Modality effects in rhythm processing: Auditory encoding of visual rhythms is neither obligatory nor automatic

J. Devin McAuley and Molly J. Henry
Michigan State University, East Lansing, Michigan
and Bowling Green State University, Bowling Green, Ohio

Modality effects in rhythm processing were examined using a tempo judgment paradigm, in which participants made *speeding-up* or *slowing-down* judgments for auditory and visual sequences. A key element of stimulus construction was that the expected pattern of tempo judgments for critical test stimuli depended on a beat-based encoding of the sequence. A model-based measure of degree of beat-based encoding computed from the pattern of tempo judgments revealed greater beat sensitivity for auditory rhythms than for visual rhythms. Visual rhythms with prior auditory exposure were more likely to show a pattern of tempo judgments similar to that for auditory rhythms than were visual rhythms without prior auditory exposure, but only for a beat period of 600 msec. Slowing down the rhythms eliminated the effect of prior auditory exposure on visual rhythm processing. Taken together, the findings in this study support the view that auditory rhythms demonstrate an advantage over visual rhythms in beat-based encoding and that the auditory encoding of visual rhythms can be facilitated with prior auditory exposure, but only within a limited temporal range. The broad conclusion from this research is that "hearing visual rhythms" is neither obligatory nor automatic, as was previously claimed by Guttman, Gilroy, and Blake (2005).

The observation that people's experience of the world depends on information received through multiple senses has led to many questions concerning interactions between visual and auditory information in perception. Much stimulus information is modality specific, in that it is conveyed only through a single sense (e.g., the color of an object). However, other stimulus information (e.g., the location of an object or the duration of an event) is amodal, in that it can be conveyed by more than one sense. One general question in research on this topic is whether the visual and auditory systems demonstrate processing specialization (preferences) for particular types of amodal information. In this respect, when spatial information received by the auditory and visual systems conflicts, there is an abundance of evidence that the visual system "wins out" over the auditory system (Hay, Pick, & Ikeda, 1965; Howard & Templeton, 1966). Conversely, when temporal information received by the auditory and visual systems conflicts, there is increasing evidence that the auditory system "wins out" over the visual system (Rencanzone, 2003; Repp & Penel, 2002).

The most well-known example of visual dominance in spatial processing is the ventriloquism effect (Stratton, 1897; Thomas, 1941). Judgments about the perceived location of a sound source are biased in the direction of the associated object's location in visual space, but the perceived location of an object in visual space is not similarly biased in the direction of the associated object's location in auditory space (Bertelson & Aschersleben, 1998; Pick, Warren, & Hay, 1969). With respect to auditory dominance in temporal processing, studies have shown that there is a tendency to perceive a single flash and single click that are in close temporal proximity as simultaneous, with misperception of the visual stimulus accounting for most of the shift (Fendrich & Corballis, 2001). Similarly, in the auditory-driving effect, a repetitive sound (auditory flutter) presented simultaneously with a repetitive visual stimulus (visual flicker) causes the perceived visual flicker rate to shift toward the auditory flutter rate, even though the auditory flutter and visual flicker rates are easily distinguishable (Rencanzone, 2003).

The focus of the present article is on the nature of auditory temporal-processing dominance in the domain of rhythm perception. Previous studies addressing this issue in the rhythm and tempo perception literature have revealed both general auditory performance advantages (Glenberg & Jona, 1991; Glenberg, Mann, Altman, Forman, & Procise, 1989; Grondin & McAuley, 2009) and asymmetric interference effects (Guttman, Gilroy, & Blake, 2005; Repp & Penel, 2002, 2004). With respect to a general auditory performance advantage, participants who are asked to recall a pattern of short and long durations comprising target auditory and visual rhythms show better recall performance for auditory rhythms than for

J. D. McAuley, dmcauley@msu.edu

visual rhythms, but notably this performance advantage is limited to temporal patterns that involve a simple integer ratio of durations (Glenberg & Jona, 1991; Glenberg et al., 1989). Glenberg and colleagues interpreted this result as support for a chunking model in which the source of auditory advantage in temporal processing can be attributed to a modality difference in the beat-based encoding of the rhythms. Further supporting this view was the finding that rhythms with slower tempi that are predicted to make beat extraction more difficult do not show the same auditory performance advantage (Glenberg & Jona, 1991).

Collier and Logan (2000) reported a similar result, but using a different paradigm in which participants made same–different judgments about pairs of rhythms. Rhythm pairs were either unimodal (auditory–auditory or visual–visual) or cross-modal (auditory–visual or visual–auditory). Consistent with the work of Glenberg and colleagues (Glenberg & Jona, 1991; Glenberg et al., 1989), an auditory performance advantage was found for unimodal auditory–auditory rhythm pairs, which was stronger at fast tempi than at slow tempi. Studies involving tapping paradigms have also reported large modality differences in synchronization performance, where visual rhythms are found to produce much worse synchronization performance than auditory rhythms and do not show the same metrical structure benefit (Patel, Iversen, Chen, & Repp, 2005).

Cross-modal comparisons involving interference paradigms have also revealed a pattern of asymmetric interference supporting auditory dominance in rhythm processing. The common assumption associated with the use of interference paradigms in this context is that the extent of task disruption reflects the extent to which the representation of the to-be-encoded information overlaps with the representation of the irrelevant information. In this respect, individuals monitoring visual rhythms in bimodal (auditory–visual) conditions have been shown to make phase corrections in response to changes in a concurrent unattended auditory rhythm during tapping tasks, but not the reverse (Repp & Penel, 2002). Similarly, when individuals were asked to tap a finger in response to a target auditory rhythm (tone sequence) or target visual rhythm (flashing light) while ignoring a distractor rhythm presented in the other modality, taps were strongly attracted to to-be-ignored auditory distractors, but to-be-ignored visual distractors had little to no effect on taps (Repp & Penel, 2004).

Recently, Guttman et al. (2005) used an interference paradigm to support the claim that when individuals make temporal judgments about rhythms, visual rhythms receive obligatory automatic auditory encoding; that is, individuals "hear" visual rhythms. Participants made same–different judgments about pairs of visual rhythms presented either in isolation or concurrently with auditory rhythms; concurrent auditory rhythms were either congruent or incongruent in relation to the timing of the visual rhythms. Consistent with the auditory encoding of visual rhythms, Guttman and colleagues found that (1) accuracy of same–different judgments for visual rhythm pairs was much worse when they were presented concurrently with incongruent auditory rhythms than with congruent

auditory rhythms, with performance intermediate for visual rhythm pairs presented in isolation (without sound); (2) both irrelevant auditory and visual information reduced accuracy, but the degree of decrement was larger for irrelevant auditory information; and (3) incongruent auditory rhythms affected visual rhythm comparison accuracy more when presented concurrently with the first to-be-encoded rhythm in the pair than when presented with the second rhythm. Moreover, two specific elements of the Guttman et al. study were designed to eliminate the possibility that the interference effects they observed were due to controlled (strategic), rather than automatic, auditory recoding of visual rhythms and to, in general, bias against the use of auditory codes. First, fast sequences were used rather than slow musiclike rhythms so that time would not allow for an effortful recoding. Second, visual rhythms were complex stimuli in which contrast reversals in a series of Gabor patches delineated the temporal structure of the rhythm; complex visual changes should have encouraged processing in a visual mode.

The present study revisits the automatic auditory encoding hypothesis but takes an approach different from that in Guttman et al. (2005) by attempting to increase the likelihood of auditory encoding, rather than biasing against it. To this end, we (1) used slow, musiclike rhythms with predictable (rather than stochastic; cf. Guttman et al., 2005) temporal structure and (2) marked time changes with discrete rhythmic flashes of a black square on a white background, rather than using more complex visual stimuli. The general rationale for this approach was that if we encouraged auditory encoding and did not find it, this would provide strong evidence against the Guttman et al. version of the auditory-encoding hypothesis, which views auditory encoding of visual rhythms as both automatic and obligatory.

Two experiments were conducted that involved making tempo judgments about auditory and visual rhythms. The use of tempo judgments provided us with a valuable window on encoding because the direction of tempo judgments (speeding up vs. slowing down) for the tested rhythms provided us with an index of the extent to which participants were encoding the rhythms they heard or saw by using an implied beat. The approach taken was motivated by recent work by McAuley and colleagues (Grahn & McAuley, 2009; McAuley, Frater, Janke, & Miller, 2006), who used the same paradigm to study individual differences in sensitivity to an implied beat, which we describe next.

Figure 1 shows the two types of rhythms of interest in this work. Four-tone (control) rhythms consist of an initial pair of elements marking out a fixed 600-msec interonset interval (IOI), followed after a 1,200-msec IOI by a pair of tones marking a variable final interval (600 msec $\pm \Delta T$). Five-tone (test) rhythms critically differ from control rhythms by the insertion of an additional element that subdivides the initial 600-msec interval into a pair of 300-msec intervals. It is important to note that, for the test rhythms, there is an implied periodic beat, which provides participants with an implicit 600-msec referent. If participants base their tempo judgments on the implied beat, the pattern of tempo judgments for test and control rhythms

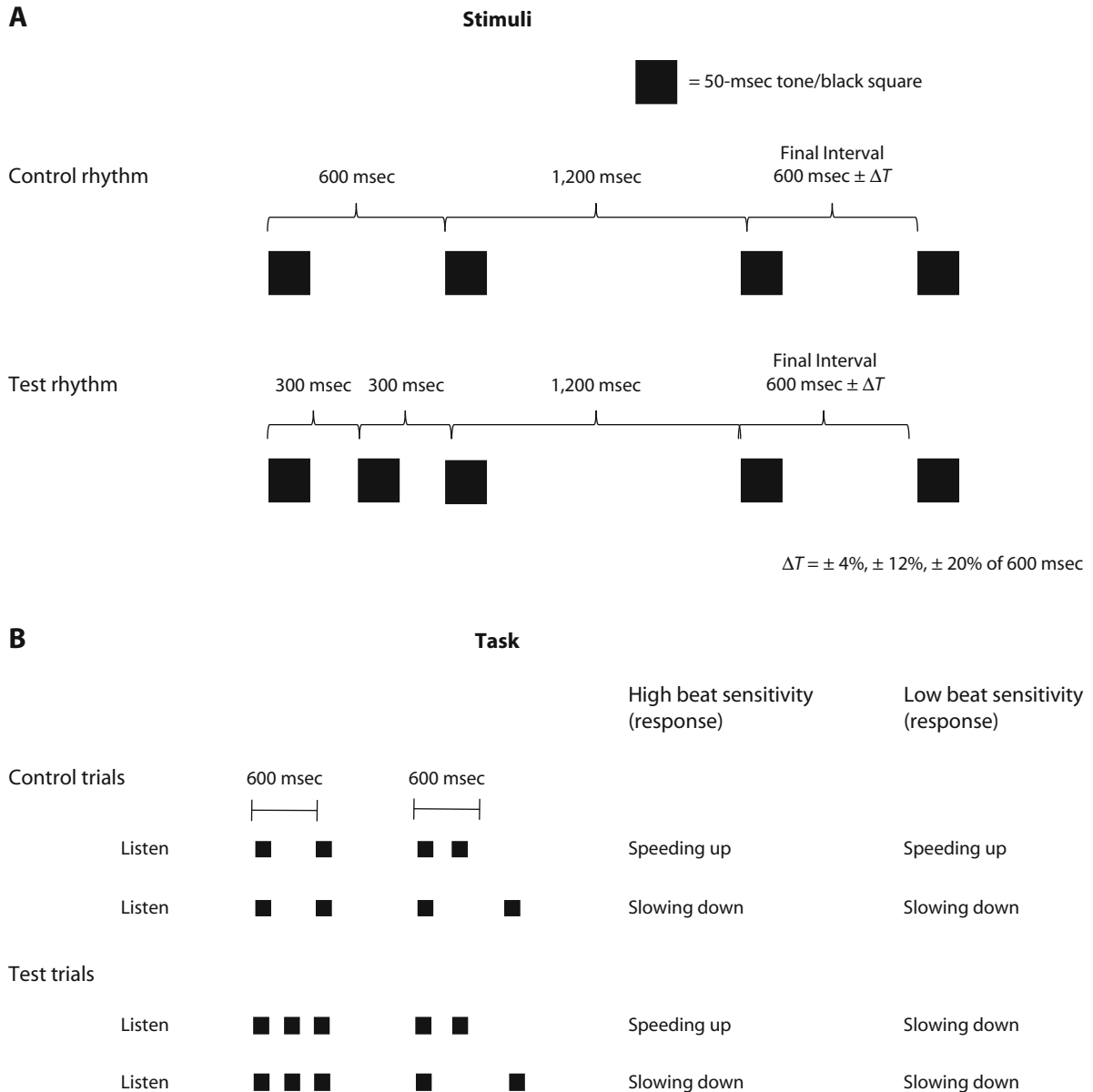**A**   **Stimuli**

Figure 1. Tempo judgment paradigm. (A) Illustration of four-element (control) and five-element (test) rhythms. (B) Task and expected pattern of tempo judgments based on four-element (control) and five-element (test) rhythms.

will be expected to be similar, since both of them involve a 600-msec referent. For control rhythms, the referent is explicitly marked, whereas for test rhythms, the referent must be inferred from the temporal structure of the sequence of elements.

Combining this paradigm with functional magnetic resonance imaging (fMRI), Grahn and McAuley (2009) observed robust individual differences in behavioral and brain responses to control and test rhythms that they interpreted as indicative of individual differences in beat sensitivity (Grahn & McAuley, 2009; McAuley, Frater, et al., 2006). Specifically, individuals who based their tempo judgments about the test rhythms on the implied 600-msec beat showed greater activation in cortical motor areas pre-

viously shown to be important for the internal generation of a beat and sensorimotor synchronization—most notably, premotor and supplementary motor areas (Grahn & Brett, 2007, 2009; Jantzen, Steinberg, & Kelso, 2005). For these participants, final intervals shorter than 600 msec tended to produce *speeding-up* responses, and final intervals longer than 600 msec tended to produce *slowing-down* responses for both the control and test rhythms. In contrast, individuals who showed less task-related activation in cortical motor areas important for beat perception responded *slowing down* the majority of the time to the test rhythms, regardless of final interval duration. Notably, although the two groups of participants responded similarly to control rhythms and showed no difference in

temporal discrimination thresholds, the observed activation differences in cortical motor regions linked to beat perception were still evident, suggesting that there was a fundamental difference in how the two groups of participants were processing the rhythms.

For the experiments reported here, the emphasis was not on examining individual differences per se but, rather, on using the tempo judgment paradigm developed by McAuley and colleagues (Grahn & McAuley, 2009; McAuley, Frater, et al., 2006) to compare auditory and visual rhythm processing. In particular, the pattern of tempo judgments about test rhythms provided us with a novel method for assessing the issue of auditory encoding of visual rhythms and, more generally, modality differences in rhythm processing. As in the earlier work by Grahn and McAuley (2009), participants in the present study were asked to make speeding-up or slowing-down judgments about control and test rhythms. Control rhythms always had an explicit temporal referent, whereas for test rhythms, the temporal referent was only implied by the temporal structure of the rhythms, providing a potential window on beat-based encoding of the rhythms. In both experiments, participants experienced both auditory and visual rhythms, with trials blocked by modality. In general, we reasoned that, if visual rhythms receive automatic auditory encoding, tempo judgments for auditory and visual rhythms with the same temporal structure should show a similar pattern of tempo judgments for both control and test rhythms. If, however, visual rhythms are not automatically encoded in an auditory form, we expected that sensitivity to the implied beat of the test rhythms (consistent with a beat-based encoding) would be lower for visual rhythms than for auditory rhythms and that the pattern of tempo judgments across modalities would differ.

In both experiments, we additionally considered the possibility that beat sensitivity for visual rhythms might vary with prior auditory exposure and tempo. If prior auditory exposure facilitates the auditory encoding of visual rhythms, beat sensitivity would be expected to be greater for visual rhythms with auditory exposure than for visual rhythms without auditory exposure. Moreover, given the previous results of Glenberg and colleagues (Glenberg & Jona, 1991; Glenberg et al., 1989) showing that encoding may be modulated by tempo, we expected that slowing down the rhythms (Experiment 2) might reduce the magnitude of any modality differences in rhythm processing.

EXPERIMENT 1

Method

Participants and Design. Twenty-nine individuals (18–48 years of age; 19 of them female) with self-reported normal hearing from the Bowling Green State University community participated in return for course credit or a monetary payment of $10. Musical training varied between participants ($M = 4$ years, $SD = 4.78$). The experiment implemented a 2 (modality: auditory, visual) $\times$ 2 (rhythm type: control, test) $\times$ 2 (order: auditory–visual, visual–auditory) $\times$ 6 (final interval) mixed-factorial design. Modality, rhythm type, and final interval were within-subjects factors, and order was a between-subjects factor. The participants completed two short sessions with modality blocked within a session and order

counterbalanced (auditory–visual order, $n = 16$; visual–auditory order, $n = 13$).

Stimuli and Equipment. Five-element (test rhythms) and four-element (control) rhythms are illustrated in Figure 1. Test and control rhythms were marked by 50-msec 440-Hz tones presented over Sennheiser HD 280 Pro headphones at a comfortable listening level (auditory condition) or by 50-msec black squares subtending ≈2.5º of visual angle presented on a white background (visual condition). Critically, the temporal characteristics of the rhythms in the auditory and visual conditions were the same. Test rhythms consisted of three initial elements marking two 300-msec IOIs, followed by two elements that specified a variable final interval (final IOI = 600 msec $\pm$ 4%, $\pm$ 12%, or $\pm$ 20%). The key aspect of the test rhythms was that a periodic 600-msec beat was implied (but not explicitly emphasized) by the temporal structure of the sequence (Povel & Essens, 1985). Control rhythms, in contrast, consisted of two initial elements that specified a 600-msec IOI, followed by two elements marking the same variable final IOI (600 msec $\pm$ 4%, $\pm$ 12%, or $\pm$ 20%). For both rhythm types (control or test), the initial group of elements (two or three) was separated from the final group of two elements by a 1,200-msec IOI.

Auditory stimuli were generated offline using Audacity Version 1.2.6 for Microsoft Windows. Stimulus presentation and response collection for both visual and auditory rhythms were controlled by E-Prime software (Psychology Software Tools). Accurate timing for visual sequences was accomplished in E-Prime through a combination of *prereleasing* events and synchronizing as closely as possible the timing of the visual stimulus sequences to multiples of the refresh rate of the monitor. Prereleasing allowed each event to be prepared during presentation of the preceding event. For a monitor with a 60-Hz refresh rate, the presentation durations for black squares, standard intervals, and gaps were accurate within ≅1–2 msec. Responses were made by pressing one of two labeled buttons on a response box.

Procedure. The participants judged on each trial whether the rhythm they heard/saw was speeding up or slowing down at the end. At the start of a session, the participants completed a single familiarization block consisting of four examples of each rhythm in the tested modality. Each session consisted of five blocks of 48 trials. Within a block, the participants made four responses to control and test rhythms for each of the six final intervals. Sessions lasted approximately 25 min, with a 10-min break between sessions. The entire experiment lasted approximately 1 h.

Data analysis. Proportions of speeding-up responses were calculated for each participant for each of the 12 trial types (2 rhythm types $\times$ 6 final intervals) per modality. Response proportions for test and control rhythms were fit with a signal detection model proposed by Grahn and McAuley (2009), in order to derive an estimate of sensitivity to the implied beat, $w \in [0, 1]$, for each participant in each modality. Most important for the present purposes, the continuous value of w estimated the extent to which tempo judgments about the test rhythms were based on the implied 600-msec beat period or the explicit 300-msec IOI marked by the first three tones of the test rhythms. Values of w closer to 1 indicate greater reliance of tempo judgments on the implied beat (corresponding to a 600-msec referent interval), whereas those closer to 0 indicate greater reliance on the explicit 300-msec referent marked by the initial three tones. Model fits allowed w and a temporal discrimination threshold parameter, σ, to vary in order to minimize the root-mean square error (RMSE) between the observed and predicted response proportions. Details of the model and the model-fitting procedure are provided in the Appendix.

Results and Discussion

Mean proportions of speeding-up responses as a function of the final interval are shown in Figure 2 (panel A, auditory rhythms; panel B, visual rhythms). Consistent with previous work, discrimination thresholds were

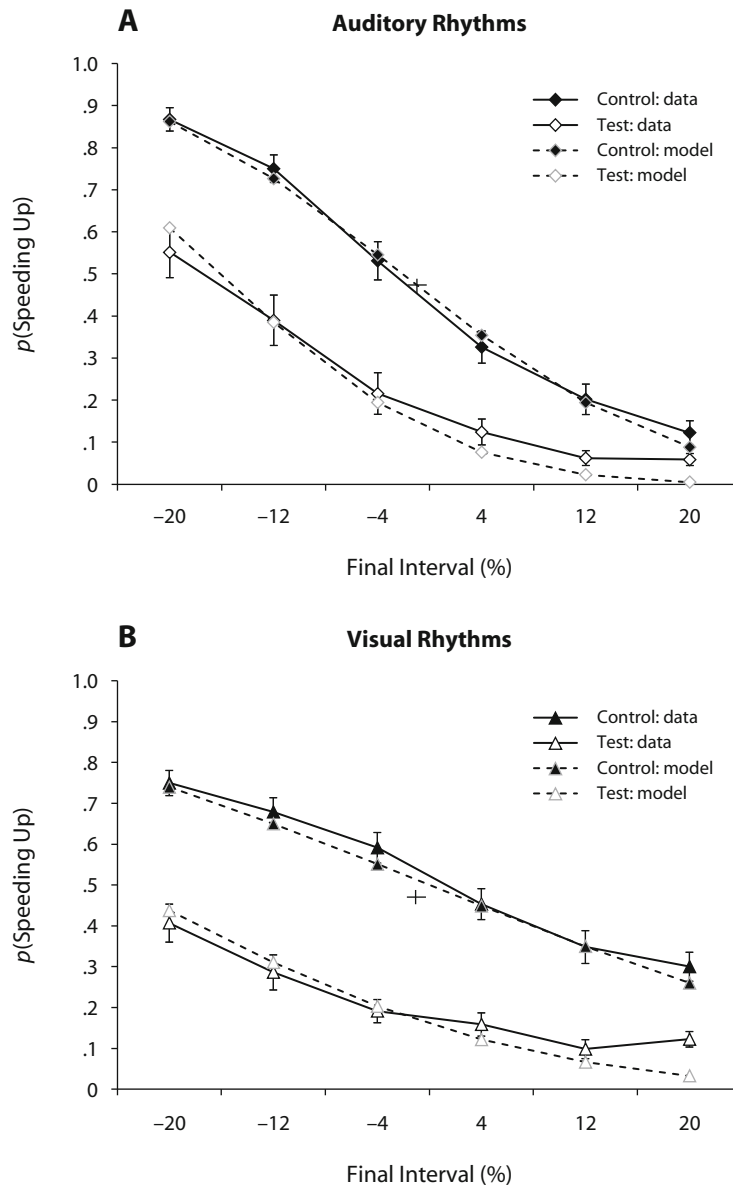**A**   **Auditory Rhythms**

B   **Visual Rhythms**

Figure 2. Experiment 1: Proportions of speeding-up responses as a function of the final interval (solid lines) and corresponding model fits (dotted lines) for (A) auditory rhythms (control and test) and (B) visual rhythms (control and test).

generally lower (greater tempo sensitivity) for auditory rhythms ($M = 13.5\% \pm 1.8\%$) than for visual rhythms ($M = 24.0\% \pm 2.5\%$) (Grondin, Meilleur-Wells, Ouellette, & Macar, 1998). Of primary interest were values of w, which, for the test rhythms, provide an index of sensitivity to the implied 600-msec beat. Higher values of w indicate greater reliance of tempo judgments on the implied 600-msec beat and, thus, greater likelihood of responding *speeding up* to final intervals shorter than 600 msec. Mean values of beat sensitivity, w, for auditory and visual control and test rhythms are shown in Figure 3 (panel A, auditory–visual order; panel B, visual–auditory order). For control rhythms, estimates of w are expected to be close to 1 for both modalities, given that most partici-

pants base their tempo judgments about control rhythms on the explicit temporal referent, which, in this case, was 600 msec. A 2 (modality) $\times$ 2 (rhythm type) repeated measures ANOVA on w revealed a main effect of modality [$F(1,28) = 10.48, p < .01\ \eta_p^2 = .27$], a main effect of rhythm type [$F(1,28) = 39.20, p < .001, \eta_p^2 = .58$], and critically, a significant modality $\times$ rhythm type interaction [$F(1,28) = 12.50, p < .001, \eta_p^2 = .31$].

Consistent with earlier work, values of w were lower for test rhythms ($M = .63 \pm .05$) with an implied 600-msec beat than for control rhythms ($M = .95 \pm .01$) with an explicit 600-msec beat [$t(28) = 6.28, p < .001$; see Grahn & McAuley, 2009]. Comparing across modalities, values of w for control rhythms approached 1, as was expected,

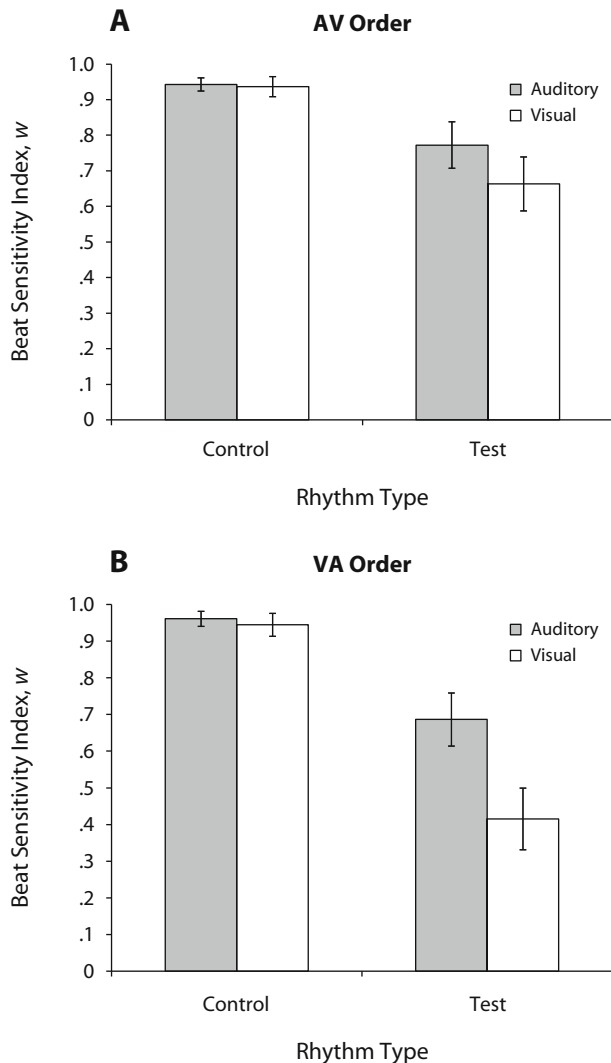## A   AV Order

B   VA Order

Figure 3. Experiment 1: Average values of beat sensitivity, *w*, for four-element (control) rhythms and five-element (test) rhythms for the (A) auditory–visual (AV) and (B) visual–auditory (VA) orders. Shaded bars illustrate auditory rhythms, and white bars show visual rhythms. Error bars indicate standard errors of the means.

and were not reliably different (auditory control rhythms, $M = .95 \pm .01$; visual control rhythms, $M = .94 \pm .02$) [$t(28) = 0.50$, $p = .62$]. Inconsistent with the automatic auditory-encoding hypothesis, however, *w* values were much greater for auditory test rhythms ($M = .73 \pm .05$) than for visual test rhythms ($M = .55 \pm .06$) [$t(28) = 3.64$, $p = .001$]. This lower value of *w* for visual test rhythms, as compared with auditory test rhythms, means that for final intervals shorter than 600 msec (negative contrasts), there was a tendency for the participants to make different tempo judgments for auditory and visual test rhythms, even though they had the same temporal structure (i.e., *speeding up* for auditory test rhythms and *slowing down* for visual test rhythms).

To consider more closely the extent to which auditory and visual test rhythms with the same temporal structure

yielded opposite tempo judgments, we examined proportions of speeding-up and slowing-down responses in both modalities for the shortest final interval (-20%) condition. This condition was selected because (1) only negative contrasts are expected to lead to opposite percepts if auditory rhythms are encoded with a beat and visual rhythms are not and (2) the -20% final interval should be the most detectable. Thus, if visual rhythms receive automatic auditory encoding, the -20% final interval will arguably be more likely to yield similar tempo perceptions across modality than when duration differences are less detectable and might differ simply because of chance levels of responding. Here, we constructed a difference metric by subtracting the proportions of speeding-up responses from .5 for each participant for each modality. Average values of this metric were positive for the auditory test rhythm ($M = .06 \pm .06$) and negative for the visual test rhythm ($M = -.09 \pm .05$), indicating that for the shortest final interval, auditory test rhythms were more likely to be judged as speeding up than as slowing down and, conversely, visual test rhythms were more likely to be judged as slowing down than as speeding up. A paired samples *t* test showed a reliable modality difference [$t(28) = 3.35$, $p < .01$, $d = 0.62$].

Next, we tested the hypothesis that the order of the auditory and visual blocks may influence the encoding of the test rhythms. Specifically, if there is an effect of prior auditory exposure on the encoding of the visual rhythms, we might expect to find that *w* values would be greater for visual test rhythms with prior auditory exposure (auditory–visual order) than for visual test rhythms without prior auditory exposure (visual–auditory order). Consistent with this hypothesis, values of *w* were significantly greater for visual test rhythms in the auditory–visual order ($M = .66 \pm .07$) than for those in the visual–auditory order ($M = .42 \pm .09$) [$t(27) = 2.19$, $p < .05$]. Critically, no effect of order was observed for auditory test rhythms [$t(27) = 0.89$, $p = .39$], suggesting that the effect of order observed for the visual test rhythms was not simply due to more practice with the task.

Finally, given earlier work using this paradigm showing individual differences in beat sensitivity, we considered the possibility that individual differences in beat sensitivity may have impacted modality differences in the encoding of the rhythms. To address this issue, a median split was performed on estimates of *w* for the auditory test rhythms in order to distinguish participants with high levels of beat sensitivity ($M = .92 \pm .02$) from those with lower levels of beat sensitivity ($M = .54 \pm .07$), and modality differences in *w* were considered separately for each group. Consistent with the results of the ANOVA involving all the participants, the participants with high levels of beat sensitivity showed similar values of *w* across modality for the control rhythms [$t(14) = 1.27$, $p = .22$] but values of *w* that were greater for auditory test rhythms than for visual test rhythms [$t(14) = 3.77$, $p < .01$]. For participants with lower levels of beat sensitivity, however, there were no modality differences in *w* values for either control rhythms [$t(13) = -0.56$, $p = .59$] or test rhythms [$t(13) = 1.61$, $p = .12$]. The latter finding suggests that

individual differences in beat sensitivity mediate the encoding of visual rhythms. To evaluate whether the effect of prior auditory exposure was influenced by sensitivity to an implied beat, we conducted 2 (modality) × 2 (rhythm type) × 2 (order) mixed measures ANOVAs separately for participants with high versus low levels of beat sensitivity. For participants high in beat sensitivity, the three-way interaction was significant [$F(1,8) = 10.16, p < .05, \eta_p^2 = .56$]; receiving the auditory condition first promoted extraction of a beat from visual test rhythms for this participant group. On the other hand, participants low in beat sensitivity did not show the same interaction [$F(1,8) = 0.26, p = .62$]; prior auditory exposure did not change perceptions of visual test rhythms.

In sum, there were four main results in Experiment 1. First, the participants showed greater sensitivity to an implied beat for rhythms presented in the auditory modality than for rhythms presented in the visual modality. A second and related result is that for what should be the most noticeably different short final interval (−20%), there was a tendency for the participants to report opposite tempo judgments across modality (speeding up for auditory test rhythms and slowing down for visual test rhythms). Third, visual beat sensitivity was modulated by prior auditory exposure. Sensitivity to the implied beat was greater for visual rhythms with prior auditory exposure than for rhythms without prior auditory exposure. Fourth, individual differences in auditory beat sensitivity modulated some of the observed effects. Taken together, these findings are inconsistent with the strong version of the Guttman et al. (2005) auditory-encoding hypothesis that visual rhythms receive obligatory and automatic auditory encoding. The present results, in contrast, support a weaker version of this hypothesis; auditory encoding of visual rhythms is not obligatory or automatic but can be encouraged by prior auditory exposure.

A question raised by this work is whether the greater beat sensitivity observed for auditory than for visual rhythms and the modulatory effect of prior auditory exposure are tempo specific or generalize to other tempi. All the test rhythms in Experiment 1 had an implied beat period of 600 msec, which is an especially salient periodicity that is close to people's preferred tempo for listening and tapping to music (McAuley, Jones, Holub, Johnston, & Miller, 2006; Parncutt, 1994). To examine whether the results of Experiment 1 would generalize to other tempi, a second experiment was conducted that was identical to the first, except that we slowed down both the auditory and visual rhythms so that the implied beat period for the test rhythms was 1,000 msec, rather than 600 msec. In general, we hypothesized that slower rhythms would reduce sensitivity to the implied beat, resulting in lower w scores for auditory rhythms, in particular. Support for this hypothesis would provide additional support for the validity of w as an index of beat sensitivity.

Another issue we were interested in exploring in Experiment 2 was the possibility that the visual rhythms in Experiment 1 were simply too fast for the participants to consistently pick up on the implied beat, as has been suggested in the earlier work on cross-modal comparisons of rhythm perception (Patel et al., 2005; Repp, 2003). Note that in Experiment 1, the shortest interval was 300 msec; however, for the slower rhythms in Experiment 2, the shortest interval was 500 msec, which may have made it easier for the participants to extract the implied beat of the visual test rhythms. If this was the case, values of w for visual test rhythms should increase for the slower tempo used in Experiment 2. Of additional interest in Experiment 2 was whether the effect of prior auditory exposure on visual rhythm perception observed in Experiment 1 would still be present for the slower rhythms in Experiment 2.

EXPERIMENT 2

Method

Participants and Design. Twenty-four undergraduates (18–25 years of age; 21 of them female) from Michigan State University with self-reported normal hearing participated in return for course credit. Musical training varied between participants ($M = 4.38$ years, $SD = 4.46$). The design for Experiment 2 was identical to that in Experiment 1. The participants completed two short sessions with modality blocked within a session and order counterbalanced (auditory–visual order, $n = 14$; visual–auditory order, $n = 10$).

Stimuli, Equipment, and Procedure. Test and control rhythms were presented at a slower tempo than in Experiment 1, using an implied beat period of 1,000 msec, rather than 600 msec. Test rhythms consisted of three initial elements marking two 500-msec IOIs, followed by two elements that specified a variable final interval (final IOI = 1,000 msec ± 4%, ± 12%, or ± 20%). The key aspect of the test rhythms was that a periodic 1,000-msec beat was implied (but not explicitly emphasized) by the temporal structure of the sequence. Control rhythms, in contrast, consisted of two initial elements that specified a 1,000-msec IOI, followed by two elements marking the same variable final IOI (1,000 msec ± 4%, ± 12%, or ± 20%). For both rhythm types (control or test), the initial group of elements (two or three) was separated from the final group of two elements by a 2,000-msec IOI. All aspects of the procedure, including methods of stimulus delivery and response collection, were the same as those in Experiment 1.

Results and Discussion

Proportions of speeding-up responses as a function of the final interval are shown in Figure 4 (panel A, auditory rhythms; panel B, visual rhythms). Consistent with previous work, thresholds were generally lower (greater tempo sensitivity) for auditory rhythms ($M = 16.2\% ± 2.5\%$) than for visual rhythms ($M = 26.0\% ± 3.0\%$) [$t(23) = -3.92, p < .01$]. As in Experiment 1, of primary interest were values of w, which, for the test rhythms, provided an index of sensitivity to the implied beat; the key difference in Experiment 2 was that all the rhythms were slower, resulting in an implied beat period of 1,000 msec, rather than 600 msec. Mean values of beat sensitivity, w, for auditory and visual control and test rhythms are shown in Figure 5 (panel A, auditory–visual order; panel B, visual–auditory order). A 2 (modality) × 2 (rhythm type) repeated measures ANOVA on w revealed a main effect of modality [$F(1,23) = 10.79, p < .01, \eta_p^2 = .32$], a main effect of rhythm type [$F(1,23) = 20.40, p < .001, \eta_p^2 = .47$], and critically, a significant modality × rhythm type interaction [$F(1,23) = 8.20, p < .01, \eta_p^2 = .26$].

Overall, w values for test rhythms ($M = .57 ± .04$) were significantly lower than w values for control rhythms
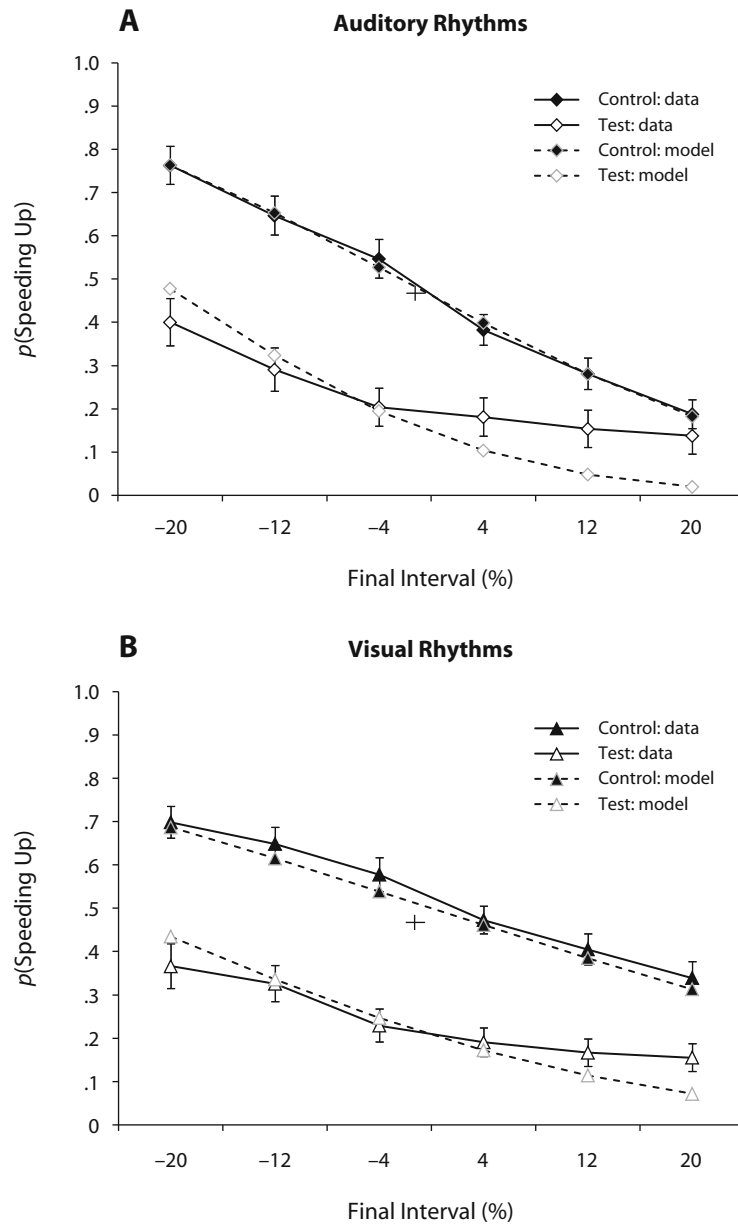
Figure 4. Experiment 2: Proportions of speeding-up responses as a function of the final interval (solid lines) and corresponding model fits (dotted lines) for (A) auditory rhythms (control and test) and (B) visual rhythms (control and test).

($M = .92 \pm .06$). Comparing across modality, values of w for control rhythms approached 1, as was expected, and were not reliably different (auditory control rhythms, $M = .93 \pm .04$; visual control rhythms, $M = .92 \pm .04$) [$t(23) = 0.72, p = .48$]. Values of w were lower for visual test rhythms ($M = .50 \pm .07$) than for auditory test rhythms ($M = .65 \pm .06$) [$t(23) = -3.71, p = .001$]. A planned comparison between the two experiments revealed that w values for auditory test rhythms were lower in Experiment 2 than in Experiment 1 [$t(58) = 1.77, p = .04$, one-tailed] but that w values for visual test rhythms did not differ between experiments [$t(58) = 0.62, p = .54$].

Similar to what was observed in Experiment 1, the lower value of w for visual test rhythms than for auditory test rhythms meant that for final intervals shorter than 1,000 msec, there was a tendency for the participants to make different tempo judgments for auditory and visual test rhythms, even though they had the same temporal structure (i.e., speeding up for auditory test rhythms and slowing down for visual test rhythms). To address this issue more directly, we examined (as in Experiment 1) proportions of speeding-up and slowing-down responses in both modalities for the shortest final interval (−20%) condition, where differences in the duration of the final

interval should be most detectable and, from the view of an automatic auditory-encoding hypothesis, would arguably be more likely to yield similar tempo perceptions across modalities, as compared with when duration differences are less detectable. Unlike in Experiment 1, average values for a difference metric, subtracting proportions of speeding-up responses from .5 for the shortest final interval for each participant for each modality, were negative for both the auditory test rhythms ($M = -.10 \pm .06$) and the visual test rhythms ($M = -.15 \pm .06$); this suggests that for even this shortest final interval, both auditory and visual test rhythms were more likely to be judged as slowing down than as speeding up.

Next, we consider whether prior auditory exposure had the same effect on visual rhythm perception as in Experiment 1. If prior auditory exposure similarly increased the extent to which the observers picked up on an implied beat in the visual rhythms, as was the case in Experiment 1, w values would be greater for visual test rhythms with prior auditory exposure (auditory–visual order) than for visual test rhythms without prior auditory exposure (visual–auditory order). In contrast to Experiment 1, we observed no effect of prior auditory exposure for rhythms presented at a slower tempo; that is, w values for visual test rhythms in the auditory–visual order ($M = .48 \pm .09$) were not significantly different from w values for visual test rhythms in the visual–auditory order ($M = .53 \pm .11$) [$t(22) = -0.36, p = .73$; see Figure 5].

As in Experiment 1, we considered the possibility that individual differences in beat sensitivity may have impacted modality differences in the encoding of the rhythms by separately examining modality differences for participants with high levels of beat sensitivity ($M = .85 \pm .03$) and low levels of beat sensitivity ($M = .45 \pm .07$). Consistent with the results of the ANOVA involving all the participants, the participants with high levels of beat sensitivity showed similar w values across modalities for the control rhythms [$t(11) = 0.32, p = .75$] but values of w that were greater for auditory test rhythms than for visual test rhythms [$t(11) = 3.03, p < .05$]. For the participants with lower levels of beat sensitivity, there were no modality differences in w values for control rhythms [$t(11) = 1.31, p = .22$]. However, in contrast to the results of Experiment 1, w values for auditory and visual test rhythms differed significantly [$t(11) = 12.87, p < .05$]. This finding suggests that individual differences in beat sensitivity for slow rhythms did not impact modality effects in the encoding of rhythms. To evaluate whether the effect of prior auditory exposure was influenced by sensitivity to an implied beat, we conducted 2 (modality) $\times$ 2 (rhythm type) $\times$ 2 (order) mixed measures ANOVAs separately for participants with high versus low levels of beat sensitivity. For the participants high in beat sensitivity, only the modality $\times$ rhythm type interaction was marginally significant [$F(1,10) = 4.77$, $p = .05, \eta_p^2 = .32$]. For the participants low in beat sensitivity, the main effects of modality [$F(1,10) = 9.46, p < .05, \eta_p^2 = .49$] and rhythm type [$F(1,10) = 106.26, p < .001, \eta_p^2 = .91$] and the modality $\times$ rhythm type interaction [$F(1,10) = 5.28, p < .05, \eta_p^2 = .35$] reached significance. Thus, in Experiment 2, receiving the auditory condition
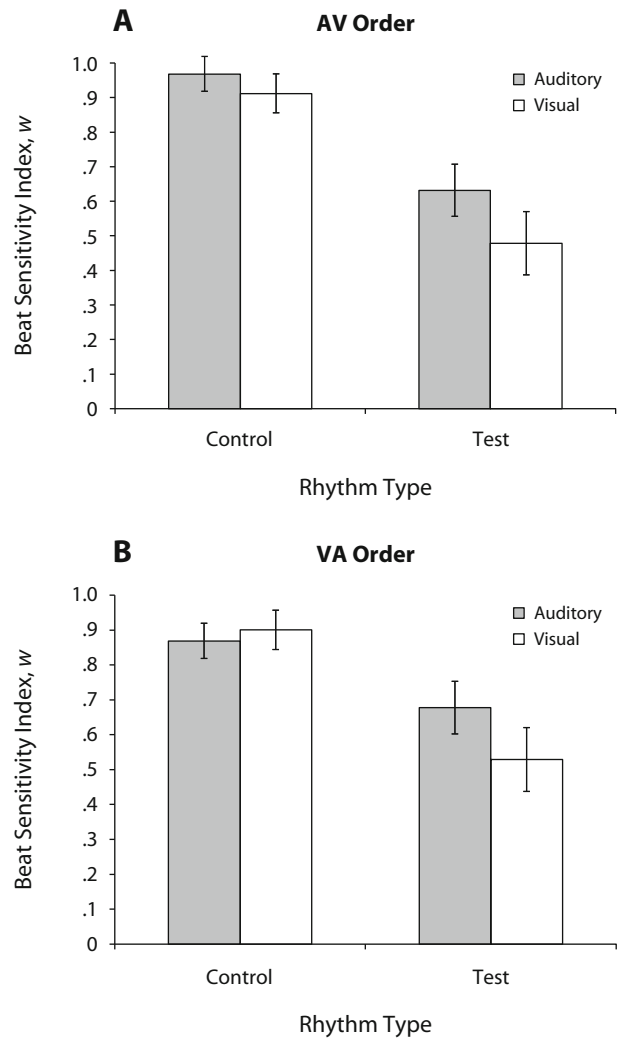

Figure 5. Experiment 2: Average values of beat sensitivity, w, for four-element (control) rhythms and five-element (test) rhythms for the (A) auditory–visual (AV) and (B) visual–auditory (VA) orders. Shaded bars illustrate auditory rhythms, and white bars show visual rhythms. Error bars indicate standard errors of the means.

first did not promote extraction of a beat from visual test rhythms for either participant group.

GENERAL DISCUSSION

Two experiments compared auditory and visual rhythm perception in order to revisit a claim by Guttman et al. (2005) that visual rhythms receive automatic and obligatory auditory encoding. Four main findings contribute to the current understanding of modality effects in rhythm processing. First, with respect to the question of encoding, a comparison of the pattern of tempo judgments across modality and the computation of a model-based index of beat sensitivity revealed greater sensitivity to an implied beat for auditory rhythms than for visual rhythms (Experiments 1 and 2). Noteworthy here was that, for the shortest final interval (-20%) in Experiment 1, auditory

test rhythms were more likely to be judged as speeding up than as slowing down and visual test rhythms were more likely to be judged as slowing down than as speeding up; thus, inconsistent with the strong version of the auditory-encoding hypothesis proposed by Guttman and colleagues, there was a tendency for auditory and visual rhythms with the same temporal structure to yield opposite tempo judgments (at least in Experiment 1).

Second, there was an asymmetric effect of presentation order, such that prior auditory exposure modulated judgments about visual rhythms, but not the reverse (Experiment 1). Specifically, hearing the auditory rhythms prior to watching visual sequences (auditory–visual order) increased visual beat sensitivity, relative to watching visual sequences without prior auditory exposure (visual–auditory order). Watching visual sequences prior to hearing auditory rhythms did not similarly impact tempo judgments for the auditory sequences; auditory beat sensitivity was generally high for both the auditory–visual and the visual–auditory orders.

Third, Experiment 2 highlighted the importance of stimulus tempo. Slowing down the rhythms in Experiment 2 so that the implied beat period was 1,000 msec, rather than 600 msec, reduced beat sensitivity for the auditory test rhythms, as compared with Experiment 1, but was not found to affect beat sensitivity for the visual rhythms. Slowing down the rhythms additionally eliminated the effect of prior auditory exposure on visual beat sensitivity.

Finally, a less central finding was some preliminary evidence that modality differences may have been modulated by individual variation in the extent to which the participants picked up on the implied beat of auditory test rhythms. Individuals showing greater sensitivity to the implied beat were more likely to make opposite tempo judgments about auditory and visual rhythms with the same temporal structure. Thus, these individuals, more so than others, seem to have shown distinct encoding of auditory and visual rhythms. The last result is perhaps not that surprising, given that the participants who were insensitive to the implied beat in the auditory rhythms would, arguably, also likely have been insensitive to the implied beat of visual rhythms.

Taken together, the findings from the two experiments support the view that auditory encoding of visual rhythms is neither obligatory nor automatic. The strongest evidence against the obligatory automatic-encoding view proposed by Guttman et al. (2005) is that visual and auditory rhythms with the same temporal structure produced reliably opposite tempo judgments for a condition in which the magnitude of final interval difference was large and, thus, not likely to be attributable to poorer temporal resolution in the visual modality. The results further suggest that although auditory encoding of visual rhythms does not appear to be obligatory or automatic, prior auditory exposure can encourage the auditory encoding of visual rhythms, at least within a limited temporal range. Relevant to this point, Guttman et al. randomized the presentation of auditory, visual, and cross-modal rhythm pairs within an experimental block. Thus, visual presentations were likely preceded, within the span of a few trials, by an auditory presentation of the same rhythm. In the context of the

present findings, this aspect of the Guttman et al. design suggests that auditory exposure would have increased the likelihood that visual rhythms received auditory encoding, rather than providing definitive support that auditory encoding of visual rhythms was obligatory and automatic. More generally, the present results are consistent with previous research on cross-modal comparisons of rhythm perception, which has revealed auditory influences on visual rhythm perception, but not the reverse (Collier & Logan, 2000; Glenberg & Jona, 1991; Glenberg et al., 1989; Repp, 2003; Repp & Penel, 2002).

Several issues raised by this research have the potential to affect the interpretation of the results and deserve careful consideration. First, one issue concerns the possibility that the participants did not fully understand the tempo judgment task and, so, differences in w values could potentially reflect the adoption of different strategies. Here, it is important to note that the participants received both auditory and visual rhythms, so to explain the observed modality differences from this perspective would require positing that the participants shifted their strategy across modalities. Nonetheless, we have reasonable evidence that the instructions were clear. For both experiments, we asked the participants to rate their understanding of the task on a scale ranging from 1 (*I did not understand at all*) to 6 (*I understood exactly what to do*). Ratings of understanding were uniformly high (Experiment 1, $M = 5.8 \pm 0.1$; Experiment 2, $M = 5.6 \pm 0.1$). In addition, in other work, we have also shown that this paradigm yields high test–retest reliability and, so, participants respond consistently across different administrations of the assessment. Also relevant with respect to this issue is that the largest modality differences are found for those participants most sensitive to the beat of the auditory rhythms. Thus, if the instructions were even clearer or more explicit about the presence of the implied beat, it would likely only further magnify the modality differences that we observed. The point here is that even if one labels picking up the beat the "correct" way to respond to the sequence, those participants who respond most "correctly" to the auditory rhythms are not likely to respond "correctly" to the visual rhythms, which is evidence against the automatic auditory encoding of visual rhythms.

Second, it is important to consider whether the pattern of differences observed across modalities can be explained simply by differences in discrimination thresholds. The observation here is that people have difficulty perceiving temporal characteristics of flashes of light, so it is, at first glance, conceivable that sensory limitations of the visual system—and not rhythm perception per se—may be responsible for the results. There are several sources of evidence that suggest that this is not the driving factor in the observed modality differences in rhythm processing. Although thresholds are worse for the visual rhythms than for the auditory rhythms, individuals do make consistent judgments for both auditory and visual control rhythms, with w values close to 1 in both cases. Moreover, the fact that the largest differences in responses to the auditory and visual test rhythms occur with the −20% final interval condition (where detection of speeding up is best for both

auditory and visual control rhythms) suggests that limitations in perceiving the temporal characteristics of the visual rhythm markers is not driving all of the modality differences in beat sensitivity.

Third, an additional possibility is that it may simply be more difficult to extract a beat from the auditory code of a visual rhythm than from an auditory rhythm.[1] That is, given that visual rhythms and auditory rhythms are processed by different sensory systems, the auditory recoding of a visual rhythm, if any, must occur at a later stage of processing. This may bypass some close links with the motor system that are involved in beat induction. Thus, the auditory code of a visual rhythm may be more abstract and disembodied than a real auditory rhythm. This view might, at first glance, appear to account for the results, even if auditory recoding of visual rhythms were obligatory. However, we do not view this interpretation as very parsimonious, since it suggests that the auditory encoding of visual rhythms is different from the auditory encoding of auditory rhythms; that is, there are two types of auditory encoding. We agree that, in theory, this is possible, but this is different from what Guttman et al. (2005) proposed with an obligatory automatic auditory-encoding hypothesis. Addressing this issue comprehensively will require additional research.

Finally, one could argue that listening to or watching the test rhythms does not require beat-based encoding in order for the pattern of tempo judgments in response to test and control rhythms to be similar. From this perspective, it is important to note that the task could be treated as a simple interval comparison, where the first interval was sometimes (in the case of test rhythms) bisected by a third tone. Although a judgment about tempo seems to require establishment of a beat tempo, it is arguably not necessary to extract a beat to compare two intervals. One potential alternative is that a tempo judgment could be derived indirectly from an interval comparison. A recent fMRI study by Grahn and McAuley (2009) argued against this possibility. This research revealed differences in patterns of brain activation for individuals who responded similarly to control and test rhythms (high beat sensitivity) and individuals who responded differently to control rhythms and test rhythms (low beat sensitivity). Supporting the validity of the beat sensitivity index, w, brain activation for individuals with high beat sensitivity was increased, relative to those with low beat sensitivity, in areas that have been implicated in the internal generation of a beat (Grahn & Brett, 2007, 2009). Nonetheless, it still could be argued that these differences in brain activation (as well as similar patterns of responses for control and test rhythms) could be attributable to the use of, for example, an interval concatenation strategy, whereby individuals make a tempo judgment by concatenating the neighboring 300-msec intervals and comparing this duration with the final interval. This interpretation still does not seem plausible to us. Notably, McAuley and Jones (2003) provided a fairly extensive test of a variety of interval models that included concatenation, and none of these were found to provide a good account of the data. More generally, it is important to emphasize that we are not making claims here about the mechanism that participants use to extract a beat. Inde-

pendently of mechanism, the fact that there are consistent differences across modalities is informative about the nature of encoding of auditory and visual rhythms.

More broadly, the present findings contribute to the growing body of evidence showing (1) auditory advantages in rhythm and tempo perception (Collier & Logan, 2000; Glenberg & Jona, 1991; Glenberg et al., 1989; Grondin & McAuley, 2009; Patel et al., 2005) and (2) asymmetric interactions (Guttman et al., 2005; Repp & Penel, 2002, 2004). A general interpretation of this work is that it supports a *modality-appropriateness hypothesis*, whereby the "best" modality for perception and encoding of a given domain receives priority (Welch & Warren, 1980). According to this view, the auditory modality receives priority for the encoding of temporal information, and the visual modality receives priority for the encoding of spatial information (Glenberg & Jona, 1991; Welch & Warren, 1980). Thus, when temporal (spatial) information is presented to the visual (auditory) modality, there is an automatic recoding into the appropriate modality, which, in cases of conflicting auditory and visual information, can lead to illusions such as the ventriloquism effect or auditory driving (Jack & Thurlow, 1973; Rencanzone, 2003).

The present results suggest that modality appropriateness by itself is insufficient and too general to explain some of the subtle differences between auditory and visual rhythm processing. When this literature on modality differences in rhythm perception is considered in its entirety, it seems clear that a number of factors have the potential to affect the auditory encoding of visual rhythms; here, we have identified two—namely, prior auditory exposure and stimulus tempo. The effects of tempo reported here are consistent with the results of earlier work by Glenberg and colleagues (Glenberg & Jona, 1991; Glenberg et al., 1989), as well as Collier and Logan (2000). The novel effect of prior auditory exposure reported here raises a number of interesting questions deserving of future study. First, does the prior auditory exposure require active task performance, or can it occur with passive listening? Second, does the observed effect of auditory exposure on visual rhythm perception require presentation of rhythms with the same absolute beat period, or can it be triggered by auditory rhythms with the same temporal structure (relative timing) but different absolute beat periods? Third, what role does auditory imagery play? That is, could the same effect of prior auditory exposure be found by having people simply imagine the tones in their head or by imaging tapping?

Neural evidence on modality differences in rhythm processing is reminiscent of an early claim by Fraisse (1948), who argued that the motor system is more responsive to auditory input than to visual input. The possibility of an auditory recoding of visually presented rhythms was also recently considered by Karabanov, Blom, Forsman, and Ullén (2009). In this study, the focus was on brain activity associated with reproduction of rhythms that were learned prior to scanning in either the auditory or the visual modality. During scanning, a pacing signal was presented auditorily or visually. The key finding was that reproduction of rhythms that were both trained and paced in the visual modality activated essentially the same auditory–motor

network of brain areas as did the reproduction of auditory rhythms. There was additional deactivation of the left angular gyrus following auditory pacing, as compared with visual pacing. On the basis of these results, Karabanov et al. suggested that the rhythms in both modalities were transformed into a common auditory–motor representation and that deactivation of the angular gyrus following auditory pacing represented cross-modal inhibition of areas involved in visual processing. The present experiments suggest that the motor system can be made to be more responsive to visual input by providing auditory exposure. Ongoing work is investigating this claim by combining auditory and visual rhythms with fMRI. Here, we have preliminary evidence that visual rhythms with prior auditory exposure produce greater activation in brain regions previously linked to perception and the production of auditory rhythms (Grahn, Henry, & McAuley, 2009).

Conclusions

Two experiments revisited a claim by Guttman et al. (2005) that visual rhythms receive obligatory auditory encoding. The approach taken in this research was to use the direction of tempo judgments about rhythms with an implied beat to make inferences about the nature of encoding. Overall, the patterns of tempo judgments about auditory and visual rhythms were different in critical ways that suggested that auditory and visual rhythms were being encoded differently. However, auditory exposure appears to increase the likelihood that tempo judgments about visual rhythms will show a pattern similar to that for tempo judgments about auditory rhythms. In sum, the findings from this work provide evidence that although auditory encoding of visual rhythms is neither obligatory nor automatic, it can be encouraged with prior auditory exposure; moreover, these effects may be found only within a limited temporal range.

REFERENCES

Bertelson, P., & Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin & Review*, **5**, 482-489.

Collier, G. L., & Logan, G. (2000). Modality differences in short-term memory for rhythms. *Memory & Cognition*, **28**, 529-538.

Fendrich, R., & Corballis, P. M. (2001). The temporal cross-capture of audition and vision. *Perception & Psychophysics*, **63**, 719-725.

Fraisse, P. (1948). Rythmes auditifs et rythmes visuels [Auditory and visual rhythms]. *L'Année Psychologique*, **49**, 21-41.

Glenberg, A. M., & Jona, M. (1991). Temporal coding in rhythm tasks revealed by modality effects. *Memory & Cognition*, **19**, 514-522.

Glenberg, A. M., Mann, S., Altman, L., Forman, T., & Procise, S. (1989). Modality effects in the coding and reproduction of rhythms. *Memory & Cognition*, **17**, 373-383.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, **19**, 893-906.

Grahn, J. A., & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex*, **45**, 54-61.

Grahn, J. A., Henry, M. J., & McAuley, J. D. (2009). Effects of prior auditory exposure on brain activity during visual rhythm perception [Abstract] (Prog. No. 94.18). *2009 Neuroscience Meeting Planner*. Chicago: Society for Neuroscience.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage*, **47**, 1894-1903.

Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. *Perception*, **38**, 1542-1559.

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. *Psychological Research*, **61**, 261-268.

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see: Auditory encoding of visual temporal sequences. *Psychological Science*, **16**, 228-235.

Hay, J. C., Pick, H. L., & Ikeda, K. (1965). Visual capture produced by prism spectacles. *Psychological Science*, **2**, 215-216.

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

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual & Motor Skills*, **37**, 967-979.

Jantzen, K. J., Steinberg, F. L., & Kelso, J. A. S. (2005). Functional MRI reveals the existence of modality and coordination-dependent timing networks. *NeuroImage*, **25**, 1031-1042.

Karabanov, A., Blom, Ö., Forsman, L., & Ullén, F. (2009). The dorsal auditory pathway is involved in performance of both visual and auditory rhythms. *NeuroImage*, **44**, 480-488.

McAuley, J. D., Frater, D., Janke, K., & Miller, N. S. (2006, August). *Detecting changes in timing: Evidence for two modes of listening*. Paper presented at the 9th International Conference on Music Perception and Cognition, Bologna.

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 1102-1125.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, **135**, 348-367.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, **11**, 409-464.

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, **163**, 226-238.

Pick, H. L., Warren, D. H., & Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception & Psychophysics*, **6**, 203-205.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, **2**, 411-440.

Rencanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology*, **89**, 1078-1093.

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits/costs of interval subdivision. *Journal of Motor Behavior*, **35**, 355-370.

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 & Performance*, **28**, 1085-1099.

Repp, B. H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, **68**, 252-270.

Stratton, G. M. (1897). Vision without inversion of the retinal image. *Psychological Review*, **4**, 341-360.

Thomas, G. J. (1941). Experimental study of the influence of vision on sound localization. *Journal of Experimental Psychology*, **28**, 163-177.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, **88**, 638-667.

NOTE

1. We thank Bruno Repp for bringing this possibility to our attention.

APPENDIX

To characterize the pattern of tempo judgments about control and test rhythms, Grahn and McAuley (2009) developed a model-based index, w, that measures the extent to which tempo judgments about test rhythms are based on the explicit temporal referent marked by the first three tones of the sequence (300 msec in Experiment 1 and 500 msec in Experiment 2) or the implied temporal referent (600 msec in Experiment 1 and 1,000 msec in Experiment 2). To produce a tempo judgment, the model assumes that participants' speeding-up/slowing-down judgments on a given trial involve simultaneous consideration of both referents, where the longer temporal referent corresponds to the implied beat. The weight parameter, $w \in [0, 1]$, quantifies the relative contribution of each temporal referent to tempo judgments and thus provides a measure of sensitivity to the implied beat. Specifically, for each referent, P, a temporal contrast metric, C_i (Grahn & McAuley, 2009; McAuley & Jones, 2003), is calculated for each final interval, T_i, where C_i is given by

$$C_i = \frac{(T_i - P)}{P}.$$

Values of temporal contrast for each referent are assumed to be normally distributed with standard deviation σ; these values are then z-transformed and combined using a simple weighted average:

$$z = (1 - w)z_{i_{\text{short}}} + wz_{i_{\text{long}}}.$$

Here, $z_{i_{\text{short}}}$ and $z_{i_{\text{long}}}$ refer to the values of z for a specific final interval, T_i, for the short and long temporal referents, respectively. Predicted proportions of speeding-up responses, P(Speeding Up), for each final interval, T_i, can then be generated using the standard cumulative normal distribution function:

$$P(\text{Speeding Up}) = 1 - \Phi(z).$$

Model fits allow both $w \in [0, 1]$ and σ to vary, minimizing the RMSE between the observed and predicted response proportions. Most important for the present purposes, the estimated values of w provide an estimate of the extent to which tempo judgments about the five-element (test) rhythms are based on the explicit temporal referent (300 msec in Experiment 1 and 500 msec in Experiment 2) or the implied temporal referent (600 msec in Experiment 1 and 1,000 msec in Experiment 2), with the latter consistent with a beat-based encoding of the rhythm.