Shared credit for shared success: Successful joint performance strengthens the sense of joint agency

Janeen D. Loehr[a]

[a]Department of Psychology, University of Saskatchewan, 9 Campus Drive, Saskatoon, Saskatchewan, Canada, S7N 5A5. E-mail: janeen.loehr@usask.ca.

Highlights

- Partners performed a joint action together and rated their feelings of joint agency
- More successful joint performance elicited stronger feelings of joint agency
- External cues to joint success enhanced the effect of joint performance on agency
- People derive joint agency from success at the level of the dyad
- Cues to joint agency may be weighted according to their salience in a given context

Abstract

When people perform joint actions together, they experience a sense of joint agency, or shared control over actions and their effects. The current study examined how internal and external cues related to the success of a joint action influence joint agency. In three experiments, partners coordinated their actions to produce eight-tone sequences that matched a metronome pace. Across experiments, more successful joint performance (closer match to required pace) elicited stronger feelings of joint agency. This relationship was evident whether participants rated their control over sequence timing or responsibility for task outcome. Furthermore, the relationship was stronger when participants received external cues to joint success compared to when participants could rely only on internal cues. These findings indicate that people derive their sense of joint agency from success at the level of the dyad and that cues to joint agency may be weighted according to their salience in a given context.

Key words: agency; joint agency; joint action; interpersonal coordination; shared control; responsibility attribution; cue integration

1. Introduction

When people perform actions alone, they typically experience a sense of self-agency, that is, a sense that they control their actions and the effects their actions have on the world (Haggard & Tsakiris, 2009). A wealth of research has established that numerous cues contribute to the sense of self-agency (e.g., Chambon, Sidarus, & Haggard, 2014; Moore & Fletcher, 2012; Synofzik, Vosgerau, & Voss, 2013), including internal cues arising from motor control processes and external, situational cues such as the context of the action. Current theoretical frameworks posit that self-agency results from a combination of cues that are weighted according to their salience and reliability in a given context (Moore & Fletcher, 2012; Synofzik et al., 2013). In contrast, much less is known about people's experiences of agency when they engage in joint action, that is, when they coordinate their actions with others to achieve a shared goal (Sebanz, Bekkering, & Knoblich, 2006). Philosophers have proposed that engaging in joint action may uniquely elicit a sense of joint agency, or shared control over actions and their consequences (Pacherie, 2012; Seemann, 2009). Recent work has established that people do experience joint agency when they coordinate their actions with a partner (Bolt, Poncelet, Schultz, & Loehr, 2016) and that the sense of joint agency depends in part on the predictability of one's own and others' actions (Bolt & Loehr, 2017). A critical next step is to elucidate the cues that contribute to the sense of joint agency and how these cues are weighted in a joint action context. The current study investigated how internal and external cues related to successful performance of a joint action influence people's experience of joint agency.

The sense of agency is comprised of multiple aspects, including a sense of being the cause of actions, of having control over actions as they unfold, and of being responsible for actions' outcomes (Moore, 2016; Pacherie, 2008). The current study focused on people's experiences of agency over continuous actions (that is, sequences of actions that unfold over time). Examples of continuous actions include driving a car from one location to another or performing a musical piece from start to finish. Actions such as these are organized in a hierarchical structure of goals and outcomes. For example, Pacherie (2008) distinguishes between the distal level (cognitive representations of the overarching goal of the task, e.g., driving to a destination; performing a musical piece), the proximal level (actions needed to achieve the distal goal, e.g., turning the steering wheel to the right; pressing a piano key), and the sensorimotor level. Goals at each level have associated outcomes, e.g., at the distal level, arriving at the destination or performing the piece as intended. Joint actions are particularly interesting because the distal goal is shared between co-actors. For example, when two people perform a piano duet, they share the distal goal of producing the musical piece (Loehr & Vesper, 2016; for philosophical discussions of shared goals, see Butterfill, 2012; Pacherie, 2012; for implementation of shared goals in hierarchical computational frameworks for joint action, see Keller, Novembre, & Loehr, 2016; Pesquita, Whitwell, & Enns, 2017). The current study examined how achieving a shared distal goal influences the sense of joint agency.

How and why might achieving a shared distal goal influence people's experience of joint agency? In the self-agency literature, theoretical accounts and empirical evidence have established that self-agency relies in part on retrospective comparisons between goals and outcomes at each level of the action hierarchy, with better matches eliciting stronger experiences of self-agency (see Pacherie, 2012, for a review). At the distal level, successful task performance indicates that the distal goal has been achieved; in other words, more successful task performance indicates a better match between the distal goal and its outcome (Sidarus, Vuorre, & Haggard, 2017). Accordingly, in the domain of solo action, greater success at achieving an individual distal goal elicits a stronger sense of self-agency. Specifically, several studies have shown that artificially boosting task performance elicits higher ratings of self-agency (e.g., Inoue, Takeda, & Kimura, 2017; Metcalfe & Greene, 2007; Wen, Yamashita, & Asama, 2015). For example, Wen et al. (2015) showed that people reported a stronger sense of self-agency when their performance was enhanced because a computer 'assisted' them by ignoring their erroneous keystrokes.

In the domain of joint action, one study has likewise shown that greater success at achieving a shared distal goal strengthens people's experience of self-agency. Dewey, Pacherie, and Knoblich (2014) had partners coordinate their actions to achieve the shared goal of keeping a cursor centred on a moving target. Each partner controlled a joystick that moved the cursor in only the left or the right direction, respectively. Participants reported stronger self-agency when their joint task performance was enhanced because both partners' joysticks elicited cursor movement, compared to when joint task performance was impaired because only the participant's joystick elicited cursor movement. Interestingly, people's ratings of their *partner's* agency were also enhanced when both partners could move the cursor. Thus, participants felt stronger self- *and* other-agency under conditions that elicited better joint task performance. This is consistent with the idea that people engaged in joint action operate in a so-called "we-mode" (Gallotti & Frith, 2013), in which individual actions are represented as contributions to the group's overall goal (Tsai, Sebanz, & Knoblich, 2011). Accordingly, people may have experienced a sense of joint agency ("we are in control") driven by success at the level of the group rather than the individual (Dewey et al., 2014).

Several studies have also assessed trial-level correlations between joint task performance and agency ratings (Bolt & Loehr, 2017; Dewey et al., 2014; van der Wel, 2015; van der Wel, Sebanz, & Knoblich, 2012). Although better joint performance has been consistently associated with enhanced agency ratings across experiments, the correlations were not always strong. For example, van der Wel et al. (2012) reported correlations that were significantly greater than zero in only two of four joint task conditions and were weaker than in a solo version of the task. Similarly, Dewey et al. (2014) reported that performance-agency correlations did not always remain significant after controlling for visuomotor coupling. However, the agency rating scales used in the aforementioned studies reflected self-agency or may have been ambiguous regarding whether they referred to self- or joint agency (Dewey et al., 2014). Consequently, Bolt et al. (2016) developed a rating scale that asks participants specifically about their experiences of joint agency. Using this rating scale, Bolt & Loehr (2017) examined the relationship between joint task performance and joint agency. In this study, participants produced tone sequences together with a partner to achieve the shared goal of matching a metronome pace. Better joint performance (a closer match to the metronome pace) elicited stronger ratings of joint agency. However, this study used a confederate design in which the timing of the partner's tones was controlled by a computer and drawn from a pre-determined distribution of inter-tap intervals. Thus, although participants believed they were coordinating with a partner, the task did not involve mutual adaptation between partners (that is, participants adapted to the partner's timing but not vice versa). Accordingly, the association between joint task performance and joint agency was driven primarily by the participant's task performance.

In sum, existing findings suggest that achieving a shared distal goal may strengthen people's experience of joint agency. However, this has not yet been tested directly, as existing studies have either examined joint actions that require mutual adaptation (Dewey et al., 2014; van der Wel, 2015; van der Wel et al., 2012) or measured people's experience of joint agency

(Bolt & Loehr, 2017), but not both. The first goal of the current research was therefore to directly test the hypothesis that more successful joint task performance would elicit stronger feelings of joint agency. This hypothesis stemmed from previous work suggesting that people operate in the we-mode during cooperative joint action and derive feelings of joint agency based on the success of the group as a whole (Dewey et al., 2014). An alternative possibility was that people's experience of agency during joint action could be driven by egocentric processes, such as the well-known self-serving bias (Mezulis, Abramson, Hyde, & Hankin, 2004) whereby people ascribe success to themselves and failure to other causes. In this case, better joint task performance would strengthen people's experience of individual agency rather than joint agency.

Most of the literature reviewed in the preceding paragraphs examined the influence of implicit cues to task performance on agency. For example, in several studies (Inoue et al., 2017; van der Wel, 2015; Wen et al., 2015), participants' goal was to move an object toward a target as quickly as possible. Participants saw the object reach the target and presumably had an implicit sense of how quickly the target was reached, but they did not receive explicit feedback about their speed. Thus, these studies show that implicit performance cues are sufficient to influence agency. One study (Metcalfe & Greene, 2007) showed that enhancing the salience of action outcomes related to the distal goal (i.e., adding auditory feedback to hits and misses in a target-catching task) strengthened the link between task performance and self-agency. This is consistent with cue integration models of self-agency, which posit that cues are weighted according to their salience or reliability in a given context (Moore & Fletcher, 2012; Synofzik et al., 2013). However, the impact of cue salience on people's experience of joint agency has yet to be examined. The second goal of the current research was to test whether providing explicit feedback about joint task performance would increase its salience as a cue to joint agency and therefore enhance its effect on joint agency.

2. Experiment 1

Participants in Experiment 1 were recruited in pairs and performed the same task as in Bolt and Loehr (2017): They produced sequences of eight tones in alternation with each other with the shared goal of matching the pace set by a metronome (see Figure 1). Previous research using this task has confirmed that when pairs of participants perform this task together (i.e., when neither is a confederate), they mutually adapt to each other's actions (Bolt et al., 2016). Half of the pairs were assigned to the *explicit-cue* group, in which participants received explicit feedback about how well they matched the metronome pace *before* they provided ratings of joint agency. The remaining pairs were assigned to the *implicit-cue* group, in which participants received feedback about how well they matched the metronome pace *after* they provided joint agency ratings. Thus, participants in the implicit-cue group could rely only on implicit cues to joint performance when making their agency ratings, but they still received performance feedback to ensure that they would realize similar performance improvements over the course of the experiment as the explicit-cue group. Performance feedback took the form of a vertical bar indicating the size of the difference between the metronome pace and the pair's actual pace, referred to as the pair's *ITI error*. Participants in both groups were expected to experience stronger feelings of joint agency as ITI error decreased (reflecting better joint task performance). In other words, a positive relationship was expected between ratings of joint agency (where smaller values indicate stronger feelings of joint agency) and ITI error (where smaller values indicate better joint performance). Furthermore, the relationship between joint agency and ITI error was expected to be stronger (more positive) in the explicit-cue group compared to the implicit-cue group. The term “ITI error” is used in place of “joint task performance” in the Methods and Results for ease of exposition.

2.1 Method

2.1.1 Design

Participants alternated their actions to produce sequences of eight tones that matched a metronome pace, as shown in Figure 1. Experiment 1 used a multi-level design with one categorical predictor variable and one continuous predictor variable. The categorical predictor variable, cue group, was manipulated at the pair level: Half of the pairs received explicit feedback about the joint performance before providing agency ratings (explicit-cue group) and half of the pairs did not receive feedback about the joint performance until after providing agency ratings (implicit-cue group). The continuous predictor variable, ITI error, was measured at the pair level on each trial (i.e., the ITI error was the same for both partners in the pair on a given trial). The outcome variable, joint agency, was measured at the participant level on each trial (i.e., each partner in a pair provided a separate rating of their feeling of joint agency). Each pair performed 40 trials.

2.1.2 Sample size estimation based on analysis of existing datasets

Before carrying out Experiment 1, a power analysis was conducted to determine the required sample size. There are at least two challenges in conducting prospective power analyses for multi-level designs: accommodating the structure of the specific multi-level model of interest, and generating estimated values for each of the model’s parameters (Lane & Hennes, 2018). A useful approach for meeting these challenges is to conduct power simulations using parameter estimates derived from analyses of existing datasets (Lane & Hennes, 2018). This approach was implemented here, as described in detail in Section S.1 of the Supplementary Material. Briefly, two existing datasets were analyzed to estimate the required parameter values, and then SIMR (Green & Macleod, 2016; Green, MacLeod, & Alday, 2017) was used to estimate power at a range of sample sizes using Monte Carlo simulation. The power analysis indicated that a sample size of 100 participants (25 pairs in each cue group) was needed to achieve power ≥ 80% for testing the hypothesis that the effect of ITI error would differ between cue groups.

2.1.3 Participants

One hundred adults (78 female, 21 male, 1 declined to indicate gender; mean age = 20.46, *SD* = 3.46) participated in the study in pairs. Participants were recruited without regard for the gender composition of the pairs; 29 of the 50 pairs consisted of two females, one of the pairs consisted of two males, 19 pairs were mixed-gender, and one pair’s gender composition is unknown. Participants provided written informed consent according to procedures reviewed by the Behavioural Research Ethics Board at the University of Saskatchewan. Participants were compensated with course credit.

2.1.4 Procedure

Partners sat on the same side of a table with a computer screen centered between them approximately 60 cm from the table edge. An Interlink force-sensitive resister (FSR; 3.81cm^2) was placed in front of each partner approximately 20 cm from the table edge. Each participant tapped their FSR with the index finger of their dominant hand. The FSR registered taps without providing auditory feedback. Each FSR was connected to an Arduino micro-controller, which signalled PsychoPy software (Peirce, 2007) when a tap was registered. PsychoPy recorded each tap and presented an 880 Hz tone with a tap-to-tone latency of approximately 12 ms. PsychoPy also presented the remaining stimuli, including the 1000 Hz metronome tones. Tones were presented through speakers on both sides of the screen. Number keypads were placed beside each FSR and covered with occluders so that participants could enter their agency ratings but could not see their partner's ratings.

The experiment began with two trials during which the experimenter (a research assistant) explained the task. Partners then performed 2 blocks of 4 training trials and 8 blocks of 5 test trials. One partner was the leader (produced the first sequence tone) for all trials in a block. Partners alternated between leader and follower across blocks (including training). Whether the participant seated on the left or right was the leader on the first block was counterbalanced across participants. At the beginning of each block, instructions presented onscreen indicated which partner was to produce the first tone.

The sequence of events that comprised a trial is shown in Figure 1. Each trial began with a 1500 ms visual cue consisting of a cartoon face with two arms, one of which was colored red to remind participants which partner would produce the first tone. A fixation cross then appeared and remained in the center of the screen until the last sequence tone was produced. Four metronome tones were presented at 500 ms intervals beginning 500 ms after fixation onset. Partners were instructed to alternate their actions to produce an 8-tone sequence while maintaining the metronome pace. They were instructed to produce the first sequence tone "where the fifth [metronome] tone would be if there was one." After the last sequence tone was produced, a black screen was presented for 700 ms. In the explicit-cue group, the black screen was followed by feedback indicating how well the pair's performance matched the metronome pace. The feedback, which is described in the next paragraph, was presented for 700 ms, after which both partners provided their agency ratings. Last, a 700 ms black screen was presented before the next trial began. In the implicit-cue group, participants made their agency ratings first, followed by the same sequence of 700 ms black screen, feedback presented for 700 ms, and 700 ms black screen before the next trial began. In Experiment 1, participants were asked to rate their "feelings of control over the timing of the sequence" on a scale from 1 (labeled "shared control") to 99 ("independent control"). Participants were asked about their experience of control over the timing of the tones because they were expected to experience strong and constant self-agency over the tones themselves. Scale endpoints were chosen based on philosophical definitions of joint agency (see Bolt et al., 2016, for more details). Participants were instructed that "The rating scale ranges from having a shared sense of control where you worked together to control the timing, to having an independent sense of control where you worked separately to control the timing of the sequence." Zero was included as the first digit for ratings ≤ 9 to prevent partners from guessing each other's ratings based on number of keystrokes. Partners entered their ratings in random order, determined separately for each trial and signalled by which side of the screen the rating instructions appeared on first. At the end of the experiment, participants filled out a demographics questionnaire and were debriefed.

An example feedback screen is shown in Figure 1. The feedback consisted of a thin horizontal line that denoted perfect performance (i.e., a mean ITI of 500 ms) plus a vertical bar presented above or below the horizontal line that represented the pair's actual performance (i.e., the pair's ITI error, defined as the mean ITI over all eight sequence tones [beginning with the interval from the last metronome tone to the first sequence tone] minus the required ITI of 500 ms). The horizontal line was approximately 11 cm wide and was presented in the centre of the screen. The vertical bar was approximately 3 cm wide and was centered on the horizontal line. The vertical bar extended above the horizontal line for positive ITI errors (indicating that the pair tapped slower than the required ITI) or below the horizontal line for negative ITI errors (indicating that the pair tapped faster than the required ITI). The height of the vertical bar was scaled to the size of the ITI error, such that errors ≥ 75 ms resulted in a bar of maximum height (reaching near the top or bottom of the screen, respectively), and errors < 75 ms resulted in a bar that was proportional (e.g., a 35 ms error would yield a bar half the maximum height). The maximum of 75 ms was chosen based on the analyses of two previous datasets that is described in the Supplementary Material. In these datasets, 95% of pairs' ITI errors fell between 0 and 65 ms. Thus, nearly all of the pairs' ITI errors were expected to fall within 75 ms.

2.1.5 Data analysis

2.1.5.1 Sequence and rating errors

Trials were removed from analysis if they contained a sequence production error (participants produced their tones in the wrong order) or a rating error (participants entered an invalid rating of <1 or >99). In total, 28 trials (1.40%) were removed due to sequence production errors and an additional 35 trials (1.75%) were removed due to rating errors. This left an average of 38.74 agency ratings per participant.

2.1.5.2 Continuous predictor variable: ITI error

Although participants received signed feedback about their ITI errors (that is, negative or positive values for sequences that were faster or slower than the required pace, respectively), they were expected to experience stronger joint agency as ITI errors decreased regardless of whether they performed too fast or too slow. Absolute ITI error was therefore used as the continuous predictor variable, and the term ITI error therefore refers to absolute values. The ITI error distributions were checked for outliers, separately for each cue group, using the same criteria as in the analyses of previous datasets (ITI error > 250 ms, falling well beyond the rest of the distribution of ITI errors; see section S.1.3.2 in the Supplementary Material). There were no such outliers in either cue group. As in the analyses of previous datasets, the range of ITI errors varied across pairs. The average of the mean ITI error within each pair was 20.29 ms (*SD* = 6.78 ms) in the explicit-cue group and 24.28 ms (*SD* = 13.59 ms) in the implicit-cue group. Across all pairs, 97.72% and 95.15% of the ITI errors were ≤ 75 ms (corresponding to the maximum height of the vertical feedback bar) in the explicit-cue and implicit-cue groups, respectively.

2.1.5.3 Joint agency analysis

The goals for statistical inference were to test three specific hypotheses: 1) that participants would report stronger joint agency as ITI error decreased in the explicit-cue group; 2) that participants would report stronger joint agency as ITI error decreased in the implicit-cue group; and 3) that the relationship between joint agency and ITI error would be stronger in the explicit-cue group compared to the implicit-cue group. Multilevel linear mixed-effects model analyses were used to test these hypotheses. Multilevel mixed-effects models include fixed effects for the predictor variables of interest (i.e., cue group and ITI error) and can also include random effects at each level of the design (i.e., at the pair, participant, and trial levels in the current experiment). However, including all possible random effects in the model risks overfitting the data (Bates, Kliegl, Vasishth, & Baayen, 2015). The planned analysis strategy was therefore to begin with the maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013) and then to iteratively refine the random effects to find the structure best supported by the data (Bates et al., 2015). Because the data for the two cue groups came from different sets of participants, the structure best supported by the data could differ between groups. (Indeed, this was the case in the analyses of previous datasets reported in the Supplementary Material). The joint agency data were therefore analyzed in two steps.

First, hypotheses 1 and 2, which concerned the effect of ITI error on agency within each cue group, were tested separately for each cue group using the random effect structure specific to that cue group. Thus, the analysis for each cue group began with a model that included the fixed effect of ITI error and a maximal random effects structure consisting of an intercept and slope for ITI error at the pair level, an intercept and slope for ITI error at the participant level, and an intercept for trial. The random effects were refined as follows, in these and all subsequent analyses. First, if the model fitting procedure failed to converge, random effects whose covariance was estimated as zero were removed. Next, in an iterative process, the random effect that accounted for the least variance was removed and the goodness of fit was assessed to determine whether it had been significantly reduced. This procedure allowed random effects not supported by the data to be removed (Bates et al., 2015). Finally, correlation parameters were added for the remaining variance components and goodness of fit was assessed to determine whether it had significantly improved (Bates et al., 2015). Changes in goodness of fit were assessed by comparing the estimated deviances (-2 log-likelihood) using a likelihood ratio test. The final model for the explicit-cue group included all of the random effects as well as covariance between random effects at the pair level. The final model for the implicit-cue group included all of the random effects except the slope of ITI error at the pair level, as well as covariance between random effects at the participant level. The *t*-test against zero was used to test the effect of ITI error on joint agency within each cue group.

Second, hypothesis 3, which concerned the difference in the effect of ITI error between the two cue groups, was tested using a model that included fixed effects of cue group, ITI error, and their interaction as well as all of the random effects that had been retained in either analysis of the two cue groups. The final model for the between-groups comparison therefore included all of the random effects and covariances between random effects at both the pair and participant levels.[1] The *t*-test comparing the effect of ITI error between cue groups (equivalent to the *F* test for the interaction between the fixed effects of cue group and ITI error) was used to test the between-group difference. All model fits were estimated using restricted maximum likelihood via the MIXED command in SPSS Version 24. Degrees of freedom were obtained by Satterthwaite approximation.

2.2 Results

The left panel of Figure 2 shows the estimated mean joint agency ratings by ITI error for each cue group in Experiment 1. The first two rows of Table 1 show the estimated raw (unstandardized) coefficients (*b*) for the effect of ITI error on joint agency within each cue group, along with their 95% CIs, the standardized coefficients (b^*) as a measure of effect size, and *t*-tests for the effects of ITI error within each group. As predicted, participants in both cue groups reported significantly more shared control as ITI error decreased. The unstandardized *b* coefficients denote the effect of a 1-ms change in ITI error on joint agency scores. Thus, for participants in the explicit-cue group, agency ratings became more shared by 4 points for every 10-ms decrease in ITI error. For participants in the implicit-cue group, agency ratings became more shared by 1.3 points for every 10-ms decrease in ITI error. The third row of Table 1 shows the estimated between-group difference in the effect of ITI error on joint agency, along with its 95% CI, the standardized coefficient, and the *t*-test for the between-group difference. As predicted, the effect of ITI error on joint agency was significantly more positive for the explicit-cue group compared to the implicit-cue group.[2]

2.3 Discussion

Experiment 1 examined whether people experience stronger feelings of joint agency when they are more successful at achieving the shared goal of a joint action. Experiment 1 yielded three main findings. First, as predicted, people experienced a stronger sense of shared (as opposed to independent) control over the timing of a jointly-produced sequence when their joint performance more closely matched the shared goal. Second, the relationship between joint task performance and joint agency was evident both when people were provided with explicit feedback about their task success and when they could rely only on implicit cues to task success. Together, these findings provide direct support for the hypothesis that people derive feelings of *joint* agency based on the success of the group as a whole, consistent with Dewey et al.'s (2014) findings that people experience stronger self- *and* other-control over more successful joint actions. Furthermore, these findings indicate that people attribute successful joint performance to the dyad rather than to themselves, contrary to what would be expected if people have a self-serving bias to attribute positive outcomes to themselves (e.g., Mezulis et al., 2004). This possibility is examined further in Experiment 2. Third, Experiment 1 showed that increasing the salience of joint performance by providing people with explicit feedback increased the impact of joint performance on people's sense of joint agency. This finding adds to previous research indicating that increasing the salience of action outcomes related to a distal goal increases the impact of task performance on self-agency (Metcalfe & Greene, 2007). This finding also provides evidence that mechanisms of cue integration, which have been proposed to underlie self-agency, might also underlie people's experience of joint agency.

3. Experiment 2

Experiment 1 demonstrated that more successful joint performance elicits stronger feelings of shared *control over the timing* of a joint action. Experiment 2 examined whether more successful joint performance also elicits stronger feelings of shared *responsibility for the outcome* of a joint action. Examining the link between successful joint performance and feelings of shared responsibility served three purposes. First, the sense of agency includes both control over one's actions and responsibility for their effects (Moore, 2016; Pacherie, 2008). Accordingly, self-agency has been measured by asking participants about their feelings of both control over actions (e.g., Metcalfe & Greene, 2007; Wen et al., 2015) and responsibility for action outcomes (e.g., Obhi & Hall, 2011; Sato, 2009). In contrast, joint agency has yet to be examined in terms of responsibility for a joint action outcome. Second, in order for participants to rate their responsibility for a known joint outcome, they were provided with explicit feedback about the joint performance before they made their ratings. Thus, Experiment 2 served as a replication of the effect of explicit performance cues on joint agency using a different rating scale. Finally, asking people to rate their feelings of responsibility provided a more direct test of whether people attribute successful joint performance to the dyad rather than to either individual within the dyad.

The rating scale used in Experiment 2 ranged from "I was responsible for the outcome" through "We were equally responsible for the outcome" to "My partner was responsible for the outcome." The mid-point of "We were equally responsible" captured feelings of shared responsibility for the joint outcome. Endpoints of "I" and "My partner" were included so that participants could indicate who was more responsible for the outcome if they felt that responsibility was not equally shared. Based on the findings from Experiment 1, participants were expected to experience a stronger sense of shared responsibility for more successful joint performance. In other words, participants were expected to attribute more successful joint performance to the dyad, providing ratings near the middle of the scale, and attribute less successful joint performance to one or the other partner, providing ratings closer to either end of the scale. The statistical analysis plan was therefore to analyze the scale collapsed across ratings of "I" and "My Partner" (see Section 3.1.4.3 for details).[3]

3.1 Method

The methods were the same as for Experiment 1, with the following exceptions.

3.1.1 Design

Pairs of participants performed the same sequence production task as in Experiment 1. All pairs received feedback about their joint performance before providing joint agency ratings (equivalent to the explicit-cue group in Experiment 1).

3.1.2 Participants

The same number of pairs was recruited for Experiment 2 as for a single group in Experiment 1 (i.e., 25 pairs). Thus, fifty adults participated in the study (34 female, 16 male; mean age = 20.54, SD = 2.74). Thirteen of the pairs consisted of two females, four of the pairs consisted of two males, and eight pairs were mixed-gender.

3.1.3 Procedure

Experiment 2 used the same procedure as Experiment 1 with one exception: Both partners rated their "feelings of responsibility for how well you matched the pace" on a scale from 1 (labeled "I was responsible for the outcome") to 50 ("We were equally responsible for the outcome") to 99 ("My partner was responsible for the outcome"). Participants were instructed that "the rating scale ranges from feeling that you were completely responsible for the outcome, through feeling that you and your partner were equally responsible for the outcome, through feeling that your partner was completely responsible for the outcome."

3.1.4 Data analysis

3.1.4.1 Sequence and rating errors

In total, 13 trials (1.30%) were removed due to sequence production errors and an additional 28 trials (2.80%) were removed due to rating errors. This left an average of 38.36 agency ratings per participant.

3.1.4.2 Continuous predictor variable: ITI error

One trial met the ITI error outlier criteria used in Experiment 1 and was removed from analysis. The average of the mean ITI error within each pair was 21.14 ms (SD = 10.36 ms). Across all pairs, 96.77% of the ITI errors were ≤ 75 ms.

3.1.4.3 Responsibility rating scale

The responsibility rating scale was collapsed so that participants' ratings fell along a continuum from equal responsibility (corresponding to original ratings of "We were equally responsible") to individual responsibility (corresponding to original ratings of either "I" or "My partner" was responsible). First, 50 was subtracted from the scores on the original scale, thus shifting the original 1 to 99 scale to a scale that ranged from −49 to +49. Next, the absolute value was taken, so that a value of zero corresponded to the "We were equally responsible" rating and values greater than zero corresponded to ratings that one partner or the other was more responsible (to a maximum value of 49). Last, 1 was added to each score so that the scores ranged from 1-50. Thus, the collapsed responsibility scale was similar to the control scale used in Experiment 1, with smaller values indicating stronger joint agency (here, shared responsibility) and larger values indicating weaker joint agency (here, individual responsibility).

3.1.4.4 Joint agency analysis

The collapsed responsibility scale was analyzed following the same strategy as in Experiment 1. The analysis began with a model that included a fixed effect of ITI error and a maximal random effects structure comprised of an intercept and slope for ITI error at the pair level, an intercept and slope for ITI error at the participant level, and an intercept for trial. After iteratively refining the random effects, the final model included all random effects but no covariances between random effects at either the pair or participant level. The *t*-test comparing the effect of ITI error against zero is reported.

3.2 Results

The solid line in the right panel of Figure 2 shows the estimated mean joint agency ratings by ITI error for Experiment 2 along the collapsed responsibility scale shown on the left ordinate axis. The middle row of Table 1 shows the estimated unstandardized coefficient for the effect of ITI error on joint agency in Experiment 2, along with its 95% CI, the standardized coefficient as a measure of effect size, and the *t*-test comparing the effect of ITI error against zero. As predicted, participants reported significantly more shared responsibility as ITI error decreased: Agency ratings became more shared by 3.5 points for every 10-ms decrease in ITI error.

3.3 Discussion

Experiment 2 showed that, as predicted, people experienced a stronger sense of shared responsibility for outcomes that more closely matched the shared goal of a joint action. This finding demonstrates that successful joint task performance elicits not only stronger feelings of shared control over the continuous joint action, but also stronger feelings of shared responsibility for the outcome of the joint action. Experiment 2 also provides direct evidence that people attribute responsibility for joint success as equally shared between members of the pair, rather than attributing responsibility to themselves or their partner. This finding provides further evidence that a self-serving bias has minimal impact on people's experiences of joint agency during joint action. Finally, Experiment 2 replicated the effects of explicit performance cues found in Experiment 1 using a different rating scale. In sum, Experiment 2 further supports the conclusion that people use external performance cues to derive feelings of joint agency based on success at the level of the group.

4. Experiment 3

Experiment 1 examined the effects of implicit and explicit performance cues on people's experiences of joint agency, and Experiment 2 further examined the effects of explicit cues. The purpose of Experiment 3 was to further examine the effects of implicit cues. Specifically, Experiment 3 tested whether people's experience of joint agency would be influenced by implicit performance cues in the absence of any feedback about the joint performance. In Experiment 1, participants in the implicit-cue group were provided with feedback about the joint performance on each trial, after they made their joint agency ratings. In Experiment 3, participants were not provided with any feedback about the joint performance at all. Participants were expected to experience stronger joint agency for more successful joint task performance even in the absence of feedback. Alternatively, if people in the implicit-cue group in Experiment 1 relied on implicit performance cues only because the salience of joint performance was enhanced by the presence of feedback on each trial, no effect of joint success would be expected in Experiment 3.

4.1 Method

The methods were the same as for Experiment 1, with the following exceptions.

4.1.1 Design

Pairs of participants performed the same sequence production task as in Experiment 1, except that no feedback about the joint performance was presented at any time.

4.1.2 Participants

The same number of pairs was recruited for Experiment 3 as for a single group in Experiment 1 (i.e., 25 pairs). Thus, fifty adults participated in the study (37 female, 13 male; mean age = 20.04, *SD* = 3.85). Twelve of the pairs consisted of two females, one of the pairs consisted of two males, and eleven pairs were mixed-gender.

4.1.3 Procedure

Experiment 3 used the same procedure as in Experiment 1 with the exception that no feedback about the joint performance was presented. Thus, after the last sequence tone was produced, a black screen was presented for 700 ms, participants made their agency ratings, and a black screen was presented for 700 ms before the next trial began. Participants rated their "feelings of control over the timing of the sequence" on the same scale used in Experiment 1.

4.1.4 Data analysis

4.1.4.1 Sequence and rating errors

In total, 10 trials (1.00%) were removed due to sequence production errors and an additional 11 trials (1.10%) were removed due to rating errors. This left an average of 39.16 agency ratings per participant.

4.1.4.2 Continuous predictor variable: ITI error

Two trials met the ITI error outlier criteria used in Experiments 1 and 2 and were removed from analysis. The average of the mean ITI error within each pair was 23.56 ms (*SD* = 19.62 ms). Across all pairs, 93.26% of the ITI errors were ≤ 75 ms.

4.1.4.3 Joint agency analysis

The joint agency data from Experiment 3 were analyzed following the same strategy as in Experiment 1. The analysis began with a model that included a fixed effect of ITI error and a maximal random effects structure comprised of an intercept and slope for ITI error at the pair level, an intercept and slope for ITI error at the participant level, and an intercept for trial. After iteratively refining the random effects, the final model included an intercept and slope for ITI error at the participant level and an intercept for trial. The *t*-test comparing the effect of ITI error against zero is reported.

4.2 Results

The dashed line in the right panel of Figure 2 shows the estimated mean joint agency ratings by ITI error for Experiment 3 along the control rating scale shown on the right ordinate axis. The last row of Table 1 shows the estimated unstandardized coefficient for the effect of ITI error on joint agency in Experiment 3, along with its 95% CI, the standardized coefficient for effect size, and the *t*-test comparing the effect of ITI error against zero. As predicted, participants reported significantly more shared control as ITI error decreased. Participants' agency ratings became more shared by 1.9 points for every 10-ms decrease in ITI error. This effect is similar in size to the effect of ITI error in the implicit-cue group in Experiment 1, in which agency ratings became more shared by 1.3 points for every 10-ms decrease in ITI error.

4.3 Discussion

Experiment 3 examined whether people use implicit performance cues to derive their sense of joint agency in the absence of any feedback about joint task performance. As in Experiments 1 and 2, in Experiment 3 people experienced stronger feelings of joint agency for performances that more closely matched the shared goal of the joint action. Thus, Experiment 3 replicated the findings from the implicit-cue group in Experiment 1. Moreover, Experiment 3 showed that people rely on implicit cues even in the absence of any external cues that enhance the salience of joint performance. Thus, Experiment 3 provides further evidence that people derive feelings of joint agency based on the success of the group as a whole.

5. General Discussion

The current study investigated how internal and external cues to joint performance influence people's experience of joint agency during a continuous joint action. In three experiments, pairs of participants coordinated a sequence of actions to achieve the shared goal of matching the pace set by a metronome. Across all three experiments, people reported feeling stronger joint agency for joint actions that were more successful, i.e., that more closely matched the pair's shared goal. Joint performance influenced joint agency whether people were provided with external cues to joint performance (Experiments 1 and 2) or could rely on only internal cues (Experiments 1 and 3). Enhancing the salience of joint performance via external cues increased the impact of joint performance on people's experience of joint agency. Furthermore, joint performance cues influenced people's feelings of both shared control over the timing of the sequence and shared responsibility for the outcome of the sequence. Together, these findings demonstrate that people's experiences of joint agency are strengthened by success at the level of the group. These findings complement previous research showing that implicit cues to the success of solo action enhance feelings of self-agency (e.g., Inoue et al., 2017; Metcalfe & Greene, 2007; Wen et al., 2015) and provide direct support for Dewey et al.'s (2014) conclusion that more successful joint performance elicits stronger feelings of shared agency.

The effects of implicit performance cues on joint agency were relatively modest in the current study. People's ratings of agency became more shared by 1.3 and 1.9 points for every 10-ms decrease in ITI error in the implicit-cue group in Experiment 1 and Experiment 3, respectively. This represents a change of approximately 10-14 points on a 99-point scale over the range of typical ITI error values (from near zero to approximately 75 ms). One possible explanation for the relatively modest effect of implicit performance cues may be the nature of the implicit cues themselves. Reinforcement learning accounts posit that in the absence of external feedback, learning may be guided by internal estimates of confidence, which may be relatively noisy, particularly early in learning (e.g., Guggenmos & Sterzer, 2017; Guggenmos, Wilbertz, Hebart, & Sterzer, 2016). Notably, these modest effects occurred in an experimental context in which there were few other competing cues to joint agency. For example, participants always performed the same task with the same partner, so situational cues to agency remained constant across all trials. Furthermore, participants performed the task under normal performance conditions, that is, in the absence of noise or unusual task conditions (e.g., unexpected computer assistance, Wen et al., 2015; or impairment of one partner's contribution, Dewey et al., 2014). The findings reported here therefore establish an important baseline for future research examining the impact of implicit performance cues relative to other types of cues and under noisier or more ambiguous conditions.

In contrast, explicit performance cues had a relatively strong effect on joint agency. For example, in Experiment 1 people's ratings of joint agency became more shared by 4.0 points for every 10-ms decrease in ITI error, which translates to a change of approximately 30 points on the 99-point agency scale over the range of typical ITI errors. The current study is the first to directly compare the impact of implicit and explicit cues to joint performance on people's experience of joint agency. Cue-integration models of self-agency (Moore & Fletcher, 2012; Synofzik et al., 2013) predict that the more salient a given cue, the greater impact it should have on self-agency. The current study provides evidence that this prediction holds true for joint agency. However, further research is necessary to firmly establish that joint agency is based on the same cue-integration principles thought to underlie self-agency. Given that the current study established the impact of veridical feedback on joint agency, a next step could be to examine the effect of reducing the reliability of explicit performance cues by providing participants with false feedback about the joint performance. Furthermore, in the current study, enhancing the salience of joint performance in a subtler way did not increase its impact on joint agency: Providing explicit feedback *after* joint agency ratings (Experiment 1, implicit-cue group) did not increase the impact of joint performance relative to providing no feedback at all (Experiment 3). Thus, there may be limits to the degree to which the salience of joint performance cues impacts joint agency.

The current study demonstrated that joint performance influenced people's feelings of both shared control over the timing of the sequence and shared responsibility for the outcome of the sequence. This finding adds to previous research on joint agency that examined people's experiences of shared control (Bolt & Loehr, 2017; Bolt et al., 2016) and aligns with research on self-agency, in which different components of agency are assessed, including control over actions and responsibility for action outcomes (Moore, 2016). The precise relationship between these components, such as the degree to which control and responsibility are correlated and whether different cues impact control and responsibility in different ways, remains an important avenue for future research (see, e.g., Frith, 2014). The current findings further demonstrate that people report feeling a sense of joint agency regardless of the alternative provided by the other endpoint(s) of the scale, that is, whether the alternative was a feeling of independence (meaning the partners worked separately rather than together; Experiment 1) or a feeling that one individual or the other had more impact on the outcome of the joint action (Experiment 2). Thus, the current study shows that joint performance cues influence the *type* of agency people experience (i.e., the strength of joint agency relative to independent/individual agency). Previous research examining agency during joint action has examined the *degree* to which people experience agency (e.g., more vs. less control), which may reflect joint or individual agency or both (Dewey et al., 2014). Another important avenue for future research is to establish whether agency type and degree vary independently, as hypothesized by Pacherie (2012).

Last, measuring people's feelings of shared responsibility in Experiment 2 provided direct evidence that people attribute responsibility for successful joint performance to the dyad rather than either individual within the dyad. Thus, self-serving attributions may have a minimal impact on joint agency. Indeed, previous research has established that group contexts sometimes elicit group-serving attributions, that is, attributions of responsibility for positive outcomes to the group rather than individuals within the group (e.g., Taylor & Doria, 1981). Most research on group-serving attributions has examined group contexts that are very different from the joint action employed here. For example, typical paradigms have had participants perform tasks individually, aggregated the results at the group level, and then asked participants to rate their feelings of responsibility for the aggregated group performance. In many cases the members of the 'group' do not interact at all (Rantilla, 2000). Thus, the current study adds to previous work by demonstrating group-serving attributions for a joint action that involved live interaction and real-time, temporally precise interpersonal coordination.

Two additional questions for future research concern the relationship between explicit ratings of joint agency and a) implicit measures of agency and b) judgments of performance. Explicit measures of self-agency, similar to the explicit measures employed in the current study, capture high-level, conceptual judgments of self-agency. In contrast, implicit measures such as intentional binding (Moore & Obhi, 2012) and sensory attenuation (Gentsch & Schütz-Bosbach, 2011) are thought to capture lower-level, pre-reflective or non-conceptual feelings of self-agency (Synofzik, Vosgerau, & Newen, 2008). An intriguing possibility is that people may experience joint agency at the pre-reflective level during joint action (Obhi & Hall, 2011). This possibility is supported by findings of comparable intentional binding and sensory attenuation for actions produced by self and other in interactive contexts (e.g., Obhi & Hall, 2011; Strother, House, & Obhi, 2010). Whether and how implicit measures of agency relate to explicit measures of joint agency remains an exciting avenue for future research. Judgments of performance are an important component of metacognition about action (Metcalfe, Eich, & Miele, 2013; Metcalfe & Greene, 2007) and are closely related to judgments of agency. Previous research has demonstrated dissociations between judgments of performance and judgments of agency in both solo action (Metcalfe et al., 2013; Metcalfe & Greene, 2007) and joint action contexts (van der Wel, 2015). However, the relationship between judgments of joint performance and judgments of joint agency remains to be tested directly.

In sum, the current research demonstrates that people feel stronger joint agency over more successful joint actions, particularly when they receive external cues indicating joint success. Thus, people experience a sense of joint agency when they successfully achieve a shared distal goal. Potential implications of these findings include the following. First, people's experiences of joint versus independent agency are likely to influence how they adapt their future behaviour after more or less successful joint action. Researchers have just begun to explore how engaging in joint action impacts error-monitoring signals that are related to adapting future behaviour (Li et al., 2010; Loehr, Kourtis, & Brazil, 2015). People's subjective experiences of agency may impact both their behavioural and neural responses to jointly-produced errors. Second, providing cues that strengthen people's experiences of joint agency may be useful in contexts that require interpersonal affiliation and group cohesion (e.g., Carron, Shapcott, & Burke, 2007; Overy & Molnar-Szakacs, 2009). The current findings suggest that providing explicit feedback about joint success may be one way to enhance joint agency in such contexts. Finally, the current study aligns with rising interest in how affective and motivational factors influence the sense of self-agency (e.g., Gentsch & Synofzik, 2014; Gentsch, Weiss, Spengler, Synofzik, & Schütz-Bosbach, 2015). The current study provides initial evidence concerning how such factors influence the sense of joint agency by showing that positive outcomes of a joint action elicit stronger feelings of joint agency.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models, 1–27. https://doi.org/arXiv:1506.04967

Bolt, N. K., & Loehr, J. D. (2017). The predictability of a partner's actions modulates the sense of joint agency. *Cognition*, *161*, 60–65. https://doi.org/10.1016/j.cognition.2017.01.004

Bolt, N. K., Poncelet, E. M., Schultz, B. G., & Loehr, J. D. (2016). Mutual coordination strengthens the sense of joint agency in cooperative joint action. *Consciousness and Cognition*, *46*, 173–187. https://doi.org/10.1016/j.concog.2016.10.001

Butterfill, S. (2012). Joint action and development. *The Philosophical Quarterly*, *62*(246), 23–47. https://doi.org/10.1111/j.1467-9213.2011.00005.x

Carron, A. V., Shapcott, K. M., & Burke, S. M. (2007). Group cohesion in sport and exercise. In M. R. Beauchamp & M. A. Eys (Eds.), *Group dynamics in exercise and sport psychology: Contemporary themes* (pp. 117–139). New York, NY: Routledge.

Chambon, V., Sidarus, N., & Haggard, P. (2014). From action intentions to action effects: How does the sense of agency come about? *Frontiers in Human Neuroscience*, *8*(May), 1–9. https://doi.org/10.3389/fnhum.2014.00320

Dewey, J. A., Pacherie, E., & Knoblich, G. (2014). The phenomenology of controlling a moving object with another person. *Cognition*, *132*(3), 383–397. https://doi.org/10.1016/j.cognition.2014.05.002

Frith, C. D. (2014). Action, agency and responsibility. *Neuropsychologia, 55*, 137-142.

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences*, *17*(4), 160–165. https://doi.org/10.1016/j.tics.2013.02.002

Gentsch, A., & Schütz-Bosbach, S. (2011). I did it: Unconscious expectation of sensory consequences modulates the experience of self-agency and Its functional signature. *Journal of Cognitive Neuroscience*, *23*(12), 3817–3828.

Gentsch, A., & Synofzik, M. (2014). Affective coding: The emotional dimension of agency. *Frontiers in Human Neuroscience*, *8*, 608. https://doi.org/10.3389/fnhum.2014.00608

Gentsch, A., Weiss, C., Spengler, S., Synofzik, M., & Schütz-Bosbach, S. (2015). Doing good or bad: How interactions between action and emotion expectations shape the sense of agency. *Social Neuroscience*, *10*(4), 418–430. https://doi.org/10.1080/17470919.2015.1006374

Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Green, P., MacLeod, C. J., & Alday, P. (2017). Package "simr." Retrieved from https://cran.r-project.org/web/packages/simr/simr.pdf

Guggenmos, M., & Sterzer, P. (2017). A confidence-based reinforcement learning model for perceptual learning. *BioRxiv.* https://doi.org/10.1101/136903

Guggenmos, M., Wilbertz, G., Hebart, M. N., & Sterzer, P. (2016). Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *ELife*, *5*(e13388), 1–19. https://doi.org/10.7554/eLife.13388.001

Haggard, P., & Tsakiris, M. (2009). The experience of agency: Feelings, judgments, and responsibility. *Current Directions in Psychological Science*, *18*(4), 242–246.

Inoue, K., Takeda, Y., & Kimura, M. (2017). Sense of agency in continuous action: Assistanceinduced performance improvement is self-attributed even with knowledge of assistance. *Consciousness and Cognition*, *48*, 246–252. https://doi.org/10.1016/j.concog.2016.12.003

Keller, P. E., Novembre, G., & Loehr, J. D. (2016). Musical ensemble performance: Representing self, other, and joint action outcomes. In S. S. Obhi & E. S. Cross (Eds.), *Shared representations: Sensorimotor foundations of social life* (pp. 280–310). Cambridge, UK: Cambridge University Press.

Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationships research. *Journal of Social and Personal Relationships*, *35*(1), 7–31. https://doi.org/10.1177/0265407517710342

Li, P., Jia, S., Feng, T., Liu, Q., Suo, T., & Li, H. (2010). The influence of the diffusion of responsibility effect on outcome evaluations: Electrophysiological evidence from an ERP study. *NeuroImage*, *52*, 1727–1733. https://doi.org/10.1016/j.neuroimage.2010.04.275

Loehr, J. D., Kourtis, D., & Brazil, I. A. (2015). It's not just my fault: Neural correlates of feedback processing in solo and joint action. *Biological Psychology*, *111*, 1–7. https://doi.org/10.1016/j.biopsycho.2015.08.004

Loehr, J. D., & Vesper, C. (2016). The sound of you and me: Novices represent shared goals in joint action. *Quarterly Journal of Experimental Psychology*, *69*(3), 535–547. https://doi.org/10.1080/17470218.2015.1061029

Metcalfe, J., Eich, T. S., & Miele, D. B. (2013). Metacognition of agency: Proximal action and distal outcome. *Experimental Brain Research*, *229*, 485–496. https://doi.org/10.1007/s00221-012-3371-6

Metcalfe, J., & Greene, M. J. (2007). Metacognition of agency. *Journal of Experimental Psychology: General*, *136*(2), 184–199. https://doi.org/10.1037/0096-3445.136.2.184

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*(5), 711–747. https://doi.org/10.1037/0033-2909.130.5.711

Moore, J. W. (2016). What is the sense of agency and why does it matter? *Frontiers in Psychology*, *7*, 1272. https://doi.org/10.3389/fpsyg.2016.01272

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, *21*, 59–68. https://doi.org/10.1016/j.concog.2011.08.010

Moore, J. W., & Obhi, S. S. (2012). Intentional binding and the sense of agency: A review. *Consciousness and Cognition*, *21*, 546–561. https://doi.org/10.1016/j.concog.2011.12.002

Obhi, S. S., & Hall, P. (2011). Sense of agency and intentional binding in joint action. *Experimental Brain Research*, *211*(3–4), 655–662. https://doi.org/10.1007/s00221-011-2675-2

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception*, *26*(5), 489–504. https://doi.org/10.1525/mp.2009.26.5.489

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, *107*, 179–217. https://doi.org/10.1016/j.cognition.2007.09.003

Pacherie, E. (2012). The phenomenology of joint action: Self-agency vs. joint-agency. In A. Seemann (Ed.), *Joint attention: New developments* (pp. 343–389). Cambridge, MA: MIT Press.

Pesquita, A., Whitwell, R. L., & Enns, J. T. (2017). Predictive joint-action model: A hierarchical predictive approach to human cooperation. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-017-1393-6

Rantilla, A. K. (2000). Collective task responsibility allocation: Revisiting the group-serving bias. *Small Group Research*, *31*(6), 739–763.

Sato, A. (2009). Both motor prediction and conceptual congruency between preview and action-effect contribute to explicit judgment of agency. *Cognition*, *110*, 74–83. https://doi.org/10.1016/j.cognition.2008.10.011

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Seemann, A. (2009). Joint agency: Intersubjectivity, sense of control, and the feeling of trust. *Inquiry*, *52*(5), 500–515. https://doi.org/10.1080/00201740903302634

Sidarus, N., Vuorre, M., & Haggard, P. (2017). Integrating prospective and retrospective cues to the sense of agency: A multi-study investigation. *Neuroscience of Consciousness*, (1), 1–10. https://doi.org/10.1093/nc/nix012

Strother, L., House, K. A., & Obhi, S. S. (2010). Subjective agency and awareness of shared actions. *Consciousness and Cognition*, *19*(1), 12–20. https://doi.org/10.1016/j.concog.2009.12.007

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), 219–239. https://doi.org/10.1016/j.concog.2007.03.010

Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: An interplay between prediction and postdiction. *Frontiers in Psychology*, *4*(127), 1–8. https://doi.org/10.3389/fpsyg.2013.00127

Taylor, D. M., & Doria, J. (1981). Self-serving and group serving bias in attribution. *The Journal of Social Psychology*, *113*, 201–211.

Tsai, J. C.-C., Sebanz, N., & Knoblich, G. (2011). The GROOP effect: Groups mimic group actions. *Cognition*, *118*, 135–140. https://doi.org/10.1016/j.cognition.2010.10.007

van der Wel, R. P. R. D. (2015). Me and we: Metacognition and performance evaluation of joint actions. *Cognition*, *140*, 49–59. https://doi.org/10.1016/j.cognition.2015.03.011

van der Wel, R. P. R. D., Sebanz, N., & Knoblich, G. (2012). The sense of agency during skill learning in individuals and dyads. *Consciousness and Cognition*, *21*(3), 1267–1279. https://doi.org/10.1016/j.concog.2012.04.001

Wen, W., Yamashita, A., & Asama, H. (2015). The sense of agency during continuous action: Performance is more important than action-feedback association. *PLoS ONE*, *10*(4), e0125226. https://doi.org/10.1371/journal.pone.0125226

Footnotes

[1]The same random effects structure would be retained if the analysis began with a model that included fixed effects of cue group, ITI error, and their interaction, as well as the maximal random effects structure, and iteratively refined it as described in the main text. Estimates of variance and covariance for the final model in each analysis in Experiments 1-3 are presented in Section S.2 of the Supplementary Material.

[2]In a previous experiment, effects of ITI error measured at the pair level were driven by the participant's ITI errors but not their partner's (when the partner was a confederate whose tones were actually produced by a computer; Bolt & Loehr, 2017). A secondary analysis was therefore conducted to confirm that both the participant's mean ITI error (calculated as the absolute difference between the participant's ITIs, from preceding tone to participant's tap, and the required 500 ms ITI) and the partner's mean ITI error (calculated likewise, from preceding tone to partner's tap) influenced joint agency in Experiment 1. This analysis began with an empty model that included a fixed effect of group and random intercepts at the pair, participant, and trial levels. Changes in model fit were assessed after each of the following steps, in the order presented, using the likelihood ratio test. Adding a fixed effect of participant ITI error along with its random slope at the participant level improved model fit. Removing the random slope for participant ITI error reduced model fit, so the random slope was retained. Adding a fixed effect for the group by participant ITI error interaction did not significantly improve model fit, although the effect of participant ITI error was numerically larger in the explicit-cue condition compared to the implicit-cue condition. Adding a fixed effect of partner ITI error along with its random slope at the participant level improved model fit. Removing the random slope for partner ITI error reduced model fit, so the random slope was retained. Adding a fixed effect for the group by partner ITI error interaction did not significantly improve model fit, although the effect of partner ITI error was numerically larger in the explicit-cue condition compared to the implicit-cue condition. Thus, the secondary analysis confirmed that both the participant's and the partner's ITI errors contributed to the effect of pair-level ITI error on joint agency. All analyses reported in the main text therefore examined the effects of pair-level ITI error on joint agency.

[3]If the results showed the opposite pattern, with ratings closer to the individual ends of the scale for successful joint performance and closer to the middle for less successful joint performance, one could then examine whether people were more likely to attribute success to themselves rather than their partner, as would be expected if a self-serving bias was at work.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of Canada [grant number 2108299] and the Canada Foundation for Innovation [grant number 32476]. The author thanks Lauren Slavik and Juan Pablo Del Castillo Cabada for assistance with data collection.

Table 1. Estimated effect of ITI error on joint agency within and between experimental groups.

Expt	Group	*b*	95% CI	b^*	*t*	*df*	*p*
1	Explicit-cue[a]	0.40	[0.19, 0.62]	0.32	3.85	23.09	.001
	Implicit-cue[a]	0.13	[0.02, 0.24]	0.12	2.39	44.03	.021
	Difference[b]	0.26	[0.02, 0.50]	0.22	2.18	44.55	.034
2	Explicit-cue[c]	0.35	[0.24, 0.46]	0.51	6.61	16.29	<.001
3	Implicit-cue (no feedback)[a]	0.19	[0.11, 0.27]	0.22	4.70	40.08	< .001

Note. Expt = Experiment.

[a]Joint agency rated on a scale from 1 = shared to 99 = independent control.

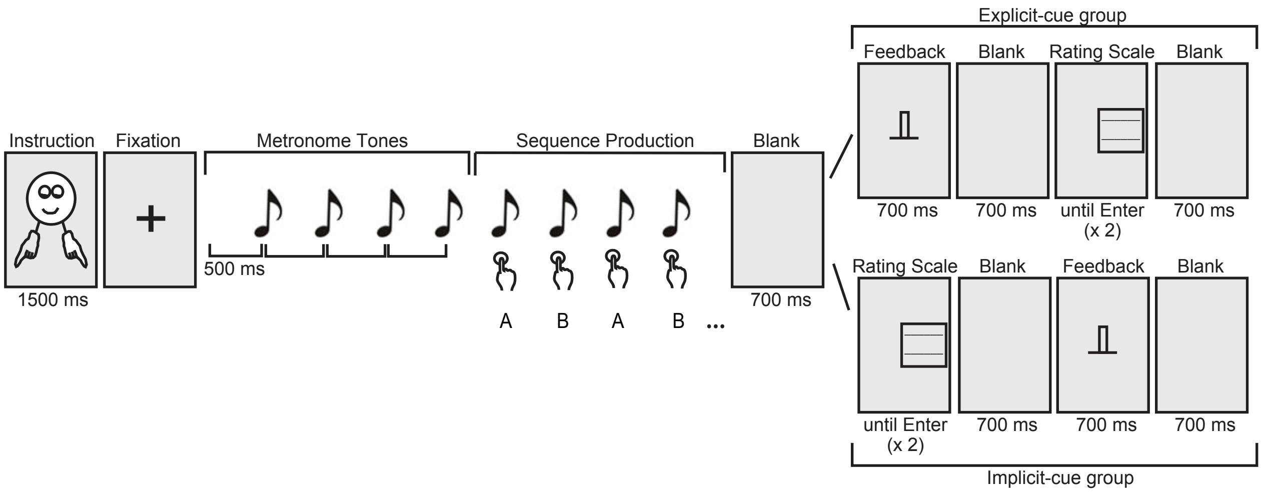
[b]Difference calculated as Expt 1 explicit-cue group minus Expt 1 implicit-cue group.

[c]Joint agency rated on a scale from 1 = equal to 50 = individual responsibility.

Figure Captions

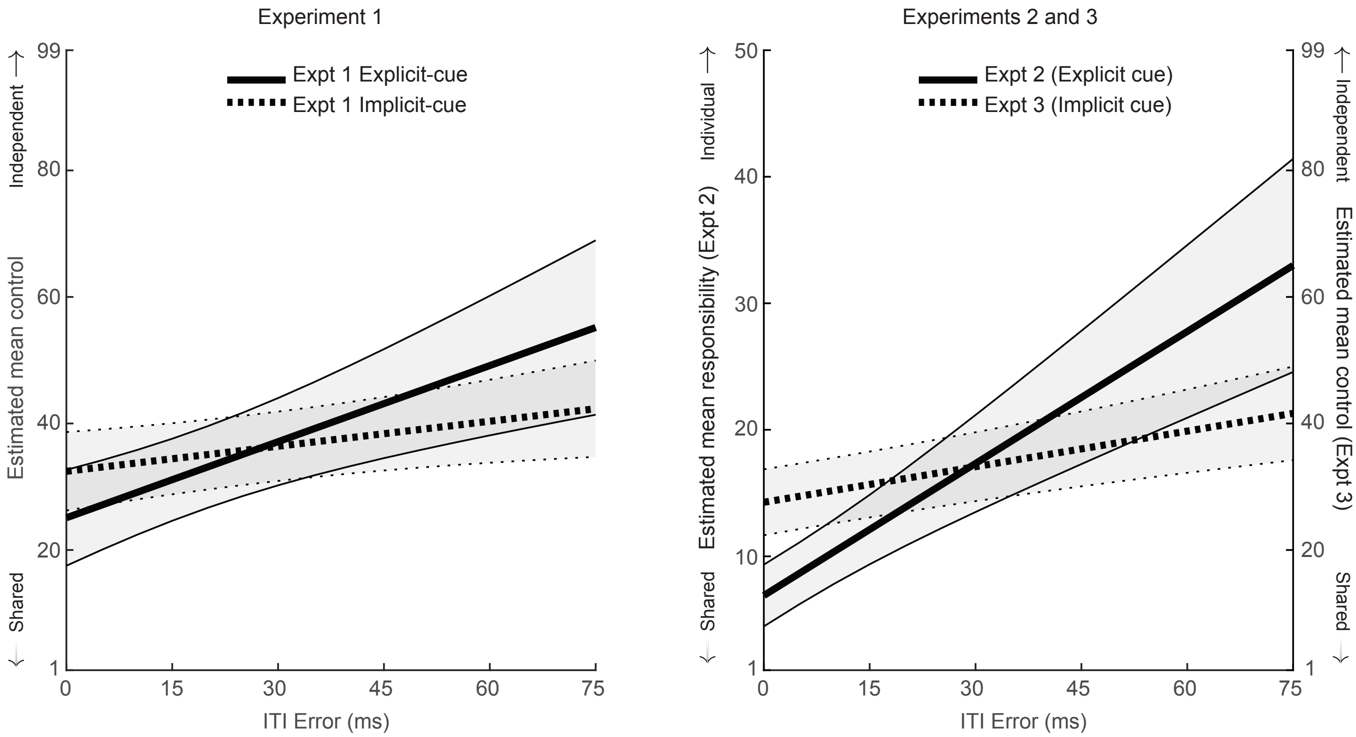
Figure 1. Schematic illustration of the sequence production task used in Experiment 1. Following instructions and fixation, participants heard a metronome (i.e., a series of isochronous pacing tones, illustrated by eighth note symbols) and then alternated their actions to produce a sequence of eight tones (illustrated by combined button press and eighth note symbols, labeled A and B for the two partners). After producing the last tone, participants either received feedback about their performance and then each made an agency rating (explicit-cue group) or each made an agency rating and then received feedback about their performance (implicit-cue group).

Figure 2. Left: Estimated mean (± 95% CI) ratings of joint agency as a function of ITI error in the explicit-cue and implicit-cue groups in Experiment 1. Right: Estimated mean (± 95% CI) ratings of joint agency as a function of ITI error in Experiments 2 and 3. The left ordinate axis shows the collapsed responsibility rating scale used in Experiment 2 and the right ordinate axis shows the control rating scale used in Experiment 3.



Explicit-cue group
Feedback
Blank
Rating Scale
Blank
Instruction
Fixation
Metronome Tones
Sequence Production
Blank
700 ms
700 ms
until Enter (x 2)
700 ms
500 ms
A
B
A
B
...
1500 ms
700 ms
Rating Scale
Blank
Feedback
Blank
until Enter (x 2)
700 ms
700 ms
700 ms
Implicit-cue group

Figure 1



Experiment 1
Expt 1 Explicit-cue
Expt 1 Implicit-cue
Independent →
← Shared
Estimated mean control
99
80
60
40
20
1
0
15
30
45
60
75
ITI Error (ms)
Experiments 2 and 3
Expt 2 (Explicit cue)
Expt 3 (Implicit cue)
Individual →
← Shared
Estimated mean responsibility (Expt 2)
50
40
30
20
10
1
Estimated mean control (Expt 3)
Independent →
← Shared
99
80
60
40
20
1
0
15
30
45
60
75
ITI Error (ms)

Figure 2

Supplementary Material

S.1 Prospective power analysis based on existing datasets

S.1.1 Overview

This section of the Supplementary Material describes the power analysis that was conducted to determine the necessary sample size for Experiment 1. First, the planned statistical analysis for Experiment 1 is briefly described, beginning with the key statistical test upon which the power analysis was based. Second, the analysis of two existing datasets is described, from which estimates for each of the parameter values required for the power analysis were derived. Finally, details of the power analysis, carried out using SIMR (Green & Macleod, 2016; Green, MacLeod, & Alday, 2017), are provided.

S.1.2 Key statistical test and planned statistical analysis

Experiment 1 was designed to test two major hypotheses: First, that participants in both the explicit-cue and implicit-cue groups would show a positive relationship between joint agency and ITI error, and second, that the relationship would be stronger in the explicit-cue group compared to the implicit-cue group. The power analysis was based on the second hypothesis. That is, the power analysis was conducted with the goal of ensuring that Experiment 1 had sufficient power to detect an interaction between cue group and ITI error.

Because Experiment 1 used a multilevel design, as described in Section 2.1.1 in the main text, the statistical analysis plan was to conduct a multilevel linear mixed-effects model analysis, as described in detail in Section 2.1.5.3 in the main text. To summarize briefly, the design included a categorical predictor variable at the pair level (cue group), a continuous predictor variable at the pair level (ITI error), and an outcome measure at the participant level (joint agency ratings). The mixed-effects model therefore included fixed effects for cue group, ITI error, and their interaction. The maximal random effects structure, which was to be iteratively refined to find the structure best supported by the data (Bates, Kliegl, Vasishth, & Baayen, 2015), consisted of a random intercept and slope for ITI error at the pair level, a random intercept and slope for ITI error at the participant level, and a random intercept at the trial level.

The key statistical test for the power analysis was the *F* test for the interaction between the fixed effects of cue group and ITI error. More specifically, H_1 was that the effect of ITI error would be larger (more positive) in the explicit-cue group compared to the implicit-cue group. H_0 was that there would be no difference in the size of the ITI error effect between the cue groups.

S.1.3 Estimates of parameter values based on analyses of existing datasets

S.1.3.1 Required parameter values

One of the challenges in performing prospective power analyses for a multilevel design is that there are a large number of parameter values that need to be estimated (Lane & Hennes, 2018). For Experiment 1, the following parameter values needed to be estimated: the minimum effect size of interest for the key hypothesis (i.e., the size of the interaction effect); the sizes of the other fixed effects; each of the random effects at the pair, participant, and trial levels; and the likely values of the continuous predictor variable, ITI error. One recommended approach to estimating a large number of parameter values is to draw on previously conducted studies using similar measures or designs (Lane & Hennes, 2018). Following this recommendation, two existing datasets were selected from our lab that most closely matched the measures and design of the planned experiment (Dataset 1: Bolt, Poncelet, Schultz, & Loehr, 2016, Experiment 1; Dataset 2: unpublished data). These datasets are described next.

S.1.3.2 Description of existing datasets

Datasets 1 and 2 both came from experiments in which pairs of participants performed the same alternating-tone sequence production task that was to be used in Experiment 1; provided ratings of joint agency on the same scale that was to be used in Experiment 1, after each of 40 trials; and received explicit feedback about the joint performance either after providing agency ratings (Dataset 1; similar to the implicit-cue group in Experiment 1) or before providing agency ratings (Dataset 2; similar to the explicit-cue group in Experiment 1). In both experiments, participants also performed other sequence production tasks interleaved throughout the experiment. Only data from the alternating-tone task were analyzed here.

In both datasets, each pair's ITI error (in ms) was recorded on each trial. However, participants were only presented with categorical feedback indicating whether the joint performance was "correct" (i.e., matched the metronome pace) or "incorrect." In Dataset 1, feedback was provided after participants' agency ratings simply to encourage participants to perform well on the task. A key consideration of the experimental design for Dataset 1 was that the percentage of trials labeled "incorrect" should be equal to 20% in all experimental conditions. Thus, feedback was determined based on whether the pair's ITI error fell within an accuracy window that changed across blocks depending on how well the pair performed. In Dataset 2, participants also received categorical feedback. The feedback was presented before participants' agency ratings, and the percentage of trials labeled "incorrect" was set to 40% in all experimental conditions.

Dataset 1 included 24 pairs of participants and Dataset 2 included 16 pairs. Both datasets were pre-processed to remove trials with sequence errors (participants tapped in the wrong order) and rating errors (participants entered an invalid rating of <1 or >99). One additional trial was removed from Dataset 1 because it had an ITI error > 250 ms, which fell well beyond the rest of the distribution of ITI errors in Dataset 1. Dataset 2 had no such obvious outliers.

S.1.3.3 Analysis strategy

The existing datasets were analyzed using a strategy as similar as possible to the planned analysis for Experiment 1. Each dataset contained one predictor variable, ITI error (the absolute value of the pair's ITI error, in ms), and the outcome measure, joint agency (measured on a scale from 1="shared control" to 99="independent control"). Each analysis therefore began with a fixed effect of ITI error and a maximal random effects structure that included an intercept and slope for ITI error at the pair level, an intercept and slope for ITI error at the participant level, and an intercept at the trial level. The random effects structure was iteratively refined by first removing random effects whose covariance was estimated as zero, then iteratively checking whether the goodness of fit was significantly reduced after the random effect that accounted for the least variance was removed, and finally testing whether goodness of fit improved by fitting correlation parameters for the remaining variance components (Bates et al., 2015; see Section 2.1.5.3 in the main text for more details). The final model for Dataset 1 included the fixed effect of ITI error, a random intercept at the pair level, a random intercept and slope for ITI error at the participant level, a random intercept at the trial level, and covariances between the random effects at the participant level. The final model for Dataset 2 included the fixed effect of ITI error, a random intercept and slope for ITI error at the participant level, a random intercept at the trial level, and covariances between the random effects at the participant level.

S.1.3.4 Results

S.1.3.4.1 Fixed effects

Figure S1 shows the effect of ITI error on joint agency in Datasets 1 and 2. As Figure S1 shows, people reported significantly more shared control as ITI error decreased in both datasets. Thus, the existing datasets provided preliminary support for the hypothesized positive relationship between ITI error and joint agency. The first two rows of Table S1 show the estimated intercept and slope (along with their 95% CIs) for the fixed effect of ITI error in each dataset. The *b* coefficient for the intercept denotes the estimated mean joint agency rating at an ITI error of 0 ms; the *b* coefficient for the slope of ITI error denotes the effect of a 1-ms change in ITI error. Thus, the slope $b = 0.12$ points/ms in Dataset 1 indicates that ratings of joint agency became more shared by 1.2 points for every 10-ms decrease in ITI error when participants received categorical feedback about the performance *after* providing their agency ratings (similar to the implicit-cue group in Experiment 1). The slope *b* = 0.18 points/ms in Dataset 2 indicates that ratings of joint agency became more shared by 1.8 points for every 10-ms decrease in ITI error when participants received categorical feedback about performance *before* providing their agency ratings (similar to the explicit-cue group in Experiment 1). Thus, the existing datasets also provided preliminary support for the hypothesis that the relationship between ITI error and joint agency would be stronger when participants received explicit feedback about their performance before making their joint agency ratings.

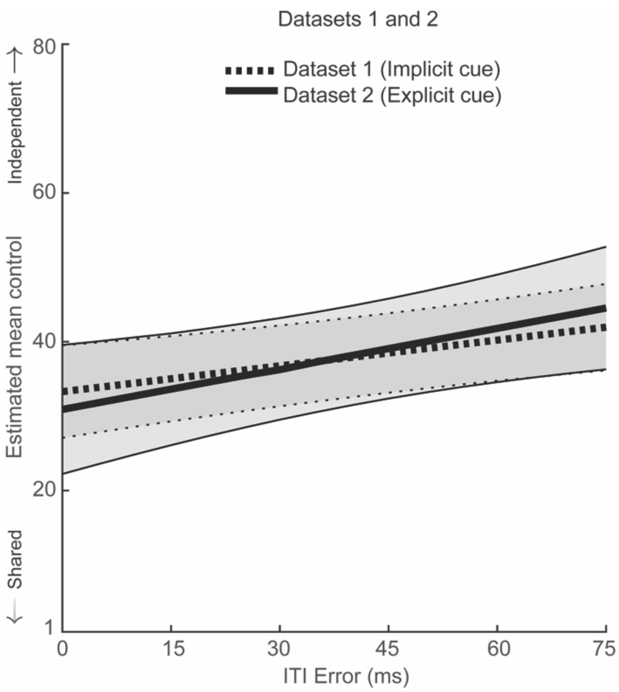


Figure S1. Estimated mean ratings of joint agency (± 95% CI) as a function of ITI error in Datasets 1 and 2.

Table S1. Estimated intercepts and slopes (along with their 95% CIs) for the effect of ITI error on joint agency in Datasets 1 and 2.

	Intercept		Slope	
Dataset	*b*	95% CI	*b*	95% CI
1 (similar to implicit-cue group)	33.30	[27.08, 39.52]	0.12	[0.04, 0.19]
2 (similar to explicit-cue group)	30.90	[22.23, 39.57]	0.18	[0.04, 0.32]
Difference[a]	-2.4		0.06	

[a]Difference is calculated as Dataset 2 minus Dataset 1.

S.1.3.4.2 Random effects

Table S2 shows the variance estimates for each of the random effects that remained in the final model for each dataset after iteratively refining the random effects structure. Table S2 also shows the percentage of the total variance accounted for by each random effect. As Table S2 shows, most of the variance occurred around the intercept at the participant level.

Table S2. Variance estimates for Datasets 1 and 2.

Parameter (level)	Dataset 1 Estimate (% of total variance)	Dataset 2 Estimate (% of total variance)
Residual	288.99 (42.75%)	344.71 (36.58%)
Intercept (Pair)	38.44 (5.69%)	-
Intercept (Participant)	344.73 (50.99%)	548.88 (58.25%)
ITI error (Participant)	0.04 (0.006%)	0.11 (0.01%)
Intercept (Trial)	3.84 (0.57%)	48.65 (5.16%)

S.1.3.4.3 Continuous predictor variable: ITI error

Visual inspection of each pair's ITI error distribution in each dataset revealed that the range of ITI errors varied across pairs, such that the mean ITI error within a pair ranged from 9.78 ms to 49.30 ms (M = 24.32 ms, SD = 12.05 ms) in Dataset 1 and from 15.07 ms to 35.54 ms (M = 22.51 ms, SD = 5.93 ms) in Dataset 2. Across all pairs in a given dataset, 95% of the ITI errors fell below 63.71 ms in Dataset 1 and 64.69 ms in Dataset 2.

S.1.3.5 Parameter estimates for the power analysis

The results from the analyses reported above were used to estimate the parameter values needed for the power analysis. The fixed effects were considered first. The full model for Experiment 1 included fixed effects for cue group, ITI error, and their interaction. In the full model, the *b* coefficient for the intercept and the *b* coefficient for the effect of ITI error represent the intercept and slope in one of the two cue groups (termed the reference group). The implicit-cue group was considered the reference group for the purpose of the power analysis. Values from the analysis of Dataset 1 (first row of Table S1) were therefore used as estimates for the intercept (b = 33) and slope of ITI error (b = 0.12) for the implicit-cue/reference group. The *b* coefficient for the effect of cue group represents the difference in intercepts between the two cue groups, and the *b* coefficient for the cue group by ITI error interaction represents the difference in slopes between the two cue groups. Difference values from the comparison between Datasets 1 and 2 (third row of Table S1) were therefore used as initial estimates of the effect of cue group (b = 2.4) and the cue group by ITI error interaction (b = 0.06). However, the difference in slopes between the implicit-cue and explicit-cue groups in Experiment 1 was expected to be larger than the difference in slopes between the existing datasets, because in Experiment 1 explicit performance feedback would be the only difference between the two groups. The power analysis was therefore conducted using a minimum effect size of b = 0.09 (1.5 times the size of the difference in the existing datasets) for the critical cue group by ITI error interaction.

The random effects were considered next. The analyses of Datasets 1 and 2 showed that random effects at the pair and trial levels accounted for very little of the total variance (6.25% of the total variance in Dataset 1 and 5.16% of the total variance in Dataset 2). A simplified random effects structure, with only a random intercept at the participant level, was therefore used for the power analysis. This amounts to estimating zero variance at the pair and trial levels, effectively treating the observations from the two participants within each pair on each trial as independent. This could slightly overestimate the prospective power if a larger percentage of the variance occurred at the pair or trial levels in Experiment 1. To balance out the potential for overestimation, relatively conservative estimates were used for the variance of the random effect, the residual variance, and the likely values of the continuous predictor variable. For the estimates of variance, the larger variance estimates from Dataset 2 were used rather than the smaller estimates from Dataset 1. For the continuous predictor variable, ITI error, the likely range of values was estimated to be 0 ms to 65 ms in equally spaced increments across the 40 trials. The maximum value of 65 ms was based on the 95th percentiles in

Datasets 1 and 2, and the equally spaced increments were implemented based on visual inspection of each pair's ITI error distribution in each dataset.

S.1.4 Power analysis

S.1.4.1 Power analysis procedure

The power analysis was conducted using the SIMR package (Green & Macleod, 2016; Green et al., 2017) in the R statistical platform (R Development Core Team, 2016). SIMR begins with a fitted multilevel mixed-effect model and estimates power by repeating three steps: 1) simulating a new dataset using the fitted model provided, 2) refitting the model to the simulated dataset, and 3) applying the statistical test of interest to the simulated fit. Power is calculated from the proportion of successful tests in Step 3.

The power analysis was implemented in SIMR as follows. First, the `makeLmer` function was used to build an artificial fitted mixed-effect model with the parameters described in Section S.1.3.5. Specifically, the artificial model was built by specifying a) the model structure (i.e., fixed effects of cue group, ITI error, and their interaction, plus a random intercept at the participant level); b) the estimated sizes of the fixed effects; c) the estimated random intercept variance and residual variance; and d) a data frame containing the predictor variables structured according to the design at the participant level (i.e., cue group varying between participants and ITI error varying across trials within participants; the design was specified at the participant level because zero variances were assumed for random effects at the pair and trial levels). Second, the `extend` function was used to increase the total number of participants in the artificial fitted model to 100 (50 per cue group). Last, the `powerCurve` function was used to run the 3-step power estimation procedure described in the preceding paragraph at 10 different sample sizes up to the maximum of 100. For each sample size, 1000 simulations were conducted that tested the interaction between cue group and ITI error at $\alpha = .05$.

S.1.4.2 Power analysis results

The simulation results indicated that a sample size of 100 participants would be required to achieve power ≥ 80% for the minimum effect size of $b = 0.09$ for the cue group by ITI error interaction. Specifically, the simulations estimated the power to be 82.90% (95% CI: [80.42, 85.18]) with a sample size of 100. Thus, 100 participants were recruited for Experiment 1 (25 pairs for each cue group).

S.2 Estimates of variance and covariance in Experiments 1-3.

Table S3. Variance and covariance estimates for the final model in each analysis in Experiments 1-3.

Analysis group	Expt 1			Expt 2	Expt 3
	Explicit-cue	Implicit-cue	Between-groups	Explicit-cue	Implicit-cue (no feedback)
Parameter (level)					
Residual	284.73	401.06	343.24	137.22	375.89
Intercept (Pair)	197.20	7.83	105.29	4.90	-
ITI error (Pair)	0.20	-	0.11	0.05	-
Covariance (Pair)	-4.71	-	-2.25	-	-
Intercept (Part.)	259.28	414.00	337.67	48.79	313.46
ITI error (Part.)	0.10	0.11	0.09	0.02	0.04
Covariance (Part.)	-	-3.26	-2.13	-	-
Intercept (Trial)	24.34	29.93	26.90	10.89	6.40

Note. Part. = participant.

References

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models, 1–27. https://doi.org/arXiv:1506.04967

Bolt, N. K., Poncelet, E. M., Schultz, B. G., & Loehr, J. D. (2016). Mutual coordination strengthens the sense of joint agency in cooperative joint action. *Consciousness and Cognition*, *46*, 173–187. https://doi.org/10.1016/j.concog.2016.10.001

Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Green, P., MacLeod, C. J., & Alday, P. (2017). Package "simr." Retrieved from https://cran.r-project.org/web/packages/simr/simr.pdf

Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationships research. *Journal of Social and Personal Relationships*, *35*(1), 7–31. https://doi.org/10.1177/0265407517710342

R Development Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria.