CrossMark

On the interplay of visuospatial and audiotemporal dominance: Evidence from a multimodal kappa effect

Karin M. Bausenhart[1] · Katrina R. Quinn[2]

Abstract When participants judge multimodal audiovisual stimuli, the auditory information strongly dominates temporal judgments, whereas the visual information dominates spatial judgments. However, temporal judgments are not independent of spatial features. For example, in the kappa effect, the time interval between two marker stimuli appears longer when they originate from spatially distant sources rather than from the same source. We investigated the kappa effect for auditory markers presented with accompanying irrelevant visual stimuli. The spatial sources of the markers were varied such that they were either congruent or incongruent across modalities. In two experiments, we demonstrated that the spatial layout of the visual stimuli affected perceived auditory interval duration. This effect occurred although the visual stimuli were designated to be task-irrelevant for the duration reproduction task in Experiment 1, and even when the visual stimuli did not contain sufficient temporal information to perform a two-interval comparison task in Experiment 2. We conclude that the visual and auditory marker stimuli were integrated into a combined multisensory percept containing temporal as well as task-irrelevant spatial aspects of the stimulation. Through this multisensory integration process, visuospatial information affected even temporal judgments, which are typically dominated by the auditory modality.

✉ Karin M. Bausenhart
karin.bausenhart@uni-tuebingen.de

[1] Department of Psychology, University of Tübingen, Schleichstrasse 4, 72076 Tübingen, Germany

[2] Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, Tübingen, Germany

Perception requires combining sensory information from different sources into a detailed representation of our external environment. Far from being a direct translation of physical stimulus input into a perceptual impression, what we perceive is rather constructed out of multiple sources of information from different sensory systems. Crucially, in order to produce a consistent and comprehensive percept of our environment, information from these multiple sensory sources has to be combined and potential conflicts between the different modalities have to be resolved. In this process of *multisensory integration*, information from different senses is typically not weighted equally, and the weights assigned to the different senses seem to depend on the reliability of the information for a particular judgment (Driver & Spence, 2000; Ernst & Banks, 2002; Ernst & Bülthoff, 2004; Welch & Warren, 1980).

For example, when making temporal judgments, the auditory system typically provides more reliable information than the visual system (Gamache & Grondin, 2010; Repp & Penel, 2002; Ulrich, Nitschke, & Rammsayer, 2006; Wearden, Edwards, Fakhri, & Percival, 1998). As a consequence, when information from the auditory and the visual modality is presented concurrently (i.e., bimodally), the auditory information typically overweighs the visual information, and thus dominates the combined audiovisual percept (Aschersleben & Bertelson, 2003; Fendrich & Corballis, 2001; Morein-Zamir,

Soto-Faraco, & Kingstone, 2003; Shams, Kamitani, & Shimojo, 2000; Shipley, 1964; Vroomen & de Gelder, 2004).[1]

Most importantly for the present study, audition strongly dominates over the visual modality also regarding perceived duration (Bausenhart, De la Rosa, & Ulrich, 2014; de Haas, Cecere, Cullen, Driver, & Romei, 2013; De la Rosa & Bausenhart, 2013; Klink, Montijn, & van Wezel, 2011; Sarmiento, Shore, Milliken, & Sanabria, 2012; Shi, Chen, & Müller, 2010; Walker & Scott, 1981). For example, Walker and Scott required their participants to reproduce the duration of auditory and visual intervals of identical duration. During unimodal stimulus presentations, auditory intervals were typically reproduced as longer than visual intervals. Crucially, when stimuli were presented bimodally, the perceived duration of the combined stimulus was similar to the perceived duration of the auditory unimodal stimulus, and significantly prolonged as compared to the perceived duration of the visual unimodal intervals. Thus, the auditory temporal information had overruled the visual temporal information.

This basic pattern of results has been confirmed and extended by several more recent studies that employed physically conflicting stimulus durations and controlled for potential confounds such as response or attentional biases (Bausenhart et al., 2014; Klink et al., 2011; Romei, De Haas, Mok, & Driver, 2011). For example, when conflicting durations are presented to the different modalities (e.g., by combining auditory and visual intervals of physically different duration), perceived duration is typically determined largely by the auditory rather than the visual input, even if participants are required to ignore the auditory stimulation and judge the visual duration only. At the same time, sensitivity for duration discrimination is not impaired, but rather enhanced by providing multisensory duration information, even if the auditory information is incongruent with the visual information (Bausenhart et al., 2014; De la Rosa & Bausenhart, 2013). In contrast, when participants have to judge the duration of auditory intervals and ignore concurrently presented visual intervals, duration judgments are typically not affected by incongruent visual stimulation, or are affected to a relatively small extent (Bausenhart et al., 2014; Klink et al., 2011). Even when participants are required to judge the combined bimodal interval (i.e., when both modalities are designated to be task-relevant), auditory information typically contributes more strongly to the

combined audiovisual percept than does visual information, and vision only dominates the percept when the auditory stimulation is very noisy or of weak intensity (Hartcher-O'Brien, Di Luca, & Ernst, 2014; Walker & Scott, 1981). In conclusion, auditory temporal information generally strongly dominates over visual temporal information, determining the combined audiovisual duration percept to a much stronger extent than the visual information does.

In contrast to the auditory system's typical dominance in temporal perception, spatial judgments show the reverse dominance, with the visual system affecting the spatial percept more than the auditory system (Bertelson & Radeau, 1981; Howard & Templeton, 1966; Slutsky & Recanzone, 2001; Vroomen, Bertelson, & de Gelder, 2001; Welch & Warren, 1980). Specifically, when participants have to localize the spatial source of an auditory stimulus, their responses are typically strongly biased toward the spatial source of a co-occurring visual stimulus. Therefore, it is assumed that in the process of multisensory integration, the perceived location of an auditory stimulus gets "pulled" toward the perceived location of the co-occurring visual stimulus. This phenomenon, termed the *spatial ventriloquist effect* (Howard & Templeton, 1966), is typically ascribed to the superior spatial resolution of the visual in comparison to the auditory system. Consistent with this notion, the influence of the auditory input on the combined spatial percept only increases when the quality of the visuospatial input is strongly degraded, and thus, its reliability for localization decreases (Alais & Burr, 2004; Battaglia, Jacobs, & Aslin, 2003).

In spite of these auditory and visual dominances that depend on the task domain at hand, temporal and spatial information per se are not processed independently of each other (Nijhawan & Khurana, 2010; Oliveri, Koch, & Caltagirone, 2009). In fact, there are considerable interactions between the temporal and the spatial domain, which can lead to marked perceptual distortions. A classical example for such distortions can be found in the *kappa effect* (Cohen, Hansel, & Sylvester, 1953; Price-Williams, 1954), in which the perceived duration of an empty temporal interval increases with increasing spatial distance between the two marker stimuli indicating the beginning and end of the interval (Cohen et al., 1953; Huang & Jones, 1982; Price-Williams, 1954; Sarrazin, Giraudo, Pailhous, & Bootsma, 2004). For example, in the study of Price-Williams (1954), participants were required to reproduce the temporal interval between two successive light flashes, which were presented from either the same location or from two different locations separated horizontally by 32 in. The participants' mean reproduced durations were significantly longer when the marker flashes were presented from different locations than when they were presented from the same location. Thus, the spatial distance between the marker stimuli had distorted their perceived temporal "distance." This

---

[1] Here and throughout the article, we employ the term "dominance" in the sense of "taking a predominant role in the unified percept," rather than "completely taking over" the resulting percept. In fact, several studies suggest that the relative weighting of the modalities in multisensory integration does not take place in an all-or-none fashion, but that there is rather a smooth transition of the relative weights of each modality, on the basis of their relative sensory reliabilities (e.g., Ernst & Bülthoff, 2004).

kappa effect has not only been demonstrated in the visual, but also in the tactile (Grondin, Kuroda, & Mitsudo, 2011; Suto, 1952) and, even though less reliably, the auditory (Grondin & Plourde, 2007; Sarrazin, Giraudo, & Pittenger, 2007) modality.[2] However, to our knowledge, the kappa effect has not been investigated in bimodal settings so far.

Yet, such an investigation would be highly informative in order to foster our understanding of the principles underlying multimodal integration. Specifically, several studies have shown that the typically observed patterns of dominance depend on the stimulus intensity or input quality in the separate modalities, presumably based on the signal-to-noise ratio of the modalities' inputs (Ernst & Banks, 2002; Hartcher-O'Brien et al., 2014; Walker & Scott, 1981). Beyond such rather low-level stimulus features, however, other factors, such as the experimental context and the type of available information, also might affect how much each modality contributes to a combined multisensory percept (Sarmiento et al., 2012). In the present experiments, we therefore investigated how visual and auditory temporal information is integrated under varying conditions of (task-irrelevant) spatial information. On the one hand, as we outlined above, audiovisual *perceived duration* is typically strongly dominated by auditory information, especially when the auditory signals are clearly perceivable. Therefore, any visual input should not—or should only to a relatively small extent—affect perceived duration (Bausenhart et al., 2014; Klink et al., 2011; Walker & Scott, 1981). On the other hand, the *perceived location* of a stimulus is dominated strongly by visual input (Bertelson & Radeau, 1981; Howard & Templeton, 1966; Vroomen et al., 2001), and perceived location is known to distort perceived duration (Cohen et al., 1953; Price-Williams, 1954). Therefore, it is also conceivable that the presence of visuospatial information might impair the typical auditory dominance in temporal perception.

The present experiments were designed to investigate such a potential interplay between audiotemporal and visuospatial information with regard to perceived duration. Specifically, participants were required to perform either reproductions (Exp. 1) or comparative judgments (Exp. 2) of the perceived duration of auditorily marked empty intervals. These could either be presented alone ("unimodal" condition) or be accompanied by visual stimulation ("bimodal congruent"). Importantly, in these conditions, interval markers were presented either from the same spatial source or from different spatial sources, and we assessed whether kappa effects would emerge under these conditions. Crucially, in an additional

("bimodal incongruent") condition, interval markers were spatially separated only in the visual but not in the (task-relevant) auditory modality. We investigated whether this visuospatial information would affect auditory perceived duration (i.e., whether it would lead to a prolonged perceived duration; i.e., a kappa effect), despite the typical auditory dominance in the temporal task domain.

Experiment 1

In this experiment, participants performed temporal reproductions of empty intervals. The magnitude of the kappa effect was assessed in three different conditions (cf. Fig. 1). First, in *unimodal* trials, participants reproduced three different temporal intervals (500, 900, and 1,600 ms), which were marked by two brief auditory markers from either the same spatial source or two different spatial sources. Second, in *bimodal congruent* trials, the two auditory marker stimuli were accompanied by synchronous and spatially congruent visual stimuli. Third, and theoretically most important, in *bimodal incongruent* trials, the two auditory markers were always presented from the same spatial source (e.g., always from the left side). The synchronously presented first and second visual stimuli originated from either the same spatial source as the auditory stimuli (e.g., also both from the left side) or from different spatial sources (e.g., the first from the left and the second from the right side). According to the typically observed auditory dominance over visual information for temporal judgments, a kappa effect would be expected in the unimodal and bimodal congruent conditions, but not in the bimodal incongruent condition—in which the first and second auditory markers were not spatially separated. In contrast, according to the typical visual dominance over audition in the spatial domain, the visual stimuli from different spatial sources might cause an illusory perceived distance between the (task-relevant) auditory markers. This, in turn, should evoke a kappa effect also in the bimodal incongruent condition.

Method

Participants The data for this experiment were collected by undergraduate students of Psychology at the University of Tübingen, in partial fulfillment of a laboratory course assignment. The students recruited twenty-eight participants (23 female, five male) among their peers. All participants took part voluntarily, were naïve to the purpose of the experiment, and received standardized instructions. The mean age of these participants was 24.5 years ($SD = 6.6$), and they received course credit for their participation.

Stimuli and apparatus Participants were seated in a dimly lit room at a viewing distance of approximately 55 cm from a 36-

---

[2] Please note that here and in the following pages, we refer exclusively to kappa effects based on the spatial distance between the marker stimuli. Empirically more established instances of kappa-like phenomena in the auditory domain are based on perceived distance on frequency or intensity continua—that is, on pitch or loudness distance (Alards-Tomalin, Leboe-McGowan, & Mondor, 2013; MacKenzie, 2007; Shigeno, 1986).
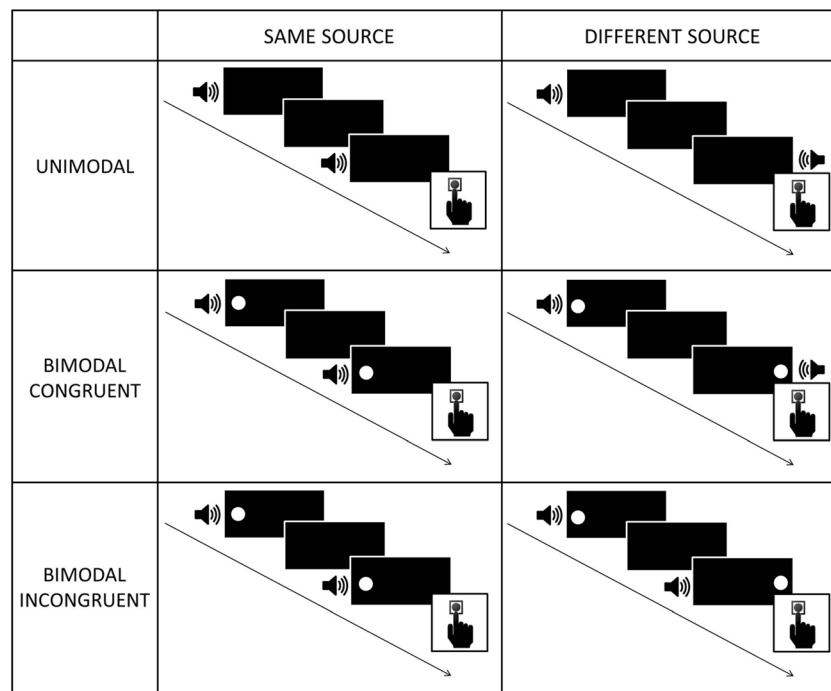
Fig. 1 Schematic illustration of the auditory and visual stimuli in Experiment 1, for the different conditions resulting from the combination of spatial source and modality condition. Depicted are exemplary trials in which the first auditory stimulus is presented on the left side

cm-wide CRT monitor running at 150 Hz. All stimuli were created and presented on a Macintosh computer running MATLAB version 7.14 (The MathWorks) and using Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007). All visual stimuli were presented on a black background (<1 cd/m²). The visual stimuli consisted of a white fixation dot (1-mm diameter), presented at the screen center, and filled white (100 cd/m²), red (35 cd/m²), and green (65 cd/m²) circles (17-mm diameter). These circles could be presented at either the left or the right edge of the monitor for 20 ms. Their midpoint was 17 cm from and horizontally aligned with the fixation point. The German expressions *Jetzt* ("now") and *Rot oder grün?* ("red or green?") served as response prompts and were presented in white at the center of the screen.

The auditory stimuli consisted of 20-ms sine tones at a frequency of 800 Hz, presented at a volume of 78 dB, with onset and offset ramps of 5 ms. Auditory stimuli were presented from one of two speakers situated to the left and right of the monitor, and were separated by 55 cm measured from the center of each speaker, and thus 27.5 cm from the fixation point.

Responses were collected using the "m," "y," and "x" keys on a standard German keyboard.

Procedure The experiment took approximately 1 h. Each trial began with the presentation of the fixation point for 750 ms. The auditory onset marker was then presented from either the

left or right speaker for 20 ms. It was followed by an empty interval with a duration of 500, 900, or 1,600 ms, which was terminated by the presentation of the auditory offset marker for 20 ms and from the left or the right speaker. Then, 100 ms after the end of the interval duration, the response prompt *Jetzt* ("now") appeared, and participants were required to reproduce the interval duration between the auditory markers via two keypresses (indicating the onset and offset of the reproduced interval) of the "m" key with their right index finger. The response prompt disappeared after the first keystroke.

The three different interval durations were presented with equal probabilities in each of the three different modality conditions (unimodal, bimodal congruent, and bimodal incongruent) and for each of the two different spatial source conditions (markers from same vs. different spatial source; also see the supplementary material for all presented stimulus combinations). Specifically, in *unimodal trials* (Fig. 1, upper panel), both the onset and offset markers were equally likely to originate from the left and right speakers, therefore creating trials in which the onset and offset of the interval were marked by signals from the same spatial source (i.e., both from the left [LL] or both from the right [RR] speaker) or from different spatial sources (i.e., left and then right speaker [LR], or vice versa [RL]).

Bimodal congruent trials (Fig. 1, middle panel) were identical to the unimodal trials, except that visual circles appeared simultaneously and were spatially congruent with the auditory

marker stimuli, that is, they originated from the same side of presentation as the auditory markers. Thus, again, the onset and offset markers were equally likely to occur from either the same spatial source (i.e., LL or RR) or from different spatial sources (i.e., LR or RL).

In *bimodal incongruent trials* (Fig. 1, lower panel), the auditory marker stimuli always originated from the same spatial source (i.e., regarding the auditory stimulus origin, there were only LL and RR trials). Again, visual stimuli appeared simultaneously with the auditory markers. The visual stimuli could be spatially congruent with the auditory markers, resulting in same-source trials (i.e., audiovisual LL or RR). Thus, same-source trials in the bimodal incongruent condition were identical to the those in the bimodal congruent condition (and thus only dummy-coded as "incongruent," in order to achieve a fully balanced experimental design). Alternatively, the visual and auditory stimuli could be spatially incongruent (i.e., visual RL accompanying auditory RR, or visual LR accompanying auditory LL). Given the typical visual dominance in spatial perception (Bertelson & Radeau, 1981; Howard & Templeton, 1966; Vroomen et al., 2001), this should result in an illusorily perceived "different" location for the audiovisual marker stimuli. Consequently, these trials were defined as "different source, bimodal incongruent."

Therefore, overall, a 3 (Interval Duration: 500, 900, 1,600 ms) × 3 (Modality Condition: unimodal, bimodal congruent, bimodal incongruent) × 2 (Spatial Source: same vs. different) design resulted, with the side of presentation of the first signal (R vs. L) being fully balanced within all conditions. Each of the resulting 18 trial types was repeated 24 times, thus creating 432 experimental trials. Participants were explicitly instructed to reproduce the duration between the auditory signals and not to rely on the visual signals for the reproduction task. To assure that they would not close their eyes during the block or completely ignore the visual stimulation, catch trials were presented to the participants. These were identical to the bimodal trials described above, with one exception: In these trials, the visual signals within a trial were either green or red (with equal probability). In these trials, participants had to reproduce the duration of the auditorily marked interval as well, but in addition, after the reproduction, they were prompted to indicate the color of the presented circles, by pressing the "y" key for red stimuli with their left middle finger or the "x" key for green stimuli with their left index finger. A total of 48 different catch trials resulted from the combination of the three different interval durations, the two bimodal conditions, the two spatial sources, two sides of presentation, and the two presentation colors.

Participants first performed a practice block of 30 trials, which were picked randomly from all possible trials. Then participants performed the 480 trials (432 experimental + 48 catch trials), which were intermixed, presented in randomized order, and divided into subblocks of 30 trials each. After each subblock, there was a break in which participants were given feedback on their performance within the last block of 30 trials. The feedback consisted of the number of errors in catch trials, as well as the absolute percentage of deviation of the reproduced durations from the presented durations. The latter index was chosen because it does not provide an indication of whether participants have over- or underestimated the interval duration. Participants were instructed to keep this score as low as possible, because lower scores would correspond to more accurate reproductions. The break could be self-terminated by a simple keypress.

Results

On average, participants correctly reported the color of the visual stimuli on 96% [range: 85–100%] of all catch trials. Therefore, it seems safe to conclude that none of the participants ignored the visual stimulation, and consequently, all participants were included in the following analysis. The practice and catch trials were then discarded from further analysis. Outlier reproductions were excluded following a two-step procedure: First, all trials in which reproductions exceeded 5,000 ms were excluded. Then, for each participant and condition, resulting from the combination of the factors of interval duration, modality condition, and spatial source, a three-sigma rule was applied: All reproductions that deviated by more than ±3 standard deviations from the mean of the respective condition for a given participant were excluded from further analysis. This procedure resulted in 0.55% of excluded trials overall.

Reproduced duration A three-way repeated measures analysis of variance (ANOVA) was then performed for reproduced duration, with the factors Interval Duration, Modality Condition, and Spatial Source (see Fig. 2). In the following results, whenever appropriate, Greenhouse–Geisser-corrected p values are reported. As one would expect, there was a main effect of interval duration, $F(2, 54) = 400.32$, $p < .001$, $\eta_p^2 = .94$, with the mean reproduced durations (556, 987, and 1,390 ms) increasing with increasing interval durations (500, 900, and 1,600 ms, respectively). Overall, participants reproduced somewhat longer durations in the unimodal condition (986 ms) than in the bimodal congruent (976 ms) and bimodal incongruent (971 ms) conditions, as indicated by a main effect of modality condition, $F(2, 54) = 5.43$, $p = .02$, $\eta_p^2 = .17$. Most importantly, longer mean reproduced durations were observed for stimuli presented from different spatial sources (1,002 ms) than for stimuli presented from the same spatial source (954 ms), $F(1, 27) = 38.40$, $p < .001$, $\eta_p^2 = .59$. Thus, a typical kappa effect was observed. Interestingly, modality condition modulated the size of this kappa effect, as indicated by an interaction of modality condition and spatial source, $F(2, 54) = 8.20$, $p = .002$, $\eta_p^2 = .23$. Also, there was an interaction of
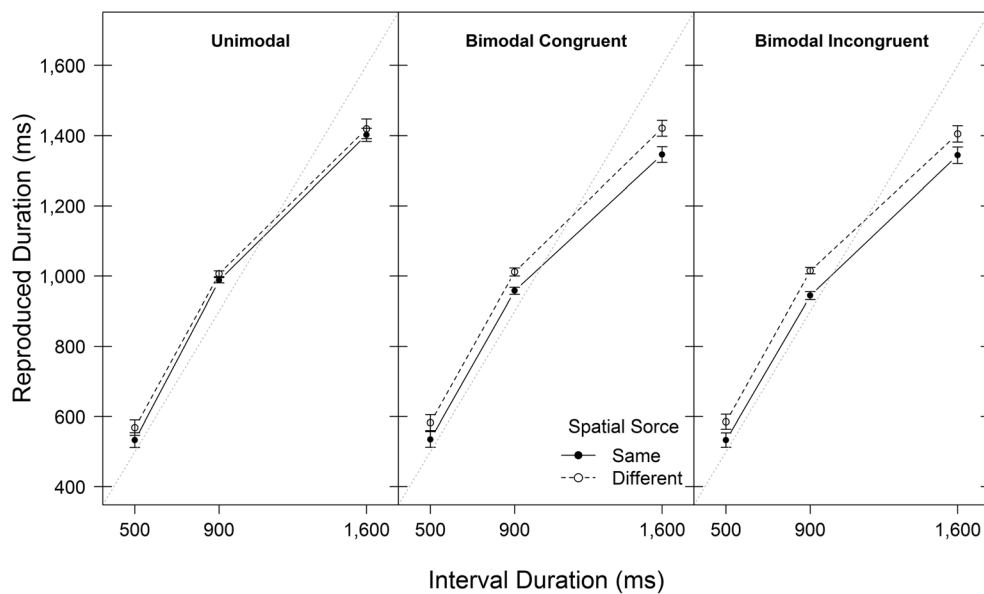
Fig. 2 Mean reproduced durations (±1 *SE*) in milliseconds in Experiment 1, as a function of interval duration and spatial source. The different panels depict the three different modality conditions. Within-subjects *SE*s are computed according to Morey (2008)

modality condition and interval duration, $F(4, 108) = 6.27, p < .001, \eta_p^2 = .19$. Neither the interaction of interval duration and spatial source ($p = .63$) nor the three-way interaction ($p = .19$) was significant.

Follow-up analyses were conducted to decompose the theoretically important interaction of spatial source and modality condition. Since interval duration did not interact with spatial source (i.e., it did not modulate the size of the kappa effect),

we aggregated reproduced duration over all interval durations and then computed the size of the kappa effect as the difference κ = reproduced duration in different-source trials – reproduced duration in same-source trials, for each of the modality conditions. Then a repeated measures ANOVA with the factor Modality Condition was conducted on the magnitudes of the kappa effect (Fig. 3). This analysis revealed a significant effect of modality condition, $F(2, 54) = 8.20, p = .002, \eta_p^2 =$
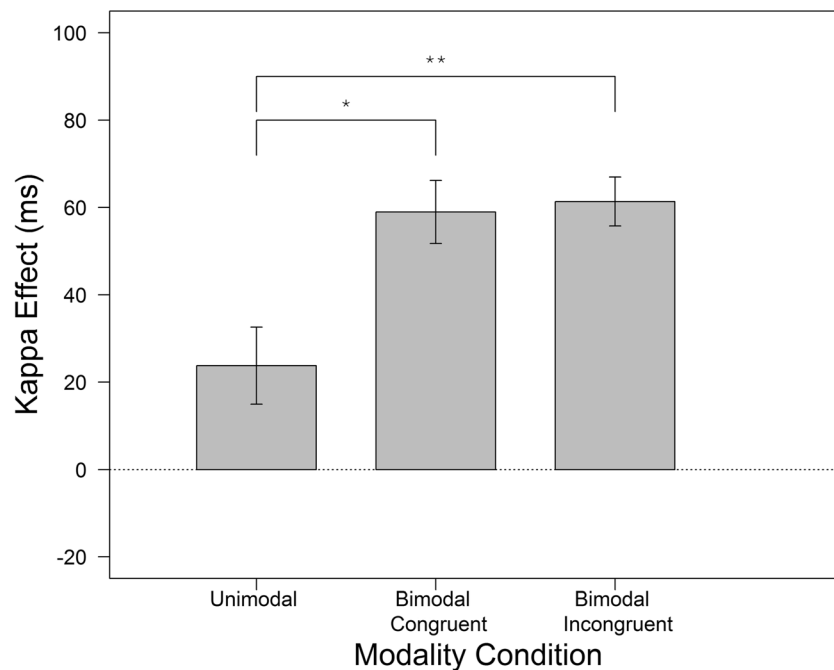

Fig. 3 Mean kappa effects (±1 *SE*) in milliseconds (calculated as the reproduced duration in different-source minus same-source trials) in Experiment 1, as a function of modality condition. Asterisks indicate

significant results from Bonferroni-corrected post hoc comparisons. Within-subjects *SE*s are computed according to Morey (2008)
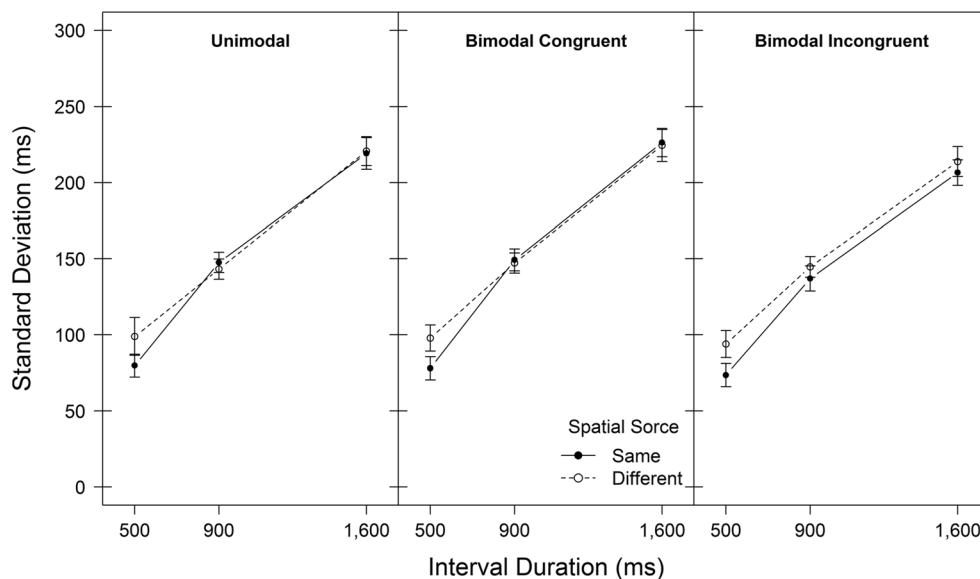
Fig. 4 Mean standard deviations of the reproduced durations (±1 *SE*) in milliseconds in Experiment 1, as a function of interval duration and spatial source. The different panels depict the three different modality conditions. Within-subjects *SE*s are computed according to Morey (2008)

.23, which corresponds to the interaction of modality condition and spatial source from the omnibus ANOVA reported above. Specifically, the kappa effects amounted to 24 ms in the unimodal condition, 59 ms in the bimodal congruent condition, and 61 ms in the bimodal incongruent condition. Pairwise Bonferroni-corrected comparisons revealed differences between the unimodal condition and both the bimodal congruent, $t(27) = 2.84$, $p = .03$, and bimodal incongruent, $t(27) = 3.57$, $p < .01$, conditions, but no difference between the two bimodal conditions, $t(27) = 0.31$, $p = 1.00$. Therefore, even though kappa effects emerged in all three modality conditions, they were more pronounced when visual information was presented in addition to the auditory information. Most interestingly, these bimodal kappa effects occurred irrespective of whether or not the visual information was congruent with the auditory information.

Variability of reproductions A repeated measures ANOVA with the factors Interval Duration, Modality Condition, and Spatial Source was carried out for the standard deviations of reproduced durations (*SD*; see Fig. 4 for the corresponding *M*s and *SE*s), which reflects the variability of duration reproductions. This analysis revealed that, unsurprisingly, the variability of the reproductions increased with interval duration, $F(2, 54) = 73.65$, $p < .001$, $\eta_p^2 = .73$. Specifically, the mean *SD*s were 87, 145, and 219 ms for the 500-, 900-, and 1,600-ms interval durations, respectively. Also, reproductions were slightly more variable in the unimodal (152 ms) and the bimodal congruent (154 ms) conditions than in the bimodal incongruent condition (145 ms), $F(2, 54) = 3.89$, $p = .03$, $\eta_p^2 = .13$. Marginally more variable reproductions were observed for intervals marked by different (154 ms) rather than same (146 ms) spatial sources, $F(1, 27) = 3.90$, $p = .06$, $\eta_p^2 = .13$. This marginal effect was presumably due to an interaction of interval duration and spatial source, $F(2, 54) = 3.87$, $p = .03$, $\eta_p^2 = .13$, such that reproductions were more variable for different than for the same spatial source at the 500-ms interval (77 vs. 97 ms), but not at the 900-ms (145 vs. 145 ms) or the 1,600-ms (217 vs. 220 ms) interval. None of the other interactions approached significance (all *F*s < 1).

Discussion

In the *unimodal* auditory condition, participants reproduced interval durations as being longer when the respective intervals were marked by stimuli presented from different rather than from the same spatial source. This kappa effect was present across a relatively wide range of interval durations (500–1,600 ms), thus extending previous evidence of a kappa effect for auditorily marked intervals of 150 and 225 ms (Grondin & Plourde, 2007).

Interestingly, the magnitude of the kappa effect more than doubled when visual signals were presented along with the auditory signals in the *bimodal congruent* condition. This seems somewhat counterintuitive, considering the spatial properties of our experimental layout: Because the loudspeakers were placed to the left and right of the monitor employed for visual stimulation, the visual stimuli in the different-source condition were actually physically less separated (34 cm) than the auditory stimuli (55 cm, measured from the centers of the speakers). Therefore, on the basis of physical distance, one might expect the kappa effect to decrease when visuospatial information is taken into account. However, due to the superior spatial resolution of the visual system, the

perceived distance between the spatial sources of the marker stimuli might be much more pronounced when visual stimuli are present rather than just unimodal auditory ones. Actually, by employing the Müller-Lyer illusion, the kappa effect has been shown to depend on the perceived illusory rather than physical distance (Lebensfeld & Wapner, 1968). Therefore, we interpret the present findings as a sign of a multisensory integration process, by which the auditory and their accompanying visual markers were integrated into a combined multisensory percept, also incorporating the task-irrelevant spatial information, which is more readily and concisely available from the visual modality. The spatial properties of this combined audiovisual percept then affected perceived duration. Thereby, visual information contributed strongly to the temporal percept. This finding contrasts with a variety of previous results demonstrating that auditory information typically strongly dominates visual information regarding perceived duration (Bausenhart et al., 2014; De la Rosa & Bausenhart, 2013; Klink et al., 2011; Romei et al., 2011; Walker & Scott, 1981).

Crucially, evidence for such a visual influence on auditory information in the temporal domain was present even in the *bimodal incongruent* condition, in which only the visual, but not the auditory, information was presented from different spatial sources. Specifically, on the basis of the typical dominance of auditory information for duration processing, a kappa effect would not be expected in the bimodal incongruent condition. In contrast to this hypothesis, however, a kappa effect was observed. Moreover, this effect was just as strong as the one observed in the bimodal congruent condition. Thus, this finding provides even stronger evidence for the notion that visuospatial information influences the perception of audiotemporal information.

Even though this has not been directly assessed in the present study, a plausible explanation for this finding is that the *physical* distance between the visual stimuli might have led to an *illusory* distance between the same-source auditory markers, which in turn evoked a kappa effect. On the basis of the similar magnitudes of the kappa effect, one might infer that this illusory perceived distance was comparable to the perceived distance in the bimodal congruent condition, in which both auditory and visual stimuli were presented from physically different locations. Hence, the conflict between spatial information from the auditory and visual modalities presumably has been resolved in favor of the visual information, just as one might expect from studies on the spatial ventriloquist effect, which have demonstrated visual dominance when clearly perceivable visual signals are employed (Alais & Burr, 2004; Bertelson & Radeau, 1981; Slutsky & Recanzone, 2001; Vroomen et al., 2001). However, while these studies typically require a response in the spatial domain (e.g., pointing to the spatial source of stimulation), the present results indicate that this phenomenon might also affect perception in the temporal domain through a process of multisensory integration.

However, there is a caveat to this explanation of our results in terms of multisensory integration processes. Specifically, one might argue that participants did not obey the instruction to reproduce the intervals marked by the auditory stimuli, but—in bimodal trials—rather reproduced the intervals on the basis of the visual stimulation. According to this strategy, kappa effects of comparable magnitude should have emerged in the bimodal congruent and in the bimodal incongruent condition. Specifically, in same-source trials, visual stimuli emerged from the same spatial location, whereas in different-source trials, visual stimuli emerged from spatially distinct locations, and this was true for both the congruent and the incongruent modality condition. Thus, if participants had ignored the auditory stimulation and responded to the visual stimulation only, this could have resulted in a result pattern similar to the one observed.

On the one hand, our data provide some evidence consistent with this notion, since overall, unimodal auditory intervals were reproduced as being slightly longer than bimodal ones (around 15 ms at the 900-ms interval). Such a difference would be expected if participants had judged duration in bimodal trials on basis of the visual information alone (Wearden et al., 1998). However, the observed difference was much smaller than what would be expected on the basis of comparable studies (De la Rosa & Bausenhart, 2013; Walker & Scott, 1981), and thus, this factor presumably plays only a minor role in explaining the present results.

On the other hand, several points clearly argue against this alternative explanation: First, sensitivity for temporal judgments is typically much higher in the auditory than in the visual modality (e.g., Gamache & Grondin, 2010; Ulrich et al., 2006). It seems very unlikely that participants would deliberately rely on the modality with inferior temporal resolution and thus increased task difficulty. However, if participants had nonetheless followed such a strategy and responded on the basis of the visual information while ignoring the auditory information, one would expect lower sensitivity, and consequently a higher variability of reproductions in bimodal than in unimodal auditory trials (Bausenhart et al., 2014; Shi et al., 2010). This was clearly not the case in the present data; if anything, the *SD*s of reproductions were even slightly lower in the bimodal incongruent than in the unimodal condition. Actually, this result rather strengthens the notion that responses were based on a combined multisensory percept, since multisensory integration is typically assumed to reduce perceptual noise, in order to achieve an optimally stable percept (Battaglia et al., 2003; Ernst & Bülthoff, 2004). Second, the relatively high proportion of unimodal auditory trials (around one third of all trials) would have made a strategy of

ignoring the auditory stimulation even more counterproductive, since a presumably effortful shift of attention between modalities would be required on a trial-by-trial basis (Grondin, Ivry, Franz, Perreault, & Metthé, 1996; Rousseau, Poirier, & Lemyre, 1983; Ulrich et al., 2006).

Nonetheless, the validity of the present conclusions depends critically on the assumption that participants obeyed the instruction to reproduce the interval marked by the auditory stimuli rather than relying on the visual information alone. Only if this assumption is met can one safely conclude that the bimodal kappa effects found here indicate the results of a multisensory integration process, which combines spatial and temporal information in order to produce a coherent temporal percept, rather than a simple response strategy or crossmodal bias. Therefore, we designed Experiment 2 to specifically rule out this alternative explanation of attending and responding to the visual stimulation only.

Experiment 2

The basic results obtained in Experiment 1 point to multisensory integration of temporal and spatial information from the visual and auditory modalities, by which visuospatial information affects the perception of the temporal information conveyed by the auditory modality. Alternatively, however, one might attribute our results to a response strategy, according to which participants ignored the auditory stimulation and responded on the basis of the visual information. To rule out this alternative, we created a situation in which the task-relevant temporal information was only provided in the auditory modality, but not in the visual modality (Bausenhart et al., 2014; Morein-Zamir et al., 2003).

To this end, we employed the method of constant stimuli in order to assess perceived duration. Specifically, rather than reproducing a single empty temporal interval marked by two brief auditory stimuli, participants had to compare the duration of two successive empty intervals marked by three brief auditory stimuli. This three-stimulus setup is sometimes even described as the standard paradigm to elicit the kappa phenomenon, and leads to more consistent evidence in favor of the kappa effect than do single-interval presentations such as those employed in Experiment 1 (Kuroda, Grondin, & Tobimatsu, 2014; Roussel, Grondin, & Killeen, 2009; Roy, Kuroda, & Grondin, 2011). In bimodal trials, visual stimulation was included that could alter the perceived spatial properties of the combined multimodal percept of the marker stimuli, yet was not sufficient for performing the temporal task. Specifically, visual stimuli only accompanied the first and third auditory markers, but not the second one, which separates the two successive intervals. Accordingly, the visual modality did not contain any information about the relative duration of the two intervals. Thus, in order to compare the

durations of the two intervals, participants had to rely on the auditory information in each trial. In this experiment, again, same-source and different-source trials were created in a unimodal auditory, a bimodal congruent, and a bimodal incongruent condition.

If the results of Experiment 1 were due to noncompliant attention allocation to the visual modality rather than multisensory integration, in Experiment 2 any observed effect of spatial information on temporal perception should be based on the information available from the auditory modality alone. Specifically, then, kappa effects of comparable magnitude should be observed in the unimodal and bimodal congruent conditions. In contrast, the kappa effect should be absent in the bimodal incongruent condition. On the other hand, if the results of Experiment 1 were due to a multisensory integration process, by which the task-irrelevant visual information affected the auditory temporal judgments, a kappa effect would be expected in all three conditions.

Method

Participants A fresh sample of 20 participants was recruited for Experiment 2. Two of the participants from this sample were removed (one only finished one of the two experimental sessions, and one reported adjusting the volume settings during the experiment). The final sample consisted of 14 women and four men with a mean age of 22.72 years ($SD = 4.07$). All participants took part voluntarily, provided written informed consent, were naïve to the purpose of the experiment, and received standardized instructions.

Stimuli and apparatus The stimuli and apparatus were identical to those in in Experiment 1, with the only exception that the "," key was also used for response collection.

Procedure Each participant took part in two sessions completed on different days, each of which lasted approximately 75 min. Participants had to compare two subsequently presented offset-to-onset intervals, indicated by three successive tones. Specifically, the first interval duration was marked by the first and the second tone, and the second interval was marked by the second and the third tone. As in Experiment 1, trials could be either unimodal, bimodal congruent, or bimodal incongruent (also see the supplementary material for all potential stimulus combinations).

Unimodal trials began with the presentation of the fixation point, which remained on the screen until the response prompt appeared. After 750 ms following the onset of the fixation point, three auditory stimuli were presented for 20 ms each, separated by the first and second empty intervals. There were two types of intervals: standard intervals (500 ms) and comparison intervals, which were equiprobable and symmetrically distributed around the standard duration (367, 400, 433, 467,

500, 533, 567, 600, or 633 ms). The comparison duration and presentation order of the standard and comparison were randomized across trials, and both presentation orders were equiprobable. The first auditory stimulus was presented with equal probability from either the left or the right speaker. The spatial source of the remaining marker stimuli was determined by the following rules: Tones marking the standard interval were always presented from the same spatial source. Tones marking the comparison interval could be presented either from the same spatial source as the previous two (thus creating same-source trials; e.g., RRR) or from a different spatial source (creating different-source trials; e.g., RRL for the presentation order *standard first*, or LRR for the presentation order *standard second*). Then, 100 ms following the offset of the third auditory marker, the response prompt appeared, and participants indicated with the index ("m" key) and middle finger ("," key) of their right hand whether they had perceived the first or second interval, respectively, as longer. After their response, the fixation point reappeared, and the next trial began 1,000 ms later.

Bimodal congruent trials were identical to unimodal trials, except that the visual stimuli appeared simultaneously and spatially congruent (i.e., either on the left or on the right) with the first and third auditory marker stimuli. The second auditory marker was never accompanied by a visual stimulus (e.g., auditory RRR accompanied by visual R_R, or auditory RRL accompanied by visual R_L). Thus, the visual modality alone did not convey sufficient information to fulfill the interval comparison task.

Finally, in *bimodal incongruent trials*, visual stimuli were also presented synchronously with the first and third auditory marker stimuli. However, they were spatially congruent with the auditory markers only in same-source trials, in which again all stimuli within a trial were presented from the same source (as in Exp. 1, these trials were nonetheless dummy-coded as bimodal incongruent trials in order to obtain a fully balanced experimental design). In different-source trials, the first auditory and visual stimuli were always spatially congruent. The second visual stimulus, however, was always temporally synchronous but spatially incongruent with the third auditory marker. Specifically, when the standard interval was presented first, all three auditory stimuli and the first visual stimulus originated from the same spatial source (e.g., auditory RRR accompanied by visual R_L). Thus, in these trials, the spatially incongruent visual stimulus presented with the third auditory marker should evoke an illusory spatial separation between the markers of the comparison interval. Therefore, the comparison interval (demarcated by the second and third auditory stimuli) might appear prolonged in comparison to the standard interval (demarcated by the first and second auditory stimuli). In trials in which the standard interval was presented second, the three auditory stimuli originated from different spatial sources. Again, the first visual stimulus was spatially congruent with the auditory stimulus, but the second

was not (e.g., auditory RLR accompanied by visual R_L). Here the perceived distance between the standard interval markers (i.e., the second and third auditory marker) should be decreased by the spatially incongruent visual stimulus. Then, again, the comparison interval should appear prolonged in comparison to the standard interval.[3]

Therefore, overall, a 3 (Modality Condition: unimodal, bimodal congruent, bimodal incongruent) × 2 (Spatial Source: same vs. different) × 9 (Comparison Interval Durations: 367, 400, 433, 467, 500, 533, 567, 600, or 633 ms) design resulted, with side of presentation of the first signal (R vs. L) and presentation order of standard and comparison fully balanced within all conditions. Within each of two sessions, each of the resulting 54 conditions was repeated 12 times, resulting in 648 experimental trials. Again, additional catch trials were presented in order to assure that participants would not close their eyes during the block or completely ignore the visual stimulation. As in Experiment 1, these catch trials were similar to the bimodal trials described above, but with the visual signals randomly colored either green or red. After participants had responded to the comparison task, they were prompted to indicate the color of the presented circles by pressing the "y" key for red stimuli and the "x" key for green stimuli with their left middle and index fingers, respectively. Eighty different catch trials resulted from the combination of the two bimodal conditions, two spatial sources, two presentation orders, two sides of first stimulus presentation, and five different comparison durations distributed symmetrically around the standard duration (393, 447, 500, 553, or 607 ms).

Participants thus completed two experimental sessions with 756 trials each, consisting of one block of 28 practice trials (picked at random from all possible trials) and then 728 trials (648 experimental and 80 catch trials) presented in random order. After each 28 trials, participants could take a self-terminated break and were given feedback on their performance within the last subblock. The feedback consisted of the number of errors in the catch-trial task and the percentage of correct responses in the interval discrimination task.

Results

In catch trials, participants achieved an average of 96% correct responses (range = 76%–100%), and thus all participants were included in the following analyses. Subsequently, all practice and catch trials were discarded from the data.

Figure 5 shows the average proportions of "comparison longer than standard" responses for each combination of

---

[3] As an anonymous reviewer pointed out, in bimodal incongruent different-source trials there were more location changes within the auditory modality for the stimulus order "comparison first" (e.g. RLR) than in all other conditions of this experiment. Additional analyses including the factor Stimulus Order, however, demonstrated that the magnitude of the kappa effect was not affected by the number of auditory location changes within a given trial.
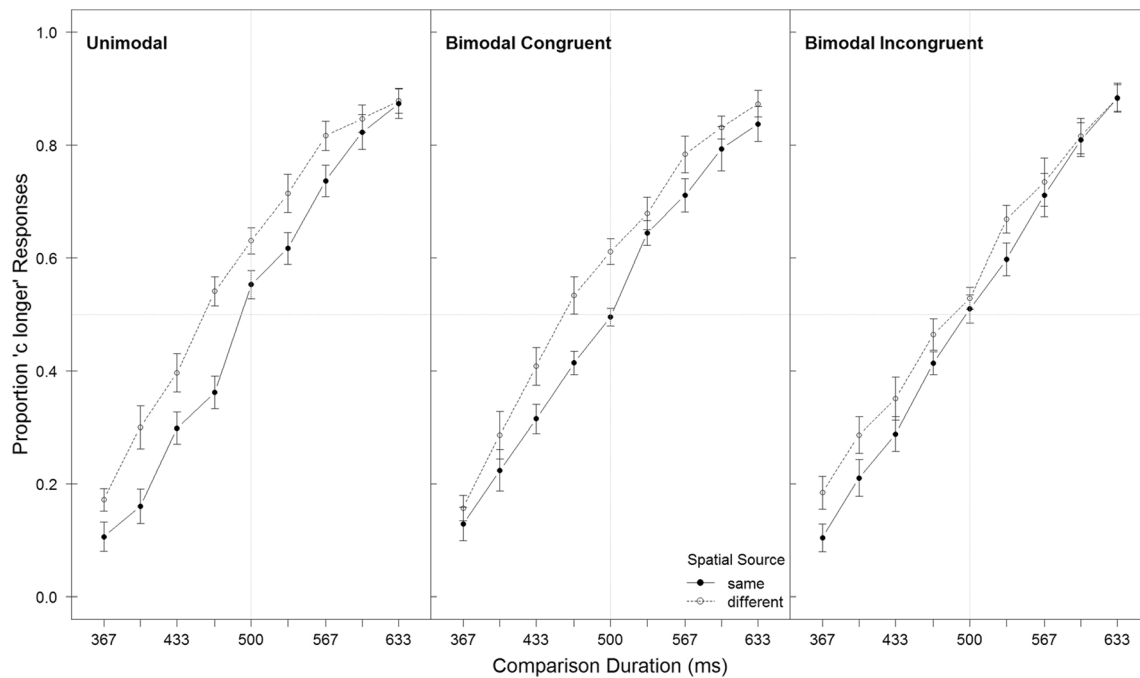
Fig. 5 Mean proportions of "comparison (c) longer" responses (±1 *SE*) in Experiment 2, as a function of comparison duration and spatial source. The different panels depict the three different modality conditions. Within-subjects *SE*s are computed according to Morey (2008)

modality condition, spatial source, and comparison duration.[4] Visual inspection of these psychometric functions suggests several important aspects of the data: First, unsurprisingly, in all conditions the probability of judging the comparison interval as longer than the standard interval increases with comparison duration. Second, when the visual and auditory markers of the standard and comparison intervals stem from the same spatial source, the point of subjective equality (PSE; i.e., the comparison duration corresponding to approximately 50% of "comparison longer" responses) seems to be very close to the duration of the standard interval (500 ms). Thus, perceived duration of the comparison stimulus is quite veridical when the marker stimuli of the standard and comparison interval are presented at the same spatial source, and this is true for all modality conditions. Third, in different-source trials, the psychometric functions seem to be shifted laterally to the left. In other words, the PSE seems smaller than the one observed in same-source trials. This reflects an overestimation of the comparison interval with respect to the 500-ms standard duration, and also with respect to the comparison intervals presented in the same-source condition (i.e., a kappa effect). Again, this is true for all modality conditions, even though the

effect seems to be somewhat less pronounced in the bimodal incongruent condition.

To test these impressions more comprehensively, logistic psychometric functions were fitted to the proportions of "comparison longer" responses (P) at all comparison durations (x_i), separately for each participant, modality condition, and spatial source:

$$P(x_i) = \frac{1}{1 + e^{-(x_i - a)/b}}$$

where a represents the location (i.e., the PSE) and b represents the spread (i.e., the inverse slope) of the psychometric function. The just noticeable difference (JND), corresponding to the interquartile range of the psychometric function, can be computed as JND = $b \times \ln(3)$ (Bush, 1963). PSE and JND indicate the judged duration of the comparison stimulus and discrimination sensitivity, respectively. The functions were fitted through a maximum-likelihood procedure by using the MATLAB function "fminsearch," which is an implementation of the Nelder–Mead algorithm (Nelder & Mead, 1965). Goodness of fit was assessed by computing the deviance D for each of the 108 estimated psychometric functions (3 Modality Conditions × 2 Sources × 18 Participants) separately. On average, D amounted to 7.45 (SD = 4.00, range = [1.12, 21.83]). Chi-square tests conducted on these values with df = 8 revealed only five (4.6%) significant results with $p < .05$. Considering the Type I error rate (α = .05), this indicates appropriate goodness of fit for the estimated psychometric functions.

---

[4] In all following analyses, the data were collapsed over presentation orders of the standard and comparison. It should be noted, however, that an additional analysis including this factor (Bausenhart, Dyjas, Vorberg, & Ulrich, 2012; Ulrich & Vorberg, 2009), as well as an analysis based on the nonparametric Spearman–Kärber method (Ulrich & Miller, 2004), yielded similar results. Most crucially, in all additional analyses, robust kappa effects were observed for all three modality conditions.

Point of subjective equality A repeated measures ANOVA with the factors Modality Condition and Spatial Source was conducted for PSE. Modality condition did not significantly affect the PSE, $F(2, 34) = 2.43$, $p = .10$, $\eta_p^2 = .12$. The PSE was smaller, however, for the different-source condition ($M = 467.4$ ms, $SD = 37.4$) than for the same-source condition ($M = 498.2$ ms, $SD = 14.5$), $F(1, 17) = 12.44$, $p < .01$, $\eta_p^2 = .42$. This corresponds to an overestimation of the comparison duration in different-source trials as compared to same-source trials, and thus, to a typical kappa effect. Also, there was an interaction of modality condition and spatial source, $F(2, 34) = 5.11$, $p = .01$, $\eta_p^2 = .23$. To investigate this interaction in more detail, the kappa effect was computed for each participant and modality condition as the difference $\kappa = $ PSE(same-source) – PSE(different-source). It should be noted that since smaller values of the PSE reflect prolonged perceived duration, here again positive values indicate the presence of a typical kappa effect. Across participants, the kappa effects amounted to 38 ms in the unimodal condition, 30 ms in the bimodal congruent, and 24 ms in the bimodal incongruent condition. Pairwise Bonferroni-corrected t tests were conducted to compare the magnitude of these effects across the different modality conditions. These comparisons showed that the kappa effect was smaller in the bimodal incongruent condition than in the unimodal condition, $t(17) = 2.94$, $p = .03$ (cf. Fig. 6). None of the other comparisons reached significance, all ps > .19. Nonetheless, even in the bimodal incongruent condition, the observed kappa effect was significantly larger than zero, $t(17) = 2.53$, $p = .02$.

Just noticeable difference A repeated measures ANOVA with the factors Modality Condition and Spatial Source revealed that modality condition alone did not affect the JND, $F(2, 34) = 2.82$, $p = .10$, $\eta_p^2 = .14$. Overall, the JND for same-source trials ($M = 88$ ms) was slightly smaller than that for different-source trials ($M = 96$ ms), $F(1, 17) = 4.68$, $p < .05$, $\eta_p^2 = .22$. A marginal interaction between the two factors reflects that this effect was mostly due to the unimodal and bimodal incongruent conditions, whereas in the bimodal congruent condition, the JND was numerically even slightly larger on same-source than on different-source trials, $F(2, 34) = 3.48$, $p = .07$, $\eta_p^2 = .17$ (cf. Fig. 7).

Discussion

Experiment 2 replicated the basic finding of an auditory kappa effect in the unimodal condition, that is, longer perceived duration in different-source trials than in same-source trials. The magnitude of this effect was numerically comparable (38.3 ms) to the one observed in Experiment 1 (35.4 ms for the unimodal 500-ms interval). This seems quite remarkable, given several differences between the two experiments. First, the tasks employed to measure perceived duration were duration reproduction in Experiment 1 and two-alternative forced choice interval comparison in Experiment 2. Therefore, the response mode (production of a motor response vs. perceptual judgment) differed between these tasks. Second, only in the latter task was a constant standard interval presented in each trial, against which the magnitude of the relevant comparison
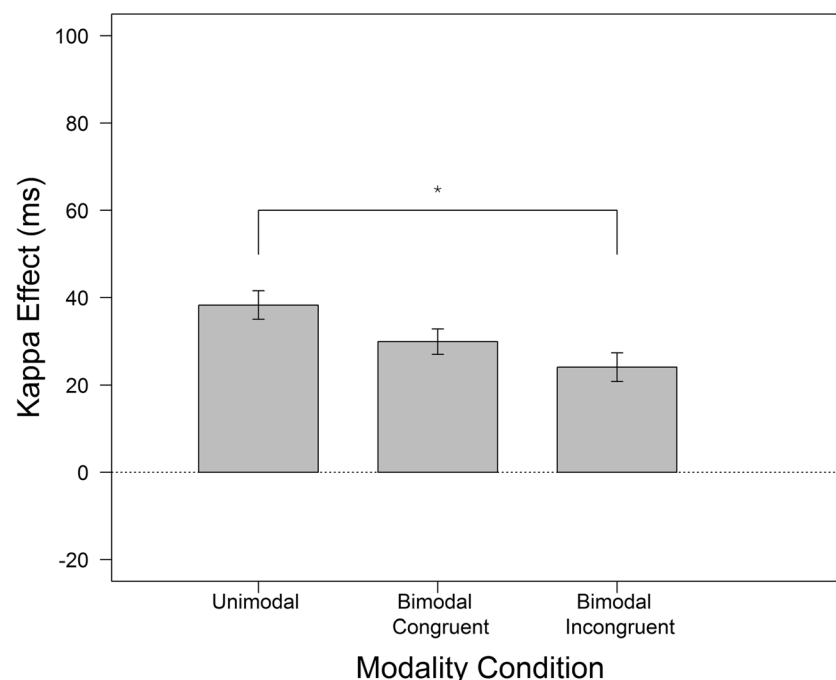

Fig. 6 Mean kappa effects (±1 SE) in milliseconds (calculated as the PSE in same-source trials minus the PSE in different-source trials) in Experiment 2 as a function of modality condition. Asterisks indicate significant results from Bonferroni-corrected post hoc comparisons. Within-subjects SEs are computed according to Morey (2008)
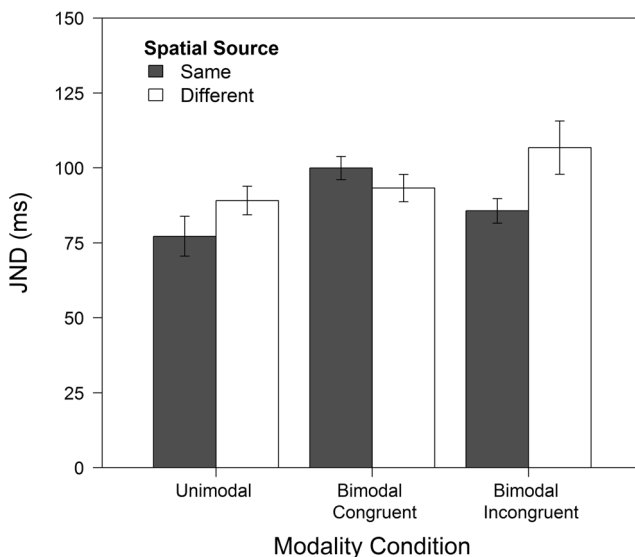
Fig. 7 Mean just noticeable differences (JNDs, ±1 *SE*) in milliseconds in Experiment 2, as a function of modality condition and spatial source. Within-subjects *SE*s are computed according to Morey (2008)

interval had to be compared directly. Third, in the reproduction task the interval durations extended across a rather large range (500–1,600 ms) and were randomly intermixed, whereas the presented intervals in Experiment 2 varied across a much smaller range (367–633 ms, ranging around a 500-ms standard).

Given the lack of a standard interval and the greater variability of the presented intervals in Experiment 1, it might have been conceivable that the responses were more prone to biases in perceived duration than were the responses in Experiment 2. Yet this was apparently not the case: Kappa effects of comparable magnitude were observed for the 500-ms intervals in the unimodal conditions. Therefore, the present kappa effect not only occurs across a range of durations (Exp. 1), but also seems quite robust against variations of the stimulation conditions und judgment mode (Roy et al., 2011).

Despite the similarity of the kappa effect's magnitude in unimodal conditions, remarkably, the observed kappa effects in the bimodal conditions of the two experiments differed quite strongly in magnitude. In Experiment 1, the kappa effects observed in bimodal congruent and incongruent trials were larger than the ones observed in unimodal auditory trials. In Experiment 2, the kappa effect observed in bimodal incongruent trials was significantly smaller than the one observed in unimodal trials. Numerically, the kappa effect in bimodal congruent trials was also smaller than the unimodal kappa effect, and larger than the bimodal incongruent kappa effect, but it did not differ significantly from either condition. There are two potential explanations for this finding of reduced bimodal kappa effects in Experiment 2.

First, this finding might indicate that the proposed alternative explanation of Experiment 1—that is, participants responding on the basis of the visual information alone rather

than of the auditory or the combined multisensory information—was indeed valid to some extent. Since the visual modality is much more sensitive for spatial information than is the auditory modality, this strategy would presumably lead to a more pronounced kappa effect in bimodal than in unimodal auditory trials. Precluding this potentially harmful decision strategy in Experiment 2, however, forced participants to base their decisions on the auditory interval markers. This, in turn, would have led to a more noisy representation of the spatial properties of stimulation, and therefore to reduced kappa effects in bimodal trials as compared to Experiment 1. Actually, this explanation might be tested more directly by assessing kappa effects also in a unimodal visual condition. Here, the kappa effect should be just as large as, or even larger than, the one in bimodal trials. However, our experiments purposely did not include such a condition, for two reasons: First, the participants were actually instructed to base their temporal reproductions and decisions on the auditory information only, and we did not want to compromise this instruction by including a visual unimodal condition. Second, this information would not have been informative with regard to our main experimental question.

A second possible explanation of this pattern of results is based on the fact that in Experiment 1, all visual marker signals in bimodal trials were accompanied by auditory stimuli, but in Experiment 2, only two of the three visual markers were accompanied by auditory stimuli. Indeed, it has been shown that the contingency of the inputs from the different modalities affects the extent to which multisensory integration takes place (Bausenhart et al., 2014; Klink et al., 2011; Sarmiento et al., 2012). For example, in Klink et al. (2011), participants judged the duration of visually presented filled intervals, which were accompanied by auditory filled intervals of incongruent duration. In several experiments, auditory duration strongly affected perception of the visual intervals. When the auditory intervals were embedded in a regular stream of auditory unimodal stimulation, however, their biasing effect on visual interval perception diminished. Presumably, these conditions violated the "unity assumption" (Welch & Warren, 1980), which has often been described as a prerequisite for multisensory integration processes to occur. Similarly, multisensory integration effects on perceived duration of empty visual intervals were much smaller when the visual interval markers were accompanied by only one rather than by two auditory pulses (Bausenhart et al., 2014). Likewise, in the present Experiment 2, which used a similar setup, multisensory biasing effects of the irrelevant auditory information on the relevant visual information were observed, even though to a reduced extent than in Experiment 1. Thus, the reduced magnitude of the bimodal kappa effects in Experiment 2 might have been caused by the reduced contingency between the auditory and visual information, and thus a reduced strength of multisensory integration. It should be noted that these two

explanations (precluding decisions based on visual information and hampered multisensory integration due to reduced contingency between the modalities) are not mutually exclusive, and could have both affected the outcome of Experiment 2 to some extent.

As we noted above, Experiment 2 ruled out the possibility that participants based their decisions on information from the irrelevant visual modality alone. However, participants might have adopted the alternative strategy to deliberately switch between the signal modalities within each trial, such that the temporal information of the first and third markers was obtained from the visual modality, and the temporal information of the second marker was obtained from the auditory modality.[5] However, since all trials in the present experiment were randomly intermixed, participants could not predict in advance whether a unimodal or a bimodal trial would be presented, making such a strategy rather unlikely. Moreover, a large body of evidence has shown that intermodal shifts in duration discrimination strongly impair discrimination performance (Grondin et al., 1996; Rousseau et al., 1983; Ulrich et al., 2006). Thus, a "switching" strategy would entail higher JNDs in bimodal than in unimodal auditory trials. In contrast to this prediction, however, JNDs did not significantly differ between the modality conditions. Therefore, the "switching" account is not a viable explanation for the results of Experiment 2.

Consequently, the effects of the irrelevant stimulation on the relevant modality can be attributed to multisensory integration processes. Even though the reduced contingency between modalities presumably goes along with a reduction of the amount of multisensory integration effects, a reliable kappa effect was still observed in the bimodal incongruent condition. Thus, the most important finding of Experiment 1 was replicated, while any effects of noncompliant attention allocation to the visual modality were precluded. Therefore, this effect seems to be the result of integration of the visuospatial properties of stimulation into the audiotemporal percept. Consequently, participants based their temporal decisions on the integrated multisensory percept, which combined spatial and temporal aspects of stimulation.

General discussion

Taken together, the present results support previous findings of a unimodal auditory kappa effect based on spatial distance (Grondin & Plourde, 2007). In addition, and more importantly for the present purpose, bimodal kappa effects have been investigated for the first time. In two experiments, we observed clear kappa effects for the presentation of congruent audiovisual interval markers, that is, longer perceived duration when

the markers are presented from distant rather than from the same spatial source. Most importantly, such kappa effects were even observed in a bimodal incongruent condition, in which the spatial distance of the marker stimuli was physically only present in the task-irrelevant visual modality. This finding demonstrates that visuospatial information can affect the processing of audiotemporal information in temporal reproduction (Exp. 1) and discrimination (Exp. 2) tasks. Moreover, Experiment 2 confirmed that this effect was not just due to strategic responding on basis of the irrelevant visual information. Thus, the observed dominance of visuospatial over audiotemporal information seems to rely on a rather automatic process of multimodal integration of the information from different senses.

Crucially, this process combines not only aspects of the stimulation that are directly relevant to the task at hand (i.e., temporal information from the auditory modality), but also other aspects that are, per instruction, task-irrelevant (i.e., spatial information from the visual modality). Such a nonselective integration of the basic features of stimulation may be taken as evidence for a rather early locus of multisensory integration mechanisms, acting before attentional selection can filter out task-irrelevant aspects of stimulation. This is in line with studies demonstrating that multisensory integration, as for example in the spatial ventriloquism effect, proceeds independently of deliberate or automatic attention allocation to the multisensory stimulation (Bertelson, Vroomen, Gelder, & Driver, 2000; Vroomen et al., 2001). Moreover, this view would be consistent with the results of various neurophysiological studies showing that multisensory integration acts relatively early in perceptual processing (e.g., Noesselt et al., 2010; Stekelenburg & Vroomen, 2005; for overviews, see Alais, Newell, & Mamassian, 2010; Driver & Noesselt, 2008), as well as behavioral studies indicating perceptual benefits caused by concurrent uninformative stimulation from other modalities (Chen & Spence, 2011; Huang, Yeh, Chen, & Spence, 2011; Morein-Zamir et al., 2003; Vroomen & Keetels, 2009).

How could such an early integration effect be brought in line with theoretical explanations of temporal perception in general and of the kappa effect in particular? For example, as outlined above, a spatial ventriloquist effect might act on the perceived location of the auditory marker stimuli, such that perceived distance between those markers is increased in the bimodal conditions in which visual markers stem from different spatial sources. These integrated multimodal marker percepts then might trigger the switch component of a pacemaker–accumulator mechanism (Gibbon, 1977; Treisman, 1963), which starts and ends the accumulation of temporal pulses. Actually, there is evidence for the notion that multisensory integration of temporally incongruent audiovisual markers modulates the perceived interval duration via this switch mechanism (De la Rosa & Bausenhart, 2013).

---

[5] We thank an anonymous reviewer for pointing this out.

Relating this mechanism to the integration of *spatially* incongruent audiovisual signals, a suggestion by Aschersleben and Müsseler (2008) becomes relevant, which attributes the kappa effect to spatial priming. Accordingly, when two stimuli are presented successively, a gradual activation pattern builds up around the spatial position of the first stimulus. Perception of the second stimulus is consequently sped up by this preexisting activation when it is presented at or near the location of the first stimulus, as compared to stimulation at more distant locations. In terms of a pacemaker–accumulator account of temporal perception, the switch mechanism, which regulates the transmission of pulses to the temporal accumulator, would open earlier after the offset of an interval marked by stimuli presented at the same location, and thus shorten the perceived duration of the demarcated interval relative to a situation in which the marker stimuli are presented from more distant spatial locations (see also Grondin, 1998, for a related suggestion based on the need to reallocate spatial attention, and consequently longer perceptual latency of the second marker, when presented at a distance from the first marker). Interestingly, smaller *SD* of reproductions (Exp. 1) and smaller JND for interval comparison (Exp. 2), especially at shorter interval durations, were observed for same-source trials than for different-source trials. Thus, higher sensitivity was observed for intervals marked by stimuli presented at the same rather than different spatial locations. Even though this interpretation must remain speculative, this result fits well with the idea that spatial priming or attentional availability might facilitate processing of the second interval marker, when presented at close vicinity of the first marker stimulus.

Other accounts of the kappa effect proceed from the notion that the spatial layout of the marker signals results in the assumption of imputed motion of a single object moving through space over time. Importantly, participants tend to attribute a constant velocity to this motion. Consequently, the temporal perception of objects "traveling" larger distances is readjusted to span longer durations than the perception of objects "traveling" smaller distances (Collyer, 1977; Jones & Huang, 1982). Although this explanation fits well with the results of Experiment 2, it seems somewhat less plausible that motion was imputed in Experiment 1, in which only two marker stimuli were presented per trial. But even though the present experiments were not designed to distinguish between the spatial priming or the imputed motion account, both explanations would be consistent with the general assumption outlined above: that the bimodal marked stimuli are combined into a single representation with a common spatial layout already at an early stage of processing. A promising approach to formalize such a mechanism might alternatively also be based on the work of Goldreich (2007), who suggested a Bayesian framework that accounts for spatiotemporal illusions in the tactile modality on basis of prior speed expectations (for applications to the visual kappa effect, see also Chen,

Zhang, & Kording, 2016; Ogata, Kuroda, Miyazaki, Grondin, & Tobimatsu, 2016).

On the other hand, it should be noted that some researchers have suggested that the interdependence of time and space might arise at a relatively late level of processing—that is, at a representational stage (Cai & Connell, 2015). Specifically, it was suggested that temporal and spatial information share a common encoding format. Therefore, initially an unbiased representation of temporal information is built and then stored for later retrieval in a memory stage. In this stage, however, the temporal representation is subject to interference—for example, from spatial information, which is supposed to share the same (presumably magnitude-based) representational format (Fabbri, Cancellieri, & Natale, 2012; Walsh, 2003). As a prerequisite for such an interference to occur, however, the spatial information must be encoded with a relatively high acuity. Specifically, the participants in Cai and Connell's study had to reproduce—varying on a trial-by-trial basis—either the duration of a sound or the length of a stick, which were presented simultaneously. Duration reproductions were not affected by stick length, when the stick could be perceived exclusively via the tactile modality. However, when the stick could be touched and seen at the same time, the—presumably more reliable—visuotactile length percept affected reproduced duration of the unrelated sound. In this sense, the observed effect resembles the effects of multisensory integration, in which the reliability of the perceptual input from different modalities determines the strength of their relative contributions to the integrated multisensory percept (Alais & Burr, 2004; Ernst & Banks, 2002; Hartcher-O'Brien et al., 2014). Nonetheless, typical prerequisites for multisensory integration to occur, such as the unity assumption (Welch & Warren, 1980), were most likely violated under these experimental conditions. For example, the stick was placed on the table by the experimenter at the beginning of each trial and thus its occurrence was not contingent on the temporal occurrence of the sound. Accordingly, it is conceivable the mechanisms underlying the effects observed by Cai and Connell differ from the ones underlying typical multisensory integration effects such as spatial or temporal ventriloquism.

Also, unlike in Cai and Connell's (2015) study, in the present experiments participants were never asked to judge or reproduce the stimulation from the irrelevant visual modality. Therefore, it seems rather unlikely that a dedicated representation of the spatial distance between the visual markers would have interfered with the audiotemporal percept only at a late memory stage, in which the separate spatial and temporal representations are stored for later retrieval. Consequently, and because of the variety of studies attributing multisensory integration effects to a rather early, perceptual stage of processing (see above), we currently favor an interpretation of the present findings as due to automatic, perceptual integration of the marker stimuli. Yet, the present experiments were not

designed to distinguish between possible loci of the effects of bimodal stimulation on perceived interval duration, and thus, this interpretation remains somewhat speculative.

Despite this open issue, our primary research question can be answered quite clearly: reliable kappa effects on perceived auditory duration were observed under conditions of audiovisual stimulation and most importantly, when the spatial separation of the marker stimuli was only present in the task-irrelevant visual modality. Therefore, even though auditory information typically dominates strongly over visual information when participants perform temporal judgments, visuospatial information can affect auditory temporal processing and thereby cause marked distortions of perceived auditory duration. Consequently, the relative contributions of auditory and visual information to perceived duration seem not only to be determined by the reliability of the unisensory signals in the relevant temporal task domain, but also by their reliability in task-irrelevant domains such as space.

References

Alais, D., & Burr, D. (2004). Ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257–262. https://doi.org/10.1016/S0960-9822(04)00043-0

Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory processing in review: From physiology to behaviour. *Seeing and Perceiving*, *23*, 3–38. https://doi.org/10.1163/187847510X488603

Alards-Tomalin, D., Leboe-McGowan, L. C., & Mondor, T. A. (2013). Examining auditory kappa effects through manipulating intensity differences between sequential tones. *Psychological Research*, *77*, 480–491. https://doi.org/10.1007/s00426-012-0438-8

Aschersleben, G., & Bertelson, P. (2003). Temporal ventriloquism: Crossmodal interaction on the time dimension: 2. Evidence from sensorimotor synchronization. *International Journal of Psychophysiology*, *50*, 157–163. https://doi.org/10.1016/s0167-8760(03)00131-4

Aschersleben, G., & Müsseler, J. (2008). On the perceived interdependence of space and time: Evidence for spatial priming in the temporal kappa effect. In R. Nijhawan & B. Khurana (Eds.), *Space and time in perception and action* (pp. 181–195). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511750540.012

Battaglia, P. W., Jacobs, R. A., & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A*, *20*, 1391–1397. https://doi.org/10.1364/JOSAA.20.001391

Bausenhart, K. M., De la Rosa, M. D., & Ulrich, R. (2014). Multimodal integration of time: Visual and auditory contributions to perceived duration and sensitivity. *Experimental Psychology*, *61*, 310–322. https://doi.org/10.1027/1618-3169/a000249

Bausenhart, K. M., Dyjas, O., Vorberg, D., & Ulrich, R. (2012). Estimating discrimination performance in two-alternative forced choice tasks: Routines for MATLAB and R. *Behavior Research*

Methods, *44*, 1157–1174. https://doi.org/10.3758/s13428-012-0207-z

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory–visual spatial discordance. *Perception & Psychophysics*, *29*, 578–584. https://doi.org/10.3758/BF03207374

Bertelson, P., Vroomen, J., Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*, 321–332. https://doi.org/10.3758/BF03205552

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. https://doi.org/10.1163/156856897X00357

Bush, R. R. (1963). Estimation and evaluation. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 429–469). New York, NY: Wiley.

Cai, Z. G., & Connell, L. (2015). Space–time interdependence: Evidence against asymmetric mapping between time and space. *Cognition*, *136* 268–281. https://doi.org/10.1016/j.cognition.2014.11.039

Chen, Y.-C., & Spence, C. (2011). The crossmodal facilitation of visual object representations by sound: Evidence from the backward masking paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1784–1802. https://doi.org/10.1037/a0025638

Chen, Y., Zhang, B., & Kording, K. P. (2016). Speed constancy or only slowness: What drives the kappa effect. *PLoS ONE*, *11*, e154013:1–14. https://doi.org/10.1371/journal.pone.0154013

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1953). A new phenomenon in time judgment. *Nature*, *172*, 901.

Collyer, C. E. (1977). Discrimination of spatial and temporal intervals defined by three light flashes: Effects of spacing on temporal judgments and of timing on spatial judgments. *Perception & Psychophysics*, *21*, 357–364. https://doi.org/10.3758/BF03199487

de Haas, B., Cecere, R., Cullen, H., Driver, J., & Romei, V. (2013). The duration of a co-occurring sound modulates visual detection performance in humans. *PLoS ONE*, *8*, e54789. https://doi.org/10.1371/journal.pone.0054789

De la Rosa, M. D., & Bausenhart, K. M. (2013). Multimodal integration of interval duration: Temporal ventriloquism or changes in pace-maker rate? *Timing and Time Perception*, *1*, 189–215. https://doi.org/10.1163/22134468-00002015

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron*, *57*, 11–23. https://doi.org/10.1016/j.neuron.2007.12.013

Driver, J., & Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. *Current Biology*. https://doi.org/10.1016/S0960-9822(00)00740-5

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433. https://doi.org/10.1038/415429a

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*, 162–169. https://doi.org/10.1016/j.tics.2004.02.002

Fabbri, M., Cancellieri, J., & Natale, V. (2012). The A Theory Of Magnitude (ATOM) model in temporal perception and reproduction tasks. *Acta Psychologica*, *139*, 111–123. https://doi.org/10.1016/j.actpsy.2011.09.006

Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics*, *63*, 718–725. https://doi.org/10.3758/BF03194432

Gamache, P.-L., & Grondin, S. (2010). Sensory-specific clock components and memory mechanisms: Investigation with parallel timing. *European Journal of Neuroscience*, *31*, 1908–1914. https://doi.org/10.1111/j.1460-9568.2010.07197.x

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*, 279–325. https://doi.org/10.1037/0033-295X.84.3.279

Goldreich, D. (2007). A Bayesian perceptual model replicates the cutaneous rabbit and other tactile spatiotemporal illusions. *PLoS ONE, 2,* e333. https://doi.org/10.1371/journal.pone.0000333

Grondin, S. (1998). Judgments of the duration of visually marked empty time intervals: Linking perceived duration and sensitivity. *Perception & Psychophysics, 60,* 319–330. https://doi.org/10.3758/BF03206040

Grondin, S., Ivry, R. B., Franz, E., Perreault, L., & Metthé, L. (1996). Markers' influence on the duration discrimination of intermodal intervals. *Perception & Psychophysics, 58,* 424–433. https://doi.org/10.3758/BF03206818

Grondin, S., Kuroda, T., & Mitsudo, T. (2011). Spatial effects on tactile duration categorization. *Canadian Journal of Experimental Psychology, 65,* 163–167. https://doi.org/10.1037/a0022785

Grondin, S., & Plourde, M. (2007). Discrimination of time intervals presented in sequences: Spatial effects with multiple auditory sources. *Human Movement Science, 26,* 702–716. https://doi.org/10.1016/j.humov.2007.07.009

Hartcher-O'Brien, J., Di Luca, M., & Ernst, M. O. (2014). The duration of uncertain times: Audiovisual information about intervals is integrated in a statistically optimal fashion. *PLoS ONE, 9,* e89339:3–10. https://doi.org/10.1371/journal.pone.0089339

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation.* Oxford: Wiley.

Huang, P.-C., Yeh, S.-L., Chen, Y.-C., & Spence, C. (2011). Synchronous sounds enhance visual sensitivity without reducing target uncertainty. *Seeing and Perceiving, 24,* 623–638. https://doi.org/10.1163/187847611X603765

Huang, Y., & Jones, B. (1982). On the interdependence of temporal and spatial judgments. *Perception & Psychophysics, 32,* 7–14.

Jones, B., & Huang, Y. L. (1982). Space-time dependencies in psychophysical judgment of extent and duration: Algebraic models of the tau and kappa effects. *Psychological Bulletin, 91,* 128–142. https://doi.org/10.1037/0033-2909.91.1.128

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception, 36*(ECVP Abstract Suppl.), 14.

Klink, P. C., Montijn, J. S., & van Wezel, R. J. A. (2011). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism, and variable internal clock rates. *Attention, Perception, & Psychophysics, 73,* 219–236. https://doi.org/10.3758/s13414-010-0010-9

Kuroda, T., Grondin, S., & Tobimatsu, S. (2014). Does spatiotemporal integration occur with single empty time intervals instead of two neighboring intervals in the visual modality? In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 2519–2524). Austin: Cognitive Science Society.

Lebensfeld, P., & Wapner, S. (1968). Configuration and space-time interdependence. *American Journal of Psychology, 81,* 106–110.

MacKenzie, N. (2007). *The kappa effect in pitch/time context.* Doctoral Dissertation, Columbus: Ohio State University.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research, 17,* 154–163. https://doi.org/10.1016/S0926-6410(03)00089-2

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7,* 308–313. https://doi.org/10.1093/comjnl/7.4.308

Nijhawan, R., & Khurana, B. (2010). *Space and time in perception and action.* Cambridge: Cambridge University Press.

Noesselt, T., Tyll, S., Boehler, C. N., Budinger, E., Heinze, H.-J., & Driver, J. (2010). Sound-induced enhancement of low-intensity vision: multisensory influences on human sensory-specific cortices and thalamic bodies relate to perceptual enhancement of visual detection sensitivity. *Journal of Neuroscience, 30,* 13609–13623. https://doi.org/10.1523/JNEUROSCI.4524-09.2010

Ogata, K., Kuroda, T., Miyazaki, M., Grondin, S., & Tobimatsu, S. (2016). The Kappa effect with only two visual markers. *Multisensory Research, 29,* 703–725. https://doi.org/10.1163/22134808-00002533

Oliveri, M., Koch, G., & Caltagirone, C. (2009). Spatial–temporal interactions in the human brain. *Experimental Brain Research, 195,* 489–497. https://doi.org/10.1007/s00221-009-1834-1

Price-Williams, D. R. (1954). The kappa-effect. *Nature, 173,* 363–364.

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1085–1099. https://doi.org/10.1037/0096-1523.28.5.1085

Romei, V., De Haas, B., Mok, R. M., & Driver, J. (2011). Auditory stimulus timing influences perceived duration of co-occuring visual stimuli. *Frontiers in Psychology, 2,* 215. https://doi.org/10.3389/fpsyg.2011.00215

Rousseau, R., Poirier, J., & Lemyre, L. (1983). Duration discrimination of empty time intervals marked by intermodal pulses. *Perception & Psychophysics, 34,* 541–548. https://doi.org/10.3758/BF03205907

Roussel, M.-È., Grondin, S., & Killeen, P. (2009). Spatial effects on temporal categorisation. *Perception, 38,* 748–762. https://doi.org/10.1068/p6165

Roy, M., Kuroda, T., & Grondin, S. (2011). Effect of space on auditory temporal processing with a single-stimulus method. In P. Strumillo (Ed.), *Advances in sound localization* (pp. 95–104). Rijeka, Croatia: InTech.

Sarmiento, B. R., Shore, D. I., Milliken, B., & Sanabria, D. (2012). Audiovisual interactions depend on context of congruency. *Attention, Perception, & Psychophysics, 74,* 563–74. https://doi.org/10.3758/s13414-011-0249-9

Sarrazin, J.-C., & Giraudo, Marie-Dominique Pittenger, J. B. (2007). Tau and Kappa effects in physical space: The case of audition. *Psychological Research, 71,* 201–218. https://doi.org/10.1007/s00426-005-0019-1

Sarrazin, J.-C., Giraudo, M.-D., Pailhous, J., & Bootsma, R. J. (2004). Dynamics of balancing space and time in memory: tau and kappa effects revisited. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 411–30. https://doi.org/10.1037/0096-1523.30.3.411

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

Shi, Z., Chen, L., & Müller, H. J. (2010). Auditory temporal modulation of the visual Ternus effect: The influence of time interval. *Experimental Brain Research, 203,* 723–735. https://doi.org/10.1007/s00221-010-2286-3

Shigeno, S. (1986). The auditory tau and kappa effects for speech and nonspeech stimuli. *Perception & Psychophysics, 40,* 9–19. https://doi.org/10.3758/BF03207588

Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science, 145,* 1328–1330. https://doi.org/10.1126/science.145.3638.1328

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *NeuroReport, 12,* 7–10. https://doi.org/10.1097/00001756-200101220-00009

Stekelenburg, J. J., & Vroomen, J. (2005). An event-related potential investigation of the time-course of temporal ventriloquism. *NeuroReport, 16,* 641–644. https://doi.org/10.1097/00001756-200504250-00025

Suto, Y. (1952). The effect of space on time estimation (S-effect) in tactual space: I. *Japanese Journal of Psychology, 22,* 45–57.

Treisman, M. M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs: General and Applied, 77,* 1–31. https://doi.org/10.1037/h0093864

Ulrich, R., & Miller, J. (2004). Threshold estimation in two-alternative forced-choice (2AFC) tasks: The Spearman-Kärber method. *Perception & Psychophysics*, *66*, 517–533. https://doi.org/10.3758/BF03194898

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Crossmodal temporal discrimination: Assessing the predictions of a general pacemaker-counter model. *Perception & Psychophysics*, *68*, 1140–1152. https://doi.org/10.3758/BF03193716

Ulrich, R., & Vorberg, D. (2009). Estimating the difference limen in 2AFC tasks: Pitfalls and improved estimators. *Attention, Perception, & Psychophysics*, *71*, 1219–1227. https://doi.org/10.3758/APP.71.6.1219

Vroomen, J., Bertelson, P., & de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics*, *63*, 651–659. https://doi.org/10.3758/BF03194427

Vroomen, J., & de Gelder, B. (2004). Temporal ventriloquism: Sound modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 513–518. https://doi.org/10.1037/0096-1523.30.3.513

Vroomen, J., & Keetels, M. (2009). Sounds change four-dot masking. *Acta Psychologica*, *130*, 58–63. https://doi.org/10.1016/j.actpsy.2008.10.001

Walker, J. T., & Scott, K. J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1327–1339. https://doi.org/10.1037/0096-1523.7.6.1327

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology*, *51*, 97–120. https://doi.org/10.1080/713932672

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*, 638–667. https://doi.org/10.1037/0033-2909.88.3.638