Short Research Article

Can You See What I Hear?

Detecting Changes in Multimodal Setting

Anna Conci[1,2], Merim Bilalić[3], and Robert Gaschler[1]

[1]FernUniversität in Hagen, Hagen, Germany
[2]Alpen-Adria-Universität Klagenfurt, Klagenfurt, Austria
[3]Northumbria University, Newcastle-upon-Tyne, United Kingdom

Abstract. Previous research on inattentional blindness (IB) has focused almost entirely on the visual modality. This study extends the paradigm by pairing visual with auditory stimuli. New visual and auditory stimuli were created to investigate the phenomenon of inattention in visual, auditory, and paired modality. The goal of the study was to assess to what extent the pairing of visual and auditory modality fosters the detection of change. Participants watched a video sequence and counted predetermined words in a spoken text. IB and inattentional deafness occurred in about 40% of participants when attention was engaged by this difficult (auditory) counting task. Most importantly, participants detected the changes considerably more often (88%) when the change occurred in both modalities rather than just one. One possible reason for the drastic reduction of IB or deafness in a multimodal context is that discrepancy between expected and encountered course of events proportionally increases across sensory modalities.

Keywords: inattentional blindness, inattentional deafness, multimodal change, attentional capacity

OPEN DATA

OPEN MATERIALS

People cannot believe they would likely fail to detect a well-visible change in the environment while looking at it, just because they are concentrating on something else. This phenomenon, called inattentional blindness (IB), has been known for decades (Mack & Rock, 1998; Neisser & Becklen, 1975; Simons & Chabris, 1999). The phenomenon takes place in everyday circumstances, for example, during driving. Given the omnipresence of the phenomenon, there is an auditory counterpart too – inattentional deafness. Inattentional deafness occurs when we fail to perceive an auditory event while listening to something else. But what happens when the modalities are combined, and the unexpected event is both visible and audible? The current investigation tests the extent to which participants can profit from changes in both modalities.

IB was first documented in an experiment by Neisser and Becklen (1975), who showed that half of the participants failed to perceive unexpected events in a video while paying attention to another video on the same screen. The study showed that gaze and attention are two separate entities, as participants were unable to perceive new stimuli despite looking at them directly. Posner (1980) showed that attention is not necessarily connected to the foveal structure of the visual system or to the movements of the eye. Mack and Rock (1998) named this phenomenon "IB" after confirming the results in their experiment, in which participants had to evaluate the length of a cross that was presented for 200 ms. A quarter of the participants did not notice that there was a small square presented in their fixation area. The authors concluded that there is no conscious perception without attention. Similarly, Simons and Chabris (1999) showed persistent IB with their famous gorilla video. Participants had to count the passes in a basketball game. In the middle of the video, a woman in a gorilla costume crosses the scene, a sequence that lasts for about 5 s. Almost half of the participants failed to notice the gorilla in the middle of the basketball game. Simons and Chabris also showed that IB depends on the overlap of the discrimination task and the changing object. Among the participants assigned to count the passes of the team wearing white T-shirts, IB was higher – compared to the group of participants paying attention to the team wearing black (the same color as the gorilla). It seems that it is easier to detect an unexpected event when it shares a basic visual feature with an object in the focus of the attention (Simons & Chabris, 1999).

Inattentional deafness, the auditory counterpart of IB, is less studied, although it was first documented in research on the cocktail party effect almost 70 years ago (Cherry, 1953). It is possible to recognize a story presented to one ear and to be unaware of a story presented simultaneously to the other ear. If an auditory stimulus changes only once

Experimental Psychology (2020), *67*(3), 186–193
https://doi.org/10.1027/1618-3169/a000487

or appears only for a certain time without being perceived by a person, then one speaks of inattentional deafness. In contrast, if a person does not perceive a constantly changing stimulus in the auditory scene, then this is called change deafness. The latter can occur even when the person has the expectation that there might be a change.

Vitevitch (2003) reported one of the first experiments on inattentional deafness where both the discrimination task and the unexpected event were auditory. Participants concentrating on a spoken word list were likely to fail to notice a change in speaker. Specifically, participants had to repeat hard and easy words as fast and accurately as possible. Almost half of them did not perceive the change in the voice speaking the words. Dalton and Fraenkel (2012) demonstrated inattentional deafness in a dynamic conversation setting. Participants were to listen to a discussion between two women, while ignoring one between two men (or the other way around). While the conversations were in progress, a man walked around the room, saying the phrase "I'm a gorilla." Among the participants assigned to follow the male conversation, only 10% were deaf to the unexpected event. However, inattentional deafness was observed among 70% of the participants instructed to listen to women's voices. Demonstrations of inattentional deafness are not limited to verbal material. Koreimann, Gula, and Vitouch (2014) showed that more than 50% of the participants did not hear an e-guitar solo in a classical piece of music when their attention was engaged by counting timpani beats.

Inattentional deafness and blindness are not restricted to the unimodal setting. A task designed to engage attention in one modality can lead to inattentional deafness (or blindness) in the other modality. For example, Macdonald and Lavie (2011) showed inattentional deafness under visual load. Participants missed a single tone while performing a visual task. They presented a cross for 150 ms, and the participants had to decide which arm was blue (low load condition) or which arm was longer (high load condition). Participants were more likely to miss the auditory stimulus under high visual load. Awareness of the task-unrelated tone was less pronounced in the high load visual attention task. This may speak to a shared pool of resource between visual and auditory perception. Raveh and Lavie (2015) investigated whether the visual load influences auditory perception if an auditory signal is presented several times during a visual search paradigm. They found a higher detection rate (for the sound) in the low-load condition (visual search task). In other words, the change was more likely to be detected when the visual task was not demanding. The results remained unchanged whether the answer was given immediately after the signal or at the end of the trial. In addition, they showed that visual perceptual load also influences auditory detection when participants were expecting an auditory signal during the visual search task. It seems that the capacity limits in perception under high perceptual load allow only the processing of the essential information – in this case, in vision and hearing – which supports the idea of a shared attentional capacity between both modalities.

Some studies suggest that change detection differs for change in source and change in content. Fenn et al. (2011) investigated inattentional deafness using a phone conversation. Since participants had to listen accurately to the message, they missed the change in their speaking partner. Therefore, at least, some unexpected changes in source are not detected automatically. Presumably during natural conversations, some voice characteristics are not monitored precisely all the time. Rather, people might direct attention toward such characteristics when expecting a change in who is talking or when identification of the speaker is crucial to process the message. Yet, the change detection rate increased drastically when one person was male and the other speaker was female. Thus, some changes in voice characteristics might draw attention automatically (see also Vitevitch & Donoso, 2011).

The studies reported so far varied the overlap between the target event (i.e., gorilla) and task-relevant features (black vs. white T-shirts, female vs. male voice) did so within a modality. In everyday life, it is uncommon to have input from only one modality at any given time. We are rarely exposed to solely unimodal stimuli and unimodal change. Therefore, we have chosen to investigate change detection in a multimodal setting testing IB with unimodal and multimodal stimuli. In order to investigate whether attention to off-task events is allocated separately in different modalities, researchers have studied multimodal changes. Wayand, Levin, and Varakin (2005) used an event in a video sequence that was visual *and* auditory. Over the course of the video, a woman crossed the scene in the background and scratched her fingernails on a blackboard. Approximately four out of 10 participants registered the bimodal stimulus. Sound volume made no significant difference. Talsma, Doty, Strowd, and Woldorff (2006) demonstrated that the attentional capacity within a modality is smaller than the attentional capacity between modalities. Accordingly, it is possible that the attentional modulation of sensory neural processing is at least partially independent for different modalities. In line with independent attention resources in different modalities, Talsma et al. (2006) suggest that attention devoted to one modality is not lacking in the other modality (e.g., Talsma et al., 2006). This view contrasts with the above-reported cross-modal studies (e.g., Macdonald & Lavie, 2011, showing inattentional deafness under visual load).

There is some evidence that there can be asymmetries with respect to whether attention engaged in one modality is lacking in the other. In the experiments of Pizzighello and Bressan (2008), participants were randomly assigned to one of five conditions. The study included a visual condition, two auditory conditions (comprehension and recall), and two dual conditions (comprehension and recall). The visual condition featured a visual discrimination task as well as an appearance of an unexpected visible object on the same screen. The auditory tasks were to either listen to a short story and answer short questions (comprehension) or recall as many words as possible from an orally presented word list (recall). The critical trial also contained a visual change. Participants in the dual condition were to count the repetitions of a visual event, and, as in the auditory condition, to listen to a short story or to memorize words. The rate of IB was higher when participants had both visual and auditory tasks, in contrast to the condition when they had only a visual discrimination task. Different than expected, there was no difference between the dual and auditory tasks. The results suggest that adding a visual task to an auditory one can engage attention more than vice versa. Paying attention to an auditory task can block the perception of an unexpected visual object. This suggests that people automatically reduce attention to the visual field when focusing on an auditory task.

Relevant to the question of what might trigger the detection of multimodal changes, Spence and Driver (1997) showed that an irrelevant cuing sound improves visual elevation judgments (target localization task). Participants were faster in deciding where the target was located when the uninformative cuing sound was on the same side as the target. It seems that irrelevant auditory cues cause rapid cross-modal shifts and can guide the spatial attention exogenously.

The above studies suggest that IB and deafness might be reduced when changes occur in two modalities. Arguably, in order to further constrain how change detection is driven by attention in different modalities, one should not merely show that multimodal changes have a (somewhat) higher detection rate than unimodal changes. A small advantage of the multimodal setup compared to the unimodal one might be taken to rule out an account claiming just one attentional resource – there must be at least some modality specificity involved. Yet, it would not permit an assertion of full independence of the attentional resources in different modalities. There could also be some mix between general and specific resources. So far, results do not clearly show the dependency between the attentional resources. Some researchers claim that the perception operates separately between the modalities and that the attentional capacity is, therefore, modality-specific (e.g., Duncan, Martens, & Ward, 1997). Some studies suggest that the resources are shared between the modalities (e.g., Raveh & Lavie, 2015; Rees, Frith, & Lavie, 2001).

Accordingly, in the current study, we aimed at comparing multimodal change detection rate against a benchmark of optimal use of information from each of the modalities. To investigate IB, we used an auditory discrimination task in three different conditions. In one video, there was only a visual change, while in the second video, there was a sole auditory change, and the third condition combined both changes, as both the visual appearance and the voice of the speaker changed. In contrast to the above studies, we made the unexpected change even more obvious. Rather than a short event lasting only for a few seconds, the video used for the task featured a change in the middle of the video, which lasted for the rest of the scene, totaling more than 48 s.

First of all, we wanted to replicate the phenomena of IB and inattentional deafness. Furthermore, comparing the multimodal change condition with the unimodal change conditions, we aimed at testing the extent to which participants could use the additional opportunities for change detection. If participants divided a uniform attentional resource across two modalities, combined changes should at best lead to a modest increase in detection rate. However, if participants were able to process the information in the two modalities with independent resources, they should be found to profit exhaustively from the added modality. Material and raw data are provided online (Conci, 2019).

Method

Participants

A total of 337 participants were tested with ages ranging from 18 to 60 years. In addition, there was a control group (no attention-engaging task provided, $n = 27$). We excluded participants who failed to give the correct answer in the discrimination task (see below) by a standard deviation (six words) or more (see the Materials section) and who answered *yes* to the control question five (see the Procedure section). Thus, the analysis was performed on the data of 295 participants (164 female) with a mean age of 35.17 years ($SD = 11.41$). The mean age in the conditions is listed in Table 1.

The sample size was informed by a pilot study with 45 participants (Conci, 2015). The effect size, as calculated by Cohen's w (Cohen, 1988) for measuring discrepancies between two proportions (given by the null hypothesis and

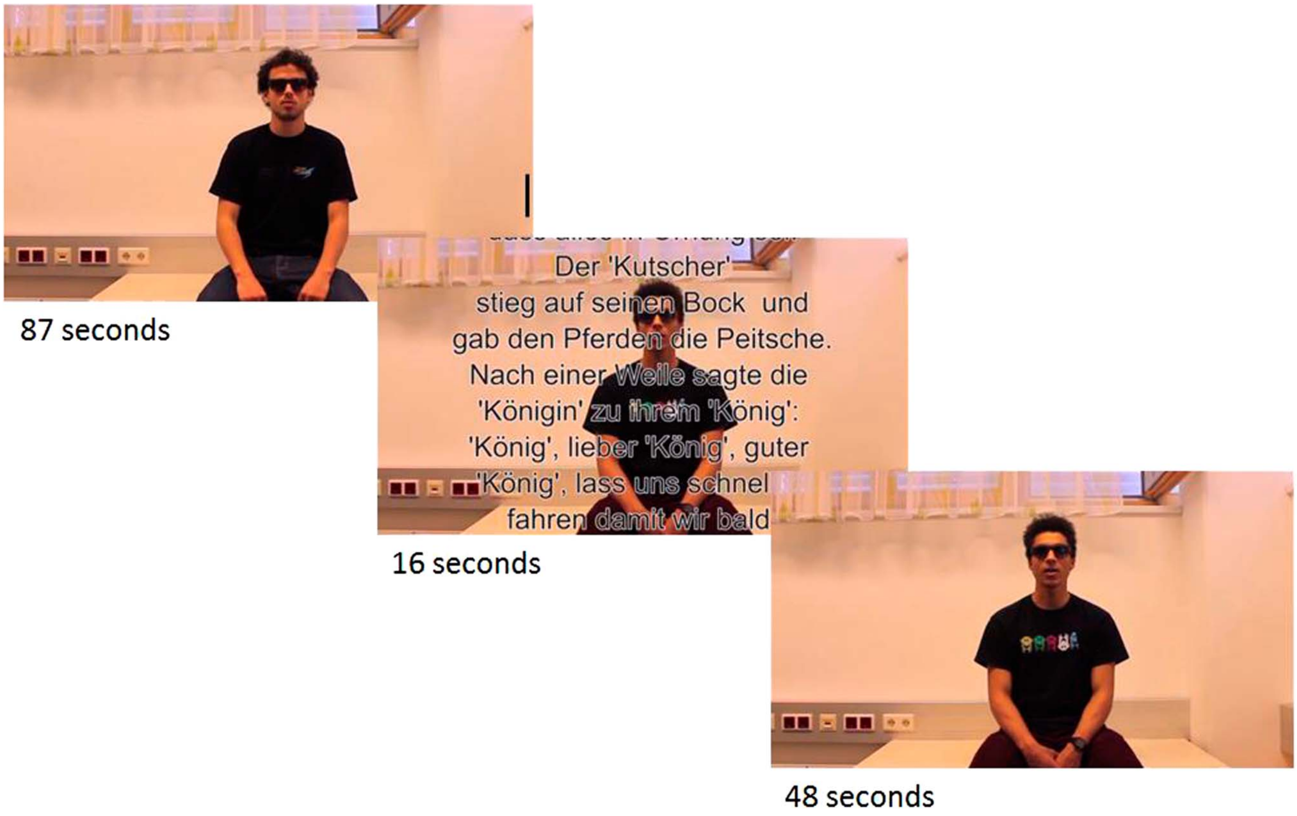

Figure 1. This figure shows three frames from videos 1 (visual condition) and 3 (multimodal condition). First, Actor 1 is visible. He recites the text passage used in the experiment in a seated position. After 87 s, part of the text is displayed on screen to distract the viewer from the change in narrator. After 16 s, the text scrolls completely upward, revealing Actor 2. Actor 2 recites the rest of the text; thus, the change is visible for the rest of the experiment, lasting for a total of 48 s.

the alternate), was 0.37 across all three groups, 0.30 for the difference between the auditory and multimodal conditions, and 0.26 between visual and multimodal. The required sample size for a power of 0.99 and an alpha error of 0.05 was 157 participants across all three conditions (w = .37), 238 for auditory and multimodal comparison (w = .30), and 317 participants for visual and multimodal conditions (w = .26; G*Power; Faul, Erdfelder, Lang, & Buchner, 2007). Although in the end our sample did not quite reach the upper limit of 317 participants due to exclusion criteria, we note that the sample was more than enough to detect a w = .26 effect with a .95 power and 0.05 alpha level (this would yield 229 participants necessary).

All participants had normal or corrected-to-normal vision and reported normal hearing.

Materials and Apparatus

For this experiment, we produced three different versions of a 148-s long video (retrievable under https://osf.io/sy8dq/, as well as the raw data). First, there is a person visible, seated, and speaking a predefined text (frequently used as a game). This text contains 42 instances in total of the German words *König* and *Königin* (king and queen). The discrimination task was to count these terms. After 87 s, a running text (part of the story) was displayed over the screen to catch the participant's attention. During this distraction, the change occurred (see Figure 1). In the first version, the speaker changed. In the second version, the voice changed. The third version combined both changes at the same time. To investigate whether the IB level decreases in a multimodal setting, we applied the same discrimination task in all conditions.

We tested 27 participants in the control group. Participants watched one of the three videos without counting the words *König* or *Königin*. The use of a control group allowed us to make sure that the event would be detected in the absence of the discrimination task. The procedure after the experiment was the same except for the reporting of the word count (see below). The change was detected correctly by 92% of the participants, suggesting that the video material is appropriate for

Table 1. Frequencies and mean age

	Visual	Auditory	Multimodal	Total
Mean age (*SD*)	35 (12)	35 (11)	36 (11)	35 (11)
Blind	41 (42%)	41 (41%)	12 (12%)	94 (32%)
Not blind	57 (58%)	59 (59%)	85 (88%)	201 (68%)
Total	98	100	97	295

Note. "Not blind" denotes the participants who reported seeing a change despite the discrimination task. Column headers "visual," "auditory," and "multimodal" indicate the condition. Inattentional blindness refers to participants failing to notice a change in the video. Additionally, we report here the mean age and standard deviation (*SD*) of each group.

comparing IB conditions. Two out of 27 persons missed the change in the visual condition without being given the discrimination task.

The videos were recorded with a Canon Rebel t2i camera and then presented on a 12.5″ Lenovo ThinkPad X220 (12.5″ screen).

Procedure

Data acquisition took place as part of 10 bachelor of science theses, all completed in the course of the same semester with exactly the same materials and instruction (see the Acknowledgments section). Participants were quasi-randomly assigned to one of the three video conditions. In particular, to ensure equal group sizes and avoid that the possibility that differences between the people signing up late for the experiment and those signing up early might confound the results, the first participant showing up for the experiment was assigned to the visual condition, the second one to the auditory condition, the third one to the multimodal condition, the fourth participant again to the visual condition, and so on. The experiment lasted for approximately 10 min. Participants were instructed to watch the video and to count the spoken words *König* (king) and *Königin* (queen).

After the video ended, the participants reported their word count. Following this, they were asked several questions adopted from the experiment by Simons and Chabris (1999): (1) Did you notice anything unusual? (2) Did something change? If yes, what and when? If the participants replied to this question in the negative, they were asked the following question: (3) Did you notice that the person/voice changed? (4) Have you participated in a similar experiment before? The control item (Wayand et al., 2005) was (5) Did you see the red star in the upper left corner? (No red star had been presented.) Following a positive reply, the participant was excluded from the analysis because this answer would cast doubt on whether the person had answered Question 3 truthfully. Six persons were excluded from the analysis as a result. Additionally, participants were asked whether they were familiar with the spoken text used for the discrimination task, the actor/s in the video, or the gorilla video from Simons and Chabris (1999).

Results

Twenty-five out of 27 participants in the control group were aware of the change in the video. Participants in the control group found it difficult to believe that the change in the video could remain unnoticed. Without the counting task, the change was clearly noticeable.

Table 1 shows the frequencies of the answers in the experimental conditions. The detection rate (change was detected successfully) in the multimodal condition (88%) was higher than in the unimodal conditions (58% and 59%).

For the main analysis, we used a χ^2 test. There is a significant difference in detection rate between the unimodal conditions and the multimodal condition; $\chi^2(1, 295) = 25.29$; $p < .001$, odds ratio = 0.2 (95% CI: 0.09, 0.4). Paired comparisons showed differences between the visual and the multimodal condition ($\chi^2(1, 195) = 21.38$; $p < .001$, Cohen's $w = .33$) as well as between the auditory and multimodal condition ($\chi^2(1, 197) = 20.52$; $p < .001$, $w = .32$). Cohen's w shows a medium effect in both cases (Cohen, 1988). There was no difference between the two unimodal changes. Note that age did not differ among the conditions (see Table 1).

We further ran an explorative analysis using the two unimodal conditions to derive a benchmark for assessing whether participants in the multimodal condition would exhaustively profit from the added option for detecting the change. Rather than just being higher by a small margin, the rate of change detection in the multimodal condition was ahead of the rate in the unimodal conditions to an extent which suggested that information in the two simultaneously presented modalities was used very efficiently to detect changes. We calculated the hypothetic rate of still being blind after two subsequent runs on one unimodal video (0.172). For this, we formed hypothetical pairs consisting of one person in the unimodal auditory condition and one person in the unimodal visual condition and determined the probability that neither of the two noticed the change; the likelihood that the participant in the visual condition remained blind was 0.418, while the likelihood that the participant in the auditory condition remained blind was 0.410. Multiplying these probabilities yielded the likelihood that neither of the participants in the hypothetical pair detected the change (0.172). This rate matched the rate of being blind in the multimodal condition (0.124). Thus, the likelihood that one person in the multimodal condition remained blind was about as low as the likelihood that both of the paired persons would have remained blind. The participants in the multimodal condition apparently used the two options to detect a change very efficiently.

Discussion

The phenomenon of IB has been known for decades. In contrast, its auditory counterpart, inattentional deafness, is not well investigated. In this study, we combined the two phenomena, IB and inattentional deafness, to investigate how detection rate changes in a multimodal setting. The unexpected event was a change in the experimental video that was either visual, auditory, or audiovisual. This combination of condition and stimulus resulted in a cross-modal, a unimodal, and a multimodal setting.

The study showed that IB is not restricted to setups where attended information and unexpected change occur within the same modality. Our two unimodal conditions showed similar IB rates. Importantly, we observed that participants in the multimodal condition could detect the change very efficiently (keeping up with pairs of participants from the unimodal conditions). The odd ratio of 0.2 shows that it is five times more unlikely for a person to be blind in the multimodal conditions than in the unimodal condition. Our results differ from Rees et al. (2001) who claimed that IB occurs within one modality but not between different modalities. Yet, Raveh and Lavie (2015) showed that visual load has an impact on the detection rate of an auditory stimulus. They claimed that attentional capacity is shared between vision and hearing. In line with this, we found IB with the auditory discrimination task and the visual change. Yet, diverging from a shared resource view, the rate of change detection was high for multimodal changes.

One way to reconcile these two findings might be to further elaborate on mechanisms underlying the change detection. For instance, changes cooccurring in two modalities might lead to a stronger surprise, and this in turn could lead to verbal knowledge about the surprising change that can be reported in the interview (cf., Rünger & Frensch, 2008). Horstmann (2015) suggested that the stream of changes in sensory stimulation is constantly being compared against expectations and that discrepancies (i.e., surprises) draw attention. Research on attention capture by surprise suggests that discrepancy detection uses combined features (e.g., Schützwohl, 1998, Exp. 3; see Horstmann, 2015, for an overview). Surprise is evoked by unexpected (schema-discrepant) events, and its intensity is determined by the degree of schema-discrepancy (cf., Reisenzein, Horstmann, & Schützwohl, 2019). Unexpected changes cooccurring in two modalities can yield larger overall discrepancy. Thus, when a discrepancy in the auditory and the visual channel occurs at the same time, this might lead to an overall higher level of discrepancy, which, in turn, yields a higher likelihood of drawing attention and leads to explicit knowledge about the unexpected event (cf., Haider & Frensch, 2009).

Work in the auditory domain is consistent with a bidirectional surprise-attention link. Attending change-relevant objects upfront might increase the strength of the surprise signal. Change detection rates in auditory scenes have been found to increase if participants were made to pay attention to change-relevant objects beforehand (cf., Eramudugolla, Irvine, McAnally, Martin, & Mattingley, 2005; Irsik, Vanden Bosch der Nederlanden, & Snyder, 2016). There are prior studies in line with better change detection if two modalities are involved. For instance, Santangelo, Ho, and Spence (2008) found evidence that bimodal cues (audio-tactile ones in their experiment) capture visuospatial attention better than unimodal cues. Wayand et al. (2005) found IB in a visual setting (42.9% missed the unexpected event) and also in the multimodal condition (between 50% and 60% inattentional blind persons). In these studies, the added modality only added partially to the probability of detecting the change. Our multimodal results, however, suggest that under some conditions, the added modality can be exploited exhaustively, yielding detection rates that are as high as can be expected from combining the unimodal rates – rather than just showing a small advantage compared to the unimodal variant.

Future studies should isolate the factors that determine whether the added modality can be used partially versus fully. Opening one route to pursuing this, Spence (2010) claimed that (in contrast to unisensory signals) multimodal cues (i.e., audiovisual cues) capture spatial attention more efficiently when the unimodal signals appear at approximately the same position. Furthermore, scrutinizing the differences between the studies, duration of presentation might be a relevant variable. In other IB studies, the unexpected stimulus has been present for just a few seconds, leaving less time to detect the change compared to the setup used in our study. For instance, Simons and Chabris (1999) showed the person in the gorilla costume for 5 s and in a supplemental experiment for 9 s. IB occurred in both experiments. Yet, the change in the experiment from Wayand et al. (2005) lasted for 30 s in a video length of 45 s. Sixty percent of the participants missed the woman in the video scratching her nails on a chalkboard. In our study, the unexpected event persisted for more than 50 s, until the very end of the video.

On the one hand, comparing different studies suggests that IB can occur with short-lived changes as well as sustained ones. On the other hand, it is as yet an open issue whether duration determines if the profit taken from a multimodal as compared to a unimodal change is partial versus exhaustive. Conceivably, additional time could be used to allocate attention to one modality or the other successively. Alternatively, options to detect the change in one modality versus the other might be harvested simultaneously by automatic bottom-up processes. The current study demonstrates that exhaustive usage of an added modality is possible. Future work will have to further detail how this constrains the mechanisms involved in change detection in IB setups. Similarly, the consequences for applications involving safety have to be elaborated further. Santangelo et al. (2008), for example, recommend a bimodal warning signal (audio-tactile) for drivers. Our study suggests that it would also be helpful to have audiovisual alerts to avoid accidents.

References

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*, 975–979. https://doi.org/10.1121/1.1907229

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: L. Erlbaum Associates.

Conci, A. (2015). *"Ich sehe was, das du nicht hörst…" Inwiefern verändert sich die Wahrnehmung bei gepaarten visuellen und auditiven Stimuli?* (Unpublished master's thesis). Klagenfurt, Austria: Universität Klagenfurt.

Conci, A. (2019). *Detecting changes in multimodal setting*. Retrieved from https://osf.io/sy8dq

Dalton, P., & Fraenkel, N. (2012). Gorillas we have missed: Sustained inattentional deafness for dynamic events. *Cognition*, *124*, 367–372. https://doi.org/10.1016/j.cognition.2012.05.012

Duncan, J., Martens, S. & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, *387*, 808–810. https://doi.org/10.1038/42947

Eramudugolla, R., Irvine, D. R. F., McAnally, K. I., Martin, R. L., & Mattingley, J. B. (2005). Directed attention eliminates "change deafness" in complex auditory scenes. *Current Biology*, *15*, 1108–1013. https://doi.org/10.1016/j.cub.2005.05.051

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https://doi.org/10.3758/bf03193146

Fenn, K. M., Shintel, H., Atkins, A. S., Skipper, J. I., Bond, V. C., & Nusbaum, H. C. (2011). When less is heard than meets the ear: Change deafness in a telephone conversation. *Quarterly Journal of Experimental Psychology*, *64*, 1442–1456. https://doi.org/10.1080/17470218.2011.570353

Haider, H., & Frensch, P. A. (2009). Conflicts between expected and actually performed behavior lead to verbal report of incidentally acquired sequential knowledge. *Psychological Research*, *73*, 817–834. https://doi.org/10.1007/s00426-008-0199-6

Horstmann, G. (2015). The surprise-attention link: A review. *Annals of the New York Academy of Sciences*, *1339*(1), 106–115. https://doi.org/10.1111/nyas.12679

Irsik, V. C., Vanden Bosch der Nederlanden, C. M., & Snyder, J. S. (2016). Broad attention to multiple individual objects may facilitate change detection with complex auditory scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1806–1817. https://doi.org/10.1037/xhp0000266

Koreimann, S., Gula, B., & Vitouch, O. (2014). Inattentional deafness in music. *Psychological Research*, *78*, 304–312. https://doi.org/10.1007/s00426-014-0552-x

Macdonald, J. S., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention Perception,* & Psychophysics, *73*, 1780–1789. https://doi.org/10.3758/s13414-011-0144-4

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: The MIT Press.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, *7*, 480–494. https://doi.org/10.1016/0010-0285(75)90019-5

Pizzighello, S., & Bressan, P. (2008). Auditory attention causes visual inattentional blindness. *Perception*, *37*, 859–866. https://doi.org/10.1068/p5723

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

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

Rees, G., Frith, C., & Lavie, N. (2001). Processing of irrelevant visual motion during performance of an auditory attention task. *Neuropsychologia*, *39*, 937–949. https://doi.org/10.1016/s0028-3932(01)00016-1

Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The cognitive-evolutionary model of surprise: A review of the evidence. *Topics in Cognitive Science*, *11*(1), 50–74. https://doi.org/10.1111/tops.12292

Rünger, D., & Frensch, P. A. (2008). How incidental sequence learning creates reportable knowledge: The role of unexpected events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1011–1026. https://doi.org/10.1037/a0012942

Santangelo, V., Ho, C., & Spence, C. (2008). Capturing spatial attention with multisensory cues. *Psychonomic Bulletin & Review*, *15*, 398–403. https://doi.org/10.3758/pbr.15.2.398

Schützwohl, A. (1998). Surprise and schema strength. *Journal of Experimental Psychology:Learning, Memory, and Cognition*, *24*, 1182–1199. https://doi.org/10.1037/0278-7393.24.5.1182

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*, 1059–1074. https://doi.org/10.1068/p281059

Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences*, *1191*(1), 182–200. https://doi.org/10.1111/j.1749-6632.2010.05440.x

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1–22. https://doi.org/10.3758/bf03206843

Talsma, D., Doty, T. J., Strowd, R., & Woldorff, M. G. (2006). Attentional capacity for processing concurrent stimuli is larger across sensory modalities than within a modality. *Psychophysiology*, *43*, 541–549. https://doi.org/10.1111/j.1469-8986.2006.00452.x

Vitevitch, M. S. (2003). Change deafness: The inability to detect changes between two voices. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 333–342. https://doi.org/10.1037/0096-1523.29.2.333

Vitevitch, M. S., & Donoso, A. (2011). Processing of indexical information requires time: Evidence from change deafness. *The Quarterly Journal of Experimental Psychology*, *64*, 1484–1493. https://doi.org/10.1080/17470218.2011.578749

Wayand, J. F., Levin, D. T., & Varakin, D. A. (2005). Inattentional blindness for a noxious multimodal stimulus. *The American Journal of Psychology*, *118*, 339–352.

History
Received July 15, 2019
Revision received June 10, 2020
Accepted June 10, 2020
Published online September 7, 2020

Acknowledgments

We are grateful to the students Jana Bolemant, Oliver Brandt, Christina Danam, Petra Gantner, Stefanie Gellner, Isabella Jocher, Sonja Riegger, Julia Schmid, Andreas-Horst Thurau, Nawina Walker, and Anna Ziske for helping with data acquisition.

Open Data

Material and raw data are provided online (Conci, 2019). https://osf.io/sy8dq/

ORCID
Anna Conci
https://orcid.org/0000-0002-6777-9122

Anna Conci
FernUniversität in Hagen
Fakultät für Psychologie
Universitätsstraße 33
58084 Hagen
Germany
anna.conci@fernuni-hagen.de