive to suppress them and less opportunity to adapt to them (Geyer et al., 2008; Müller et al., 2009; Won et al., 2019). Thus, by presenting an auditory oddball stimulus on a low percentage of trials, we should dramatically increase the likelihood of observing auditory distractor interference with visual-search performance, if it exists at all.

Method

Sixteen participants took part in Experiment 3 (median age: 26 years; range: 21–39 years; 10 female). Three participants (two female) who participated in Experiment 1 also participated in Experiment 3.[3]

The visual search task was almost identical to the previous experiments but was accompanied by a sequence of auditory stimuli consisting of 50 ms of pink noise (high pass filtered at 200 Hz, with 5 ms rise and fall, presented at 75 dB(A) SPL) repeated every 1,000 ms (see Figure 3) and the on- and offsets of the visual-search displays were synchronized to this 1 Hz auditory rhythm, so that, deviating from the preceding experiments, display offset was not triggered by the response. Hence, the visual search display was presented until the 1-s cycle in which the response was given was completed (e.g., if the response in a trial was at 1.5-s, the visual display was presented until 2-s), or until the response deadline of 4-s as in the previous experiments. The intertrial interval lasted either 1, 2, or 3 1-s cycles. Twenty-five percent of the trials started with a (subjectively quite annoying) sawtooth tone (deviant)—generated at 500 Hz, 50 ms duration with 5 ms rise and fall, presented at 75 dB(A) SPL replacing 8.23 ± 1.57% (*Mdn* ± range; min. = 6.73%) of the pink noise standard stimuli. This deviant (oddball) tone served as the auditory distractor in Experiments 3 and 4, whereas pink noise (similar to the distractor in Experiment 1 and the accessory stimulus in Experiment 2) now acted as the standard tone in an oddball sequence. Two trials with deviant tones (distractors) were never presented one after the other to further increase the oddball effect. The experiment consisted of 15 blocks of 96 trials each. At the beginning of each block, at least four trials did not contain a distractor. Note that because of the randomized trial sequence and the variable intertrial interval (1–3 1-s cycles), distractor presentation was temporally unpredictable as in all other experiments reported here.

Results

Still, there was no significant change in RTs (Figure 4a) when the distractor was present ($M = 671$ ms, $Mdn \pm \text{MAD} = 621 \pm 63$ ms) compared to when it was absent ($M = 680$ ms, $Mdn \pm \text{MAD} = 633 \pm 54$ ms), $W = 38.00$, $p = .130$, $r = -.44$, $BF_{10} = 0.67$. No significant change in error rates was observed between distractor-present ($M = 2.89\%$, $Mdn \pm \text{MAD} = 2.68 \pm 1.64\%$) and distractor-absent trials ($M = 2.57\%$, $Mdn \pm \text{MAD} = 2.53 \pm .85\%$), $W = 78.50$, $p = .605$, $r = .15$, $BF_{10} = 0.30$. Additionally, BIS showed no change in performance (Figure 4b) between distractor-present ($M = -0.07$, $Mdn \pm \text{MAD} = 0.05 \pm 1.04$) and distractor-absent trials ($M = 0.07$, $Mdn \pm \text{MAD} = -0.14 \pm 0.92$), $W = 64.00$, $p = .860$, $r = -.06$, $BF_{10} = 0.27$. Once again, the BF for BIS indicates moderate evidence for the absence of a distractor effect.

Discussion

In Experiment 3, we used a rare auditory oddball distractor (~8% of deviants in a steady auditory sequence) and presented this distractor on only 25% of all visual-search trials. Despite the use of these two established experimental techniques to render the auditory distractor more distracting, it still did not notably interfere with visual search.

Experiment 4

As a last attempt to induce auditory distraction in the additional-singleton paradigm, we allotted a temporal processing advantage to the rare oddball distractor by presenting it before the search-display onset. Most of the earlier works on the audio-visual oddball paradigm (Escera et al., 1998, 2001, 2002, 2003) presented the auditory stimuli 300 ms prior to the visual stimuli and indeed observed interference on the discrimination of a single visual stimulus. Hence, Experiment 4 was identical to Experiment 3, except that we introduced a temporal difference of 300 ms between the auditory and visual stimulus (onset-to-onset), providing the distractor with a temporal advantage to see whether this renders it potent enough to interfere with visual search.

Method

Sixteen participants took part in Experiment 4 (median age: 25 years; range: 21–41 years; 11 female). Experiment 4 was identical to Experiment 3, except that synchronization of search-display onset to the ongoing auditory stream was changed so that at the beginning of each trial, the 50-ms auditory stimulus preceded the visual search displays by 300 ms instead of occurring simultaneously.

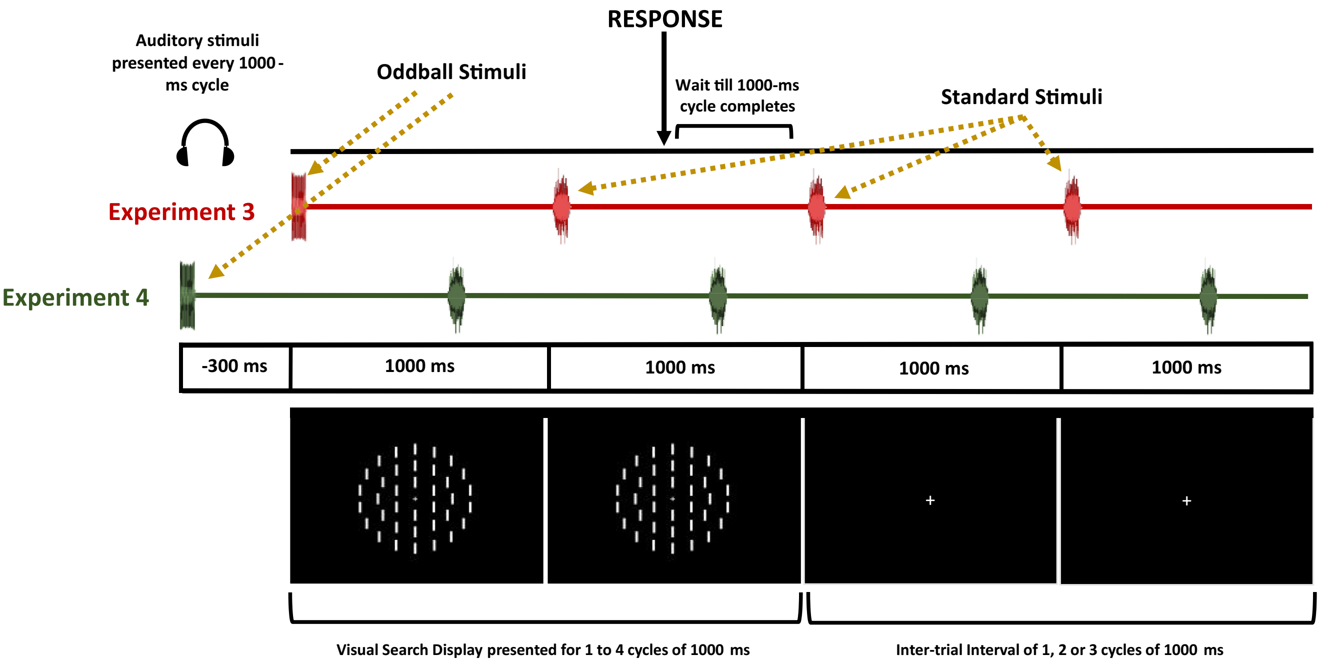
Results

As in Experiment 1, we observed a significant decrease in RTs (by 27 ms; Figure 4c) when the distractor was present ($M = 626$ ms, $Mdn \pm \text{MAD} = 564 \pm 44$ ms) compared to distractor-absent trials ($M = 654$ ms, $Mdn \pm \text{MAD} = 587 \pm 49$ ms), $W = 0.00$, $p < .001$,

[3] Given that participants were fully informed about the presence of distracting sounds and that they are supposed to ignore these as well as due to the extensive training they received before starting with the main task, we do not believe that participating in multiple studies had any effect on distractor handling. Nevertheless, we also re-ran the analyses in this and all following experiments with those participants excluded. None of these re-analyses indicated a different pattern of effects than those reported here. See Supplement I in the online supplemental materials for details.

Figure 3

Exemplary Distractor-Present Trial in Experiments 3 and 4 Using the Auditory Oddball Paradigm



Note. The oddball was always synchronized with search-display onset (occurring at the same time, Experiment 3, or 300 ms earlier, Experiment 4) and served as the salient distractor occurring on 25% of the search trials, whereas the standard stimulus was also synchronized with display onset but was also played at other (regularly spaced) times during a trial. See the online article for the color version of this figure.

$r = -1.00$, $BF_{10} = 2{,}053.55$. Error rates showed a nonsignificant increase on distractor-present ($M = 2.51\%$, $Mdn \pm MAD = 1.79 \pm 0.89\%$) compared to distractor-absent trials ($M = 2.06\%$, $Mdn \pm MAD = 1.49 \pm 0.50\%$), $W = 89.00$, $p = .106$, $r = .48$, $BF_{10} = 1.35$. This pattern of results again indicates a distractor-presence effect on the SAT. Indeed, the BIS did not show any significant difference in performance (Figure 4d) between the distractor-present ($M = -0.09$, $Mdn \pm MAD = 0.21 \pm 0.93$) and the distractor-absent trials ($M = 0.09$, $Mdn \pm MAD = 0.48 \pm 0.83$), $W = 61.00$, $p = .744$, $r = -.10$, $BF_{10} = 0.31$.

Discussion

As a final attempt to demonstrate interference by an auditory distractor on visual search, Experiment 4 endowed a 300-ms head start to the rarely presented (25% of trials) oddball distractor (∼8% of sounds) employed in Experiment 3. Despite all these measures, the distractor still did not cause any notable interference.

Intriguingly, the effect on the SAT observed in Experiment 1 that we eliminated in Experiment 2 by presenting a sound on all trials, and which was also not observed in the highly similar Experiment 3, reemerged here. In Experiment 1, we interpreted this pattern as an accessory-stimulus effect, which is an automatic reaction to a salient stimulus in a task-irrelevant modality. We were not surprised that this effect did not occur in Experiment 3, because there as well, we presented a sound (standard or deviant) on every trial. However, the same is true for Experiment 4. Note though that here, the presentation of the deviant (the auditory distractor) was predictive of the upcoming search display, because it would only occur before search-display onset and not during the intertrial interval. This renders an interpretation in terms of an alerting signal more likely (Matthias et al., 2010; Posner et al., 1973; Posner & Petersen, 1990; see General Discussion for details). Interesting as this speculation might be, our main goal was to examine whether an auditory distractor can interfere with visual-search performance. Results from Experiment 4 demonstrate that even a rare oddball stimulus with a temporal processing advantage does not reliably interfere with visual-search performance.

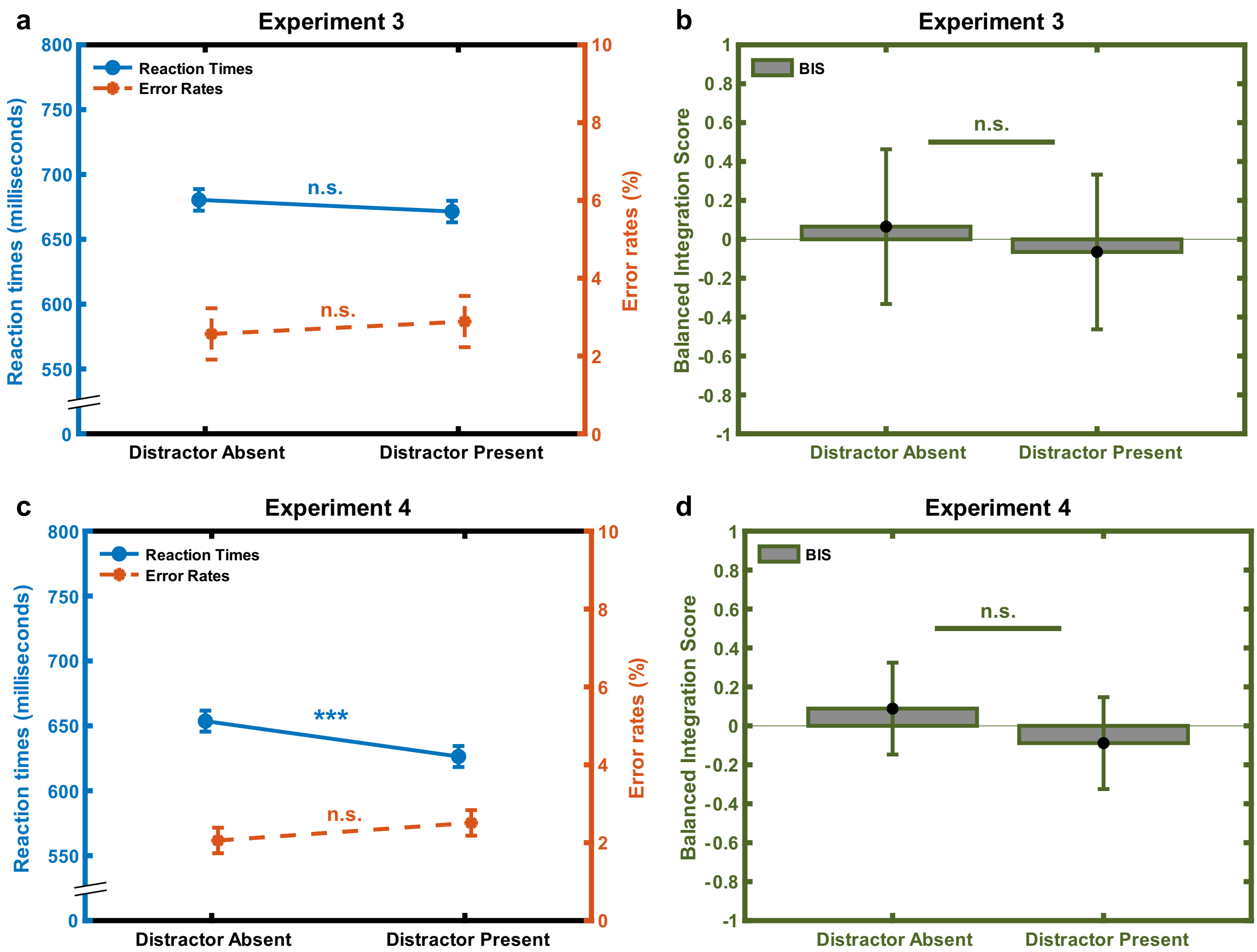
Accumulated Evidence for the Null Hypothesis

Across multiple experiments, we failed to observe interference by auditory distractors on overall visual-search performance, with each additional experiment strengthening the case for the absence of an interference effect. Here, we analyze the accumulated evidence for the null hypothesis (BF_{01}) over the four experiments with auditory distractors (Experiments 1–4) by performing pairwise comparison of the BIS for distractor-present and distractor-absent conditions with the standard deviations used in the calculation of BIS calculated across all experiments. We obtain a $BF_{10} = 6.90$ for the Bayesian Wilcoxon signed-rank test, which once again indicates moderate evidence for the null hypothesis.

Experiment 5

The preceding experiments show that auditory distractors do not interfere with the allocation of focused attention. This stands in sharp contrast to decades of observing interference by visual

Figure 4
Results for Experiments 3 (a, b) and 4 (c, d)



Note. (a, c) Mean (averaged across participants) of median RTs and error rates. (b, d) BIS, a combined performance measure. Error bars represent 95% within-participant confidence intervals. RTs = reaction times; BIS = balanced integration score. See the online article for the color version of this figure.

distractors in the additional-singleton paradigm and to our expectations based on previous work with cross-modal effects on attention and experiences in daily life. However, not only were the distractors irrelevant to the task, but also sound, in general, was irrelevant during the whole testing session. In real life, this hardly ever occurs because survival heavily depends on attending to alarming sounds (such as acoustic warnings or approaching predators). Also note that, in the standard additional-singleton paradigm, the distractor modality is always relevant, because both target and distractor are from the same (visual) modality. To test this explanation of our results, we made the auditory modality relevant while maintaining its irrelevance to the visual-search task. In particular, we introduced an additional task of counting the occurrence of a, sparsely presented counting sound. Trials with a counting sound (and no distractor; ~4.7% of trials) were entered into the random sequence of distractor-present and distractor-absent trials to ensure that the auditory modality is globally being attended to. Thus, the main search task remained the same as before, with the auditory distractor still fully irrelevant. The extra trials with the to-be-counted sound were excluded from the analysis.

In Experiment 5a, we finally observe an interference effect because of the auditory distractor. To establish that this effect is robust, especially after repeating evidence for the absence of distraction from four experiments, we replicated Experiment 5a with a new set of participants (Experiment 5b). To make full use of these data ($N = 2 \times 16 = 32$) and to quantify the accumulated evidence, we additionally report the BFs for the combined data set ($BF_{10}^{combined}$).

Method

Experiment 5a

Sixteen participants took part in Experiment 5a (median age: 26 years; range: 22–38 years; 7 female). Two participants (one female) participated in Experiment 1, one participant (female) participated in Experiment 3 and one participant (female) participated in Experiment 4 earlier. The design was identical to Experiment 1, except that a few additional trials (ranging from 3 to 6 in number) were presented in each block—which contained a sound different from the auditory distractor that had to be counted. These additional trials never contained a distractor. Thus, each block now consisted of $96 + x$ trials, with x being the number of trials with the counting sound and ranging from 3 to 6. As in Experiment 1, we ran 15 blocks of trials with the first block being a (nonanalyzed) practice phase. For half of the participants the auditory distractor was a high-frequency tone (520 Hz pure-tone presented at ~75 dB(A))—instead of a pink noise (as in Experiment 1)—with the counting sound being a low-frequency tone (440 Hz pure-tone presented at ~75 dB(A)). For the remaining participants, the auditory distractor was the low-frequency tone, with the counting sound being the high-frequency tone. Along with the visual-search task, participants were asked to count the number of times the additional sound (counting sound) was presented and report this number at the end of each block.

Experiment 5b

Experiment 5b is an exact replication of Experiment 5a with a new set of 16 participants (median age: 27 years; range: 21–44 years; 10 female). Two participants (1 female) participated in Experiment 1, one participant (female) participated in Experiment 2, and one participant (female) participated in Experiment 3 earlier.

Results

Experiment 5a

To ensure that we analyzed only periods where participants actually attended to the auditory modality, only trials from the blocks with correct answers for the counting task were used (12.37 blocks on average across participants, with the participant performing worst on the counting task contributing nine blocks and five participants without any counting errors contributing all 14 blocks). We analyzed only trials without a counting sound, which were comparable to Experiment 1. Contrary to Experiment 1, there was no significant change in RTs for distractor-present trials ($M = 592$ ms, $Mdn \pm \text{MAD} = 579 \pm 65$ ms) compared to distractor-absent trials ($M = 584$ ms, $Mdn \pm \text{MAD} = 581 \pm 56$ ms), $W = 83.00$, $p = .464$, $r = .22$, $\text{BF}_{10} = 0.42$ and error rates significantly increased (Figure 5a) on distractor-present ($M = 2.87\%$, $Mdn \pm \text{MAD} = 2.46 \pm 0.57\%$) compared to distractor-absent trials ($M = 2.19\%$, $Mdn \pm \text{MAD} = 2.25 \pm 0.89\%$), $W = 109.00$, $p = .034$, $r = .60$, $\text{BF}_{10} = 3.85$. Most importantly, the BIS revealed a significant decrease in performance (Figure 5b) when the distractor was present ($M = -.28$, $Mdn \pm \text{MAD} = -.12 \pm 0.88$) compared to when it was absent ($M = .28$, $Mdn \pm \text{MAD} = .39 \pm 1.22$), $W = 21.00$, $p = .013$, $r = -.69$, $\text{BF}_{10} = 10.60$.

Experiment 5b

Experiment 5b fully replicates the pattern of results observed in Experiment 5a. For the trials in blocks with a correct counting response (11.81 blocks on average across participants, min = 10, max = 14 [for three participants]), there was no significant change in RTs for distractor-present trials ($M = 593$ ms, $Mdn \pm \text{MAD} = 559 \pm 62$ ms) compared to distractor-absent trials ($M = 597$ ms, $Mdn \pm \text{MAD} = 572 \pm 57$ ms), $W = 46.00$, $p = .274$, $r = -.32$, $\text{BF}_{10} = 0.43$. Error rates showed a significant increase (Figure 5c) on distractor-present ($M = 4.01\%$, $Mdn \pm \text{MAD} = 3.42 \pm 1.68\%$) compared to distractor-absent trials ($M = 2.87\%$, $Mdn \pm \text{MAD} = 2.35 \pm 1.32\%$), $W = 117.00$, $p = .009$, $r = .72$, $\text{BF}_{10} = 12.05$. Here again, BIS revealed a significant decrease in performance (Figure 5d) when the distractor was present ($M = -.23$, $Mdn \pm \text{MAD} = 0.32 \pm 0.84$) compared to when it was absent ($M = .23$, $Mdn \pm \text{MAD} = 0.54 \pm 0.86$), $W = 22.00$, $p = .016$, $r = -.68$, $\text{BF}_{10} = 8.48$.

The combined data set revealed moderate evidence for the null hypothesis for RTs, $\text{BF}_{10}^{\text{combined}} = 0.19$, and strong evidence for the alternative hypothesis for error rates, $\text{BF}_{10}^{\text{combined}} = 144.32$, and BIS, $\text{BF}_{10}^{\text{combined}} = 327.21$.

Discussion

Experiment 5 finally reveals an interference effect on visual-search performance because of the auditory distractor. Note that, since the design was similar to that of Experiment 1, the well-established accessory-stimulus effect (Hackley & Valle-Inclán, 1999; Jepma et al., 2009) should have resulted in a decrease in RTs when the auditory distractor was present (as observed in Experiment 1). However, we do not observe any significant effect on RTs in Experiments 5a and 5b individually and even moderate evidence in favor of the absence of such an effect in the data combined across both experiments. This indicates that the accessory-stimulus effect (which speeds up responses) and the interference effect (which slows down responses) cancel each other. Consequently, there is no observable difference in RTs. Fortunately, BIS allows to analyze performance corrected for condition-specific SATs and demonstrates that, in contrast to Experiment 1, the distractor now reliably produced interference.

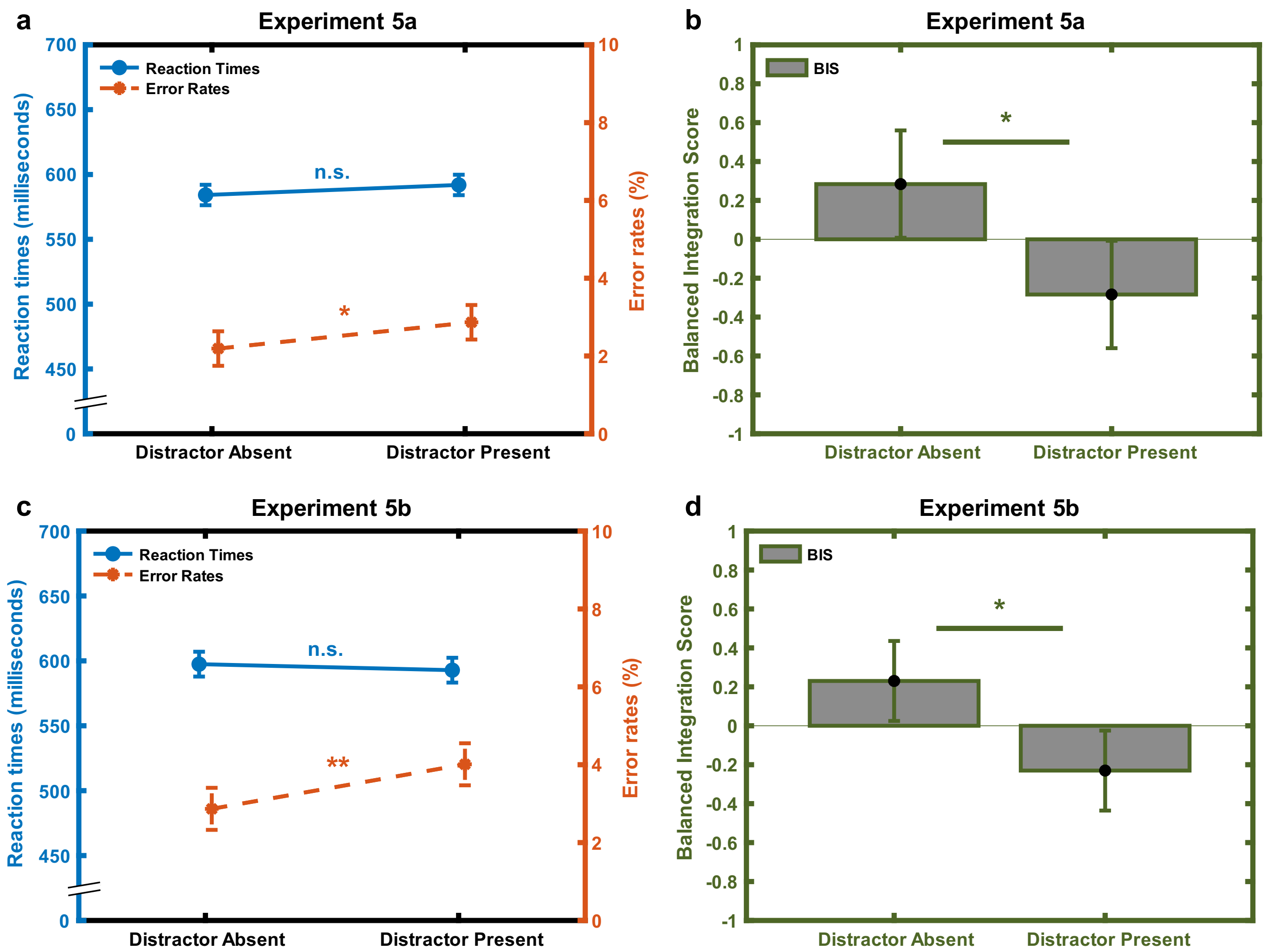
Hence, BIS (and error rates) finally demonstrate a significant decrease in performance because of the presence of the auditory distractor that emerges only when the auditory modality is globally being attended to. Taking the results of the current experiment in the context of the previous experiments, it looks like there exists an attentional mechanism which allows ignoring the auditory distractors completely during visual search.

Robustness to the Choice of Priors for BFs

To check for the robustness of the results to the choice of the priors, we conducted a sensitivity analysis for the effect of distractor presence on BIS for all experiments. Table 1 shows the resulting BFs, obtained from the Bayesian Wilcoxon signed-rank test, under three different choices of diffused priors (as recommended by Depaoli et al., 2020; Kruschke, 2021, in the absence of an informed prior). The first choice, which was also used in the tests reported throughout the article, was the default prior in JASP, which is a standard Jeffreys-Zellner-Siow Cauchy prior with the scale factor (r) of 0.707 ($=\sqrt{2}/2$). The second prior we used is a wide prior with $r = 1.0$ (recommended as default by Rouder et al., 2009), and the third prior is an ultrawide prior with $r = 1.4$. The choice of the wide and ultrawide priors is as recommended by JASP as robustness checks. The BFs in Table 1 do not vary drastically, demonstrating that our conclusions do not depend on the choice of priors.

Figure 5

Results From Experiment 5a (a, b) and Experiment 5b (c, d)



Note. (a, c) Mean (average across participants) of median reaction and error rates. (b, d) BIS, a combined performance measure. Error bars represent 95% within-participant confidence intervals. BIS = balanced integration score. See the online article for the color version of this figure.

General Discussion

In a series of experiments, we tried extensively to induce interference by an auditory distractor on visual search. No notable distraction was observed when sound was globally irrelevant (Experiments 1 and 2), even when the distractor was presented as a rare oddball stimulus (Experiment 3) or, additionally, with a temporal advantage of 300 ms (Experiment 4). However, once we made the auditory modality globally relevant (while maintaining its irrelevance to the visual-search task; Experiment 5), we finally observed significant distraction with decreased performance on distractor-present compared to distractor-absent trials.

The visual-attention literature typically differentiates two classes of distractor interference: attentional capture and (nonspatial) filtering costs. "Attentional capture" refers to the phenomenon that focused attention is initially drawn to the distractor and can only afterward move on to the search target (Liesefeld & Müller, 2019; Liesefeld et al., 2023; Theeuwes, 2010; see below for other uses of the term "attentional capture" in the auditory-distraction literature). There is an ongoing debate on whether behavioral distractor-interference effects in visual search must necessarily reflect attentional capture, or alternatively, can (also) be interpreted as delayed attention allocation to the target for other reasons ("nonspatial filtering costs"; Becker, 2007; Liesefeld et al., 2019). In terms of biased competition, the latter would reflect that the distractor entered but did not win the competition for attention, so that it does not actually draw attention. For example, the competitor might hamper attending the target because it produces another peak on the priority map (see Introduction) and this ambiguity needs to be resolved before focused attention can be allocated to the target. Obviously, in Experiments 1–4, neither form of distraction—attentional capture or nonspatial filtering costs—was induced by an auditory distractor when sound was globally irrelevant. Based on our broader theoretical position (see below), we tend to believe that the interference observed in Experiment 5 reflects (nonspatial) filtering costs, because even visual distractors do not typically produce attentional capture in

Table 1

BFs in Favor of the Presence (BF_{10}, Alternative Hypothesis) or Absence (BF_{01}, Null Hypothesis) of the Effect of Distractor-Presence on BIS With Three Different Choices of Prior Widths (r_{scale})

		r_{scale}		
Experiment	Favored hypothesis	0.707	1.0	1.4
1	Null	3.81	3.79	3.78
2	Null	3.76	3.79	3.75
3	Null	3.67	3.68	3.67
4	Null	3.21	3.20	3.21
1, 2, 3, and 4 (accumulated evidence)	Null	6.90	6.88	6.89
5a	Alternative	10.60	10.28	9.16
5b	Alternative	8.48	8.32	8.68
5a and 5b (accumulated evidence)	Alternative	327.21	327.97	352.52

Note. BFs are calculated using the Bayesian Wilcoxon signed-rank test. BF = Bayes factor; BIS = balanced integration score.

comparable visual searches and interference effects of the magnitude observed here typically go along with evidence for successfully prevented attentional capture (Liesefeld et al., 2017, 2019, 2023). However, more direct markers of attention allocation (e.g., eye movements or electrophysiology) are needed to substantiate this claim (see Liesefeld & Müller, 2019). For the present purposes, it suffices to recognize that (any form of) auditory distraction on visual search occurs only if sound is globally relevant.

A trivial explanation for our failure to observe distraction in Experiments 1–4 could be that our distractors were (still) not salient enough. However, a sound presented in isolation (our distractor in Experiment 1) is highly salient according to the definition of salience commonly accepted in research on visual search and detailed in the introduction (high local feature contrast) and a particularly high distractor salience is not even required to observe (some) interference in the additional-singleton paradigm (Zehetleitner et al., 2013). Furthermore, theories on auditory salience consider sounds with gaps and/or when presented on top of background noise (Experiment 2) as salient (Kayser et al., 2005). Additionally, oddball stimuli (Experiments 3 and 4) are known to be particularly attention grabbing (Escera et al., 1998, 2002). One could certainly further increase the distractor's salience, for example, by increasing its loudness until it must attract attention; as expressed in the popular quote by Sully "One would like to know the fortunate (or unfortunate) man who could receive a box on the ear and not attend to it" (Sully, 1892, p. 146). Still, short of harming the ear, our results show that auditory distractors do not interfere with visual search when the auditory modality is irrelevant. Actually, our study seems to provide an appropriate response to Sully's implied riddle: if the sound (or the box) harms the ear, the respective modality arguably becomes relevant to the receiver and is therefore attended.

How can we reconcile our findings with the observation that "as any parent can attest, their visual attention to the stimuli beyond the windshield can be disrupted if their attention is captured by the auditory signals from the back seat of the car" (Wolfe, 2021)? Given that the auditory signals in this example are emitted from the parents' children, we would argue that—in contrast to the completely uninformative distractors used in the present study—they are never fully task-irrelevant (with the long-term goal of raising happy children in mind). Furthermore, even if the parents could convince themselves that sounds emitted from their children are temporarily irrelevant (because they sit safely in their seats), they would probably still not want to ignore the auditory modality completely, because it also informs their task of steering their vehicle. As demonstrated in Experiment 5, the global relevance of auditory stimuli would suffice to hamper full suppression of distracting sound from the back of the car. More generally, whenever sound might signal reward or danger (Folyi et al., 2016; Kim & Anderson, 2021; Kim et al., 2022)—which are inevitably relevant to human survival—it obviously is always a bad idea to completely ignore salient sounds globally. Nevertheless, our cognitive machinery seems to have this extreme option in store for focusing on other modalities when needed.

Current-day examples where ignoring sounds globally is advantageous include cognitively demanding tasks in safe environments such as participating in a psychological experiment or working on a train or airplane. Now that we have developed a laboratory task that can in principle induce auditory distraction of concurrent visual search (Experiment 5), future research can start examining (a) under which conditions sound can be blocked globally and (b) how more selective distractor handling is possible, for example, by adapting to statistical regularities (Ferrante et al., 2018; Goschy et al., 2014; Sauter et al., 2018; Wang & Theeuwes, 2018) or by adapting to the specific relationship between the counting sound and the auditory distractor (e.g., in terms of similarity, feature dimension, or relative salience; see Gaspar & McDonald, 2014; Liesefeld et al., 2019, 2022; van Zoest & Donk, 2004).

It is important to keep in mind, though, that any predictions regarding auditory distraction that are derived from research on visual distraction are necessarily highly speculative, because of the many glaring differences between the two modalities. Here, we will highlight only two of these that appear of special importance for the present study. First, it is currently impossible to equate the salience of visual and auditory stimuli, as there is not even an established method to equate salience across different visual dimensions (for promising approaches, see Nothdurft, 2000; Zehetleitner et al., 2013). Any differences between distractors from different modalities can therefore always be explained by hypothetical differences in salience. This is part of the reason why the present study used various different distractor sounds, focusing on sounds that are thought to be highly salient (see above). Second, space plays a fundamentally different role for audition than it does for vision. Whereas visual stimuli are necessarily perceived to occupy a specific location, the location of auditory stimuli is generally more ambiguous so that a listener can easily be misled about it (spatial ventriloquism effect; Chen & Vroomen, 2013) or even perceive sounds as "nonspatial" (Van der Burg et al., 2008). At the very least, the spatial resolution of audition is much poorer than that of vision. One implication for research on auditory distraction of visual search is that it is hard to control where even a sound emitted from a specific position in space (e.g., a loudspeaker installed inside the screen) is perceived as being emitted from. Given the typical visual-search displays, the spatial resolution of audition is simply insufficient to tell whether a sound comes from the target versus any of the nontarget stimuli in the display or even from outside of the visual display. This stands in sharp contrast to visual distractors in the additional-singleton paradigm, which are typically presented within the search array at a potential target location, but not at the location of the current-trial target (see Supplement III in the online supplemental materials for an example of an additional-singleton experiment that is closely modeled after the present study design). Thus, it is unclear whether space-based suppression mechanisms (e.g., Gaspelin & Luck, 2018; Liesefeld & Müller, 2021; Liesefeld et al., 2017) can be applied to auditory distractors at all and whether an auditory distractor falls within or outside of the hypothetical spatial "attentional window" (e.g., Theeuwes, 2004). This difference between audition and vision is particularly important for the present line of research, because one could explain the absence of interference in Experiments 1–4 by suppression/ignoring the perceived location of the diotic sound (if it has a location at all, see Introduction section). If the counting sound in Experiment 5 was perceived as being emitted from the same location as the auditory distractor, it rendered this location relevant for the additional counting task and this might be the reason why participants were unable to suppress it. We find this explanation unlikely for two reasons: (a) in other studies using visual distractors, even distractors at completely irrelevant locations (e.g., Di Caro et al., 2019; Forster & Lavie, 2008a, 2008b) or at suppressed locations (Ferrante et al., 2018; Sauter et al., 2018; Wang & Theeuwes, 2018) have produced at least some interference, indicating that spatial suppression is never complete and we should have observed at least some interference in Experiments 1–4. (b) The low spatial resolution of audition renders it implausible that spatial suppression mechanisms are particularly effective, because much larger areas of space would need to be suppressed in order to prevent the spatially unspecific auditory stimuli from evading this suppression. Still, this theoretical possibility needs to be tested in future research, because spatial suppression of auditory distractors might turn out to be more rather than less efficient or auditory distractors might never become quite as salient as visual distractors and are therefore easier to suppress spatially.

Auditory Distraction in Other Paradigms

As auditory distraction on visual search has—to our knowledge—never been systematically examined before, in the following, we aim to relate our findings to effects of irrelevant auditory stimuli in several other, rather disparate, research traditions. First, our results seem to directly contradict studies showing that spoken distractor letters can affect search for visual target letters (Tellinghuisen & Nowak, 2003), and spoken distractor color words can affect search for colored shapes (Matusz et al., 2015). However, the relevant "modality" in these studies might be letters and colors in a more abstract categorical rather than a visual sense and it is the specific category (rather than the presentation domain) that enables distraction. Thus, the robustness of visual search against auditory distraction observed here might be purely sensory while another route of interference is opened when the visual target and the auditory distractor (sometimes) share categorical features. Other everyday experiences might be explained by this assumption (i.e., that cross-modal interference is more dependent on the content rather than on the modality): it is hard to focus on reading or writing a text when people are chatting in the same room. Relatedly, even visually presented verbal material (e.g., letters or digits) might engage the auditory system (Baddeley, 2012), thus potentially enabling auditory distraction as observed in our Experiment 5.

Second, a large share of research on auditory distraction more generally, employs tasks with serial presentation of memoranda that have to be recalled in the correct order. Results from this task are often explained by the duplex-mechanism account (Hughes, 2014). According to this influential theoretical stance, there are two general types of distraction. Sounds can either cause distraction when they share a relevant feature with the focal task—interference-by-process—or when the sound draws (nonspatial) attention away from the focal task—attentional capture. One well-established instance of interference-by-process occurs when background sounds that repeatedly change over time (changing-state sounds) impair serial recall of visually presented stimuli (typically letters or words; irrelevant sound effect; e.g., Colle & Welsh, 1976; Jones et al., 1992; Salamé & Baddeley, 1982). It is assumed that the successive changes in sound give rise to cues pertaining to the order of sound, which in turn interferes with the rehearsing of the visually presented memoranda for later serial recall (Jones & Macken, 1993). Notably, serial recall is typically not affected by otherwise similar nonchanging ("steady-state") sounds, which would not produce any ordering cues that might interfere with the serial recall. Hughes (2014) distinguishes interference-by-process from another form of distraction which is termed as "attentional capture." In contrast to the visual-distraction community (see above), Hughes conceives of attentional capture as a withdrawal of processing resources from the current task. One might argue that processing the auditory distractor in Experiment 5 shares processes with processing stimuli on the (secondary) counting task. However, as there is no apparent reason to believe that processing of our auditory distractor shares any specific processes with processing of the visual target in the main task, the interference we observed there would probably be classified as "attentional capture" in the duplex-mechanism account (recall again that the meaning of the term differs across research communities).

SAT Effect as an Index of Distractor Processing

Although they did not hamper target processing, the distractor sounds in Experiments 1 and 4 were processed, as evidenced by a significant reduction of RTs. This effect alone would typically be interpreted as evidence of counterintuitive search facilitation by these distractors. However, BIS results suggested differently—with no significant increase or decrease in actual performance because of the distractor sounds (i.e., they did not distract). Rather the pattern of results in RTs, error rates and BIS taken together indicates that the distractors affected the SAT: processing (of the target) was speeded at the cost of higher error rates. Similar effects on the SAT by irrelevant spatial sounds (noninformative cues) have been observed before and these have been interpreted in terms of signal-detection theory as "criterion shifts" (Spence & Driver, 1997). In current models of decision making, SATs are typically thought to affect the decision threshold (Liesefeld & Janczyk, 2019, 2023; Ratcliff & McKoon, 2008; but see Lerche & Voss, 2018), whereas true distraction would affect either the nondecision time (when processing of the target starts) or the rate of evidence accumulation (how fast the target is processed). Either of the latter two is affected (in addition) when we make the auditory modality relevant in Experiment 5, where BIS indicates a significant decrease in performance because of the distractor's presence. Thus a methodological lesson that can be learned from the current study is that interpreting results of RTs and error rates individually—without considering a performance measure that is relatively insensitive to SATs, like the BIS (Liesefeld & Janczyk, 2019, 2023)—can lead to erroneous conclusions.

Of importance from a cognitive perspective, the observed SAT effect indicates that sounds in these experiments were processed to a certain degree. Thus, the absence of interference by our auditory distractors is not because of the sound not being processed, but rather must be explained by a mechanism that shields the allocation of focused attention from being hampered by this irrelevant information as long as the auditory modality is completely irrelevant (we will speculate on this mechanism further below). Here, we can only speculate on the nature of this unexpected SAT effect by relating it to other phenomena known from other strands of research. Two (potentially related) phenomena known as alerting- (or warning-) signal and accessory-stimulus effect, respectively, seem to be relevant in this regard (although these are sometimes equated, see Low et al., 1996). An alerting signal is a stimulus that is informative with regard to the imminent onset of the task-relevant stimuli. It is supposed to bring the attentional system into a readied state and therefore improve performance (e.g., Matthias et al., 2010). While an alerting stimulus is typically played in advance, an accessory stimulus is a salient-but-irrelevant stimulus that is presented together with an "imperative" stimulus (i.e., a stimulus requiring a response) from another modality. Response times (and sometimes also proportion of correct responses) are decreased when an accessory stimulus is present. This effect seems to be particularly strong for visual imperative (task-relevant) stimuli combined with auditory accessory (task-irrelevant) stimuli as in the present study (Hackley & Valle-Inclán, 1999; Jepma et al., 2009).

In Experiment 4, the auditory distractor was presented before display onset and was always followed by a search display, so that the former can be said to be informative of the latter. Therefore, this distractor might have acted as an alerting signal. In Experiment 1, by contrast, the auditory distractor was presented concurrently with the visual-search display and is, thus, more readily interpreted as an accessory stimulus. Thus, it appears possible that different phenomena explain the similar-looking SAT effects in the two experiments. Alternatively, one could argue that the onset of a sound is detected slightly earlier than the onset of a visual search display (given that RTs for auditory stimuli are typically 20–60 ms shorter than for visual stimuli; Lewald & Guski, 2003; Shelton & Kumar, 2010—likely as a result of faster transduction time in the inner ear compared to the retina; Recanzone, 2009; Stein & Meredith, 1993). Therefore, the concurrently presented auditory distractor in Experiment 1 might have acted as a warning signal as well: the attentional system was readied slightly earlier on trials compared to trials without a distractor. If that is the case, it becomes difficult to distinguish warning-signal and accessory-stimulus effects.

Several explanations for the accessory-stimulus effect have been brought forward (see Jepma et al., 2009, for a concise summary) and these might apply to the warning-signal effect as well. According to the energy integration hypothesis, the irrelevant stimulus increases the perceived intensity of the task-relevant stimulus (Stein et al., 1996). As in the present case, this would also increase the perceived intensity of the nontargets surrounding the target, the guidance toward the target would not improve, and we therefore think that this is unlikely to be the correct account for our findings. Alternatively, the accessory stimulus might affect parameters of evidence integration or motor execution. In terms of the drift-diffusion model (Ratcliff, 1978; Ratcliff et al., 2016), these parameters would be the drift rate (speed of information uptake), the decision threshold (amount of information needed to make a decision), or the nondecision time (all processes unrelated to the decision proper, including the motor response), respectively. Given that changes in the threshold parameter yield SAT effects (e.g., Liesefeld & Janczyk, 2019, 2023), our data seem to be most parsimoniously explained by an accessory-stimulus effect on the decision threshold.

In a study quite similar to ours, Gao et al. (2021) obtained a different pattern of results. Participants were asked to find a target defined by a vertical or horizontal line segment among distractors tilted to the left and to the right. Target and distractors were enclosed by white circles. A diotic auditory stimulus accompanied half of the trials and was presented with the onset of the visual display. They also did not observe any effect of the auditory stimulus on visual search performance and, in contrast to our results, they did not even observe the SAT effect discussed in the preceding paragraphs. This discrepancy might be explained by considering that they observed a significant set-size effect, indicating inefficient search (likely because of their use of heterogeneous distractors, namely different tilt directions; Duncan & Humphreys, 1989). As detailed in the introduction, there is reason to believe that in order to examine interference by a salient distractor it is important to use an efficient search task.

Theoretical Implications for Theories of Visual Search

If sounds were processed, but do not interfere, some attentional mechanism at a relatively late processing stage (Deutsch & Deutsch, 1963) must have prevented interference in the present study. To understand what attentional mechanism this might be, it is useful to consider the massive knowledge base on visual distraction. Using the N2pc component of the event-related potential as a marker of covert shifts of attention (Constant et al., 2023; Eimer, 2014), Liesefeld et al. (2017) have shown that a distractor standing out in the same dimension as the search target (orientation) reliably captures attention (see also Liesefeld et al., 2019; Sauter et al., 2021, for respective behavioral and eye-movement evidence). This was the case even though the distractor feature was predictable and categorically different from the target feature (e.g., left- vs. right-tilted), and the distractor occurred with high prevalence (2/3rds of all trials). By contrast, when a predictable salient distractor stands out in a different feature dimension (e.g., a color distractor during search for an orientation target), it does not usually capture attention (while still producing behavioral interference; Jannati et al., 2013; Liesefeld et al., 2019, 2023). This discrepancy is predicted by the long-standing dimension weighting account (Found & Müller, 1996; Liesefeld & Müller, 2019): during integration on a search-guiding priority map, saliency signals from the distractor dimension are down-weighted with respect to those from the target dimension (see Liesefeld & Müller, 2021, for computational details), so that, after this integration of bottom-up (saliency) and top-down (dimension weighting) influences, the target achieves the highest priority. Such a differential weighting is not possible for same-dimension distractors, so that initial attention allocations in such situations are exclusively determined by saliency (see also, van Heusden et al., 2022; Van Zoest & Donk, 2008).

On that background, it appears that relative down-weighting of salience signals emitted by auditory distractors during search for a visual target is highly efficient: even though they do not capture attention, down-weighted visual distractors usually produce residual interference as measurable via RTs (see Liesefeld et al., 2019, 2023), which, compared to same-dimension distraction, is reduced by an order of magnitude. Across four experiments (Experiments 1–4), we here did not even observe this residual interference, indicating that the down-weighting of different-modality distractors is even more efficient than the down-weighting of different-dimension distractors. This would be predicted by an extension of the dimension-weighting account, the modality-weighting account (Nasemann et al., 2023; Töllner et al., 2009). A recent study by Tsai et al. (2023) has provided evidence toward this prediction for tactile distractors. Comparable to our study, they also show an absence of interference effects of their tactile distractors on visual search, indicating a down-weighting of the tactile modality. Conversely, when we made the auditory modality relevant in Experiment 5—by introducing the additional counting task—its priority (and thus the modality weight) increased because of the top-down demand of the additional task. Thus, we argue, the global relevance of the auditory modality prevents the (full) down-weighting of salience signals emitted from the auditory distractor, resulting in a performance cost because of the auditory distractor.

Such (full) down-weighting of the auditory modality can be thought of as an attentional barrier for the auditory system, which might help guard us against interference by auditory information if need be. Not all our senses are equipped with a physical barrier to filter out all information from a specific modality when one wishes to focus fully on another modality—like the eyelids are for the eye. The presence of eyelids is highly efficient to guard us against the entirety of visual information if need be—and focus on what is important to us in that moment (e.g., listening to music or meditating). By shielding cognitive mechanisms involved in the allocation of attention from auditory distraction, this attentional barrier in the auditory modality can potentially save us from drowning in the sea of irrelevant sounds we are confronted with in our everyday life. As this barrier would fulfill a similar function as the eyelids do for the visual modality, but only attentionally, we call this proposed mechanism the Attentional Earlid. Shutting or opening this attentional earlid might be voluntary (goal-driven) or it might emerge as a consequence of experience with the task (selection history; Awh et al., 2012).

References

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics*, *55*(5), 485–496. https://doi.org/10.3758/BF03205306

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Becker, S. I. (2007). Irrelevant singletons in pop-out search: Attentional capture or filtering costs? *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 764–787. https://doi.org/10.1037/0096-1523.33.4.764

Bernstein, I. H., Clark, M. H., & Edelstein, B. A. (1969). Effects of an auditory signal on visual reaction time. *Journal of Experimental Psychology*, *80*(3, Pt.1), 567–569. https://doi.org/10.1037/h0027444

Best, V., Baumgartner, R., Lavandier, M., Majdak, P., & Kopčo, N. (2020). Sound externalization: A review of recent research. *Trends in Hearing*, *24*, Article 2331216520948390. https://doi.org/10.1177/2331216520948390

Bisley, J. W., & Mirpour, K. (2019). The neural instantiation of a priority map. *Current Opinion in Psychology*, *29*, 108–112. https://doi.org/10.1016/j.copsyc.2019.01.002

Chelazzi, L., Marini, F., Pascucci, D., & Turatto, M. (2019). Getting rid of visual distractors: The why, when, how, and where. *Current Opinion in Psychology*, *29*, 135–147. https://doi.org/10.1016/j.copsyc.2019.02.004

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics*, *75*(5), 790–811. https://doi.org/10.3758/s13414-013-0475-4

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*, 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning & Verbal Behavior*, *15*(1), 17–31. doi.org/10.1016/S0022-5371(76)90003-7

Constant, M., & Liesefeld, H. R. (2021). Massive effects of saliency on information processing in visual working memory. *Psychological Science*, *32*(5), 682–691. https://doi.org/10.1177/0956797620975785

Constant, M., Mandal, A., Asanowicz, D., Yamaguchi, M., Gillmeister, H., Kerzel, D., Luque, D., Pesciarelli, F., Fehr, T., Mushtaq, F., Pavlov, Y. G., & Liesefeld, H. R. (2023). *A multilab investigation into the N2pc as an indicator of attentional selectivity: Direct replication of Eimer (1996)*. PsyArXiv. https://doi.org/10.31234/osf.io/3472y

Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods*, *46*(4), 1149–1151. https://doi.org/10.3758/s13428-013-0441-z

Dalton, P., & Lavie, N. (2004). Auditory attentional capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(1), 180–193. https://doi.org/10.1037/0096-1523.30.1.180

Dalton, P., & Spence, C. (2007). Attentional capture in serial audiovisual search tasks. *Perception & Psychophysics*, *69*(3), 422–438. https://doi.org/10.3758/BF03193763

Depaoli, S., Winter, S. D., & Visser, M. (2020). The importance of prior sensitivity analysis in Bayesian statistics: Demonstrations using an interactive shiny app. *Frontiers in Psychology*, *11*, Article 608045. https://doi.org/10.3389/fpsyg.2020.608045

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*(1), 80–90. https://doi.org/10.1037/h0039515

Di Caro, V., Theeuwes, J., & Della Libera, C. (2019). Suppression history of distractor location biases attentional and oculomotor control. *Visual Cognition*, *27*(2), 142–157. https://doi.org/10.1080/13506285.2019.1617376

Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *353*(1373), 1319–1331. https://doi.org/10.1098/rstb.1998.0286

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*(10), 526–535. https://doi.org/10.1016/j.tics.2014.05.005

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, *10*(5), 590–604. https://doi.org/10.1162/089892998562997

Escera, C., Corral, M.-J., & Yago, E. (2002). An electrophysiological and behavioral investigation of involuntary attention towards auditory frequency, duration and intensity changes. *Cognitive Brain Research*, *14*(3), 325–332. https://doi.org/10.1016/S0926-6410(02)00135-0

Escera, C., Yago, E., & Alho, K. (2001). Electrical responses reveal the temporal dynamics of brain events during involuntary attention switching. *European Journal of Neuroscience*, *14*(5), 877–883. https://doi.org/10.1046/j.0953-816x.2001.01707.x

Escera, C., Yago, E., Corral, M.-J., Corbera, S., & Nuñez, M. I. (2003). Attention capture by auditory significant stimuli: Semantic analysis follows attention switching. *European Journal of Neuroscience*, *18*(8), 2408–2412. https://doi.org/10.1046/j.1460-9568.2003.02937.x

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382–390. https://doi.org/10.1016/j.tics.2006.06.011

Ferrante, O., Patacca, A., Di Caro, V., Della Libera, C., Santandrea, E., & Chelazzi, L. (2018). Altering spatial priority maps via statistical learning of target selection and distractor filtering. *Cortex*, *102*, 67–95. https://doi.org/10.1016/j.cortex.2017.09.027

Folk, C. L., & Remington, R. W. (2008). Bottom-up priming of top-down attentional control settings. *Visual Cognition*, *16*(2–3), 215–231. https://doi.org/10.1080/13506280701458804

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030–1044. https://doi.org/10.1037/0096-1523.18.4.1030

Folyi, T., Liesefeld, H. R., & Wentura, D. (2016). Attentional enhancement for positive and negative tones at an early stage of auditory processing. *Biological Psychology*, *114*, 23–32. https://doi.org/10.1016/j.biopsycho.2015.12.001

Forster, S., & Lavie, N. (2008a). Attentional capture by entirely irrelevant distractors. *Visual Cognition*, *16*(2–3), 200–214. https://doi.org/10.1080/13506280701465049

Forster, S., & Lavie, N. (2008b). Failures to ignore entirely irrelevant distractors: The role of load. *Journal of Experimental Psychology. Applied*, *14*(1), 73–83. https://doi.org/10.1037/1076-898X.14.1.73

Found, A., & Müller, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a "dimension-weighting" account. *Perception & Psychophysics*, *58*(1), 88–101. https://doi.org/10.3758/BF03205479

Gao, M., Chang, R., Wang, A., Zhang, M., Cheng, Z., Li, Q., & Tang, X. (2021). Which can explain the pip-and-pop effect during a visual search: Multisensory integration or the oddball effect? *Journal of Experimental Psychology: Human Perception and Performance*, *47*(5), 689–703. https://doi.org/10.1037/xhp0000905

Gaspar, J. M., & McDonald, J. J. (2014). Suppression of salient objects prevents distraction in visual search. *Journal of Neuroscience*, *34*(16), 5658–5666. https://doi.org/10.1523/JNEUROSCI.4161-13.2014

Gaspelin, N., & Luck, S. J. (2018). Combined electrophysiological and behavioral evidence for the suppression of salient distractors. *Journal of Cognitive Neuroscience*, *30*(9), 1265–1280. https://doi.org/10.1162/jocn_a_01279

Geyer, T., Müller, H. J., & Krummenacher, J. (2008). Expectancies modulate attentional capture by salient color singletons. *Vision Research*, *48*(11), 1315–1326. https://doi.org/10.1016/j.visres.2008.02.006

Goschy, H., Bakos, S., Müller, H. J., & Zehetleitner, M. (2014). Probability cueing of distractor locations: Both intertrial facilitation and statistical learning mediate interference reduction. *Frontiers in Psychology*, *5*, Article 1195. https://doi.org/10.3389/fpsyg.2014.01195

Hackley, S. A., & Valle-Inclán, F. (1999). Accessory stimulus effects on response selection: Does arousal speed decision making? *Journal of Cognitive Neuroscience*, *11*(3), 321–329. https://doi.org/10.1162/089892999563427

Hanning, N. M., Szinte, M., & Deubel, H. (2019). Visual attention is not limited to the oculomotor range. *Proceedings of the National Academy of Sciences*, *116*(19), 9665–9670. https://doi.org/10.1073/pnas.1813465116

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*, Article 150. https://doi.org/10.3389/fnins.2014.00150

Hillyard, S. A., Störmer, V. S., Feng, W., Martinez, A., & McDonald, J. J. (2016). Cross-modal orienting of visual attention. *Neuropsychologia*, *83*, 170–178. https://doi.org/10.1016/j.neuropsychologia.2015.06.003

Hughes, R. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, *3*(1), 30–41. https://doi.org/10.1002/pchj.44

Jannati, A., Gaspar, J. M., & McDonald, J. J. (2013). Tracking target and distractor processing in fixed-feature visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(6), 1713–1730. https://doi.org/10.1037/a0032251

JASP Team. (2021). *JASP* (Version 0.16.0) [Computer software]. https://jasp-stats.org/

Jepma, M., Wagenmakers, E.-J., Band, G. P. H., & Nieuwenhuis, S. (2009). The effects of accessory stimuli on information processing: Evidence from electrophysiology and a diffusion model analysis. *Journal of Cognitive Neuroscience*, *21*(5), 847–864. https://doi.org/10.1162/jocn.2009.21063

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 369–381. https://doi.org/10.1037/0278-7393.19.2.369

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *44*(4), 645–669. https://doi.org/10.1080/14640749208401304

Kahneman, D., Treisman, A., & Burkell, J. (1983). The cost of visual filtering. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(4), 510–522. https://doi.org/10.1037/0096-1523.9.4.510

Kayser, C., Petkov, C. I., Lippert, M., & Logothetis, N. K. (2005). Mechanisms for allocating auditory attention: An auditory saliency map. *Current Biology*, *15*(21), 1943–1947. https://doi.org/10.1016/j.cub.2005.09.040

Kerzel, D., & Huynh Cong, S. (2022). Biased competition between targets and distractors reduces attentional suppression: Evidence from the positivity posterior contralateral and distractor positivity. *Journal of Cognitive Neuroscience*, *34*(9), 1563–1575. https://doi.org/10.1162/jocn_a_01877

Kim, A. J., & Anderson, B. A. (2021). How does threat modulate the motivational effects of reward on attention? *Experimental Psychology*, *68*(3), 165–172. https://doi.org/10.1027/1618-3169/a000521

Kim, A. J., Grégoire, L., & Anderson, B. A. (2022). Value-biased competition in the auditory system of the brain. *Journal of Cognitive Neuroscience*, *34*(1), 180–191. https://doi.org/10.1162/jocn_a_01785

Kruschke, J. K. (2021). Bayesian Analysis reporting guidelines. *Nature Human Behaviour*, *5*(10), Article 10. https://doi.org/10.1038/s41562-021-01177-7

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review*, *13*(1), 132–138. https://doi.org/10.3758/BF03193824

Leiva, A., Parmentier, F. B. R., & Andrés, P. (2015). Aging increases distraction by auditory oddballs in visual, but not auditory tasks. *Psychological Research*, *79*(3), 401–410. https://doi.org/10.1007/s00426-014-0573-5

Leonard, C. J. (2021). Consensus emerges and biased competition wins: A commentary on Luck et al. (2021). *Visual Cognition*, *29*(9), 560–562. https://doi.org/10.1080/13506285.2021.1912236

Lerche, V., & Voss, A. (2018). Speed–accuracy manipulations and diffusion modeling: Lack of discriminant validity of the manipulation or of the parameter estimates? *Behavior Research Methods*, *50*(6), 2568–2585. https://doi.org/10.3758/s13428-018-1034-7

Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research*, *16*(3), 468–478. https://doi.org/10.1016/S0926-6410(03)00074-0

Liesefeld, H. R., Fu, X., & Zimmer, H. D. (2015). Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1140–1151. https://doi.org/10.1037/xlm0000081

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs? *Behavior Research Methods*, *51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Liesefeld, H. R., & Janczyk, M. (2023). Same same but different: Subtle but consequential differences between two measures to linearly integrate speed and accuracy (LISAS vs. BIS). *Behavior Research Methods*, *55*(3), 1175–1192. https://doi.org/10.3758/s13428-022-01843-2

Liesefeld, H. R., Lamy, D., Gaspelin, N., Geng, J., Kerzel, D., Schall, J., Allen, H. A., Anderson, B. A., Boettcher, S., Busch, N., Carlisle, N. B., Colonius, H., Draschkow, D., Egeth, H., Leber, A. B., Müller, H. J., Röer, J. P., Schubö, A., Slagter, H. A., … Wolfe, J. (2023). *Terms of debate: Consensus definitions to guide the scientific discourse on visual distraction*. PsyArXiv. https://doi.org/10.31234/osf.io/4b2gk

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2019). Distractor-interference reduction is dimensionally constrained. *Visual Cognition*, *27*(3–4), 247–259. https://doi.org/10.1080/13506285.2018.1561568

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2021). Attentional capture: An ameliorable side-effect of searching for salient targets. *Visual Cognition*, *29*(9), 600–603. https://doi.org/10.1080/13506285.2021.1925798

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2022). Preparatory control against distraction is not feature-based. *Cerebral Cortex*, *32*(11), 2398–2411. https://doi.org/10.1093/cercor/bhab341

Liesefeld, H. R., Liesefeld, A. M., Töllner, T., & Müller, H. J. (2017). Attentional capture in visual search: Capture and post-capture dynamics revealed by EEG. *NeuroImage*, *156*, 166–173. https://doi.org/10.1016/j.neuroimage.2017.05.016

Liesefeld, H. R., Moran, R., Usher, M., Müller, H. J., & Zehetleitner, M. (2016). Search efficiency as a function of target saliency: The transition from inefficient to efficient search and beyond. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 821–836. https://doi.org/10.1037/xhp0000156

Liesefeld, H. R., & Müller, H. J. (2019). Distractor handling via dimension weighting. *Current Opinion in Psychology*, *29*, 160–167. https://doi.org/10.1016/j.copsyc.2019.03.003

Liesefeld, H. R., & Müller, H. J. (2020). A theoretical attempt to revive the serial/parallel-search dichotomy. *Attention, Perception, & Psychophysics*, *82*(1), 228–245. https://doi.org/10.3758/s13414-019-01819-z

Liesefeld, H. R., & Müller, H. J. (2021). Modulations of saliency signals at two hierarchical levels of priority computation revealed by spatial statistical distractor learning. *Journal of Experimental Psychology: General*, *150*(4), 710–728. https://doi.org/10.1037/xge0000970

Liesefeld, H. R., & Müller, H. J. (2023). Target salience and search modes: A Commentary on Theeuwes (2023). *Journal of Cognition*, *6*(1), Article 1. https://doi.org/10.5334/joc.279

Low, K. A., Larson, S. L., Burke, J., & Hackley, S. A. (1996). Alerting effects on choice reaction time and the photic eyeblink reflex. *Electroencephalography and Clinical Neurophysiology*, *98*(5), 385–393. https://doi.org/10.1016/0013-4694(96)95085-3

Luck, S. J., Gaspelin, N., Folk, C. L., Remington, R. W., & Theeuwes, J. (2021). Progress toward resolving the attentional capture debate. *Visual Cognition*, *29*(1), 1–21. https://doi.org/10.1080/13506285.2020.1848949

Marsh, J. E., Campbell, T. A., Vachon, F., Taylor, P. J., & Hughes, R. W. (2020). How the deployment of visual attention modulates auditory distraction. *Attention, Perception, & Psychophysics*, *82*(1), 350–362. https://doi.org/10.3758/s13414-019-01800-w

Matthias, E., Bublak, P., Müller, H., Schneider, W., Krummenacher, J., & Finke, K. (2010). The influence of alertness on spatial and nonspatial components of visual attention. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(1), 38–56. https://doi.org/10.1037/a0017602

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

Matusz, P. J., & Eimer, M. (2011). Multisensory enhancement of attentional capture in visual search. *Psychonomic Bulletin & Review*, *18*(5), 904–909. https://doi.org/10.3758/s13423-011-0131-8

Moher, J., Abrams, J., Egeth, H. E., Yantis, S., & Stuphorn, V. (2011). Trial-by-trial adjustments of top-down set modulate oculomotor capture. *Psychonomic Bulletin & Review*, *18*(5), 897–903. https://doi.org/10.3758/s13423-011-0118-5

Müller, H. J., Geyer, T., Zehetleitner, M., & Krummenacher, J. (2009). Attentional capture by salient color singleton distractors is modulated by top-down dimensional set. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 1–16. https://doi.org/10.1037/0096-1523.35.1.1

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, *320*, 264–265. https://doi.org/10.1038/320264a0

Nasemann, J., Töllner, T., Müller, H. J., & Shi, Z. (2023). Hierarchy of intra- and cross-modal redundancy gains in visuo-tactile search: Evidence from the posterior contralateral negativity. *Journal of Cognitive Neuroscience*, *35*(4), 543–570. https://doi.org/10.1162/jocn_a_01971

Nothdurft, H.-C. (1993). The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Research*, *33*(14), 1937–1958. https://doi.org/10.1016/0042-6989(93)90020-W

Nothdurft, H.-C. (2000). Salience from feature contrast: Variations with texture density. *Vision Research*, *40*(23), 3181–3200. https://doi.org/10.1016/S0042-6989(00)00168-1

Parmentier, F. (2016). Deviant sounds yield distraction irrespective of the sounds' informational value. *Journal of Experimental Psychology Human Perception & Performance*, *42*(6), 837–846. https://doi.org/10.1037/xhp0000195

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Posner, M. I., Klein, R., Summers, J., & Buggie, S. (1973). On the selection of signals. *Memory & Cognition*, *1*(1), 2–12. https://doi.org/10.3758/BF03198062

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*(1), 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time. *Hearing Research*, *258*(1), 89–99. https://doi.org/10.1016/j.heares.2009.04.009

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian T tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning & Verbal Behavior*, *21*(2), 150–164. https://doi.org/10.1016/S0022-5371

Sauter, M., Hanning, N. M., Liesefeld, H. R., & Müller, H. J. (2021). Post-capture processes contribute to statistical learning of distractor locations in visual search. *Cortex*, *135*, 108–126. https://doi.org/10.1016/j.cortex.2020.11.016

Sauter, M., Liesefeld, H. R., Zehetleitner, M., & Müller, H. J. (2018). Region-based shielding of visual search from salient distractors: Target detection is impaired with same-but not different-dimension distractors. *Attention, Perception, & Psychophysics*, *80*(3), 622–642. https://doi.org/10.3758/s13414-017-1477-4

Shelton, J., & Kumar, G. P. (2010). Comparison between auditory and visual simple reaction times. *Neuroscience & Medicine*, *1*(1), 30–32. https://doi.org/10.4236/nm.2010.11004

Spence, C. (2010). Crossmodal attention. *Scholarpedia*, *5*(5), Article 6309. https://doi.org/10.4249/scholarpedia.6309

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1–22. https://doi.org/10.3758/BF03206843

Stein, B. E., London, N., Wilkinson, L. K., & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *Journal of Cognitive Neuroscience*, *8*(6), 497–506. https://doi.org/10.1162/jocn.1996.8.6.497

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. MIT Press.

Stilwell, B. T., & Gaspelin, N. (2021). Attentional suppression of highly salient color singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(10), 1313–1328. https://doi.org/10.1037/xhp0000948

Sully, J. (1892). *The human mind: A text-book of psychology*. Longmans, Green.

Tellinghuisen, D. J., & Nowak, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics*, *65*(5), 817–828. https://doi.org/10.3758/BF03194817

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, *50*(2), 184–193. https://doi.org/10.3758/BF03212219

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review*, *11*(1), 65–70. https://doi.org/10.3758/BF03206462

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Theeuwes, J. (2023). The attentional capture debate: When can we avoid salient distractors and when not? *Journal of Cognition*, *6*(1), Article 35. https://doi.org/10.5334/joc.251

Theeuwes, J., Bogaerts, L., & van Moorselaar, D. (2022). What to expect where and when: How statistical learning drives visual selection. *Trends in Cognitive Sciences*, *26*(10), 860–872. https://doi.org/10.1016/j.tics.2022.06.001

Töllner, T., Gramann, K., Müller, H. J., & Eimer, M. (2009). The anterior N1 component as an index of modality shifting. *Journal of Cognitive Neuroscience*, *21*(9), 1653–1669. https://doi.org/10.1162/jocn.2009.21108

Townsend, J. T. (1972). Some results concerning the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, *25*(2), 168–199. https://doi.org/10.1111/j.2044-8317.1972.tb00490.x

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Tsai, S.-Y., Nasemann, J., Qiu, N., Töllner, T., Müller, H. J., & Shi, Z. (2023). Little engagement of attention by salient distractors defined in a different dimension or modality to the visual search target. *Psychophysiology*, Advance online publication. https://doi.org/10.1111/psyp.14375.

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(5), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E.-J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*(3), 813–826. https://doi.org/10.3758/s13423-020-01798-5

van Heusden, E., van Zoest, W., Donk, M., & Olivers, C. N. L. (2022). An attentional limbo: Saccades become momentarily non-selective in between saliency-driven and relevance-driven selection. *Psychonomic Bulletin & Review*, *29*(4), 1327–1337. https://doi.org/10.3758/s13423-022-02091-3

van Zoest, W., & Donk, M. (2004). Bottom-up and top-down control in visual search. *Perception*, *33*(8), 927–937. https://doi.org/10.1068/p5158

Van Zoest, W., & Donk, M. (2008). Goal-driven modulation as a function of time in saccadic target selection. *Quarterly Journal of Experimental Psychology*, *61*(10), 1553–1572. https://doi.org/10.1080/17470210701595555

Wang, B., & Theeuwes, J. (2018). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 13–17. https://doi.org/10.1037/xhp0000472

Wolfe, J. M. (1994). Guided search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*(1), 33–39. https://doi.org/10.1111/1467-9280.00006

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), Article 3. https://doi.org/10.1038/s41562-017-0058

Won, B.-Y., Kosoyan, M., & Geng, J. J. (2019). Evidence for second-order singleton suppression based on probabilistic expectations. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(1), 125–138. https://doi.org/10.1037/xhp0000594

Wykowska, A., & Schubö, A. (2011). Irrelevant singletons in visual search do not capture attention but can produce nonspatial filtering costs. *Journal of Cognitive Neuroscience*, *23*(3), 645–660. https://doi.org/10.1162/jocn.2009.21390

Zehetleitner, M., Koch, A. I., Goschy, H., & Müller, H. J. (2013). Salience-based selection: Attentional capture by distractors less salient than the target. *PLoS ONE*, *8*(1), Article e52595. https://doi.org/10.1371/journal.pone.0052595

Zhaoping, L. (2023). Peripheral and central sensation: Multisensory orienting and recognition across species. *Trends in Cognitive Sciences*, *27*(6), 539–552. https://doi.org/10.1016/j.tics.2023.03.001

Received December 20, 2022
Revision received August 24, 2023
Accepted September 13, 2023 ■

Supplementary Materials

The Surprising Robustness of Visual Search Against Concurrent Auditory Distraction

Supplement – I

Here, we present the results of Experiment 3, 5a and 5b without the data of the participants who took part in a previous experiment.

Experiment 3

Here again (n = 13), there was no significant change in RTs when the distractor was present (M = 694 ms, $Mdn \pm MAD$ = 633 ± 89 ms) compared to when it was absent (M = 701 ms, $Mdn \pm MAD$ = 666 ± 87 ms), W = 32.00, p = .376, r = -.29, BF_{10} = 0.39. No significant change in error rates was observed between distractor-present (M = 3.09 %, $Mdn \pm MAD$ = 3.27 ± 2.38 %) and distractor-absent trials (M = 2.57 %, $Mdn \pm MAD$ = 2.49 ± 1.38 %), W = 62.50, p = .249, r = .37, BF_{10} = 0.51. Additionally, BIS showed no change in performance between distractor-present (M = -0.12, $Mdn \pm MAD$ = -0.10 ± 0.96) and distractor-absent trials (M = 0.12, $Mdn \pm MAD$ = -0.05 ± 0.91), W = 32.00, p = .376, r = -.29, BF_{10} = 0.43.

Experiment 5a

There was no significant change in RTs (n = 12) when the distractor was present (M = 600 ms, $Mdn \pm MAD$ = 579 ± 75 ms) compared to when it was absent (M = 589 ms, $Mdn \pm MAD$ = 573 ± 66 ms), W = 56.00, p = .204, r = .44, BF_{10} = 0.84. No significant change in error rates was observed between distractor-present (M = 2.61 %, $Mdn \pm MAD$ = 2.27 ± 0.46 %) and distractor-absent trials (M = 1.99 %, $Mdn \pm MAD$ = 1.76 ± 0.86 %), W = 63.00, p = .064, r = .61, BF_{10} = 1.99. However, as expected BIS showed a significant change in performance between distractor-present (M = -0.28, $Mdn \pm MAD$ = -0.23 ± 1.10) and distractor-absent trials (M = 0.28, $Mdn \pm MAD$ = -0.65 ± 1.01), W = 9.00, p = .016, r = -.77, BF_{10} = 7.99.

Experiment 5b

Again, there was no significant change in RTs (n = 11) when the distractor was present (M = 585 ms, $Mdn \pm MAD$ = 554 ± 57 ms) compared to when it was absent (M = 589 ms, $Mdn \pm MAD$ = 559 ± 44 ms), W = 23.00, p = .413, r = -.30, BF_{10} = 0.38. There was a significant increase in error rates for distractor-present (M = 4.09 %, $Mdn \pm MAD$ = 3.47 ± 1.19 %) compared to distractor-absent trials (M = 2.64 %, $Mdn \pm MAD$ = 1.76 ± 1.44 %), W = 62.00, p = .007, r = .88, BF_{10} = 20.56. Again, as expected BIS showed a significant change in performance between distractor-present (*M* = -0.31, *Mdn* ± *MAD* = -0.13 ± 1.32) and distractor-absent trials (*M* = 0.31, *Mdn* ± *MAD* = 0.94 ± 0.76), *W* = 3.00, *p* = .005, *r* = -.91, BF_{10} = 35.04.

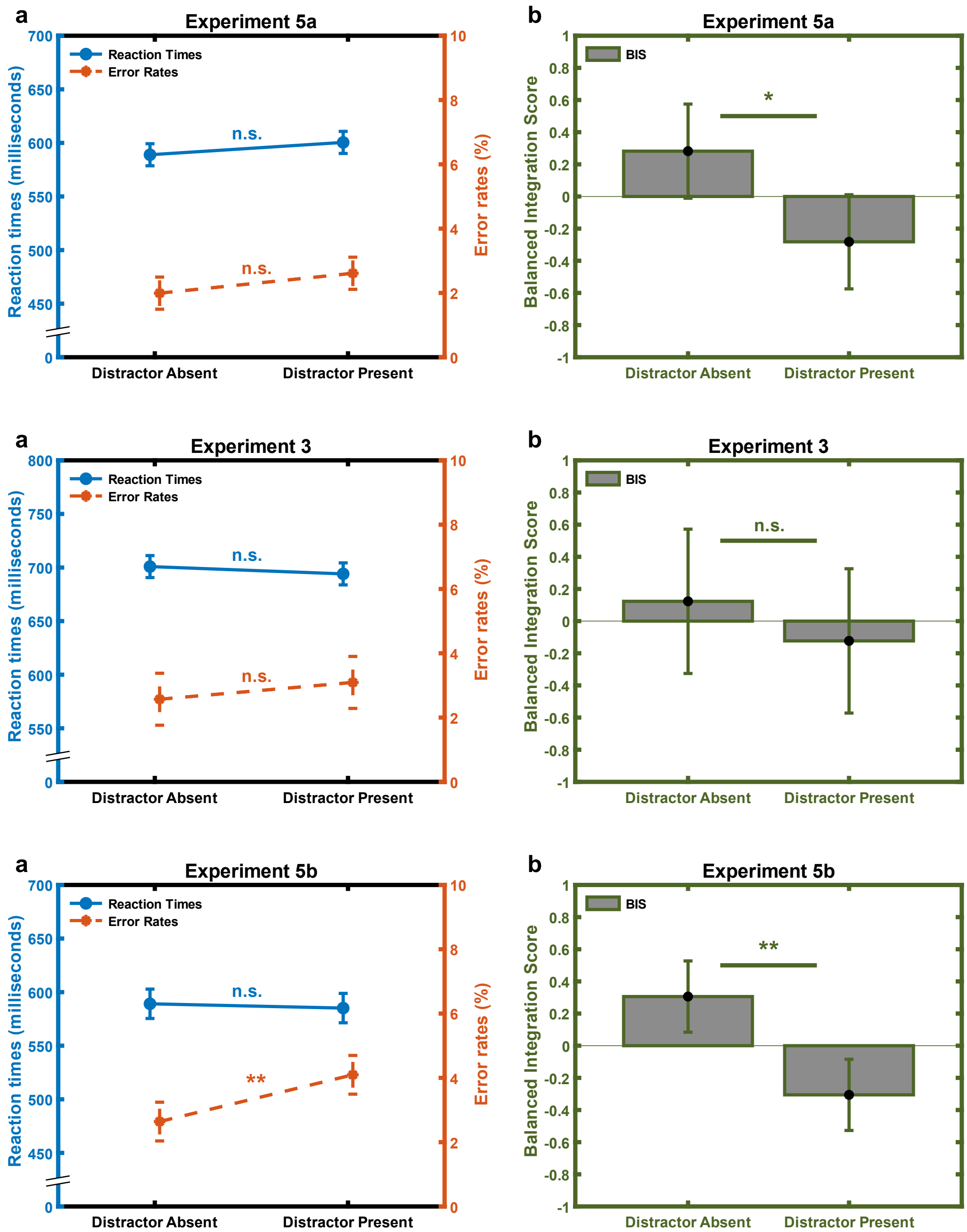


Figure S1. Results for Exp. 3, 5a and 5b without the participants who took part in a previous experiment.

Supplement – II

Here, we report the results of the test for Normality (Shapiro-Wilk test) for all comparisons. Consequently, we report the results for Student *t* tests for the comparisons which do not violate the assumption of Normality. All these *t* tests confirm the statistical decision reached from the Wilcoxon tests reported in the manuscript. Note: significant result in the Shapiro-Wilk test indicates deviation from Normality.

Experiment 1

The difference between the distractor-present and the distractor-absent condition violates the assumption of Normality as tested using the Shapiro-Wilk test for the RTs (W = .82, p = .006), error rates (W = .75, p < .001) and the BIS (W = .76, p < .001).

Experiment 2

The difference in RTs between the distractor-present and the distractor absent condition violate the assumption of Normality (W = .85, p = .013). The difference in error rates does not violate the assumption of Normality (W = .90, p = .093), and the *t* test does not show any significant change in the error rates, $t(15)$ = -0.54, p = .598, d_z = -0.14. The difference in BIS does not violate the assumption of Normality (W = .97, p = .874), and the *t* test does not show any significant change in BIS, $t(15)$ = -0.07, p = .706, d_z = -0.02.

Experiment 3

The difference in RTs does not violate the assumption of Normality (W = .90, p = .086), the *t* test does not show any significant change in the RTs, $t(15)$ = -1.62, p = .126, d_z = -0.405. The difference in error rates similarly does not violate the assumption of Normality (W = .91, p = .116), and the *t* test does not show any significant change in the error rates, $t(15)$ = 0.73, p = .479, d_z = 0.18. Similarly, the difference in BIS does not violate the assumption of Normality (W = .93, p = .208), and the *t* test does not show any significant change in BIS, t = -0.49, p = .628, d_z = -0.12.

Experiment 4

The difference in RTs between the distractor-present and the distractor-absent condition violates the assumption of Normality (W = .75, p < .001). And so do the differences in error rates (W = .86, p = .021) and the BIS (W = .87, p = .025).

Experiment 5a

The difference in RTs violates the assumption of Normality (W = .84, p = .011). The difference in error rates does not violate the assumption of Normality (W = .97, p = .830), and the t test shows a significant increase in the error rates for the distractor-present condition, $t(15)$ = 2.29, p = .037, d_z = 0.57. The difference in BIS does not violate the assumption of Normality (W = .97, p = .810), and the t test shows a significant decrease in BIS for the distractor-present condition, $t(15)$ = -3.09, p = .007, d_z = -0.77.

Experiment 5b

The difference in RTs does not violate the assumption of Normality (W = .93, p = .266), and the t test does not show any significant change in the RTs, $t(15)$ = -0.73, p = .476, d_z = -0.18. The difference in error rates similarly does not violate the assumption of Normality (W = .93, p = .224), the t test shows a significant increase in the error rates for the distractor-present condition, $t(15)$ = 3.19, p = .006, d_z = 0.80. The difference in BIS does not violate the assumption of Normality (W = .91, p = .120), and the t test shows a significant decrease in BIS for the distractor-present condition, $t(15)$ = -3.37, p = .004, d_z = -0.84.

Supplement – III

Supplementary Experiment – with visual distractor

A critique might come up with a trivial explanation for not observing distractor interference in Experiments 1 to 4, namely that our task design is not suited to measure the consequences of interference, or our analyses are not sensitive enough to detect it. There have been reports of contrasting evidence about the interference effect in visual-search tasks comparable to ours (reporting whether a target is present or absent; Chan & Hayward, 2009; Kumada, 1999). Zehetleitner et al. (2009) pointed out that the absence of distractor effects in these studies was probably due to the distractors being highly predictive, occurring on 100% of trials in distractor-present blocks. Indeed, intermixing distractor-present and distractor-absent trials (with 50% distractor prevalence) as in the current study resulted in reliable interference effects in Zehetleitner et al. (2009). Also note that our Experiment 5 demonstrates that our task is sensitive to measure interference effects. Nevertheless to resolve any doubt that the specific visual-search task employed here might be unsuited to resolve any interference effects, we ran a control experiment replacing the auditory distractor with a commonly used visual distractor: a salient red stimulus (Jannati et al., 2013; Liesefeld et al., 2022; Theeuwes, 1991, 1992).

Methods

Sixteen participants performed the experiment (median age: 25.5 years; range: 18 - 33 years; 8 females). The search displays and the general design of the experiment was similar to Experiment 1, except that the auditory distractor was replaced by a visual distractor, by drawing one of the nontarget bars in red (see Fig. S2a). No auditory stimuli were presented in this experiment. To avoid any biases regarding the processing of colored stimuli, the feedback for incorrect trials was a cross sign (X) presented for 1000 ms signifying 'wrong', whereas the feedback for the delayed response was a drawing of an hourglass presented for 1000 ms, instead of the fixation cross turning red and blue (see Exp. 1), respectively.

Results

As expected, the salient visual distractor significantly increased RTs (by 14 ms; see Fig. S2) for distractor-present (M = 593 ms, $Mdn \pm MAD$ = 596 ± 75 ms) compared to distractor-absent trials (M = 579 ms, $Mdn \pm MAD$ = 580 ± 68 ms), W = 136.00, p < .001, r = 1.00, BF_{10} = 2500.21. Error rates showed no significant difference between distractor-present (M = 2.91 %, $Mdn \pm MAD$ = 2.16 ± 1.05 %) and distractor-absent trials (M = 2.43 %, $Mdn \pm MAD$ = 2.31 ± 1.04 %), W = 84.00, p = .182, r = .40, BF_{10} = 0.73. Crucially, BIS showed a significant decrease in performance when the distractor was present (M = -0.21, $Mdn \pm MAD$ = -0.15 ± 1.01) compared to when the distractor was absent (M = 0.21, $Mdn \pm MAD$ = 0.13 ± .0.82), W = 25.00, p = .025, r = -.63, BF_{10} = 5.49.

Discussion

By demonstrating reliable distractor-presence effects on visual-search performance, this experiment confirms that the specific visual-search task used here is suitable to measure distractor interference. Hence, we can exclude the possibility that the absence of interference by an auditory distractor in Exps. 1 to 4 are due to any idiosyncrasies of the incorporated visual-search task. Furthermore, this shows that our sample size and analysis protocol is suitable to detect even relatively small interference effects in the order of 14 ms (see Liesefeld et al., 2019, for comparison). Any interference below this low magnitude would likely have little practical relevance and would not be perceived subjectively as distracting.

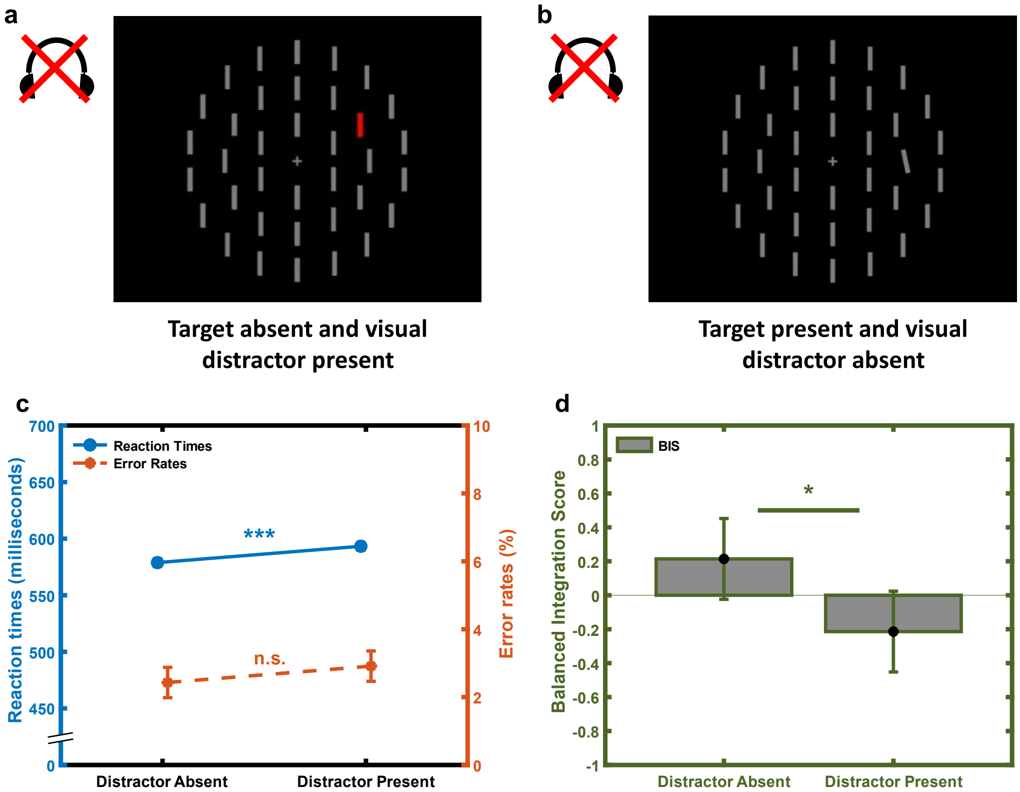


Figure S2. Example search displays **(a, b)** and results **(c, d)** for experiment with visual distractor.

2.2 Auditory distractors are processed but do not interfere with visual search of any difficulty when sound is irrelevant.

This manuscript has been published as:

Mandal, A., Röer, J. P., & Liesefeld, H. R. (2024). Auditory distractors are processed but do not interfere with visual search of any difficulty when sound is irrelevant. *Visual Cognition*, 1–17. https://doi.org/10.1080/13506285.2024.2397825

Author contributions

Ananya Mandal: conceptualization (equal), data curation, formal analysis, investigation (equal), methodology (equal), project administration, software, visualization, writing – original draft (supporting), and writing – review and editing (equal).

Jan Philipp Röer: conceptualization (equal), investigation (equal), methodology (equal), supervision (supporting), and writing–review and editing (supporting).

Heinrich R. Liesefeld: conceptualization (equal), formal analysis (supporting), methodology (equal), project administration (supporting), funding acquisition, resources, supervision, validation, visualization (supporting), writing – original draft, and writing – review and editing (equal).

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Auditory distractors are processed but do not interfere with visual search of any difficulty when sound is irrelevant

Ananya Mandal [a,b], Jan Philipp Röer [c] and Heinrich R. Liesefeld [b,d]

[a]General and Experimental Psychology, Ludwig-Maximilians-Universität Munich (LMU), Munich, Germany; [b]Graduate School for Systemic Neurosciences, LMU, Munich, Germany; [c]Department of Psychology and Psychotherapy, Witten/Herdecke University, Witten, Germany; [d]Department of Psychology, University of Bremen, Bremen, Germany

ABSTRACT
People often report being distracted during their visual tasks, such as monitoring the road ahead, by task-irrelevant sounds, for example, a baby crying on the backseat. When we first tried to study the effect of auditory distraction on visual-search performance in the laboratory using the highly sensitive additional-singleton paradigm – to our surprise – several types of auditory distractors reliably failed to cause any substantial interference. We explained these findings with a powerful attentional filtering mechanism that can shield visual search from interference by irrelevant sounds. In the present study, we examine conditions under which this mechanism might break down as suggested by insights from research on auditory distraction. It has been shown that whether an auditory distractor causes interference often hinges on the difficulty of the employed task. However, across three levels of search-task difficulty, we here reliably replicate the pattern we had observed before: At each difficulty level of Experiment 1, an unpredictably presented auditory distractor was processed to some extent (as indicated by an effect on the speed-accuracy tradeoff) but did not interfere with overall search performance. The same distractor reliably impeded search performance when the attentional shielding mechanism was experimentally disabled by a secondary task in Experiment 2.

ARTICLE HISTORY
Received 26 May 2023
Accepted 21 August 2024

KEYWORDS
Visual search; auditory distraction; selective attention; perceptual load; task difficulty

Have you ever been annoyed by an occasional sound? Especially when people try to focus on an attention-demanding task such as driving under bad viewing conditions or playing a challenging video game, they often request silence to avoid auditory distraction. Trying to examine this phenomenon in the laboratory by adapting a standard visual-distraction paradigm to study auditory distraction, we recently found – to our surprise – that an *auditory distractor* does not affect how well participants perform a visual-search task of medium difficulty. In particular, in a modification of the classical *additional-singleton paradigm* (Liesefeld et al., 2019; Theeuwes, 1991; for examples, see Figure 1(a) and (b)), we asked participants to indicate whether a 12°-tilted target bar was present in an array of irrelevant vertical bars and on half of the trials these search displays were accompanied by a salient-but-irrelevant sound (see Figure 1(c)). In this paradigm, interference is measured as *distractor-presence costs* in terms of reduced visual-search performance on distractor-present compared to distractor-absent trials (see Liesefeld et al., 2024, for definitions of terms).

Across multiple experiments using different distractor sounds and techniques to increase the disruptive potential of these sounds, we did not observe any distractor-presence costs (Mandal et al., 2024). Even rare auditory oddballs, which are considered to be highly distracting (Escera et al., 1998, 2002), did not interfere with visual search. They did not distract even though they occurred on only 25% of trials (cf. Müller et al., 2009) and even when they were presented 300 ms before the search display (i.e., being endowed with a temporal processing advantage). This consistent failure to observe distraction was surprising to us, because an extensive research tradition

CONTACT Ananya Mandal Ananya.Mandal@psy.lmu.de General and Experimental Psychology, Ludwig-Maximilians-Universität Munich, Leopoldstr. 13, 80802 Munich, Germany; Heinrich R. Liesefeld Heinrich.Liesefeld@uni-bremen.de Department of Psychology, University of Bremen, Hochschulring 18, 28359 Bremen, Germany

Supplemental data for this article can be accessed online at https://doi.org/10.1080/13506285.2024.2397825.

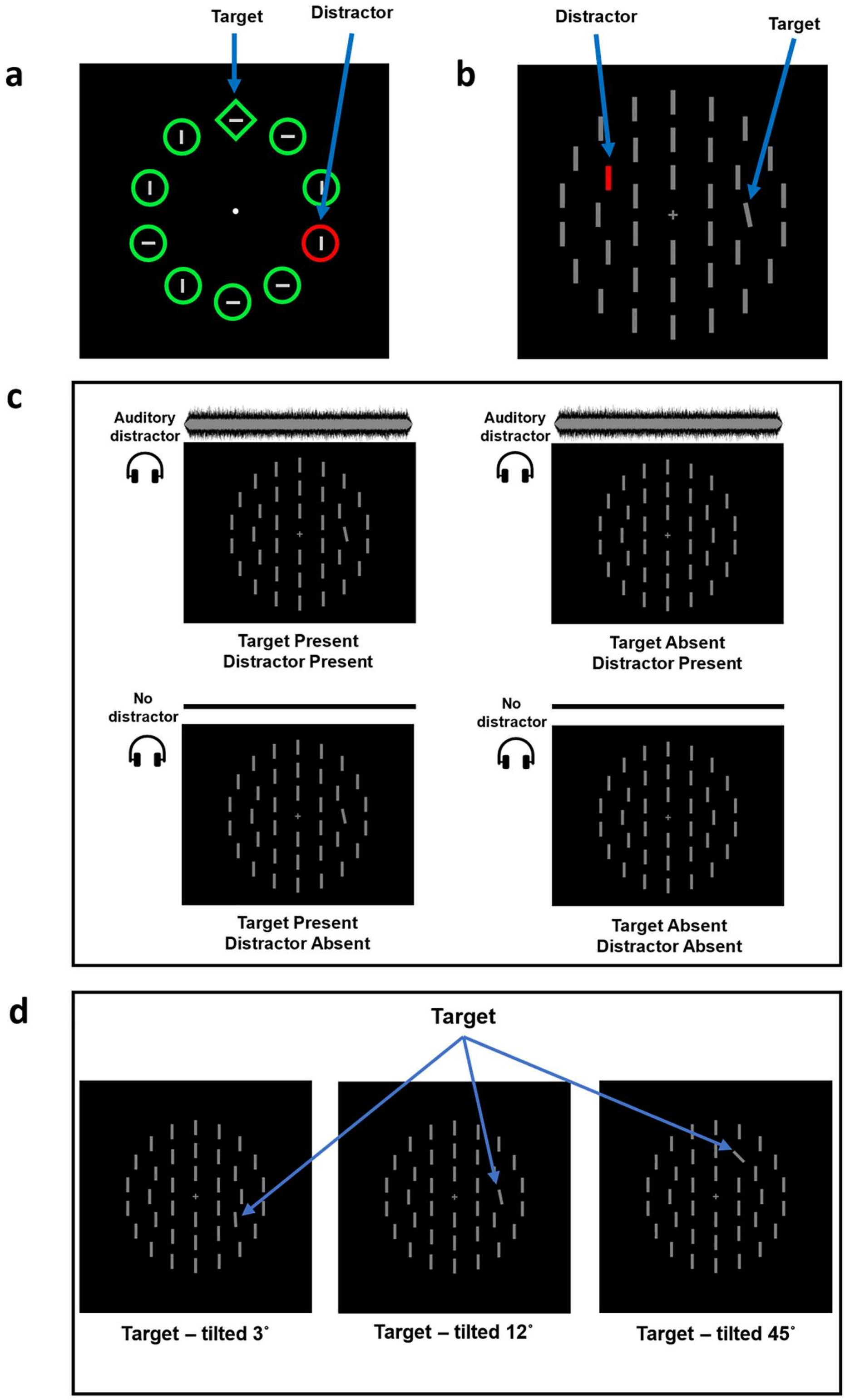


Figure 1. The additional-singleton paradigm. (a) In the standard additional-singleton paradigm (e.g., McDonald et al., 2013) participants look for a shape singleton (the diamond) and report the orientation of the bar inside. (b) In the adapted version employed by Mandal et al. (2024) participants indicate whether a target singleton (the 12°-tilted bar) is present or absent among a dense array of *nontargets* (vertical bars). In (a) and (b) an additional salient visual singleton (the red distractor) causes interference as indicated by distractor-presence costs. (c) All four possible combinations of target and auditory-distractor presence employed in Mandal et al. (2024) and in the current study. (d) The present study manipulates search difficulty across blocks by employing targets that are tilted by either 3°, 12°, or 45°. Please visit https://figshare.com/s/469e9f33b943e892e734 for example trials from each block.

on visual distraction had shown that all kinds of salient visual distractors reliably produce such costs in the additional-singleton paradigm (for reviews, see Chelazzi et al., 2019; Liesefeld & Müller, 2019; Theeuwes, 2010). In fact, visual distractors interfere with visual search even if conditions are ideal to prepare against these distractors, namely if they occur on the majority of trials (Geyer et al., 2008; Moher et al., 2011) and stand out in a different visual feature dimension than the target (Liesefeld et al., 2019, 2022; Liesefeld & Müller, 2019). Only when Mandal et al. (2024) made the auditory modality task-relevant by introducing a secondary auditory task, they observed interference by an auditory distractor on the primary visual-search task. We interpreted this pattern of results as indicating that, as long as the auditory modality is completely irrelevant[1], people are able to block it from interfering with visual search, much as if they had shut their "attentional earlid."

The concept of the "attentional earlid" (Mandal et al., 2024) serves as a metaphor for an attentional shielding mechanism within the auditory system. This envisioned mechanism helps protect against unwanted auditory interference and focus on non-auditory tasks when needed, much like how eyelids can physically block out most of the visual input, allowing us to focus on non-visual tasks. The "attentional earlid" is speculated to function as an attentional barrier and helps in preventing irrelevant auditory stimuli from interfering with non-auditory tasks. This attentional barrier can either be kept opened or closed (much like the eyelids) depending on current behavioural goals and task requirements (e.g., see Experiment 5 of Mandal et al., 2024 and Experiment 2 in present study) – which either allows potentially distracting auditory stimuli to pass through, or not.

Based on the unexpected results from Mandal et al. (2024), we became interested in what it takes to break through the attentional shielding mechanism, or under which conditions it lets information through despite the auditory modality being irrelevant. The extensive literature on auditory distraction might provide clues as to why our observers were able to resist auditory distraction during visual search, although this research tradition typically employs quite different tasks, with serial-recall of (visually or auditorily presented) verbal material being most common (e.g., Alikadic & Röer, 2022; Kaiser et al., 2021; Kattner & Bryce, 2022; Marsh et al., 2020; Wöstmann & Obleser, 2016).

According to the *duplex-mechanism account of auditory distraction* (Hughes, 2014), distractors hold the potential to impede task performance when they share certain features with task-relevant stimuli (*interference-by-process*) or when they possess attention-diverting power. Distractors possess this power either by virtue of their content (e.g., your own name; *specific attentional capture*) or because they violate expectations (e.g., an oddball in a sequence of standards; *aspecific attentional capture*). In classic visual-distraction paradigms, the distractor-defining feature is rather abstract and, most importantly, differs from the target-defining feature. Therefore, according to Hughes' (2014) classification, this paradigm does not typically induce interference-by-process or, respectively, specific attentional capture. Aiming to bridge the gap between the visual-distraction and the auditory-distraction literature, we decided to employ a sine-wave tone with no meaning attached to it as an auditory distractor. Thus, what Mandal et al. (2024) have observed with the auditory modality being relevant was "aspecific attentional capture," which can be avoided by greater task engagement according to the duplex-mechanism account. It seems possible that task engagement in Mandal et al. (2024) was relatively high and that chances to observe attentional capture would increase as the visual-search task becomes easier. The search task employed by Mandal et al. was already fairly easy (searching a 12°-tilted target among vertical bars, see Figure 1(a)), but further increasing the target-nontarget contrast would make it even easier (Liesefeld et al., 2016), and would thus potentially sufficiently reduce task engagement to observe distraction without making the auditory modality relevant by a secondary auditory task.

Similarly, *Perceptual-Load Theory* (Lavie, 1995, 2005; Lavie & Tsal, 1994) suggests that distraction occurs for efficient searches (*low perceptual load*), but not for inefficient searches (*high perceptual load*). Under high perceptual load, attentional resource is depleted so that none is left for processing the distractor. In fact, even an entirely irrelevant visual distractor presented outside the search array causes interference under low load (Forster & Lavie, 2008a, 2008b). Perceptual load in one modality can also affect processing of stimuli in a different modality. Importantly, interference due to auditory distractors is reduced for visual tasks with high perceptual load compared to tasks with low perceptual load (Tellinghuisen & Nowak, 2003). Similarly, Raveh and Lavie (2015) found that high perceptual load in the visual modality can reduce detection sensitivity for auditory stimuli (*inattentional deafness*). As demonstrated by Liesefeld et al. (2016), search for a tilted target is efficient (as indicated by the absence of set-size effects) for target contrasts above 6-10°. Still, maybe search for a 12°-contrast target in Mandal et al. (2024) was not yet easy enough to count as low in perceptual load. Thus, just like the duplex-mechanism account of auditory distraction seems to suggest, Perceptual-Load Theory might also predict that auditory distraction can be observed in an even easier visual-search task.

Then again, most studies in the context of Perceptual-Load Theory do not actually examine whether a distractor impedes overall (search) performance. That is, these studies do not typically compare distractor-present against distractor-absent trials, which is the main comparison of interest in the additional-singleton task (see above; see Forster & Lavie, 2008b, for an exception and discussion of this point). Instead of examining whether the distractor *causes interference*, this research tradition typically focuses on empirical markers indicating that it *is processed*. Based on the classical Eriksen Flanker Task (Eriksen & Eriksen, 1974), this is typically tested by comparing distractors that are congruent vs. incongruent with the respective reported target feature (for a review, see Lavie & Tsal, 1994). If a congruency effect is observed, this indicates that the distractor was processed. It is not possible to examine congruency effects in the Mandal et al. (2024) task, because we purposefully avoided to employ modality-independent features (e.g., meaning such as letter identity, which might yield interference-by-process, see Hughes, 2014) that could be shared between the visual-search target and the auditory distractor.

However, an aspect of our findings might provide a leverage point to examine whether the distractor was processed. In particular, even though actual search performance was unaffected by the distractor sound in Mandal et al. (2024), the sound did affect participants' speed-accuracy tradeoff (SAT). That is, when the auditory distractor was present, responses became faster and error rates increased (the typical SAT-effect pattern, Heitz, 2014) compared to when the distractor was absent. A combined performance measure controlling for SAT effects (the *Balanced Integration Score, BIS*; Liesefeld & Janczyk, 2019, 2023) was unaffected by distractor presence, indicating that the RT and error-rate pattern is best explained as a SAT effect. If we take this SAT effect as a marker of distractor processing, this would resolve the conundrum from the perspective of Perceptual Load Theory: The efficient search for a 12°-contrast target produces low perceptual load and therefore the distractor was processed. To explain from this perspective why Mandal et al. (2024) did not observe distractor-presence costs, one only needs to additionally assume that processing a distractor does not inevitably yield interference. If that is the correct explanation for our previous findings, we would expect the SAT effect to vanish when we increase perceptual load by employing an inefficient visual-search task. Indeed, Gao et al. (2021) employed an inefficient visual-search task (as indicated by an effect of set size) and their participants' search performance was completely unaffected by an auditory distractor.

In contrast to all these tentative predictions from current theories on auditory distraction, Experiment 1 shows no impedance of task performance by distractor presence at various levels of search difficulty and the SAT effect was present throughout, thus fully replicating the results of Mandal et al. (2024) multiple times. This was the case for all examined target tilts of 3° (inefficient search), 12° (intermediate difficulty in the efficient-search range), and 45° (very easy, efficient search; see Figure 1(d) and Liesefeld et al., 2016). Thus, as implied by our proposal of an "attentional earlid," participants can block auditory distraction from interfering with visual search of any difficulty. Experiment 2 demonstrates that the employed distractor can indeed cause interference when the auditory modality becomes relevant due to a secondary task, which in our interpretation means that the "attentional earlid" has to remain open in order to accomplish that task.

Experiment 1

Methods

Participants

Our power calculations were based on Experiment 5 of Mandal et al. (2024), where we had observed auditory distraction in almost exactly the task design used here. The only difference is that in Experiment 1 we do not employ a secondary counting task, which Mandal et al. had used to make the auditory modality relevant (see Experiment 2 for details). We refrained from using such a secondary task, because we wanted to see whether auditory distractors can cause interference under variations of task difficulty when the auditory modality is completely irrelevant. Our power analysis is based on the size of the distractor-effect on BIS, which was our main performance measure in that previous study and here. Across the two replications reported by Mandal et al., the smaller effect size was $d_z = -0.78$. To achieve a statistical power of $1 - \beta = .80$ at $\alpha = .05$, two-tailed, the required sample size is $n = 15$ (calculated using G*power v3.1.9.7; Faul et al., 2009).

In total, 20 participants at Witten/Herdecke University took part in the experiment. They reported normal hearing and normal or corrected-to-normal vision. Written informed consent was obtained from all participants, and they received course credit or were paid for their participation. This type of study falls under the umbrella ethic approval for research involving human participants given by Witten/Herdecke University. For two participants, none or incomplete data were stored, so that our final sample consisted of $n = 18$ participants. We did not record any demographic information, but participants were recruited from the standard psychology-student population.

Stimuli and design

Stimulus presentation and response collection was controlled using PsychoPy (Peirce et al., 2019). Search displays were presented on a monitor (screen resolution: 1680 × 1050 pixels, refresh rate: 60 Hz) at a viewing distance of 70 cm. Exactly as in Mandal et al. (2024), search arrays consisted of 36 grey bars (1.35° × 0.25° of visual angle in size) presented on a black background (see Figure 1). The bars were arranged in three concentric rings (radii of 2.1°, 4.2°, and 6.3°) around a central fixation cross (0.5° × 0.5°). On half of the trials, search displays consisted of vertically oriented bars only (*target-absent condition*). On the other half of the trials, one target bar was tilted 3°, 12°, or 45° to the left (*target-present condition*).

Along with the presentation of the bars, an irrelevant auditory stimulus (distractor) was randomly presented in half of the trials (*distractor-present condition*), with no sound presented in the remaining half (*distractor-absent condition*). The distractor was a 440 Hz sine tone that was presented binaurally using beyerdynamic DT 100 headphones at 51 dB(A) SPL, measured using a Brüel & Kjær handheld type 2250 analyzer. We used an artificial ear made of Styrofoam and rubber bands to couple the handheld analyzer to the headphones. The experiment was divided into 18 blocks, with 96 trials in each block. Across blocks, we manipulated search difficulty by adapting the tilt of the target bar from 3° (difficult) to 12° (medium) to 45° (easy). Difficulty was randomized across blocks with 6 blocks per difficulty level and each block being preceded by a 10-trial non-analyzed practice phase to accustom participants to the respective target tilt. Thus, results are based on 1728 trials in total, with 144 trials in each cell of the 3 (search difficulty) × 2 (target presence) × 2 (distractor presence) design. As our analyses combine data across target presence, we achieve highly precise performance estimates based on a relatively high number of 288 trials per relevant condition (Baker et al., 2021).

Procedure

The search displays remained on the screen until participants made their response, or until the response deadline of 4000 ms. For trials with auditory distractors, the sound was played from display onset to display offset. The task was to respond whether the tilted target bar was present or absent in each trial. Responses were given on a keyboard using the left and right index fingers (counterbalanced across participants) and triggered visual display and sound offset. If the response was incorrect or delayed, the central fixation cross turned red or blue, respectively, for 1000 ms. The trials were separated by an interstimulus interval of randomly jittered duration between 800 and 1200 ms, during which only the fixation cross was presented on the screen. Participants were instructed to respond accurately and as quickly as possible.

Analyses

As we were interested in whether an auditory distractor interferes with visual search, the reported analyses focus on the effect of distractor presence. As performance measures, we extracted median correct reaction times (RTs), error rates and Balanced Integration Scores (BIS, Liesefeld & Janczyk, 2019, 2023). Pairwise comparisons between distractor-absent and distractor-present conditions were performed on these aggregate measures using Wilcoxon signed-rank test (two-tailed comparisons). Consequently, we also report the mean (*M*) and median values (*Mdn*) per condition with the respective MAD (Median Absolute Deviation: a measure of dispersion). Results graphs nevertheless show the more commonly employed means and within-participant confidence intervals (Cousineau & O'Brien, 2014). Additionally, we report Bayes factors (*BF*s), for the Bayesian Wilcoxon signed-rank test (result based on data augmentation algorithm with 5 chains of 10,000 iterations), quantifying evidence for the alternative over the null hypothesis (BF_{10}). BFs were calculated using the standard JZS prior with a scale factor of $r = \sqrt{2}/2$. To classify the strengths of evidence through the Bayes Factors, we used Jeffreys's criterion (van Doorn et al., 2021) – which states that for the alternative hypothesis, *BF*s between 1 and 3 are weak evidence, *BF*s between 3 and 10 are moderate evidence, and BFs greater than 10 are strong evidence. Following these criteria, BF_{10} between 1 and 1/3 (= 0.33) is considered weak evidence for the null hypothesis, BF_{10} between 1/3 and 1/10 is considered as moderate evidence for the null hypothesis, and BF_{10} below 1/10 (= 0.10) is considered strong evidence for the null hypothesis. All analyses were performed using JASP v0.16.4 (JASP Team, 2021) and using custom scripts in Matlab.

Results

Replicating Mandal et al. (2024) correct responses were significantly *faster* on distractor-present ($M = 1029$ ms, $Mdn \pm MAD = 1030 \pm 198$ ms) compared to distractor-absent trials ($M = 1080$ ms, $Mdn \pm MAD = 1093 \pm 230$ ms), $W = 14.00$, $p < .001$, $r = -.84$, $BF_{10} = 148.35$, for the 3°-target-contrast condition (see Figure 2(a)). That is, instead of interfering, the distractor increased response speed. The results were similar for the 12°-target-contrast condition (Figure 2(c)) with responses being significantly faster on distractor-present ($M = 599$ ms, $Mdn \pm MAD = 596 \pm 44$ ms) compared to distractor-absent trials ($M = 614$ ms, $Mdn \pm MAD = 613 \pm 44$ ms), $W = 24.00$, $p = .006$, $r = -.72$, $BF_{10} = 24.34$. Finally, the 45°-target-contrast condition (Figure 2(e)) also showed a significant decrease in RTs with the distractor-present ($M = 542$ ms, $Mdn \pm MAD = 538 \pm 56$ ms) compared to distractor-absent trials ($M = 559$ ms, $Mdn \pm MAD = 545 \pm 51$ ms), $W = 24.00$, $p = .006$, $r = -.72$, $BF_{10} = 22.26$.

As in most experiments of Mandal et al. (2024), there was a difference in error rates for the 3°-target condition when the distractor was present ($M = 19.54\%$, $Mdn \pm MAD = 16.56 \pm 9.04\%$) compared to when the distractor was absent ($M = 19.11\%$, $Mdn \pm MAD = 15.10 \pm 8.71\%$) – which would indicate some interference. However, this difference was not significant, $W = 107.00$, $p = .369$, $r = .25$, $BF_{10} = 0.30$. By contrast, the difference in error rates for the 12°-target condition between distractor-present ($M = 5.32\%$, $Mdn \pm MAD = 5.04 \pm 2.08\%$) and distractor-absent trials ($M = 3.63\%$, $Mdn \pm MAD = 2.09 \pm 1.05\%$), was significant, $W = 144.00$, $p = .012$, $r = .68$, $BF_{10} = 15.19$. Similarly, the 45°-target condition showed a significant difference in the error rates when the distractor was present ($M = 4.77\%$, $Mdn \pm MAD = 3.99 \pm 1.56\%$) compared to when the distractor was absent ($M = 3.44\%$, $Mdn \pm MAD = 2.62 \pm 1.22\%$), $W = 132.50$, $p = .043$, $r = .55$, $BF_{10} = 2.83$.

To sum up, in all difficulty (target-tilt) conditions, the presence of an auditory distractor speeds up responses at the cost of increased error rates. As already argued in Mandal et al. (2024), interpreting the non-significance of the error-rate effect in the 3° condition as evidence for the absence of an effect would amount to accepting the null hypothesis, which is not supported by frequentist tests (*p* values). However, in contrast to Mandal et al. (2024), the Bayes factor for the distractor-effect on error rates this time provided moderate evidence for the null hypothesis in the 3°-target condition ($BF_{01} = 3.28$), strong evidence for the alternative in the 12°-target condition ($BF_{10} = 15.19$) and weak evidence for the alternative in the 45°-target condition ($BF_{10} = 2.83$; although the frequentist test was still significant, $p = .043$, except if corrected for multiple comparisons). Especially the significant error-rate effect for 12° targets confirms that we had correctly interpreted the pattern of decreased response times and (non-significantly) increased error rates in Mandal et al. (2024) as a speed-accuracy tradeoff (SAT).

Thus, following Mandal et al. (2024), we combined speed and accuracy to control for speed-accuracy tradeoffs by calculating the *Balanced Integration Score* (*BIS*; Liesefeld et al., 2015; Liesefeld & Janczyk, 2019, 2023) based on aggregated data per 18 participant × 2 distractor-presence cell, separately for each task difficulty. A negative BIS indicates a lesser than average performance and a positive BIS indicates a

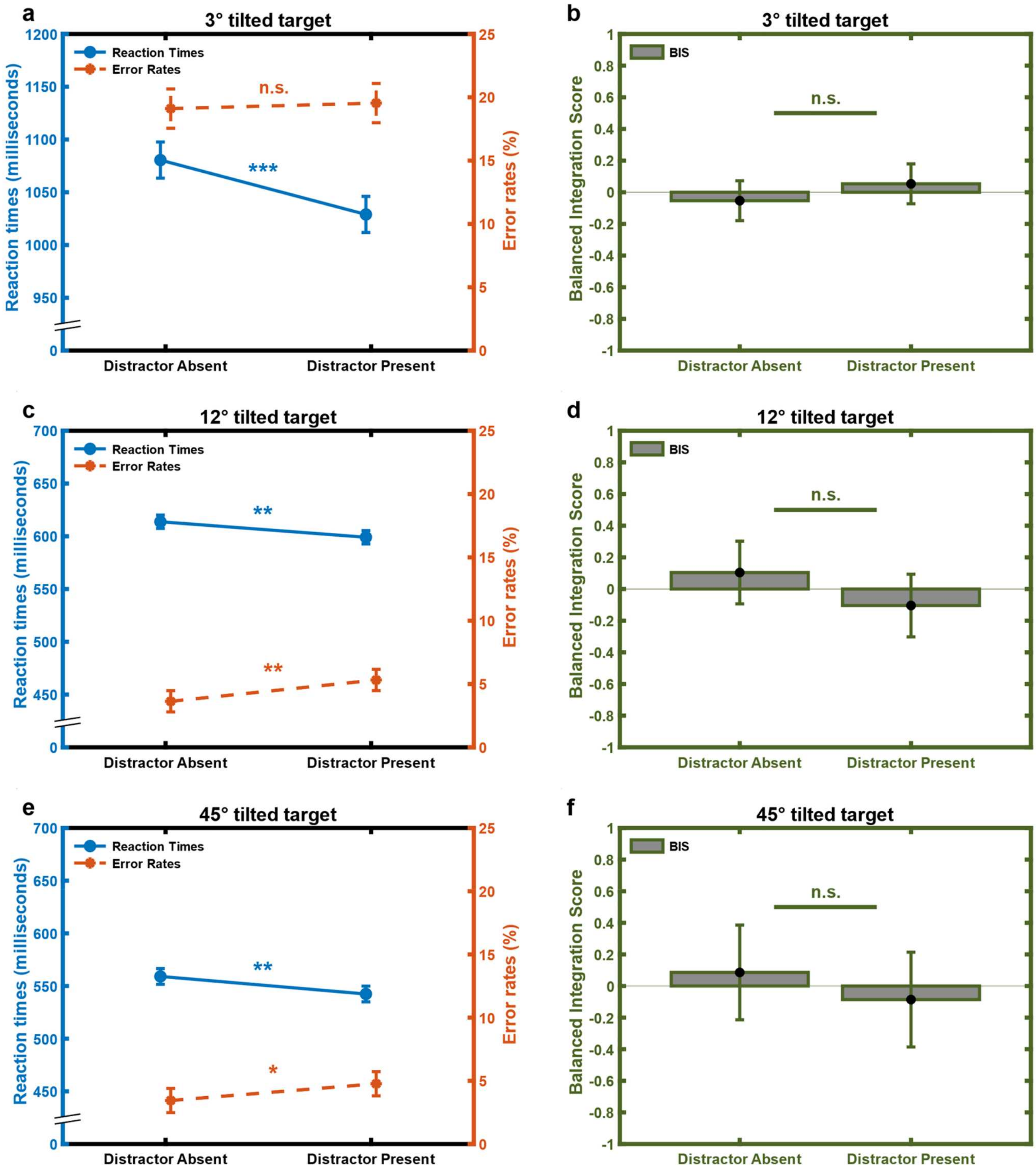


Figure 2. Visual-search performance for blocks with 3° – (a, b), 12° – (c, d), and 45° – (e, f) tilted targets. (a, c, e) Average median reaction times and error rates. (b, d, f) Balanced Integration Score (BIS). Error bars represent 95% within-participant confidence intervals. The lines connecting the distractor absent and distractor present conditions (a, c, d) are not supposed to imply that there are an infinite number of (potential) measurements in between the two conditions. Rather, for easy visual comparisons of the patterns across panels and figures, here and in Figure 3 the lines' direction of tilt can be used as a visual guide to distinguish which condition has a higher/lower value for the RTs and error rates. n.s. $p > .05$, *$p < .05$, **$p < .01$, ***$p < .001$.

greater than average performance. Analyses on BIS (Figure 2(b, d, f)) did not provide evidence for or against any distractor effect for the 3° and 12° target-contrast conditions, but weak evidence for the absence of an effect for the 45° target-contrast condition (Table 1).

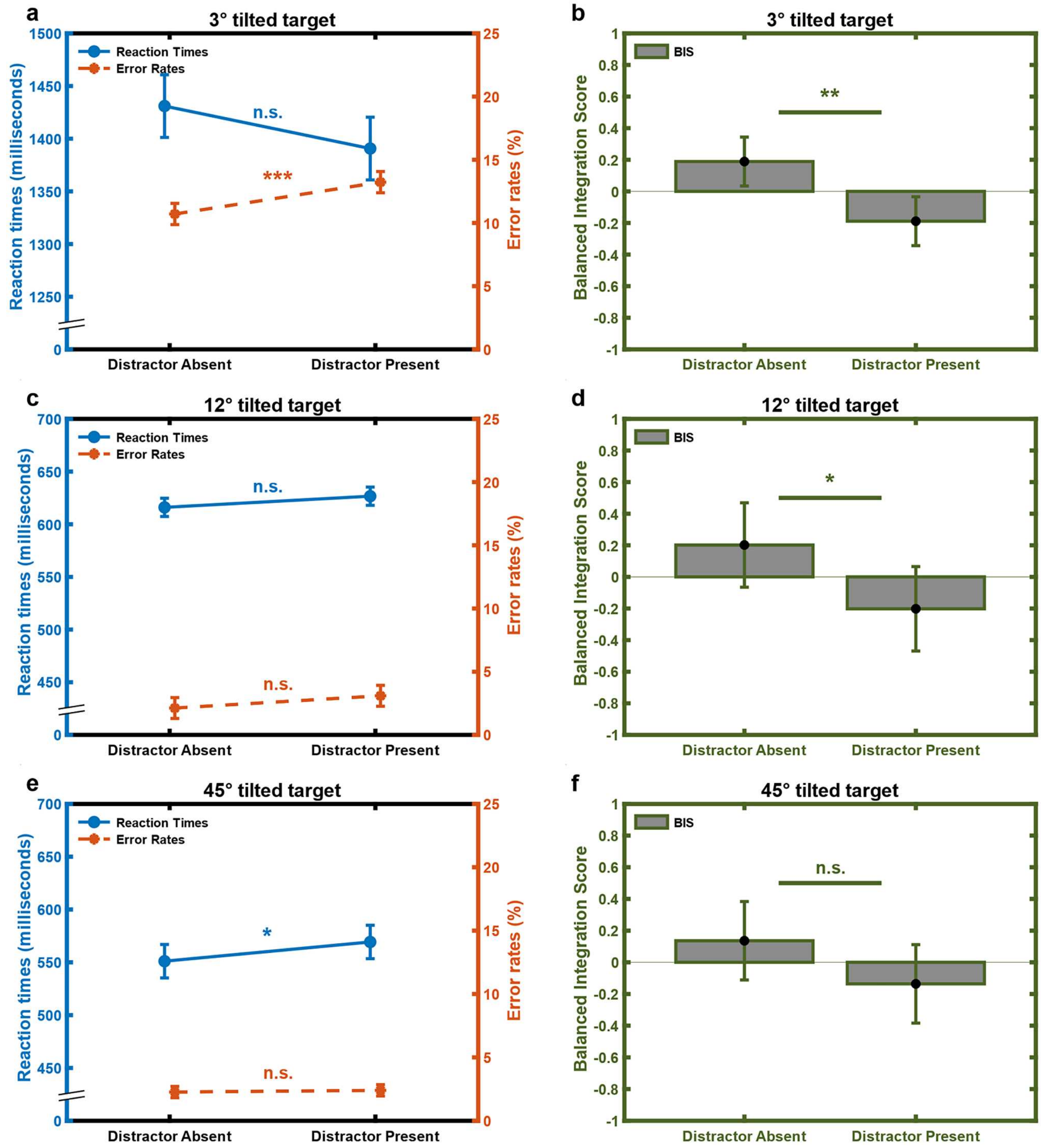


Figure 3. Visual-search performance for blocks with 3° – (a, b), 12° – (c, d), and 45° – (e, f) tilted targets in Experiment 2. (a, c, e) Average median reaction times and error rates. (b, d, f) Balanced Integration Score (BIS). Error bars represent 95% within-participant confidence intervals. n.s. $p > .05$, *$p < .05$, **$p < .01$, ***$p < .001$.

Table 1. Effects of distractor presence on combined performance (BIS) in the three target-contrast conditions in Experiment 1

	Distractor-present			Distractor-absent						
Contrast	*M*	*Mdn*	*MAD*	*M*	*Mdn*	*MAD*	*W*	*p*	*r*	BF_{10}
3°	0.05	0.27	0.69	−0.05	0.03	0.76	122.00	.119	.43	0.80
12°	−0.10	0.29	0.99	0.10	0.44	0.90	52.00	.154	-.39	0.67
45°	−0.09	0.03	0.69	0.09	0.07	0.89	72.00	.580	-.16	0.30

Experiment 2

A trivial explanation for the absence of interference effects in Experiment 1 would be that the distractor does not have the potential to interfere. The distractor might just not be potent enough or interference might be too short-lived to affect visual search. Consequently, Experiment 2 was designed to demonstrate that the employed distractor can in principle cause interference in our visual-search task. This was achieved by using the approach of Mandal et al.'s (2024) Experiment 5. In particular, we required the participants to process auditory stimuli by adding a secondary task involving the counting of a sound that was clearly distinguishable from the distractor sound. In terms of our "attentional earlid" interpretation, this means that task demands caused the "attentional earlid" to remain open. We kept the search-difficulty manipulation to explore whether this would modulate the anticipated interference effect.

Methods

In Experiment 2 we made the auditory modality generally relevant, by employing the same procedure as used in Experiment 5 of Mandal et al. (2024). Consequently, along with the visual stimuli and the irrelevant auditory distractor (440 Hz tone; presented in half of the trials as in Experiment 1), a counting sound (520 Hz sine tone) was presented in a few additional trials (ranging from 3 to 6 in number). Participants were instructed to count the number of times the counting sound was presented in each block and report this number at the end of the block. The counting sound (520 Hz sine tone) is easily distinguishable from the auditory distractor (440 Hz sine tone) and the additional trials (with the counting sound) never contained the auditory distractor and did not contribute to our visual-search performance measures (RT, error rate, and BIS). Everything else was kept similar to Experiment 1, including the task-difficulty manipulation. The sound stimuli (auditory distractor and counting sound) were presented via headphones (Sennheiser HD 600, Sennheiser, Wedemark, Germany) at ~50–52 dB(A) SPL measured at the eardrum, using the miniDSP Earphone Audio Response System device (miniDSP, Hong Kong, China). Each block in the Experiment consisted of 96 + x trials, with x being the number of trials with the counting sound and ranging from 3 to 6. These additional trials with the counting sound were excluded from the main analysis. Reporting the wrong counting number would indicate that the participants did not pay attention to the auditory modality and we consequently remove these blocks from the analysis. If more than 2 blocks out of the 6 blocks for any search-task-difficulty level had to be excluded based on this criterion, the whole data set was removed from the analysis.

In total, 20 participants at Ludwig-Maximilians-Universität Munich (LMU) took part in Experiment 2. They reported normal hearing and normal or corrected-to-normal vision. Written informed consent was obtained from all participants, and they received course credit or were paid for their participation. Procedures were approved by the ethics committee of the Department of Psychology and Pedagogics at LMU. Three participants had to be excluded as their counting responses (at the end of each block) were incorrect in more than 2 out of 6 blocks in at least one search-task-difficulty level. Thus, our final sample consisted of $n = 17$ participants. The experiment was preregistered, and the preregistration file is available at: https://osf.io/gbtx8.

Results

We did not observe a significant difference in RTs in the 3°-target condition (see Figure 3(a)) for distractor-present ($M = 1391$ ms, $Mdn \pm MAD = 1381 \pm 317$ ms) compared to distractor-absent trials ($M = 1431$ ms, $Mdn \pm MAD = 1381 \pm 234$ ms), $W = 39.00$, $p = .080$, $r = -.49$, $BF_{10} = 1.63$. The results were similar for the 12°-target condition (Figure 3(c)), with no significant differences for distractor-present ($M = 627$ ms, $Mdn \pm MAD = 606 \pm 50$ ms) compared to distractor-absent trials ($M = 616$ ms, $Mdn \pm MAD = 589 \pm 50$ ms), $W = 117.00$, $p = .057$, $r = .53$, $BF_{10} = 2.26$. Finally, for the 45°-target condition (Figure 3(e)), we observed a significant increase in RT for distractor-present ($M = 569$ ms, $Mdn \pm MAD = 555 \pm 17$ ms) compared to distractor-absent trials ($M = 551$ ms, $Mdn \pm MAD = 555 \pm 49$ ms), $W = 128.00$, $p = .013$, $r = .67$, $BF_{10} = 11.52$.

Contrary to Experiment 1, there was a significant difference in error rates for the 3°-target condition when the distractor was present ($M = 13.23\%$, $Mdn \pm MAD = 14.93 \pm 3.47\%$) compared to when it was absent ($M = 10.71\%$, $Mdn \pm MAD = 11.81 \pm 2.78\%$), $W = 145.00$, $p < .001$, $r = .89$, $BF_{10} = 188.69$. By contrast, in the 12°-target condition there was no significant difference in the error rate between distractor-present ($M = 3.09\%$, $Mdn \pm MAD = 1.25 \pm 0.90\%$) and distractor-absent trials ($M = 2.13\%$, $Mdn \pm MAD = 1.56 \pm 0.87\%$), $W = 91.00$, $p = .245$, $r = .34$, $BF_{10} = 0.58$; in the 45°-target condition no significant difference was observed either between distractor-present ($M = 2.40\%$, $Mdn \pm MAD = 2.08 \pm 1.56\%$) and distractor-absent trials ($M = 2.27\%$, $Mdn \pm MAD = 1.56 \pm 1.15\%$), $W = 46.00$, $p = 1.000$, $r = .01$, $BF_{10} = 0.26$.

Following the same procedure as used in Experiment 1 we calculated the BIS in order to control for variation in speed-accuracy tradeoff. Here, performance significantly decreased in the presence of the auditory distractor ($M = -0.19$, $Mdn \pm MAD = -0.29 \pm 0.94$) compared to the distractor's absence ($M = 0.19$, $Mdn \pm MAD = 0.21 \pm 0.87$), $W = 18.00$, $p = .004$, $r = -.77$, $BF_{10} = 37.71$ in the 3°-target condition (Figure 3(b)). This distraction effect in the overall search performance was also observed in the 12°-target condition (Figure 3(d), also see Experiment 5 of Mandal et al., 2024) where the performance was significantly lower for distractor-present ($M = -0.20$, $Mdn \pm MAD = -0.12 \pm 0.53$) compared to distractor-absent trials ($M = 0.20$, $Mdn \pm MAD = 0.19 \pm 0.54$), $W = 29.00$, $p = .023$, $r = -.62$, $BF_{10} = 4.92$. However, for the 45°-target condition (Figure 3(f)) the BIS difference between distractor-present ($M = -0.14$, $Mdn \pm MAD = -0.11 \pm 0.42$) and distractor-absent trials ($M = 0.14$, $Mdn \pm MAD = 0.01 \pm 0.68$) was in the anticipated direction but not significant, $W = 49.00$, $p = .207$, $r = -.36$, $BF_{10} = 0.72$. Thus overall, the results indicate that our distractor has the potential to distract as demonstrated in two out of the three examined conditions.

To directly test whether search difficulty modulates the distractor effect on the BIS we ran pair-wise comparisons on the differences between the distractor-present and distractor-absent condition for each level of difficulty reported above. There was no significant difference in distractor interference between any pair of difficulty conditions: 3° target ($M = -0.38$, $Mdn \pm MAD = -0.28 \pm 0.26$) vs. 12° target ($M = -0.40$, $Mdn \pm MAD = -0.15 \pm 0.20$), $W = 63.00$, $p = .548$, $r = -0.18$, $BF_{10} = 0.27$; 12° target vs. 45° target ($M = -0.27$, $Mdn \pm MAD = -0.19 \pm 0.41$), $W = 70.00$, $p = .782$, $r = -0.09$, $BF_{10} = 0.29$; 3° target vs. 45°-target, $W = 68.00$, $p = .712$, $r = -0.11$, $BF_{10} = 0.28$.

Discussion

After visual search of intermediate difficulty had turned out to be surprisingly robust against auditory distraction (Mandal et al., 2024), we here made another attempt to induce such distraction by manipulating task difficulty. In previous studies with different focal tasks, task-difficulty was an important determinant of cross-modal auditory distraction (Halin et al., 2014; Hughes et al., 2013; Sörqvist et al., 2012; Sörqvist & Marsh, 2015). Consequently, the aim of the present study was to see whether auditory distractors might interfere with visual search at other levels of search difficulty. This was clearly not the case in Experiment 1: Counterintuitively, the presence of an auditory distractor, *improved* visual-search performance in terms of reaction times. A beneficial effect of task-irrelevant information seems somewhat unusual at first, but similar effects have been reported in the literature. Visual localization of a target stimulus is faster when the auditory distractor is semantically consistent (e.g., a meow sound when searching for the image of a cat) compared to a quiet control condition (Iordanescu et al., 2008; see also Iordanescu et al., 2010) and a non-spatial sound can increase search efficiency for an otherwise hard-to-find visual-search target if sound and target flicker synchronously (Van der Burg et al., 2008). However, just as in the comparable Experiment 1 of Mandal et al. (2024) this improvement was in the present study (more or less) balanced by a corresponding worsening in terms of error rates – a typical effect on the speed-accuracy tradeoff (SAT; Heitz, 2014). Even though there was some variation and statistical uncertainty across difficulty levels regarding the effect on error rates and on an SAT-insensitive combined performance measure (BIS, Liesefeld & Janczyk, 2019, 2023), the overall results pattern clearly does not provide any indication for interference by the auditory distractor.

When we started this project, we were somewhat confident that we would observe auditory distraction, because the employed additional-singleton task has proven highly sensitive for measuring distraction in visual search (for reviews, see, e.g., Anderson et al., 2021; Chelazzi et al., 2019; Theeuwes et al., 2022). Our task would mimic everyday situations in which people often feel distracted by sounds, such as when a parent working from home tries to focus on their computer screen, while the children are playing in the adjacent room. To our surprise, the only condition under which we could observe auditory distraction in visual search until now, is when we made the auditory modality relevant by a secondary task. This dependence of interference by an auditory distractor in visual search was replicated in the present study. In particular, Experiment 2 showed that when an auditory secondary task renders the auditory modality relevant in general, the same auditory distractor causes interference. Mandal et al. (2024) have interpreted this pattern as indication that a highly effective attentional mechanism exists that is able to block out auditory interference. They called this mechanism the "attentional earlid." According to our speculations, this mechanism acts like a very strong filter that suppresses (and maybe also distorts) input from the auditory modality whenever this modality is irrelevant. At which level of the processing hierarchy this filter takes effect needs to be examined in future research. The present study indicates that this filter is effective under any level of visual-search task difficulty.

There might be other types of auditory distractors that are able to break through the "attentional earlid." For example, future research might find that highly meaningful stimuli, such as your own name, or even more salient sounds can cross that barrier (e.g., Liang et al., 2022). Another potential factor would be the complexity of the distractor material. A common finding in the auditory distraction literature is that the disruptive effect of a sequence increases with its acoustic complexity (Bell et al., 2019; Jones et al., 1993; Schlittmeier et al., 2008). In the present study and others that failed to find an effect of auditory distraction in visual search (Mandal et al., 2024; Wood et al., 2006), simple sounds such as sine-tones or white noise were presented. In future studies, the complexity of the auditory distractor material could be manipulated experimentally or auditory distractors could be used that have been shown to be particularly attention-grabbing (Kattner & Ellermeier, 2018; Perham et al., 2023; Röer et al., 2013, 2017, 2019, 2022; Vachon et al., 2020). We deliberately chose an abstract auditory distractor, because it is the auditory equivalent to abstract visual distractors commonly employed in research on visual distraction and therefore better serves to build a bridge between these two largely disconnected lines of research. In principle, however, the sine-tone used in the present study does hold the potential to interfere with visual search as demonstrated in Experiment 2 (see also Experiment 5 of Mandal et al., 2024). The same tone causes interference in the same visual-search task when a secondary auditory task makes the auditory modality relevant globally while maintaining its irrelevance for the visual-search task itself.

Although they do not produce the anticipated distractor-presence costs on overall visual-search performance, the auditory distractors in our Experiment 1 were also not fully ignored either. Rather, across the multiple experiments of Mandal et al. (2024) and the three tasks-difficulty conditions of the present study, we reliably observe evidence that the auditory distractor speeds up responses. To have such an effect, it must be processed to a certain degree. That this behavioural marker of distractor processing was observed even in an inefficient search task (3°-target contrast), contradicts the predictions that we had made based on Perceptual Load Theory (Lavie, 2005). This finding also conflicts with that of an earlier study by Gao et al. (2021), where the auditory distractor had no effect on an inefficient visual-search task. Gao et al.'s study differs from ours not only in the distractor used, but also in how the search task was rendered inefficient (low target-contrast here and heterogenous distractors in their study; see Duncan & Humphreys, 1989). Future research will tell which of these differences between studies determines whether an auditory distractor is processed during visual search.

Mandal et al. (2024) have speculated that their auditory distractor might have acted as a warning signal or as an accessory stimulus to explain the effect on the SAT. Warning signals provided prior to a task-relevant event are supposed to put the observer in a heightened (alerted) state, ready to perform a task (Posner, 2008). This seems counterintuitive at first, because in the present experiments the onset of the auditory distractor did not precede the visual-search display, but rather coincided with its presentation. However, auditory stimuli are known to be processed faster than visual stimuli (Jose & Gideon Praveen, 2010; Lewald & Guski, 2003; Stein & Meredith, 1993) and might therefore exert an alerting effect on visual search. Another line of research indicates that playing any sound at the same time as the presentation of a to-be-processed visual stimulus speeds responses and this phenomenon has been termed the *accessory-stimulus effect* (e.g., Jepma et al., 2009). In fact, accessory-stimulus effects on the SAT as observed in our studies (faster, but more error-prone responding) have been reported for other tasks before (Low et al., 1996; Posner et al., 1973). Thus, we currently assume that due to the functioning of the "attentional earlid," irrelevant sounds cannot exert their potential to interfere and instead act as an accessory stimulus to affect the SAT. More research is needed to more firmly relate these phenomena.

Our study was motivated by findings from the auditory-distraction literature indicating that task difficulty has an effect on whether an auditory distractor causes interference (for an authoritative review, see Hughes, 2014). Neither did we find any interference at any task-difficulty level in Experiment 1, nor did task-difficulty significantly modulate the interference effect in Experiment 2. The largest effect size for comparisons of task difficulty in Experiment 2 was an $r = -0.18$ for the comparison of 3° vs. 12°, which could be interpreted as high vs. low perceptual load based on the results of Liesefeld et al. (2016). In contrast to the predictions we had derived from the duplex-mechanism account of auditory distraction (Hughes, 2014) and Perceptual-Load Theory (Lavie, 1995, 2005; Lavie & Tsal, 1994), interference was somewhat *higher* in the harder 3°-target task in terms of medians. However, the comparison was not significant, and it (slightly) pointed into the opposite direction for means, so that there is no firm basis to interpret differences in interference across task-difficulty conditions in the present study and our interpretations instead focus on the presence vs. absence of interference effects within each task-difficulty condition.

Obviously, there is a multitude of differences between the tasks traditionally used in research on auditory distraction and the additional-singleton task employed here, so that any comparisons between our studies and this literature should be done only with much caution. Serial recall of (sometimes visually presented) memoranda is probably the most common task to study auditory distraction (e.g., Alikadic & Röer, 2022; Marsh et al., 2020). Some researchers have also used a rapid-serial-auditory-presentation (RSAP) paradigm (Dalton & Lavie, 2004; Dalton & Spence, 2007) which is also interpreted as a search situation. Maybe a crucial difference is that, in these paradigms, stimuli are presented sequentially, whereas in a typical visual-search task all stimuli are presented concurrently. In fact, it is the direct competition between multiple concurrently presented stimuli that is of core interest for theories of visual search (e.g., Duncan & Humphreys, 1989; Wolfe, 2021). Furthermore, even if relevant stimuli in research on auditory distraction are sometimes visually presented, these are often verbalizable material (such as words, digits, or letters). Our hunch is that – in line with classic theories of working memory (Baddeley, 2012; Baddeley & Hitch, 1974) – it makes a difference whether participants are to process such verbalizable material vs. engaging in a task that, like visual search, does obviously not benefit from verbalization. It will be interesting to see whether verbalizable visual material in the primary visual-search or a secondary task (which might engage the "phonological loop" in Baddeley's model), will cause the "attentional earlid" to open just as the auditive counting task employed in our Experiment 2 did.

Generally, there are obvious differences between audition and vision that render it difficult to compare results across the two modalities. For example, while visual stimuli are perceived to occupy specific locations, the perception of spatial positions of auditory stimuli is more ambiguous – with sounds' location being potentially misled by visual information (*spatial ventriloquism effect*; Chen & Vroomen, 2013), or sounds even being perceived as "non-spatial" (Van der Burg et al., 2008). Clearly, the spatial resolution of sounds is much poorer compared to that of visual stimuli. Consequently, space-based suppression mechanisms that are effective against visual distraction (Chelazzi et al., 2019; Gaspelin & Luck, 2018; Liesefeld et al., 2017; Liesefeld & Müller, 2021; Theeuwes et al., 2022) may not effectively apply to auditory distractors, because it is less evident whether a sound comes from the location of the target, non-target, or even from outside of the visual-search display (see Mandal et al., 2024 for a more detailed discussion).

Future research may investigate how more selective distractor handling can be achieved by manipulating the relationships between the auditory distractor on the one hand and the visual-search target or the counting sound on the other hand. Dalton and Spence (2007), for example, found increased distraction if the auditory distractor and the visual target stood out in the same feature dimension (duration) compared to different feature dimensions. Indeed, the dimensional relationship between target and distractor has a strong impact within the visual domain (e.g., Gaspar & McDonald, 2014; Liesefeld et al., 2019, 2022).

Our findings are in line with the *Modality-Weighting Account* (Nasemann et al., 2023; Töllner et al., 2009), an extension of the well-established *Dimension-Weighting Account* of various visual-search phenomena (Found & Müller, 1996; Krummenacher & Mueller, 2012; Liesefeld et al., 2019). It proposes that weights are assigned to perceptual input from different modalities in a way that sensory input from a relevant modality receives increased priority. The Modality-Weighting Account has not yet been applied to distraction, but dimensional weighting is known to play an important role in the handling of visual distractors by enhancing target and/or reducing distractor processing (Liesefeld et al., 2019, 2022; Liesefeld & Müller, 2019, 2021). Strikingly, the complete absence of distractor interference that we interpret as an "attentional earlid" would mean that modality weights can go down to zero, thus effectively avoiding any attentional interference by signals from the respective modality. Such extreme down-weighting of task-irrelevant/-interfering dimensions within the visual modality seems to occur only if there is a high incentive and much opportunity to adapt to the distractor, because the distractor is presented on a huge majority (~80%) of trials (Geyer et al., 2008; Müller et al., 2009; Zehetleitner et al., 2009; see also Liang et al., 2022, for compatible findings of distractor prevalence in the auditory modality). With fewer distractors (50% or even up to 66%, Liesefeld et al., 2022), it is probably not worth the effort or impossible (due to lack of opportunity to adapt) to fully down-weight the distractor dimension within the visual modality. By contrast, the shielding mechanism studied here seems in full effect already with distractors occurring on as few as 25% of trials in Mandal et al. (2024)'s Experiments 3 and 4. From this perspective, the present study demonstrates that signals from the auditory modality can be down-weighted to (almost) zero, that is in our interpretation, the "attentional earlid" can remain firmly closed, across various levels of difficulty of the visual-search task.

Note

1. More precisely, "the auditory modality is irrelevant" here means that there was no information arriving through the auditory modality that was needed or helpful to solve the current experimental task (see the Stimuli section in Liesefeld et al. (2024), for a more detailed terminological treatment of irrelevance in the context of an experimental task).

Acknowledgements

We thank Anna M. Liesefeld for helpful discussions on Experiment 1 and for careful reading and commenting on a previous version of this paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This project was made possible by funding of the Research Group "Handling Visual Distraction" by the Center for Advanced Studies at Ludwig-Maximilians-University München (CAS[LMU]). A.M. is supported by the German Research Foundation (DFG) via the RTG 2175 "Perception in Context and its Neural Basis."

Data availability statement

The manuscript is available as a preprint at PsyArXiv (https://doi.org/10.31234/osf.io/9fv8k). Experiment 1 was not preregistered; Experiment 2 was preregistered (https://osf.io/gbtx8). The datasets generated during the current study and the analysis files are available on OSF (https://osf.io/7v4nf).

ORCID

Ananya Mandal http://orcid.org/0000-0002-2627-0733
Jan Philipp Röer http://orcid.org/0000-0001-7774-3433
Heinrich R. Liesefeld http://orcid.org/0000-0003-4551-8607

References

Alikadic, L., & Röer, J. P. (2022). Loud auditory distractors are more difficult to ignore after all: A preregistered replication study with unexpected results. *Experimental Psychology*, *69*(3), 163–171. https://doi.org/10.1027/1618-3169/a000554

Anderson, B. A., Kim, H., Kim, A. J., Liao, M.-R., Mrkonja, L., Clement, A., & Grégoire, L. (2021). The past, present, and future of selection history. *Neuroscience & Biobehavioral Reviews*, *130*, 326–350. https://doi.org/10.1016/j.neubiorev.2021.09.004

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 8, pp. 47–89). Academic Press.

Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2021). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods*, *26*(3), 295. https://doi.org/10.1037/met0000337

Bell, R., Röer, J. P., Lang, A.-G., & Buchner, A. (2019). Distraction by steady-state sounds: Evidence for a graded attentional model of auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(4), 500–512. https://doi.org/10.1037/xhp0000623

Chelazzi, L., Marini, F., Pascucci, D., & Turatto, M. (2019). Getting rid of visual distractors: The why, when, how, and where. *Current Opinion in Psychology*, *29*, 135–147. https://doi.org/10.1016/j.copsyc.2019.02.004

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics*, *75*(5), 790–811. https://doi.org/10.3758/s13414-013-0475-4

Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods*, *46*(4), 1149–1151. https://doi.org/10.3758/s13428-013-0441-z

Dalton, P., & Lavie, N. (2004). Auditory attentional capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(1), 180–193. https://doi.org/10.1037/0096-1523.30.1.180

Dalton, P., & Spence, C. (2007). Attentional capture in serial audiovisual search tasks. *Perception & Psychophysics*, *69*(3), 422–438. https://doi.org/10.3758/BF03193763

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, *10*(5), 590–604. https://doi.org/10.1162/089892998562997

Escera, C., Corral, M.-J., & Yago, E. (2002). An electrophysiological and behavioral investigation of involuntary attention towards auditory frequency, duration and intensity changes. *Cognitive Brain Research*, *14*(3), 325–332. https://doi.org/10.1016/S0926-6410(02)00135-0

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Forster, S., & Lavie, N. (2008a). Attentional capture by entirely irrelevant distractors. *Visual Cognition*, *16*(2–3), 200–214. https://doi.org/10.1080/13506280701465049

Forster, S., & Lavie, N. (2008b). Failures to ignore entirely irrelevant distractors: The role of load. *Journal of Experimental Psychology: Applied*, *14*(1), 73–83. https://doi.org/10.1037/1076-898X.14.1.73

Found, A., & Müller, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a "dimension-weighting" account. *Perception & Psychophysics*, *58*(1), 88–101. https://doi.org/10.3758/BF03205479

Gao, M., Chang, R., Wang, A., Zhang, M., Cheng, Z., Li, Q., & Tang, X. (2021). Which can explain the pip-and-pop effect during a visual search: Multisensory integration or the oddball effect? *Journal of Experimental Psychology: Human Perception and Performance*, *47*(5), 689–703. https://doi.org/10.1037/xhp0000905

Gaspar, J. M., & McDonald, J. J. (2014). Suppression of salient objects prevents distraction in visual search. *The Journal of Neuroscience*, *34*(16), 5658–5666. https://doi.org/10.1523/JNEUROSCI.4161-13.2014

Gaspelin, N., & Luck, S. J. (2018). Combined electrophysiological and behavioral evidence for the suppression of salient distractors. *Journal of Cognitive Neuroscience*, *30*(9), 1265–1280. https://doi.org/10.1162/jocn_a_01279

Geyer, T., Müller, H. J., & Krummenacher, J. (2008). Expectancies modulate attentional capture by salient color singletons. *Vision Research*, *48*(11), 1315–1326. https://doi.org/10.1016/j.visres.2008.02.006

Halin, N., Marsh, J. E., Hellman, A., Hellström, I., & Sörqvist, P. (2014). A shield against distraction. *Journal of Applied Research in Memory and Cognition*, *3*(1), 31–36. https://doi.org/10.1016/j.jarmac.2014.01.003

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*, article no. 150. https://doi.org/10.3389/fnins.2014.00150

Hughes, R. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, *3*(1), 30–41. https://doi.org/10.1002/pchj.44

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & Jones, D. M. (2013). Cognitive control of auditory distraction: Impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(2), 539–553. https://doi.org/10.1037/a0029064

Iordanescu, L., Grabowecky, M., Franconeri, S., Theeuwes, J., & Suzuki, S. (2010). Characteristic sounds make you look at target objects more quickly. *Attention, Perception & Psychophysics*, *72*(7), 1736–1741. https://doi.org/10.3758/APP.72.7.1736

Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, *15*(3), 548–554. https://doi.org/10.3758/PBR.15.3.548

JASP Team. (2021). *JASP* (Version 0.16.0) [Computer Software]. https://jasp-stats.org/

Jepma, M., Wagenmakers, E.-J., Band, G. P. H., & Nieuwenhuis, S. (2009). The effects of accessory stimuli on information processing: Evidence from electrophysiology and a diffusion model analysis. *Journal of Cognitive Neuroscience*, *21*(5), 847–864. https://doi.org/10.1162/jocn.2009.21063

Jones, D. M., Macken, W. J., & Murray, A. C. (1993). Disruption of visual short-term memory by changing-state auditory stimuli: The role of segmentation. *Memory & Cognition*, *21*(3), 318–328. https://doi.org/10.3758/BF03208264

Jose, S., & Gideon Praveen, K. (2010). Comparison between auditory and visual simple reaction times. *Neuroscience and Medicine*, *1*(01), 30–32. https://doi.org/10.4236/nm.2010.11004

Kaiser, S., Buchner, A., & Bell, R. (2021). Positive and negative mood states do not influence cross-modal auditory distraction in the serial-recall paradigm. *PLoS One*, *16*(12), e0260699. https://doi.org/10.1371/journal.pone.0260699

Kattner, F., & Bryce, D. (2022). Attentional control and metacognitive monitoring of the effects of different types of task-irrelevant sound on serial recall. *Journal of Experimental Psychology: Human Perception and Performance*, *48*(2), 139–158. https://doi.org/10.1037/xhp0000982

Kattner, F., & Ellermeier, W. (2018). Emotional prosody of task-irrelevant speech interferes with the retention of serial order. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(8), 1303–1312. https://doi.org/10.1037/xhp0000537

Krummenacher, J., & Mueller, H. (2012). Dynamic weighting of feature dimensions in visual search: Behavioral and psychophysiological evidence. *Frontiers in Psychology*, *3*, article no. 221. https://doi.org/10.3389/fpsyg.2012.00221

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451–468. https://doi.org/10.1037/0096-1523.21.3.451

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197. https://doi.org/10.3758/BF03213897

Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. *Cognitive Brain Research*, *16*(3), 468–478. https://doi.org/10.1016/S0926-6410(03)00074-0

Liang, W., Brown, C. A., & Shinn-Cunningham, B. G. (2022). Catastrophic effects of sudden interruptions on spatial auditory attention. *The Journal of the Acoustical Society of America*, *151*(5), 3219–3233. https://doi.org/10.1121/10.0010453

Liesefeld, H. R., Fu, X., & Zimmer, H. D. (2015). Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1140–1151. https://doi.org/10.1037/xlm0000081

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods*, *51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Liesefeld, H. R., & Janczyk, M. (2023). Same same but different: Subtle but consequential differences between two measures to linearly integrate speed and accuracy (LISAS vs. BIS). *Behavior Research Methods*, *55*(3), 1175–1192. https://doi.org/10.3758/s13428-022-01843-2

Liesefeld, H. R., Lamy, D., Gaspelin, N., Geng, J. J., Kerzel, D., Schall, J. D., Allen, H. A., Anderson, B. A., Boettcher, S., Busch, N. A., Carlisle, N. B., Colonius, H., Draschkow, D., Egeth, H., Leber, A. B., Müller, H. J., Röer, J. P., Schubö, A., Slagter, H. A., … Wolfe, J. (2024). Terms of debate: Consensus definitions to guide the scientific discourse on visual distraction. Attention. *Perception, & Psychophysics*. Advance online publication. https://doi.org/10.3758/s13414-023-02820-3

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2019). Distractor-interference reduction is dimensionally constrained. *Visual Cognition*, *27*(3–4), 247–259. https://doi.org/10.1080/13506285.2018.1561568

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2022). Preparatory control against distraction is not feature-based. *Cerebral Cortex*, *32*(11), 2398–2411. https://doi.org/10.1093/cercor/bhab341

Liesefeld, H. R., Liesefeld, A. M., Töllner, T., & Müller, H. J. (2017). Attentional capture in visual search: Capture and post-capture dynamics revealed by EEG. *NeuroImage*, *156*, 166–173. https://doi.org/10.1016/j.neuroimage.2017.05.016

Liesefeld, H. R., Moran, R., Usher, M., Müller, H. J., & Zehetleitner, M. (2016). Search efficiency as a function of target saliency: The transition from inefficient to efficient search and beyond. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 821–836. https://doi.org/10.1037/xhp0000156

Liesefeld, H. R., & Müller, H. J. (2019). Distractor handling via dimension weighting. *Current Opinion in Psychology*, *29*, 160–167. https://doi.org/10.1016/j.copsyc.2019.03.003

Liesefeld, H. R., & Müller, H. J. (2021). Modulations of saliency signals at two hierarchical levels of priority computation revealed by spatial statistical distractor learning. *Journal of Experimental Psychology: General*, *150*(4), 710–728. https://doi.org/10.1037/xge0000970

Low, K. A., Larson, S. L., Burke, J., & Hackley, S. A. (1996). Alerting effects on choice reaction time and the photic eyeblink reflex. *Electroencephalography and Clinical Neurophysiology*, *98*(5), 385–393. https://doi.org/10.1016/0013-4694(96)95085-3

Mandal, A., Liesefeld, A. M., & Liesefeld, H. R. (2024). The surprising robustness of visual search against concurrent auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *50*(1), 99–118. https://doi.org/10.1037/xhp0001168

Marsh, J. E., Campbell, T. A., Vachon, F., Taylor, P. J., & Hughes, R. W. (2020). How the deployment of visual attention modulates auditory distraction. *Attention, Perception, & Psychophysics*, *82*(1), 350–362. https://doi.org/10.3758/s13414-019-01800-w

McDonald, J. J., Stormer, V. S., Martinez, A., Feng, W., & Hillyard, S. A. (2013). Salient sounds activate human visual cortex automatically. *The Journal of Neuroscience*, *33*(21), 9194–9201. https://doi.org/10.1523/JNEUROSCI.5902-12.2013

Moher, J., Abrams, J., Egeth, H. E., Yantis, S., & Stuphorn, V. (2011). Trial-by-trial adjustments of top-down set modulate oculomotor capture. *Psychonomic Bulletin & Review*, *18*(5), 897–903. https://doi.org/10.3758/s13423-011-0118-5

Müller, H. J., Geyer, T., Zehetleitner, M., & Krummenacher, J. (2009). Attentional capture by salient color singleton distractors is modulated by top-down dimensional set. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 1–16. https://doi.org/10.1037/0096-1523.35.1.1

Nasemann, J., Töllner, T., Müller, H. J., & Shi, Z. (2023). Hierarchy of intra- and cross-modal redundancy gains in visuo-tactile search: Evidence from the posterior contralateral negativity. *Journal of Cognitive Neuroscience*, *35*(4), 543–570. https://doi.org/10.1162/jocn_a_01971

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perham, N., Begum, F., & Marsh, J. E. (2023). The categorical deviation effect may be underpinned by attentional capture: Preliminary evidence from the incidental recognition of distracters. *Auditory Perception & Cognition*, *6*(0), 20–51. https://doi.org/10.1080/25742442.2023.2167448

Posner, M. I. (2008). Measuring alertness. *Annals of the New York Academy of Sciences*, *1129*(1), 193–199. https://doi.org/10.1196/annals.1417.011

Posner, M. I., Klein, R., Summers, J., & Buggie, S. (1973). On the selection of signals. *Memory & Cognition*, *1*(1), 2–12. https://doi.org/10.3758/BF03198062

Raveh, D., & Lavie, N. (2015). Load-induced inattentional deafness. *Attention, Perception, & Psychophysics*, *77*(2), 483–492. https://doi.org/10.3758/s13414-014-0776-2

Röer, J. P., Bell, R., & Buchner, A. (2013). Self-relevance increases the irrelevant sound effect: Attentional disruption by one's own name. *Journal of Cognitive Psychology*, *25*(8), 925–931. https://doi.org/10.1080/20445911.2013.828063

Röer, J. P., Bell, R., Buchner, A., Saint-Aubin, J., Sonier, R.-P., Marsh, J. E., Moore, S. B., Kershaw, M. B. A., Ljung, R., & Arnström, S. (2022). A multilingual preregistered replication of the semantic mismatch effect on serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(7), 966–974. https://doi.org/10.1037/xlm0001066

Röer, J. P., Bell, R., Körner, U., & Buchner, A. (2019). A semantic mismatch effect on serial recall: Evidence for interlexical processing of irrelevant speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(3), 515–525. https://doi.org/10.1037/xlm0000596

Röer, J. P., Körner, U., Buchner, A., & Bell, R. (2017). Attentional capture by taboo words: A functional view of auditory distraction. *Emotion (Washington, D.C.)*, *17*(4), 740–750. https://doi.org/10.1037/emo0000274

Schlittmeier, S. J., Hellbrück, J., Thaden, R., & Vorländer, M. (2008). The impact of background speech varying in intelligibility: Effects on cognitive performance and perceived disturbance. *Ergonomics*, *51*(5), 719–736. https://doi.org/10.1080/00140130701745925

Sörqvist, P., & Marsh, J. E. (2015). How concentration shields against distraction. *Current Directions in Psychological Science*, *24*(4), 267–272. https://doi.org/10.1177/0963721415577356

Sörqvist, P., Stenfelt, S., & Rönnberg, J. (2012). Working memory capacity and visual–verbal cognitive load modulate auditory–sensory gating in the brainstem: Toward a unified view of attention. *Journal of Cognitive Neuroscience*, *24*(11), 2147–2154. https://doi.org/10.1162/jocn_a_00275

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. MIT Press.

Tellinghuisen, D. J., & Nowak, E. J. (2003). The inability to ignore auditory distractors as a function of visual task perceptual load. *Perception & Psychophysics*, *65*(5), 817–828. https://doi.org/10.3758/BF03194817

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, *50*(2), 184–193. https://doi.org/10.3758/BF03212219

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Theeuwes, J., Bogaerts, L., & van Moorselaar, D. (2022). What to expect where and when: How statistical learning drives visual selection. *Trends in Cognitive Sciences*, *26*(10), 860–872. https://doi.org/10.1016/j.tics.2022.06.001

Töllner, T., Gramann, K., Müller, H. J., & Eimer, M. (2009). The anterior N1 component as an index of modality shifting. *Journal of Cognitive Neuroscience*, *21*(9), 1653–1669. https://doi.org/10.1162/jocn.2009.21108

Vachon, F., Marsh, J. E., & Labonté, K. (2020). The automaticity of semantic processing revisited: Auditory distraction by a categorical deviation. *Journal of Experimental Psychology: General*, *149*(7), 1360–1397. https://doi.org/10.1037/xge0000714

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š, Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E.-J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*(3), 813–826. https://doi.org/10.3758/s13423-020-01798-5

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

Wood, J., Chaparro, A., Hickson, L., Thyer, N., Carter, P., Hancock, J., Hoe, A., Le, I., Sahetapy, L., & Ybarzabal, F. (2006). The effect of auditory and visual distracters on the useful field of view: implications for the driving task. *Investigative Opthalmology & Visual Science*, *47*(10), 4646. https://doi.org/10.1167/iovs.06-0306

Wöstmann, M., & Obleser, J. (2016). Acoustic detail but not predictability of task-irrelevant speech disrupts working memory. *Frontiers in Human Neuroscience*, *10*, article no. 538. https://doi.org/10.3389/fnhum.2016.00538.

Zehetleitner, M., Proulx, M. J., & Müller, H. J. (2009). Additional-singleton interference in efficient visual search: A common salience route for detection and compound tasks. *Attention, Perception & Psychophysics*, *71*(8), 1760–1770. https://doi.org/10.3758/APP.71.8.1760

Supplementary Materials

Auditory distractors are processed but do not interfere with visual search of any difficulty when sound is irrelevant.

Increasing task-difficulty decreases performance.

Experiment 1

To check whether our manipulation of task difficulty was effective, we compared RTs and error rates (averaged across distractor conditions) between the different levels of complexity (target contrast). In line with Liesefeld et al. (2016), overall task performance improved with higher target contrast. In particular, we observed a significant decrease in RTs (by 445 ms) from 3°-target trials (*M* = 1052 ms, *Mdn ± MAD* = 1058 ± 200 ms) to 12°-target trials (*M* = 608 ms, *Mdn ± MAD* = 607 ± 43 ms) *W* = 165.00, *p* < .001, *r* = .93, BF_{10} = 849.09, and (by 56 ms) from 12°-target trials to 45°-target trials (*M* = 551 ms, *Mdn ± MAD* = 540 ± 49 ms), *W* = 171.00, *p* < .001, *r* = 1.00, BF_{10} = 5481.81. Error rates were significantly higher for 3°-target trials (*M* = 19.33 %, *Mdn ± MAD* = 16.49 ± 9.36 %) compared to 12°-target (*M* = 4.48 %, *Mdn ± MAD* = 3.13 ± 1.39 %), *W* = 168.00, *p* < .001, *r* = .97, BF_{10} = 584.81. However, there was no significant difference in error rates when comparing 12°-target to 45°-target trials (*M* = 4.10 %, *Mdn ± MAD* = 3.39 ± 1.56 %), *W* = 101.00, *p* = .513, *r* = .18, BF_{10} = 0.36.

Additionally, BIS (recalculated for each pairwise comparison of task difficulty levels i.e., BIS calculated for 3°-target vs. 12°-target and same for 12°-target and 45°-target) revealed a significant increase in performance for the 12°-target (*M* = 1.22, *Mdn ± MAD* = 1.34 ± 0.17) compared to the 3°-target (*M* = -1.22, *Mdn ± MAD* = -1.12 ± 0.84), *W* = 0.00, *p* < .001, *r* = -1.00, BF_{10} = 3224.18, along with a significant increase in performance for the 45°-target (*M* = 0.50, *Mdn ± MAD* = 0.57 ± 0.73) compared to the 12°-target (*M* = -0.50, *Mdn ± MAD* = 0.01 ± 0.86), *W* = 0.00, *p* < .001, *r* = -1.00, BF_{10} = 13032.53. Thus, confirming that the overall search performance indeed improved with decreases in task-difficulty (as manipulated using target contrast).

Experiment 2

As in Exp. 1, here again we check whether there the task-difficulty manipulation was effective. Hence, just as we did in Exp. 1, we compared RTs and error rates (averaged across distractor conditions) between the different levels of complexity (target contrast). In line with Exp. 1, and Liesefeld et al. (2016), we observe that the overall task performance improved with higher target contrast (i.e., with decreased task difficulty). In particular, we observe that there is a significant decrease in RT for the 12°-target (*M* = 623 ms, *Mdn ± MAD* = 589 ± 51 ms) by 795 ms compared to the 3°-target (*M* = 1418 ms, *Mdn ± MAD* = 1381 ± 259 ms), *W* =153.00, $p < .001$, *r* = 1.00, BF_{10} = 2945.56, and for the 45°-target (*M* = 558 ms, *Mdn ± MAD* = 555 ± 35 ms) by 64 ms compared to the 12°-target, *W* =153.00, $p < .001$, *r* = 1.00, BF_{10} = 1764.50. Error rates were significantly higher for 3°- target trials (*M* = 11.97 %, *Mdn ± MAD* = 13.77 ± 3.55 %) compared to 12°-target (*M* = 2.61 %, *Mdn ± MAD* = 1.39 ± 0.87 %), *W* = 153.00, $p < .001$, *r* = 1.00, BF_{10} = 2369.31. However, just as also observed in Exp. 1, there was no significant difference in error rates when comparing 12°-target to 45°-target trials (*M* = 2.33 %, *Mdn ± MAD* = 1.46 ± 1.04 %), *W* = 97.00, *p* = .141, *r* = .43, BF_{10} = 0.63.

Further, comparison of BIS between the task difficulties reveals that the overall performance of the 3°-target (*M* =-1.54, *Mdn ± MAD* = -1.52 ± 0.71) was indeed worse than that of the 12°-target (*M* = 1.54, *Mdn ± MAD* = 1.64 ± 0.23), *W* = 0.00, $p < .001$, *r* = -1.00, BF_{10} = 5414.56. Similarly, the performance of the 12°-target (*M* = -0.36, *Mdn ± MAD* = -0.34 ± 0.37) was worse than the 45°-target (*M* = 0.36, *Mdn ± MAD* = 0.39 ± 0.33), *W* = 5.00, $p < .001$, *r* = -0.94, BF_{10} = 223.73. Thereby, further confirming that the visual search performance indeed decreases as the task-difficulty increases.

2.3 Tracking the misallocation and reallocation of spatial attention toward auditory stimuli.

This manuscript has been published as:

Mandal, A., Liesefeld, A. M., & Liesefeld, H. R. (2024). Tracking the misallocation and reallocation of spatial attention toward auditory stimuli. *Journal of Neuroscience*, 44(30).
https://doi.org/10.1523/JNEUROSCI.2196-23.2024

Author contributions

Ananya Mandal: conceptualization (equal), data curation, formal analysis, investigation, methodology (equal), project administration, software, visualization, writing – original draft, and writing – review and editing (equal).

Anna M. Liesefeld: conceptualization (supporting), methodology (supporting), supervision (supporting), and writing–review and editing (supporting).

Heinrich R. Liesefeld: conceptualization (equal), formal analysis (supporting), methodology (equal), project administration (supporting), funding acquisition, resources, supervision, validation, visualization (supporting), writing – original draft (supporting), and writing – review and editing (equal).

Behavioral/Cognitive

Tracking the Misallocation and Reallocation of Spatial Attention toward Auditory Stimuli

Ananya Mandal,[1,2] **Anna M. Liesefeld,**[1] **and Heinrich R. Liesefeld**[2,3]

[1]General and Experimental Psychology, Ludwig-Maximilians-Universität Munich, Munich 80802, Germany, [2]Graduate School for Systemic Neurosciences, Ludwig-Maximilians-Universität Munich, Planegg 82152, Germany, and [3]Department of Psychology, Universität Bremen, Bremen 28359, Germany

Completely ignoring a salient distractor presented concurrently with a target is difficult, and sometimes attention is involuntarily attracted to the distractor's location (attentional capture). Employing the N2ac component as a marker of attention allocation toward sounds, in this study we investigate the spatiotemporal dynamics of auditory attention across two experiments. Human participants (male and female) performed an auditory search task, where the target was accompanied by a distractor in two-third of the trials. For a distractor more salient than the target (Experiment 1), we observe not only a distractor N2ac (indicating attentional capture) but the full chain of attentional dynamics implied by the notion of attentional capture, namely, (1) the distractor captures attention before the target is attended, (2) allocation of attention to the target is delayed by distractor presence, and (3) the target is attended after the distractor. Conversely, for a distractor less salient than the target (Experiment 2), although responses were delayed, no attentional capture was observed. Together, these findings reveal two types of spatial attentional dynamics in the auditory modality (distraction with and without attentional capture).

Key words: **auditory attention capture; auditory distraction; auditory search; ERP component latency; N2ac; selective attention**

Significance Statement

Oftentimes, we find it hard to avoid attending to a salient sound that distracts us from our current tasks. Although a common everyday experience, little is known about how spatial distraction unfolds at the neural level in the auditory modality. Using electrophysiological markers of attention allocations, we report comprehensive evidence of spatial attentional capture by a salient auditory distractor, indicating that attention is first misallocated to the distractor and only afterward reallocated toward the target. Similar patterns were observed earlier only in vision, and their discovery in the auditory modality indicates toward the existence of domain-general spatial attentional dynamics consistent across sensory modalities. We also demonstrate that only a distractor more salient than the target reliably captures attention.

Introduction

People are constantly bombarded with a wide range of sensory experiences. To select and attend the important ones among multiple simultaneous stimuli is a massive computational challenge (Tsotsos, 1990; Bronkhorst, 2000). Likely as a side effect of these highly efficient selection mechanisms (Liesefeld et al., 2021), attention is sometimes misallocated toward a salient-but-irrelevant stimulus—leading to behavioral costs. Although almost exclusively studied in vision, this "problem" of involuntary attentional capture is also relevant for audition. Consider a typical office scenario from an auditory perspective—the sounds from mouse clicks or keyboard typing by a busy coworker, grinding of the coffee machine, footsteps of someone walking by, or even the fan noise of an overheated computer. All these sounds are perceived simultaneously and emerge from various locations—although people actively pay attention to only a small subset of them. A knock on the office door (a salient event which is relevant to the office worker) would immediately draw attention to the door. However, if a bird pecked on the window or a coworker accidentally dropped their coffee mug (an irrelevant-but-salient event), it would also capture attention.

There is little research on such spatial attentional capture by auditory stimuli, but similar scenarios have been extensively studied with visual stimuli using the additional-singleton paradigm—where a salient-but-irrelevant stimulus (distractor) occurring simultaneously with a to-be-found target stimulus

Received Nov. 27, 2023; revised May 4, 2024; accepted May 6, 2024.

Author contributions: A.M., A.M.L., and H.R.L. designed research; A.M. performed research; A.M. analyzed data; A.M. and H.R.L. wrote the paper.

This project was funded by the German Research Association (DFG) via the RTG 2175 "Perception in Context and its Neural Basis."

The authors declare no competing financial interests.

Correspondence should be addressed to Ananya Mandal at ananya.mandal@psy.lmu.de or Heinrich R. Liesefeld at heinrich.liesefeld@uni-bremen.de.

https://doi.org/10.1523/JNEUROSCI.2196-23.2024

involuntarily captures attention under certain conditions (Hickey et al., 2006; Kiss et al., 2012; Burra and Kerzel, 2013; Gaspar et al., 2016; Liesefeld et al., 2022). To study the attentional dynamics involved in visual attentional capture, the N2pc component of the event-related potential (ERP) has proven highly useful. The N2pc component is a transient negative increase in activity over posterior electrode sites contralateral to the attended stimulus—and it is often used as a marker of spatial attention allocations (Eimer, 1996, 2014; Constant et al., 2023). It has been used to measure the timing of attention allocations (Töllner et al., 2011; Grubert and Eimer, 2016) and, in particular, to uncover the temporal dynamics of attentional capture by salient visual distractors (Liesefeld et al., 2017, 2022; Liesefeld and Müller, 2019). Strictly speaking, the notion of attentional capture implies a series of events, involving the initial misallocation of attention to the salient distractor followed by reallocating attention toward the target. The notion of sequential attention allocations also implies that the presence of distractors delays attention allocation to the target compared with when the distractor is absent. These spatiotemporal dynamics predict a very specific N2pc pattern: (1) the occurrence of a distractor N2pc, reflecting attentional capture by the distractor, (2) a delay in target N2pc when the distractor is present compared with distractor absence, and (3) a delay in target N2pc compared with the distractor N2pc (Liesefeld and Müller, 2019).

There is evidence of an auditory analog to the N2pc which has been termed *N2ac* (Gamble and Luck, 2011). The N2ac component is also a transient negative potential contralateral to the location of an attended stimulus, but it is elicited by auditory stimuli and observed at more anterior electrode sites. This component is thought to be functionally similar to the N2pc component (Gamble and Woldorff, 2015; Lewald and Getzmann, 2015; Klatt et al., 2018). Thus, if it is possible to induce auditory attentional capture, we should be able to conceptually replicate the above-described pattern of spatiotemporal dynamics that has been taken as indicative of visual attentional capture. This would allow to further validate the functional interpretation of the N2ac component and to prove the existence of spatial attentional capture in the auditory domain. Most importantly, demonstrating such a relatively complex pattern of spatiotemporal attentional dynamics for the auditory modality in light of previous research in vision would be a strong evidence for more fundamental, modality-overarching principles of spatial dynamics of attentional (mis)allocation.

Experiment 1

Materials and methods

In order to build a basic auditory search scene that would allow to disentangle various attentional dynamics, we adapted the design of Hickey et al. (2009). This design provides several advantages for our purposes. Hickey et al. used a visual-search display with two objects (a square and a line), one stimulus being defined as the target while the other served as the distractor. Participants were asked to report on a feature of the target (e.g., whether the square target was a square or diamond—when the square is rotated 90°—or if the line target was a small or big line). Such tasks in which the target is defined by one feature (e.g., its shape) and participants are to report one of its features (e.g., its size) are called discrimination, categorization, or compound search tasks (Liesefeld et al., 2024) and are commonly employed in the additional-singleton paradigm (Theeuwes, 1991). Notably, by placing one stimulus on the midline and the other lateralized, Hickey et al. were able to disentangle attentional dynamics related to the respective lateralized stimulus, namely, either the target or the distractor, in the presence of the midline stimulus. Just as the N2pc studied in Hickey et al., the N2ac employed here is a component emerging contralateral to the attended stimulus, and stimuli on the midline consequently cannot elicit such lateralized activity. The additional advantage of the Hickey et al. (2009) design for our study is that only very few stimuli were presented simultaneously (also see Eimer, 1996; Hilimire et al., 2012). Presenting many identical (nontarget) stimuli simultaneously (as would be the default for visual-search studies) is not feasible for studies using spatial sounds, because the auditory system cannot spatially segregate more than a few concurrently presented sounds, in particular if these stimuli are quite abstract and similar (Bregman, 1990), like those employed in typical visual-search displays.

On this background, the auditory scene employed in the present study consisted of just two clearly distinguishable sounds (sine waves vs square-tone bursts)—where the instructions defined one as the target and the other as the distractor. Participants were required to report on each trial whether the target stimulus was of high frequency or low frequency indicated through pressing either of two keys.

Participants. Sixteen participants took part in Experiment 1 (median age, 26 years; range, 23–38 years; nine males). This sample size is in the upper range of previous N2pc and N2ac studies and sufficient to detect effects of size $d_z = 0.75$ and above, with a probability of $1—\beta = 0.80$ ($\alpha = 0.05$, two-tailed). In both the experiments, participants reported normal hearing. Written informed consent was obtained from all the participants, and they received course credit or were paid for their participation. No participant had to be excluded. Procedures were approved by the ethics committee of the Department of Psychology and Pedagogics at Ludwig-Maximilians-Universität Munich.

Experimental design. Stimulus presentation and response collection were controlled using PsychoPy (Peirce et al., 2019). The auditory search stimuli were presented using three studio sound monitors (Genelec 8020D, Genelec) that were placed in the frontal field (Fig. 1*a*). One sound monitor was placed directly in front of the participants at 0° angle, while the other two sound monitors were at ∼40° to the left and right, respectively, at 110 cm from the participants. The sound monitors were visible to participants throughout the experiment. Participants were asked to rest their head on a chinrest throughout the experiment such that there are minimal head movements and distance of the head from the sound source is kept constant.

On each trial, either one (distractor-absent condition; target-only) or two (distractor-present condition) sounds were presented from different locations. The target sounds were sine-wave tones of either 440 Hz or 470 Hz. Participants' task was to indicate whether the target was the high-frequency (i.e., the 470 Hz) or the low-frequency (the 440 Hz) tone in each trial. In distractor-present trials, an additional sound, from a different location than the target was presented. The distractor sounds were square-tone bursts (comprising of three 50 ms bursts separated by 50 ms of silence; see Fig. 1*b* for a visual representation of the stimuli's waveforms), which could either be of low frequency (410 Hz) or of high frequency (500 Hz). The temporal gaps in the distractor sounds render them more salient than the continuous sine-wave targets (Kayser et al., 2005). Both the distractor and the target sounds were presented for 250 ms at ∼70 dB(A) SPL,

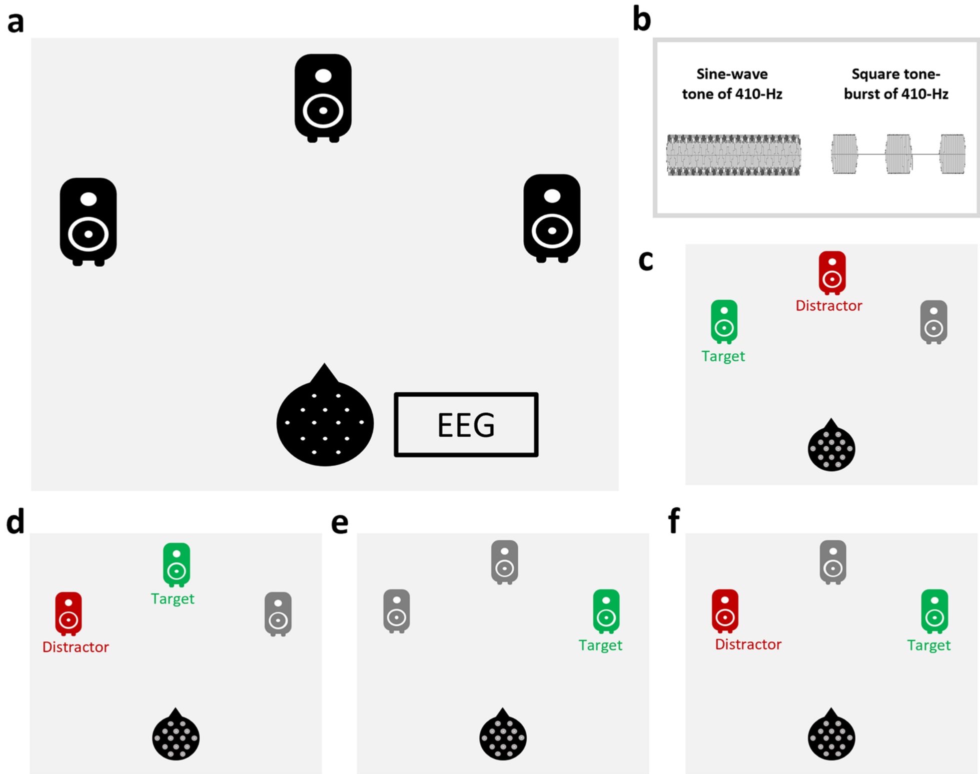


Figure 1. Auditory search scene. ***a***, Setup for the auditory search. ***b***, Representation of the waveform of the two kinds of sounds used. ***c***, Distractor-midline/target-lateral condition. ***d***, Distractor-lateral/target-midline condition. ***e***, Distractor-absent/target-lateral condition. ***f***, Both-lateral condition.

measured at the eardrum, using the miniDSP EARS device (miniDSP). Participants were instructed to maintain eye fixation at a fixation cross presented on a monitor in front of them while performing the task. The monitor was placed such that it was below the sound monitor at 0° and did not obstruct the sound from it. Participants had to respond within 3,000 ms after the onset of the sound (response deadline). In case of incorrect or delayed responses, the fixation cross changed to red or blue for 1,000 ms, respectively. No feedback was provided for correct responses. The intertrial interval was jittered between 800 and 1,200 ms.

The main part of the experiment consisted of 20 blocks of 108 trials each (2,160 trials in total), with 720 distractor-absent trials and 1,440 distractor-present trials. An additional practice block was provided at the beginning of the experiment, which consisted of 108 unanalyzed trials.

The experiments were designed to isolate attentional processing of target and distractor through lateralized ERPs—specifically the N2ac component (Gamble and Luck, 2011). The design follows a similar logic as used to dissect the attentional dynamics of target and distractor processing in visual search using the N2pc component (Liesefeld et al., 2022). The distractor-midline/target-lateral condition (Fig. 1*c*) contained a lateralized target with a distractor on the midline and served to isolate target-related activity. The distractor-lateral/target-midline condition (Fig. 1*d*) contained a lateralized distractor with a target on the midline and served to isolate distractor-related activity. The distractor-absent/target-lateral condition (Fig. 1*e*) contained only a lateralized target without a distractor and served as a baseline for target-related activity in the absence of a distractor. Finally, in the both-lateral condition, the distractor was presented on one and the target on the other side (Fig. 1*f*). Within each block, all display configurations as well as frequency of distractor and target were completely balanced and randomized in distractor-present trials, which were randomly intermixed with distractor-absent trials in which target frequency was balanced and randomized as well.

Electrophysiological recording and analysis. The EEG was recorded via 60 preamplified Ag/AgCl electrodes positioned according to the international 10–10 system. Horizontal ocular artifacts were monitored via two additional electrodes on the outer canthi of both eyes. All impedances were kept below 15 kΩ. Signals were amplified (250 Hz low-pass filter, 10 s time constant; BrainAmp DC, Brain Products) and sampled at 1,000 Hz. EEG data were processed with custom-written MATLAB scripts using functions from EEGLAB v2023.0 (Delorme and Makeig, 2004) and ERPLAB v10.0 (Lopez-Calderon and Luck, 2014) and the “latency.m” function (Liesefeld, 2018).

Signals were rereferenced off-line to the average of both mastoids. A 0.5 Hz high-pass and a 40 Hz low-pass FIR filter (EEGLAB default) were applied, after which independent component analysis was run on the signal. The independent components were classified using the ICLabel v1.4 (Pion-Tonachini et al., 2019) tool within EEGLAB, and components representing blinks or horizontal eye movements (prob. >80%) or muscle artifacts (prob. >90%) were then removed from the continuous EEG data (resulting in the removal of $M = 3.94$, min = 2, max = 10 components in Experiment 1 and $M = 4.88$, min = 1, max = 9 components in Experiment 2). After this, the data were segmented into epochs from −200 to 800 ms relative to the search-stimuli onset and baseline-corrected relative to the prestimulus interval. The trials with artifacts in the analyzed channels (FC5/6; voltage steps larger than 50 µV per sampling point, activity changes <0.5 µV within a 500 ms time window, or absolute amplitude exceeding ±80 µV; equal to $M = 0.37\%$, min = 0%

and max = 3.61% of trials in Experiment 1 and $M = 2.85\%$, min = 0% and max = 17.23% of trials in Experiment 2) or incorrect responses were excluded (at least 367 trials remaining per individual in each condition after trial exclusion). The electrode positions FC5/6 were chosen (instead of the anterior electrode cluster used in Gamble and Luck, 2011) as the N2ac has been found to be the most prominent at these locations in previous research (Lewald and Getzmann, 2015; Lewald et al., 2016).

In order to determine the analysis window for the component of interest, the on- and offsets of the strongest component of the respective polarity were identified as the time points where the ERP in the grand-average (GA) difference wave crossed 30% of the component's peak amplitude (detected in the time window of 0–700 ms after the stimulus onset). The 50% area latency was used for the component latency estimation, where the component area was bounded by the ERP, a threshold set at 30% of the respective component's peak amplitude searched in a common time window encompassing GA on- and offsets of all analyzed components. Here and in previous work (Liesefeld et al., 2017, 2022; Liesefeld, 2018), we chose 50% area latency as our latency measure since it is more robust than other latency measures and approximates the median of the distribution of component latencies across trials (Luck, 2005; Liesefeld, 2018). Peak latency approximates a mode of the distribution, whereas percent–amplitude (onset) latency is biased toward the earliest latencies, and both earliest trials and the mode are arguably less representative of the latency distribution than the median (Rousselet and Wilcox, 2020). Both alternative measures are more prone to noise than 50% area latency, and therefore we employ a jackknife procedure (Smulders, 2010) when we calculate them for comparison. We report and interpret these alternative latency measures, i.e., jackknifed 50% amplitude (onset) latency and jackknifed peak latency, whenever at least one of their results are in conflict with results from our preferred 50% area latency.

Statistical analyses. For amplitudes, we report p_{perm} and p_{permAdj} values obtained from two permutation methods modeled after Sawaki et al. (2012). The procedure followed for p_{perm} is the same as that used in Liesefeld et al. (2022). For each of the 10,000 permutations, n_L and n_R trials were randomly assigned from the respective display configuration to the left and right ERP. Here, n_L and n_R represent the number of trials that went into the respective original individual ERPs after any trial rejection. The GA waveform was built from these, and the signed area amplitude (average of amplitudes in the predicted direction across the analysis window) was extracted in the time range of 0–500 ms (fixed across participants and conditions). This is 100 ms longer than that employed by Liesefeld et al. (2022; who used a 100–500 ms time window) to encompass potential earlier components which might occur due to the higher processing speed of the auditory modality (Stein and Meredith, 1993). For p_{permAdj} we used the same procedure as p_{perm} except that the analysis time range was individually adjusted to each component. In particular, comparable to the nonpermutation amplitude analysis in Liesefeld et al. (2017, 2022), we calculated the amplitude for each permutation from a 30 ms time window centered around the respective component's 50% area latency in the GA waveform in each run. Thus, this analysis time window was fixed across participants for a particular component (and run) but varied across components (and runs). The p_{perm} and p_{permAdj} values indicate the proportion of runs in the random permutation that yielded an amplitude larger than or equal to that of the correct assignment of left and right trials for the analysis. In other words, p_{perm} and p_{permAdj} can be interpreted as the probability of observing a value larger or equal to the observed amplitude merely due to random fluctuations. While robust for ERP waveforms containing a single component, p_{perm}s Type II error rate (false negatives) might be inflated by an additional component of opposite polarity within the analysis time window (e.g., two N2pcs for stimuli on different sides in the both-lateral condition). Maybe the signed-amplitude measure of a considerable number of random permutations will be affected by the component of opposite polarity and thereby inflate the proportion of runs with a larger amplitude. Conversely, any component of the same amplitude would inflate p_{perm}s Type I error rate (false positives), because with a broad time window, it will be counted into the signed-amplitude measure of the targeted component. Our new p_{permAdj}, by focusing on individualized (comparatively smaller) time windows, eliminates the need for a fixed broad time window encompassing all components that causes these problems. This permutation approach still ensures the desired Type I error rate (5%), which might even be more closely met than with p_{perm} in situations with multiple components of the same sign. Some of these advantages of our new method are exemplified on the present data in the analyses below.

Pairwise comparisons between distractor-absent and distractor-present conditions for median-correct reaction times (RTs) and error rates were performed using Wilcoxon signed-rank tests (two-tailed comparisons). Result graphs show the mean of medians and within-participant confidence intervals (Cousineau and O'Brien, 2014). Additionally, we report Bayes factors (BFs), for the Bayesian Wilcoxon signed-rank test (result based on data augmentation algorithm with five chains of 10,000 iterations), quantifying evidence for the alternative over the null hypothesis (BF_{10}). BFs were calculated using the standard JZS Cauchy prior with a scale factor of $r = \sqrt{2}/2$. To classify the strengths of evidence through the BFs, we used Jeffreys's criterion (van Doorn et al., 2021)—which states that for the alternative hypothesis, BFs between 1 and 3 are weak evidence, BFs between 3 and 10 are moderate evidence, and BFs greater than 10 are strong evidence. Following these criteria, $1/3\ (=0.33) < BF_{10} < 1$ is considered weak evidence for the null hypothesis, $1/10\ (=0.10) < BF_{10} < 1/3\ (=0.33)$ is considered as moderate evidence for the null hypothesis, and $BF_{10} < 1/10\ (=0.10)$ is considered strong evidence for the null hypothesis. Accordingly, all within-experiment latency differences in the ERP components were compared using frequentist and Bayesian Wilcoxon signed-rank tests (two-tailed comparisons), while for cross-experiment comparisons, frequentist and Bayesian Mann–Whitney U tests were used. All analyses were performed using JASP v0.17.2 (JASP Team, 2023) and using custom scripts in MATLAB. The processed data, analysis scripts, and results are available on OSF (https://doi.org/10.17605/OSF.IO/QR86K).

Results

Distractor interference on behavior

Compared to distractor-absent trials ($M = 526$ ms; Mdn ± MAD = 516 ± 73 ms), distractor presence ($M = 585$ ms; Mdn ± MAD = 573 ± 87 ms) significantly delayed responses by 59 ms; $W = 136.00$; $p < 0.001$; $r = 1.00$; $BF_{10} = 2{,}582.80$. The $r = 1.00$ means that each individual participant was delayed by distractor presence. The error rates for the distractor-present trials ($M = 9.3\%$; Mdn ± MAD = 8.55 ± 4.45%) were also significantly higher (by 3.17%) than those for the distractor-absent trials ($M = 6.13\%$; Mdn ± MAD = 5.94 ± 3.92%); $W = 121.00$; $p = 0.004$; $r = 0.78$;

$BF_{10} = 32.61$. Thus, overall, we observe a significant distractor interference effect in both RTs and error rates (Fig. 2*a*).

Distractor presence delays attention allocation toward the target

As expected, spatial attention was (eventually) directed to the location of the target irrespective of whether the distractor was present or absent. This was evidenced by the N2ac component which was observed for the target in both the distractor-absent ($p_{perm} < 0.001$; $p_{permAdj} < 0.001$) and the distractor-midline ($p_{perm} < 0.001$; $p_{permAdj} < 0.001$) condition (Fig. 3*a*). Importantly, distractor presence ($M = 250.63$ ms; Mdn ± MAD = 250.5 ± 16 ms) delayed the target N2ac by 46.56 ms ($W = 117.5$; $p = 0.011$; $r = 0.73$; $BF_{10} = 5.56$) compared with distractor absence ($M = 204.06$ ms; Mdn ± MAD = 206.5 ± 25.5 ms). This pattern of results indicates that the salient distractor indeed delayed the allocation of attention to the target. Such delay in attention allocation to the target due to a distractor is well documented in visual search using the N2pc component (Liesefeld et al., 2017, 2022). This delay is comparable with that induced by visual distractors in Liesefeld et al. (2017), 59 ms, and Liesefeld et al. (2022), 70–88 ms.

Attentional capture by a salient distractor: distractor N2ac

The interference due to the distractor (as observed through the behavioral results and the delay in target N2ac due to distractor presence) could be due to the distractor entering the competition for attention and either winning this competition—thereby capturing attention or (barely) failing to win the competition and only delaying the attention toward the target—"nonspatial filtering costs" (Becker, 2007; Liesefeld et al., 2019). If attention is indeed captured by the salient distractor, then we should observe a significant N2ac toward the lateral distractor. To resolve for this, we analyzed the distractor-lateral/target-midline condition and observed a significant N2ac ($p_{perm} = 0.025$; $p_{permAdj} < 0.001$) for the distractor (Fig. 3*a*), thus indicating that the distractor indeed captured attention.

Attention shifts from distractor to the target

As a final criterion to consider the observed interference as attentional capture, we predicted that spatial attention as indicated by the N2ac first goes to the distractor and only afterward moves on to the target (Liesefeld and Müller, 2019). To examine this, we compared the distractor N2ac latency in the distractor-lateral/target-midline condition to the latency of the N2ac in the distractor-midline/target-lateral condition. Indeed, the N2ac toward the distractor ($M = 128.19$ ms; Mdn ± MAD = 135 ± 12 ms) was 122.44 ms earlier than the N2ac toward the target ($M = 250.50$ ms; Mdn ± MAD = 250.63 ± 16 ms); $W = 136.00$; $p < 0.001$; $r = 1.00$; $BF_{10} = 3{,}241.99$ (Fig. 3*a*).

Another (independent) way to test for this attention shift is possible with data from the both-lateral condition (Fig. 4*a*; plotted and analyzed relative to the target side). There, we should observe an N2ac first toward the distractor followed by an N2ac toward the target in the same difference wave. In a difference wave calculated with respect to the target, the former should show up as a positivity, because the distractor is opposite to the target. Indeed, there was first a pronounced positivity, indicating an N2ac toward the distractor, which was followed by a negativity, N2ac toward the target (Fig. 4*a*). This switch in polarity of the N2ac closely resembles the N2pc polarity flip reported first by Hickey et al. (2006; but see McDonald et al., 2013) and later demonstrated more robustly by Liesefeld et al. (2017). For research on visual attentional capture, this flip has been of high theoretical significance (Theeuwes, 2010). While both N2ac components observed here were significant with our new method ($p_{permAdj} < 0.001$), only the target N2ac ($p_{perm} < 0.001$), but not the distractor N2ac ($p_{perm} = 0.061$), was significant with the more common approach.

This exemplifies the advantages of our improved permutation approach when components of opposing sign occur within the same ERP: (1) from the midline conditions, we know that there is a distractor N2ac preceding the target N2ac, and (2) from the composite ERP (see the next section), we know that two components together (but with opposite sign) produce the characteristic pattern observed in the both-lateral condition. Thus, not observing a significant distractor N2ac in the latter condition must be a false negative due to a limitation of p_{perm}. We assume that the smaller positivity (distractor N2ac) is missed, because the analysis window also includes the target N2ac, which is larger and therefore dominates the outcome of p_{perm} (see Materials and Methods). By contrast, $p_{permAdj}$ avoids this problem by taking a smaller time window targeted at the component of interest, which is less likely to be confounded with other components.

In any case, both sets of findings (midline vs both-lateralized conditions) clearly indicate that attention shifts from the distractor to the target. The estimated time required to shift attention from one sound to the other (122.44 ms) is within the estimates for the time required for relocating attention in visual search (100–150 ms; Woodman and Luck, 2003) and is close to the estimates empirically observed for visual targets and distractors by

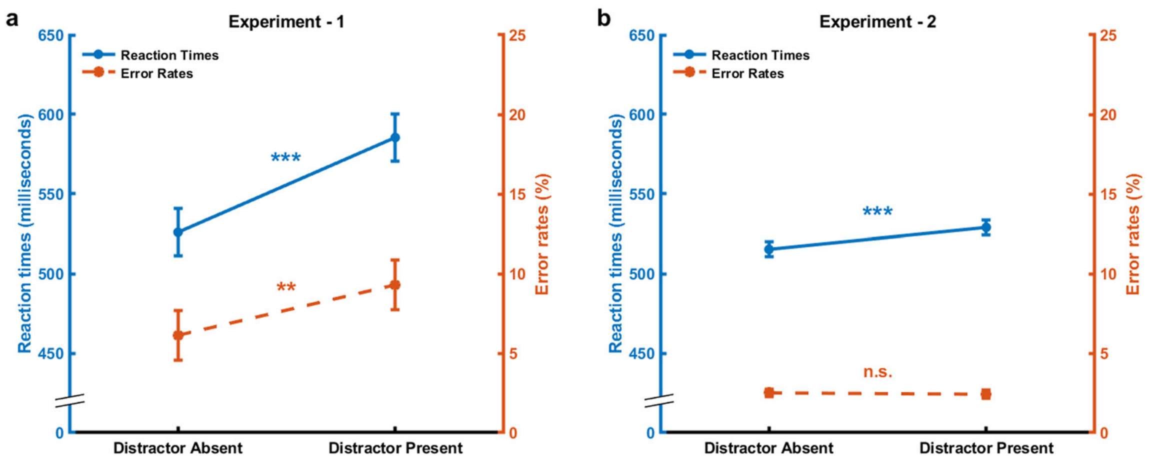


Figure 2. Distractor interference in Experiments 1 and 2. ***a***, Mean of the median RTs and error rates for the distractor-absent and distractor-present conditions in Experiment 1. ***b***, RTs and error rates in Experiment 2. Error bars represent 95% within-participant confidence intervals. $^{n.s.}p > 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

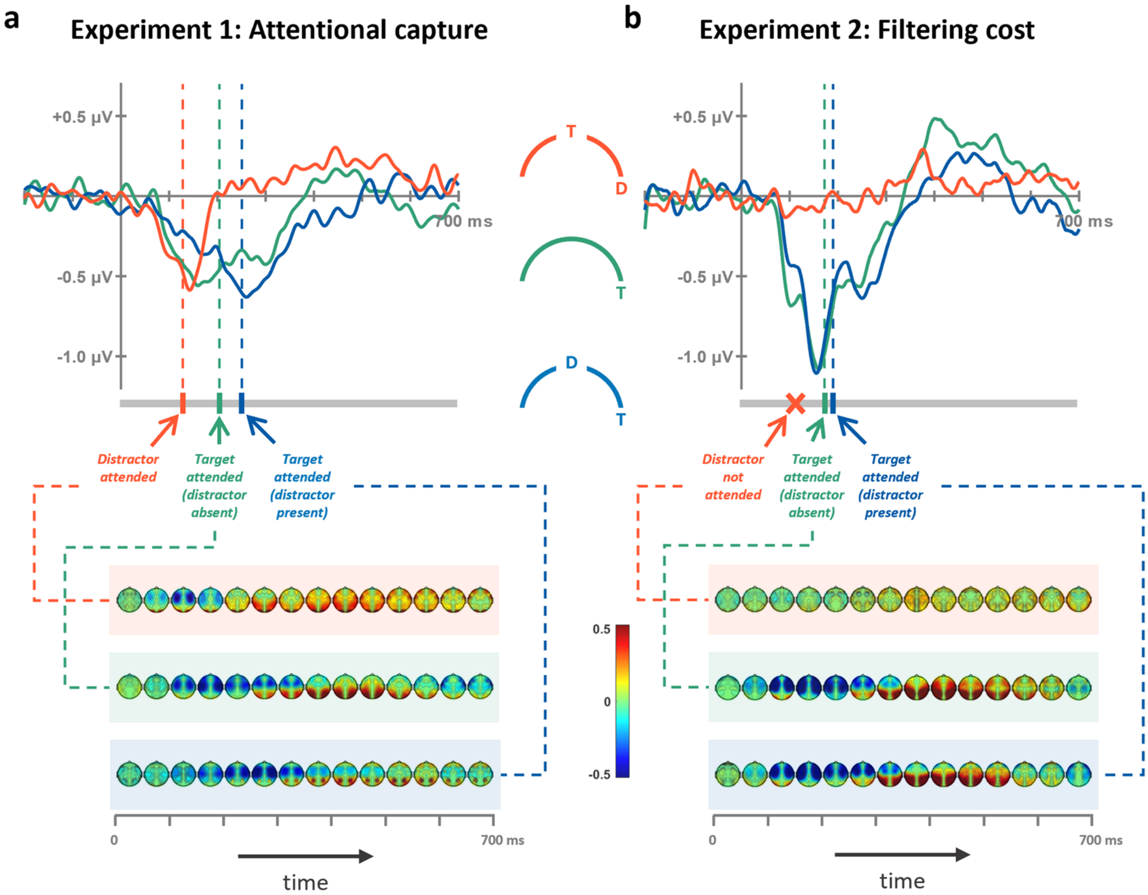


Figure 3. Tracking the dynamics of attentional capture and filtering costs using N2ac. Top panels demonstrate the exact pattern of results as obtained from Experiments 1 (***a***) and 2 (***b***) using the N2ac component, while the bottom panels show the respective scalp topography of the contralateral minus ipsilateral waveforms at each electrode location calculated for 50 ms time windows from 0 to 700 ms. Vertical dashed lines in the ERPs (top panels) indicate mean 50% area latencies. A smoothing filter (Savitzky–Golay filter; order, 3; frame length, 51) was applied to this and subsequent waveforms before visualization to improve the visibility of effects.

Liesefeld et al. (2017), 100 ms, and Liesefeld et al. (2022), 99–119 ms.

Lateralized ERPs sum up arithmetically

Presenting one stimulus in a lateral position along with another stimulus on the midline serves to isolate the attentional dynamics related to the lateral stimulus in the presence of the other stimulus. If this logic holds, the ERP for displays with the target on one and the distractor on the other side (both-lateral) should simply be the difference between the two midline conditions (Gaspar and McDonald, 2014; Liesefeld et al., 2017). Figure 4*a* shows this difference (composite ERP) plotted along with the actually observed both-lateral ERP. To quantify the obvious overlap between these two waveforms, we correlated both GA ERPs across the (characteristic) time window of 0–500 ms after the stimulus onset. The obtained $R^2 = 0.80$ confirms a strong overlap between the observed and composite ERP. The validity of this approach is further confirmed by the close resemblance of the scalp topography for the observed and composite ERPs (Fig. 4*a*, bottom panel).

Experiment 2

Materials and methods

One possible critique of the findings in Experiment 1 could be that, although completely in line with attentional capture, the N2ac pattern observed is due to low-level sensory imbalances caused by the presentation of a stimulus in one auditory hemifield and not due to attentional processes. To rule out this possibility and to also examine distractor interference without attentional capture (filtering costs), we performed a control experiment where we swap the roles of the target and the distractor from Experiment 1 (Gaspar and McDonald, 2014; Barras and Kerzel, 2017; Gaspelin et al., 2023). If the N2ac pattern observed in Experiment 1 was due to sensory imbalances in the auditory hemifield, we would expect to observe the exact same pattern in Experiment 2. If, however, the N2ac in Experiment 1 really indicates attentional capture, we should not replicate it when switching target and distractor identities. This is because the square-tone bursts employed as a distractor in Experiment 1 and as a target in Experiment 2 are much more salient than the sine-wave sounds, such that it should be much easier to select the target and to ignore the distractor (Zehetleitner et al., 2013; Gaspar and McDonald, 2014) in this experiment.

A new sample of 16 participants (median age, 24 years; range, 19–44 years; eight males) took part in Experiment 2. The general setup and the task design were similar to that of Experiment 1, except that the target- and distractor-defining features were swapped (while the reported features, i.e., the base frequencies, remained the same). That is, the target was now a square-tone burst (three 50 ms tone burst separated by 50 ms silence in between) of either 440 Hz (low frequency) or 470 Hz (high frequency). The distractors were continuous sine tones of either 410 Hz (low frequency) or 500 Hz (high frequency). Both the

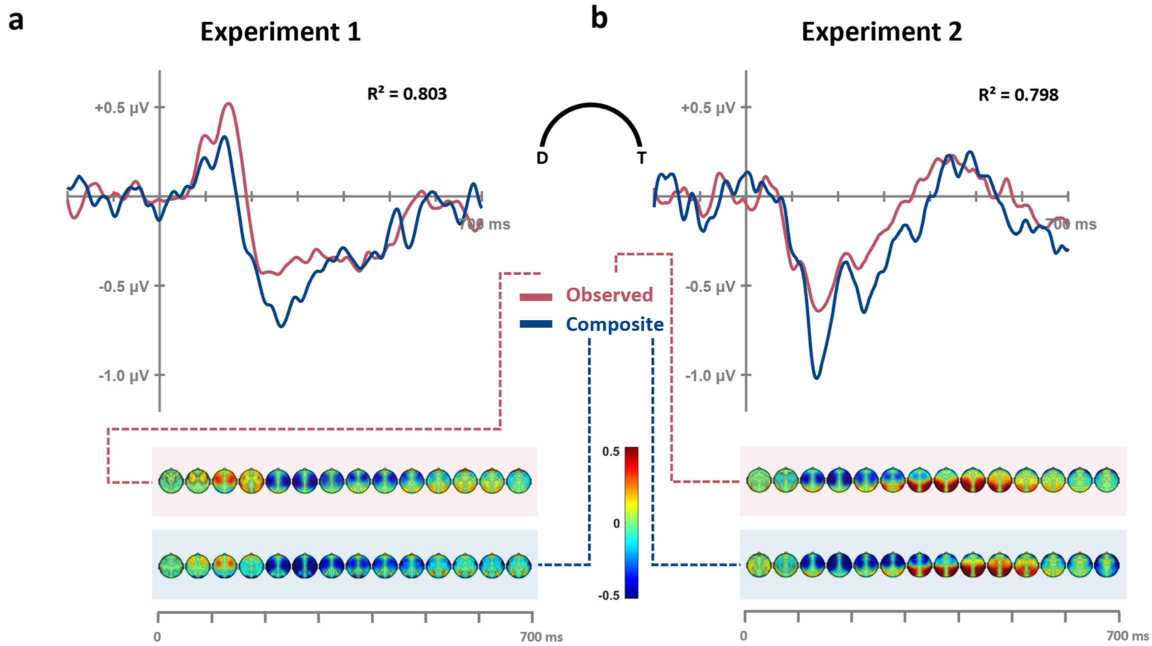


Figure 4. Observed and composite ERP for both-lateral conditions plotted relative to the target side for Experiment 1 (*a*) and Experiment 2 (*b*). Composite ERPs for both-lateral conditions were constructed by subtracting the distractor-lateral/target-midline condition from the distractor-midline/target-lateral condition. Top panel shows different waves, and bottom panel shows the corresponding topoplots. Interestingly, the composite ERPs seem to accentuate the patterns replicated in the actual both-lateral ERPs, which might indicate that they extract the underlying signals more accurately.

target and distractor were presented for 250 ms as in Experiment 1. Here again, participants were instructed to report whether the target was of high frequency or low frequency. Importantly, the switch of target- and distractor-defining features was supposed to render the target more salient than the distractor.

Results

Distractor interference on behavior

Responses on distractor-present trials ($M = 529$ ms; Mdn ± MAD = 500 ± 61 ms) were significantly delayed (by 14 ms; Fig. 2*b*) compared with distractor-absent trials ($M = 515$ ms; Mdn ± MAD = 497 ± 66 ms); $W = 130.00$; $p < 0.001$; $r = 0.91$; $BF_{10} = 362.14$. Error rates showed no significant difference (Fig. 2*b*) between distractor-present ($M = 2.42\%$; Mdn ± MAD = 1.32 ± 0.76%) and distractor-absent trials ($M = 2.5\%$; Mdn ± MAD = 1.74 ± 1.04%); $W = 38.00$; $p = 0.222$; $r = -0.37$; $BF_{10} = 0.48$.

Delay in attention allocation to the target due to distractor presence?

As in Experiment 1, here also attention was allocated to the target location—as indicated by an N2ac component that occurred when the distractor was present ($p_{perm} < 0.001$; $p_{permAdj} < 0.001$) or absent ($p_{perm} < 0.001$; $p_{permAdj} < 0.001$). According to our 50% area latency measure, there was a delay of 17.69 ms ($W = 96.50$; $p = 0.041$; $r = 0.61$, $BF_{10} = 3.45$) for the target N2ac (Fig. 3*b*) when the distractor was present ($M = 189.94$; Mdn ± MAD = 179.0 ± 24.5) compared with when the distractor was absent ($M = 172.25$; Mdn ± MAD = 169.5 ± 11). This delay was significantly smaller than that observed for Experiment 1 (difference in delay between Experiments 1 and 2; $M = 28.88$ ms; $U = 189.50$; $p = 0.021$; $r = 0.48$; $BF_{10} = 1.46$).

Although significant when using our preferred 50% area latency measure, this effect is not so clearly discernable from the respective ERP waveforms (Fig. 3*b*). Furthermore, subsequent investigation of other latency measures revealed that while jackknifed 50% amplitude (onset) latency confirms this result ($p < 0.001$), the jackknifed peak latency ($p = 0.388$) did not replicate this effect for Experiment 2. No such dependency of results on the choice of latency measure was observed in Experiment 1, for which the latency differences are also visually more evident (Fig. 3*a*). Thus, latency effects in Experiment 2 depend on the choice of latency measure and should be interpreted accordingly. In particular, readers not (yet) convinced that 50% area latency represents the distribution of single-trial component latencies best (see Materials and Methods) or believe that this particular latency measure is biased in some other way can still safely interpret the latency effects reported for Experiment 1, but not those for Experiment 2.

No evidence for attentional capture by the distractor: absence of distractor N2ac

Unlike in Experiment 1, where the distractor clearly captured attention as indicated by a distractor N2ac, in Experiment 2, we do not observe a statistically reliable N2ac to the distractor ($p_{perm} = 0.306$; $p_{permAdj} = 0.577$; Fig. 3*b*). Note that the nonsignificant $p_{permAdj}$ value here demonstrates that the new method does not have a highly inflated Type 1 error probability. The absence of a distractor N2ac is further confirmed by the absence of a positivity preceding the target N2ac in the observed and composite waveform of the both-lateral condition (which again showed a strong overlap; $R^2 = 0.80$; Fig. 4*b*). In the absence of electrophysiological evidence for attentional capture by the less salient distractor, the distractor-presence effect on behavior (and potentially on the target N2ac latency) is best interpreted as non-spatial filtering costs (in contrast to attentional capture in Experiment 1).

N2ac latency as measure of relative salience?

As discussed earlier, the N2ac component is believed to be the auditory analog of the lateralized N2pc component (sometimes also referred to as PCN; Töllner et al., 2008). Stimulus salience is known to modulate the latency of the N2pc component elicited by a visual-search target, such that with increasing salience N2pc latency decreases (Töllner et al., 2011). To get a first impression of whether N2ac latency might also reflect relative salience, we compared distractor-absent target N2ac latencies across the two experiments since the square-tone burst (target in Experiment 2) is more salient than the sine tone (target in Experiment 1). As expected, in the distractor-absent/target-lateral condition, we indeed see a significant decrease in the target N2ac latency (Fig. 5) by 31.81 ms for Experiment 2 (more salient target) compared with that of Experiment 1 (less salient target; $U = 195.00$; $p = 0.012$; $r = 0.52$; $BF_{10} = 1.96$).

This result, however, also hinges on the validity of our preferred 50% area latency measure. Visual inspection of the respective difference waves (Fig. 5) might not be fully convincing, and other latency measures such as the jackknifed onset ($p = 0.203$) or peak latency ($p = 0.955$) do not confirm this latency difference—indicating that the earlier N2acs and the mode of the N2ac distribution are not affected by our salience manipulation. Alternatively, the failure to observe a significant effect in the alternative latency measures might be due to their lower reliability (Luck, 2005; Liesefeld, 2018). Nevertheless, in light of these discrepancies, further research, ideally employing a within-participant design along the lines of Töllner et al. (2011), is needed to substantiate the sensitivity of N2ac latency for differences in salience.

Discussion

The present study investigated the dynamics of spatial attention in auditory search using the N2ac component. In Experiment 1, an auditory distractor of higher salience than the target captured spatial attention as indicated by an N2ac to the distractor. After this initial misallocation, attention was reallocated to the target as indicated by an ensuing target N2ac, which was delayed compared with a distractor-absent condition. This signature N2ac pattern is the most comprehensive ERP evidence for attentional capture, as explained by Liesefeld and Müller (2019). In contrast, an auditory distractor less salient than the target in Experiment 2 failed to capture attention, although it still delayed responses (and maybe attention allocation toward the target), a phenomenon that is described as nonspatial filtering costs in visual search. This is the first report of both these attentional phenomena in the auditory domain.

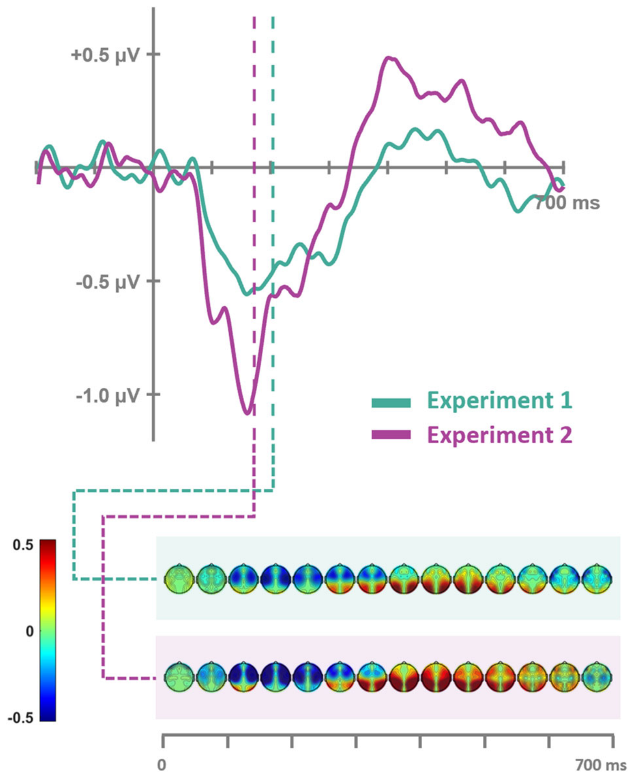


Figure 5. N2ac latency might indicate relative salience of stimuli. The target N2ac 50% area latency when the distractor is absent decreases for Experiment 2 compared with that for Experiment 1, indicating an increased relative salience of the target stimuli in Experiment 2 (square-tone bursts) compared with those in Experiment 1 (sine tones).

Indication of a domain-general attention allocation mechanism

The N2pc component has been extensively studied in visual search over the last decades and has been established as a marker for the spatial allocation of attention in vision (Luck and Hillyard, 1994; Eimer, 1996, 2014; Constant et al., 2023). Similar lateralized ERP components have been discovered for other modalities, such as the N2ac component (Gamble and Luck, 2011) for audition (which we study here) and the N2cc component (Katus and Eimer, 2019; Tsai et al., 2023) for touch. Both components are believed to be functional analogs of the N2pc component in their respective sensory modalities and are used to index spatial allocation of attention.

The temporal dynamics of misallocation and consequent reallocation of spatial attention due to salient distractors has been earlier demonstrated only in vision through the N2pc component (Liesefeld et al., 2017, 2022). That the N2ac component indicates similar attentional dynamics for audition not only strengthens the functional interpretation of the N2ac component as an auditory analog to the N2pc component but also points toward the existence of domain-general spatiotemporal attentional dynamics—indicating a more fundamental principle of attentional misallocation and reallocation across all sensory modalities (see also Shinn-Cunningham, 2008 for a related discussion).

Differences in spatial attentional allocation for detection and discrimination tasks

In the original N2ac study (Gamble and Luck, 2011), the target N2ac emerged only when the target was accompanied by a distractor, whereas it emerged with and without a concurrently presented distractor in the present study (Fig. 3). This discrepancy is likely due to differences in the task designs employed in the two studies. Gamble and Luck (2011) employed a detection task, wherein participants were tasked with determining presence or absence of the target in each trial. In contrast, we employed a discrimination task, which involves the differentiation of a specific target feature (reported feature). It is well established that detection tasks impose comparatively lower attentional demands compared with discrimination tasks (Lavie, 2005) and that detection might occur in a preattentive manner (Treisman and Gelade, 1980; Luck et al., 1997). In scenarios involving target detection, an exhaustive focal–attentional engagement with the target stimuli might not be obligatory, particularly when the salience of the target is high (which is the case for the targets in Gamble and Luck, 2011). Instead, target presence can in principle be deduced from an inhomogeneity on the priority map induced by a salience signal emitted from the target (Müller et al., 2004; Liesefeld et al., 2016). This signal can inform the correct response without necessitating spatial attentional allocation and thus without eliciting an N2ac. Only when there are competing salience signals in search-detection tasks focal–attentional analysis might be required for confirming target presence (Hoffman, 1979). In contrast, discrimination of the target's reported feature (high vs low frequency in our case) in search-discrimination tasks might make a spatial shift of attention toward the target mandatory, yielding a pronounced N2ac component even in the absence of any competing stimulus.

Absence of distractor suppression (P_D) component

Along with the target N2pc component, Hickey et al. (2009) observed a P_D component—reflecting inhibition of the salient visual distractor (see Gaspelin et al., 2023 for a comprehensive review). Given the similarity of our design with Hickey et al. (2009), except for the modality difference, one could expect to observe signatures of spatial suppression in the distractor-lateral/target-midline condition. Such a P_D component (termed P_{AD}) has been recently reported for auditory distractors in the time range of 100–300 ms at a set of electrodes close to site FC5/6 analyzed here (Lunn et al., 2023). In this study, attention was not captured (no distractor N2ac), so that the P_D would be interpreted as proactive or stimulus-triggered suppression (Liesefeld et al., 2024) that potentially serves to avoid attentional capture (Gaspelin and Luck, 2018, 2019). When attention is captured—as was the case in our Experiment 1—such a P_D should only occur after the distractor N2ac, indicating reactive suppression (Gaspelin et al., 2023; Liesefeld et al., 2024), comparable to what has been observed in visual search (Liesefeld et al., 2017, 2022; Liesefeld and Müller, 2019).

From Figure 3, we observe some positivities contralateral to the distractor occurring in a time range of 300–600 ms in both Experiments 1 and 2. However, these positivities emerge at a time point when the target must already have been found and thus too late to be of functional significance for auditory search. Furthermore, their topography does not conform with that of the P_{AD} observed by Lunn et al. (2023). Rather, topoplots in Figure 3 indicate a prominent positivity over posterior electrodes in this time range in most conditions. Interestingly, a similar posterior contralateral positivity for auditory stimuli has also been observed in comparable previous studies (Gamble and Luck, 2011; Gamble and Woldorff, 2015; Lewald et al., 2016) and has been tentatively interpreted as a shift in visual attention following the auditory attentional shift, reorientation of spatial attention to a neutral position after target localization (Gamble and Luck, 2011; Gamble and Woldorff, 2015), or a posterior contralateral version of the late positive component (Lewald et al., 2016).

Differences between audition and vision

To readers familiar with the physiology of the auditory system, the observation of sound-induced lateralization might be somewhat bewildering. In contrast to the retinotopically organized visual system or the somatotopically organized haptic system, the auditory system only reconstructs the spatial source of a sound via indirect cues such as the interaural level or time differences and is generally thought to be weakly lateralized. The observed sound-induced ERP lateralization therefore indicates that attention-induced lateralization might not depend on the spatial organization of the underlying sensory system.

Gamble and Luck (2011) had argued that the lesser degree of contralaterality of the auditory system compared with the visual system might explain why the N2ac (here and in other studies) is smaller in peak amplitude (0.5–1 μV) than the typical N2pc (1–2 μV). Another reason for small peak amplitudes might be that the N2acs observed so far are less tightly locked to the sound onset and therefore smeared out across time (therefore of lower peak amplitude and longer in duration than the typical N2pc). In fact, stretched-out N2pcs of smaller peak amplitude and higher latency (i.e., smeared out N2pcs) have been observed when they likely reflect the second attention allocation (Liesefeld et al., 2017) or when the target is harder to find (Töllner et al., 2011; Dowdall et al., 2012). In the present study, the N2acs in Experiment 1 (where the target was relatively harder to find) are also relatively less in peak amplitude and longer in duration (i.e., more smeared) than those observed in Experiment 2 (where the target was relatively easier to find), thus pointing toward similar factors (in their respective modalities) influencing the amplitude and duration of both the N2ac and N2pc components.

Also note that the N2ac seems to emerge earlier than the typical N2pc. In the ERP graphs shown in Figure 3, for example, the N2ac seems to emerge (onset) already at ∼70 ms, whereas the typical N2pc does not occur before ∼130 ms [see Liesefeld et al. (2017, 2022) for examples from the same lab and Hickey et al. (2009) for search displays of comparable complexity]. Similarly, while the dynamics of N2ac and N2pc flip (in both-lateral condition) are similar, the whole complex occurs earlier in audition compared with that in vision (Hickey et al., 2006; Liesefeld et al., 2017, 2022). In line with these observations, simple reaction times to auditory stimuli are typically 20–60 ms shorter than to visual stimuli (Lewald and Guski, 2003; Shelton and Kumar, 2010) which is likely the result of faster physiological processing time in the auditory system than the visual system (Stein and Meredith, 1993). In fact, already the very first steps of perceptual processing (the transduction process) are of a magnitude faster in audition compared with that in vision (Recanzone, 2009). Thus, it is not surprising that attentional dynamics also occur earlier when triggered by auditory events compared with visual events. This would be the case even if the attentional process itself is highly comparable or even functionally identical.

References

Barras C, Kerzel D (2017) Salient-but-irrelevant stimuli cause attentional capture in difficult, but attentional suppression in easy visual search. Psychophysiology 54:1826–1838.

Becker SI (2007) Irrelevant singletons in pop-out search: attentional capture or filtering costs? J Exp Psychol Hum Percept Perform 33:764–787.

Bregman AS (1990) *Auditory scene analysis: the perceptual organization of sound*. Cambridge, MA: The MIT Press.

Bronkhorst A (2000) The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. Acta Acust United Acust 86:117–128.

Burra N, Kerzel D (2013) Attentional capture during visual search is attenuated by target predictability: evidence from the N2pc, Pd, and topographic segmentation. Psychophysiology 50:422–430.

Constant M, et al. (2023) A multilab investigation into the N2pc as an indicator of attentional selectivity: direct replication of Eimer (1996). Available at: https://psyarxiv.com/3472y/ [Accessed May 13, 2023].

Cousineau D, O'Brien F (2014) Error bars in within-subject designs: a comment on Baguley (2012). Behav Res Methods 46:1149–1151.

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

Dowdall JR, Luczak A, Tata MS (2012) Temporal variability of the N2pc during efficient and inefficient visual search. Neuropsychologia 50:2442–2453.

Eimer M (1996) The N2pc component as an indicator of attentional selectivity. Electroencephalogr Clin Neurophysiol 99:225–234.

Eimer M (2014) The neural basis of attentional control in visual search. Trends Cogn Sci 18:526–535.

Gamble ML, Luck SJ (2011) N2ac: an ERP component associated with the focusing of attention within an auditory scene. Psychophysiology 48: 1057–1068.

Gamble ML, Woldorff MG (2015) The temporal cascade of neural processes underlying target detection and attentional processing during auditory search. Cereb Cortex 25:2456–2465.

Gaspar JM, Christie GJ, Prime DJ, Jolicœur P, McDonald JJ (2016) Inability to suppress salient distractors predicts low visual working memory capacity. Proc Natl Acad Sci U S A 113:3693–3698.

Gaspar JM, McDonald JJ (2014) Suppression of salient objects prevents distraction in visual search. J Neurosci 34:5658–5666.

Gaspelin N, et al. (2023) The distractor positivity component and the inhibition of distracting stimuli. J Cogn Neurosci 35:1693–1715.

Gaspelin N, Luck SJ (2018) Combined electrophysiological and behavioral evidence for the suppression of salient distractors. J Cogn Neurosci 30: 1265–1280.

Gaspelin N, Luck SJ (2019) Inhibition as a potential resolution to the attentional capture debate. Curr Opin Psychol 29:12–18.

Grubert A, Eimer M (2016) All set, indeed! N2pc components reveal simultaneous attentional control settings for multiple target colors. J Exp Psychol Hum Percept Perform 42:1215–1230.

Hickey C, Di Lollo V, McDonald JJ (2009) Electrophysiological indices of target and distractor processing in visual search. J Cogn Neurosci 21:760–775.

Hickey C, McDonald JJ, Theeuwes J (2006) Electrophysiological evidence of the capture of visual attention. J Cogn Neurosci 18:604–613.

Hilimire MR, Hickey C, Corballis PM (2012) Target resolution in visual search involves the direct suppression of distractors: evidence from electrophysiology: target resolution and distractor suppression. Psychophysiology 49: 504–509.

Hoffman JE (1979) A two-stage model of visual search. Percept Psychophys 25:319–327.

JASP Team (2023) JASP software: JASP (Version 0.17.2)[Computer software]. Available at: https://jasp-stats.org/.

Katus T, Eimer M (2019) The N2cc component as an electrophysiological marker of space-based and feature-based attentional target selection processes in touch. Psychophysiology 56:e13391.

Kayser C, Petkov CI, Lippert M, Logothetis NK (2005) Mechanisms for allocating auditory attention: an auditory saliency map. Curr Biol 15:1943–1947.

Kiss M, Grubert A, Petersen A, Eimer M (2012) Attentional capture by salient distractors during visual search is determined by temporal task demands. J Cogn Neurosci 24:749–759.

Klatt L-I, Getzmann S, Wascher E, Schneider D (2018) The contribution of selective spatial attention to sound detection and sound localization: evidence from event-related potentials and lateralized alpha oscillations. Biol Psychol 138:133–145.

Lavie N (2005) Distracted and confused? Selective attention under load. Trends Cogn Sci 9:75–82.

Lewald J, Getzmann S (2015) Electrophysiological correlates of cocktail-party listening. Behav Brain Res 292:157–166.

Lewald J, Guski R (2003) Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. Cogn Brain Res 16:468–478.

Lewald J, Hanenberg C, Getzmann S (2016) Brain correlates of the orientation of auditory spatial attention onto speaker location in a 'cocktail-party' situation. Psychophysiology 53:1484–1495.

Liesefeld HR (2018) Estimating the timing of cognitive operations with MEG/EEG latency measures: a primer, a brief tutorial, and an implementation of various methods. Front Neurosci 12:765.

Liesefeld HR, et al. (2024) Terms of debate: consensus definitions to guide the scientific discourse on visual distraction. Atten Percept Psychophys.

Liesefeld HR, Liesefeld AM, Müller HJ (2019) Distractor-interference reduction is dimensionally constrained. Vis Cogn 27:247–259.

Liesefeld HR, Liesefeld AM, Müller HJ (2021) Attentional capture: an ameliorable side-effect of searching for salient targets. Vis Cogn 29:600–603.

Liesefeld HR, Liesefeld AM, Müller HJ (2022) Preparatory control against distraction is not feature-based. Cereb Cortex 32:2398–2411.

Liesefeld HR, Liesefeld AM, Töllner T, Müller HJ (2017) Attentional capture in visual search: capture and post-capture dynamics revealed by EEG. Neuroimage 156:166–173.

Liesefeld HR, Moran R, Usher M, Müller HJ, Zehetleitner M (2016) Search efficiency as a function of target saliency: the transition from inefficient to efficient search and beyond. J Exp Psychol Hum Percept Perform 42: 821–836.

Liesefeld HR, Müller HJ (2019) Distractor handling via dimension weighting. Curr Opin Psychol 29:160–167.

Lopez-Calderon J, Luck SJ (2014) ERPLAB: an open-source toolbox for the analysis of event-related potentials. Front Hum Neurosci 8:213.

Luck SJ (2005) *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Luck SJ, Girelli M, McDermott MT, Ford MA (1997) Bridging the gap between monkey neurophysiology and human perception: an ambiguity resolution theory of visual selective attention. Cogn Psychol 33:64–87.

Luck SJ, Hillyard SA (1994) Electrophysiological correlates of feature analysis during visual search. Psychophysiology 31:291–308.

Lunn J, Berggren N, Ward J, Forster S (2023) Irrelevant sights and sounds require spatial suppression: ERP evidence. Psychophysiology 60:e14181.

McDonald JJ, Green JJ, Jannati A, Di Lollo V (2013) On the electrophysiological evidence for the capture of visual attention. J Exp Psychol Hum Percept Perform 39:849–860.

Müller H, Krummenacher J, Heller D (2004) Dimension-specific intertrial facilitation in visual search for pop-out targets: evidence for a top-down modulable visual short-term memory effect. Vis Cogn 11:577–602.

Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, Kastman E, Lindeløv JK (2019) PsychoPy2: experiments in behavior made easy. Behav Res Methods 51:195–203.

Pion-Tonachini L, Kreutz-Delgado K, Makeig S (2019) ICLabel: an automated electroencephalographic independent component classifier, dataset, and website. Neuroimage 198:181–197.

Recanzone GH (2009) Interactions of auditory and visual stimuli in space and time. Hear Res 258:89–99.

Rousselet GA, Wilcox RR (2020) Reaction times and other skewed distributions: problems with the mean and the median. Meta-Psychol 4. Available at: https://open.lnu.se/index.php/metapsychology/article/view/1630 [Accessed March 6, 2024].

Sawaki R, Geng JJ, Luck SJ (2012) A common neural mechanism for preventing and terminating the allocation of attention. J Neurosci 32:10725–10736.

Shelton J, Kumar GP (2010) Comparison between auditory and visual simple reaction times. Neurosci Med 1:30–32.

Shinn-Cunningham BG (2008) Object-based auditory and visual attention. Trends Cogn Sci 12:182–186.

Smulders FTY (2010) Simplifying jackknifing of ERPs and getting more out of it: retrieving estimates of participants' latencies. Psychophysiology 47: 387–392.

Stein BE, Meredith MA (1993) *The merging of the senses*. Cambridge, MA: MIT Press.

Theeuwes J (1991) Cross-dimensional perceptual selectivity. Percept Psychophys 50:184–193.

Theeuwes J (2010) Top–down and bottom–up control of visual selection. Acta Psychol 135:77–99.

Töllner T, Gramann K, Müller HJ, Kiss M, Eimer M (2008) Electrophysiological markers of visual dimension changes and response changes. J Exp Psychol Hum Percept Perform 34:531–542.

Töllner T, Zehetleitner M, Gramann K, Müller HJ (2011) Stimulus saliency modulates pre-attentive processing speed in human visual cortex. PLoS One 6:e16276.

Treisman AM, Gelade G (1980) A feature-integration theory of attention. Cogn Psychol 12:97–136.

Tsai S-Y, Nasemann J, Qiu N, Töllner T, Müller HJ, Shi Z (2023) Little engagement of attention by salient distractors defined in a different dimension or modality to the visual search target. Psychophysiology 12: e14375.

Tsotsos JK (1990) Analyzing vision at the complexity level. Behav Brain Sci 13:423–445.

van Doorn J, et al. (2021) The JASP guidelines for conducting and reporting a Bayesian analysis. Psychon Bull Rev 28:813–826.

Woodman GF, Luck SJ (2003) Serial deployment of attention during visual search. J Exp Psychol Hum Percept Perform 29:121–138.

Zehetleitner M, Koch AI, Goschy H, Müller HJ (2013) Salience-based selection: attentional capture by distractors less salient than the target. PLoS One 8:e52595.

3. General Discussion

3.1 Summary of the studies

Study 1 (Chapter 2.1) explored the potential for auditory distractors to affect performance in a visual search task (cross-modal distraction). Our initial experiments (Experiments 1 and 2) showed no significant distraction when sounds were generally irrelevant. This lack of distraction persisted even when the sounds were presented as a rare oddball stimulus (Experiment 3) or had a 300 ms temporal advantage (Experiment 4). However, a significant distraction effect emerged in Experiment 5 when the auditory modality was made globally relevant through an additional simple auditory task – despite remaining irrelevant to the visual search task. This additional auditory task effectively disabled the auditory attentional shielding mechanism, leading to considerable cross-modal interference.

Study 2 (Chapter 2.2) builds on the findings from Study 1 by examining the potential for auditory distractors to interfere with visual search tasks with different levels of difficulty. Consistently, across three levels of search-task difficulty, we replicated the pattern observed in Study 1, showing no interference from completely irrelevant auditory distractors. However, when the attentional shielding mechanism was experimentally disabled by introducing a secondary task in Experiment 2, the same distractor significantly hindered search performance.

Study 3 (Chapter 2.3) focused on the dynamics of spatial attention during auditory search, utilizing the N2ac component. In the first experiment, an auditory distractor of higher salience than the target captured spatial attention towards itself, as shown by an N2ac response to the distractor. This initial misallocation was followed by a delayed N2ac response to the target, indicating that attention was eventually redirected to the target, but later than in those conditions without the distractor. This pattern is considered the most detailed ERP evidence for attentional capture (Liesefeld & Müller, 2019) as previously observed only in the visual modality (Liesefeld et al., 2017, 2022). Conversely, in the second experiment, a less salient auditory distractor did not capture attention but still caused delays in response time and possibly in attention shifting to the target, similar to non-spatial filtering costs observed in visual searches. This study is the first to document both of these attentional effects and their associated spatio-temporal dynamics for the auditory modality.

3.1.1 The Attentional Earlid

In Studies 1 and 2, auditory distractions impacted visual search only when the auditory modality was made globally relevant by introducing a secondary auditory task. This suggests that when auditory information is entirely irrelevant—meaning no auditory input is needed or useful for completing any task (Liesefeld et al., 2024) —people can effectively prevent auditory stimuli from interfering with visual search, effectively as if they had closed their "*attentional earlid*". This concept of an “attentional earlid” (Mandal, Liesefeld, et al., 2024a) refers to an attentional barrier within the auditory system, protecting against unwanted auditory interference when necessary.

Unlike the eye, which has physical filter, the eyelids to block visual input and help focus on other senses, most senses lack such a mechanism. Eyelids are highly efficient in shielding us from the entirety of visual information if necessary – enabling concentration on non-visual activities like listening to music or meditating. Just as eyelids can block out visual stimuli, the "attentional earlid" serves to block out irrelevant auditory stimuli, helping to maintain focus on other tasks. This attentional barrier is not a physical one but rather an attentional mechanism that filters out auditory distractions, potentially preventing us from being overwhelmed by irrelevant sounds in daily life. For instance, in a busy office environment, the Attentional Earlid would allow an individual to focus on reading a report despite the background noise of conversations and ringing phones. Similarly, in a crowded café, it would help someone concentrate on their book or work on their laptop, filtering out the surrounding clatter and chatter. This mechanism, akin to the function of eyelids but on an attentional level, supports our ability to manage simultaneous input from several sensory modalities and maintain focus in the face of potential distractions.

Further, it suggests that our attentional system is highly adaptable, capable of dynamically adjusting the prioritization of sensory inputs based on their relevance to the task at hand. The existence of the Attentional Earlid also highlights the importance of stimuli relevance in determining whether a particular sensory modality will interfere with another. Further research into the Attentional Earlid could explore its underlying neural mechanisms and how it is influenced by factors such as individual differences in sensory processing, and the nature of the auditory stimuli.

3.1.2 Common spatio-temporal dynamics for visual and auditory attention

In the field of visual search research, the N2pc component has long been recognized as a key electrophysiological indicator of spatial attention allocation (Eimer, 1996; Luck & Hillyard, 1994). This concept is not limited to vision alone, and similar lateralized ERP components have been identified in other sensory modalities. For instance, the N2ac component for auditory spatial attention (Gamble & Luck, 2011), and the N2cc component (also known as CCN; Katus & Eimer, 2019) for tactile stimuli are considered functionally analogous to the N2pc. These ERP components serve as markers for spatial attention allocation within their respective sensory modalities. The presence of N2pc, N2ac, and N2cc components – each representing spatial attention allocation through a transient negative increase in activation in electrodes contralateral to the attended stimuli compared to ipsilateral stimuli – already hints towards a commonality in the mechanism for spatial attentional allocation across different sensory modalities, despite the differences in spatial processing of visual, auditory, and tactile stimuli.

The temporal dynamics of attention, particularly the processes of misallocation and subsequent reallocation of attention caused by salient distractors i.e., events associated with attentional capture, have been well documented in vision through the N2pc component (Liesefeld et al., 2017, 2022; Liesefeld & Müller, 2019). Study 3 (Chapter 2.3) extends this to the auditory modality, where the N2ac reflects similar attentional dynamics. This not only supports the interpretation of the N2ac as an auditory counterpart to the N2pc but also points to a broader, domain-general principle governing spatio-temporal attentional dynamics, suggesting that the mechanisms underlying attentional shifts are consistent across different sensory modalities.

The presence of analogous ERP components across various sensory modalities, each exhibiting similar spatio-temporal dynamics for attentional shifts, indicate that the brain might employ a common mechanism for managing spatial attention across visual, or auditory stimuli. This unified approach highlights the fundamentals of spatial attentional processing that transcends individual sensory modalities, pointing to a more integrated dynamic of how attention is allocated and reallocated in response to spatial stimuli presented concurrently.

3.2 Methodological insights

3.2.1 Speed-accuracy trade-off (SAT) and Balanced Integration Score (BIS)

Despite not affecting target processing, the distractor sounds in Study 1 (Experiments 1 and 4) and Study 2 (Experiment 1) led to significantly faster reaction times (RTs). This significant reduction in RTs observed in the presence of distractor sounds might initially be misinterpreted as the distractors aiding in faster response times – especially in light of non-significant effects on error rates. However, when examining the Balanced Integration Score (BIS; Liesefeld & Janczyk, 2019, 2022) – a performance measure which is relatively insensitive to speed-accuracy trade-offs, it became apparent that there was no significant change in overall performance due to the sounds, suggesting they did not function as distractions in the conventional sense. Instead, a combined analysis of RTs, error rates, and BIS provides a clearer picture, indicating that the distractors influenced the speed-accuracy trade-off (SAT; Heitz, 2014). Specifically, while response to target was quicker, it came at the expense of an increased error rate. This highlights that the presence of distractor sounds did not improve overall task performance but rather shifted the balance between speed and accuracy. Thus, while reaction times alone might suggest a facilitation effect, the overall performance data reveals a more complex interaction where speed improvements were counterbalanced by reduced accuracy.

Further, the observed speed-accuracy trade-off (SAT) effect indicates that the auditory stimuli in these experiments were indeed processed to some extent. Therefore, the lack of distraction observed in these studies is not because the sounds were not processed. Instead, it points to the existence of a (attentional) mechanism that prevents irrelevant auditory information from interfering with the task, as long as the auditory modality is completely irrelevant (the Attentional Earlid, see Chapter 3.1.1). Overall, an essential methodological insight from these studies is that interpreting RTs and error rates separately can be misleading. It is crucial to consider a performance measure that is less sensitive to SATs, such as BIS, to avoid drawing incorrect conclusions.

3.2.2 Modified non-parametric method for ERP: Permutation adjusted ($P_{permAdj}$)

ERP (Event-Related Potential) datasets are inherently complex due to their high temporal resolution and the multiple of electrode sites involved in the measurements, capturing millisecond level changes in brain activity with great detail. This complexity can provide valuable insights into brain activity but also poses significant challenges for analysis and interpretation. The richness of the data, encompassing a vast array of time points and electrode sites, coupled with the possibility of influence from other (non-task related) brain activities such as mind-wandering, means that random variations can easily lead to statistically significant results if inappropriate analyses are conducted. These variations often manifest at certain time points and electrode sites, creating effects that appear noteworthy but are, in reality, spurious or bogus (Luck & Gaspelin, 2017). This makes it easy to obtain significant results in ERP experiments, yet difficult to discern which findings are genuine and replicable, as the data's inherent noise can be mistaken for meaningful patterns (see Luck & Gaspelin, 2017 for a detailed discussion).

One frequently employed method for analyzing ERP data involves visually inspecting the Grand Average (GA) waveform to determine the appropriate analysis time window. Subsequently, explicit statistical analyses are typically conducted solely for the time points and electrode sites that exhibited differences during this visual inspection. This technique introduces additional biases since eyeballing the

GA waveform leads to multiple implicit comparisons (Luck, 2014; Luck & Gaspelin, 2017). To circumvent this problem, it is generally recommended to perform the analysis for the ERP component within a pre-determined time window, established prior to any visual inspection of the GA waveform. This approach helps to minimize biases and ensures a more objective analysis, thereby enhancing the reliability and validity of the findings.

However, using a pre-determined analysis time window is not always feasible. When there is insufficient prior literature on the time window of an ERP component – such as for sufficiently novel ERP studies – or when the ERP component varies widely over a broad range of time intervals, it becomes challenging to define a time window in advance. For example, the P_D component can appear within a considerably broad range of time windows (approximately 100 - 400 ms) depending on the task and stimuli (see Figure 2 of Gaspelin et al., 2023 , for a comprehensive overview of the variation in time window and shape of the P_D component). For such ERP components, with a large range of time-window in which the component occurs, some deflection (here positive deflection in the context of the P_D component) can be expected in this window even in the absence of a consistent/true component. If a narrow measurement window is chosen – based on the presence of a deflection in the observed GA waveforms, this could easily be noise, leading to a biased but statistically significant effect even if no real effect exists. Conversely, using a broad measurement window that includes the entire possible range of the ERP component latencies (for P_D approximately 100 - 400 ms) would mix noise with the actual ERP effect, reducing statistical power.

To address this issue, Sawaki et al. (2012) developed a novel non-parametric permutation approach to estimate the amplitude of an ERP component. This method involves measuring the signed area (e.g., positive for the P_D component and negative for the N2pc component) over a relatively long time-interval in which the ERP component can occur (such as the 100 - 400 ms range for the P_D component, see above). To account for the bias caused by only taking the signed area (as this would only include non-zero values in the chosen direction), a distribution of noise using the noise from the actual data was constructed. To achieve this, the original dataset was repeatedly shuffled, and the signed area was calculated in each run of the permutation, providing us with a distribution of signed area amplitudes of noise in the direction of the ERP (positive for P_D, negative for N2pc or N2ac). This distribution was then used to estimate the likelihood of whether the observed ERP's signed area/amplitude was a result of random fluctuations in the data rather than a real physiological response. This estimate of the likelihood (P_{perm}, also see Liesefeld et al., 2022) indicate the proportion of runs in the random permutation that yielded an amplitude larger than or equal to that of the correct assignment/real ERP. That is, this can be interpreted as the probability of obtaining an amplitude equal to or greater than the observed value purely due to noise. Consequently, depending on the chosen α-level (typically: .05), if the P_{perm} value is less than α-level, it is interpreted as a significant/true ERP component.

The P_{perm} method is generally robust for analyzing ERP waveforms containing a single component, but it can face challenges when additional components of opposite polarity are present within the analysis time window leading to an increased Type-II error rate (false negatives). For instance, in scenarios like the both-lateral condition (when a target and distractor is presented concurrently in opposite hemifields, see Chapter 2.3), two N2acs can occur for stimuli on different sides one after the other (see Figure 4 in Chapter 2.3). In such cases, the presence of the opposite polarity component can skew the signed-amplitude measured in many of the runs of the random permutations, artificially boosting the proportion of runs with larger amplitudes. Similarly, when components of the same polarity are present one after the other, they can increase the Type-I error rate (false positives), since their amplitudes may be counted within the broad time window and affect the target component's signed amplitude measure.

To tackle this problem, in Study 3 (Chapter 2.3, Mandal et al., 2024b) we propose a new modified permutation method to calculate the $P_{permAdj}$. The procedure is similar to the computation of P_{perm} in that this measure is also computed using a non-parametric permutation approach, with the only difference being that instead of using a broad time window, the analysis time window was specifically tailored to each individual component. Specifically, the amplitude in each run of the permutation was calculated from a 30-ms time-window centered around the 50%-area latency of the component in the GA waveform. By using individualized, smaller time windows rather than a broad, fixed time window that can potentially encompass multiple subsequent components, the new measure helps to avoid the aforementioned problems. $P_{permAdj}$ still maintains the desired α-level (Type-I error rate) at 5%, and in scenarios with multiple components of the same sign, it may even provide more accurate control over the error rate compared to the original P_{perm} method.

The data presented in Study 3 (Chapter 2.3, Mandal et al., 2024b) nicely illustrates some of these advantages of the $P_{permAdj}$ over the P_{perm} method. We observe an instance of this in the result for both-lateral condition in Exp. 1 (Figure 4a in Chapter 2.3), where attention shifts from the distractor to the target (both of which are on opposite sides). This results in a flip in polarity of the ERP for the distractor N2ac (when the ERP waveform is calculated with respect to the target) – thereby, resulting in the target N2ac (negative) and the distractor N2ac (positive) of opposite polarity. Consequently, as foreseen – informed through the limitations of P_{perm}, we do not observe a significant distractor N2ac with the P_{perm} and it is only revealed as a significant component through the $P_{permAdj}$. We know that the distractor N2ac exists/ is a real component in this experiment, since (a) we clearly observed it in the distractor-lateral/target-midline condition (Figure 3a in Chapter 2.3), and (b) the composite ERPs (Figure 4a in Chapter 2.3) reproduces the characteristic pattern observed in both-lateral condition. This clearly demonstrates an instance where the Type-II error (false negatives) are increased due to the P_{perm} method in these scenarios and can be circumvented by using the $P_{permAdj}$. That Type-I error rate (false positives) are not inflated in the $P_{permAdj}$ method, is demonstrated by the non-significance of a distractor N2ac in Experiment 2 (Figure 3b in Chapter 2.3) in both $P_{permAdj}$ and P_{perm} – highlighting the absence of a distractor N2ac (which is indeed the case as independently observed through the absence of a positivity in both the observed and composite waveform of the both-lateral condition) in this experiment.

3.3 Processing spatial stimuli in the visual and auditory modality

The way spatial information is processed in visual and auditory modalities differs fundamentally. Vision is primarily spatial, while hearing is more temporally oriented (Kubovy, 1988). Visuo-spatial information is processed directly through retinotopic mapping, where the position of an object is represented based on its location on the retina. In contrast, sound localization relies on more indirect methods, such as interpreting interaural time differences (ITD) and interaural level differences (ILD). Despite these differences, humans can still localize sounds with considerable accuracy, particularly when auditory stimuli are widely separated spatially. Spatial separation plays a crucial role in sound perception as exemplified through studies showing that spatial separation of sound sources enhances the ability to process individual sounds amidst multiple concurrent sounds (Kawashima & Sato, 2015). Additionally, the human auditory cortex represents spatial separation between sounds, helping to segregate them into distinct streams (Shiell et al., 2018).

Horizontal sound localization in humans primarily relies on the above mentioned interaural time and level differences (ITD/ILD) – and an integrated form of these codes for sound laterality is retained in the human auditory cortex (Edmonds & Krumbholz, 2014). The greater the spatial separation between sounds, the more significant is the differences in ITD and ILD, leading to improved localization accuracy. Although the visual modality can perceive spatial positions of stimuli even with minimal separation, the auditory modality also does the job, but requires a larger degree of separation. In other words, the spatial acuity of the visual system is higher compared to the auditory system (Callan et al., 2015; Recanzone, 2013).

Consequently, many studies presenting concurrent spatial auditory stimuli, including our Study 3, present auditory stimuli with a higher angular separation between the auditory stimuli, than would be typically for visual stimuli. In Study 3, we present the concurrent auditory stimuli with an angular difference of ~40° instead of presenting stimuli with a separation of less than 10° – as is the case for most visual search studies (also see Gamble & Luck, 2011; Lewald & Getzmann, 2015; Klatt et al., 2018 for comparable studies). This approach accommodates the lower spatial resolution of the auditory system compared to the visual system, ensuring that the auditory stimuli are distinguishable and accurately localized.

3.4 Similarities in the visual and auditory attention

Despite involving different sensory modalities (with their respective strength and weaknesses as discussed above) – auditory and visual attention share significant similarities in how they function and are processed in the brain. A clear example of this (as discussed in Chapter 3.1.2) is the presence of ERP markers that function similarly as index of visual and auditory spatial attention i.e., N2pc and N2ac respectively – in spite of the obvious differences in how both the modalities process spatial stimuli (see Chapter 3.3). Additionally, the mirroring of temporal dynamics of attention mis-allocation and reallocation in the auditory modality (Chapter 2.3; Mandal et al., 2024b) to those only previously observed in the visual modality (Liesefeld et al., 2017, 2022) further exposes the similarities in spatial attentional mechanisms amongst the two modalities.

Similarities between auditory and visual attention are also observed in the forms of distraction during the concurrent presentation of target and distractor stimuli. As already discussed in Chapter 1.2.2, when target and distractor are presented simultaneously, they compete for attention. This competition can lead to distraction either through attentional capture (where the distractor wins the competition, and attention is misallocated towards the distractor), or filtering-cost (where the target wins but attentional allocation towards the target is delayed none-the-less). Both types of distraction have been earlier observed only in visual search (Becker, 2007; Liesefeld et al., 2017). The presence of these distinct forms of distraction also in the auditory modality (Chapter 2.3; Mandal et al., 2024b) further evidences the similarities between auditory and visual attention.

The ability to learn and utilize spatial regularities is another shared feature of auditory and visual attention. Spatial probability learning, a phenomenon which can be thought of as a long-term effect of selection history (Addleman & Jiang, 2019a) – can help in guiding attention towards a (implicitly) learned spatial region where a target is biased to occur more frequently. Such spatial probability learning has been extensively studied in visual search (Ferrante et al., 2018; Jiang et al., 2013). However, a comparable effect had been shown for auditory search by Addleman & Jiang (2019b), providing evidence for implicit spatial attentional bias in the auditory modality through spatial probability learning. Their study shows that participants can learn to allocate attention more efficiently to locations where targets are more likely to appear, demonstrating the potential for learning spatial regularities also in the auditory modality.

That visual attentional mechanisms can be extended to auditory attention (see Chapter 1.3.2), or auditory and visual attention might have some commonalities – are not a new concept. Indeed, Shinn-Cunningham (2008) reviews several such commonalities observed in the visual and auditory object perception in complex scene – indicating a similarity in attentional mechanisms in the two modalities. Shinn-Cunningham further suggests that these similarities indeed indicate common neural mechanisms controlling attention across modalities. Our findings in Study 3 provides empirical evidence towards this – by highlighting the similarities in attentional dynamics, as observed in the brain for both visual and auditory modalities.

3.5 Limitations and future directions

3.5.1 Study 1 and 2

While the research presented in this thesis helps uncover important aspects of both cross-modal and unimodal auditory attentional mechanisms, there still remains a few questions which future research might help answer. In Studies 1 and 2, the introduction of an additional simple auditory task – the counting of sounds in each block – revealed a cross-modal auditory distractor interference effect. This effect is argued to result from the increased priority of the auditory modality due to the top-down demand of the additional task. Consequently, the system is hindered from down-weighting inputs from the auditory modality (as seen in other experiments without the additional counting task) while integrating activations in the master priority map.

However, an alternate (though perhaps unlikely) explanation could be that performing the additional counting task induced interference due to the higher cognitive load of managing an additional task alongside the main visual search task. In this regard, it must be noted that the additional counting sound was (*a*) presented only sparsely in each block (between 3 to 6 times), and (*b*) was comparatively easier to distinguish from the distractor sound. These factors rendered the counting task relatively easy to perform. If the counting task itself caused the interference, the interference effect should have been visible only when the counting sound was presented, rather than during the distractor sound as well. Additionally, if the interference was indeed due to increased cognitive load, then increasing the difficulty of the main task – thereby, increasing the cognitive load, would have resulted in an interference due to the distractor, even in the experiment with the counting sound absent. However, we fail to observe any such distraction effect by varying difficulty of the visual search task (see Exp. 1 in Chapter 2.2). Nevertheless, future research might aim to disentangle these potential explanations by designing experiments that can isolate such effects from those of top-down attentional demands in the experiment with the additional counting sound. This will further help clarify the mechanisms underlying cross-modal auditory distractor interference and contribute to a more comprehensive understanding.

Another factor worth exploring in future research is the relationship between the distractor sound and the counting sound. Currently, both sounds are sine wave tones that differ only in their frequencies. This could be interpreted as both sounds belonging to the same feature dimension – tone frequency (comparable to a visual target being tilted 12-deg or 45-deg, thereby, belonging to the dimension orientation in vision). It is possible that varying the sound properties so that the counting sound and the distractor sound belong to different dimensions, or are very different from each other in general, could increase or decrease the distraction effect.

3.5.2 Study 3

In Study 3, as an additional peripheral result, we see some evidence of the increase in latency of N2ac with decrease in salience, just as previously observed for the visual N2pc component (Töllner et al., 2011). To get these results, the distractor-absent conditions in both Experiments 1 and 2 were compared, since the target in Experiment 1 was less salient than the target in Experiment 2. This resulted in an earlier N2ac latency for Experiment 2 compared to Experiment 1. Consequently, this constitutes a between-participant comparison for N2ac latency, which is generally not ideal. Thus, to have a robust comparison of the sensitivity of the N2ac latency future research may be conducted which ideally employs a within-participants design similar to the approach used by Töllner et al. (2011).

Another interesting observation from Study 3 is the presence of the late positivities over the posterior electrodes after the time range of the N2acs. Previous comparable research has also reported similar posterior contralateral positivities in response to auditory stimuli (Gamble & Luck, 2011; Gamble & Woldorff, 2015; Lewald et al., 2016). Although, it is not completely clear what this represents, these late positivities have been tentatively interpreted in various ways: such as, visual attentional shift following auditory attention, or re-orientation of visuo-spatial attention after target detection (Gamble & Luck, 2011; Gamble & Woldorff, 2015). Alternatively, they have been considered a posterior contralateral manifestation of the late positive component (LPCpc; Lewald et al., 2016). Given the limited understanding of this component's functionality, further research is necessary to gain more insight.

References (General Introduction and Discussion)*

Addleman, D. A., & Jiang, Y. V. (2019). The influence of selection history on auditory spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(4), 474–488. https://doi.org/10.1037/xhp0000620

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*(25), 10367–10371.

Anderson, P. W. (1972). More Is different: Broken symmetry and the nature of the hierarchical structure of science. *Science*, *177*(4047), 393–396.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Becker, S. I. (2007). Irrelevant singletons in pop-out search: Attentional capture or filtering costs? *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 764–787. https://doi.org/10.1037/0096-1523.33.4.764

Blauert, J. (1996). *Spatial hearing: The psychophysics of human sound localization*. The MIT Press. https://doi.org/10.7551/mitpress/6391.001.0001

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. The MIT Press. https://doi.org/10.7551/mitpress/1486.001.0001

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press. https://doi.org/10.1037/10037-000

* This list of references is only for the General Introduction (Chapter 1) and General Discussion (Chapter 3). The references for the individual studies contained in this thesis are listed at the end of the corresponding chapters (Chapters 2.1 – 2.3).

Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica United with Acustica*, *86*(1), 117–128.

Callan, A., Callan, D., & Ando, H. (2015). An fMRI Study of the Ventriloquism Effect. *Cerebral Cortex*, *25*(11), 4248–4258. https://doi.org/10.1093/cercor/bhu306

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. https://doi.org/10.1016/j.visres.2011.04.012

Chan, L. K. H., & Hayward, W. G. (2009). Feature integration theory revisited: Dissociating feature detection and attentional guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 119–132. https://doi.org/10.1037/0096-1523.35.1.119

Chelazzi, L., Marini, F., Pascucci, D., & Turatto, M. (2019). Getting rid of visual distractors: The why, when, how, and where. *Current Opinion in Psychology*, *29*, 135–147. https://doi.org/10.1016/j.copsyc.2019.02.004

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*(1), 28–71.

Constant, M., & Liesefeld, H. R. (2023). Effects of salience are long-lived and stubborn. *Journal of Experimental Psychology. General*. https://doi.org/10.1037/xge0001420

Constant, M., Mandal, A., Asanowicz, D., Yamaguchi, M., Gillmeister, H., Kerzel, D., Luque, D., Pesciarelli, F., Fehr, T., Mushtaq, F., Pavlov, Y. G., & Liesefeld, H. R. (2023). *A multilab investigation into the N2pc as an indicator of attentional selectivity: Direct replication of Eimer (1996).* PsyArXiv. https://doi.org/10.31234/osf.io/3472y

Crawford, M. B. (2015). *The world beyond your head: On becoming an individual in an age of distraction*. Farrar, Straus and Giroux.

Dalton, P., & Lavie, N. (2004). Auditory attentional capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(1), 180–193. https://doi.org/10.1037/0096-1523.30.1.180

Donk, M., & van Zoest, W. (2008). Effects of salience are short-lived. *Psychological Science*, *19*(7), 733–739. https://doi.org/10.1111/j.1467-9280.2008.02149.x

Draschkow, D., & Võ, M. L.-H. (2017). Scene grammar shapes the way we interact with objects, strengthens memories, and speeds search. *Scientific Reports*, *7*(1), 16471.

Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *353*(1373), 1319–1331.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Edmonds, B. A., & Krumbholz, K. (2014). Are interaural time and level differences represented by independent or integrated codes in the human auditory cortex? *Journal of the Association for Research in Otolaryngology*, *15*(1), 103–114. https://doi.org/10.1007/s10162-013-0421-0

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology*, *99*(3), 225–234. https://doi.org/10.1016/0013-4694(96)95711-9

Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*(10), 526–535. https://doi.org/10.1016/j.tics.2014.05.005

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: A priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382–390. https://doi.org/10.1016/j.tics.2006.06.011

Fiecconi, E. C. (2021). Aristotle on attention. *Archiv Für Geschichte Der Philosophie*, *103*(4), 602–633. https://doi.org/10.1515/agph-2018-0014

Found, A., & Müller, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a "dimension-weighting" account. *Perception & Psychophysics*, *58*(1), 88–101. https://doi.org/10.3758/BF03205479

Gamble, M. L., & Luck, S. J. (2011). N2ac: An ERP component associated with the focusing of attention within an auditory scene. *Psychophysiology*, *48*(8), 1057–1068. https://doi.org/10.1111/j.1469-8986.2010.01172.x

Gamble, M. L., & Woldorff, M. G. (2015). The temporal cascade of neural processes underlying target detection and attentional processing during auditory search. *Cerebral Cortex*, *25*(9), 2456–2465. https://doi.org/10.1093/cercor/bhu047

Gaspelin, N., Lamy, D., Egeth, H. E., Liesefeld, H. R., Kerzel, D., Mandal, A., Müller, M. M., Schall, J. D., Schubö, A., & Slagter, H. A. (2023). The distractor positivity component and the inhibition of distracting stimuli. *Journal of Cognitive Neuroscience*, *35*(11), 1693–1715.

Golob, E. J., Venable, K. B., Scheuerman, J., & Anderson, M. T. (2017). Computational modeling of auditory spatial attention. *CogSci*.

Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, *8*. https://www.frontiersin.org/article/10.3389/fnins.2014.00150

Hickey, C., Di Lollo, V., & McDonald, J. J. (2009). Electrophysiological indices of target and distractor processing in visual search. *Journal of Cognitive Neuroscience*, *21*(4), 760–775. Journal of Cognitive Neuroscience. https://doi.org/10.1162/jocn.2009.21039

Hillyard, S. A., Störmer, V. S., Feng, W., Martinez, A., & McDonald, J. J. (2016). Cross-modal orienting of visual attention. *Neuropsychologia*, *83*, 170–178. https://doi.org/10.1016/j.neuropsychologia.2015.06.003

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, *117*(4), 500.

Ihlefeld, A., & Shinn-Cunningham, B. (2008). Spatial release from energetic and informational masking in a selective speech identification task. *The Journal of the Acoustical Society of America*, *123*(6), 4369–4379. https://doi.org/10.1121/1.2904826

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203.

James, W. (1890). *The principles of psychology, Vol. 2.* Henry Holt and Company. https://doi.org/10.1037/11059-000

Jannati, A., Gaspar, J. M., & McDonald, J. J. (2013). Tracking target and distractor processing in fixed-feature visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(6), 1713–1730. https://doi.org/10.1037/a0032251

Katus, T., & Eimer, M. (2019). The N2cc component as an electrophysiological marker of space-based and feature-based attentional target selection processes in touch. *Psychophysiology*, *56*(9), e13391. https://doi.org/10.1111/psyp.13391

Kawashima, T., & Sato, T. (2015). Perceptual limits in a simulated "Cocktail party." *Attention, Perception, & Psychophysics*, *77*(6), 2108–2120. https://doi.org/10.3758/s13414-015-0910-9

Kaya, E. M., & Elhilali, M. (2012). A temporal saliency map for modeling auditory attention. *2012 46th Annual Conference on Information Sciences and Systems (CISS)*, 1–6.

Kayser, C., Petkov, C. I., Lippert, M., & Logothetis, N. K. (2005). Mechanisms for allocating auditory attention: An auditory saliency map. *Current Biology*, *15*(21), 1943–1947. https://doi.org/10.1016/j.cub.2005.09.040

Kiss, M., Grubert, A., Petersen, A., & Eimer, M. (2012). Attentional capture by salient distractors during visual search is determined by temporal task demands. *Journal of Cognitive Neuroscience*, *24*(3), 749–759. https://doi.org/10.1162/jocn_a_00127

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*(4), 219–227.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 318–320. https://doi.org/10.1037/0096-1523.14.2.318

Kumada, T. (1999). Limitations in attending to a feature value for overriding stimulus-driven interference. *Perception & Psychophysics*, *61*(1), 61–79. https://doi.org/10.3758/BF03211949

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*(3), 143–148.

Lewald, J., Hanenberg, C., & Getzmann, S. (2016). Brain correlates of the orientation of auditory spatial attention onto speaker location in a “cocktail-party” situation. *Psychophysiology*, *53*(10), 1484–1495. https://doi.org/10.1111/psyp.12692

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*(1), 9–16.

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods*, *51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Liesefeld, H. R., & Janczyk, M. (2022). Same same but different: Subtle but consequential differences between two measures to linearly integrate speed and accuracy (LISAS vs. BIS). *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-01843-2

Liesefeld, H. R., Lamy, D., Gaspelin, N., Geng, J. J., Kerzel, D., Schall, J. D., Allen, H. A., Anderson, B. A., Boettcher, S., Busch, N. A., Carlisle, N. B., Colonius, H., Draschkow, D., Egeth, H., Leber, A. B., Müller, H. J., Röer, J. P., Schubö, A., Slagter, H. A., … Wolfe, J. (2024). Terms of debate: Consensus definitions to guide the scientific discourse on visual distraction. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-023-02820-3

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2019). Distractor-interference reduction is dimensionally constrained. *Visual Cognition*, *27*(3–4), 247–259. https://doi.org/10.1080/13506285.2018.1561568

Liesefeld, H. R., Liesefeld, A. M., & Müller, H. J. (2022). Preparatory control against distraction is not feature-based. *Cerebral Cortex*, *32*(11), 2398–2411. https://doi.org/10.1093/cercor/bhab341

Liesefeld, H. R., Liesefeld, A. M., Pollmann, S., & Müller, H. J. (2019). Biasing allocations of attention via selective weighting of saliency signals: Behavioral and neuroimaging evidence for the dimension-weighting account. *Current Topics in Behavioral Neurosciences*, *41*, 87–113. https://doi.org/10.1007/7854_2018_75

Liesefeld, H. R., Liesefeld, A. M., Töllner, T., & Müller, H. J. (2017). Attentional capture in visual search: Capture and post-capture dynamics revealed by EEG. *NeuroImage*, *156*, 166–173. https://doi.org/10.1016/j.neuroimage.2017.05.016

Liesefeld, H. R., Moran, R., Usher, M., Müller, H. J., & Zehetleitner, M. (2016). Search efficiency as a function of target saliency: The transition from inefficient to efficient search and beyond. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 821–836. https://doi.org/10.1037/xhp0000156

Liesefeld, H. R., & Müller, H. J. (2019). Distractor handling via dimension weighting. *Current Opinion in Psychology*, *29*, 160–167. https://doi.org/10.1016/j.copsyc.2019.03.003

Liesefeld, H. R., & Müller, H. J. (2021). Modulations of saliency signals at two hierarchical levels of priority computation revealed by spatial statistical distractor learning. *Journal of Experimental Psychology: General*, *150*(4), 710–728. https://doi.org/10.1037/xge0000970

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146–157.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291–308. https://doi.org/10.1111/j.1469-8986.1994.tb02218.x

Luck, S. J., & Vecera, S. P. (2002). Attention. *Stevens' Handbook of Experimental Psychology*, *1*, 235–286.

Mandal, A., Liesefeld, A. M., & Liesefeld, H. R. (2024a). The surprising robustness of visual search against concurrent auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *50*(1), 99–118. https://doi.org/10.1037/xhp0001168

Mandal, A., Liesefeld, A. M., & Liesefeld, H. R. (2024b). Tracking the misallocation and reallocation of spatial attention toward auditory stimuli. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.2196-23.2024

Maxwell, J. C. (1865). VIII. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, *155*, 459–512.

McDonald, J. J., Störmer, V. S., Martinez, A., Feng, W., & Hillyard, S. A. (2013). Salient sounds activate human visual cortex automatically. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(21), 9194–9201. https://doi.org/10.1523/JNEUROSCI.5902-12.2013

Müller, H. J., & Krummenacher, J. (2006). Visual search and selective attention. *Visual Cognition*, *14*(4–8), 389–410.

Nagel, E. (1961). *Problems in the Logic of Scientific Explanation*. Routledge & Kegan Paul.

Nasemann, J., Töllner, T., Müller, H. J., & Shi, Z. (2023). Hierarchy of intra- and cross-modal redundancy gains in visuo-tactile search: Evidence from the posterior contralateral negativity. *Journal of Cognitive Neuroscience*, *35*(4), 543–570. https://doi.org/10.1162/jocn_a_01971

Recanzone, G. H. (2013). The role of auditory cortex in spatial processing. In Y. E. Cohen, A. N. Popper, & R. R. Fay (Eds.), *Neural Correlates of Auditory Cognition* (pp. 83–114). Springer. https://doi.org/10.1007/978-1-4614-2350-8_4

Sawaki, R., Geng, J. J., & Luck, S. J. (2012). A common neural mechanism for preventing and terminating the allocation of attention. *Journal of Neuroscience*, *32*(31), 10725–10736. https://doi.org/10.1523/JNEUROSCI.1864-12.2012

Shiell, M. M., Hausfeld, L., & Formisano, E. (2018). Activity in human auditory cortex represents spatial separation between concurrent sounds. *Journal of Neuroscience*, *38*(21), 4977–4984. https://doi.org/10.1523/JNEUROSCI.3323-17.2018

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(5), 182–186. https://doi.org/10.1016/j.tics.2008.02.003

Spence, C., & Driver, J. (1997). On measuring selective attention to an expected sensory modality. *Perception & Psychophysics*, *59*(3), 389–403. https://doi.org/10.3758/BF03211906

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, *50*(2), 184–193. https://doi.org/10.3758/BF03212219

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, *51*(6), 599–606. https://doi.org/10.3758/BF03211656

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. https://doi.org/10.1016/j.actpsy.2010.02.006

Töllner, T., Gramann, K., Müller, H. J., & Eimer, M. (2009). The anterior N1 component as an index of modality shifting. *Journal of Cognitive Neuroscience*, *21*(9), 1653–1669. https://doi.org/10.1162/jocn.2009.21108

Töllner, T., Zehetleitner, M., Gramann, K., & Müller, H. J. (2011). Stimulus saliency modulates pre-attentive processing speed in human visual cortex. *PLOS ONE*, *6*(1), e16276. https://doi.org/10.1371/journal.pone.0016276

Tsotsos, J. K. (1990). Analyzing vision at the complexity level. *Behavioral and Brain Sciences*, *13*(3), 423–445. https://doi.org/10.1017/S0140525X00079577

Van der Burg, E., Olivers, C. N. L., Bronkhorst, A. W., & Theeuwes, J. (2008). Pip and pop: Nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(5), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053

van der Heijden, K., Rauschecker, J. P., de Gelder, B., & Formisano, E. (2019). Cortical mechanisms of spatial hearing. *Nature Reviews Neuroscience*, *20*(10), 609–623. https://doi.org/10.1038/s41583-019-0206-5

Wetzel, N., & Schröger, E. (2014). On the development of auditory distraction: A review. *PsyCh Journal*, *3*(1), 72–91.

Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M. (2010). Visual search. *Current Biology: CB*, *20*(8), R346-349. https://doi.org/10.1016/j.cub.2010.02.016

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092. https://doi.org/10.3758/s13423-020-01859-9

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), Article 3. https://doi.org/10.1038/s41562-017-0058

Wöstmann, M., Störmer, V. S., Obleser, J., Addleman, D. A., Andersen, S. K., Gaspelin, N., Geng, J. J., Luck, S. J., Noonan, M. P., & Slagter, H. A. (2022). Ten simple rules to study distractor suppression. *Progress in Neurobiology*, *213*, 102269.

Zehetleitner, M., Proulx, M. J., & Müller, H. J. (2009). Additional-singleton interference in efficient visual search: A common salience route for detection and compound tasks. *Attention, Perception & Psychophysics*, *71*(8), 1760–1770. https://doi.org/10.3758/APP.71.8.1760

Acknowledgements

Firstly, I would like to express my heartfelt gratitude to my PhD supervisor, Dr. Heinrich Liesefeld. Your unwavering support, both scientifically and administratively, has been an essential and integral part of my progress. Heinrich, thank you for your exceptional guidance and for training me in the art of doing science. Your encouragement to explore a wide range of topics and your patience in guiding me through each of them have truly shaped my development as a researcher. Your mentorship has made this challenging journey not only possible but also deeply rewarding.

I am deeply grateful to my thesis committee members, Prof. Hermann Müller and Prof. Benedikt Grothe for your helpful discussions, and valuable suggestions throughout my PhD journey. Additionally, many thanks to Dr. Anna Liesefeld for your continuous support, and helpful suggestions.

I express my special gratitude to Dr. Artyom Zinchenko, whose kindness and dependability have made him an invaluable colleague and friend. Artyom, working with you has been an absolute joy. I have thoroughly enjoyed our engaging brainstorming sessions, and I truly appreciate all the support you have provided. You have shown me how fun and seamless collaborations can be, and for that, I am deeply thankful.

I extend my deepest gratitude to Verena Winkler for your constant administrative support throughout the years, right from the beginning of my time with the RTG and up until the end. Your kindness and support have been indispensable, ensuring that everything ran smoothly such that I could focus on my research. Additionally, I would like to thank Gabriella Zopcsak for your invaluable assistance with all matters related to the psychology department.

Many thanks to Nan, Hao, Werner, Cemre, Shao-yang, and Chengyu for being exceptional companions in this journey – your presence has made this endeavour much more enjoyable. Additionally, Martin, thank you for your invaluable help in getting me acquainted with everything during my initial days in Munich; your support made the transition effortless and pleasant.

Finally, I would like to express my gratitude to my parents. Maa and Baba, your unwavering support and encouragement have been the bedrock of my journey. You have been the two strongest pillars of strength in my life, and I owe all that I am today to your love and guidance. To Avishek, thank you for standing by me through the ups and downs. Your ability to find humour amidst the chaos I sometimes brought back from work has been a constant source of comfort and strength. Your support and reassurance have been invaluable, and I am truly blessed to have you by my side.

List of publications

- **Mandal, A**., Röer, J. P., & Liesefeld, H. R. (2024). Auditory distractors are processed but do not interfere with visual search of any difficulty when sound is irrelevant. *Visual Cognition*, 1–17. https://doi.org/10.1080/13506285.2024.2397825

- **Mandal, A.**, Liesefeld, A. M., & Liesefeld, H. R. (2024). Tracking the misallocation and reallocation of spatial attention toward auditory stimuli. *Journal of Neuroscience*, 44(30). https://doi.org/10.1523/JNEUROSCI.2196-23.2024

- **Mandal, A.**, Liesefeld, A. M., & Liesefeld, H. R. (2024). The surprising robustness of visual search against concurrent auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *50*(1), 99. https://doi.org/10.1037/xhp0001168 – *APA Editor's Choice Article*

- Gaspelin, N., Lamy, D., Egeth, H. E., Liesefeld, H. R., Kerzel, D., **Mandal, A.**, ... & van Moorselaar, D. (2023). The distractor positivity component and the inhibition of distracting stimuli. *Journal of Cognitive Neuroscience*, 35(11), 1693-1715. https://doi.org/10.1162/jocn_a_02051

- Constant, M., **Mandal, A.**, Asanowicz, D., Yamaguchi, M., Gillmeister, H., Kerzel, D., … Liesefeld, H. R. (in principle acceptance). A multilab investigation into the N2pc as an indicator of attentional selectivity: Direct replication of Eimer (1996). *PCI RR (to be published in Cortex)*

- **Mandal, A.**, & Liesefeld, H. R. (2022). Visual pop-out search is robust to auditory distraction. *Journal of Vision*, 22(14), 3887-3887. https://doi.org/10.1167/jov.22.14.3887 (Conference abstract)

DECLARATION OF AUTHOR CONTRIBUTIONS

The surprising robustness of visual search against concurrent auditory distraction

Ananya Mandal: conceptualization (equal), data curation, formal analysis, investigation, methodology (equal), project administration, software, visualization, writing – original draft (equal), and writing – review and editing (equal).
Anna M. Liesefeld: conceptualization (supporting), methodology (supporting), supervision (supporting), writing–original draft (supporting), and writing–review and editing (supporting).
Heinrich R. Liesefeld: conceptualization (equal), formal analysis (supporting), methodology (equal), project administration (supporting), funding acquisition, resources, supervision, validation, visualization (supporting), writing – original draft (equal), and writing – review and editing (equal).

Auditory distractors are processed but do not interfere with visual search of any difficulty when sound is irrelevant

Ananya Mandal: conceptualization (equal), data curation, formal analysis, investigation (equal), methodology (equal), project administration, software, visualization, writing – original draft (supporting), and writing – review and editing (equal).
Jan Philipp Röer: conceptualization (equal), investigation (equal), methodology (equal), supervision (supporting), and writing–review and editing (supporting).
Heinrich R. Liesefeld: conceptualization (equal), formal analysis (supporting), methodology (equal), project administration (supporting), funding acquisition, resources, supervision, validation, visualization (supporting), writing – original draft, and writing – review and editing (equal).

Tracking the misallocation and reallocation of spatial attention toward auditory stimuli

Ananya Mandal: conceptualization (equal), data curation, formal analysis, investigation, methodology (equal), project administration, software, visualization, writing – original draft, and writing – review and editing (equal).
Anna M. Liesefeld: conceptualization (supporting), methodology (supporting), supervision (supporting), and writing–review and editing (supporting).
Heinrich R. Liesefeld: conceptualization (equal), formal analysis (supporting), methodology (equal), project administration (supporting), funding acquisition, resources, supervision, validation, visualization (supporting), writing – original draft (supporting), and writing – review and editing (equal).

______________________________ ______________________________

Ananya Mandal Heinrich R. Liesefeld