Double Flash Illusions: Current Findings and Future Directions

Julian Keil*

Biological Psychology, Christian-Albrechts-Universität zu Kiel, Kiel, Germany

Twenty years ago, the first report on the sound-induced double flash illusion, a visual illusion induced by sound, was published. In this paradigm, participants are presented with different numbers of auditory and visual stimuli. In case of an incongruent number of auditory and visual stimuli, the influence of auditory information on visual perception can lead to the perception of the illusion. Thus, combining two auditory stimuli with one visual stimulus can induce the perception of two visual stimuli, the so-called fission illusion. Alternatively, combining one auditory stimulus with two visual stimuli can induce the perception of one visual stimulus, the so-called fusion illusion. Overall, current research shows that the illusion is a reliable indicator of multisensory integration. It has also been replicated using different stimulus combinations, such as visual and tactile stimuli. Importantly, the robustness of the illusion allows the widespread use for assessing multisensory integration across different groups of healthy participants and clinical populations and in various task setting. This review will give an overview of the experimental evidence supporting the illusion, the current state of research concerning the influence of cognitive processes on the illusion, the neural mechanisms underlying the illusion, and future research directions. Moreover, an exemplary experimental setup will be described with different options to examine perception, alongside code to test and replicate the illusion online or in the laboratory.

Keywords: double flash illusion, multisensory integration, crossmodal influence, perception, congruence, sound-induced flash illusion

THE SOUND-INDUCED FLASH ILLUSION

Multisensory integration is a fundamental perceptual process, by which information arriving from different senses is combined to a unified percept, and numerous studies showed that multisensory integration is beneficial for perception. For example, redundant multisensory information reduces reaction times (Miller, 1982; Cappe et al., 2009; Pomper et al., 2014). Moreover, the cocktail party effect, in which auditory perception is supported by visual cues, indicates that multiple streams of information can support decoding relevant information (Zion-Golumbic et al., 2013). However, multisensory information can also have detrimental effects, in which the perception of one sensory modality is affected by conflicting information from a second modality. Often, these incongruent information streams will be perceptually integrated, resulting in subjective illusions. Examples of such illusions due to incongruent multisensory information include the McGurk Effect (McGurk and MacDonald, 1976) or the Ventriloquist Effect (Choe et al., 1975). In both examples, visual information influences auditory perception. Interestingly, the unisensory information underlying these illusions is salient and easily perceived in isolation.
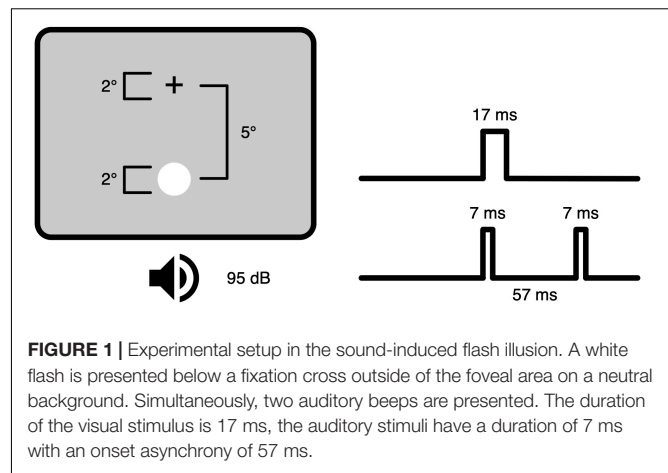
The above-mentioned examples illustrate an influence of visual information on auditory perception. However, the reverse influence has also been observed: Twenty years ago, Shams, Kamitani and Shimojo (Shams et al., 2000) published an – in their words – "striking visual illusion" indicating that visual perception can be influenced by other sensory modalities. The authors described that pairing a single visual stimulus with multiple auditory stimuli will lead to the illusory perception of multiple visual stimuli. This phenomenon was later coined the "sound-induced illusory flash effect" or "sound-induced flash illusion" (SIFI) (Bhattacharya et al., 2002; Shams et al., 2005a). The SIFI is a highly reliable effect that has been replicated in numerous studies. Interestingly, it is not specific to audiovisual stimuli, but the visual illusion can be induced by tactile stimuli as well (Violentyev et al., 2005). Yet, despite the overall robustness, there is a large inter-individual variability in the susceptibility to the illusion (de Haas et al., 2012). Across samples, the average likelihood of the illusion has often been reported to be around 50% for audiovisual (e.g., Keil et al., 2014) and visuotactile (e.g., Lange et al., 2013) stimuli. Thus, pairing one visual stimulus, which is easily detected in isolation and readily distinguished from two visual stimuli, with two auditory or tactile stimuli, renders visual perception bistable, but the individual likelihood to perceive the illusion varies.

A large number of empirical studies have explored the phenomenological change in perception, the underlying computational principles and the neural mechanisms associated with the perception of the illusion. After 20 years, it is time to summarize the current state of the research in a comprehensive review. Therefore, the aims of this review are to outline the proposed explanations of the SIFI, to compile current lines of research and to provide an update on future directions. Moreover, it will describe an example procedure to induce the SIFI alongside reproducible code for easy replication of the behavioral phenomenon.

AN EXAMPLE SIFI EXPERIMENT

The original publication on the SIFI (Shams et al., 2000) described an experiment comprising white disks subtending 2° at an eccentricity of 5° on black background spaced 50 ms in time and undefined beeps spaced 57 ms in time. Critically, observers reported multiple flashes, when one disk was accompanied by multiple beeps. A follow-up publication (Shams et al., 2002) further specified the luminance of the visual stimuli as 108 cd/m^2 with a duration of 17 ms and the loudness of the auditory stimuli as 95 dB SPL at 3500 Hz with a duration of 7 ms (**Figure 1**). Thus, both the visual and auditory stimuli were highly salient. Interestingly, the authors claim that the specific stimulus characteristics should not influence the illusion perception, but the likelihood to perceive the illusion can be influenced by various stimulus and task characteristics (see section "Principles of Multisensory Integration and the SIFI").

An examination of the optimal temporal spacing of the first and second auditory and visual stimuli indicated that this illusion declined above an offset of 70 ms. However, changing the offset between auditory and visual stimuli within ±70 ms did not

FIGURE 1 | Experimental setup in the sound-induced flash illusion. A white flash is presented below a fixation cross outside of the foveal area on a neutral background. Simultaneously, two auditory beeps are presented. The duration of the visual stimulus is 17 ms, the auditory stimuli have a duration of 7 ms with an onset asynchrony of 57 ms.

affect the strength of the illusion (Shams et al., 2002). Therefore, multiple experiments used stimulus offsets of 180 ms to control for reaction tendencies not due to multisensory integration (Mishra et al., 2007; Michail and Keil, 2018). As a measure of perception, participants are usually asked to report the number of perceived flashes using a N-alternative forced-choice task, where N refers to the possible visual stimuli used in the experiment (e.g., 3-alternative forced-choice task: 0, 1, and 2). From this, the susceptibility to the illusion can be computed, as described in the next section ("Examining the Participants' Perception").

Based on this information, an example experiment was built using the open-source online experimental platform lab.js[1]. This experiment comprises 10 ms 1000 Hz tones in an individually chosen loudness in combination with visual stimuli subtending 1.6° at 4.1° eccentricity at a viewing distance of 60 cm with a duration of 16 ms (one screen refresh cycle at 60 Hz) on a neutral gray background. Importantly, due to the limitations of the lab.js platform, the onsets of the first auditory and visual stimuli are asynchronous, with the visual stimulus leading by one screen refresh cycle. The experiment can be tested at https://open-lab.online/test/sound-induced-flash-illusion-example/, and the code is available at https://github.com/juliankeil/SIFI_Example alongside an example analysis script for R.

In the last years, LED monitors have replaced CRT monitors in many laboratories. LED monitors in general provide less exact onset timing and stimulus duration (Elze, 2010; Cooper et al., 2013; Ghodrati et al., 2015). It is therefore important to consider the limits of the experimental setup and to design the experiment accordingly. Whereas it is recommended to use monitors and sound cards with precise timing in order to exactly control the onset and duration of the auditory and visual stimuli, empirical research shows that the SIFI is tolerant to asynchronies within ±70 ms (Shams et al., 2002). Extending psychophysical research to online studies will likely lead to less controlled stimulation environments due to the use of various combinations of hardware and software (Bridges et al., 2020). Therefore, presenting uni- and multisensory control conditions in laboratory and online studies is critically important, as these stimuli offer the possibility

---

[1]https://labjs.readthedocs.io/en/latest/

to check whether the participants can differentiate single and multiple auditory and visual stimuli and are actually reporting their perception.

Examining the Participants' Perception

The presentation of different combinations of flashes and beeps can either result in the so-called fission or fusion illusion (Andersen et al., 2004; Kaposvari et al., 2014). The fission illusion refers to the SIFI as described in the original publications by Shams et al. (2000, 2001, 2002). Here, two auditory stimuli are paired with one visual stimulus, which can result in the perception of two visual stimuli. The fusion illusion occurs following the presentation of one auditory stimulus together with two visual stimuli. Here, the illusion consists of the perception of only one visual stimulus.

Examining the participants' subjective perception of the incongruent multisensory stimuli can be accomplished by asking the participants to rate their perception using a forced-choice task. From this rating, the individual susceptibility to the illusion can be computed. One often-used straightforward approach to this is the computation of the perception rate, i.e., the fraction of responses to a given stimulus combination indicating a certain perception relative to the number of presented stimuli (e.g., number of times the participant reported "2" relative to the number of presentations of one visual and two auditory stimuli). Alternatively, measures of perception based on the signal detection theory have been proposed (Watkins et al., 2006; Whittingham et al., 2014). Therein, information from different congruent and incongruent stimulus combinations is combined to compute sensitivity (d') and response criteria. Vanes et al. (2016) describe how information from congruent trials with two beeps in combination with two flashes and incongruent trials with two beeps combined with one flash can be used in the analysis of the fission illusion. Importantly, participants report the number of perceived flashes and the authors consider "false alarms" as the illusion. Thus, the response of "2" in congruent trials (2 flashes-2 beeps) is a "hit," and the response of "1" is a "miss." Accordingly, the response of "2" in incongruent trials (1 flash-2 beeps) is a "false alarm" (i.e., fission illusion), and the response of "1" is a "correct rejection." All these values are then considered relative to the number of presented trials in each condition. From these values, sensitivity can be computed as

$$d' = z(\text{hit rate}) - z(\text{false alarm rate}),$$

with z as the inverse of the standard normal cumulative distribution function. The criterion can be computed as,

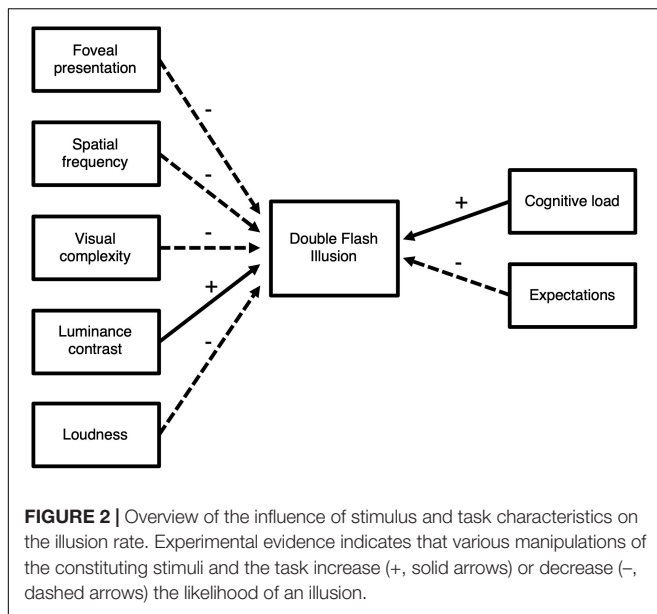$$\ln(\beta) = [z(\text{false alarm rate})^2 - z(\text{hit rate})^2]/2$$

Similarly, information from congruent trials with one beep in combination with one flash, and incongruent trials with one beep and two flashes can be used in the analysis of the fusion illusion. In this case, the response of "1" in congruent trials (1 flash-1 beep) is a "hit," and the response of "2" is a "miss." Accordingly, the response of "1" in incongruent trials (2 flashes-1 beep) is a "false alarm" (i.e., fusion illusion), and the response of "2" is a "correct rejection."

In summary, presenting various combinations of one and two auditory and visual stimuli allows examining the fission and fusion illusion. The exact stimulus properties appear to be less critical. However, the presentation of the visual stimuli in the periphery and the temporal spacing of stimuli within ±70 ms are important. From the different stimulus categories, the response rate, the sensitivity, and the criterion can be computed as outcome parameters.

PRINCIPLES OF MULTISENSORY INTEGRATION AND THE SIFI

When we integrate multisensory information and perceive our environment, we have to solve two problems. On the one hand, we need to decide whether two signals come from a common source and integrate the signals accordingly. Over the course of a wide range of studies, three basic principles of multisensory integration have been established that guide our perception: the spatial principle, the temporal principle, and the principle of inverse effectiveness (Stein and Meredith, 1993; Stein et al., 2014). In short, these principles state that multisensory integration is strongest when the input modalities are spatially concordant, temporally aligned, and when the neural responses to the presented stimuli are weak. On the other hand, we need to estimate the reliability of the signals with respect to a given feature. In addition to the three principles of multisensory integration, the modality appropriateness hypothesis has been proposed (Welch and Warren, 1980). This hypothesis has been extended in a maximum-likelihood-estimation framework, which suggests that information from each sensory modality is weighted based on its relative reliability (Ernst and Bülthoff, 2004). Similarly, the information reliability hypothesis proposes the dominance of the modality providing the most reliable information (Andersen et al., 2004). With respect to the SIFI, the auditory system has a higher temporal resolution than the visual system. Because the auditory modality is more reliable, it should dominate the overall percept in the SIFI. Finally, Andersen et al. (2004) argue, that all these principles "should be considered as factors which contribute to the relative dominance of each modality and not as all-or-nothing principles." At the same time, these authors highlight the role of cognitive processes such as attention in the SIFI and question the automatic multisensory integration. As a further extension of the maximum-likelihood-estimation framework, Shams et al. (2005b) compared the behavior of human participants to an ideal observer model and found that participants used Bayesian inference to decide whether, to what degree, and how to integrate the audiovisual information.

With respect to the reliability of the input modalities, it should be noted that whereas presenting the visual stimulus in the fovea reduces the illusion rate (Shams et al., 2001), minor changes in size, eccentricity or luminance result in a similar perception of the illusion (e.g., Keil et al., 2014; Balz et al., 2016a; Michail and Keil, 2018). However, characteristics of the visual stimulus do appear to influence the illusion.

FIGURE 2 | Overview of the influence of stimulus and task characteristics on the illusion rate. Experimental evidence indicates that various manipulations of the constituting stimuli and the task increase (+, solid arrows) or decrease (–, dashed arrows) the likelihood of an illusion.

For example, increased spatial frequency and visual complexity of the visual stimulus reduce the illusion perception, whereas increased luminance contrast has the opposite effect (Takeshima and Gyoba, 2013; Gyoba and Takeshima, 2015; Pérez-Bellido et al., 2015). Moreover, reducing the loudness of the auditory stimuli appears to reduce the illusion rate (Andersen et al., 2004; **Figure 2**).

In the SIFI, the constituting stimuli in the SIFI are presented in close temporal and spatial proximity. Moreover, compared to the auditory system, the temporal resolution of the visual system is relatively low (Vanrullen, 2016). In the SIFI, visual perception of short flashes is influenced by concurrent information from sensory modalities with a higher temporal resolution, such as auditory or tactile stimuli. Thus, multisensory integration in the SIFI follows the basic principles summarized above.

Cognitive Influences on Multisensory Integration

The implicit assumption in studies on the principles of multisensory integration is that properties of the stimuli and the perceptual system guide conscious perception, and that this perception can be modeled using computational processes. However, Andersen et al. (2004) highlight the role of attention in the perception of the SIFI and recent behavioral studies underline that the perception of the SIFI can be influenced by cognitive factors.

Using the fusion and the fission illusion, Andersen et al. (2004) examined the influence of task instructions on behavior. They found that the integration of auditory and visual information was not automatic but depended on whether participants were instructed to count beeps or flashes. They thus interpreted their findings as support for the directed attention hypothesis, stating that the attended modality dominates perception. Using a task-independent modulation of cognitive
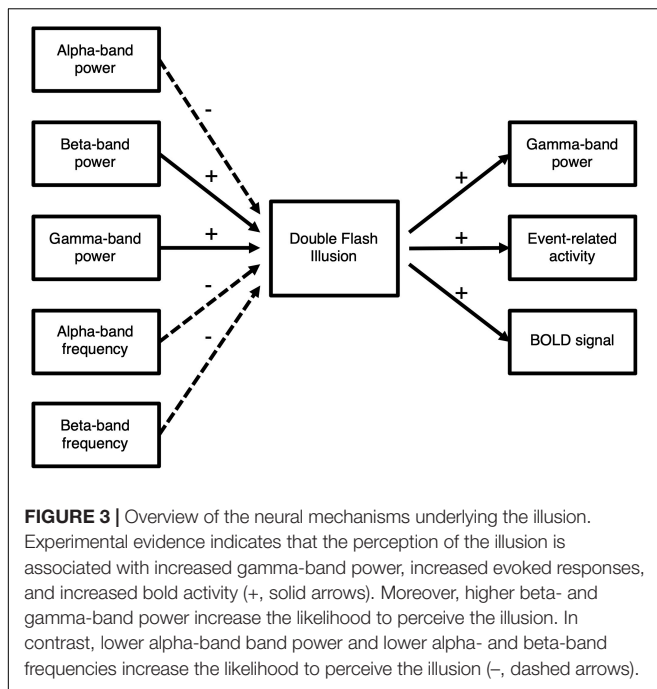
resources, Michail and Keil (2018) examined the influence of cognitive load on the perception of the SIFI. The authors found that increased cognitive load induced by an n-back task leads to higher illusion rates. They argue that their findings provide strong evidence that audiovisual integration can be modulated by the amount of available cognitive resources and it therefore argues against a pre-attentive account of multisensory integration. A recent study examined the influence of expectations regarding the presented stimuli on the perception of the SIFI (Wang et al., 2019). In short, the authors show that expectations regarding the proportion of SIFI trials shape perception, indicating an influence of task-related cognitive processes.

Taken together, the available evidence supports the idea that multisensory integration in the SIFI is not an automatic and rigid process, but that the stimulus characteristics, task instructions and cognitive processes such as attention and expectations shape multisensory integration (**Figure 2**).

NEURAL MECHANISMS

The cognitive mechanisms summarized above suggest that the influence of auditory information on visual perception should occur at a later, decision-level stage rather than on an early sensory processing stage. Similarly, based on findings from the ventriloquist illusion, Rohe and Noppeney (2015) argue that unisensory stimuli are processed under the assumption of independence in early sensory processing stages, and that assumptions regarding the reliabilities of the signals are taken into account at higher processing stages. This idea is in line with findings from electroencephalography (EEG) experiments on the SIFI, which indicate that the perception of the SIFI is associated with increased gamma-band power. Gamma-band power has been interpreted as a signature of multisensory integration (Senkowski et al., 2005). Importantly, the increased gamma-band power related to the perception of the SIFI occurred relatively late, i.e., after the initial stimulus processing (Bhattacharya et al., 2002; Mishra et al., 2007; Balz et al., 2016a). Similarly, magnetencephalography (MEG) and EEG studies found increased evoked responses at longer latencies (Shams et al., 2001; Keil et al., 2014). Further support for a central role of higher-order cortical areas in the SIFI comes from studies using transcranial magnetic stimulation (TMS). Stimulating the angular gyrus resulted in a reduced likelihood to perceive the illusion and thus reduced multisensory integration (Kamke et al., 2012; Hamilton et al., 2013).

In contrast, some MEG and EEG studies also found earlier modulations of evoked responses and neural oscillations, indicating crossmodal influences on the level of early sensory cortical areas (Shams et al., 2005a; Mishra et al., 2007; Lange et al., 2011; Balz et al., 2016b). Similarly, an fMRI study found increased BOLD activity in the primary visual cortex (Watkins et al., 2006). However, the authors also note an involvement of the superior temporal sulcus and the inferior colliculi.

FIGURE 3 | Overview of the neural mechanisms underlying the illusion. Experimental evidence indicates that the perception of the illusion is associated with increased gamma-band power, increased evoked responses, and increased bold activity (+, solid arrows). Moreover, higher beta- and gamma-band power increase the likelihood to perceive the illusion. In contrast, lower alpha-band band power and lower alpha- and beta-band frequencies increase the likelihood to perceive the illusion (–, dashed arrows).

Taken together, these findings on the neural substrate of the SIFI underline a multisensory integration process involving early and late processing stages at different hierarchies, in which crossmodal influences can influence perception at multiple stages (Senkowski et al., 2008; Keil and Senkowski, 2018; **Figure 3**).

NEURAL ACTIVITY INFLUENCES MULTISENSORY PERCEPTION

In everyday life, we constantly and effortlessly integrate the sensory inputs of our environment. In this regard, we are often able to predict future events. How we predict a future event has been discussed extensively at the theoretical level (Summerfield and Egner, 2009), and according to the predictive coding framework (Friston, 2005), we continuously uses available information to predict forthcoming events and to reduce sensory uncertainty. Importantly, this framework posits that top-down processes prior to stimulus presentation act upon primary sensory cortices. Recent studies have shown that sensory information is continuously sampled and that low-frequency oscillatory activity likely mediates this sampling (Busch and Vanrullen, 2010; Vanrullen et al., 2011). In addition, human studies have demonstrated that the amplitude and phase of oscillatory activity, as well as neural connectivity in cortical networks, relates to cognitive processes, sensory representation, attentional selection, and dynamic routing of information (Van Dijk et al., 2008; Mazaheri et al., 2009; Jensen and Mazaheri, 2010). These findings are in remarkable agreement with the results of animal studies (Fries, 2005; Schroeder and Lakatos, 2009).

Similar to unisensory stimulation, a number of studies have indicated that ongoing oscillations in cortical networks affect the processing of forthcoming multisensory stimuli. An MEG study examined the differences in oscillatory power and functional connectivity between SIFI trials in which the fission illusion was perceived and those trials in which only one visual stimulus was perceived (Keil et al., 2014). The authors report that increased beta-band power in the STG preceded a multisensory illusion, and that increased beta-band functional connectivity between STG and primary auditory cortex was related to illusion perception on a single trial level. Similarly, Kaiser et al. (2019) analyzed single-trial power prior to the SIFI and found that increased beta- and gamma-band power in occipital electrodes predicted the illusion perception. Moreover, using visual and tactile stimuli, Lange et al. (2013) found that reduced alpha-band power in visual cortical areas and increased gamma-band power in parietal and temporal cortical areas preceded the illusion, and the authors argued that this reflects cortical excitability. Two further studies highlighted the role of the alpha-band phase as a temporal window of integration (TWI) for the shaping of audiovisual perception (Cecere et al., 2015; Keil and Senkowski, 2017). The former authors found a correlation between the individual alpha-band frequency and illusion rate, which indicates that alpha-band oscillations provide a TWI in which the crossmodal influence could induce an illusion. Importantly, modulating the individual alpha-band frequency using transcranial alternating current stimulation modulated the probability of an illusion perception. Keil and Senkowski (2017) confirmed the relationship between the individual alpha-band frequency and the SIFI perception rate and localized this effect to the occipital cortex. Importantly, a recent EEG study on the auditory and tactile induced double flash illusion further confirmed the relationship between neural oscillations and the TWI (Cooke et al., 2019). However, whereas the authors replicated the relationship between the individual alpha-band frequency the SIFI, the TWI in the tactile-induced flash illusion was defined by the individual beta-band frequency. Thus, it appears that the neural oscillations recorded in visual cortex reflect task-depended functional connectivity networks. Interestingly, an MRI study found a correlation between the individual susceptibility of the illusion and the gray matter volume in the primary visual cortex (de Haas et al., 2012).

In agreement with studies on unisensory perception, a number of studies indicate that neural oscillations influence multisensory processing. Therein, alpha-band power indicates excitability in primary sensory areas and the phase of neural oscillations provides a TWI for crossmodal influence. Increased beta-band and gamma-band power in multisensory cortical areas might indicate increased readiness to integrate information (**Figure 3**).

CLINICAL APPLICATIONS AND FUTURE DIRECTIONS

Empirical research over the last 20 years has established the SIFI as a robust and reliable tool to study multisensory integration

in various settings and contexts. From this research, we can distill a standardized experimental setup and we have a mature sense of the average response to incongruent multisensory stimulation across various populations. Moreover, findings on the neuroscientific studies give us a detailed insight into the neural mechanisms underlying the illusion and their influence on perception. Based on this empirical background, we can now start to look into the factors influencing conscious perception. This includes brain states as well as cognitive factors and inter-individual differences in healthy and clinical populations.

In the last decade, empirical findings have indicated that the brain state influences information processing (Van Dijk et al., 2008; Busch and Vanrullen, 2010; Samaha et al., 2017). Importantly, multisensory studies have shown that cortical activity in one sensory area can influence information processing in other sensory cortical areas (Lakatos et al., 2007; Kayser et al., 2008; Mercier et al., 2015). Simultaneously, recent electrophysiological studies point toward a central role of neural oscillations and functional connectivity for multisensory integration and conscious perception (Lange et al., 2014; Keil and Senkowski, 2018). However, only few studies have examined causal manipulations of neural activity, for example using brain stimulation (Kamke et al., 2012; Hamilton et al., 2013; Cecere et al., 2015). Future studies on the SIFI could now be directed at changes in local cortical activity and network configuration associated with the illusion, and subsequent directed manipulation of these parameters.

Recent studies examined the influence of cognitive factors on multisensory integration. However, the role of attention therein is hotly debated (Andersen et al., 2004, 2009; Talsma et al., 2010; Macaluso et al., 2016; Tang et al., 2016), and the role of expectations for multisensory integration has only recently been examined (Gau and Noppeney, 2015; Wang et al., 2019). Behavioral data suggest that participants can adjust the susceptibility to the SIFI according to the expectations of stimulus timing (Chan et al., 2020). Moreover, concurrent emotional stimuli reduce multisensory integration, but information on the underlying neural mechanisms is scarce (Maiworm et al., 2012). Finally, the role of cognitive load for multisensory perception has only received little attention (Michail and Keil, 2018). Thus, future studies could aim to examine multisensory integration using a standardized experimental setup while manipulating cognitive influences. Concurrent examination of neural activity can then indicate the involvement of various cortical areas therein.

In addition, a number of studies have shown age-related changes in multisensory integration and perception (Murray et al., 2016; Hirst et al., 2018, 2020). Overall, these studies indicate that multisensory integration is plastic and dynamic across development from early life to adulthood. With respect to the SIFI, Hernández et al. (2019) found that the susceptibility to the illusion increased with age and declining cognitive status. Similar results indicate that cognitive impairment increases the susceptibility to the illusion across longer SOAs (Chan et al., 2015). It has been shown that cognitive performance varies as a function of the time of day (Monk et al., 1997; Carrier and Monk, 2000), but

the influence of time on multisensory integration remains largely unexplored. Preliminary evidence indicates that neural oscillations, especially in the alpha band, vary with increased time on task (Benwell et al., 2019), yet the effects of these changes on multisensory integration have not been explored. Similar to the influence of cognitive processes on multisensory integration, using a standardized experimental setup in different age groups, at different times of the day, and at multiple replications can indicate developmental changes in multisensory integration and the influence of environmental factors therein.

In recent years, a number of studies have examined aberrant multisensory integration in various clinical populations. In autism spectrum disorder (ASD), deficits in perceiving the temporal relationship between different sensory inputs might impair multisensory integration. Accordingly, there is evidence of increased TWI in ASD, which is, however, more pronounced in speech stimuli than in the simpler SIFI (Stevenson et al., 2014). Similarly, along a psychosis continuum, increased proneness to the SIFI has been suggested to be linked to reduced temporal sensitivity. In a healthy population, Ferri et al. (2018) found increased susceptibility to the illusion in participants with high schizotypal scores, related to an increased TWI. Similarly, patients suffering from schizophrenia show an increased susceptibility to the illusion at longer SOAs, which is indicative of a lager TWI (Haß et al., 2016). However, no difference in the illusion perception appears at a short SOA, although changes in neural activity suggest aberrant multisensory processing in schizophrenia (Balz et al., 2016b).

Research on the SIFI uncovered basic neural and behavioral processes underlying multisensory integration. We now have a robust and reliable tool at our disposal to examine multisensory integration and perception in various settings. Future studies can now build upon these findings to further examine the influence of cognitive and emotional processes, development and aging, as well as fatigue and time on multisensory integration and perception.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Andersen, T. S., Tiippana, K., Laarni, J., Kojo, I., and Sams, M. (2009). The role of visual spatial attention in audiovisual speech perception. *Speech Commun.* 51, 184–193. doi: 10.1016/j.specom.2008.07.004

Andersen, T. S., Tiippana, K., and Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Brain Res. Cogn. Brain Res.* 21, 301–308. doi: 10.1016/j.cogbrainres.2004.06.004

Balz, J., Keil, J., Romero, Y. R., Mekle, R., Schubert, F., Aydin, S., et al. (2016a). GABA concentration in superior temporal sulcus predicts gamma power and perception in the sound-induced flash illusion. *Neuroimage* 125, 724–730. doi: 10.1016/j.neuroimage.2015.10.087

Balz, J., Roa Romero, Y., Keil, J., Krebber, M., Niedeggen, M., Gallinat, J., et al. (2016b). Beta/gamma oscillations and event-related potentials indicate aberrant multisensory processing in schizophrenia. *Front. Psychol.* 7:1896. doi: 10.3389/fpsyg.2016.01896

Benwell, C. S. Y., London, R. E., Tagliabue, C. F., Veniero, D., Gross, J., Keitel, C., et al. (2019). Frequency and power of human alpha oscillations drift systematically with time-on-task. *Neuroimage* 192, 101–114. doi: 10.1016/j.neuroimage.2019.02.067

Bhattacharya, J., Shams, L., and Shimojo, S. (2002). Sound-induced illusory flash perception: role of gamma band responses. *Neuroreport* 13, 1727–1730. doi: 10.1097/00001756-200210070-00007

Bridges, D., Pitiot, A., MacAskill, M. R., and Peirce, J. (2020). The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PsyArXiv [Preprint]* doi: 10.31234/osf.io/d6nu5

Busch, N. A., and Vanrullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16048–16053. doi: 10.1073/pnas.1004801107

Cappe, C., Thut, G., Romei, V., and Murray, M. M. (2009). Selective integration of auditory-visual looming cues by humans. *Neuropsychologia* 47, 1045–1052. doi: 10.1016/j.neuropsychologia.2008.11.003

Carrier, J., and Monk, T. H. (2000). Circadian rhythms of performance: new trends. *Chronobiol. Int.* 17, 719–732. doi: 10.1081/cbi-100102108

Cecere, R., Rees, G., and Romei, V. (2015). Individual differences in alpha frequency drive crossmodal illusory perception. *Curr. Biol.* 25, 231–235. doi: 10.1016/j.cub.2014.11.034

Chan, J., Kaiser, J., Brandl, M., Matura, S., Prvulovic, D., Hogan, M., et al. (2015). Expanded temporal binding windows in people with mild cognitive impairment. *Curr. Alzheimer Res.* 12, 61–68. doi: 10.2174/1567205012666141218124744

Chan, J. S., Connolly, S. K., and Setti, A. (2020). The number of stimulus-onset asynchronies affects the perception of the sound-induced flash illusion in young and older adults. *Multisens. Res.* 31, 175–190. doi: 10.1163/22134808-00002605

Choe, C. S., Welch, R. B., Gilford, R. M., and Juola, J. F. (1975). The "ventriloquist effect": visual dominance or response bias? *Percept. Psychophys.* 18, 55–60. doi: 10.3758/bf03199367

Cooke, J., Poch, C., Gillmeister, H., Costantini, M., and Romei, V. (2019). Oscillatory properties of functional connections between sensory areas mediate cross-modal illusory perception. *J. Neurosci.* 39, 5711–5718. doi: 10.1523/JNEUROSCI.3184-18.2019

Cooper, E. A., Jiang, H., Vildavski, V., Farrell, J. E., and Norcia, A. M. (2013). Assessment of OLED displays for vision research. *J. Vis.* 13:16. doi: 10.1167/13.12.16

de Haas, B., Kanai, R., Jalkanen, L., and Rees, G. (2012). Grey matter volume in early human visual cortex predicts proneness to the sound-induced flash illusion. *Proc. R. Soc. B Biol. Sci.* 279, 4955–4961. doi: 10.1098/rspb.2012.2132

Elze, T. (2010). Achieving precise display timing in visual neuroscience experiments. *J. Neurosci. Methods* 191, 171–179. doi: 10.1016/j.jneumeth.2010.06.018

Ernst, M. O., and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169. doi: 10.1016/j.tics.2004.02.002

Ferri, F., Venskus, A., Fotia, F., Cooke, J., and Romei, V. (2018). Higher proneness to multisensory illusions is driven by reduced temporal sensitivity in people with high schizotypal traits. *Conscious. Cogn.* 65, 263–270. doi: 10.1016/j.concog.2018.09.006

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480. doi: 10.1016/j.tics.2005.08.011

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

Gau, R., and Noppeney, U. (2015). How prior expectations shape multisensory perception. *Neuroimage* 124, 876–886. doi: 10.1016/j.neuroimage.2015.09.045

Ghodrati, M., Morris, A. P., and Price, N. S. C. (2015). The (un)suitability of modern liquid crystal displays (LCDs) for vision research. *Front. Psychol.* 6:303. doi: 10.3389/fpsyg.2015.00303

Gyoba, J., and Takeshima, Y. (2015). Spatial frequency modulates the degree of illusory second flash perception. *Multisens. Res.* 28, 1–10. doi: 10.1163/22134808-00002484

Hamilton, R. H., Wiener, M., Drebing, D. E., and Coslett, H. B. (2013). Gone in a flash: manipulation of audiovisual temporal integration using transcranial magnetic stimulation. *Front. Psychol.* 4:571. doi: 10.3389/fpsyg.2013.00571

Haß, K., Sinke, C., Reese, T., Roy, M., Wiswede, D., Dillo, W., et al. (2016). Enlarged temporal integration window in schizophrenia indicated by the double-flash illusion. *Cogn. Neuropsychiatry* 22, 145–158. doi: 10.1080/13546805.2017.1287693

Hernández, B., Setti, A., Kenny, R. A., and Newell, F. N. (2019). Individual differences in ageing, cognitive status, and sex on susceptibility to the sound-induced flash illusion: a large-scale study. *Psychol. Aging* 34, 978–990. doi: 10.1037/pag0000396

Hirst, R. J., McGovern, D. P., Setti, A., Shams, L., and Newell, F. N. (2020). What you see is what you hear: 20 years of research using the Sound-Induced Flash Illusion. *PsyArXiv [Preprint]* doi: 10.31234/osf.io/7m586

Hirst, R. J., Stacey, J. E., Cragg, L., Stacey, P. C., and Allen, H. A. (2018). The threshold for the McGurk effect in audio-visual noise decreases with development. *Sci. Rep.* 8:12372. doi: 10.1038/s41598-018-30798-8

Jensen, O., and Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* 4:186. doi: 10.3389/fnhum.2010.00186

Kaiser, M., Senkowski, D., Busch, N. A., Balz, J., and Keil, J. (2019). Single trial prestimulus oscillations predict perception of the sound-induced flash illusion. *Sci. Rep.* 9:5983. doi: 10.1038/s41598-019-42380-x

Kamke, M. R., Vieth, H. E., Cottrell, D., and Mattingley, J. B. (2012). Parietal disruption alters audiovisual binding in the sound-induced flash illusion. *Neuroimage* 62, 1334–1341. doi: 10.1016/j.neuroimage.2012.05.063

Kaposvari, P., Bognár, A., Csibri, P., Utassy, G., and Sáry, G. (2014). Fusion and fission in the visual pathways. *Physiol. Res.* 63, 625–635.

Kayser, C., Petkov, C. I., and Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cereb. Cortex* 18, 1560–1574. doi: 10.1093/cercor/bhm187

Keil, J., Müller, N., Hartmann, T., and Weisz, N. (2014). Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. *Cereb. Cortex* 24, 1278–1288. doi: 10.1093/cercor/bhs409

Keil, J., and Senkowski, D. (2017). Individual alpha frequency relates to the sound-induced flash illusion. *Multisens. Res.* 30, 565–578. doi: 10.1163/22134808-00002572

Keil, J., and Senkowski, D. (2018). Neural oscillations orchestrate multisensory processing. *Neuroscientist* 24, 609–626. doi: 10.1177/1073858418755352

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., and Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292. doi: 10.1016/j.neuron.2006.12.011

Lange, J., Keil, J., Schnitzler, A., Van Dijk, H., and Weisz, N. (2014). The role of alpha oscillations for illusory perception. *Behav. Brain Res.* 271, 294–301. doi: 10.1016/j.bbr.2014.06.015

Lange, J., Oostenveld, R., and Fries, P. (2011). Perception of the touch-induced visual double-flash illusion correlates with changes of rhythmic neuronal activity in human visual and somatosensory areas. *Neuroimage* 54, 1395–1405. doi: 10.1016/j.neuroimage.2010.09.031

Lange, J., Oostenveld, R., and Fries, P. (2013). Reduced occipital alpha power indexes enhanced excitability rather than improved visual perception. *J. Neurosci.* 33, 3212–3220. doi: 10.1523/JNEUROSCI.3755-12.2013

Macaluso, E., Noppeney, U., Talsma, D., Vercillo, T., Hartcher-O'Brien, J., and Adam, R. (2016). The curious incident of attention in multisensory integration: bottom-up vs. top-down. *Multisens. Res.* 29, 557–583. doi: 10.1163/22134808-00002528

Maiworm, M., Bellantoni, M., Spence, C., and Röder, B. (2012). When emotional valence modulates audiovisual integration. *Attent. Percept. Psychophys.* 74, 1302–1311. doi: 10.3758/s13414-012-0310-3

Mazaheri, A., Nieuwenhuis, I., Van Dijk, H., and Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Hum. Brain Mapp.* 30, 1791–1800. doi: 10.1002/hbm.20763

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746–748. doi: 10.1038/264746a0

Mercier, M. R., Molholm, S., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H., and Foxe, J. J. (2015). Neuro-oscillatory phase alignment drives speeded multisensory response times: an electro-corticographic investigation. *J. Neurosci.* 35, 8546–8557. doi: 10.1523/JNEUROSCI.4527-14. 2015

Michail, G., and Keil, J. (2018). High cognitive load enhances the susceptibility to non-speech audiovisual illusions. *Sci. Rep.* 8:11530. doi: 10.1038/s41598-018-30007-6

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* 14, 247–279. doi: 10.1016/0010-0285(82)90010-x

Mishra, J., Martinez, A., Sejnowski, T. J., and Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *J. Neurosci.* 27, 4120–4131. doi: 10.1523/JNEUROSCI.4912-06. 2007

Monk, T. H., Buysse, D. J., Reynolds, C. F., Berga, S. L., Jarrett, D. B., Begley, A. E., et al. (1997). Circadian rhythms in human performance and mood under constant conditions. *J. Sleep Res.* 6, 9–18. doi: 10.1046/j.1365-2869.1997. 00023.x

Murray, M. M., Lewkowicz, D. J., Amedi, A., and Wallace, M. T. (2016). Multisensory processes: a balancing act across the lifespan. *Trends Neurosci.* 39, 567–579. doi: 10.1016/j.tins.2016.05.003

Pérez-Bellido, A., Ernst, M. O., Soto-Faraco, S., and López-Moliner, J. (2015). Visual limitations shape audio-visual integration. *J. Vis.* 15:5. doi: 10.1167/15. 14.5

Pomper, U., Brincker, J., Harwood, J., Prikhodko, I., and Senkowski, D. (2014). Taking a call is facilitated by the multisensory processing of smartphone vibrations, sounds, and flashes. *PLoS One* 9:e103238. doi: 10.1371/journal.pone. 0103238.s002

Rohe, T., and Noppeney, U. (2015). Cortical hierarchies perform bayesian causal inference in multisensory perception. *PLoS Biol.* 13:e1002073. doi: 10.1371/ journal.pbio.1002073.s009

Samaha, J., Iemi, L., and Postle, B. R. (2017). Prestimulus alpha-band power biases visual discrimination confidence, but not accuracy. *Conscious. Cogn.* 54, 47–55. doi: 10.1016/j.concog.2017.02.005

Schroeder, C. E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18. doi: 10.1016/j.tins. 2008.09.012

Senkowski, D., Schneider, T., Foxe, J., and Engel, A. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci.* 31, 401–409. doi: 10.1016/j.tins.2008.05.002

Senkowski, D., Talsma, D., Herrmann, C. S., and Woldorff, M. G. (2005). Multisensory processing and oscillatory gamma responses: effects of spatial selective attention. *Exp. Brain Res.* 166, 411–426. doi: 10.1007/s00221-005-2381-z

Shams, L., Iwaki, S., Chawla, A., and Bhattacharya, J. (2005a). Early modulation of visual cortex by sound: an MEG study. *Neurosci. Lett.* 378, 76–81. doi: 10.1016/j.neulet.2004.12.035

Shams, L., Kamitani, Y., and Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature* 408:788. doi: 10.1038/35048669

Shams, L., Kamitani, Y., and Shimojo, S. (2002). Visual illusion induced by sound. *Brain Res. Cogn. Brain Res.* 14, 147–152. doi: 10.1016/s0926-6410(02)00069-1

Shams, L., Kamitani, Y., Thompson, S., and Shimojo, S. (2001). Sound alters visual evoked potentials in humans. *Neuroreport* 12, 3849–3852. doi: 10.1097/ 00001756-200112040-00049

Shams, L., Ma, W. J., and Beierholm, U. (2005b). Sound-induced flash illusion as an optimal percept. *Neuroreport* 16, 1923–1927. doi: 10.1097/01.wnr.0000187634. 68504.bb

Stein, B. E., and Meredith, M. A. (1993). *The Merging of the Senses.* Cambridge, MA: MIT Press.

Stein, B. E., Stanford, T. R., and Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nat. Rev. Neurosci.* 15, 520–535. doi: 10.1038/nrn3742

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., et al. (2014). Multisensory temporal integration in autism spectrum disorders. *J. Neurosci.* 34, 691–697. doi: 10.1523/JNEUROSCI.3615-13.2014

Summerfield, C., and Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* 13, 403–409. doi: 10.1016/j.tics.2009.06.003

Takeshima, Y., and Gyoba, J. (2013). Complexity of visual stimuli affects visual illusion induced by sound. *Vis. Res.* 91, 1–7. doi: 10.1016/j.visres.2013.07.013

Talsma, D., Senkowski, D., Soto-Faraco, S., and Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* 14, 400–410. doi: 10.1016/j.tics.2010.06.008

Tang, X., Wu, J., and Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neurosci. Biobehav. Rev.* 61, 208–224. doi: 10.1016/j.neubiorev.2015.11.002

Van Dijk, H., Schoffelen, J.-M., Oostenveld, R., and Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J. Neurosci.* 28, 1816–1823. doi: 10.1523/JNEUROSCI.1853-07.2008

Vanes, L. D., White, T. P., Wigton, R. L., Joyce, D., Collier, T., and Shergill, S. S. (2016). Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia. *Psychiatry Res.* 245, 58–65. doi: 10.1016/j.psychres.2016.08.016

Vanrullen, R. (2016). Perceptual cycles. *Trends Cogn. Sci.* 20, 723–735. doi: 10.1016/ j.tics.2016.07.006

Vanrullen, R., Busch, N. A., Drewes, J., and Dubois, J. (2011). Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. *Front. Psychol.* 2:60. doi: 10.3389/fpsyg.2011.00060

Violentyev, A., Shimojo, S., and Shams, L. (2005). Touch-induced visual illusion. *Neuroreport* 16, 1107–1110. doi: 10.1097/00001756-200507130-00015

Wang, A., Sang, H., He, J., Sava-Segal, C., Tang, X., and Zhang, M. (2019). Effects of cognitive expectation on sound-induced flash illusion. *Perception* 48, 1214–1234. doi: 10.1177/0301006619885796

Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D., and Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *Neuroimage* 31, 1247–1256. doi: 10.1016/j.neuroimage.2006.01.016

Welch, R. B., and Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychol. Bull.* 88, 638–667. doi: 10.1037/0033-2909. 88.3.638

Whittingham, K. M., McDonald, J. S., and Clifford, C. W. G. (2014). Synesthetes show normal sound-induced flash fission and fusion illusions. *Vis. Res.* 105, 1–9. doi: 10.1016/j.visres.2014.08.010

Zion-Golumbic, E., Cogan, G. B., Schroeder, C. E., and Poeppel, D. (2013). Visual input enhances selective speech envelope tracking in auditory cortex at a "cocktail party." *J. Neurosci.* 33, 1417–1426. doi: 10.1523/JNEUROSCI.3675-12. 2013