

Running head: JOINT ACTION OUTCOMES

Monitoring individual and joint action outcomes in duet music performance

Janeen D. Loehr<sup>1</sup>, Dimitrios Kourtis<sup>2</sup>, Cordula Vesper<sup>3</sup>, Natalie Sebanz<sup>3,4</sup>, and Günther  
Knoblich<sup>4</sup>

*Journal of Cognitive Neuroscience*, 25, 1049-1061. © 2013 Massachusetts Institute of Technology  
doi: 10.1162/jocn\_a\_00388

<sup>1</sup>Department of Psychology, University of Saskatchewan, Saskatoon, Canada

<sup>2</sup>Department of Experimental Psychology, Ghent University, Ghent, Belgium

<sup>3</sup>Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour,  
Nijmegen, The Netherlands

<sup>4</sup>Department of Cognitive Science, Central European University, Budapest, Hungary

Address correspondence to:

Janeen Loehr

Department of Psychology

University of Saskatchewan

9 Campus Drive

Saskatoon, Saskatchewan, Canada, S7N 5A5

janeen.loehr@usask.ca

## Abstract

We investigated whether people monitor the outcomes of their own and their partners' individual actions as well as the outcome of their combined actions when performing joint actions together. Pairs of pianists memorized both parts of a piano duet. Each pianist then performed one part while their partner performed the other; EEG was recorded from both. Auditory outcomes (pitches) associated with keystrokes produced by the pianists were occasionally altered in a way that either did or did not affect the joint auditory outcome (i.e., the harmony of a chord produced by the two pianists' combined pitches). Altered auditory outcomes elicited a feedback-related negativity whether they occurred in the pianist's own part or the partner's part, and whether they affected individual or joint action outcomes. Altered auditory outcomes also elicited a P300 whose amplitude was larger when the alteration affected the joint outcome compared to individual outcomes, and when the alteration affected the pianist's own part compared to the partner's part. Thus, musicians engaged in joint actions monitor their own and their partner's actions as well as their combined action outcomes, while at the same time maintaining a distinction between their own and others' actions and between individual and joint outcomes.

Efficient and flexible behavior requires that people monitor the outcomes of their actions to ensure that they achieve their intended goals. Much research has been devoted to understanding the cognitive and neural mechanisms underlying action monitoring and control (see Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004, for a review). This work has focused almost exclusively on people's behavioral and neural responses to errors when they perform tasks alone. Little is known about how action-monitoring mechanisms operate during joint actions, that is, when two or more people coordinate their actions to achieve a shared goal. Compared to individual actions, joint actions may pose several additional challenges for action monitoring. First, joint actions often involve simultaneous actions by different individuals and may thus create the necessity to monitor one's own as well as partners' actions in parallel. Second, many joint action outcomes are not simply the sum of individual action outcomes. For instance, the same tones produced by an individual musician may become part of different harmonies depending on the tones another musician is simultaneously producing. This raises the question of whether people monitor their own or their partners' actions with respect to individual action goals (those necessary to achieve each individual's part of the joint action) or with respect to shared action goals (the combined outcome of their coordinated actions). The current study addresses these questions using duet music performance, in which pairs of performers produce complementary sequences of action that are precisely coordinated in time to produce a joint outcome, the musical piece.

### *EEG Markers of Individual Action Monitoring*

Investigations of action monitoring have identified several event-related potentials (ERPs) that arise in response to errors and to feedback about action outcomes. Of particular interest in the current study is the feedback-related negativity (FRN), which has a frontocentral scalp distribution and peaks approximately 250 ms after people receive feedback indicating that they have produced an error (Holroyd & Coles, 2002; Miltner,

Braun, & Coles, 1997) or feedback indicating an unfavourable outcome, such as monetary loss (Gehring & Willoughby, 2002; Hajcak, Holroyd, Moser, & Simons, 2005). Some researchers have argued that the FRN reflects a mismatch between the expected and actual outcome of an action, regardless of whether the outcome is positive or negative (Oliveira, McDonald, & Goodman, 2007). Recent theories have postulated that the FRN reflects action monitoring activity in the anterior cingulate cortex, an area of the posterior medial frontal cortex that is involved in the detection of errors, response conflict, and unfavourable action outcomes (Botvinick, Cohen, & Carter, 2004; Carter & van Veen, 2007; Nieuwenhuis, Holroyd, Mol, & Coles, 2004; Ridderinkhof et al., 2004).

The FRN is often followed by a P300 potential, a positivity that peaks 300-600 ms after action-related feedback, or more generally after any stimulus that is task-relevant (or “motivationally significant”; Nieuwenhuis, Aston-Jones, & Cohen, 2005). The most popular account of the P300’s functional significance holds that it indexes the revision of a mental model of the environmental context (Donchin & Coles, 1988), and that its amplitude is proportional to the change in the model. A more recent account proposes that the P300 reflects noradrenergic facilitation of the response to a stimulus, which scales with the significance of the stimulus (Nieuwenhuis et al., 2005). Both accounts converge on the ideas that the process indexed by the P300 is preceded by an evaluation of stimulus significance, and that P300 amplitude scales according to this significance (Nieuwenhuis et al., 2005). Consistent with these ideas, P300 amplitude scales with the magnitude of reward or loss indicated by feedback about action outcomes (Sato et al., 2005; Yeung & Sanfey, 2004). The P300 may thus reflect a later stage of feedback processing that is related to the evaluation of the significance of the feedback.

*Action Monitoring During Joint Action*

Many joint actions, including duet music performance, require continuous coordination of complementary actions in order to achieve a jointly intended outcome. However, such paradigmatic cases of joint action have hardly been addressed in previous cognitive neuroscience research; instead, researchers have focused almost exclusively on situations in which two people take turns performing similar tasks (Knoblich, Butterfill, & Sebanz, 2011). Turn-taking paradigms have been used to show that action-monitoring processes can be applied to other people's actions in addition to one's own. For example, Yu and Zhou (2006) showed that FRNs were elicited by negative outcomes of both one's own and another person's actions in a gambling task. Similarly, the error-related negativity (ERN) is elicited by both one's own errors and observed errors (de Bruijn, 2012; van Schie, Mars, Coles, & Bekkering, 2004). Like the FRN, the ERN has a frontocentral scalp distribution; however, the ERN is elicited by response errors (e.g., incorrect movements) and peaks approximately 80 ms after the error has occurred.<sup>1</sup> Both the ERN and the FRN are thought to reflect action-monitoring processes that are elicited by the first indication that an action is incorrect, whether this arises from internal information (incorrect movement, eliciting an ERN) or external information (feedback about the action outcome, eliciting a FRN; Holroyd & Coles, 2002; Stahl, 2010).

To date, only one study has examined whether people apply action-monitoring processes to their own and another person's actions when they must act simultaneously (Picton, Saunders, & Jentsch, 2012). Pairs of participants performed separate but simultaneous choice reaction-time tasks, and received feedback about the accuracy of their responses after every trial. Because participants sat side-by-side and were not instructed to directly observe each other's actions, they had access to different indicators that their own and their partners' actions were incorrect: the movement itself in the case of their own errors, and feedback about the action outcome in the cases of both their own and their partners'

errors. Consistent with the hypothesis that action-monitoring processes are elicited by the first indication of an error, each person's own errors elicited ERNs. Furthermore, feedback indicating that the partner had made an error elicited the FRN, indicating that people do monitor their partners' action outcomes when they perform tasks simultaneously. However, Picton et al.'s (2012) paradigm does not allow people's neural responses to feedback about their own and their partners' action outcomes to be directly compared, because own errors comprised incorrect movements in addition to outcome-related feedback. The first goal of the current study was to provide further evidence that people monitor their own and their partners' action outcomes in parallel, and at the same time to directly compare people's neural responses to these action outcomes. This was accomplished by manipulating the auditory outcomes resulting from correct movements made by a pianist and his or her duet partner, as will be described in more detail below.

The second goal of the current study was to investigate whether people monitor each person's individual part in a joint action and/or the combined outcome of their coordinated actions. Recent theory suggests that each person involved in a joint action must minimally represent (and monitor) his or her own part in the joint action and the goal of the joint action (Vesper, Butterfill, Knoblich, & Sebanz, 2010), but must not necessarily represent their partner's part in the joint action. For example, in the case of a musical duet one performer may simply represent her part and the goal of coordinating her actions in time with her partner's. Recent empirical work suggests that people can indeed form representations of joint task goals when they take turns performing actions (Tsai, Sebanz, & Knoblich, 2011), but it has not yet been established whether people monitor the joint outcome of coordinated actions. Duet music performance affords a clear distinction between the individual parts and shared goals of a joint action, as each performer is required to produce their own individual part which, when combined, creates the whole musical piece (Keller, 2008). Next, we discuss how

research on action monitoring in solo music performance can be extended to duet music performance in order to answer our research questions.

### *Action Monitoring in Music Performance*

Successful music performance requires that musicians monitor the auditory consequences of their actions. Years of training on an instrument lead to strong associations between a given movement or set of movements and a given auditory outcome (Drost, Rieger, Brass, Gunter, & Prinz, 2005; Haueisen & Knösche, 2001; see Zatorre, Chen, & Penhune, 2007, for a review). Consequently, manipulating the auditory outcomes associated with musicians' movements so that they do not match action-based expectations disrupts solo performance (Keller & Koch, 2006; Pfordresher, 2003; see Pfordresher, 2006, for a review) and elicits ERP components associated with action monitoring (Katahira, Abla, Masuda, & Okanoya, 2008; Maidhof, Vavatzanidis, Prinz, Rieger, & Koelsch, 2009).

Maidhof et al. (2009) asked pianists to perform musical sequences on a keyboard while the pitches associated with their keystrokes were occasionally altered to produce a mismatch between the actual and expected action outcome. EEG activity in response to altered pitches revealed both a FRN and a P300 compared to EEG activity elicited by correct pitches. These findings indicate that pianists monitor not only their movements but also the auditory consequences of their actions, and respond to feedback indicating that their intended auditory outcomes have not been achieved. We reasoned that if musicians are able to monitor their co-performers' actions and the joint outcome of their combined actions during duet performance, then similar components should be elicited by altered pitches indicating that their co-performers' individual intended outcomes or jointly intended outcomes have not been achieved (i.e., by mismatches between the actual and expected pitches produced by their partners or by their combined actions).

In sum, we aimed first to examine whether people monitor their own and their partner's actions during joint actions, and second to determine whether people monitor each person's individual action outcomes as well as the joint outcome of their combined actions. We asked pairs of pianists to memorize two-part piano duets. Each pianist then performed one part while their partner performed the other, while EEG was recorded from both. During the duet performances, we occasionally altered the pitches elicited by one or the other pianist's keystrokes to create mismatches between actual and expected action outcomes.

The musical piece as a whole consisted of a sequence of four-pitch chords. Each pianist produced two of the four pitches in each chord. When combined, each set of four pitches created a specific harmony (musical relationship between the notes of a chord). This allowed us to examine pianists' responses to two types of pitch alteration. The first changed a pitch in one of the pianists' parts without changing the harmony of the chord to which the pitch belonged; thus, this type of alteration affected only one performer's individual part. The second type of pitch alteration changed both a pitch and the harmony of the chord; thus, this type of alteration affected not only one performer's individual part but also the joint outcome, that is, the musical harmony produced by the two parts combined. We predicted that pitch alterations would elicit a FRN and a P300 regardless of whether they occurred in the pianist's part or the partner's part, indicating that pianists monitored both their own and their co-performers' parts of the performance. We also predicted that altered pitches that affected the joint outcome would elicit larger responses than altered pitches that affected only individual outcomes, particularly at the later stage of processing captured by the P300, reflecting the significance of shared goals in joint task performance.

## Method

### *Participants*

12 pianists (7 male, mean age = 21.58, SD = 3.71) were recruited from music schools in the Netherlands and participated in pairs. Four of the six pairs had never played music together before the experiment. All but two of the pianists were right-handed. All pianists had at least 7 years of private piano lessons, except one who had only 4 years ( $M = 10.50$ ,  $SD = 3.60$ ). Only pianists who could successfully perform the stimuli from memory were included in the study.

### *Equipment*

Melodies were performed on a Yamaha P-95B weighted-key digital piano. Presentation of metronome pulses and auditory feedback was implemented via Max/MSP 5.1.7 software run on a Macintosh computer. Piano tones were generated using a piano timbre and metronome pulses were generated using a drum timbre from the built-in internal sound card on an iMAC 8.1 computer. Participants heard the metronome pulses and performances over two speakers placed in front of the keyboard and set at a comfortable volume.

### *Stimuli and Design*

Two piano duets were composed for the study. The first half of one of the duets is shown in Figure 1. Each duet consisted of 32 four-voice chords (voices from highest to lowest frequency: soprano, alto, tenor, bass). The duets were composed so that two of the voices (soprano and alto) could be performed with the right hand and two of the voices (tenor and bass) with the left hand. One chord occurred on each quarter-note beat of the piece, which was notated in 4/4 time. The chords were separated into four 8-chord phrases, each of which was marked by a final fermata indicating that pianists should pause at the end of the phrase. One duet also contained eighth notes between three of the chords, which served to link the chords musically.

The duets were composed so that the harmonic transitions between chords conformed to the rules of harmony in Western classical music. Within each piece, eight chords were identified whose harmony could be altered so that it was either musically expected (e.g., a major chord built on the fourth scale degree that followed a major chord built on the second scale degree, which is a typical transition in Western music) or musically less expected (e.g., a minor chord built on the second scale degree that followed a major chord built on the second scale degree, which is a less typical transition in Western music). These chords served as *alteration* chords (chords whose pitches were occasionally altered). The musically less expected chords were included in the musical score, which pianists were required to memorize and perform during the experiment. Thus, pitch alterations that changed the musical harmony (and thus affected the joint outcome) resulted in musically more expected chords. This ensured that participants' responses to these pitch alterations were not due to encountering a musically unexpected chord. Half of the alteration chords occurred on strong beats and half on weak beats, and none occurred on a chord preceded by an eighth note.

Pitch alterations were composed for the soprano note in the right-hand part and the bass note in the left-hand part of each alteration chord (see Figure 1). The outermost voices were chosen so as to maximize the salience of the altered pitches. There were two types of pitch alterations. *Individual* alterations changed the pitch that was heard but did not change the harmony of the chord to which the pitch belonged. These alterations required shifting the pitch up or down by 4.6 semitones on average. *Joint* alterations changed both the pitch that was heard and the harmony of the chord to which the pitch belonged. These alterations required shifting the pitch up or down by 1.8 semitones on average.<sup>2</sup> Individual and joint pitch alterations were distributed across the eight chords on which alterations could occur such that individual alterations occurred in the soprano part (produced by one member of the pair) and joint alterations occurred in the bass part (produced by the other member of the

pair) for half of the chords, and individual alterations occurred in the bass part and joint alterations in the soprano part for the other half of the chords. For the participant producing the soprano part, an alteration of the soprano pitch functioned as a *self* alteration, whereas the same alteration functioned as an *other* alteration for the participant producing the bass part. The opposite was true for the person producing the bass part (soprano = *other*, bass = *self*). Thus, we compared participants' responses to pitch alterations in a 2 (person: self, other) x 2 (outcome: individual, joint) within-subjects design.

### *Procedure*

Participants were tested in pairs. Each pair was randomly assigned one of the two duets, which they were asked to memorize before coming to the lab for EEG recording. Participants were told they would be performing the piece as a duet, and they were asked to memorize both the right- and left-hand parts so that they could perform either one while their partner performed the other. They were free to practice both hands together if they wished to do so. Participants were sent the musical score and a set of 6 audio recordings (3 of the right-hand part and 3 of the left-hand part), and were asked to practice until they could perform the right- and left-hand parts along with the audio recordings of the other part in succession without any errors. This extensive practice ensured that participants would be able to perform the pieces without any errors when they arrived at the lab and that participants were very familiar with both duet parts.

After arriving at the lab, participants were given a few minutes to warm up. Each participant then performed the two parts of the duet from memory to verify that they had correctly memorized both parts. All were able to perform from memory with no errors. They were then informed that they would be allowed to see the score during performance (pilot testing indicated that pianists had difficulty performing numerous error-free repetitions of the piece without the support of the score). A copy of the score was then placed approximately

90 cm in front of each participant. Participants were asked to move their eyes as little as possible during the recorded performances. They then practiced performing the duet together. This was followed by paced practice trials in which a metronome was sounded four times (three times for the piece that began with an upbeat) at 800 ms interonset intervals (IOIs) at the beginning of each trial and was then turned off. Participants were instructed to perform the piece at the pace set by the initial metronome. Participants were then informed that they would occasionally hear incorrect pitches in their own or their partner's part and were asked to continue performing in spite of the incorrect pitches. They completed four practice trials with different pitch alterations than those employed in the test trials.

Participants were then fitted with EEG caps, after which they completed four blocks of 30 experimental trials, which were also paced by an initial metronome. Participants were required to perform each trial without any errors. If an error was committed, the performance was stopped and the trial was repeated at the end of the block. Within each block, each of the eight alteration chords was altered six times (three alterations in the soprano and three in the bass; never both in the same chord). Thus, altered pitches were presented in 20% of the performances of each violation chord. The violations were randomly distributed across the performances in each block with the constraint that each performance contained at most four chords with altered pitches. In total, the performances contained 48 tones with altered pitch and 384 corresponding tones with correct pitch per condition. The experiment took approximately 4 hours to complete, and participants were paid €60.

#### *Data Acquisition*

The musical performances (including key press times, velocities, and pitches) were recorded using the Max/MSP software, which also sent trigger signals to the EEG acquisition computer concurrently with the auditory feedback (correct or altered) associated with the soprano and bass notes in each alteration chord.<sup>3</sup> EEG was recorded continuously from both

participants using 32 active electrodes (Acticap, Brain Products GmbH, Germany) per participant, arranged according to an extended version of the 10-20 system at F7, F3, Fz, F4, F8, FC5, FC1, FCz, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, and O2, using carefully positioned nylon caps. All electrodes were referenced to the left mastoid during recording. Vertical eye movements were monitored using pairs of bipolar electrooculography (EOG) electrodes positioned directly above and beneath the right eye, and horizontal eye movements were monitored using pairs of bipolar EOG electrodes positioned at the outer canthi of the eyes. Impedance was kept below 10 k $\Omega$ . EEG and EOG signals were amplified within a bandwidth of 0.05-100 Hz and digitized with a sampling frequency of 1000 Hz.

#### *Data Processing and Analysis*

EEG data processing was performed off-line using Brain Vision Analyzer software (V. 1.05, Brain Products GmbH, Germany). EEG data were first re-referenced to the mean of both mastoid electrodes. Automated ocular correction was performed using the procedure by Gratton, Coles, & Donchin (1983) to eliminate artifacts induced by horizontal or vertical eye movements. The data were filtered using a high-pass filter of 0.01 Hz (24 dB/oct) and a low-pass filter of 40 Hz (24 dB/oct) in order to remove slow drifts and excessive noise, respectively. The corrected EEG data were then segmented into epochs from 100 ms before to 800 ms after tone onset. Individual trials were removed if they contained further artifacts possibly induced by head, body, or arm movements, as indicated by a difference between the maximum and the minimum value within a given segment that exceeded 150  $\mu$ V. Averages were calculated separately for each subject and each condition. The 100 ms before tone onset was used as the baseline period.

Difference waves were computed on individual averages by subtracting the ERP waveforms elicited by correct pitches from the ERPs elicited by altered pitches. The FRN

was defined on this difference wave using a peak-to-peak analysis in which the most positive peak within 80 to 140 ms after tone onset was subtracted from the most negative peak within 200 to 300 ms after tone onset. This analysis was conducted on electrodes Fz, FC1, FCz, FC2, and Cz, where FRN amplitudes were maximal in both the current study and in previous FRN studies (e.g., Miltner et al., 1997). Time windows for the analysis were chosen based on grand-average peak latencies. The P300 was defined as the mean amplitude of the difference wave between 400 and 600 ms after tone onset. Consistent with previous research, the P300 exhibited a parietocentral scalp distribution (Nieuwenhuis et al., 2005; Polich, 2007).

Because the P300 was slightly lateralized over the right hemisphere in the current study, we conducted the analysis on electrodes Pz, CP2, and P4, where amplitudes were maximal. Time windows for the analysis were chosen based on grand-average peak latencies. Values for the FRN and P300 were compared across conditions by repeated-measures ANOVAs with factors person (self, other) and outcome (individual, joint). FRN and P300 values were also compared against zero within conditions to determine whether the difference between responses to correct and altered pitches reached significance.

For the behavioral data, interonset intervals (IOIs) were calculated between the onset of an altered pitch and the subsequent pitch within the same voice, for each of the four conditions. IOIs were also calculated between the onset of each correct pitch and the subsequent pitch within the same voice. The mean IOI following correct pitches was then compared to the mean IOI following altered pitches in each of the four conditions. IOIs exceeding 1000 ms (i.e., those following a fermata indicating a pause in performance) were excluded from analysis.

## Results

### *Behavioural Analyses*

IOIs following correct pitches ( $M = 721.78$  ms,  $SD = 25.87$ ) did not differ from IOIs following altered pitches in any of the four conditions ( $M_{\text{self,joint}} = 722.79$  ms,  $SD = 28.66$ ;  $M_{\text{self,individual}} = 718.97$  ms,  $SD = 24.00$ ;  $M_{\text{other,joint}} = 721.80$  ms,  $SD = 26.40$ ;  $M_{\text{other,individual}} = 721.10$  ms,  $SD = 28.31$ ,  $t_s < 1.5$ ,  $p_s > .18$ ). Thus, there was no evidence of post-error slowing (Rabbitt, 1966) in any of the conditions, consistent with previous work showing a lack of post-error slowing in response to altered auditory outcomes in solo piano performance (Maidhof et al., 2009). The average IOI of 721.29 ms ( $SD = 26.23$ ) was faster than the prescribed tempo of 800 ms per IOI,  $t(11) = 10.47$ ,  $p < .001$ , consistent with previous work on duet performance paced by an initial metronome (Loehr & Palmer, 2011).

### *ERP Analyses*

We examined whether altered auditory outcomes elicited a FRN, and if so, whether FRN amplitude differed across conditions. Figure 2A shows the grand-average waveforms and difference waves, pooled over electrode sites Fz, FC1, FCz, FC2, and Cz, time-locked to the onset of correct and altered pitches, for each condition. Figure 2A also shows the scalp voltage distribution of the difference wave for each condition within the time window of analysis. As expected, altered pitches (compared to correct pitches) elicited a negative deflection with a frontocentral scalp distribution in all four conditions. As shown in Figure 2B, the mean voltage difference between responses to correct and altered pitches was significantly different from zero in all four conditions,  $t_s > 10.65$ ,  $p_s < .001$ , and did not differ across conditions. An ANOVA comparing peak-to-peak amplitude across conditions showed no significant main effects or interaction (main effect of person:  $F(1, 11) = 0.86$ ,  $p = .37$ ; main effect of outcome:  $F(1, 11) = 0.022$ ,  $p = .89$ ; interaction:  $F(1, 11) = 0.019$ ,  $p = .89$ ).<sup>4</sup>

We next examined whether altered auditory outcomes elicited a P300 and if so, whether P300 amplitude differed across conditions. Figure 3A shows the grand-average

waveforms and difference waves, pooled over electrode sites Pz, CP2, and P4, time-locked to the onset of correct and altered pitches, for each condition, as well as the scalp voltage distribution of the difference wave for each condition within the time window of analysis. Compared to correct pitches, altered pitches elicited a positive deflection with a right-lateralized parietal scalp distribution. Figure 3B shows the mean voltage difference between responses to correct and altered pitches for each condition. The positive deflection was larger when the altered pitch occurred in the pianist's own part (self conditions) than when it occurred in the partner's part (other conditions). The deflection was also larger when the altered pitch affected the joint outcome than when it affected only one pianist's individual outcome. An ANOVA comparing the voltage differences across conditions confirmed a main effect of person,  $F(1, 11) = 32.40, p < .001$ , a main effect of outcome,  $F(1, 11) = 15.24, p = .002$ , and no interaction,  $F(1, 11) = 0.066, p = .80$ . The voltage difference was significantly greater than zero for the self-joint condition,  $t(11) = 7.34, p < .001$ , the self-individual condition,  $t(11) = 4.40, p = .001$ , and the other-joint condition,  $t(11) = 3.38, p = .006$ , but not for the other-individual condition,  $t(11) = 0.96, p = .36$ .

### Discussion

The current study examined whether people monitor their own and their partner's action outcomes during joint actions, and if so, whether they monitor each other's individual action outcomes as well as the joint outcome of combined actions. Pairs of pianists performed musical duets while the pitches associated with one or the other pianist's actions were occasionally altered so that the auditory outcome of one individual's action (i.e., a single pitch in one pianist's part) was altered or, in addition, the joint outcome of the two pianists' combined actions (i.e., the harmony of a chord jointly produced by the two pianists) was altered. Compared to correct auditory outcomes, all types of altered outcomes elicited a feedback-related negativity (FRN), the amplitude of which did not differ across conditions.

Altered outcomes also elicited a P300 whose amplitude was larger when the alterations occurred in the pianist's own part and when the alterations affected the joint outcome of the pianists' combined actions. These findings indicate that skilled performers are able to monitor the outcomes of their own actions, their co-performers' actions, and their combined actions when they perform joint actions together. They also indicate that performers nevertheless differentiate between their own and others' action outcomes, and between individual and joint action outcomes.

#### *Feedback-Related Negativity*

The negativity elicited by altered auditory outcomes in the current study resembles the FRN in terms of both scalp distribution and latency. The FRN is thought to index the detection of an error based on feedback about an action's outcome (Holroyd & Coles, 2002; Miltner et al., 1997) or the detection of a mismatch between the actual and expected outcome of an action (Oliveira et al., 2007). Years of musical training result in strong associations between actions and their auditory consequences (Drost et al., 2005; Haueisen & Knösche, 2001; Repp & Knoblich, 2009; Zatorre et al., 2007). These learned associations allow an internal forward model to predict the outcomes of the actions using efference copies of the motor command (Miall & Wolpert, 1996; Wolpert & Kawato, 1998). Our findings suggest that pianists formed expectations not only about the auditory outcomes of their own actions but also of their partners' actions, as the FRN was elicited whether the altered pitches occurred in the pianist's own part or the partner's part.

These findings are consistent with the hypothesis that people use internal forward models to predict not only the outcomes of their own actions but also those of their co-performers' actions when they perform joint actions together (Keller, Knoblich, & Repp, 2007; Knoblich & Jordan, 2003; Wolpert, Doya, & Kawato, 2003). Further support for this interpretation is gained by comparing the current results to those of Maidhof et al. (2009),

who showed that the FRN in response to altered auditory outcomes was larger when pianists produced the musical sequences themselves compared to when they merely heard the sequences. If pianists' expectations about their partner's actions were based solely on perceptual processes in the current study, the FRN elicited by alterations of the partner's part should have been smaller than the FRN elicited by alterations of the pianist's own part. Furthermore, it could be argued that participants generated predictions for the outcomes of their partners' actions based on forward model simulation of performing the partner's part themselves (facilitated by extensive practice of both parts of the piece) rather than simulation of the partner's actions per se. However, the finding that P300 amplitude differed depending on whether the altered auditory outcome occurred in the pianist's own part or the partner's part (discussed in more detail in the next section) suggests a distinction between the pianist and the partner that would not be possible if pianists simulated performing both parts themselves. Thus, the current findings are consistent with the hypothesis that pianists monitored action-based expectations for the auditory outcomes of their partner's actions in addition to their own actions.

However, there are alternative interpretations of the negativity that should be considered. One possibility is that the negativity is not a FRN but rather a mismatch negativity (MMN), which reflects the detection of deviant events in an otherwise invariant context (Alho, 1995; Giard, Perrin, Pernier, & Bouchet, 1990). However, pitch alterations in the current study cannot be considered deviants from an invariant context because they were comprised of the same pitches that formed the context. Likewise, chords that served as altered harmonies were taken from the same set of chords that comprised the context.

A second possibility is that altered pitches may have violated pianists' expectations based on the musical structure of the sequence. When people who are familiar with (Western) tonal music listen to a sequence of chords, they generate expectations for upcoming chords

based on implicit knowledge of musical structure (Bharucha & Krumhansl, 1983; Krumhansl, Bharucha, & Kessler, 1982). Perceiving a chord that is musically unexpected relative to the preceding context elicits a (sometimes) right-lateralized frontocentral negativity that peaks around 180 ms and is thought to rely on representations of music-syntactic regularities held in long-term memory (Koelsch, 2005; Koelsch, Gunter, Friederici, & Schroger, 2000).

However, it is unlikely that the negativities elicited by altered auditory outcomes in the current study reflect violations of musical expectancy. The stimuli were designed such that the harmony of every chord in which pitch alterations occurred was relatively less expected musically, and alterations that changed the harmony created a musically more expected chord. Thus, if musical expectancy drove the current effects, there should have been a) no difference between correct pitches and altered pitches that affected only an individual's part (i.e., entailed no harmony change), because the chord would have been unexpected in both cases, and b) a larger negativity in response to correct pitches than to altered pitches that affected the joint outcome (i.e., changed the harmony), because the chord was unexpected when it contained the correct pitch but expected when it contained an altered pitch. Thus, musical expectancy cannot explain the pattern of results obtained in this study.

A third possibility is that pianists may have generated expectancies for auditory events based on visual perception of the musical score, which was available to the pianists at all times. Trained musicians are able to generate "auditory images" of tones based on visual perception of a score (Schön & Besson, 2005; Yumoto et al., 2005). When they concurrently perceive tones that do not match this image, a negative component termed the imagery mismatch negativity (iMMN) is elicited. However, studies that have demonstrated the iMMN have presented participants with melodic sequences (i.e., sequences of single pitches) rather than chord sequences. Thus, it is not clear that an iMMN would also be elicited by a mismatch between a single pitch presented concurrently with three other pitches (i.e., in a 4-

pitch chord) and a visual representation of the chord to which the pitch belongs, as occurred in the current study.

A final possibility is that pianists generated auditory images not based on visual perception of the score but on long-term memory representations of the musical piece. Herholz, Lappe, Knief, and Pantev (2008) showed that an iMMN is elicited by a mismatch between a perceived pitch and the auditory image of a remembered melody (i.e., in the absence of a score). Equivalent negativities in response to pitch alterations in the pianist's own and the partner's part are consistent with this possibility, because pianists had memorized both their own and their partner's part of the score equally well as a prerequisite for participating in the study. However, it is likely that pianists formed integrated auditory-motor representations of the musical piece, given that they practiced the pieces with auditory feedback. Integrated auditory-motor representations are formed even without extensive practice or musical training (Bangert & Altenmüller, 2003; Lahav, Saltzman, & Schlaug, 2007), and coupling between auditory and motor systems is particularly strong in trained musicians (Bangert et al., 2006; Haueisen & Knösche, 2001; Zatorre et al., 2007). The negativities elicited in the current study might therefore reflect mismatches between perceived pitches and integrated auditory-motor representations of the musical piece; this is not incompatible with our interpretation that the negativities reflect violations of performers' action-based expectancies.

### *P300*

Altered auditory outcomes also elicited a parietal, right-lateralized P300 whose amplitude was larger when the alteration affected the pianist's own action outcome compared to the partner's outcome. Given that the P300 amplitude scales with evaluation of stimulus significance (Donchin & Coles, 1988; Nieuwenhuis et al., 2005; Sato et al., 2005; Yeung & Sanfey, 2004), this finding suggests that in joint action tasks, own action outcomes are more

significant than a co-performer's outcomes. Own outcomes may be evaluated as more significant than a partner's because only own outcomes can be subject to correction (in future performances if not the current one). Heightened salience of one's own action outcomes is also consistent with previous work showing that self-relevant stimuli such as one's own name or face elicit larger P300s than stimuli that refer to other people (Perrin et al., 2005).

Differentiation between own and others' action outcomes is consistent with previous music performance research showing differences in corticospinal excitability depending on whether an action representation was associated with the self (solo performance) or with a partner (joint performance; Novembre, Ticini, Schutz-Bosbach, & Keller, in press). This differentiation may also explain the right-lateralization of the P300. The main generators of the P300 are thought to be located in parietal and temporal areas (Bledowski et al., 2004; Linden, 2005), particularly around the temporo-parietal junction (TPJ; see Polich, 2007; Verleger, 2008), and activity in the right TPJ is associated with maintaining a distinction between self and other (Blakemore & Frith, 2003; Decety & Grèzes, 2006). However, it is also possible that the right-lateralization of the P300 is due to right-hemisphere specialization for pitch or spectral acoustical processing (Zatorre, Belin, & Penhune, 2002). Although this specialization is clearest in the auditory cortex, it has also been observed in higher-order processing areas (e.g., the right intraparietal sulcus during melody transposition; Foster & Zatorre, 2010).

The P300 was also larger when the alteration affected the joint outcome compared to either individual's action outcome. Previous EEG studies of joint turn-taking tasks have shown that monitoring a co-actor's task is reflected in enlarged P300 amplitudes (Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai, Kuo, Hung, & Tzeng, 2008). Our findings suggest that combined action outcomes are monitored and evaluated as more significant than outcomes associated with only one individual's actions. There are at least two reasons why

this might be the case: either because the joint outcome reflects two goals belonging to a single individual (i.e., the pianist's goals for his own part of the performance and for the joint outcome of the performance) or because the joint outcome reflects two people's goals (i.e., the pianist's goal for the joint outcome and the partner's goal for the joint outcome).

Although these two possibilities cannot be unequivocally disentangled in the present study, the latter possibility is consistent with fMRI research showing stronger activation in posterior parietal areas when one's own errors have consequences for both oneself and another person compared to when one's errors have consequences only for oneself (Radke, de Lange, Ullsperger, & De Bruijn, 2011). Note that in Radke et al.'s (2011) study, activity in the posterior medial frontal cortex, thought to be the main generator of the FRN, did not differ depending on whether errors affected another person in addition to oneself. Thus, these fMRI findings are also consistent with the fact that FRN amplitudes did not differentiate between individual and joint action outcomes in the current study.

One possible alternative explanation for the larger P300 responses elicited by pitch alterations that affected the joint outcome compared to individual action outcomes is that the former entailed changes to musical harmony whereas the latter did not. Thus, P300 amplitudes may reflect differences in the degree to which musical expectancies were violated. To our knowledge, listeners' responses to pitch alterations that do or do not change the harmony of chords within a previously-learned sequence have not yet been compared. However, previous research that compared listeners' responses to chords whose harmony was more or less unexpected has shown effects earlier in the ERP (i.e., the early right anterior negativity discussed above; Koelsch, 2005). We found no such effects on the FRN that preceded the P300; therefore, a purely perceptual explanation for the current P300 findings seems unlikely.

### *Monitoring Joint Actions*

The current study furthers understanding of action monitoring in joint action contexts in several ways. First, our findings indicate that people monitor both their own and another person's actions in parallel when they have to precisely coordinate their actions in time to achieve a common goal. This is consistent with previous work showing that people monitor feedback indicating whether or not their partner made an error when they perform independent choice reaction-time tasks at the same time (Picton et al., 2012). The present study shows that monitoring a partner's actions also occurs when people perform complex sequences of complementary actions together. Furthermore, the direct comparison between people's responses to their own and their partners' action outcomes, made possible by manipulating the auditory outcomes associated with correct movements for both performers, revealed no differences in FRN amplitude as a function of agency. This is consistent with previous work showing ERNs of similar amplitude in response to own and others' errors (de Bruijn, 2012), as well as with fMRI data showing equivalent activation in the posterior medial frontal cortex in response to own and others' errors (de Bruijn, de Lange, von Cramon, & Ullsperger, 2009), in turn-taking tasks. Although one previous study showed larger FRNs in response to one's own action outcomes than a partner's action outcomes (Yu & Zhou, 2006), this study used a gambling task in which action outcomes could not be predicted in advance. In contrast, in the current study the pianist's own action outcomes and the partner's action outcomes could be predicted equally well, as pianists had extensive practice with both parts of the task. Together, then, these findings indicate that people are equally able to monitor their own and their partners' action outcomes when they can predict what those outcomes should be.

Second, the current findings shed light on an important question that arises from the growing body of research showing that people represent and monitor each other's actions when they perform tasks together: namely, how a distinction between self and other is

maintained despite these shared representations and monitoring processes (Decety & Sommerville, 2003). Here, we show that despite the similarity of earlier neural responses to one's own and others' actions outcomes (the FRN), later processing of action outcomes differentiates between the two (the P300). Thus, the current findings demonstrate a time-course of processing that includes both shared monitoring processes and a self-other distinction, both of which may be crucial for success at joint action tasks.

Finally, the current findings expand on previous work examining people's ability to form representations of the shared goals of their combined actions when performing actions with another person. Consistent with previous work showing that people form joint task representations (Tsai et al., 2011), we show that people represent and monitor the joint goal of their combined actions in addition to the outcomes of their own actions and their partners' actions. We also show that people's neural responses to feedback indicating that a joint goal has not been achieved are stronger than their neural responses to feedback indicating that either person's individual goals for the task have not been achieved. As with the distinction between one's own and others' action outcomes, the distinction between individual and joint action outcomes is evident at the later stages of feedback processing reflected in the P300. Together, then, our findings show that people can monitor all the components of a joint action, while at the same time distinguishing between their own action outcomes and their partners', as well as between action outcomes resulting from one individual's actions and from both partners' combined actions.

### Conclusion

The current findings indicate that people monitor not only their individual contributions to a joint action, but also their partner's actions and the combined outcome of their coordinated actions. They also suggest that action outcomes that affect the shared goal of a joint action are perceived as more significant than those that affect only one individual's

contribution to the shared goal; likewise, one's own action outcomes are more significant than one's co-performers'. Thus, successful joint action relies not only on monitoring one's own actions but also the shared goal of coordinated actions. Moreover, when people perform joint actions together, they are able to apply action-monitoring processes to their own and another person's actions, while at the same time maintaining a distinction between the two.

## References

- Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing, 16*, 38-51.
- Bangert, M., & Altenmuller, E. O. (2003). Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience, 4*, 36-40.
- Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H. J., & Altenmuller, E. O. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage, 30*, 917-926.
- Bharucha, J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13*, 63-102.
- Blakemore, S. J., & Frith, C. D. (2003). Self-awareness and action. *Current Opinion in Neurobiology, 13*, 219-224.
- Bledowski, C., Prvulovic, D., Hoechstetter, K. S., M., Wibral, M., Goebel, R., & Linden, D. E. J. (2004). Localizing P300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study. *The Journal of Neuroscience, 24*, 9353-9360.
- Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8*, 539-546.
- Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience, 7*, 367-379.
- de Bruijn, E. R. A. (2012). Is your error my concern? An event-related potential study on own and observed error detection in cooperation and competition. *Frontiers in Neuroscience, 6*, 1-9.

- de Bruijn, E. R. A., de Lange, F. P., von Cramon, D. Y., & Ullsperger, M. (2009). When errors are rewarding. *The Journal of Neuroscience*, *29*, 12183-12186.
- Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*, 4-14.
- Decety, J., & Sommerville, J. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*, 527-533.
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357-374.
- Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). Action-effect coupling in pianists. *Psychological Research*, *69*, 233-241.
- Foster, N. E. V. & Zatorre, R. J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex*, *20*, 1350-1359.
- Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*, 2279-2282.
- Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysics*, *27*, 627-640.
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology*, *55*, 468-484.
- Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*, 161-170.
- Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, *13*, 786-792.

- Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imagery and the effect of musical expertise. *European Journal of Neuroscience*, *28*, 2352-2360.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679-709.
- Katahira, K., Abla, D., Masuda, S., & Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: An ERP study. *Neuroscience Research*, *61*, 120-128.
- Keller, P. E. (2008). Joint action in music performance. In F. Morganti, A. Carassa & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective on the study of interactions* (pp. 205-221). Amsterdam, The Netherlands: IOS Press.
- Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Consciousness and Cognition*, *16*, 102-111.
- Keller, P. E., & Koch, I. (2006). The planning and execution of short auditory sequences. *Psychonomic Bulletin & Review*, *13*, 711-716.
- Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. *The Psychology of Learning and Motivation*, *54*, 59-101.
- Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1006-1016.
- Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*, x1-x6.

- Koelsch, S., Gunter, T., Friederici, A. D., & Schroger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience, 12*, 520-541.
- Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 24-36.
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience, 27*, 308-314.
- Linden, D. E. (2005). The P300: Where in the brain is it produced and what does it tell us? *Neuroscientist, 11*, 563-576.
- Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *Quarterly Journal of Experimental Psychology, 64*, 2153-2167.
- Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., & Koelsch, S. (2009). Processing expectancy violations during music performance and perception: An ERP study. *Journal of Cognitive Neuroscience, 22*, 2401-2413.
- Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks, 9*, 1265-1279.
- Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience, 9*, 788-798.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision-making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin, 131*, 510-532.

- Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neuroscience and Biobehavioral Reviews*, 441-448.
- Novembre, G., Ticini, L. F., Schutz-Bosbach, S., & Keller, P. E. (in press). Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cerebral Cortex*.
- Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate cortex is not all error related: Expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, 19, 1994-2004.
- Perrin, F., Maquet, P., Peigneux, P., Ruby, P., Degueldre, C., Balteau, E., . . . Laureys, S. (2005). Neural mechanisms involved in the detection of our first name: A combined ERPs and PET study. *Neuropsychologia*, 43, 12-19.
- Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 949-964.
- Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology*, 2, 183-198.
- Picton, L., Saunders, B., & Jentsch, I. (2012). "I will fix only my own mistakes": An ERP study investigating error processing in a joint choice-RT task. *Neuropsychologia*, 50, 777-785.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128-2148.
- Rabbitt, P. M. A. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology* 71, 264-272.

- Radke, S., de Lange, F. P., Ullsperger, M., & De Bruijn, E. R. A. (2011). Mistakes that affect others: An fMRI study on processing of own errors in a social context. *Experimental Brain Research, 211*, 405-413.
- Repp, B. H., & Knoblich, G. (2009). Performed or observed keyboard actions affect pianists' judgements of relative pitch. *Quarterly Journal of Experimental Psychology, 62*, 2156-2170.
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306*, 443-447.
- Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport, 4*, 407-411.
- Schön, D., & Besson, M. (2005). Visually induced auditory expectancy in music reading: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 17*, 694-705.
- Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in coacting individuals. *Journal of Cognitive Neuroscience, 18*, 859-870.
- Stahl, J. (2010). Error detection and the use of internal and external error indicators: An investigation of the first-indicator hypothesis. *International Journal of Psychophysiology, 77*, 43-52.
- Tsai, C. C., Kuo, W. J., Hung, D. L., & Tzeng, O. J. L. (2008). Action co-representation is tuned to other humans. *Journal of Cognitive Neuroscience, 20*, 2015-2024.
- Tsai, C. C., Sebanz, N., & Knoblich, G. (2011). The GROOP effect: Groups mimic group actions. *Cognition, 118*, 138-143.

- van Schie, H. T., Mars, R. B., Coles, M. G., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, *7*, 549-554.
- Verleger, R. (2008). P3b: Towards some decisions about memory. *Clinical Neurophysiology*, *119*, 968-970.
- Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, *23*, 998-1003.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London B*, *358*, 593-602.
- Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317-1329.
- Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience*, *24*, 6258-6264.
- Yu, R., & Zhou, X. (2006). Brain responses to outcomes of one's own and other's performance in a gambling task. *Neuroreport*, *17*, 1747-1751.
- Yumoto, M., Matsuda, M., Itoh, K., Uno, A., Karino, S., Saitoh, O., . . . Kaga, K. (2005). Auditory imagery mismatch negativity elicited in musicians. *Neuroreport*, *16*, 1175-1178.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*, 37-46.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*, 547-558.

Author Note

This research was supported in part by a Marie Curie International Incoming Fellowship held by the first author (Project number 254419 within the European Union's 7<sup>th</sup> Framework Programme). The authors thank two anonymous reviewers for helpful comments on an earlier version of this manuscript.

## Footnotes

<sup>1</sup>The ERN elicited by *observed* response errors has a latency of approximately 250 ms. However, this component is referred to in the literature as an “observed ERN” rather than an FRN because it is elicited by an (observed) response error rather than by feedback indicating whether or not a response is correct (de Bruijn, 2012; van Schie et al., 2004).

<sup>2</sup>The number of semitones by which pitches were shifted was determined by musical constraints. Four-note chords typically contain notes that are separated by at least 3 semitones. Altering a pitch without changing the harmony of the chord typically requires exchanging one note from within the chord for another, resulting in a change of 3 or more semitones. In contrast, altering the harmony of a chord typically requires changing one of the chord’s notes so that it is one semitone closer to its nearest neighbor.

<sup>3</sup>Due to MIDI transmission times, there was a constant 20 ms ( $\pm 3$  ms) delay between the trigger sent to the EEG software and tone onset. All analyses corrected for this delay.

<sup>4</sup>The same ANOVA conducted on the mean amplitude of the difference wave derived from pooled electrodes between 200 and 300 ms after tone onset yielded the same pattern of results. There were no significant main effects or interaction,  $F_s < 2.00$ ,  $p_s > .18$ , but significant differences between each FRN and zero,  $t_s > 4.45$ ,  $p_s < .001$ .

## Figure captions

*Figure 1.* The first half of one duet with two pitch alterations. Symbols immediately below chords (labeled “learned harmony”) indicate the harmony given in the score and memorized by participants. Symbols labeled “alteration harmony” indicate the harmony introduced by the pitch alteration. A) *Individual* pitch alteration in the soprano voice (*self* condition for right-hand part, *other* condition for left-hand part), which does not alter the harmony of the chord. B) *Joint* pitch alteration in the bass voice (*other* condition for the right-hand part, *self* condition for the left-hand part), which alters the harmony of the chord.

*Figure 2.* A. Grand-average waveforms time-locked to the onset of correct (dotted line) and altered (dashed line) pitches and the average difference waveform (solid line) for each condition, along with the scalp voltage distributions for the difference waveforms. Waveforms are derived from pooled electrode sites (Fz, FC1, FCz, FC2, and Cz, highlighted as white circles). Arrows indicate the feedback-related negativity elicited by altered pitches. B. Mean peak-to-peak voltage differences (and standard deviations) between responses to altered and correct pitches pooled over electrode sites Fz, FC1, FCz, FC2, and Cz.

*Figure 3.* A. Grand-average waveforms time-locked to the onset of correct (dotted line) and altered (dashed line) pitches and the average difference waveform (solid line) for each condition, along with the scalp voltage distributions for the difference waveforms. Waveforms are derived from pooled electrode sites (Pz, CP2, and P4, highlighted as white circles). Arrows indicate the P300 elicited by altered pitches. B. Mean voltage differences (and standard deviations) between responses to altered and correct pitches pooled over electrode sites Pz, CP2, and P4.

Learned harmony: I viio<sup>6</sup> I<sup>6</sup> V<sup>6</sup>/vi vi V/V V I vi vi IV I V<sup>6</sup>/<sub>5</sub>IV IV<sup>6</sup> V I

Alteration harmony:

A. Individual

B. Joint

Figure 1

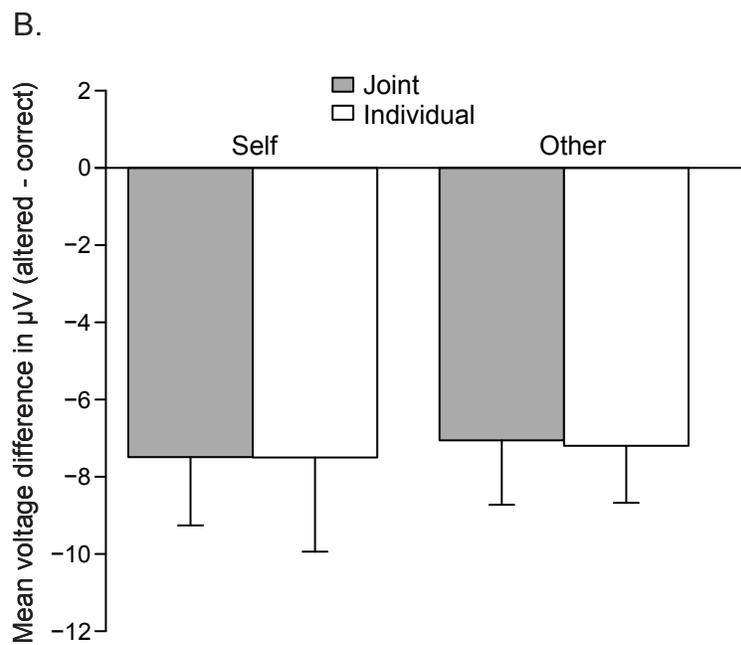
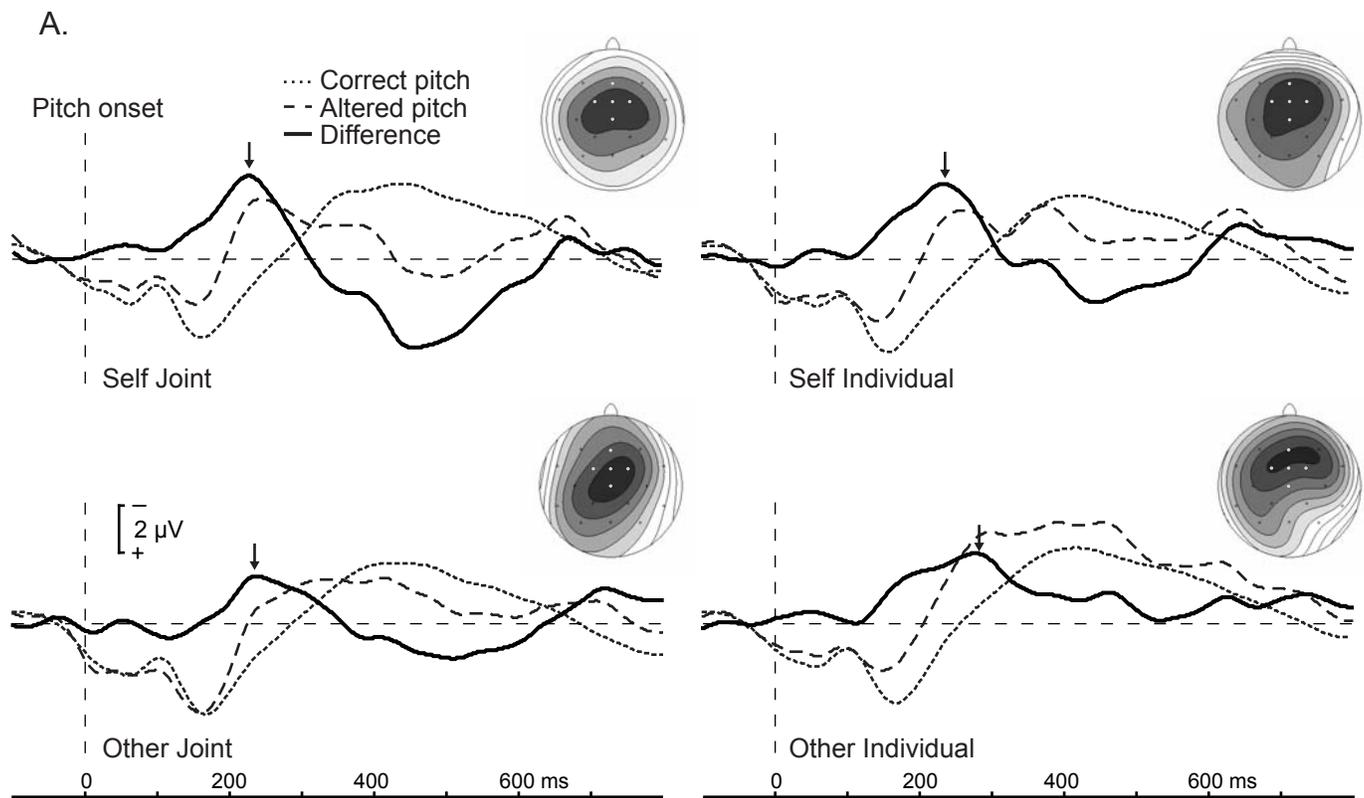


Figure 2

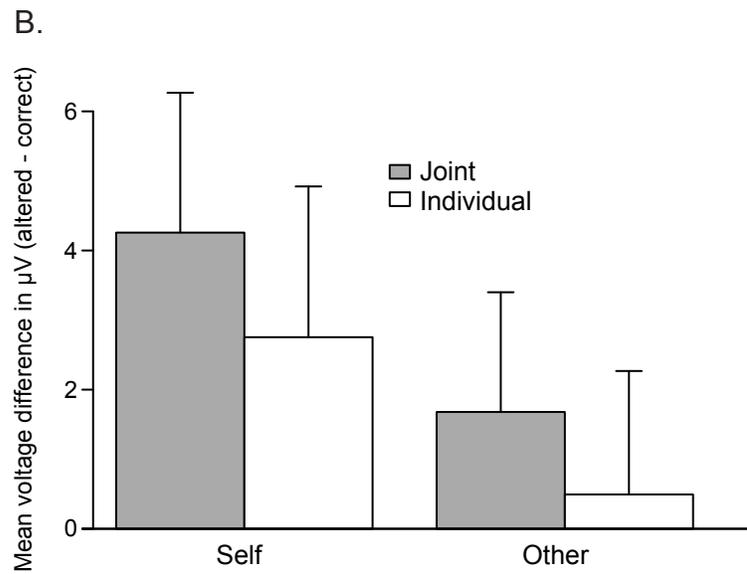
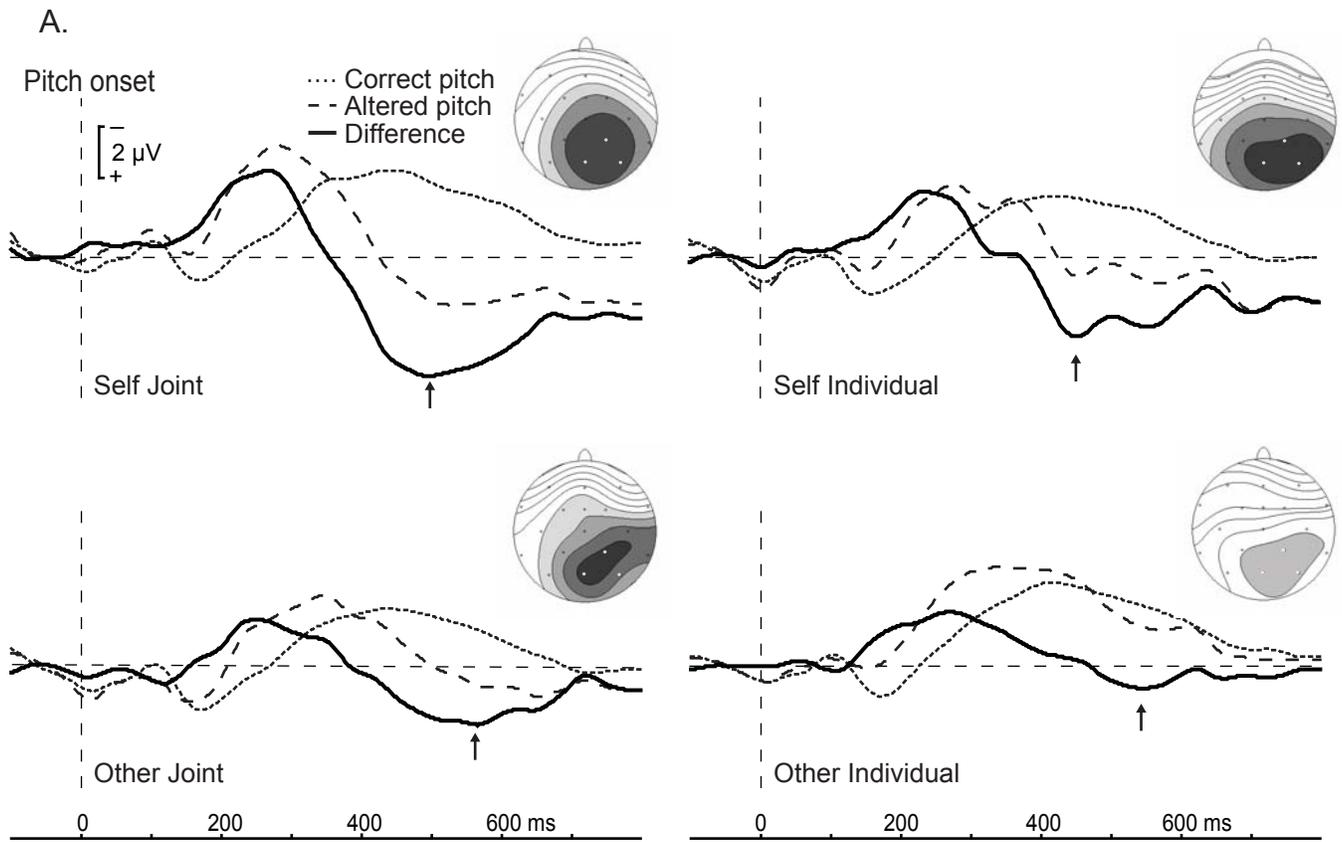


Figure 3