Psychophysiology, 53 (2016), 198–208. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/psyp.12559

How modality specific is processing of auditory and visual rhythms?

AMANDA C. PASINSKI,[a] J. DEVIN MCAULEY,[b] AND JOEL S. SNYDER[a]

[a]Department of Psychology, University of Nevada Las Vegas, Las Vegas, Nevada, USA
[b]Department of Psychology, Michigan State University, East Lansing, Michigan, USA

Abstract

The present study used ERPs to test the extent to which temporal processing is modality specific or modality general. Participants were presented with auditory and visual temporal patterns that consisted of initial two- or three-event beginning patterns. This delineated a constant standard time interval, followed by a two-event ending pattern delineating a variable test interval. Participants judged whether they perceived the pattern as a whole to be speeding up or slowing down. The contingent negative variation (CNV), a negative potential reflecting temporal expectancy, showed a larger amplitude for the auditory modality compared to the visual modality but a high degree of similarity in scalp voltage patterns across modalities, suggesting that the CNV arises from modality-general processes. A late, memory-dependent positive component (P3) also showed similar patterns across modalities.

Descriptors: Modality specific, Modality general, Auditory rhythms, Visual rhythms, CNV, P3

A wide range of everyday activities depends on having access to accurate and precise information about event timing. A large number of studies have been conducted on the nature of temporal processing (Grondin, 2001), and whether time perception is best explained by modality-specific or modality-general processes (Ivry & Schlerf, 2008). However, data have been equivocal due to a number of conflicting findings in the literature involving both behavioral and neuroimaging techniques (Bueti, Bahrami, & Walsh, 2008; Bueti, Walsh, Frith, & Rees, 2008; Lapid, Ulrich, & Rammsayer, 2009; Rao, Mayer, & Harrington, 2001; Wright, Buonomano, Mahncke, & Merzenich, 1997). One reason previous studies may have come to different conclusions about the modality specificity of timing is that the techniques used generally have not been able to distinguish between different stages of processing. For example, some aspects of performing a timing task (e.g., calculating time intervals vs. remembering time intervals) may be modality specific, whereas other stages of processing may be modality general, as has been suggested in some behavioral studies (Gamache & Grondin, 2010; Lustig & Meck, 2011). Thus, using a technique with high temporal precision, such as ERPs, might better distinguish between those aspects of timing that are modality specific and those that are modality general.

Toward that end, the present study used a well-studied tempo-discrimination paradigm (Grahn, Henry, & McAuley, 2011; Grahn & McAuley, 2009; Snyder, Pasinski, & McAuley, 2011) to examine auditory and visual ERP components that have been shown in previous research to index different aspects of timing. Auditory and visual stimuli consisted of an initial two- or three-event pattern that delineated a constant standard time interval of 600 ms, followed by two events delineating a variable test interval (see Figure 1) that was either longer or shorter than 600 ms. Participants' task was to judge whether they perceived the pattern as speeding up or slowing down, effectively comparing the variable test interval to the initial standard (referent) interval. These four- and five-event patterns have been used in the past to examine individual differences in rhythm processing strategies (Grahn et al., 2011; Grahn & McAuley, 2009; Snyder et al., 2011). In particular, while the four-event pattern can be considered a simple interval discrimination task, the five-event pattern has an inherent ambiguity because participants can either find the overall 600-ms duration of the first three events most salient or instead attend to the individual 300-ms intervals. Beat-based listening (Large & Jones, 1999; Large & Snyder, 2009; McAuley, 1995) is a strategy in which temporal encoding is facilitated by perception of a quasiperiodic pattern of accentuations or beats, and intervals are encoded in terms of these beats. This strategy can explain the perception of a 600-ms interval in the first three events of the five-event pattern. Interval-based listening (Ivry, 1996) is the alternative strategy in which listeners attend to the 300-ms interval. Importantly, with the five-event pattern, beat-based listening results in a steep psychophysical

---

The authors would like to thank Erin Hannon, Jefferson Kinney, Kim Barchard, Elisa Kim Fromboluti, and Kate Jones for their helpful suggestions on this project, and Dr. Molly Henry for her contributions to the initial stages of this project.

Address correspondence to: Joel S. Snyder, Department of Psychology, University of Nevada Las Vegas, 4505 Maryland Parkway—Mail Stop 5030, Las Vegas, Nevada 89154-5030, USA. E-mail: Joel.Snyder@unlv.edu

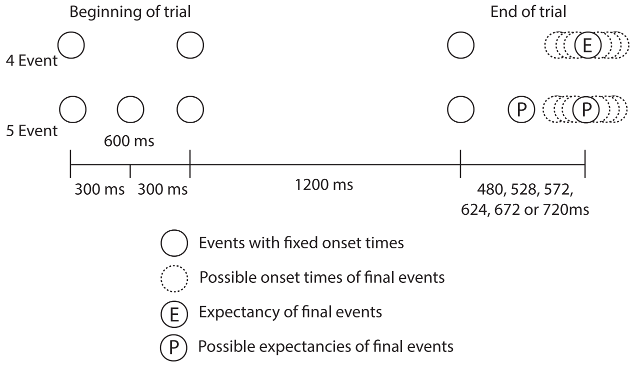

Figure 1. Five-event and four-event patterns. Final intervals reflect ± 4%, ± 12%, or ± 20% of the implied 600 ms IOI.

function, with less perception of speeding up when the final interval is longer in duration. In contrast, interval-based listening results in a flatter function as long as the 300-ms intervals of the initial three events are shorter than the final interval, which is usually the case. Given that there is only limited evidence for beat-based processing in the visual domain (also see Grahn, 2012; Patel, Iversen, Chen, & Repp, 2005), we expect a relatively flat function for the visual five-event pattern, which is more typical of interval-based processing.

In a previous ERP study using this paradigm with auditory stimuli, we found a contingent negative variation (CNV), a slow negative potential that occurred starting after the penultimate tone and ending around the onset of the final tone (Snyder et al., 2011). This was consistent with previous research implicating the CNV in temporal expectation and the encoding of time intervals (N'Diaye, Ragot, Garnero, & Pouthas, 2004; Pfeuty, Ragot, & Pouthas, 2003; Walter, Cooper, Aldridge, McCallum, & Winter, 1964). We also found a modulation of the P3 response—a late positive potential—after the onset of the final tone. This P3 modulation likely indicates an effect on processes related to change detection, stimulus categorization, and/or the updating of the memory trace of the encoded time interval (Linden, 2005; Polich, 2007). Both the CNV and P3 were found to reflect the extent to which listeners engaged in beat- versus interval-based processing. The goal of the current study is to use these two ERPs as indicators, not of beat versus interval processing, but rather the extent to which listeners engage in modality-specific or modality-general processing of timing during rhythm discrimination.

The CNV in particular could be a useful index of modality differences in temporal processing because it is specifically found during timing tasks. For example, during a time discrimination task, the CNV was sustained in amplitude for the duration of the standard interval, but on trials in which a longer interval was presented, it rapidly returned to a baseline level (Macar & Vidal, 2003), suggesting that it indexes durations stored in memory rather than the duration of a stimulus on a given trial. In another study, the CNV increased in amplitude during the encoding of a repeating time interval (i.e., a metronome-like pattern) but the amplitude increasing had ceased after a specific number of repetitions corresponding to the interval that generally results in the maximal perceptual benefit of repetition (Pfeuty et al., 2003). Importantly, some previous studies have compared the CNV for time intervals defined by events from different sensory modalities. Macar and Vidal (2003) measured the CNV for auditory and tactile intervals and for both modalities found the CNV to be maximal in amplitude at the same central-parietal midline scalp location (i.e., CPz). In contrast, N'Diaye et al. (2004) compared auditory and visual intervals and found that the CNV had larger amplitudes at frontocentral electrodes for the auditory intervals and larger amplitudes at posterior electrodes for the visual intervals, suggesting modality-specific sources of the CNV that were partially confirmed by source modeling. However, one important caveat is that this study examined filled intervals (i.e., durations that were defined by the time from onset to offset) rather than empty intervals (i.e., durations defined by onset to onset of two brief events), which could have resulted in sustained sensory activity in addition to a true CNV response (also see Pfeuty, Ragot, & Pouthas, 2008). More recent studies found better discrimination of empty intervals marked by auditory events than visual events, and further found that the CNV was larger and increased over time more for auditory intervals compared to visual intervals at frontal electrodes but not at parietal electrodes (Gontier, Hasuo, Mitsudo, & Grondin, 2013; Hasuo et al., 2014).

In the current study, we were interested in comparing ERPs during auditory and visual rhythm discrimination in order to distinguish between modality-specific and modality-general timing processes. Importantly, this paradigm uses empty time intervals, which are helpful for reducing the likelihood of recording sustained sensory responses (cf. N'Diaye et al., 2004). Given the evidence that some timing tasks involve modality-specific processing, we expected that the CNV, which is closely tied to temporal processing, would show topographic scalp voltage distributions that are distinct for the two modalities. For example, one study found that visual motion-processing areas are only involved in visual time interval discrimination but not auditory discrimination (Bueti, Bahrami, & Walsh, 2008), while another study found that rhythm reproduction of auditory and visual rhythms activated sensory association cortex areas specific for the two modalities (Penhune, Zatorre, & Evans, 1998). In contrast, we expected the P3 to show a more modality-general topographic pattern because it is thought to index general cognitive mechanisms that are not specific to timing tasks (Polich, 2007; Snyder, Yerkes, & Pitts, 2015).

In addition to examining CNV and P3 amplitude for auditory and visual modalities, we also examined scalp topographic distributions of ERP voltage in order to determine whether auditory and visual responses arise from distinct processes, as has been done in previous studies (e.g., Macar & Vidal, 2003; N'Diaye et al., 2004). In order to provide quantitative evidence about differences across modalities, we entered amplitude-normalized data from all of our electrodes into statistical analyses, with the expectation that modality differences would be indicated by an interaction between modality and electrode. Furthermore, we used the same data in correlation analyses across and within modalities with the expectation that similarities would be indicated by significant positive correlation coefficients. Larger within-modality correlations compared to across-modality correlations can provide evidence for modality specificity, along with the analyses of variance (ANOVAs). These topography analyses were performed with both voltage and current source density (CSD) transformations. Unlike voltage, CSD mapping is reference free; furthermore, it sharpens and simplifies EEG topographies, and can also result in measures that more closely represent the underlying current generators (Nunez & Srinivasan, 2006; Tenke & Kayser, 2005).

Method

Participants

Thirty-one participants (14 male) with normal hearing ($\leq$ 30 dB from 250–4000 Hz) were recruited from the University of Nevada, Las Vegas psychology participant pool. Participants were 18–47 years old ($M = 22.9$) with no prior history of substance abuse. Participants received course credit for their participation after providing written informed consent.

Materials and Procedure

The auditory stimulus used in the rhythmic patterns was generated offline in MATLAB (The MathWorks, Inc., Natick, MA) and consisted of a single pure tone (70 dB SPL, 440 Hz, 50 ms in duration, including 5-ms rise/fall times). Two types of auditory patterns were created using this one tone (Figure 1). The five-event pattern consisted of three initial tones separated by two 300-ms interonset intervals (IOIs) followed by two tones separated by a variable final IOI (600 ms $\pm \Delta T$). This resulted in a five-event pattern with a periodic 600-ms beat that was implied (but not explicitly emphasized) by the temporal structure of the first three tones of the pattern (Povel & Essens, 1985). A four-event pattern consisted of two initial tones separated by a 600-ms IOI followed by two tones separated by the same variable final IOI (600 ms $\pm \Delta T$) as the five-event pattern. Thus, the only physical difference between the two patterns was that the second tone from the five-event pattern was missing in the four-event pattern. For both patterns, the initial group of tones was separated from the final group of tones by an IOI of 1,200 ms. Final IOIs of the patterns were variable 600 ms $\pm \Delta T$, where ΔT equaled $\pm 4\%, \pm 12\%$, or $\pm 20\%$ of the implied 600 ms IOI (480, 528, 576, 624, 672, or 720 ms), and presented in a random order within each block.

The visual stimulus used to construct the corresponding visual patterns consisted of a black 60 × 47 pixel box flashed on a white screen for 50 ms. Participants were seated 80 cm in front of the monitor while stimuli were presented at a visual angle of 0.7162 radians. Visual pattern structure and IOIs were identical to those of both the auditory patterns, except there was no rise/fall time.

Participants were seated in a comfortable chair in a single-walled, sound-attenuated room and were asked to maintain fixation on a black cross on a white background in the center of a computer screen for the auditory conditions. No fixation cross was provided for the visual conditions. Participants were asked to listen to or watch the stimuli during EEG recording, and to avoid moving their eyes, head, or other body parts while the stimuli were presented. At the end of each pattern, participants indicated by pressing one of two buttons whether they perceived the pattern "slowing down" or "speeding up" at the end. Participants were allowed a 2-s intertrial interval to make their responses before a new trial began. Behavioral responses were made using an RB-830 button response box.

Participants were assigned to perform auditory blocks first or visual blocks first. For each participant, six blocks were presented. For those performing auditory blocks first, participants heard two blocks of auditory five-event patterns followed by two visual five-event blocks, an auditory four-event block, and a visual four-event block. Participants performing the visual blocks first saw two blocks of visual five-event patterns followed by two auditory five-event blocks, a visual four-event block, and an auditory four-event block. Each block contained 132 trials, with 22 trials of each of the six trial types ($\pm$ 4%, $\pm$ 12% or $\pm$ 20%) for that condition. Eight practice trials (using final IOIs of $\pm$ 40%) were presented prior to the start of the experiment. All aspects of stimulus presentation and behavioral response collection were controlled by a custom program written in Presentation.

EEG Recording

EEG signals were digitized continuously (512 Hz sampling rate and a 104 Hz bandwidth) using a BioSemi ActiveTwo system (http://biosemi.com). The EEG was recorded from an array of 72 electrodes, with an Ag-AgCl common mode sense (CMS) active electrode and an Ag-AgCl driven right leg (DRL) passive electrode serving as ground, placed at 64 points based on the 10/20 system in a BioSemi electrode cap and eight additional points below the hair line (both mastoids, both preauricular points, outer canthus of each eye, and inferior orbit of each eye), and recorded onto a PC desktop computer for offline analysis. Before EEG recording, conducting gel was applied to the skin at each electrode site with the cap on and sintered Ag-AgCl pin-type electrodes were fitted into place at each site in the cap. Sintered Ag-AgCl flat-type electrodes were attached with adhesive to sites below the hairline. Voltage offsets were adjusted to below 40 mV prior to recording, and the resting EEG was checked for any problematic electrodes prior to and throughout the recording session.

Data Analysis

Proportions of speeding up responses were calculated for each participant by averaging each participant's responses for each of the 24 trial types (4 Conditions × 6 Final IOIs) and dividing that by the total number of trials for that type. A 2 (Modality: auditory vs. visual) × 2 (Pattern Type: four-event vs. five-event) × 6 (Final IOI) repeated measures ANOVA was conducted on response proportions. Relative just noticeable differences (JNDs) and points of subjective equality (PSEs) were calculated for each participant using the *z*-transform method prescribed by MacMillan and Creelman (2005) for the four-event patterns to quantify individual differences in the ability to discriminate variations in the final IOI and compare discrimination thresholds across modalities. JNDs and PSEs were not estimated for the five-event patterns because previous research using this paradigm has shown that there is an inherent ambiguity in the perceived referent interval at the beginning of the pattern that precludes JND estimates for some individuals. Specifically, if participants judge the speed of rhythms using the explicit 300-ms referent marked by the first two intervals of the pattern, then they respond almost entirely slowing down to all final IOIs.

For the ERP data, repeated measures ANOVAs were performed for both the CNV and P3 amplitudes. Factors included whether the pattern had five or four events, the modality of the pattern, and whether the final event occurred earlier or later than the 600-ms referent. Collapsing across the three shortest final IOIs or the three longest final IOIs created this final time-change direction variable. This was done in order to have better signal-to-noise ratios for the ERPs. Because differences in relative JNDs (i.e., discrimination thresholds) for four-event patterns represents a possible confound in assessing brain responses as a function of modality, all reported ERP analyses were run both without and with the difference between auditory and visual thresholds as a covariate.

All off-line ERP processing except statistical analyses were performed using Brain Electrical Source Analysis software (BESA). Electrodes that were noted during the recording as being noisy throughout the experiment were automatically interpolated prior to analysis. Ocular artifacts (blinks, saccades, and smooth movements) were corrected automatically with a spatial-filtering method (Ille, Berg, & Scherg, 2002). Epochs contaminated by artifacts (amplitude exceeding 150 uV, gradient exceeding 75 uV, or signal below 0.10 uV) were automatically rejected before averaging. EEG

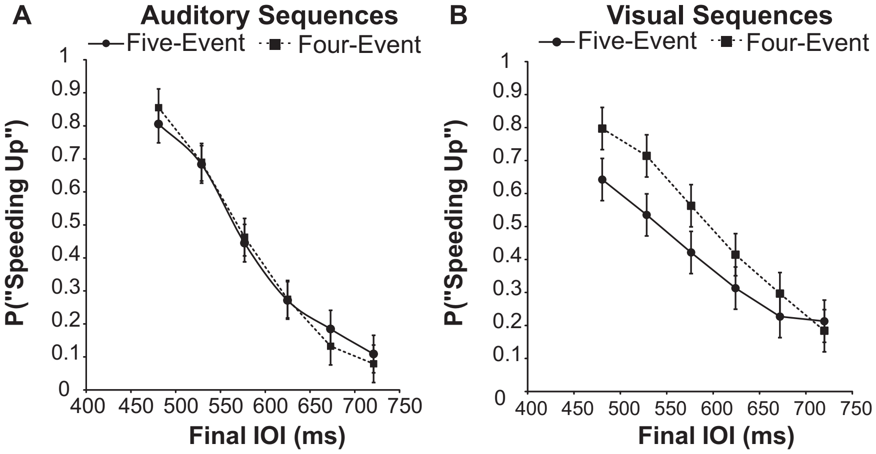

Figure 2. Proportion of speeding up responses by final IOI. Final IOIs reflect ± 4%, ± 12%, or ± 20% of the implied 600 ms IOI. A: Auditory patterns. B: Visual patterns.

epochs were averaged separately across all nonartifact trials for each of the 24 trial types and for each electrode site, and rereferenced to the average of all electrodes. We used this reference to enable adequate detection of auditory and visual ERPs at multiple scalp locations, including frontocentral, temporal, parietal, and occipital electrodes, and also for comparison with our previous study, which also referenced to the average (Snyder et al., 2011).

To examine ERPs related to processing the final two events of the pattern, epochs were segmented with time 0 at the onset of the last tone of the pattern, with a 1,226-ms pretrigger baseline period and a 1,000-ms posttrigger active period, and baseline corrected by subtracting the mean of the -26 to 0 ms portion of the baseline from each point in the epoch. These epochs were digitally band-pass filtered to attenuate frequencies below 0.01 Hz (6 dB/octave attenuation, forward) and above 40 Hz (24 dB/octave attenuation, symmetrical). To quantify the CNV (which occurred leading up to the final tone), those epochs were digitally band-pass filtered to attenuate frequencies below 0.01 Hz (6 dB/octave attenuation, forward) and above 40 Hz (24 dB/octave attenuation, symmetrical) and baseline corrected by subtracting the mean of the -752 to -726 ms portion of the baseline (time before penultimate event onset) from each point in the epoch.

ERP mean amplitudes were calculated in time ranges and electrodes based on known latencies and scalp locations of CNV and P3b components, with specific selections including maximal differences in the grand-averaged waveforms between conditions of interest at electrodes showing the maximal difference (i.e., from -100 to 0 ms for the CNV and from 350 to 630 ms for the P3, based on known latencies for these components). Mean amplitudes were averaged across the same electrode sites for each participant where ERP differences were most prominent in the grand average (at electrodes F1, Fz, F2, FC1, FCz, FC2, C1, Cz, and C2 for the CNV and at electrodes P1, Pz, P2, PO3, POz, and PO4 for the P3) and submitted to repeated measures ANOVAs with Greenhouse-Geisser corrections when appropriate.

Scalp topographies were obtained in BESA at single time points showing the maximal difference between conditions (-100 ms for the CNV and 400 ms for the P3). After range normalizing the mean amplitude data for both CNV and P3 based on the method described by McCarthy and Wood (1985), a repeated measures ANOVA was run to identify modality differences in activity patterns across all electrodes by testing for Electrode × Modality interactions. Pearson's correlations with these data were also calculated to quantify the degree of similarity in scalp topography across and within modalities. Correlation coefficients were Fisher z transformed prior to any inferential analysis based on these values.

Results

Since this study counterbalanced whether participants were exposed to the visual or auditory stimuli first, all analyses mentioned below were also initially run with modality order as a between-subjects factor; however, order was not a significant factor in either the behavioral or ERP analyses, so we report analyses without this factor for simplicity.

Behavioral Data

Figure 2 shows proportion of speeding up responses as a function of the final IOI for the auditory and visual patterns in the four-event and five-event conditions. A 2 (Modality) × 2 (Pattern Type) × 6 (Final IOI) repeated measures ANOVA was conducted on response proportions. Confirming that participants correctly performed the task, the ANOVA revealed a main effect of final IOI, $F(5,150) = 178.01$, $p < .001$, $\eta_p^2 = .856$, with participants responding that patterns sped up for final IOIs that were shorter than the implied 600-ms referent and responding that patterns slowed down for final IOIs that were longer than the implied 600-ms referent.

Consistent with previous research, there was also a main effect of modality, $F(1,30) = 8.68$, $p = .006$, $\eta_p^2 = .224$, and a significant interaction of Modality × Pattern Type, $F(1,30) = 11.68$, $p = .002$, $\eta_p^2 = .280$, suggesting a difference in salience of the 300-ms and 600-ms intervals between four- and five-event patterns for the visual patterns, but not auditory patterns. For the auditory patterns, participants responded almost identically to the four-event and five-event auditory patterns (Figure 2A), revealing a consistent use of a 600-ms referent for both pattern types. However, for the visual patterns, five-event patterns resulted in less salience for the 600-ms interval than did the four-event patterns (Figure 2B).

There were also significant interactions between modality and final IOI, $F(5,150) = 27.62$, $p < .001$, $\eta_p^2 = .479$ and pattern type and final IOI, $F(5,150) = 4.95$, $p = .003$, $\eta_p^2 = .142$, indicative of differences in the slope of the psychometric curves. In general, shallower slopes (worse temporal discrimination) were observed with visual patterns compared to auditory patterns. Consistent with this interpretation, relative JNDs in the visual four-event condition

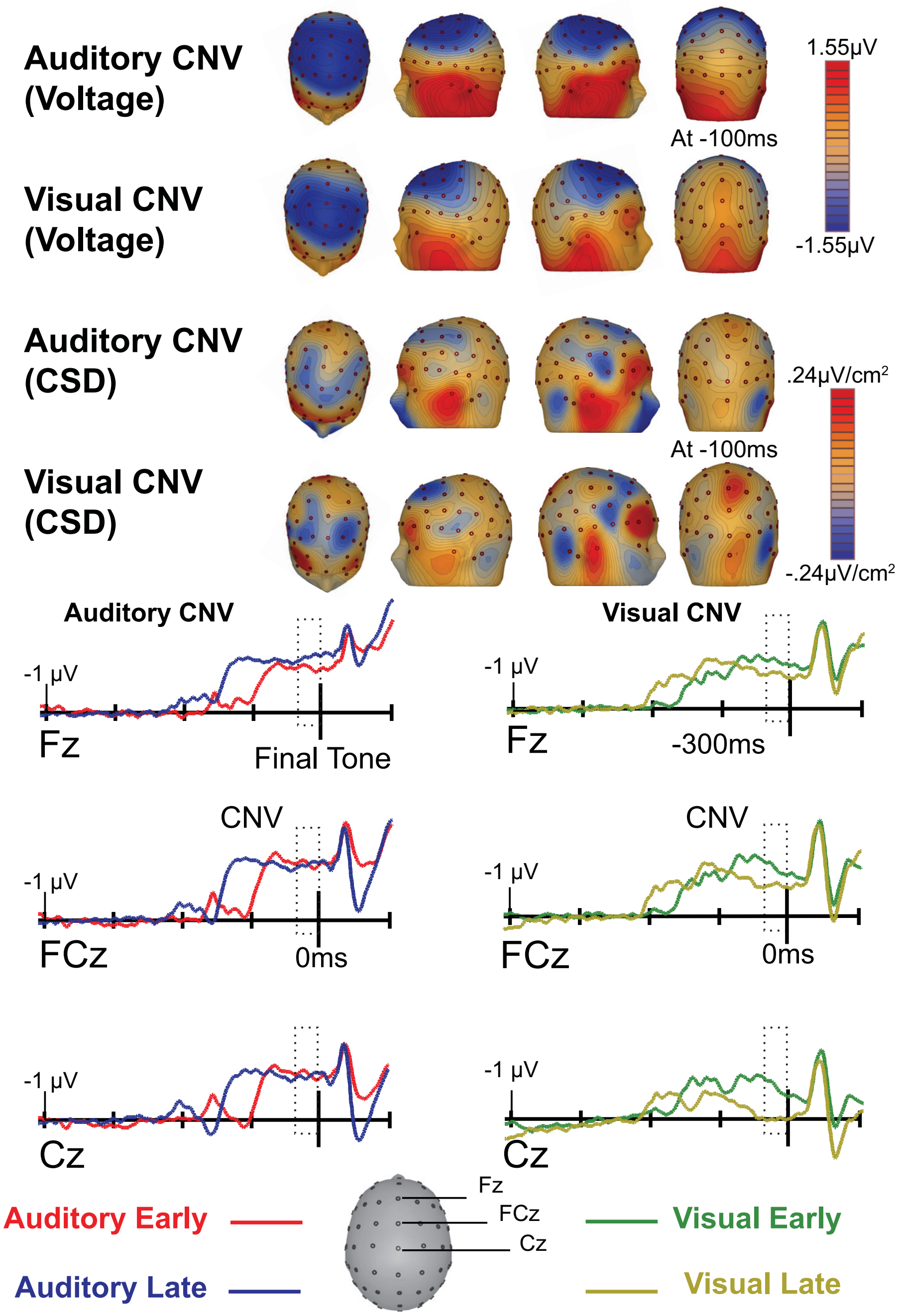

Figure 3. ERP traces and topographies for auditory and visual CNVs at frontocentral electrodes. ERPs have been collapsed across four- and five-event conditions, while topographies have been collapsed across four and five events, as well as early and late conditions.

($M = 30.1\%$, $SD = 37.0\%$) were larger than in the auditory four-event condition ($M = 11.6\%$, $SD = 6.6\%$), $t(30) = 2.89$, $p = .007$. PSEs were not significantly different between the visual four-event condition ($M = 622$ ms, $SD = 254$ ms) and the auditory four-event condition ($M = 579$ ms, $SD = 42$ ms), $t(30) = 1.07$, $p = .295$, although there was a trend for visual PSEs to be larger.

EEG Data

Because the modality differences reported above in relative JNDs for four-event patterns represent a possible confound in assessing brain responses as a function of modality, the ERP analyses were run both without and with the difference between auditory and visual JNDs as a covariate. A 2 (Modality) × 2 (Pattern Type) × 2 (Time-Change Direction: early, late) repeated measures ANOVA was conducted on both the CNV and P3 mean amplitudes. For CNV (averaged across electrodes F1, Fz, F2, FC1, FCz, FC2, C1, Cz, and C2), there was a main effect of modality, $F(1,30) = 9.13$, $p = .005$, $\eta_p^2 = .239$, with larger amplitudes in the auditory modality than in the visual modality (see Figure 3). Macar and Vidal (2004) have proposed that increases in CNV amplitude might reflect more precise temporal estimates. Our behavioral results in which we observed lower JNDs in the auditory modality than in the visual modality are consistent with this view and our own CNV findings.

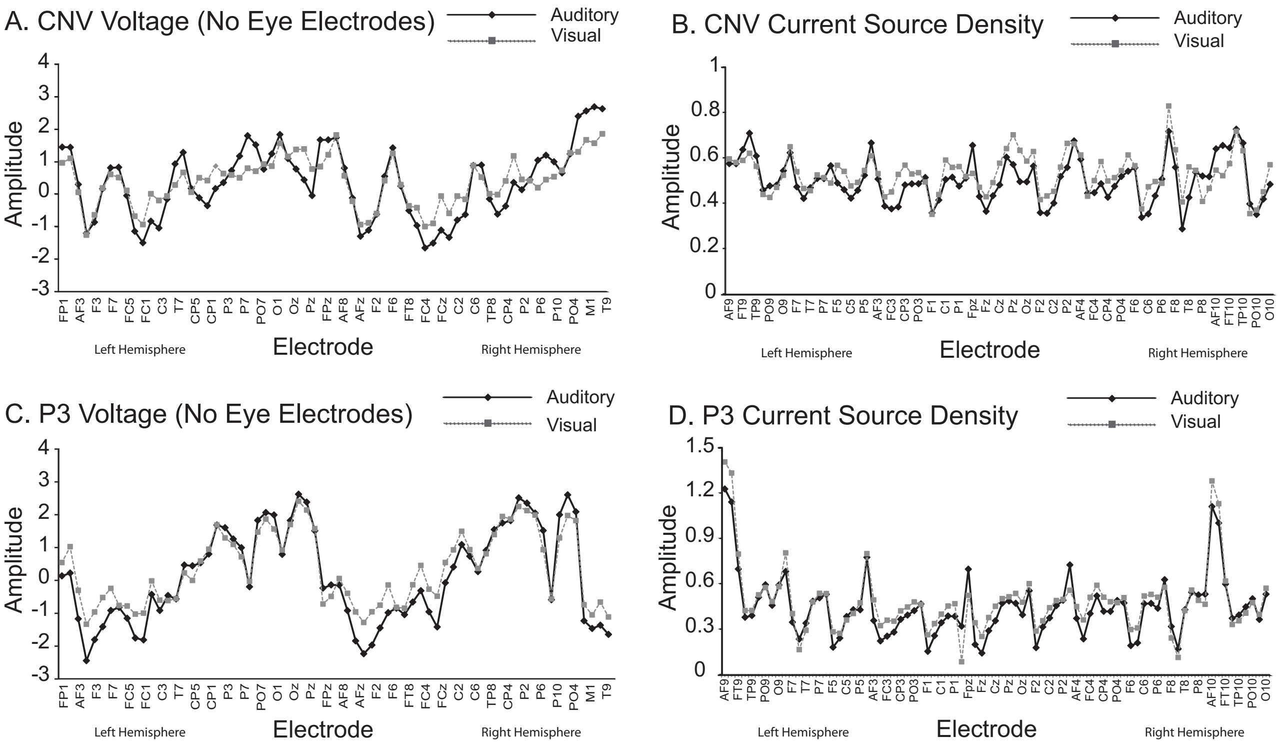

Figure 4. A: Mean amplitude of the CNV across the 68 electrodes (64 scalp electrodes plus four nonocular face electrodes). B: Mean amplitude of the P3 across the 68 electrodes (64 scalp electrodes plus four nonocular face electrodes).

There was no main effect of time-change direction, $F(1,30) = 3.57$, $p = .069$, $\eta_p^2 = .110$, but there was a significant interaction between modality and time-change direction, $F(1,30) = 10.42$, $p = .003$, $\eta_p^2 = .264$. Importantly, effects of time-change direction are due to the differences in how much time has elapsed since the penultimate event rather than sensitivity to the timing of the final event, which is unpredictable by design. Differences in CNV amplitude in the early and late condition were larger in the visual modality than in the auditory modality (see Figure 3), with larger CNV amplitudes when the final event was early (i.e., final IOI < 600 ms) compared to when it was late (i.e., final IOI > 600 ms). This may be due to a difference of perceptual strategy; if some participants were using a 300-ms referent in the auditory condition, all of the events marking the end of the final interval occurred "too late." The CNV for these individuals may be decreasing earlier than those using a different strategy.

When observing the voltage pattern across the scalp, both auditory and visual CNVs showed negative voltage at frontocentral regions and reverses in polarity beneath the superior temporal scalp regions. To determine whether there were modality differences in the pattern of scalp activity independent of amplitude differences, despite the apparent similarity shown in Figure 4A, the data from all 72 electrodes were range normalized. A repeated measures ANOVA was run to identify differences across the scalp by testing for a Modality × Electrode interaction on the range-normalized data. There was a significant interaction between modality and electrode, $F(1,30) = 5.15$, $p < .001$, $\eta_p^2 = .147$, which remained significant when the eye electrodes were removed (to exclude a possible confound of ocular artifacts), $F(1,30) = 5.59$, $p < .001$ $\eta_p^2 = .157$ (see Figure 4A). This remained significant when the difference between auditory and visual thresholds was used as a covariate for analyses both with and without eye electrodes, $F(1,30) = 3.22$, $p = .001$, $\eta_p^2 = .100$, and $F(1,30) = 3.51$, $p < .001$, $\eta_p^2 = .108$, respectively.

To determine whether there is similarity in the scalp patterns across modalities, a correlation between all 72 electrodes in the auditory and visual modality was run for each individual participant. A single-sample *t* test was then run to test whether the mean of all the participants' correlations was different than 0 (after Fisher *z* transformation). For the CNV, the mean of the participants' correlations was different from 0, suggesting that there were similarities in the CNV across modalities (see Table 1). These results suggest that there are modality-general contributions that drive the overall correlations between patterns.

Table 1. *Statistical Comparisons for CNV and P3 Components with Regard to Voltage and CSD Maps*

	CNV			
	Mean	*SD*	*t*	*p*
Voltage	.613	.223	11.490	< .001
CSD	.505	.085	26.262	< .001
	P3			
	Mean	*SD*	*t*	*p*
Voltage	.736	.237	11.839	< .001
CSD	.442	.089	23.075	< .001

Note. Correlations between auditory and visual modalities are statistically different from 0, after Fisher *z* transformation of *r* values. CNV = contingent negative variation; CSD = current source density.

Table 2. *Comparisons of Within- and Across-Modality Pairings of Correlations for CNV Voltage and CSD Maps*

	CNV			
	Voltage *M*(*SD*)	*t*	CSD *M*(*SD*)	*t*
Within auditory	.715 (.192)	4.804**	.635 (.035)	4.976**
Across modality (early)	.540 (.047)		.437 (.046)	
Within visual	.648 (.040)	4.374**	.629 (.042)	4.408**
Across modality (late)	.510 (.046)		.475 (.044)	
Within auditory	.715 (.192)	5.737**	.635 (.035)	4.900**
Across modality (late)	.510 (.046)		.475 (.044)	
Within visual	.648 (.040)	2.988*	.629 (.042)	4.467**
Across modality (early)	.540 (.047)		.437 (.046)	

Note. Within auditory and within visual refer to conditions including both auditory early and late and both visual early and late, respectively. Across-modality conditions include both modalities for that specific time change interval. Within-modality correlations are significantly different from across-modality correlations, after Fisher *z* transformation of *r* values. CNV = contingent negative variation; CSD = current source density.
$^{**}p < .001$. $^{*}p < .05$ (after Bonferonni correction).

One way to find evidence for modality-specific contributions in the current data is to directly correlate scalp topographies from different types of time change of the same modality (e.g., $-\Delta$T and late from the auditory modality) and compare the result to correlations between the same conditions but across modality (e.g., early from the auditory and visual modalities). If there is no difference between the resulting correlations, this would suggest that the activity in all the conditions being compared (regardless of whether within- or across-modality comparisons are being made) are comparable, and therefore there are no modality-specific contributions. However, if the within-modality correlations are larger, despite the fact that different types of time change are being compared, this would suggest that there are modality-specific contributions.

In order to determine if there were differences in the correlations across modalities or within modalities, the individual participants' transformed correlations (described above) were averaged based on condition, and paired sample *t* tests were run for four pairings of conditions. An ANOVA was not possible here because of the nonfactorial nature of the comparisons (i.e., the within-modality correlations were used twice, once to compare with early/across-modality and once to compare with late/across-modality). Since there were two conditions for modality (auditory and visual) and two conditions for time change direction (early and late), four pairs were needed to contrast within-modality (across time change) and across-modality (within time change) topographies. For all four of these pairings, there was a significant difference between the within-modality correlations and the across-modality conditions, with larger correlations within modalities (see Table 2).

In order to provide converging evidence for the above results, the CSD was also calculated, in addition to voltage. The CSD mean amplitudes were also range normalized. A repeated measures ANOVA was run to identify differences across the CSD maps by testing for a Modality × Electrode interaction. There was a significant interaction between modality and electrode, $F(1,30) = 2.03$, $p = .032$, $\eta_p^2 = .063$. This did not remain significant when the difference between auditory and visual thresholds was used as a covariate, $F(1,30) = 1.62$, $p = .106$, $\eta_p^2 = .053$. The CSD maps are shown with their voltage counterparts in Figure 3.

As with the CNV voltages, the correlation between all electrodes in the auditory and visual modality was run for each individual participant. A single-sample *t* test was then run to test whether the mean of all the participants' correlations was different than 0. For the CSD of the CNV, the mean of the participants' correlations was statistically different from 0, suggesting that the scalp distributions of the CNV were similar across modalities (see Table 1). As with the correlations for the voltages, the individual participant correlations across the CSD electrodes were averaged based on condition, and paired-sample *t* tests with Bonferonni corrections were run for the same four pairings of conditions. For each of the four pairings, there was a significant difference between the within-modality correlations and the across-modality conditions, with larger correlations within modalities (see Table 2), consistent with modality-specific contributions.

Whereas the CNV appears to be at least partially influenced by modality-specific contributions, the later P3 might be expected to be modality general. After the final event, participants must decide if the pattern speeds up or slows down, and this requirement to make a decision results in late positive components (Polich, 2007; Sutton, Braren, Zubin, & John, 1965). A distinct P3 occurred at the parietal electrodes (maximally at POz) for both auditory and visual patterns after the final event. As with the CNV, a repeated measures ANOVA was performed with time-change direction, pattern type (four-event vs. five-event), and the modality of the pattern as factors, but with the mean amplitude measured at parietal electrodes (P1, Pz, P2, PO3, POz, and PO4). As shown in Figure 5, there was a main effect of time-change direction, $F(1,30) = 9.34$, $p = .005$, $\eta_p^2 = .244$, with larger P3 amplitudes when the final event was early compared to when it was late, potentially suggesting a novelty enhancement for events that occur before the expected time (i.e., when the final IOI is less than 600 ms). Alternatively, this could reflect a decrease in demand for memory and cognitive resources for events that occur after the expected time (i.e., when the final IOI is greater than 600 ms), due to a reduced uncertainty for when the final event will occur.

There was no main effect of modality, $F(1,30) = 1.69$, $p = .204$, $\eta_p^2 = .055$ (see Figure 5). As with CNV, data from all 72 electrodes were range normalized, and a repeated measures ANOVA was run to test for a Modality × Electrode interaction. There was a significant interaction between modality and electrode, $F(1,30) = 3.13$, $p = .013$, $\eta_p^2 = .094$, which remained significant when the eye electrodes were removed, $F(1,30) = 2.91$, $p = .014$, $\eta_p^2 = .088$ (see Figure 4B), but this interaction did not remain significant when the difference between auditory and visual thresholds was used as a covariate for analyses both with and without eye electrodes, $F(1,30) = 2.11$, $p = .074$, $\eta_p^2 = .068$ and $F(1,30) = 2.14$, $p = .062$, $\eta_p^2 = .069$, respectively. Consistent with these findings, as shown in Figure 4B, the differences in scalp topography across modalities were not great. As with the CNV, the correlation between all 72 electrodes in the auditory and visual modality was run for each individual participant. A single-sample *t* test was then run to test whether the mean of all the participants' correlations was different than 0. For the P3, the mean of the participants' correlations was statistically different from 0, suggesting that there were similarities in the scalp distributions of the P3 across modalities (see Table 1).

As discussed previously, while these results suggest that there are modality-general contributions that drive the overall correlations between patterns, they do not rule out the possibility that there are modality-specific contributions to scalp topographies. As with the CNV, the individual participant correlations were averaged

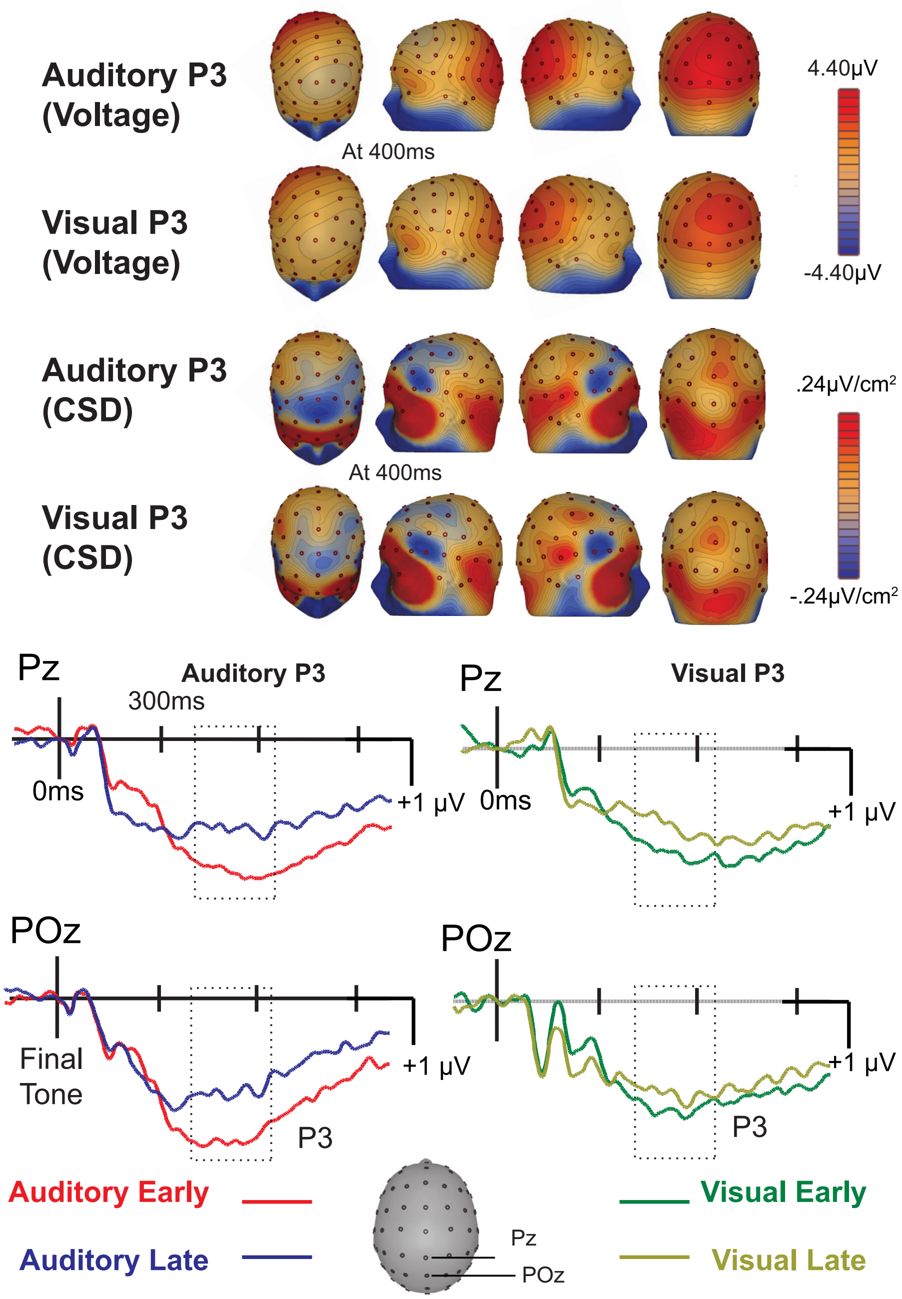

Figure 5. ERP traces and topographies for auditory and visual P3s. ERPs have been collapsed as with the CNV.

based on condition, and paired-sample *t* tests with Bonferonni corrections were run for four pairings of conditions (after Fisher *z* transformation). For three out of four of these pairings, there was a significant difference between the within-modality correlations and the across-modality conditions, with larger correlations within modalities (see Table 3).

The mean amplitudes of CSD waveforms were also calculated and then range normalized. A repeated measures ANOVA was run to identify differences across the CSD maps by testing for a Modality × Electrode interaction. There was only a marginal interaction between modality and electrode, $F(1,30) = 2.29$, $p = .085$, $\eta_p^2 = .071$. This was also not significant when the difference between auditory and visual thresholds was used as a covariate, $F(1,30) = 1.91$, $p = .136$, $\eta_p^2 = .062$. The CSD maps are shown with their voltage counterparts in Figure 5. As with the P3 voltages, the correlation between all electrodes in the auditory and visual modality was run for each individual participant. A single-sample *t* test was then run to test whether the mean of all the participants' correlations was different than 0. For the CSD of the P3, the mean of the participants' correlations was statistically different from 0, suggesting that the scalp distributions of the P3 were similar across modalities (see Table 1). As with the correlations for the voltages, the individual participant correlations across the CSD electrodes were averaged based on condition, and paired-sample *t* tests were run for the same four pairings of conditions. For all four of these pairings, there was a significant difference between the within-modality correlations and the across-modality conditions, with larger correlations within modalities (see Table 3).

Table 3. *Comparisons of Within- and Across-Modality Pairings of Correlations for P3 Voltage and CSD Maps*

	P3			
	Voltage *M*(*SD*)	*t*	CSD *M*(*SD*)	*t*
Within auditory	.817 (.037)	3.888*	.766 (.020)	4.279**
Across modality (early)	.662 (.048)		.607 (.041)	
Within visual	.801 (.026)	2.342	.756 (.031)	3.510*
Across modality (late)	.721 (.037)		.633 (.033)	
Within auditory	.817 (.037)	3.755*	.766 (.020)	4.711**
Across modality (late)	.721 (.037)		.633 (.033)	
Within visual	.801 (.026)	3.066*	.756 (.031)	4.144*
Across modality (early)	.662 (.048)		.607 (.041)	

Note. Within auditory and within visual refer to conditions including both auditory early and late and both visual early and late, respectively. Across-modality conditions include both modalities for that specific time change interval. Within-modality correlations are significantly different from across-modality correlations, after Fisher *z* transformation of *r* values. CSD = current source density.
$^{**}p < .001$. $^{*}p < .05$ (after Bonferonni correction).

Discussion

The current study used ERPs to help determine whether temporal processing is primarily modality specific or modality general. The results for the CNV, a negative component shown to reflect the temporal expectancy for the final event of a pattern (Pfeuty, Ragot, & Pouthas, 2003; Walter et al., 1964), suggest that it is to a large extent modality general but is at least partially modality specific. This took the form of an Electrode × Modality interaction and larger within-modality correlations than across-modality correlations. We also found larger CNV amplitudes in the auditory modality than in the visual modality, suggesting that the brain is more responsive during the encoding of auditory timing information compared to during encoding of visual timing. Visual CNVs with smaller amplitudes also seemed to decay more quickly than their auditory counterparts, suggesting that the visual modality may be less well equipped to form temporal expectancies. The P3 response showed a similar pattern of topographical results, suggesting that it too at least partially indexes modality-specific processing, although there was no difference in amplitude between auditory or visual modalities and the pattern of correlations showed less modality specificity than with the CNV.

Combined with the behavioral results, which showed better temporal discrimination in the auditory modality than in the visual modality, the present findings are consistent with the view that the auditory system is superior to the visual system in terms of temporal processing (Grondin, Meilleur-Wells, Ouellette, & Macar, 1998; Patel et al., 2005; Repp, 2003; Repp & Penel, 2002). The modality specificity of the CNV may also be seen in the observed interaction between time-change direction (whether the final event was early or late) and modality. The CNV amplitudes were closer in size for the auditory modality, but showed a larger voltage difference in the visual modality as a function of time-change direction.

Despite the number of conflicting findings in previous research as to the modality specificity of timing (Ivry & Schlerf, 2008), the topographical differences and similarities shown by the CNV support a small degree of modality-specific timing process. N'Daiye and colleagues (2004) looked at CNVs using EEG and magnetoencephalography in response to long-duration (490–910 ms) auditory and visual filled intervals (rather than empty intervals as in the current study) and found a distinct CNV for each modality. Specifically, they found a difference in the scalp distribution of sustained activity between visual and auditory conditions, but ultimately concluded that the CNV was modality general. The authors stated that, while sensory-driven activity was present concurrently with the sustained CNV, it was difficult to separate their individual contributions.

By using brief sensory events in the present study that defined empty time intervals, we were able to remove sustained sensory responses from the epochs we used to examine the CNV, thus providing clearer evidence that the CNV, a reflection of temporal expectancy, is in part modality specific. The modality-general contributions could arise from frontal sources: both the supplementary motor area (SMA) and pre-SMA appear to be active during auditory and visual duration estimation tasks and tasks that elicit the CNV (Lewis, Wing, Pope, Praamstra, & Miall, 2004; Macar et al., 2002; Macar, Vidal, & Casini, 1999; Pouthas et al., 2005; Schwartze, Rothermich, & Kotz, 2012). There is evidence that the later portion of the CNV reflects decision making (Rockstroh, Elbert, Canavan, Lutzenberger, & Birbaumer, 1989); this cognitive activation may be overlapping with modality-specific activation.

The temporal precision of ERPs allowed us to identify that the CNV reflects a combination of modality-specific and modality-general timing-related processing. Previous studies testing individuals with brain lesions (Harrington & Haaland, 1999) or studies using fMRI (Bueti, Bahrami, & Walsh, 2008; Bueti, Walsh et al., 2008; Grahn & McAuley, 2009; Rao et al., 2001) and transcranial magnetic stimulation (Bueti, van Dongen, & Walsh, 2008) in healthy individuals have come to differing conclusions as to the contributions of sensory and nonsensory brain areas to time estimation and perception. For example, studies of people with damage to the cerebellum or basal ganglia with time perception and production deficits have been used to argue for the general role of these structures in timing (e.g., Harrington, Haaland, & Hermanowicz, 1998; Ivry & Keele, 1989). In contrast, more recent studies using neuroimaging and brain stimulation have been used to argue that sensory association cortex areas, in particular, seem to be involved in timing for only one modality (e.g., Bueti, van Dongen, & Walsh, 2008; Bueti, Walsh et al., 2008). One possible reason these discrepant findings have arisen is that these studies may have differed in their ability to specifically pinpoint timing processes as opposed to more general cognitive processes (decision making, memory) due to the poor temporal resolution of their methods.

Studies that examined the transfer of temporal learning have also come to differing conclusions. This is possibly because, like neuroimaging and brain stimulation studies, the inability to discern when particular brain areas are active during processing makes it difficult to identify how those brain areas contribute to the different aspects/stages of timing (i.e., event encoding, duration calculation, memory storage, comparison). In some studies, learning generalized to the same interval duration defined by unlearned auditory frequencies or tactile locations (Karmarkar & Buonomano, 2003; Nagarajan, Blake, Wright, Byl, & Merzenich, 1998; Wright et al., 1997) or the same interval presented in a different modality (Nagarajan et al., 1998). These findings were consistent with the existence of modality-general mechanisms. In contrast, more recent findings have failed to show transfer of temporal learning across modalities (Lapid et al., 2009), which instead suggest modality-specific mechanisms.

Other recent research using the same time discrimination paradigm as used in the current study with auditory and visual rhythms is consistent with these more recent findings from transfer of learning

(Grahn et al., 2011; McAuley & Henry, 2010). In particular, both of these studies showed robust order effects such that, when blocks of auditory trials were presented before blocks of visual trials, the visual rhythms were more likely to be encoded using beat-based mechanisms that are more associated with auditory than visual rhythm perception, at least when the beat period was 600 ms in duration (also see Grahn, 2012; Patel et al., 2005). The fact that the encoding of auditory and visual rhythms differs from each other and can be influenced by each other is consistent with the idea of modality-specific timing. For example, differences in performance between modalities could reflect the operation of separate timing mechanisms that have different levels of precision, while cross-modal influence suggests that the output of one modality's timing mechanism can override or influence the mechanism of the other modality.

Our CNV findings suggest that temporal processing per se is encoded in part in a modality-specific manner, consistent with the framework promoted by Buonomano and colleagues (Buonomano, 2000; Buonomano, Bramen, & Khodadadifar, 2009; Karmarkar & Buonomano, 2007). These authors have argued that the encoding of time intervals is accomplished using general mechanisms of plasticity and cortical dynamics that are present in sensory areas of the cortex. However, dedicated modality-general mechanisms may also be necessary for additional memory, comparison, and decision-making processes that are commonly required for timing tasks, especially with longer intervals (Buonomano et al., 2009; Ivry & Schlerf, 2008; Meck, Penney, & Pouthas, 2008). One or more of these modality-general processes may be indexed by the CNV and P3 responses observed in the current study.

The P3 may reflect a somewhat more modality-general process than the CNV (Linden, 2005; Polich, 2007). Unlike the CNV, there were no main effects of modality for P3. The P3 measured in the current study is consistent with previous research on rhythm perception (Besson, Faita, Czternasty, & Kutas, 1997; Jongsma, Desain, & Honing, 2004; Jongsma, Meeuwissen, Vos, & Maes, 2007), appearing after the onset of the final event in time discrimination tasks, indicating categorization, updating of the memory trace of the encoded time interval, or conscious perception of a temporal change (Linden, 2005; Polich, 2007; Snyder et al., 2015). There was a significant effect of whether the final event occurred earlier or later than expected. The P3 amplitudes were larger when the final event occurred early, which might reflect a decreased processing load: more easily discriminated targets result in larger P3 amplitudes, as resources are not being taxed by working memory. Polich's (2007) context-updating theory of the P3b states that, as the strain on memory resources increases, the P3 amplitude decreases. Participants might have been holding the initial time interval in memory, and therefore final tones that occurred too late only had a small pool of resources left, resulting in the P3 amplitude decreasing. There was little difference between P3 amplitude to events occurring after the expected time across modalities. However, the auditory P3 amplitudes were larger for auditory events occurring before the expected time. Because both auditory and visual P3s index high-level cognitive processing, both suffer from decreased cognitive resources when events occur later than expected. However, because the exact mechanisms of the P3 are still unknown, further research should focus on the exact contributions of memory, decision making, and change detection to this modality-general timing component.

In summary, we measured behavioral performance and ERP components during a well-studied rhythm discrimination paradigm using auditory and visual rhythms. Overall, discrimination thresholds were lower in the auditory modality than in the visual modality. Correspondingly, the CNV component, which has been strongly linked to temporal encoding and anticipation, showed greater amplitude in the auditory modality than in the visual modality, whereas the P3 component did not show this type of modality difference. Most importantly, scalp voltage patterns for the CNV and the P3 indicated modality-general timing in the brain.

References

Besson, M., Faita, F., Czternasty, C., & Kutas, M. (1997). What's in a pause: Event-related potential analysis of temporal disruptions in written and spoken sentences. *Biological Psychology*, *46*, 3–23. doi: 10.1016/S0301-0511(96)05215-5

Bueti, D., Bahrami, B., & Walsh, V. (2008). Sensory and association cortex in time perception. *Journal of Cognitive Neurosciece*, *20*, 1054–1062. doi: 10.1162/jocn.2008.20060

Bueti, D., van Dongen, E. V., & Walsh, V. (2008). The role of superior temporal cortex in auditory timing. *PLoS ONE*, *3*, e2481. doi: 10.1371/journal.pone.0002481

Bueti, D., Walsh, V., Frith, C., & Rees, G. (2008). Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neurosciece*, *20*, 204–214. doi: 10.1162/jocn.2008.20017

Buonomano, D. V. (2000). Decoding temporal information: A model based on short-term synaptic plasticity. *Journal of Neuroscience*, *20*, 1129–1141.

Buonomano, D. V., Bramen, J., & Khodadadifar, M. (2009). Influence of the interstimulus interval on temporal processing and learning: Testing the state-dependent network model. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 1865–1873. doi: 10.1098/rstb.2009.0019

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

Gontier, E., Hasuo, E., Mitsudo, T., & Grondin, S. (2013). EEG investigations of duration discrimination: The intermodal effect is induced by an attentional bias. *PLoS ONE*, *8*, e74073. doi: 10.1371/journal.pone.0074073

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, *220*, 51–61. doi: 10.1007/s00221-012-3114-8

Grahn, J. A., Henry, M. J., & McAuley, J. D. (2011). fMRI investigation of cross-modal interactions in beat perception: Audition primes vision, but not vice versa. *NeuroImage*, *54*, 1231–1243. doi: 10.1016/j.neuroimage.2010.09.033

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage*, *47*, 1894–1903. doi: 10.1016/j.neuroimage.2009.04.039

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin*, *127*, 22–44. doi: 10.1037/0033-2909.127.1.22

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. *Psychological Research*, *61*, 261–268. doi: 10.1007/s004260050030

Harrington, D. L., & Haaland, K. Y. (1999). Neural underpinnings of temporal processing: A review of focal lesion, pharmacological, and functional imaging research. *Reviews in the Neurosciences*, *10*, 91–116. doi: 10.1515/REVNEURO.1999.10.2.91

Harrington, D. L., Haaland, K. Y., & Hermanowicz, N. (1998). Temporal processing in the basal ganglia. *Neuropsychology*, *12*, 3–12. doi: 10.1037/0894-4105.12.1.3

Hasuo, E., Gontier, E., Mitsudo, T., Nakajima, Y., Tobimatsu, S., & Grondin, S. (2014). An electrophysiological study on intra- and intermodal duration discrimination: Effects of performance level. *Proceedings of the Annual Conference of the Cognitive Science Society*, *36*, 590–594.

Ille, N., Berg, P., & Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology*, *19*, 113–124.

Ivry, R. B. (1996). The representation of temporal information in perception and motor control. *Current Opinion in Neurobiology*, *6*, 851–857. doi: 10.1016/S0959-4388(96)80037-7

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, *1*, 136–152. doi: 10.1162/jocn.1989.1.2.136

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences*, *12*, 273–280. doi: 10.1016/j.tics.2008.04.002

Jongsma, M. L., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and non-musicians. *Biological Psychology*, *66*, 129–152. doi: 10.1016/j.biopsycho.2003.10.002

Jongsma, M. L., Meeuwissen, E., Vos, P. G., & Maes, R. (2007). Rhythm perception: Speeding up or slowing down affects different subcomponents of the ERP P3 complex. *Biological Psychology*, *75*, 219–228. doi: 10.1016/j.biopsycho.2007.02.003

Karmarkar, U. R., & Buonomano, D. V. (2003). Temporal specificity of perceptual learning in an auditory discrimination task. *Learning and Memory*, *10*, 141–147. doi: 10.1101/lm.55503

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron*, *53*, 427–438. doi: 10.1016/j.neuron.2007.01.006

Lapid, E., Ulrich, R., & Rammsayer, T. (2009). Perceptual learning in auditory temporal discrimination: No evidence for a cross-modal transfer to the visual modality. *Psychonomic Bulletin & Review*, *16*, 382–389. doi: 10.3758/PBR.16.2.382

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*, 119–159. doi: 10.1037/0033-295X.106.1.119

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, *1169*, 46–57. doi: 10.1111/j.1749-6632.2009.04550.x

Lewis, P. A., Wing, A. M., Pope, P. A., Praamstra, P., & Miall, R. C. (2004). Brain activity correlates differentially with increasing temporal complexity of rhythms during initialisation, synchronisation, and continuation phases of paced finger tapping. *Neuropsychologia*, *42*), 1301–1312. doi: 10.1016/j.neuropsychologia.2004.03.001

Linden, D. E. (2005). The P300: Where in the brain is it produced and what does it tell us? *Neuroscientist*, *11*, 563–576. doi: 10.1177/1073858405280524

Lustig, C., & Meck, W. H. (2011). Modality differences in timing and temporal memory throughout the lifespan. *Brain and Cognition*, *77*, 298–303. doi: 10.1016/j.bandc.2011.07.007

Macar, F., Lejeune, H., Bonnet, M., Ferrara, A., Pouthas, V., Vidal, F., & Maquet, P. (2002). Activation of the supplementary motor area and of attentional networks during temporal processing. *Experimental Brain Research*, *142*, 475–485. doi: 10.1007/s00221-001-0953-0

Macar, F., & Vidal, F. (2003). The CNV peak: An index of decision making and temporal memory. *Psychophysiology*, *40*, 950–954. doi: 10.1111/1469-8986.00113

Macar, F., & Vidal, F. (2004). Event-related potentials as indices of time processing: A review. *Journal of Psychophysiology*, *18*, 89–104. doi: 10.1027/0269-8803.18.2-3.89

Macar, F., Vidal, F., & Casini, L. (1999). The supplementary motor area in motor and sensory timing: Evidence from slow brain potential changes. *Experimental Brain Research*, *125*, 271–280. doi: 10.1007/s002210050683

MacMillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

McAuley, J. D. (1995). *Perception of time as phase: Toward an adaptive-oscillator model of rhythmic pattern processing*. Bloomington, IN: Indiana University Press.

McAuley, J. D., & Henry, M. J. (2010). Modality effects in rhythm processing: Auditory encoding of visual rhythms is neither obligatory nor automatic. *Attention, Perception, & Psychophysics*, *72*, 1377–1389. doi: 10.3758/APP.72.5.1377

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, *62*, 203–208. doi: 10.1016/0168-5597(85)90015-2

Meck, W. H., Penney, T. B., & Pouthas, V. (2008). Cortico-striatal representation of time in animals and humans. *Current Opinion in Neurobiology*, *18*, 145–152. doi: 10.1016/j.conb.2008.08.002

N'Diaye, K., Ragot, R., Garnero, L., & Pouthas, V. (2004). What is common to brain activity evoked by the perception of visual and auditory filled durations? A study with MEG and EEG co-recordings. *Cognitive Brain Research*, *21*, 250–268. doi: 10.1016/j.cogbrainres.2004.04.006

Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N., & Merzenich, M. M. (1998). Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. *Journal of Neuroscience*, *18*, 1559–1570.

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: The neurophysics of EEG* (2nd ed.). New York, NY: Oxford University Press.

Patel, A. D., Iversen, J. R., Chen, Y. Q., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*, 226–238. doi: 10.1007/s00221-004-2159-8

Penhune, V. B., Zatorre, R. J., & Evans, A. C. (1998). Cerebellar contributions to motor timing: A PET study of auditory and visual rhythm reproduction. *Journal of Cognitive Neuroscience*, *10*, 752–765. doi: 10.1162/089892998563149

Pfeuty, M., Ragot, R., & Pouthas, V. (2003). Processes involved in tempo perception: A CNV analysis. *Psychophysiology*, *40*, 69–76. doi: 10.1111/1469-8986.00008

Pfeuty, M., Ragot, R., & Pouthas, V. (2008). Brain activity during interval timing depends on sensory structure. *Brain Research*, *1204*, 112–117. doi: 10.1016/j.brainres.2008.01.022

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148. doi: 10.1016/j.clinph.2007.04.019

Pouthas, V., George, N., Poline, J. B., Pfeuty, M., VandeMoorteele, P. F., Hugueville, L., . . . Renault, B. (2005). Neural network involved in time perception: An fMRI study comparing long and short interval estimation. *Human Brain Mapping*, *25*, 433–441. doi: 10.1002/hbm.20126

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*, 411–440. doi: 10.2307/40285311

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience*, *4*, 317–323. doi: 10.1038/85191

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and 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 and Performance*, *28*, 1085–1099. doi: 10.1037/0096-1523.28.5.1085

Rockstroh, B., Elbert, T., Canavan, A., Lutzenberger, W., & Birbaumer, N. (1989). *Slow cortical potentials and behaviour II*. Munich, Germany: Urban and Schwarzenberg.

Schwartze, M., Rothermich, K., & Kotz, S. A. (2012). Functional dissociation of pre-SMA and SMA-proper in temporal processing. *NeuroImage*, *60*, 290–298. doi: 10.1016/j.neuroimage.2011.11.089

Snyder, J. S., Pasinski, A. C., & McAuley, J. D. (2011). Listening strategy for auditory rhythms modulates neural correlates of expectancy and cognitive processing. *Psychophysiology*, *48*, 198–207. doi: 10.1111/j.1469-8986.2010.01053.x

Snyder, J. S., Yerkes, B. D., & Pitts, M. A. (2015). Testing domain-general theories of perceptual awareness with auditory brain responses. *Trends in Cognitive Sciences*, *19*, 295–297. doi: 10.1016/j.tics.2015.04.002

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*, 1187–1188. doi: 10.1126/science.150.3700.1187

Tenke, C. E., & Kayser, J. (2005). Reference-free quantification of EEG spectra: Combining current source density (CSD) and frequency principal components analysis (fPCA). *Clinical Neurophysiology*, *116*, 2826–2846. doi: 10.1016/j.clinph.2005.08.007

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature*, *203*, 380–384. doi: 10.1038/203380a0

Wright, B. A., Buonomano, D. V., Mahncke, H. W., & Merzenich, M. M. (1997). Learning and generalization of auditory temporal-interval discrimination in humans. *Journal of Neuroscience*, *17*, 3956–3963.

(RECEIVED October 23, 2014; ACCEPTED September 15, 2015)