Corepresentation During Joint Action in Marmoset Monkeys (Callithrix jacchus).

— Source link ↗

Fabia Miss, Judith M. Burkart

Institutions: University of Zurich

Related papers:

- The evolutionary origin of human hyper-cooperation

- Cooperative breeding and human cognitive evolution

- Social and Physical Cognition in Marmosets and Tamarins

- Marmoset prosociality is intentional.

- Attention in common marmosets : implications for social-learning experiments

Corepresentation During Joint Action in Marmoset Monkeys (Callithrix jacchus)

Miss, Fabia M ; Burkart, Judith M

Abstract: Behavioral coordination is a fundamental element of human cooperation. It is facilitated when individuals represent not only their own actions but also those of their partner. Identifying whether action corepresentation is unique to humans or also present in other species is therefore necessary to fully understand the evolution of human cooperation. We used the auditory joint Simon task to assess whether action corepresentation occurs in common marmosets, a monkey species that engages extensively in coordinated action during cooperative infant care. We found that marmosets indeed show a joint Simon effect. Furthermore, when coordinating their behavior in the joint task, they were more likely to look at their partner than in a joint control condition. Corepresentation is thus not unique to humans but also present in the cooperatively breeding marmosets. Since marmosets are small-brained monkeys, our results suggest that routine coordination in space and time, rather than complex cognitive abilities, plays a role in the evolution of corepresentation.

Corepresentation During Joint Action in Marmoset Monkeys (*Callithrix jacchus*)

Fabia M. Miss and Judith M. Burkart
Department of Anthropology, University of Zurich

Abstract
Behavioral coordination is a fundamental element of human cooperation. It is facilitated when individuals represent not only their own actions but also those of their partner. Identifying whether action corepresentation is unique to humans or also present in other species is therefore necessary to fully understand the evolution of human cooperation. We used the auditory joint Simon task to assess whether action corepresentation occurs in common marmosets, a monkey species that engages extensively in coordinated action during cooperative infant care. We found that marmosets indeed show a joint Simon effect. Furthermore, when coordinating their behavior in the joint task, they were more likely to look at their partner than in a joint control condition. Corepresentation is thus not unique to humans but also present in the cooperatively breeding marmosets. Since marmosets are small-brained monkeys, our results suggest that routine coordination in space and time, rather than complex cognitive abilities, plays a role in the evolution of corepresentation.

Keywords
common marmoset monkeys, coordinated joint action, corepresentation, evolution, joint Simon task, social cognition

Received 5/9/17; Revision accepted 1/5/18

Humans are a hyper-cooperative species, and the evolution of human cooperation and its underlying mechanisms continue to be a matter of debate (Richerson et al., 2016; Tomasello & Gonzalez-Cabrera, 2017). Our goal was therefore to investigate whether one such mechanism, namely action corepresentation during joint action, may be present in another highly cooperative primate species, the common marmoset monkey (*Callithrix jacchus*).

When human social partners are jointly engaged in a cooperative task, they corepresent each other's actions—for instance, in the joint Simon task (Sebanz, Knoblich, & Prinz, 2003). This task is based on the individual Simon effect (Simon & Rudell, 1967), which is the compatibility effect arising when an irrelevant feature of a test stimulus interferes with the response, as when a stimulus is played back from one side, whereas the required response is on the opposite side. For instance, in the auditory version (Ruys & Aarts, 2010) of the Simon task (full task), subjects have to learn to discriminate between two sound stimuli, "L" and "R," and choose the corresponding left- or right-hand response option. If the stimuli are played back

from either the left-hand or the right-hand side of the subject, the task is easier in compatible trials (i.e., when stimulus "L," requiring a left-hand response, is played back from the left-hand side, and stimulus "R," requiring a right-hand response, is played back from the right-hand side) than in incompatible trials (i.e., when stimulus "L" is played back from the right-hand side, and stimulus "R" is played back from the left-hand side). This compatibility effect is referred to as the Simon effect. However, when subjects have to solve only half of the task because one response option is not available (i.e., the half task), the corresponding stimulus can be ignored, and the compatibility or Simon effect disappears. Thus, when only one response option is available (e.g., on the right-hand side) and both stimuli are used, the task is equally difficult regardless of whether the stimuli are played back from the right- or left-hand side.

Corresponding Author:
Judith M. Burkart, Department of Anthropology, University of Zurich,
Winterthurerstrasse 180, Zurich, 8057, Switzerland
E-mail: judith.burkart@aim.uzh.ch

Intriguingly, the Simon effect reappears in a joint task condition. Here, subjects are still responsible for only one half of the task, while their partner must solve the other half. Subjects thus behave as if they were responsible not only for their own half of the task but also for their partner's half, suggesting that they represent not only their own actions but also their partner's (Butterfill, 2012; Ruissen & de Bruijn, 2016; Vesper, Butterfill, Knoblich, & Sebanz, 2010).

The joint Simon effect is modulated by several social factors. It is stronger among friends compared with nonfriends (Hommel, Colzato, & Van Den Wildenberg, 2009), among in-group compared with out-group members (McClung, Jentzsch, & Reicher, 2013), and after nasal oxytocin application (Ruissen & de Bruijn, 2015), and it is positively correlated with empathy among friends (Ford & Aberdein, 2015). Furthermore, it appears when participants are told they are playing with a partner in a different room but not when they are told they are playing with a computer (Tsai, Kuo, Hung, & Tzeng, 2008; see also Atmaca, Sebanz, & Knoblich, 2011, for results from the similar Eriksen flanker task) and when they believe partners have voluntary control over their actions compared with when partners' actions are being controlled by a machine (Atmaca et al., 2011).

Some work suggests a link with theory-of-mind understanding. There is no evidence for corepresentation in 2- to 3-year-old children (Milward, Kita, & Apperly, 2014), but it has been shown in 4- and 5-year-olds (Milward et al., 2014; Saby, Bouquet, & Marshall, 2014), where individual differences are linked to inhibitory control and theory-of-mind abilities (Milward, Kita, & Apperly, 2016). Likewise, Humphreys and Bedford (2011) found no joint effect in patients with theory-of-mind deficits. However, it is also likely that the perception of agency, defined as perceiving an agent as the initiator or causal source of an action, rather than full-fledged intentionality and thus theory of mind, is sufficient for the emergence of a joint Simon effect (Stenzel et al., 2014).

At the same time, however, low-level perceptual processes can be sufficient to elicit joint Simon effects (Dolk et al., 2014; Klempova & Liepelt, 2016). For instance, a joint Simon effect has been reported when the "partner" was an inanimate Japanese waving cat (Lien, Pedersen, & Proctor, 2015). This gave rise to the referential-response-coding approach, a nonsocial account of this coaction effect (Dolk et al., 2014).

To explore the evolutionary origin of the joint Simon effect, we investigated whether it would also be present in common marmoset monkeys. Common marmosets meet a suite of prerequisites identified by Sebanz, Bekkering, and Knoblich (2006) as relevant for corepresentation in joint tasks. First, as the only true cooperative breeders among primates besides humans, common marmosets arguably engage more in joint actions that require fine-grained action coordination than do other nonhuman primates. For instance, all group members carry infants, and the transfer of an infant from one carrier to another is a frequent and highly coordinated joint action (Snowdon, 2001). Marmosets also engage extensively in cooperative vocal turn taking, which requires that they take each other's responses into account (Takahashi, Narayanan, & Ghazanfar, 2013). Furthermore, these monkeys engage in action observation and imitation (Voelkl & Huber, 2007) and joint attention (Burkart & Heschl, 2006, 2007), and they perceive agency (Burkart, Kupferberg, Glasauer, & van Schaik, 2012). Since marmosets satisfy these requirements for corepresentation in a joint task as identified by Sebanz et al. (2006), we hypothesized that they would show a joint Simon effect. To further help disentangle whether such an effect was better explained as a true social effect or by low-level perceptual processes, we included a social control condition in which a conspecific partner was present but not engaged in the cooperative task.

Another important aspect of human cooperation is that it is often accompanied by social monitoring. For instance, communication allows children to effectively coordinate their joint efforts (Duguid, Wyman, Bullinger, Herfurth-Majstorovic, & Tomasello, 2014), which may be reflected in subtle communicative signals, such as gaze alternation, back checking, or mutual gaze (Wyman, Rakoczy, & Tomasello, 2013). We therefore also coded the gazing behavior of the subjects during the social task conditions (joint and joint control task) to further disentangle low-level perceptual and social explanations of a potential joint Simon effect. If the subjects perceived the joint but not the joint control condition as a shared task, we predicted that the focal individual would check back at (i.e., look at) its partner more in the joint task than in the joint control task. Furthermore, a particularly strong indication for experiencing the joint task as a shared task would be if the focal individual also engaged more in mutual gaze with its partner prior to responding to the cue.

Method

Subjects

We tested 10 (5 female, 5 male) adult common marmosets (*Callithrix jacchus*) from three family groups (Table 1). All individuals were born in captivity and housed in heated indoor enclosures. They were tested either in the morning or in the afternoon in their home enclosure in between regular feedings. As rewards, we used various special treats that were highly preferred by the

Table 1. Description of Test Subjects

Name	Age (in years) at time of testing	Sex	Social status	Partner in the joint and joint control task
	Family group A			
Jojoba	9	Female	Breeder	Marvin
Marvin	8	Male	Breeder	Jojoba
Jupie	6	Female	Helper	Jet
Jet	6	Male	Helper	Jupie
Joyce	2	Female	Helper	James
James	2	Male	Helper	Joyce
	Family group B			
Lea	8	Female	Breeder	Kyros
Kyros	7	Male	Breeder	Lea
Lima	0.6	Female	Infant	Not tested
	Family group C			
Jaja	6	Female	Breeder	Membo
Membo	6	Male	Breeder	Jaja
Jandira	0.7	Female	Infant	Not tested
Jala	0.7	Female	Infant	Not tested

particular subject, as established prior to the tests, mostly mealworms, nuts, crickets, or sweet fruits. The

dyads in the joint and the joint control task were chosen according to observations of behavior during the training phase, such as spending time together in the testing cage, participating next to each other in trials, and occasionally sharing food. The research was approved by the Kantonales Veterinäramt Zürich (License No. ZH183/13, 24826).

Apparatus and stimuli

We tested the common marmosets with an auditory version of the Simon task. The tests were conducted in a testing cage (74 cm × 40 cm × 40 cm) that was placed inside the marmosets' indoor enclosure, attached to the wire mesh such that subjects could reach and pull the drawers (Figs. 1 and 2). The subjects could be separated in three different compartments of the testing cage by using sliding doors, but visual contact between individuals was always possible. The response device consisted of two drawers that were attached to an apparatus that was placed in front of the testing cage and was connected to a laptop. When the correct drawer was pulled, the subjects could retrieve a reward from the food bowl. In the joint task condition (see below), both individuals were rewarded if the correct drawer was pulled, in order to signal the cooperative nature of the task. Accordingly, after the drawer was pulled within
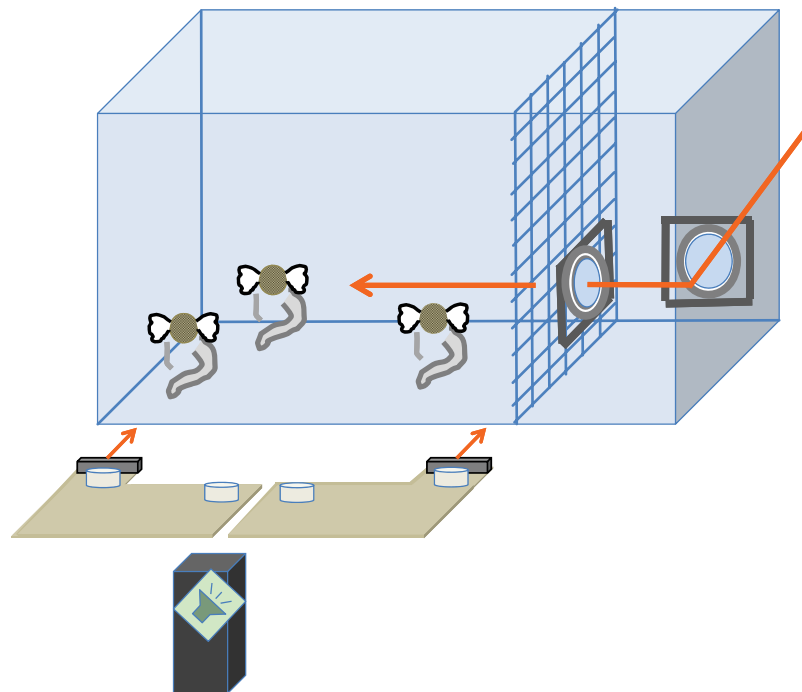

Fig. 1. Schematic drawing of the setup for the training phase, with the loudspeaker placed in a central position in front of the subjects. Depending on the sound stimulus, the individuals had to make the correct choice between the right-hand and the left-hand drawer in order to obtain the bait.
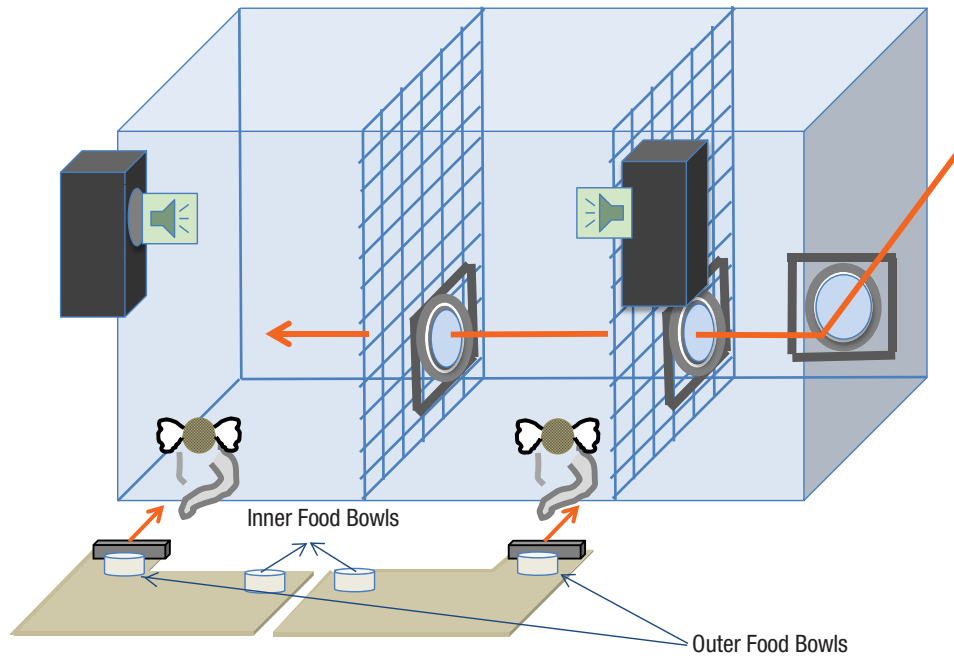
Fig. 2. Schematic drawing of the setup for the joint task situation.

reach, the outer food bowl was available to the active puller, and the inner food bowl was available to its partner (Fig. 2).

The trials were presented with OpenSesame (Mathôt, Schreij, & Theeuwes, 2012), a Python-based experimental testing environment that recorded the subjects' answers as soon as one drawer was pulled. For the testing setup, a screen was added so that the drawers could not be reached by the subjects before the screen was lifted, and the baiting procedure as well as the bait itself were not visible to the subjects.

As auditory stimuli, we used two different piano tone sequences "L" and "R" (see the Supplemental Material available online). They were considered as appropriate stimuli because their frequencies were within the most common hearing range of common marmosets (M. Osmanski, personal communication, February 26, 2015; Osmanski, Song, & Wang, 2013; Osmanski & Wang, 2011).

Experimental design

The task consisted of two experimental (*full task* and *joint task*) and two control (*half task and joint control task*) conditions (Fig. 3). If the marmosets corepresented each other's actions, they should show a compatibility effect (i.e., Simon effect) in the experimental conditions but not in the control conditions. We added the joint control task condition to control for the possibility that the joint Simon effect may exclusively result from low-level perceptual processes, independently of

whether a partner was indeed engaged in the joint task or not. In this condition, a marmoset partner was present but could not collaborate in the task (Fig. 3d). We therefore expected no compatibility effect in this joint control condition.

Procedure

Training. The training started with a habituation phase, in which the subjects were familiarized with the experimenter, the testing cage, and the whole testing procedure, including separation and opening and closing of sliding doors. The second phase consisted of training the marmosets to the full task with the sound stimuli coming from a central position (loudspeaker in front of the individuals; Fig. 1), thus without eliciting stimulus incompatibility.

In a first training step, we made sure that all the subjects were able to correctly pull the drawers and that they made the association of pulling and getting a reward for it. We allowed the whole group to enter the testing cage and participate together (with open sliding doors) so that individuals could learn from each other how to pull and retrieve the rewards from the drawer. The criterion for each individual to move on with the next training step was to successfully pull at least seven times in a session.

In the second training step, we made sure that the subjects learned the association of the particular sound stimulus (tone sequence "L" or "R") with the correct side (the left-hand or the right-hand drawer). We
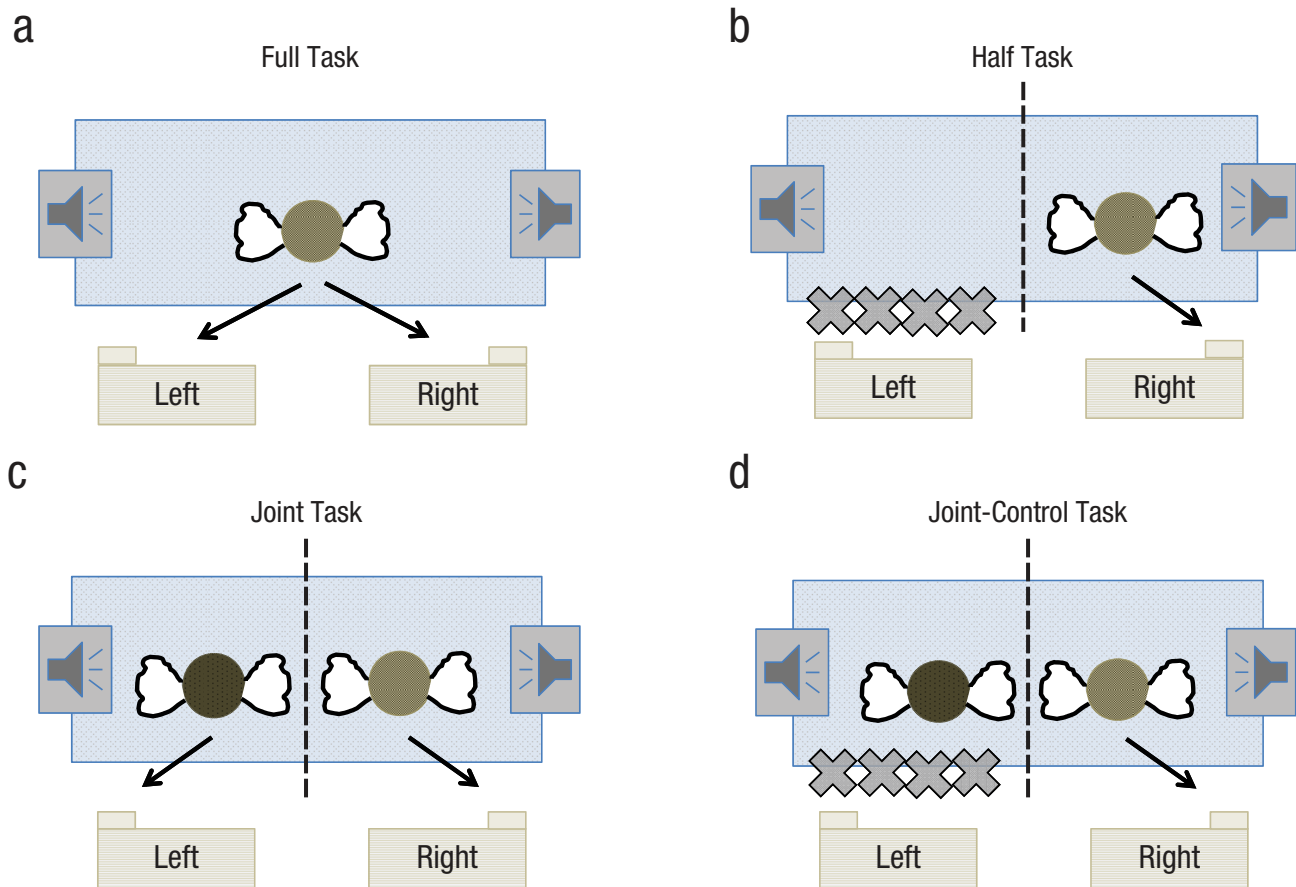
Fig. 3. Schematic drawing of the four different test conditions: (a) full task, (b) half task, (c) joint task, and (d) joint control task. The auditory stimuli "L" and "R" were broadcast from lateral speakers: Stimulus "L" signaled that the left-hand drawer was baited, whereas stimulus "R" signaled that right-hand drawer was baited. Broadcasting the stimuli from either the right- or left-hand side resulted in compatible trials (e.g., "L" played back from the left-hand side) and incompatible trials (e.g., "L" played back from the right-hand side). All tasks contained compatible and incompatible trials. In the full task (a), the focal individual had to solve the task alone. Here, we expected a compatibility effect (i.e., a Simon effect). In the half task (b), the focal individual had access to only half of the task, and we expected no Simon effect. In the joint task (c), the situation for the focal individual was the same as in the half task, but the second half of the task was solved by a second marmoset. Here, we expected a joint Simon effect. The joint control task (d) was identical to the joint task, but now the partner was prevented from responding to the stimuli because the access to its drawer was blocked by a wooden board. Here, we expected no Simon effect in the focal individual because the two subjects were no longer working together on the task.

alternated sessions in which individuals were tested alone, in dyads, and in the entire group. In the beginning, no screen was used so that the subjects could see the reward, and we played back the corresponding sound as soon as the subject pulled the baited drawer. The drawers were baited randomly with the constraint that the reward was not present more than twice in a row on the same side.

In the third training step, the screen was added to make sure the drawers could not be reached before the start of a trial and that the subjects could not observe which drawer was baited. The subjects were now tested individually, and a trial started with the sound stimulus. Simultaneously, the screen was lifted, but the bait was still invisibly covered. If the subject chose the correct

drawer, it was allowed to retrieve the reward and consume it, after which the next trial started. If the subject chose the incorrect drawer, the subject could check that the food bowl was empty, after which the screen was lowered. Training sessions lasted for 10 to 24 trials (maximum 15 min), depending on the motivation of the tested subject, and always ended with a correct choice by the subjects so they could finish with a positive experience. The criterion for starting the actual tests was to reach six sessions with at least 75% correct choices or 15 sessions with at least 70% correct choices.

Testing. The test sessions were identical to the training session (third step), except that the stimuli were no longer broadcast from the central position but from either

the left- or the right-hand side. Each subject was tested in all four conditions. In each test session (consisting of 10–24 trials depending on the motivation of the tested subjects), the subjects were presented with one condition. One round of testing consisted of four test sessions corresponding to the four conditions, and the subjects were tested over five (family group) or four (breeding pairs) rounds. The interval between the sessions was 2 to 3 days.

The order of the test conditions within a round, the side from which the stimulus was emitted, and the side of the focal individual and the partner during the joint and the joint control sessions were counterbalanced. The order of the two sound stimuli in a session was randomized to prevent side biases.

Each session started with a few warm-up pretrials, in which the sound stimuli were emitted from a central position (like in the training phase). Simultaneously, the experimenter pushed both drawers within reach of the subjects, allowing them to explore the food bowls on both sides and to retrieve the reward from the drawer. Next, the experimenter performed motivation pretrials (with sound stimuli still coming from a central position). In these motivation pretrials, the subjects had to pull the drawer themselves. If the subject chose the correct drawer, it was allowed to retrieve the reward and consume it. If the subject chose the incorrect drawer, the subject could check that the food bowl was empty. The criterion for starting the session on the particular day was to perform at least two consecutive correct choices in these motivation pretrials. If this criterion was not met, the subject was tested the following day.

Before starting a trial, the subject was attracted to the front and the middle of the testing cage. The sound stimulus was then played back from either the left or the right loudspeaker, and simultaneously, the screen was lifted but the bait was still covered. A trial ended either when the test subject (or one of the two test subjects) pulled one of the two drawers, thus asking for an exclusive choice in each trial, or after 30 s. If they made the correct choice, they were allowed to retrieve the rewards and consume them, after which the next trial started. If they made the incorrect choice, the subjects could check that the food bowls were empty, after which the screen was lowered, and there was a pause of 10 s to 15 s before the next trial started.

Data coding. The choices of the subjects were automatically recorded with OpenSesame. In addition, all sessions were video-recorded. Gazing behavior was quantified from the video recordings with Mangold INTERACT software (Mangold International, 2017). We coded the direction of the first orienting response and the looking behavior. To assess the direction of the first orienting response, we coded for each trial whether the focal individual had first headed to the left-hand or the right-hand side, defined as (a) turning the head toward that side and making at least two steps in that direction or (b) extending the arm toward a food bowl (whichever happened first). For the trials of the joint and the joint control task, we further coded for the time span between the sound stimulus and pulling (thus prior to responding to the cue), whether (a) the focal individual had looked at its partner and (b) the two individuals had mutually gazed at each other simultaneously. We coded occurrences (yes or no) per trial. We included only trials where it was the focal individual's turn to pull (indicated by the respective sound stimulus) to exclude instances of looking at the partner that were elicited by the partner's intention movement to pull.

We assessed the interrater reliabilities for 20% of the recorded trials of all tasks for the subjects' first orienting response directions (Cohen's $\kappa = .86$), as well as for all the recorded trials of the joint and the joint control task for the gaze data (Cohen's $\kappa = .72$).

Data analysis. We used two response variables to quantify a potential compatibility effect: the choice of a subject, which could either be correct or incorrect, and the direction of the first orienting response of the subject (i.e., the first two steps toward or the first reach for a drawer), which could also be correct or incorrect. We used generalized linear mixed models to test these effects and analyzed a total of 2,605 trials. We included compatibility (compatible, incompatible) and task condition (full, half, joint, joint control) as fixed factors, with contrasts set to compare control task conditions (i.e., half and joint control task) against experimental task conditions (i.e., full and joint task). We included the sound stimuli and the individuals nested within their family groups as random factors. We further included sex, age, session number, sound stimulus, and side as fixed factors in the model to control for any additional explanatory variables such as learning effects over sessions, preferences for one type of sound stimulus, or side biases.

For the subjects to react correctly in the half and the joint control task, answering included pulling the accessible drawer (i.e., possible trials) when the corresponding sound stimulus was emitted and not pulling when the sound stimulus of the blocked side (i.e., impossible trials) was played back. In the full and the joint task, both response options were accessible, and thus a correct answer was always possible. Coding a correct no pull in impossible trials of the half and the joint control task as a correct choice could arguably bias the results toward too many correct answers. However, such an effect would bias the results against our prediction and thus make the analysis more conservative. Additional

analyses furthermore suggested that such an effect was highly unlikely, because the pulling frequencies across the other tasks show that, when possible, the subjects would pull one of the response drawers in virtually every trial. A repeated measures analysis of variance on percentages of total no responses in all tasks revealed a significant difference between the experimental task conditions (full and joint task) and the control task conditions (half and joint control task), $F(3, 36) = 10.42$, $p < .001$, $\omega^2 = .41$. Within the half and the joint control task, respectively, there was almost double and fourfold no responses occurring in the impossible trials compared with the possible ones. Wilcoxon signed-rank tests on percentages of total no pulls in the half and the joint control task showed that no responses occurred significantly more often in impossible ($M = 10.16\%$) than in possible ($M = 5.24\%$) trials ($p = .04$, $r = −.64$). Nevertheless, no responses occurred in clearly less than 50% of all cases, which was consistent with the difficulty for nonhuman primates to inhibit prepotent action tendencies.

Finally, we corroborated our findings by also analyzing the direction of the first orienting response of the subjects. Note that for this response variable, no impossible trials existed because the subjects could also orient toward the drawer toward which they had no access. The pattern of results for both response variables were identical.

Results

Simon effect across conditions

When using the choice of the subject as a binary response variable, we found that the full model explained the data better than the null model, $\chi^2(16) = 68.42$, $p < .001$, and we found a significant main effect of compatibility (estimate = 0.21, $SE = 0.04$, $z = 5.23$, $p < .001$) as well as a significant Task × Compatibility interaction (estimate = 0.20, $SE = 0.04$, $z = 4.92$, $p < .001$; Fig. 4). All other fixed factors had no effect on the subjects' choices (Table 2). We found the same pattern of results when using the direction of the subjects' first orienting response as a response variable, $\chi^2(16) = 72.37$, $p < .001$ (Fig. 5; Table 3). Thus, both the choice behavior of the subjects as well as the direction of their first orienting response when approaching the apparatus were consistent with a compatibility effect in the full and the joint condition but not in the half and the joint control condition. The compatibility effects in the full and the joint tasks, but not in the half task or the joint control task, were a first indication that marmosets may corepresent each other's actions and thus perceive the joint task as a task they share with each other.
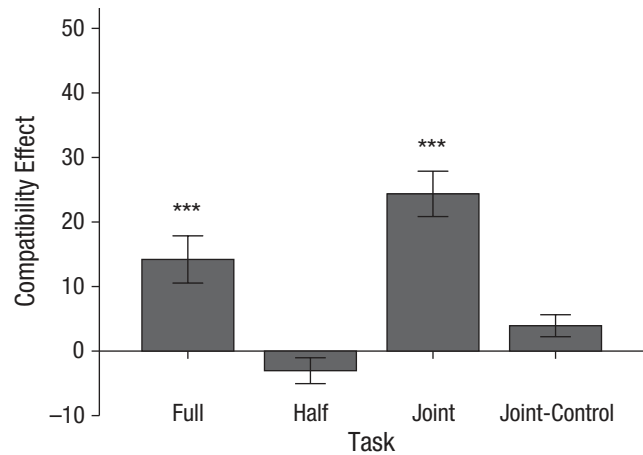

Fig. 4. Compatibility effect (i.e., Simon effect) across the four task conditions, calculated as the difference in the percentage of incorrect choices between incompatible trials and compatible trials per individual and session. Error bars represent standard errors of the mean. Asterisks indicate results that the compatibility effect was significantly stronger in the full and the joint task compared with the control tasks ($p < .001$).

Social monitoring

To further analyze whether the subjects really perceived the joint but not the joint control condition as a shared task, we tested whether the focal individual would (a) look at its partner more in the joint task than in the joint control task and (b) engage more in mutual gaze with the partner prior to responding to the cue.

We found that before pulling a drawer, the focal individuals indeed looked more often at their partner in the joint than in the joint control task, $\chi^2(5) = 66.38$,
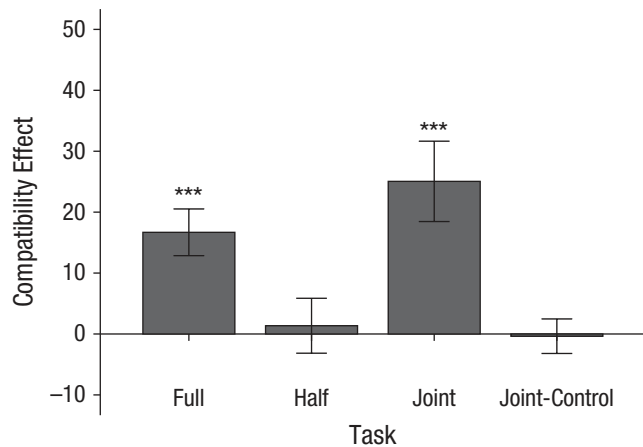

Fig. 5. Compatibility effect (i.e., Simon effect) quantified for the direction of the first orienting response across the four task conditions. The compatibility effect was calculated as the difference in the percentage of incorrect directions of first orienting responses between incompatible trials and compatible trials per session and sound stimulus. Error bars represent standard errors of the mean. Asterisks indicate that the compatibility effect was significantly stronger in the full and the joint task compared with the control tasks ($p < .001$).

Table 2. Results of the Generalized Linear Mixed Model With Choice of the Subject as the Response Variable

Fixed effect	Estimate	*SE*	Odds ratio	95% CI	*z*	*p*
Intercept	0.23	0.16				
Task						
Experimental (full, joint) vs. control (half, joint control)	0.06	0.04	1.06	[0.98, 1.15]	1.38	.17
Experimental (full, joint)	−0.02	0.06	0.98	[0.87, 1.10]	−0.34	.74
Control (half, joint control)	0.05	0.06	1.05	[0.93, 1.18]	0.86	.39
Compatibility	0.21	0.04	1.23	[1.14, 1.33]	5.23	1.71E–07
Side	−0.09	0.08	0.91	[0.78, 1.07]	−1.06	.29
Sound stimulus	−0.07	0.04	0.93	[0.86, 1.01]	−1.64	.10
Sex	0.03	0.08	1.03	[0.88, 1.21]	0.37	.71
Age	0.01	0.02	1.01	[0.97, 1.05]	0.52	.60
Session						
Linear trend	−0.09	0.10	0.91	[0.75, 1.11]	−0.95	.34
Quadratic trend	−0.04	0.09	0.96	[0.81, 1.15]	−0.45	.65
Cubic trend	0.13	0.09	1.14	[0.95, 1.36]	1.38	.17
Quartic trend	−0.11	0.09	0.90	[0.75, 1.07]	−1.27	.20
Compatibility × Task interactions						
Compatibility × Experimental Tasks vs. Control Tasks	0.20	0.04	1.22	[1.13, 1.32]	4.92	8.83E–07
Compatibility × Full Task vs. Joint Task	−0.11	0.06	0.90	[0.80, 1.01]	−1.92	.06

Note: CI = confidence interval.

Table 3. Results of the Generalized Linear Mixed Model With Direction of the First Orienting Response as the Response Variable

Fixed effect	Estimate	*SE*	Odds ratio	95% CI	*z*	*p*
Intercept	−.025	0.13				
Task						
Experimental (full, joint) vs. control (half, joint control)	−0.09	0.04	0.91	[0.85, 0.99]	−2.10	.04*
Experimental (full, joint)	0.05	0.06	1.05	[0.93, 1.18]	0.83	.40
Control (half, joint control)	0.02	0.06	1.02	[0.91, 1.15]	0.37	.71
Compatibility	−0.22	0.04	0.80	[0.74, 0.87]	−5.21	1.86E–07
Side	0.08	0.08	1.08	[0.93, 1.27]	0.99	.32
Sound stimulus	−0.14	0.08	0.87	[0.74, 1.02]	−1.74	.08
Sex	0.01	0.08	1.01	[0.86, 1.18]	0.15	.88
Age	−0.01	0.02	0.99	[0.95, 1.03]	−0.41	.69
Session						
Linear trend	−0.04	0.10	0.96	[0.79, 1.17]	−0.43	.67
Quadratic trend	−0.03	0.09	0.97	[0.81, 1.16]	−0.28	.78
Cubic trend	−0.06	0.09	0.94	[0.79, 1.12]	−0.67	.50
Quartic trend	0.04	0.09	1.04	[0.87, 1.24]	0.45	.65
Compatibility × Task interactions						
Compatibility × Experimental Tasks vs. Control Tasks	−0.23	0.04	0.79	[0.73, 0.86]	−5.44	5.44E–08
Compatibility × Full Task vs. Joint Task	0.09	0.06	1.09	[0.97, 1.23]	1.59	.11
Compatibility × Half Task vs. Joint Control Task	0.00	0.06	1	[0.89, 1.12]	0.05	.96

Note: CI = confidence interval.

Table 4. Results of the Generalized Linear Mixed Model With Looking at Partner as the Response Variable

Fixed effect	Estimate	*SE*	Odds ratio	95% CI	*z*	*p*
Intercept	−0.01	0.20				
Joint task vs. joint control task	−0.95	0.15	0.39	[0.29, 0.52]	−6.12	9.15E−10
Compatibility	−0.35	0.08	0.70	[0.60, 0.82]	−4.23	2.30E−05
Sex	−0.25	0.14	0.78	[0.59, 1.02]	−1.82	.07
Age	0.02	0.03	1.02	[0.96, 1.08]	0.57	.57
Compatibility × Task interaction	−0.05	0.15	0.95	[0.71, 1.28]	−0.33	.74

Note: CI = confidence interval.

$p < .001$ (Fig. 6a; Table 4). Moreover, they were also more likely to engage in mutual gaze in the joint task, $\chi^2(5) = 13.99$, $p < .05$ (Figs. 6a and 6b; Table 5).

Discussion

We found that common marmosets show a robust joint Simon effect. This suggests that action corepresentation is not unique to humans but is also present in another highly cooperative primate. The effect was detectable only when the partner was coengaged in the task and was not elicited by its mere presence (in the joint control task). Furthermore, before making a behavioral decision, the marmosets were more likely to look back at their partner and engage in mutual gaze in the joint task, and they were less likely to do so in the joint control task, where the partner was not coengaged in the task. Together, these results suggest that the mere presence of

a partner was not sufficient to elicit the joint Simon effect. Rather, what was needed was a partner that was coengaged in the task. Therefore, the joint Simon effect in this study appears not to be the result of low-level perceptual processes only but of coengagement in the task.

The presence of a joint Simon effect in marmosets may suggest that action corepresentation does not require having a full-fledged theory of mind, because this is unlikely in these monkeys (Burkart & Heschl, 2007). Nevertheless, they have some understanding of others as intentional agents (Burkart et al., 2012), which thus supports the proposal of Stenzel et al. (2014) that such a perception of agency is sufficient for a joint Simon effect. However, recent results also suggest that primate theory-of-mind understanding may have been underestimated in the past. For instance, Krupenye, Kano, Hirata, Call, and Tomasello (2016) found false-belief understanding in great apes in an anticipatory-looking
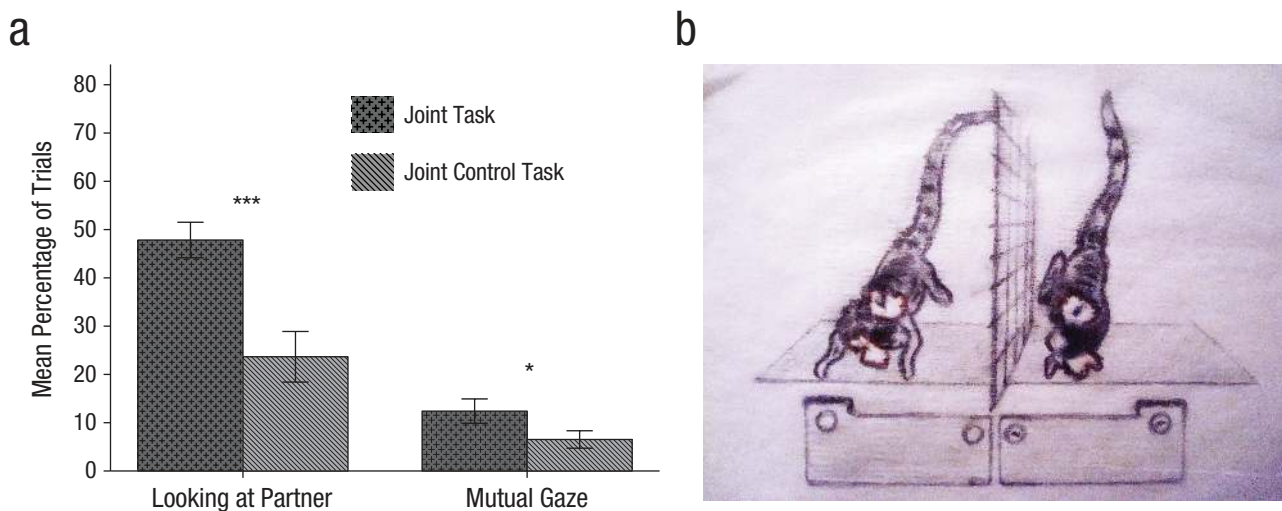
Fig. 6. Mutual gazing. The mean percentage of trials in which focal individuals looked at their partner or both partners looked at each other simultaneously (a) is shown as a function of task. Error bars represent standard errors of the mean. Asterisks indicate significant differences between tasks (*p < 05, ***p < .001). An example of mutual gazing prior to responding to the cue in a trial of the joint task (drawn from a still frame from a test video clip) is also shown (b).

Table 5. Results of the Generalized Linear Mixed Model With Mutual Gazing as the Response Variable

Fixed effect	Estimate	SE	Odds ratio	95% CI	z	p
Intercept	−1.74	0.63				
Joint task vs. joint control task	−0.67	0.20	0.51	[0.35, 0.76]	−3.36	7.9E−04
Compatibility	0.03	0.12	1.03	[0.81, 1.30]	0.29	.77
Sex	−0.38	0.29	0.68	[0.39, 1.21]	−1.28	.20
Age	−0.03	0.07	0.97	[0.85, 1.11]	−0.47	.64
Compatibility × Task interaction	0.12	0.20	1.13	[0.76, 1.67]	0.62	.53

Note: CI = confidence interval.

paradigm, similar to findings in young human children. To what extent primates, including marmosets, possess theory-of-mind abilities therefore continues to be an open question, as well as the role of such abilities in producing the joint Simon effect.

The marmosets in our task obtained a reward for each correct response. In human studies, participants are given a verbal instruction and then solve the task without reward or sometimes with an acoustic and often with a visual feedback stimulus indicating a correct or wrong response (Lien et al., 2015; Stenzel et al., 2012; joint flanker: Atmaca et al., 2011). In the joint condition, the marmosets were jointly rewarded for a correct response. This was done to increase the perception of the task as a joint task and may differ from studies with humans, in particular when no feedback stimulus is used. If visual or acoustic feedback is used, it is available to both partners, which is arguably closer to the marmoset situation in our study. Both with and without feedback, however, a correct solution may well be perceived as rewarding by both human partners in the joint task, namely when the verbal instruction or task setting implicitly or explicitly prompts the perception of the task as a shared goal. To fully understand the influence of the reward structure on the Simon effect in both humans and marmosets, researchers will need to conduct additional studies that systematically vary this factor.

The result that marmosets in the joint compared with the joint control task looked more at their partner and engaged in mutual gaze prior to responding is intriguing because a directed gaze is typically perceived as a threat in nonhuman primates (Coss, Marks, & Ramakrishnan, 2002). The gaze pattern in marmoset monkeys, however, may suggest the working hypothesis that, like humans (Tomasello, Hare, Lehmann, & Call, 2007; Wyman et al., 2013) but unlike great apes (Bullinger, Wyman, Melis, & Tomasello, 2011; Duguid et al., 2014), marmosets may use mutual gaze to coordinate their behaviors. Such a propensity to pay close attention to a cooperation partner and its attention could have evolved in marmosets to support the routine coordination of behavior in space and time in their daily lives, in particular during infant transfers (Snowdon, 2001). Our results, together with the propensity of marmosets to engage in cooperative vocal turn taking (Takahashi et al., 2013), suggest that they are more than accidental cooperators but indeed engage jointly in a cooperative task.

In sum, we demonstrated that the joint Simon effect and thus corepresentation during joint task performance is not unique to humans. This raises the question of the evolutionary origin of corepresentation and why we find it in both humans and marmoset monkeys. One possibility is that it may have evolved convergently in humans and marmosets, because they are both true cooperative breeders and therefore routinely engage in cooperative activities that require coordinated actions (Burkart, Hrdy, & van Schaik, 2009; Burkart & van Schaik, 2016).

Another possibility is that it simply represents a primate universal and that dyads of any nonhuman primate species would also show a joint Simon effect, given that they are able to engage in a joint task. This scenario is not unlikely given the contribution of low-level attentional processes to the joint Simon task (Dolk et al., 2014). To further disentangle these alternatives, it is thus crucial to use a joint Simon task to test independently breeding nonhuman primates who do not regularly engage in close coordinated and cooperative interactions.

Action Editor

Steven W. Gangestad served as action editor for this article.

Author Contributions

F. M. Miss and J. M. Burkart conceived and designed the study. F. M. Miss collected the data. Both authors analyzed

the data, wrote the manuscript, and approved the final version of the manuscript for submission.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797618772046

Open Practices

Data and materials for this study have not been made publicly available. The design and analysis plans were not preregistered.

References

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

Bullinger, A. F., Wyman, E., Melis, A. P., & Tomasello, M. (2011). Coordination of chimpanzees (*Pan troglodytes*) in a stag hunt game. *International Journal of Primatology*, *32*, 1296–1310.

Burkart, J. M., & Heschl, A. (2006). Geometrical gaze following in common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, *120*, 120–130.

Burkart, J. M., & Heschl, A. (2007). Understanding visual access in common marmosets, *Callithrix jacchus*: Perspective taking or behaviour reading? *Animal Behaviour*, *73*, 457–469.

Burkart, J. M., Hrdy, S. B., & van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology*, *18*, 175–186.

Burkart, J. M., Kupferberg, A., Glasauer, S., & van Schaik, C. P. (2012). Even simple forms of social learning rely on intention attribution in marmoset monkeys (*Callithrix jacchus*). *Journal of Comparative Psychology*, *126*, 129–138.

Burkart, J. M., & van Schaik, C. P. (2016). Revisiting the consequences of cooperative breeding. *Journal of Zoology*, *299*, 77–83.

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

Coss, R. G., Marks, S., & Ramakrishnan, U. (2002). Early environment shapes the development of gaze aversion by wild bonnet macaques (*Macaca radiata*). *Primates*, *43*, 217–222.

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., & Liepelt, R. (2014). The joint Simon effect: A review and theoretical integration. *Frontiers in Psychology*, *5*, Article 974. doi:10.3389/fpsyg.2014.00974

Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K., & Tomasello, M. (2014). Coordination strategies of chimpanzees and human children in a stag hunt game. *Proceedings of the Royal Society B: Biological Sciences*, *281*. doi:10.1098/rspb.2014.1973

Ford, R. M., & Aberdein, B. (2015). Exploring social influences on the joint Simon task: Empathy and friendship. *Frontiers in Psychology*, *6*, Article 962. doi:10.3389/fpsyg.2015.00962

Hommel, B., Colzato, L. S., & Van Den Wildenberg, W. P. M. (2009). How social are task representations? *Psychological Science*, *20*, 794–798.

Humphreys, G. W., & Bedford, J. (2011). The relations between joint action and theory of mind: A neuropsychological analysis. *Experimental Brain Research*, *211*, 357–369.

Klempova, B., & Liepelt, R. (2016). Do you really represent my task? Sequential adaptation effects to unexpected events support referential coding for the joint Simon effect. *Psychological Research*, *80*, 449–463.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*, 110–114.

Lien, M.-C., Pedersen, L., & Proctor, R. W. (2015). Stimulus–response correspondence in go–nogo and choice tasks: Are reactions altered by the presence of an irrelevant salient object? *Psychological Research*, *80*, 912–934.

Mangold International. (2017). *INTERACT user guide*. Arnstorf, Germany: Author.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–324.

McClung, J. S., Jentzsch, I., & Reicher, S. D. (2013). Group membership affects spontaneous mental representation: Failure to represent the out-group in a joint action task. *PLOS ONE*, *8*(11), Article e79178. doi:10.1371/journal.pone.0079178

Milward, S. J., Kita, S., & Apperly, I. A. (2014). The development of co-representation effects in a joint task: Do children represent a co-actor? *Cognition*, *132*, 269–279.

Milward, S. J., Kita, S., & Apperly, I. A. (2016). Individual differences in children's corepresentation of self and other in joint action. *Child Development*, *88*, 964–978.

Osmanski, M. S., Song, X., & Wang, X. (2013). The role of harmonic resolvability in pitch perception in a vocal nonhuman primate, the common marmoset (*Callithrix jacchus*). *Journal of Neuroscience*, *33*, 9161–9168.

Osmanski, M. S., & Wang, X. (2011). Measurement of absolute auditory thresholds in the common marmoset (*Callithrix jacchus*). *Hearing Research*, *277*, 127–133.

Richerson, P., Baldini, R., Bell, A., Demps, K., Frost, K., Hillis, V., . . . Zefferman, M. (2016). Cultural group selection

plays an essential role in explaining human cooperation: A sketch of the evidence. *Behavioral and Brain Sciences*, *39*, 1–68.

Ruissen, M. I., & de Bruijn, E. R. A. (2015). Is it me or is it you? Behavioral and electrophysiological effects of oxytocin administration on self–other integration during joint task performance. *Cortex*, *70*, 146–154.

Ruissen, M. I., & de Bruijn, E. R. A. (2016). Competitive game play attenuates self-other integration during joint task performance. *Frontiers in Psychology*, *7*, Article 274. doi:10.3389/fpsyg.2016.00274

Ruys, K. I., & Aarts, H. (2010). When competition merges people's behavior: Interdependency activates shared action representations. *Journal of Experimental Social Psychology*, *46*, 1130–1133.

Saby, J. N., Bouquet, C. A., & Marshall, P. J. (2014). Young children co-represent a partner's task: Evidence for a joint Simon effect in five-year-olds. *Cognitive Development*, *32*, 38–45.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*, 70–76.

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

Simon, J. R., & Rudell, A. P. (1967). Auditory S–R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*, 300–304.

Snowdon, C. T. (2001). Social processes in communication and cognition in callitrichid monkeys: A review. *Animal Cognition*, *4*, 247–257.

Stenzel, A., Chinellato, E., Bou, M. A., del Pobil, Á. P., Lappe, M., & Liepelt, R. (2012). When humanoid robots become human-like interaction partners: Corepresentation of robotic actions. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 1073–1077.

Stenzel, A., Dolk, T., Colzato, L. S., Sellaro, R., Hommel, B., & Liepelt, R. (2014). The joint Simon effect depends on perceived agency, but not intentionality, of the alternative action. *Frontiers in Human Neuroscience, 8*, Article 595. doi:10.3389/fnhum.2014.00595

Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. A. (2013). Coupled oscillator dynamics of vocal turn-taking in monkeys. *Current Biology*, *23*, 2162–2168.

Tomasello, M., & Gonzalez-Cabrera, I. (2017). The role of ontogeny in the evolution of human cooperation. *Human Nature*, *28*, 274–288. doi:10.1007/s12110-017-9291-1

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution*, *52*, 314–320.

Tsai, C.-C., Kuo, W.-J., Hung, D. L., & Tzeng, O. J. L. (2008). Action co-representation is tuned to other humans. *Journal of Cognitive Neuroscience*, *20*, 2015–2024.

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

Voelkl, B., & Huber, L. (2007). Imitation as faithful copying of a novel technique in marmoset monkeys. *PLOS ONE*, *2*(7), Article e611. doi:10.1371/journal.pone.0000611

Wyman, E., Rakoczy, H., & Tomasello, M. (2013). Non-verbal communication enables children's coordination in a "stag hunt" game. *European Journal of Developmental Psychology*, *10*, 597–610.