The sound of you and me: Novices represent shared goals in joint action

Janeen D. Loehr[1,2] and Cordula Vesper[1,3]

[1]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands
[2]Department of Psychology, University of Saskatchewan, Canada
[3]Department of Cognitive Science, Central European University, Hungary

This is an Author's Accepted Manuscript of an article published in the Quarterly Journal of Experimental Psychology, Vol. 69 (3), pp. 535-547, copyright Taylor & Francis, available online at 10.1080/17470218.2015.1061029.

Authors' addresses:

Janeen Loehr (corresponding author): Department of Psychology, University of Saskatchewan, 9 Campus Drive, Saskatoon, Saskatchewan, Canada, S7N 5A5. Email: janeen.loehr@usask.ca.

Cordula Vesper: Department of Cognitive Science, Central European University, Frankel Leó út 30-34, 1023 Budapest, Hungary. Email: vesperc@ceu.edu

Abstract

People performing joint actions coordinate their individual actions with each other to achieve a shared goal. The current study investigated the mental representations that are formed when people learn a new skill as part of a joint action. In a musical transfer-of-learning paradigm, piano novices first learned to perform simple melodies in the joint action context of coordinating with an accompanist to produce musical duets. Participants then performed their previously-learned actions with two types of auditory feedback: while hearing either their individual action goal (the melody) or the shared action goal (the duet). As predicted, participants made more performance errors in the individual goal condition compared to the shared goal condition. Further experimental manipulations indicated that this difference was not due to different coordination requirements in the two conditions or perceptual dissimilarities between learning and test. Together, these findings indicate that people form representations of shared goals in contexts that promote minimal representations, such as when learning a new action together with another person.

Keywords: joint action, shared action goals, transfer of learning, duet music performance, auditory-motor integration

The sound of you and me: Novices represent shared goals in joint action

Joint actions, in which two or more people coordinate their actions to achieve a shared goal (Knoblich, Butterfill, & Sebanz, 2011), abound in daily life. Examples range from exchanging money with a cashier or helping a friend move furniture to playing sports, having conversations, or playing music together. The nature of the mental representations underlying joint actions is emerging as a rich area of investigation (e.g., Butterfill, 2012; Gallotti & Frith, 2013; Pacherie, 2011). Recent theoretical work suggests that people represent the shared goal of a joint action (e.g., move a couch from one location to another; Vesper, Butterfill, Knoblich, & Sebanz, 2010). However, empirical work to date has focused almost exclusively on whether people represent each person's individual contributions to the joint action (e.g., lift one end of the couch and move to the left; Knoblich et al., 2011). Whether and how people represent shared goals therefore remains poorly understood. The current study investigated mental representations of shared goals using an experimental analog of duet piano performance, a quintessential joint action that requires precise interpersonal coordination yet maintains a clear distinction between individual goals (the pitches produced by each person's keystrokes) and the shared goal (the complementary pitches that comprise a musical duet; Keller, 2008; Loehr, Kourtis, Vesper, Sebanz, & Knoblich, 2013; Novembre & Keller, 2014).

Vesper et al. (2010) proposed an architecture for joint action that specifies the mental representations that may be formed when people engage in joint action. According to this account, when people engage in intentional joint action, they will at a minimum represent their own task and the shared goal of the joint action, including that the shared goal cannot be achieved by acting alone. In Vesper et al.'s (2010) model, the shared goal is captured by the formula "ME+X", where "ME" stands for a person's own contribution and "X" stands for the contribution not produced by the person herself. It is challenging to test the minimal representations underlying joint action because people often form rich representations that include their own and their partners' tasks (Sebanz, Knoblich, & Prinz, 2003, 2005; Tsai, Kuo, Jing, Hung, & Tzeng, 2006). However, minimal representations may be formed when a novice learns to perform a new joint action, particularly if the novice's task is cognitively demanding and leaves insufficient resources to form rich representations that include all components of the joint action (Vesper et al., 2010). This is likely to be the case when a person learns a new team sport, a complex couples dance routine, or, as implemented in the current study, when a novice with no music performance experience learns to perform musical duets for the first time.

Consider a simple piano duet, like the one shown in Figure 1, from the perspective of a novice with no piano performance experience. The novice's task is to produce a sequence of keystrokes (shown as a sequence of finger numbers at the top of the figure); the shared goal is to produce a sequence of complementary pitches that are coordinated in time. It is well established that when people perform actions alone, the actions are represented in terms of their desired effects or goals (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1997). Thus, if a novice learned to perform a melody alone, she would quickly form mental representations in which her keystrokes were integrated with the pitches they elicit (Bangert & Altenmüller, 2003). In a joint action such as a piano duet, however, the overall sound comprises not only the melody elicited by the novice's keystrokes but also the partner's accompaniment. According to the minimal account of joint action, the novice will minimally form a representation of the shared goal as the combined musical outcome resulting from her own and someone else's contribution. Thus, when the novice learns to perform the duet with her partner, her keystrokes should be represented in terms of the shared goal, i.e., the complementary pitches elicited by both partners' coordinated keystrokes, and not or to a lesser extent in terms of her individual goal, i.e., the melody elicited by her own keystrokes.

How does the representation of a joint action in which the complementary tones are produced by another person differ from the representation of an individual action in which a person's own keystrokes elicit both tones without another person's contribution? Intuitively, what distinguishes these two cases is knowing that another contribution is necessary to achieve the desired outcome. Accordingly, the minimal account of joint action also posits that the novice must realize that the shared goal can only be achieved with the support of another agent or other force (who, in the duet example, produces the accompaniment). Note that this does not presuppose that the novice must represent the other's precise task (i.e., the keystrokes and pitches that comprise the accompaniment), or represent the other as an intentional agent (i.e., represent the partner as the agent who produces the accompaniment), as long as the accompaniment is produced in some way (Vesper et al., 2010). Indeed, it seems unlikely that novices who lack musical experience would form representations of the specific keystrokes and pitches that comprise the accompaniment. However, they could be expected to represent their partner as the agent producing the accompaniment (cf. Wenke et al., 2011).

The current study adapted a transfer-of-learning paradigm previously used with experienced pianists (Palmer & Meyer, 2000) to investigate whether novices learning to perform musical duets for the first time form representations of a) the shared goal of their own and their partner's coordinated actions and b) their partner as the agent of the complementary actions. In transfer-of-learning paradigms, people's ability to transfer what they have learned in one context to other, systematically altered contexts is tested. This provides insight into the mental representations underlying performance (e.g., Milanese, Iani, & Rubichi, 2010; Palmer & Meyer, 2000). Participants in the current study first learned to perform musical duets with a partner, i.e., they learned a new action within a joint action context. Their performance was then assessed under test conditions in which their previously-learned keystrokes elicited either the complete duet (the shared goal) or only the melody (the participant's individual goal). In line with the minimal account of joint action, we expected novices to represent their actions in terms of the shared goal, resulting in poorer performance under test conditions in which only the individual goal was produced compared to test conditions in which the shared goal was produced.

We crossed the manipulation of elicited goals with a manipulation of whether participants performed their keystrokes alone or together with their partner. If novices represent the partner as the agent responsible for producing the accompaniment, then their performance should be worse when their keystrokes elicit the shared goal while performing alone (i.e., without the partner as the agent of the accompaniment) than while performing together. Alternatively, if novices represent the shared goal independently of the means by which another contribution is provided, then manipulating whether the partner produces keystrokes at test should not impact performance.

Experiment 1

In Experiment 1, participants completed the transfer-of-learning task illustrated in Figure 2. Participants first learned to perform simple piano duets consisting of a melody (produced by the participant) and an accompaniment (produced by the experimenter). Their performance was then assessed in test conditions that manipulated the goal achieved by the participants' actions (either the melody, the participant's individual goal, or the melody and accompaniment, the shared goal) and whether participants performed their actions alone or together with the experimenter. In Experiment 1a, instructions to participants emphasized accuracy during the test conditions, whereas in Experiment 1b, the instructions emphasized speed. This manipulation allowed us to examine participants' performance independently of their prioritization of speed or accuracy.

Method

Participants. Thirty-two adults (16 male; mean age = 24.75, *SD* = 4.35) participated in Experiment 1a and 32 adults (13 male; mean age = 21.47, *SD* = 2.52) participated in Experiment 1b. No participants reported prior training in piano performance. All participants provided written informed consent in accordance with the Declaration of Helsinki and as approved by the local medical ethics committee. Participation was compensated with course credit or €5.

Stimuli. Five isochronous 7-tone melodies and accompaniments were constructed for the experiment. An example is shown in Figure 1. The melodies were designed to be performed with the right hand, using the thumb and all fingers except the ring finger, which is the weakest and least biomechanically independent of the fingers (Loehr & Palmer, 2007, 2009). All melodies were composed in G major and used only the first five notes of the scale so that participants were not required to reposition their hands during performance. The accompaniments consisted of complementary pitches notated one octave lower than the melodies. The accompaniments were performed by the experimenter (a trained pianist) with the right hand. One melody was designated the practice melody, and the remaining four melodies were used in the test conditions described below. Melodies were presented to participants as sequences of numbers from 1 (thumb) to 5 (pinkie finger). The numbers were color-coded, and the corresponding keys were labeled with the appropriate number and color.

Equipment. Melodies were performed on a Yamaha EZ-200 piano keyboard. Presentation of auditory feedback and MIDI data acquisition were implemented via Max/MSP 5.1.7 software run on a Macintosh computer. Transmission time in this setup resulted in a constant delay of 20 ms between key presses and tone onsets.[1] Piano tones were generated using a piano timbre from the computer's internal sound card at the same volume for both melody and accompaniment. Tones were played via speakers placed next to the keyboard and the speaker volume was adjusted to a comfortable level for each participant.

Design. The design of Experiments 1a and 1b is illustrated in Figure 2a. Participants always learned to perform the melodies in a joint context (i.e., together with an accompaniment produced by the experimenter). Participants then performed the melodies under one of four test conditions in a 2 (elicited goal: individual vs. shared) x 2 (performance setting: alone vs. together) within-subjects design. Participants' keystrokes elicited only the melody in the *individual goal* condition; their keystrokes elicited both the melody and accompaniment in the *shared goal* condition. Participants produced their keystrokes alone in the *alone* condition or together with the experimenter's keystrokes in the *together* condition.

Procedure. Participants were first given the practice melody and the mapping between finger numbers, colors, and keys was explained. Participants then practiced the melody while the experimenter played along with them until they felt comfortable with the task. To ensure that the participant and experimenter began to play at the same time, participants indicated when they were ready to begin each performance, at which time the experimenter pressed a button on the keyboard to begin recording the trial. The participant then gave a verbal "go" signal, and the participant and experimenter performed their sequences together. This procedure was followed for every performance throughout the experiments. Once participants were comfortable with the task, they completed the *memorization* and *learning phases* described in the next paragraph with the practice melody. Finally, the general procedure for the *test phase* was explained, but not performed.

Participants completed the following procedure, illustrated in Figure 2b, for each of the four remaining melodies. First, in the *memorization phase*, participants practiced the currently assigned melody until it was memorized, at their own pace and accompanied by the experimenter. The notation was then removed and the duet was performed twice more as a memory check. Next, in the *learning phase*, participants performed the same melody 20 times, still accompanied by the experimenter. In Experiment 1a, participants were instructed that they should play as quickly as possible without making any errors; the goal of producing error-free performances was emphasized. This instruction was repeated again at the beginning of the *test phase*, during which the participants performed the same melody four times under one of the four test conditions. In Experiment 1b, participants were instructed to perform the melody at a slow rate during the memorization phase. For the learning phase, they were instructed to perform without making any errors and to gradually speed up their performances so that they could play at their maximum speed by the end of the learning phase. For the test phase, they were instructed to play at their maximum speed, and no mention of errors was made. After each trial in the test phase they were told to “play faster if you can.” The computer software indicated to the experimenter which test condition was to be performed only after the learning phase was complete. A different melody and accompaniment was used for each test condition (so each participant played all four melodies, one each per test condition). The order of the four conditions and the assignment of melodies to conditions were determined using a balanced Latin square design, resulting in 16 different counterbalancing orders, each of which was run for two participants.

Figure 2a shows the instructions given in the four test conditions, which were implemented as follows. For both individual goal conditions, each of the participant’s keystrokes elicited a single (melody) tone. For both shared goal conditions, each of the participant’s keystrokes elicited two tones (melody + accompaniment). In the alone conditions, the experimenter sat beside the participant with her hands in her lap to provide participants with visual evidence that the experimenter was indeed not playing along. Participants were told that they would perform the melody alone in the individual goal condition, and that the computer would fill in the experimenter’s tones in the shared goal condition. In the together conditions, the experimenter pressed the appropriate keys on the keyboard to produce the same sensation of accompaniment as in the learning and memorization phases. Participants were told that the experimenter’s tones would be muted in the individual goal condition, and that the experimenter would play along with them as before in the shared goal condition. In reality, the experimenter’s tones were muted, and participants’ key presses elicited both the melody and accompaniment tones in the shared goal condition. There were therefore no timing asynchronies between the melody and accompaniment tones, which differed from the memorization and learning phases that entailed natural asynchronies caused by small differences in the timing of the participant’s and experimenter’s keystrokes.[2] One participant reported noticing this difference in Experiment 1a and was excluded from the analysis. An occluder (height 35 cm) between the participant and the experimenter allowed the participant to see the experimenter’s upper body and arms but not her hand on the keyboard.

Data analysis. Participants’ performance was examined in terms of error rate, which was defined as the number of incorrect trials divided by the total number of trials and was calculated separately for the learning and test phases. Trials were considered incorrect if any of the participant's keystrokes deviated from those set out in the score. Error rates were measured at the trial level because errors often reflected a general breakdown of performance; for example, participants often stopped playing altogether after they made an error (typically after the fourth or fifth note in the sequence). Learning trials with errors in the accompaniment (1.29% and 1.01% in Experiments 1a and 1b, respectively) were excluded from analysis.[3] One participant's data were excluded from Experiment 1b because the error rate during learning was greater than two standard deviations above the mean. Four learning trials (0.16%) in Experiment 1a and one learning trial and one test trial in Experiment 1b were lost due to technical error. Error rates were analyzed with 2 (elicited goal: individual, shared) x 2 (performance setting: alone, together) ANOVAs. Effects were considered significant at the level of $\alpha < .05$.

Results

Learning phase. The mean error rate in the learning phase was 16.85% in Experiment 1a and 14.00% in Experiment 1b. Error rates produced during the learning phase did not differ across test conditions in either experiment, as expected given that participants did not know which test condition would follow learning. An ANOVA that included experiment as a variable revealed no difference in learning phase error rates between the two experiments (see the supplemental material available online for a detailed analysis).

Test phase. Figure 3 shows that as predicted, participants made more errors in the individual goal conditions than in the shared goal conditions in both experiments, $F(1, 30) = 7.20$, $MSE = 0.023$, $p = .012$, $\eta_P^2 = 0.19$, Hedges' $g_{av} = 0.54$ in Experiment 1a, $F(1, 30) = 8.04$, $MSE = 0.032$, $p = .008$, $\eta_P^2 = 0.21$, Hedges' $g_{av} = 0.50$ in Experiment 1b.[4] There was no effect of performance setting or interaction in either experiment, *F*s < 0.5, *MSE*s < .05, *p*s > .5 in Experiment 1a; *Fs* < 2.7, *MSE*s < .05, *p*s > .1 in Experiment 1b. An ANOVA that included experiment as a variable revealed that participants made more errors in Experiment 1b than in Experiment 1a, $F(1, 60) = 15.29$, $MSE = 0.075$, $p < .001$, $\eta_P^2 = .20$, Hedges' $g_s = 0.98$, consistent with Experiment 1b instructions that pushed participants to perform at their maximum speed. There were no significant interactions with experiment, *F*s < 2.0, *MSE*s < 0.05, *p*s > .1.

Discussion

After learning to produce a simple melody as part of a musical duet performed together with a partner, participants subsequently made more errors when their actions elicited only the melody (their individual goal) compared to when their actions elicited the duet (the shared goal). This finding suggests that novices who learn to perform a new task as part of a joint action form representations of their actions in terms of the shared goal of their own and their partner's combined actions. This supports the theoretical claim that people represent the shared goal of a joint action in contexts that promote minimal representations (Vesper et al., 2010). In contrast, participants' error rates in the shared goal condition did not differ depending on whether they played together with their partner at test or not. This suggests that participants may not have represented their partner as the agent responsible for producing the accompaniment, expecting only that the accompaniment would be produced in some way. In addition, the lack of effect of partner's performance at test rules out the possibility that higher error rates in the individual goal condition occurred simply because participants were required to shift from coordinating with a partner during learning to performing alone at test.

In Experiment 1, participants always heard two tones for every action they produced during learning, whereas they heard only one tone for every action they produced in the individual goal test conditions. It could therefore be argued that participants produced more errors in the individual goal conditions because of perceptual dissimilarity between the number of tones heard during learning and test, independently of whether or not the tones were learned in the context of a joint action. We conducted Experiment 2 to investigate this possibility.

Experiment 2

A new group of participants performed the same transfer-of-learning task as in Experiment 1, except that these participants did not perform their actions together with the experimenter during learning or at test. Instead, the computer filled in an accompaniment tone whenever the participant produced a melody tone during learning. If the pattern of error rates we observed in Experiment 1 occurred because of perceptual dissimilarity between the number of tones elicited by participants' actions during learning compared to test, and not because the sounds were represented as the shared goal of a joint action, then higher error rates should likewise occur when participants' actions elicit only the melody tones (equivalent to the individual goal) in Experiment 2, in which the sounds are learned in an individual action context.

Method

Participants. Thirty-two adults (7 male; mean age = 22.56, *SD* = 2.83) participated in Experiment 2. No participants reported prior training in piano performance.

Stimuli and equipment. Experiment 2 used the same stimuli and equipment as Experiment 1.

Design and procedure. Figure 4 shows the design of Experiment 2. Participants completed the same practice, memorization, and learning phases as in Experiment 1, with the exception that the experimenter did not play the accompaniment along with the participant. Instead, participants were instructed that they would hear their own playing along with a second voice filled in by the computer. Participants pressed the keys corresponding to the right-hand melody, and each key press elicited two simultaneous tones (melody and accompaniment). Thus, as in Experiment 1, participants always heard pairs of tones during learning; however, both tones were produced by the participant rather than by the participant together with an accompanist. We refer to the elicited duet as the *shared goal** to highlight its similarity to the shared goal in Experiment 1 while acknowledging that participants produced the duet alone in Experiment 2. The remaining instructions were the same as in Experiment 1b (i.e., participants were told to gradually speed up performance over learning so as to perform at maximum speed during test).

Participants performed the melodies under two test conditions in a within-subjects design. Elicited goals were manipulated so that participants heard only the melody (equivalent to the individual goal condition in Experiment 1) or both the melody and accompaniment (equivalent to the shared goal condition in Experiment 1). In the *individual goal* condition, participants' keystrokes elicited only the melody tones, and participants were told that the computer would no longer fill in the second voice. In the *shared goal** condition, participants' keystrokes elicited both melody and accompaniment tones, and participants were told that the computer would fill in a second voice as before. The experimenter sat beside the participant throughout the experiment. The order of the two conditions and the assignment of melodies to conditions were counterbalanced similarly to Experiment 1.

Data analysis. One participant's data were excluded from analysis because the error rate during learning was greater than two standard deviations above the mean. One learning trial was lost due to technical error. Error rates under individual and shared goal* conditions were compared with paired-samples *t*-tests.

Results

Learning phase. The mean error rate during the learning phase was 14.70% and did not differ between test conditions (see the supplemental material available online for a detailed analysis).

Test phase. Figure 5 shows that there was no difference in error rate between the individual and shared goal* conditions, $t(30) = 0.399$, $p = .69$.

Experiment 1b vs. Experiment 2. Error rates in the shared and individual goal conditions were compared for Experiments 1b and 2, both of which employed the same instructions emphasizing speed but differed in terms of the joint vs. individual nature of the task. We conducted a 2 (elicited goal) x 2 (experiment: 1b, 2) ANOVA on error rates during the test phases. The ANOVA confirmed a significant goal by experiment interaction, $F(1, 60) = 4.59$, $MSE = 0.044$, $p = .036$, $\eta_P^2 = 0.07$, Hedges' $g_s = 0.54$. Post-hoc *t*-tests confirmed that the difference between individual and shared goal conditions was only significant in Experiment 1b, $t(30) = -3.07$, $p = .0045$.

Discussion

In a version of the transfer-of-learning task in which novices learned their keystrokes in an individual action context, there was no difference in participants' performance at test depending on whether they heard only the melody (equivalent to the individual goal in Experiment 1) or heard both melody and accompaniment (equivalent to the shared goal in Experiment 1). Thus, higher error rates in conditions that entailed only participants' individual goals in Experiment 1 cannot be explained by perceptual dissimilarity between the number of tones elicited by each action during learning and test.

General Discussion

The current study investigated whether people represent their actions in terms of the shared goal of their own and a partner's combined actions when learning to perform a joint action for the first time. In a transfer-of-learning paradigm, novices with no piano performance experience first learned to produce simple melodies in the joint action context of performing musical duets with a partner. As predicted, novices subsequently produced more errors when their previously-learned actions elicited only the melody (the goal of their individual actions) compared to the duet (the shared goal of their own and their partner's combined actions). Importantly, however, higher error rates were not simply caused by shifting from a coordination context at learning to a solo context at test, as they occurred regardless of whether or not participants performed their actions together with their partner at test (Experiment 1). Higher error rates in the melody-only conditions were also not simply the result of perceptual dissimilarity between the number of tones elicited by participants' actions during learning compared to test (Experiment 2). Together, these findings indicate that novices represent their actions in terms of a shared goal in joint action contexts that promote minimal representations. This provides support for the theoretical proposal that people engaged in a joint action will represent the shared goal of the task (Vesper et al., 2010).

The current study tested representations of shared goals in a minimal joint action context by asking piano novices to perform piano duets for the first time. It is already well-established that in less minimal contexts, people readily form representations of each other's individual action goals (Atmaca, Sebanz, & Knoblich, 2011; Knoblich et al., 2011; for examples specific to music performance, see Loehr & Palmer, 2011; Novembre, Ticini, Schutz-Bosbach, & Keller, 2012; Novembre, Ticini, Schütz-Bosbach, & Keller, 2013). However, few attempts have been made to tease representations of shared and individual goals apart in joint action. Two studies have used expert music performance paradigms to examine both shared and individual goals.

Keller and Burnham (2005) showed that musicians are able to attend to and remember both their own part and the composite of their own and a complementary part when performing duets. Loehr et al. (2013) showed that duetting pianists perceive shared goals (the musical harmony produced by the two pianists' combined pitches) as more important than individual action goals (the individual pitches produced by each pianist), as evidenced by stronger neural responses to pitch errors that affect the former compared to the latter. The findings of the current study indicate that shared goals are also more salient than individual goals in novel joint actions performed by non-experts. This provides evidence that representing joint actions in terms of shared goals is not specific to expert music performance, and suggests that representing shared goals could be a defining characteristic of all joint tasks (cf. Gallotti & Frith, 2013).

Vesper et al.'s (2010) minimal account of joint action also proposed that novices do not necessarily represent their partner as the intentional agent responsible for the complementary task required to achieve a shared goal. Instead, their representations might only entail the fact that someone or something else has to make a contribution to achieve the shared goal. Consistent with this proposal, participants' performance at test in Experiment 1 did not differ depending on whether they produced the shared goal together with their partner or with the accompaniment filled in by the computer. This suggests that, once a shared goal representation was formed during learning, novices may not have subsequently represented their partner as the agent of the accompaniment. Thus, although a joint learning context is critical for forming a representation of the shared goal (as evident when comparing Experiments 1 and 2), the importance of acting with a specific intentional agent may decline over the course of learning. This interpretation is consistent with previous work showing that stimulus-response associations learned by acting jointly with one partner transfer to subsequent joint actions performed with a different partner (Milanese, Iani, Sebanz, & Rubichi, 2011). Future work should investigate the factors that determine whether or not people represent their partner as the agent of the additional contribution to the shared goal, as well as how representations of the partner may change over the course of learning a joint action.

Consistent with ideomotor accounts of action (Hommel et al., 2001; Prinz, 1997; Shin, Proctor, & Capaldi, 2010), novices readily form representations in which keystrokes are associated with the tones they produce when learning to perform melodies alone (Bangert & Altenmüller, 2003; Engel et al., 2012; Lahav, Saltzman, & Schlaug, 2007; see Novembre & Keller, 2014, for a review). The current findings indicate that novices also readily incorporate the auditory effects of others' actions into their sensorimotor action representations when learning to perform musical pieces together with a partner. These findings complement previous work indicating that people are able to predict the consequences of others' actions simultaneously with their own (Loehr et al., 2013; Van der Steen & Keller, 2013; Vesper, Knoblich, & Sebanz, 2014; Wolpert, Doya, & Kawato, 2003) and incorporate predictions about others' actions when planning and performing their part of a joint action (Knoblich & Jordan, 2003; Kourtis, Sebanz, & Knoblich, 2013; Loehr & Palmer, 2011; Vesper, van der Wel, Knoblich, & Sebanz, 2013). However, participants in the current study did not simply learn to associate each action with two combined effects (melody tone plus accompaniment tone) independently of the social nature of the joint action. Instead, representing actions in terms of combined effects was specific to the joint action learning context in which two people coordinated their actions to achieve a shared goal, indicating a top-down modulation of action-effect integration.

Two differences between the joint and solo learning contexts suggest that participants may have represented their actions in terms of combined effects in the joint context because this context allowed the effects to be represented independently, i.e., as multiple independent effects that are combined rather than as a single unitary effect (cf. Snyder & Alain, 2007). First, in the joint context participants' task knowledge included the fact that two people produced the shared goal together, which may have provided top-down evidence for independent but combined effects. Second, the joint context included asynchronies between the melody and accompaniment, which may have provided perceptual evidence for independent but combined effects. The asynchronies were relatively small (~40 ms on average) and fell well within the critical time window for forming integrated action-effect representations (i.e., less than 2 s; Elsner & Hommel, 2004). Correlational analyses showed no associations between the size of asynchronies during learning and performance in any of the test conditions (see Table 1), suggesting that asynchronies may not have played a critical role in forming representations of the shared goal. However, the degree to which task knowledge and asynchronies contribute to people's ability to represent actions in terms of shared goals remains an important question that should be tested in future research.

(Table 1 about here)

In sum, the current study provides evidence that novices form representations of the shared goal of a joint action when they learn to perform joint actions for the first time. These findings support the theoretical predictions made by the minimal architecture for joint action (Vesper et al., 2010). On a broader scale, comparing Experiment 1 (joint action) and Experiment 2 (individual action) suggests that representations of individual actions are different when they have been learned in the context of joint compared to individual performance; people's mental representations of actions go beyond their individual action goals when they work together with a partner to achieve a shared goal. The current findings also add to a growing body of literature delineating how social learning transfers across tasks, partners, and social contexts (Ferraro et al., 2012; Milanese et al., 2010; Milanese et al., 2011). Although further research is required to fully understand the preconditions for representing actions in terms of shared goals, the present study highlights how social interaction can shape our most basic action, perception, and cognitive processes.

Acknowledgements

This research was supported by the European Union's 7th Framework Programme (Marie Curie International Incoming Fellowship) under Project 254419. The authors thank Maria Aracama for assistance with data collection and Günther Knoblich for helpful discussions.

References

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint Flanker effect: Sharing tasks with real and imagined co-actors. *Experimental Brain Research, 211*, 371-385.

Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience, 4*, 26.

Butterfill, S. (2012). Joint action and development. *The Philosophical Quarterly, 62*, 23-47.

Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychological Research, 68*, 138-154.

Engel, A., Bangert, M., Horbank, D., Hijmans, B. S., Wilkens, K., Keller, P. E., & Keysers, C. (2012). Learning piano melodies in visuo-motor or audio-motor training conditions and the neural correlates of their cross-modal transfer. *Neuroimage, 63*, 966-978.

Ferraro, L., Iani, C., Mariani, M., Nicoletti, R., Gallese, V., & Rubichi, S. (2012). Look what I am doing: Does observational learning take place in evocative task-sharing situations? *PLoS ONE, 7*, e43311.

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences, 17*, 160-165.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849-937.

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., & Burnham, D. K. (2005). Musical meter in attention to multipart rhythm. *Music Perception, 22*, 629-661.

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.

Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: An EEG study. *Social Neuroscience, 8*, 31-42.

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.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863.

Loehr, J. D., Kourtis, D., Vesper, C., Sebanz, N., & Knoblich, G. (2013). Monitoring individual and joint action outcomes in duet music performance. *Journal of Cognitive Neuroscience, 25*, 1049-1061.

Loehr, J. D., & Palmer, C. (2007). Cognitive and biomechanical influences in pianists' finger tapping. *Experimental Brain Research, 178*, 518-528.

Loehr, J. D., & Palmer, C. (2009). Sequential and biomechanical factors constrain timing and motion in tapping. *Journal of Motor Behavior, 41*, 128-136.

Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *Quarterly Journal of Experimental Psychology, 64*, 2153-2167.

Milanese, N., Iani, C., & Rubichi, S. (2010). Shared learning shapes human performance: Transfer effects in task sharing. *Cognition, 116*, 15-22.

Milanese, N., Iani, C., Sebanz, N., & Rubichi, S. (2011). Contextual determinants of the social-transfer-of-learning effect. *Experimental Brain Research, 211,* 415-422.

Novembre, G., & Keller, P. E. (2014). A conceptual review on action-perception coupling in the musicians' brain: What is it good for? *Frontiers in Human Neuroscience, 8*, 603.

Novembre, G., Ticini, L. F., Schutz-Bosbach, S., & Keller, P. E. (2012). Distinguishing self and other in joint action. Evidence from a musical paradigm. *Cerebral Cortex, 22,* 2894-2903.

Novembre, G., Ticini, L. F., Schütz-Bosbach, S., & Keller, P. E. (2013). Motor simulation and the coordination of self and other in real-time joint action. *Social Cognitive and Affective Neuroscience, 9*, 1062-1068.

Pacherie, E. (2011). Framing joint action. *Review of Philosophy and Psychology, 2,* 173-192.

Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science, 11*, 63-68.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9,* 129-154.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88,* B11-B21.

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1234-1246.

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin, 136*, 943-974.

Snyder, J. S., & Alain, C. (2007). Toward a neurophysiological theory of auditory stream segregation. *Psychological Bulletin, 133,* 780-799.

Tsai, C. C., Kuo, W. J., Jing, J. T., Hung, D. L., & Tzeng, O. J. L. (2006). A common coding framework in self-other interaction: Evidence from joint action task. *Experimental Brain Research, 175*, 353-362.

Van der Steen, M. C., & Keller, P. E. (2013). The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience, 7*, 253.

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks, 23*, 998-1003.

Vesper, C., Knoblich, G., & Sebanz, N. (2014). Our actions in my mind: Motor imagery of joint action *Neuropsychology, 55,* 115-121.

Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2013). Are you ready to jump? Predictive mechanisms in interpersonal coordination. *Journal of Experimental Psychology: Human Perception & Performance, 39*, 48-61.

Wenke, D., Atmaca, S., Holländer, A., Liepelt, R., Baess, P., & Prinz, W. (2011). What is shared in joint action? Issues of co-representation, response conflict, and agent identification. *Review of Philosophy and Psychology, 2*, 147-172.

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.

Footnotes

[1]None of the participants reported noticing this delay, nor was it noticeable to the expert accompanist.

[2]Despite the fact that participants were not given specific instructions about the rhythm with which to perform the melodies, participant-experimenter asynchronies during learning were relatively small: the mean absolute asynchrony was 45.43 ms (*SD* = 16.44) in Experiment 1a and 37.23 ms (*SD* = 9.38) in Experiment 1b, and did not differ across test conditions in either experiment.

[3]No accompanist errors occurred during the test phase because the accompanist's keystrokes were muted.

[4]Based on Lakens' (2013) recommendations, we report η_P^2 for use in power analyses and for ease of comparison with other literature that reports this measure, and Hedges' *g* (g_{av} and g_s for within- and between-subject comparisons, respectively) for use in meta-analyses. Hedges' *g* is a *d* family corrected effect size and can be interpreted similarly to Cohen's *d* (e.g., 0.5 indicates a medium effect size). Hedges' *g* values were calculated using the spreadsheet provided in Lakens (2013).

Table 1. Correlations between mean absolute asynchrony during learning trials (calculated separately for each test condition) and error rate during test trials, for each test condition in Experiments 1a and 1b.

	Experiment 1a		Experiment 1b	
Test Condition	*r*(29)	*p*	*r*(29)	*p*
Individual Goal				
Alone	.18	.33	.11	.56
Together	.016	.93	.01	.96
Shared Goal				
Alone	.14	.46	.053	.78
Together	-.22	.24	-.089	.64

Figure Captions

Figure 1. Example piano duet and experimental set-up. The piano duet consists of a simple melody (top line of musical notation, produced by novice participant sitting on the right) and accompaniment (bottom line of musical notation, produced by partner sitting on the left). Numbers above and below musical notation indicate right-hand finger sequences required to produce the melody and accompaniment, respectively. An occluder prevents participants and partners from receiving visual feedback about each other's keystrokes.

Figure 2. (a) Experiment design and abbreviated instructions to participants for Experiment 1. Symbols represent manipulations of elicited goal (individual = ♪; shared = ♪ ♪) and performance setting (alone = ✋; together = ✋✋). (b) Sketch of the transfer-of-learning procedure.

Figure 3. Mean error rates by elicited goal and performance setting in Experiments 1a and 1b. Insets: The mean difference and 95% CI for the main effect of elicited goal (individual goal minus shared goal) in each experiment.

Figure 4. Experiment design and abbreviated instructions to participants for Experiment 2. Symbols represent manipulations of elicited goal (individual = ♪; shared* = ♪ ♪) and performance setting (alone = ✋).

Figure 5. Mean error rates by elicited goal in Experiment 2. Inset: The mean difference and 95% CI for the effect of elicited goal (individual goal minus shared goal).

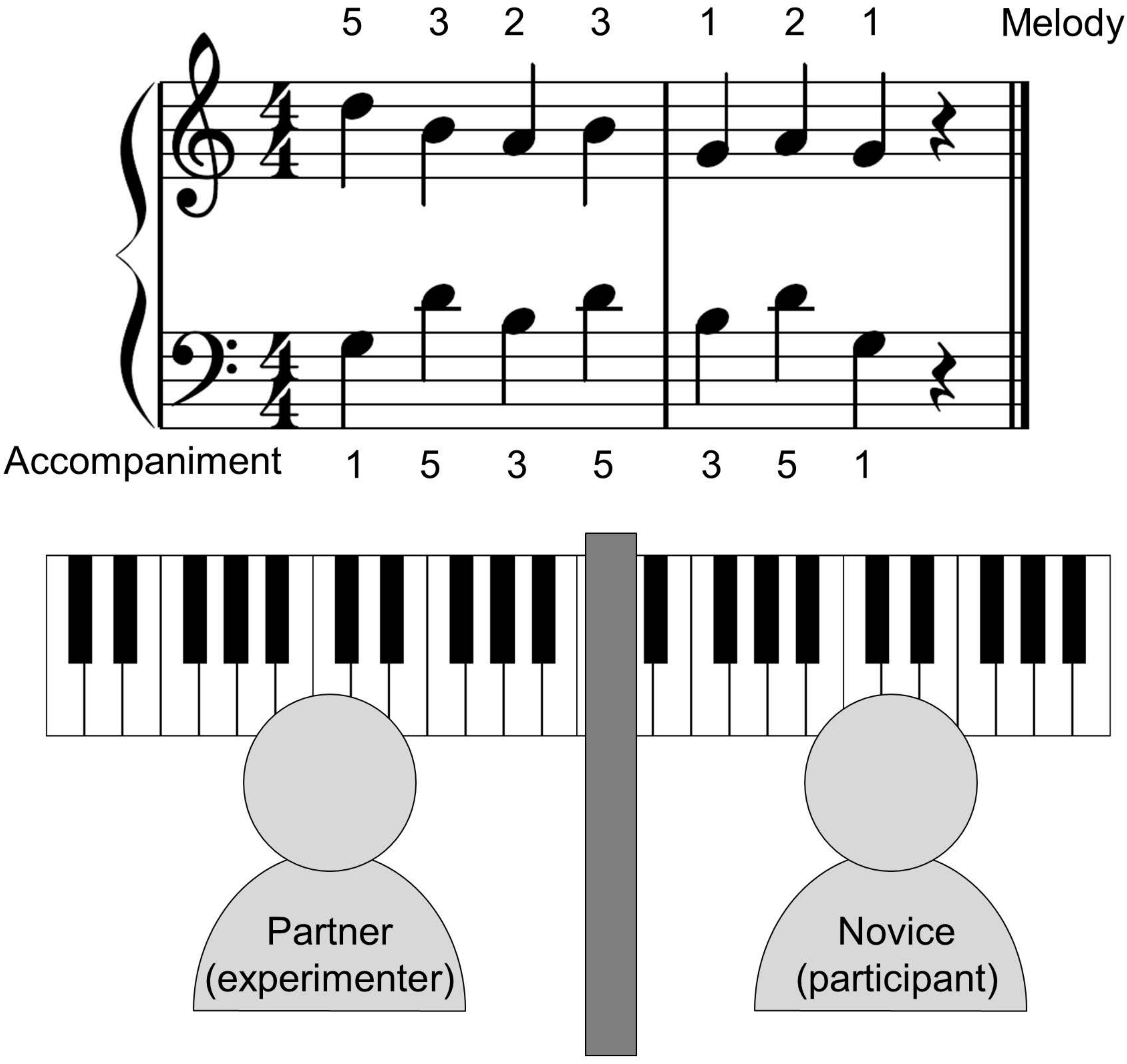

5 3 2 3 1 2 1
Melody
Accompaniment
1 5 3 5 3 5 1
Partner
(experimenter)
Novice
(participant)

Figure 1

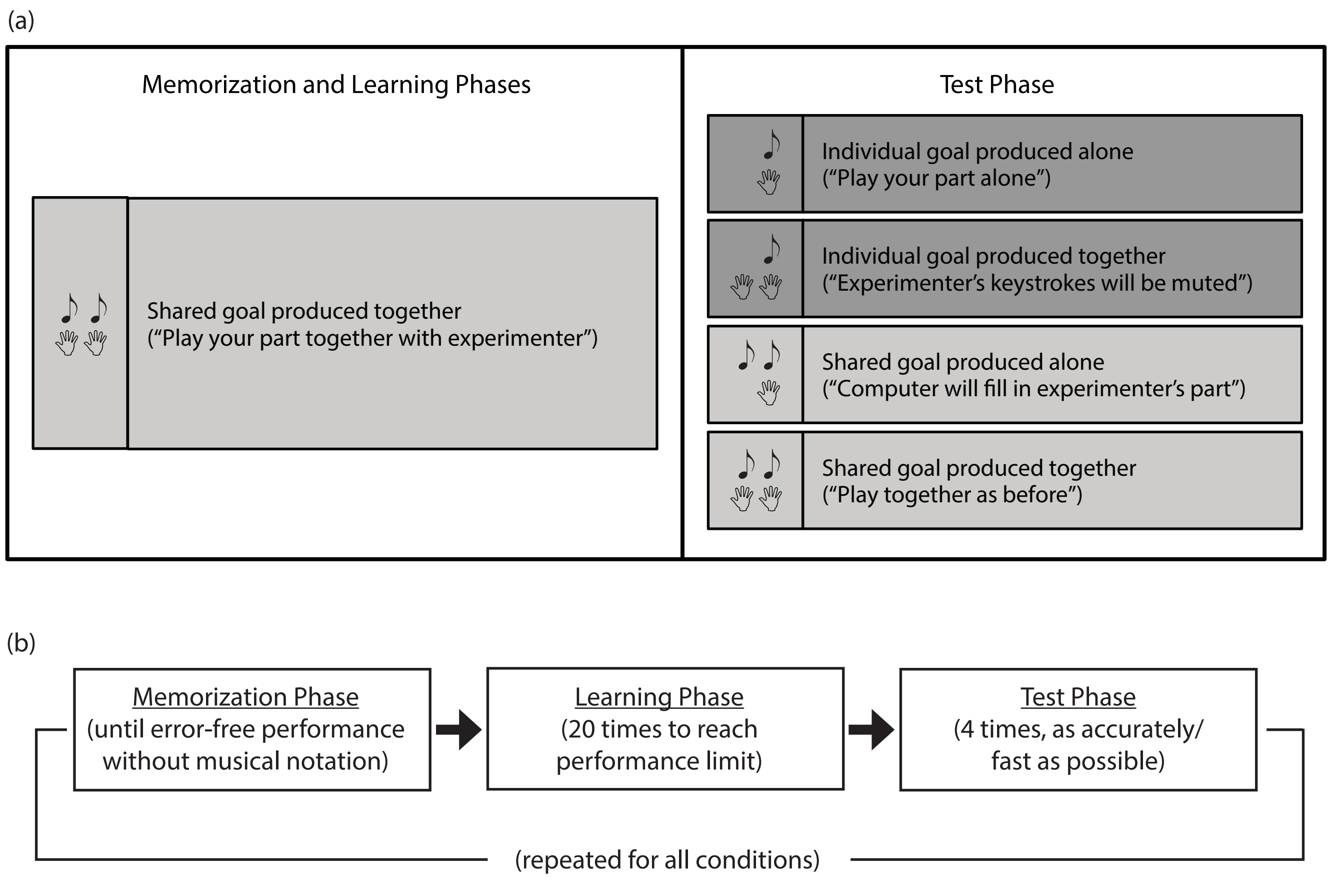
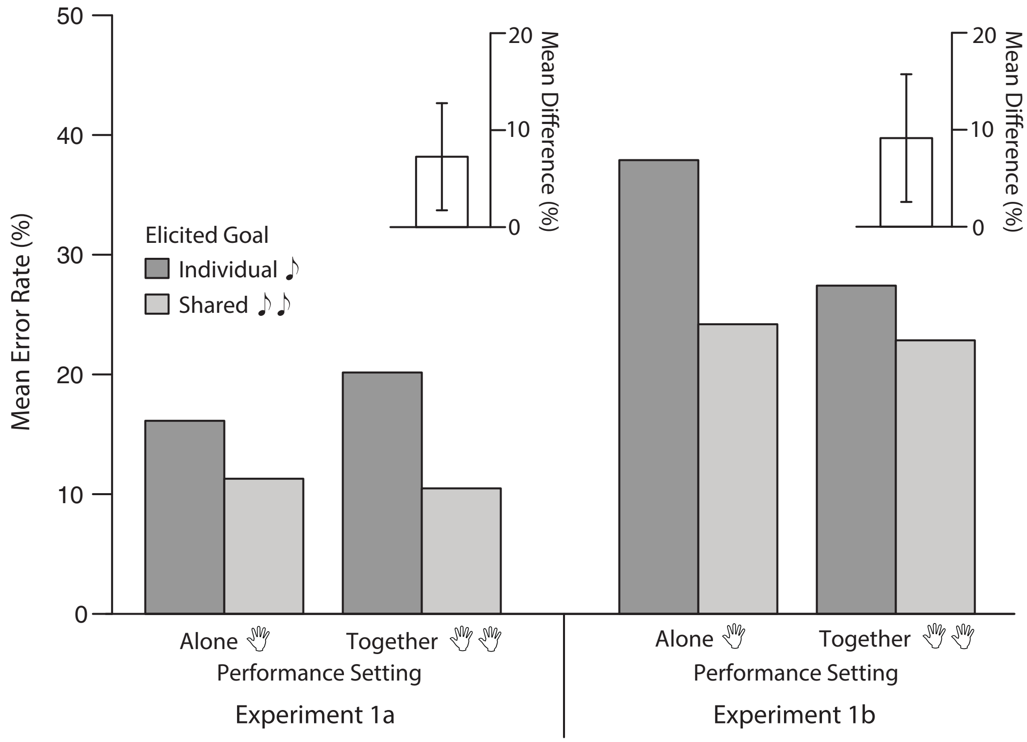

(a)
Memorization and Learning Phases
Shared goal produced together
("Play your part together with experimenter")
Test Phase
Individual goal produced alone
("Play your part alone")
Individual goal produced together
("Experimenter's keystrokes will be muted")
Shared goal produced alone
("Computer will fill in experimenter's part")
Shared goal produced together
("Play together as before")
(b)
Memorization Phase
(until error-free performance without musical notation)
Learning Phase
(20 times to reach performance limit)
Test Phase
(4 times, as accurately/ fast as possible)
(repeated for all conditions)

Figure 2



Mean Error Rate (%)
50
40
30
20
10
0
Elicited Goal
Individual
Shared
20
10
0
Mean Difference (%)
Alone
Together
Performance Setting
Experiment 1a
Alone
Together
Performance Setting
Experiment 1b

Figure 3

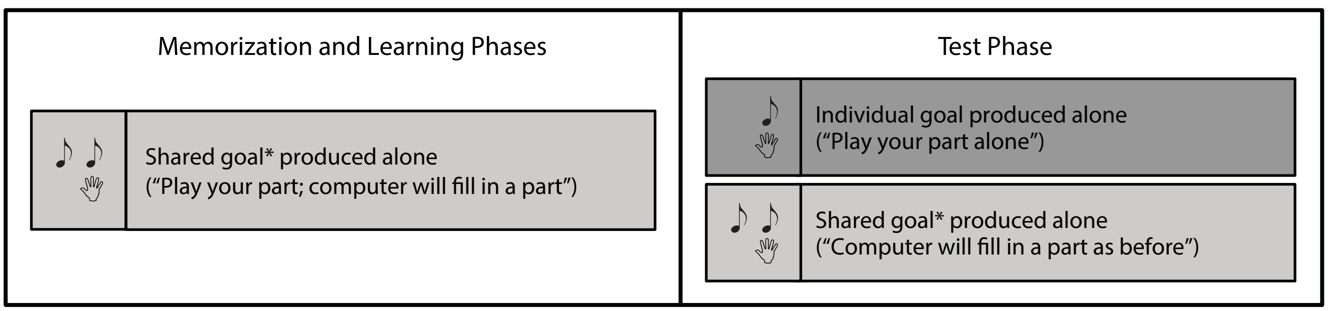
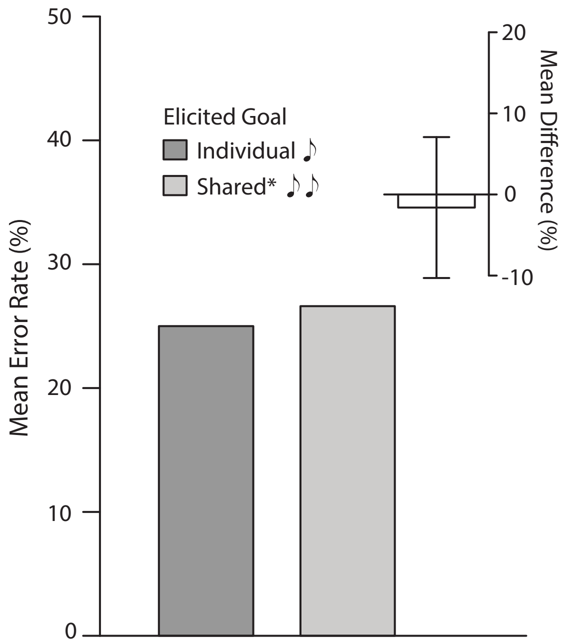

Memorization and Learning Phases
Shared goal* produced alone
("Play your part; computer will fill in a part")
Test Phase
Individual goal produced alone
("Play your part alone")
Shared goal* produced alone
("Computer will fill in a part as before")

Figure 4



50
40
30
20
10
0
Mean Error Rate (%)
Elicited Goal
Individual ♪
Shared* ♪♪
20
10
0
-10
Mean Difference (%)

Figure 5

Supplemental Material

Experiment 1 Learning Phase

We conducted a detailed analysis of participants' performance in the learning phase to determine whether participants' accuracy or speed changed over the course of learning and to confirm that participants followed the instructions in Experiments 1a and 1b. We examined both error rates and melody durations, which were calculated as the number of seconds elapsed between the first and last note of the melody. Melody durations were calculated only for correct trials. To examine changes over the course of learning, trials were binned into groups of four (trials 1-4, 5-8, …, 17-20). To examine whether participants followed instructions, experiment was included as a variable in the ANOVA. Thus, error rates and melody durations were analyzed with 2 (elicited goal) x 2 (performance setting) x 5 (trial bin) x 2 (experiment) ANOVAs.

Figure S1 shows the mean error rates and melody durations for each trial bin in each experiment. The four-way ANOVA on error rates revealed no significant main effects or interactions, indicating that error rates did not change over the course of learning and did not differ depending on instructions to participants.

The four-way ANOVA on melody durations revealed a main effect of trial bin, $F(4, 240) = 114.22$, $MSE = 27792.68$, $p < .001$, $\eta_P^2 = 0.66$, and a trial bin by experiment interaction, $F(4, 240) = 9.93$, $MSE = 27792.68$, $p < .001$, $\eta_P^2 = 0.14$. No other main effects or interactions were significant. As Figure S1 shows, melody duration decreased across learning trials in both experiments, but the decrease was larger in Experiment 1b than Experiment 1a. This confirms that participants followed the instructions to increase their performance speed toward maximum over the learning trials in Experiment 1b.

Experiment 2 Learning Phase

Participants' performance in the learning phase was examined in terms of both error rates and melody durations, which were compared with 2 (elicited goal) x 2 (trial bin) ANOVAs.

Figure S2 shows the mean error rates and melody durations for each trial bin. The ANOVA on error rates revealed that error rates differed across trial bins, $F(4, 120) = 2.62$, $MSE = 0.027$, $p = .038$, $\eta_P^2 = 0.082$. No other effects were significant. Paired-samples *t*-tests (Bonferonni-corrected $\alpha = .005$) indicated that participants made significantly more errors in the third trial bin compared to the first, $t(30) = 3.24$, $p = .003$.

The ANOVA on melody durations[1] revealed a main effect of trial bin, $F(4, 112) = 57.37$, $MSE = 48038.96$, $p < .001$, $\eta_P^2 = 0.67$, and no other effects. As Figure S2 shows, melody durations decreased across learning trials, similar to Experiment 1.

Footnote

[1]Two participants were excluded from this analysis because they had a 100% error rate during one trial bin, which precluded calculation of the mean melody duration for that trial bin.

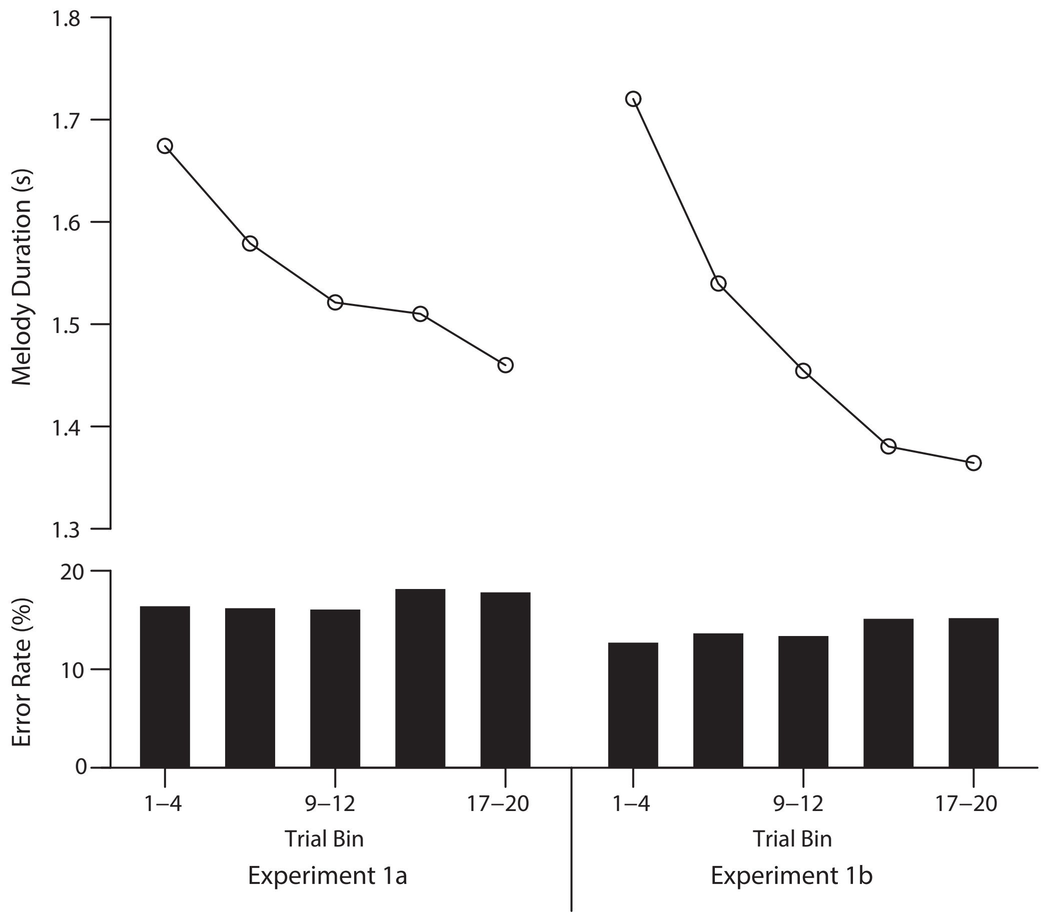


Figure S1. Mean melody duration (top) and error rate (bottom) by trial bin in Experiments 1a and 1b.

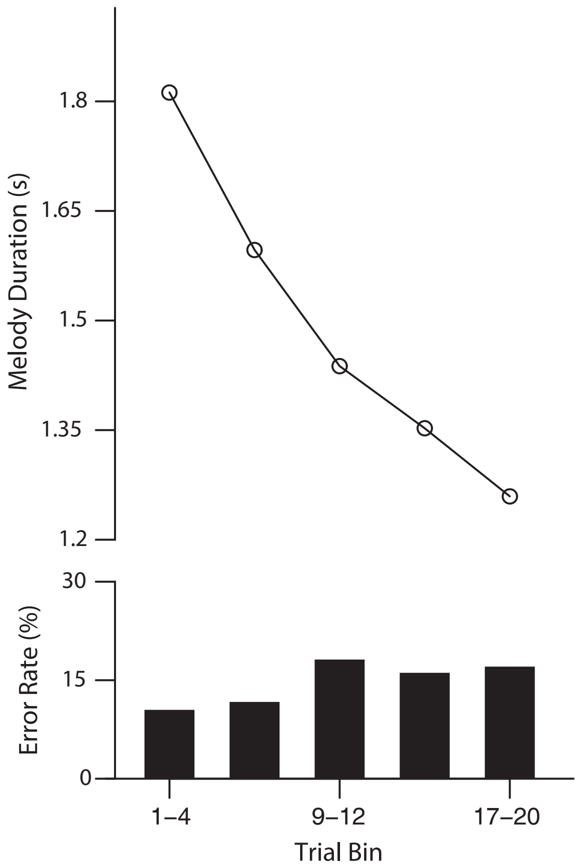


Figure S2. Mean melody duration (top) and error rate (bottom) by trial bin in Experiment 2.