Joint Simon effects for non-human co-actors

Anna Stenzel[1] · Roman Liepelt[1]

Abstract Social interactions with non-biological agents and interactions with technical devices have become increasingly important over the last years. Recent studies investigating the interactions between humans and non-human agents showed rather inconsistent results. While the joint Simon effect (cSE) was found to be absent for non-human co-actors like virtual wooden hands, other studies showed pronounced cSEs when the co-actor was a real event-producing object. However, an often overlooked difference between these studies is the way these co-actors delivered response events. Studies replacing the co-actor by event-producing objects used a continuous response mode, while in studies using wooden hands, the co-actor always produced action effects in a task-related, turn-taking mode. In a series of four experiments, we systematically tested the effects of the response mode on the size of the cSE. The cSE was larger when the co-actor produced events in a turn-taking response mode than in a continuous response mode. Furthermore, we consistently found reliable cSEs for different kinds of virtual non-human co-actors (including a Japanese waving cat, scrambled patterns, and a wooden hand), and found no difference in the size of the cSE between human and non-human co-actors. We discuss possible mechanisms explaining why a cSE might be present or absent when sharing tasks with virtual non-human co-actors.

Interacting with other humans is an essential part of our daily life, and even great apes understand the basics of intentional action of their conspecifics (Tomasello, Carpenter, Call, Behne, & Moll, 2005). Due to the fast technological progress, interaction situations with electronic devices, virtual agents, and even with real robots have become an increasing part of our life. The investigation of the cognitive processes underlying interactions with human and non-human agents or devices has therefore become a new focus of psychological research.

Many studies investigating joint action between two humans or between a human and a non-human agent used a socialized version of the standard Simon task (Simon, 1969; Simon & Rudell, 1967) – the joint go/nogo Simon task (Sebanz, Knoblich, & Prinz, 2003). In the standard Simon task, one of two stimuli (e.g., a green or a red dot) is presented on the left or right side of a monitor. The task for the participant is to press a left button when one of both stimuli (e.g., the green dot) appears, and to press a right button when the other stimulus (e.g., the red dot) appears. Responses are usually faster when the position of the stimulus and the response position correspond (compatible trial) than when both positions do not correspond (incompatible trial) even though stimulus position is fully irrelevant for solving the task. The difference in response times between compatible and incompatible trials is called the Simon effect (Simon, 1990). In a joint go/nogo Simon task, two participants share a standard Simon task so that each of them performs a go/nogo task responding to only one of both stimuli. For example, the person sitting on the left side of a monitor (pressing the left response button) responds to the green stimulus, whereas the person on the right (pressing the right response button) responds to the red stimulus. In

✉ Anna Stenzel
anna.stenzel@uni-muenster.de

✉ Roman Liepelt
roman.liepelt@uni-muenster.de

[1] University of Muenster, Institute for Psychology, Fliednerstraße 21, 48149 Münster, Germany

this joint go/nogo task, a Simon effect is usually observed. This has been called the joint go/nogo Simon effect (JSE), or "cSE" (Dolk, Hommel, Prinz, & Liepelt, 2013), as go/nogo tasks are classified as "type c" tasks according to a typology proposed by Donders (1969). Remarkably, no Simon effect is observed when one single participant performs the same go/nogo task alone (individual go/nogo Simon task). As a Simon effect is typically present in joint and standard Simon tasks but not in individual go/nogo Simon tasks, Sebanz et al. (2003) concluded that each person in the joint condition represents the action of the co-actor just as if being in charge of the action him or herself (action corepresentation account). Therefore, performing a go/nogo Simon task together with another person is functionally equivalent to performing the standard Simon task alone (Sebanz et al., 2003; Sebanz, Knoblich, & Prinz, 2005). Based on later results, Sebanz et al. (2005) further suggested that each participant not only corepresents the action the other person performs, but also corepresents the other person's task rule. This leads to an internal simulation of the other person's action whenever the target stimulus of the other person is perceived (task corepresentation account). It was speculated that action or task corepresentation might have evolved in order to facilitate interactions with other humans, so that both accounts predict that corepresentation (action or task) should only take place when sharing a task with another human co-actor. However, according to the recently proposed referential coding account (Dolk et al., 2011, 2013; Dolk, Hommel, Colzato, et al., 2014; Dolk, Hommel, Prinz, et al., 2014) a representation of the other's action or task rule is not needed in order to induce a cSE. Based on the theoretical framework of the theory of event coding (TEC; Hommel, Müsseler, Aschersleben, & Prinz, 2001), the referential coding account assumes that the action events generated by the participant and those generated by the other person are cognitively represented by the same kind of codes. Due to common coding, both persons sharing a task (e.g., the joint go/nogo Simon task) have the problem to cognitively discriminate between internally (one's own) and externally (the other person's) activated action events, as both events (typically simple button presses) are highly similar in a typical joint go/nogo Simon task. In this situation, response features that clearly discriminate between actions in a given task context need to be weighted more strongly (Memelink & Hommel, 2013) in the internal representation of actions. As the spatial position of actions (left and right) is a clearly discriminating feature in a spatial Simon task, actions are coded more strongly as left and right in the presence of the other responding person than when the other person is absent. Therefore, there is a higher dimensional overlap (Kornblum, Hasbroucq, & Osman, 1990) between spatial stimulus and response dimensions when performing the task together with another person than when performing the task alone, leading to a larger cSE in the joint than in the individual condition. Opposed to the action/task corepresentation account, the referential coding account predicts a cSE when sharing the task with a non-human agent or with an agent who does not perform the complementary part of the task (i.e., whose actions are not related to the alternative stimulus). According to referential coding, for a cSE to occur it is crucial that the perceived or imagined action of the other agent is sufficiently similar to one's own action to require action discrimination (Dolk, Hommel, Colzato, et al., 2014).

The action/task corepresentation account were related to the history of experimental evidence on the cSE, showing significant cSEs when sharing a Simon task with another human agent but not when performing the task with a non-human agent. In the core study on this subject by Tsai and Brass (2007), participants performed a joint go/nogo Simon task together with a human co-actor (i.e., a human hand looking like a real biological body part) and a non-human co-actor (i.e., a wooden hand in the form of a biological body part). Tsai and Brass (2007) used a virtual version of the joint go/nogo Simon task, in which participants placed their own right response hand on a response button located on the right side of a monitor while either a human or a wooden hand was displayed on the left side of the monitor. The participant and the virtual hand responded in a turn-taking mode to color stimuli presented at a left or right position of a central arrangement on the monitor. A cSE was only observed when the task was performed together with the human hand, but not with the wooden hand. The authors concluded that humans only corepresent the actions of biological agents, which is in line with the idea that corepresentation evolved in order to facilitate social interactions. Müller et al. (2011) replicated the basic finding of this study showing a cSE when participants performed the go/nogo Simon task together with a virtual human hand, but no cSE when performing the task with a virtual wooden hand. However, this was only true when participants had watched a short movie about a human actor prior to performing the task. After having watched a scene from the Disney movie *Pinocchio*, portraying Pinocchio acting in a human-like way, a cSE could even be observed for the wooden hand. These findings pointed to the importance of top-down factors, like the attribution of intentionality to an agent in modulating the cSE. Tsai, Knoblich, and Sebanz (2011) further specified the initial results obtained by Tsai and Brass (2007). In Experiment 1 of their study, participants performed a virtual version of the joint go/nogo Simon task in which participants placed their own hand on a response button located in front of a monitor. With this button they controlled a virtual hand that was displayed on the right side of the monitor (i.e., whenever the participant pressed the response button, the virtual hand on the monitor also pressed a button). On the left side of the monitor, another virtual hand was displayed that was controlled by the computer program. Both virtual hands (representing the participant and the virtual co-actor) had

either a human or wooden form. In line with the results by Tsai and Brass (2007), a cSE was observed when participants who controlled a human hand performed the task together with a human hand, but not when they performed the task with a wooden hand. However, when participants controlled a wooden hand, a cSE was observed when sharing the task with a virtual wooden partner, but was absent for the human hand. Tsai et al. (2011) interpreted their findings with regards to the perceptual similarity between both effectors. Whenever both effectors were perceptually similar a cSE was induced, while no cSE was observed for dissimilar effectors.

More recent studies showed that a cSE can, however, also be induced when replacing the human co-actor in a joint go/ nogo Simon task by a real object that does not look like a human hand and produces visual and/or auditory action events. The objects that have been used so far were a rotating wheel (Dolk et al., 2011, Experiment 3), a Japanese waving cat, a rotating clock, and a ticking metronome (Dolk et al., 2013). These results seem to contradict the results by Tsai and colleagues (2007, 2011) and Müller et al. (2011), who found no significant cSE when participants performed a go/nogo Simon task together with a non-human, wooden hand. In contrast, the study of Dolk et al. (2013) found reliable cSEs when participants performed the task in the presence of an object, that was clearly non-human and even less similar (with respect to the form information) to the participant's own hand than the wooden hand. Dolk et al. (2013) therefore explained their findings with the referential coding account, and not with the assumption of social action or task corepresentation.

A more detailed look at the studies investigating cSEs for non-human agents or event-producing objects shows that these studies do not only differ with respect to the body part information of the co-actor (biological vs. non-biological form), but also with respect to the way these agents delivered their respective response events. In studies using wooden hands (Müller et al., 2011; Tsai & Brass, 2007; Tsai et al. 2011), the co-actor produced action effects in a task-related, turn-taking mode responding whenever the alternative stimulus appeared (just like the typical response of a human co-actor). In contrast, in studies replacing the co-actor by a more abstract event-producing object (Dolk et al., 2011; Dolk et al., 2013), a continuous response mode was used. For example, the Japanese waving cat continuously waved its arm throughout the whole task, and the metronome constantly produced beep tones. This makes a noticeable difference in the frequency of event production. While the wooden hand responded on average only once in every second trial, objects like the Japanese waving cat produced multiple action events per trial. A higher frequency of action events might make externally perceived events more distracting as compared to a rare action event production. For example, a metronome ticking continuously might be

harder to ignore than a metronome only ticking from time to time. On the one hand, more distracting action events could potentially enhance the cSE by increasing the saliency of the spatial response dimension, and hence lead to a higher dimensional overlap between the stimulus and the response dimension.

On the other hand, a co-actor responding in a turn-taking mode is more similar to the participant – who also responds in a turn-taking mode – than a continuously event producing object. The higher similarity of actor and co-actor for turn-taking responses might not only be due to perceptual features (i.e., the frequency and stimulus-relatedness of responses), but also to more abstract features like agency. Humans have an innate sensitivity to communicative cues such as contingent, turn-taking reactions, which lead to the attribution of communicative agency (Csibra & Gergely, 2006; Gergely, Egyed, & Király, 2007). Therefore, turn-taking responses might enhance the ascription of agency to the co-actor, making him or her more similar to oneself with regards to agency, which has been shown to enhance the cSE (Stenzel et al., 2014). When actor and co-actor are more similar, the action discrimination problem is increased. This can be resolved by a spatial recoding of one's own action in the joint go/nogo Simon task (Dolk et al., 2013; Stenzel et al., 2014), which increases the cSE. Therefore, one could expect the opposite pattern of results: a larger cSE for turn-taking than for continuous responding.

The finding of significant cSEs for non-human co-actors would favor the referential coding account predicting cSEs for any (social or nonsocial) action event that is somehow similar to the participant's own action events. In contrast, the action/ task corepresentation account would predict an absence of the cSE for non-human co-actors based on its core assumption that only actions or task rules of other humans (or similar co-actors) are represented.

Another crucial difference between earlier studies testing the cSE with non-human co-actors concerns the virtual character of the co-actor. While the wooden hand stimulus was always of virtual nature (i.e., it was presented on a monitor), the objects used in the studies by Dolk et al. (2011, 2013) were always physically present.

The findings of studies using a wooden hand as a co-actor, and studies using objects have been taken as evidence either in favor of action/task corepresentation or referential coding accounts explaining the cSE by making strong claims about the underlying mechanisms of joint action. We think that a clearer answer with respect to the mechanisms underlying the cSE can only be drawn when understanding the factors that produced differential findings in previous studies. In the present study, we therefore aimed to systematically investigate the effects of the response mode (turn-taking vs. continuous responses) on the size of the cSE for virtual non-human and human co-actors.

Experiment 1

In Experiment 1, participants performed a modified version of the virtual joint go/nogo Simon task used by Tsai et al. (2011, Experiment 1). We orthogonally manipulated body form information of the co-actor (human vs. non-human) and the response mode (turn-taking vs. continuous responses). On the left side of the monitor, either a human hand or a Japanese waving cat was displayed as a virtual co-actor (see Fig. 1a). The human and non-human co-actor either responded in a continuous response mode like the objects used in the study by Dolk et al. (2013; e.g., the sequence showing

the button press of the human hand was continuously repeated throughout the whole task) or in a turn-taking response mode, which is the usual response mode in a joint go/nogo Simon task (e.g., the response sequence was only displayed when the co-actor's go stimulus appeared). On the right side of the monitor, a human hand was displayed that was controlled by the participant's button press (see Fig. 1a). That is, the right virtual hand pressed a button whenever the participant responded to his or her imperative stimulus.

The aim of Experiment 1 was to test if the presence of a cSE for non-human co-actors in previous studies (Dolk et al., 2013) was due to the use of a real task setup and the use of a
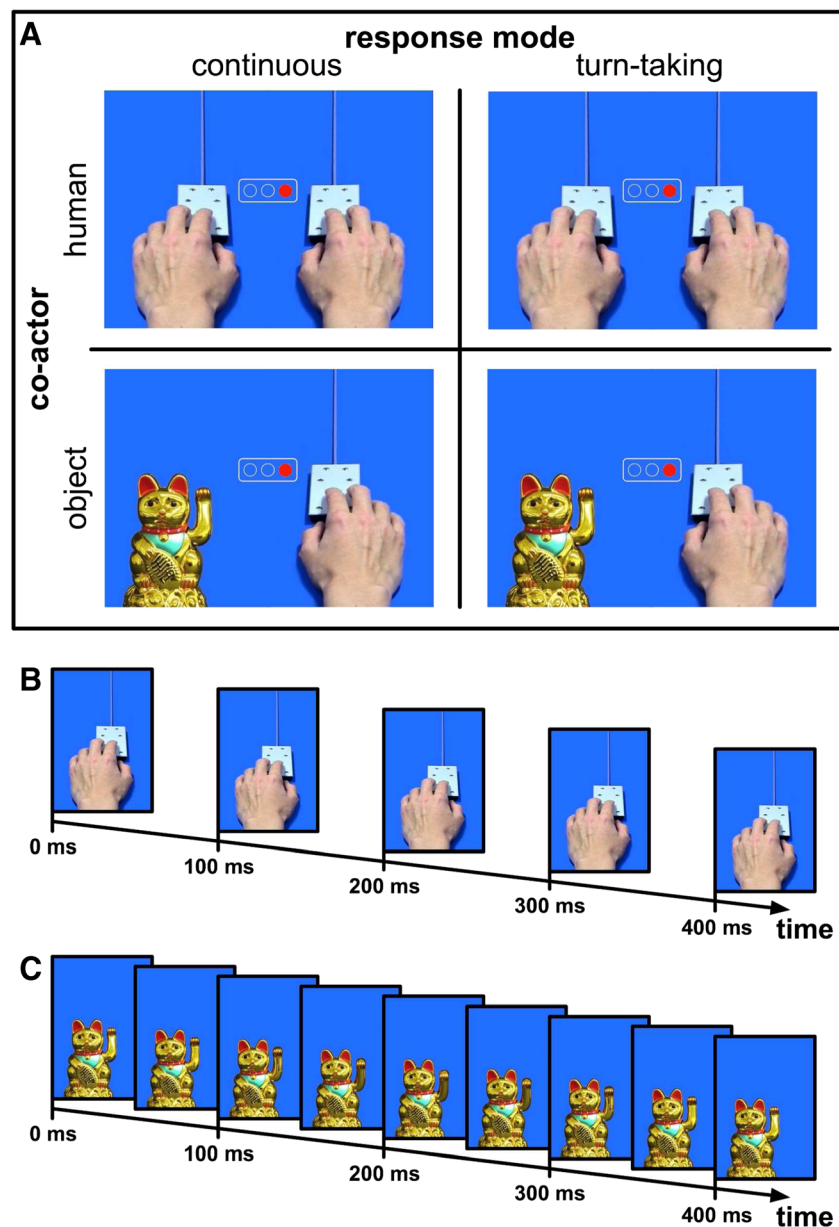
Fig. 1 Experimental design and stimuli used in Experiment 1 (**a**). Response sequence of the human hand participants controlled with their own button presses (**b**). The response sequence of the human co-actor was a mirror image thereof. Response sequence of the non-human co-actor (Japanese waving cat) (**c**)

continuous response mode. While the absence of the cSE for inanimate co-actors in virtual task setups has been taken as evidence for the action/task corepresentation account, the presence of the cSE for inanimate co-actors in real setups has been taken as evidence for the referential coding account.

If the presence of the cSE for inanimate co-actors was only due to the continuous responses mode applied in previous studies, one would expect a larger cSE for the continuous response mode than for the turn-taking response mode. Furthermore, one might expect a larger cSE for the human co-actor than for the non-human co-actor, as the human co-actor is perceptually more similar to the action events produced by the participant (controlling a virtual human hand) than the non-human co-actor. Based on previous results (Müller et al., 2011; Tsai & Brass, 2007, Tsai et al. 2011), one would even expect an absence of the cSE for the non-human co-actor responding in a turn-taking mode.

Method

Participants Thirty-two students of the University of Muenster participated in the experiment (6 male, M_{age} = 21.4 years, SD_{age} = 3.3 years, ranging from 18 to 30 years). The sample size was chosen based on previous studies including between-subject comparisons showing differences in the size of cSEs for sample sizes ranging from 16 (e.g., Dolk et al., 2013) to 38 participants ("biological condition" of the study by Müller et al., 2011), with most studies choosing sample sizes lying within this range (e.g., n = 18 in Experiment 1 by Tsai et al., 2011; n = 20 in the study by Tsai & Brass, 2007; n = 24 and n = 32 in the experiments by Liepelt, 2014). All participants were right-handed, had normal or corrected-to-normal vision, and were compensated with course credit points for their participation. They gave written informed consent to participate in the study, which was conducted in accordance with the ethical standards laid down in the 1975 Declaration of Helsinki.

Apparatus and stimuli For stimulus presentation we used Presentation software (www.neurobs.com). The stimulus program ran on a computer that was connected to a flat screen monitor. Stimuli were presented in a blue rectangular area of the monitor (27.6 cm × 19.9 cm), while the rest of the monitor remained black. Responses were recorded with a conventional response key that was placed 12 cm in front of the monitor and 5.5 cm to the right of its midline. Just like in the study by Tsai et al. (2011), the response hand of the participant was occluded from view by a white box while performing the task. The viewing distance was 60 cm.

Target stimuli and the human hands were taken from the study of Tsai and Brass (2007; see Fig. 1a). Target stimuli were a red and a green dot (1.1 cm in diameter) that were presented in a right or left location of a central gray

arrangement (4.9 cm × 2 cm). The left and right stimulus positions were located 1.5 cm away from the center. On the right side of the monitor, participants saw a right human hand (9 cm × 12 cm) resting on a response button. On the left side of the monitor, participants either saw a left human hand that was a mirror image of the right hand (human co-actor condition) or a Japanese waving cat (11.2 cm × 7.1 cm) from a front view (non-human co-actor condition; see Fig. 1a). The index finger of the human hands and the left paw of the Japanese cat were located at the same position on the screen (5.5 cm from the midline of the monitor).

The response sequence of the left and right human hand was a button press consisting of four images, starting with the index finger in an upper position, moving to a middle, and to the lowest position. After reaching the lowest position, the finger moved to the upper position again (see Fig. 1b). In the continuous response mode condition, this last upper position was the first position of the new action sequence. The response sequence of the cat was a waving cycle of its paw consisting of eight images, starting with an image of the paw in a back position moving consecutively to a front position and to the back position again (see Fig. 1c). Again, the last back position was the starting position of the new action sequence for the continuous response mode.

Task and procedure Participants performed a go/nogo Simon task together with another human hand (human co-actor condition) and a Japanese waving cat (non-human co-actor condition), which they saw on the left side of the monitor (see Fig. 1a). The action sequence of the co-actor on the left side was either looped so that the co-actor responded continuously (continuous response mode) or the co-actor only responded when the green dot appeared (turn-taking response mode). Each image of the human response sequence (consisting of four images) was presented for 100 ms, and each image of the non-human response sequence (consisting of eight images) was presented for 50 ms, so that the total duration of both response sequences was 400 ms.

The participant placed his or her own right hand on a response button that was located in front of the monitor on a table. The response button was placed 5.5 cm to the right of the midline of the monitor, so that it was located directly underneath the response button of the right hand displayed on the monitor. The participant's left hand was placed on his or her thigh. Whenever the participant pressed the response button, the response sequence of the right virtual hand was initiated. The hand of the participant was occluded from view by a white box. The red dot was the go signal for the participant, and the green dot was the go signal for the left co-actor in the turn-taking response mode conditions.

In the turn-taking response mode conditions, each trial started with the presentation of the left and right co-actor in their initial position for 500 ms. Afterwards, the target

stimulus (a red or green dot) appeared for 150 ms. Next, there was a response window of 2000 ms showing both co-actors in their initial position. When the green dot appeared, the response sequence of the left co-actor started after a random interval of 300, 350, 400, or 450 ms after stimulus offset. When the participant pressed his or her button during the response window, the response sequence of the right virtual hand was initiated. After the response window, there was a randomly chosen intertrial-interval (ITI) of 500, 1000, 1500, or 2000 ms, in which a white fixation cross was displayed on a black background. In the continuous response mode conditions, the trial sequence was the same as in the turn-taking response mode conditions except for the response sequence of the left co-actor, which was continuously repeated from the beginning of the trial to the end of the response window. Here, the co-actor did not respond to a certain stimulus but continuously performed the action (e.g., the human co-actor pressed the button once every 400 ms).

Each participant performed all four conditions (turn-taking/ human, continuous/human, turn-taking/non-human, continuous/non-human) block wise. The order of conditions was counterbalanced across the first 24 participants. For the remaining eight participants, eight orders were randomly selected with the constraint that the first condition of these orders was distributed uniformly (i.e., two out of these eight orders started with the human/continuous condition, two started with the human/turn-taking condition, etc.). Each block consisted of 160 trials in total, among them 80 go-trials for the participant (50 % stimulus-response [S-R] compatible and 50 % S-R incompatible). In the middle of each block and between blocks, participants were allowed to have a short break. At the beginning of each block, 14 practice trials were administered in order to acquaint participants with the task.

Participants were instructed to respond as fast and as accurately as possible. During the instruction phase of the experiment, the real Japanese cat, which had been photographed to create the stimuli, and the human hand of the experimenter were presented to the participant in order to connect the virtual co-actors on the screen to their real-life equivalents.

Results

Response time analysis Prior to statistical response time (RT) analysis, all go trials for the participant in which responses were incorrect (0.10 %), or 2.5 standard deviations faster or slower than the mean RT of each participant and condition (2.2 %) were excluded. In all experiments, we calculated a repeated measures analysis of variance (ANOVA) including the within-subjects factors co-actor (human, non-human), response mode (continuous, turn taking) and compatibility (compatible, incompatible). The results are shown in Fig. 2.

A significant main effect of compatibility, $F(1, 31) = 14.34$, $p < .001$, $\eta_p^2 = .32$, indicated that responses were faster for
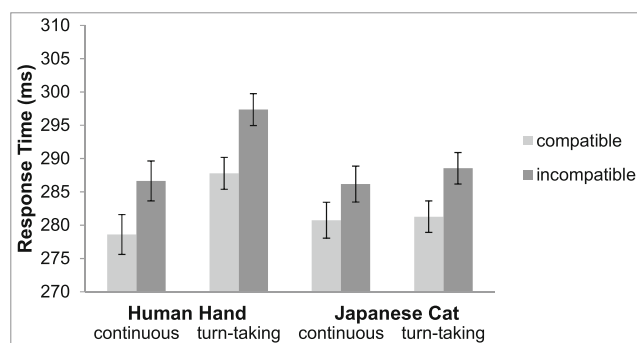
Fig. 2 Mean response times in Experiment 1 for compatible (*light gray*) and incompatible (*dark gray*) trials for the human hand (*left side*) and the Japanese cat (*right side*) plotted separately for continuous and turn-taking response modes. *Error bars* represent standard errors of the mean differences (Pfister & Janczyk, 2013)

compatible trials (282 ms) than for incompatible trials (290 ms). The interactions of co-actor × compatibility, $F(1, 31) = 1.47$, $p = .24$, $\eta_p^2 = .05$, and response mode × compatibility, $F(1, 31) = .80$, $p = .38$, $\eta_p^2 = .03$, were not significant, indicating no statistical difference between the compatibility effects of both co-actors (human: 9 ms; non-human: 6 ms), and both response mode types (turn taking: 8 ms, continuous: 7 ms). The main effect of response mode was significant, $F(1, 31) = 5.95$, $p = .02$, $\eta_p^2 = .16$, with slower RTs for the turn-taking response mode (289 ms) as compared to the continuous response mode (283 ms). The main effect of co-actor, $F(1, 31) = 2.78$, $p = .11$, $\eta_p^2 = .08$, as well as the interaction of co-actor × response mode, $F(1, 31) = 3.13$, $p = .09$, $\eta_p^2 = .09$, and the three-way interaction, $F(1, 31) = .01$, $p = .93$, $\eta_p^2 < .01$, were not significant.

Planned post hoc t tests revealed significant compatibility effects in the human/continuous: 8 ms, $t(31) = 2.71$, $p = .01$, the human/turn-taking: 10 ms, $t(31) = 3.99$, $p < .001$, and the non-human/turn-taking condition: 7 ms, $t(31) = 3.07$, $p < .01$. In the non-human/continuous condition, the compatibility effect missed the significance level: 5 ms, $t(31) = 2.02$, $p = .05$ (see Fig. 2).

Error analysis In all experiments, we calculated the same ANOVA used for the RT analysis for error rates. All main effects, two-way interactions, and the three-way interaction were not significant, all Fs$(1, 31) < 2.00$, all ps $> .16$, all $\eta_p^2 < .07$.

Discussion

In Experiment 1, compatibility effects were statistically comparable for both types of co-actors. We found a significant cSE for the human, as well as for the non-human co-actor. This finding is not in line with previous studies showing an absence of the cSE when the partner in a joint go/nogo task was non-human (Müller et al., 2011; Tsai & Brass, 2007) or when

two effectors presented on the screen were perceptually dissimilar (Tsai et al., 2011). Regarding the type of response mode, we observed statistically comparable compatibility effects for continuous and turn-taking responses. The finding of a significant cSE for a virtual non-human co-actor, which we observed for both response modes, suggests that the presence of a cSE for non-human co-actors is not restricted to continuous responses and a realistic task setup as used by Dolk et al. (2011; Dolk et al., 2013).

However, the Japanese cat might be a special kind of object as it does include some social features (a face, a body, and a response with its paw), which might make it more similar to the participant than other objects. As observed by Dolk et al. (2013), the cSE decreased in size the less features the action events of the object shared with the action events of the participant. Descriptively, Dolk et al. (2013) found the largest cSE for the Japanese cat, which included social, visual and auditory response features, and the smallest cSE for a metronome producing auditory events only. In Experiment 2, we aimed to generalize the findings of Experiment 1 to more abstract objects containing no social features.

Experiment 2

In Experiment 2, participants performed the same joint go/nogo Simon task as in Experiment 1. We replaced the Japanese cat by a scrambled pattern of the same size and color properties as the left human hand without containing the typical human body form (see Fig. 3). The response event of the scrambled pattern was produced by mixing the pattern's elements, so that the non-human response event was perceptually clearly dissimilar from the human response event.

Method

Participants A new set of thirty-two students of the University of Muenster participated in the experiment (6 male, M_{age} = 22.0 years, SD_{age} = 3.8 years, ranging from 19 to 34 years). All participants fulfilled the same criteria and were treated in the same way as participants in Experiment 1.

Stimuli The human hand stimuli were the same as in Experiment 1 (see Fig. 3a). For the non-human co-actor we used an oval section of a rectangular pattern of 19 × 25 small squares with similar color properties as the left human hand. The square was based on a rectangular picture of the left human hand. The rectangular picture was divided into 19 × 25 squares of equal size, the average color value was calculated for each of the squares, and afterwards the position of the squares was shuffled to produce an abstract response event. A total of 40 shuffled images were generated according to this procedure. For each image, only an oval section was taken

(see Fig. 3b). The final pictures were 8.9 cm × 11.8 cm in size, and were displayed at the same position as the left human hand.

The response sequence of the non-human scrambled pattern consisted of eight pictures that were shortly presented one after another for 50 ms each (see Fig. 3b). Each of these pictures in the response sequence was randomly chosen among the 40 shuffled pictures. The response sequence was either displayed after a variable interval of 300, 350, 400, or 450 ms after stimulus offset (turn-taking response mode) or was continuously repeated (continuous response mode).

Apparatus, task and procedure Apparatus, task and procedure were the same as in Experiment 1, with the exception that a printed image of the scrambled pattern was shown to participants instead of the Japanese waving cat during the instruction phase of the experiment. Participants were told that the position of the small squares within the pattern would be mixed either continuously or shortly after the green dot appeared. We did not inform participants, that the image was generated on the basis of the image of the left human hand.

Results

Response time analysis We again excluded all go trials for the participant in which responses were incorrect (0.05 %), or 2.5 standard deviations faster or slower than the mean RT of each participant and condition (2.3 %). The results are shown in Fig. 4.

A significant main effect of compatibility, $F(1, 31)$ = 41.88, p < .001, η_p^2 = .58, indicated that responses were faster in compatible trials (288 ms) than incompatible trials (297 ms). The size of the compatibility effect was significantly modulated by the response mode, $F(1, 31)$ = 5.12, p = .03, η_p^2 = .14, with larger compatibility effects for turn-taking (11 ms) than for continuous responses (7 ms). The compatibility effect was, however, not modulated by the type of co-actor, $F(1, 31)$ = .08, p = .78, η_p^2 < .01, showing statistically comparable compatibility effects for the human co-actor (9 ms) and the scrambled pattern (9 ms). The main factors of co-actor and response mode, as well as the interaction of co-actor × response mode, and the three-way interaction were not significant, all Fs(1, 31) < 1.75, all ps > .20, all η_p^2 < .06.

Planned post hoc t tests revealed significant compatibility effects in all four conditions-human/continuous: 7 ms, $t(31)$ = 3.88, p < .001; human/turn-taking: 12 ms, $t(31)$ = 5.80, p < .001; non-human/continuous: 7 ms, $t(31)$ = 2.85, p < .01; non-human/turn-taking: 10 ms, $t(31)$ = 4.92, p < .001 (see Fig. 4).

Error analysis The main effect of co-actor, $F(1, 31)$ = 3.69, p = .06, η_p^2 = .11, as well as the main effects of compatibility and response mode, all two-way interactions and the three-
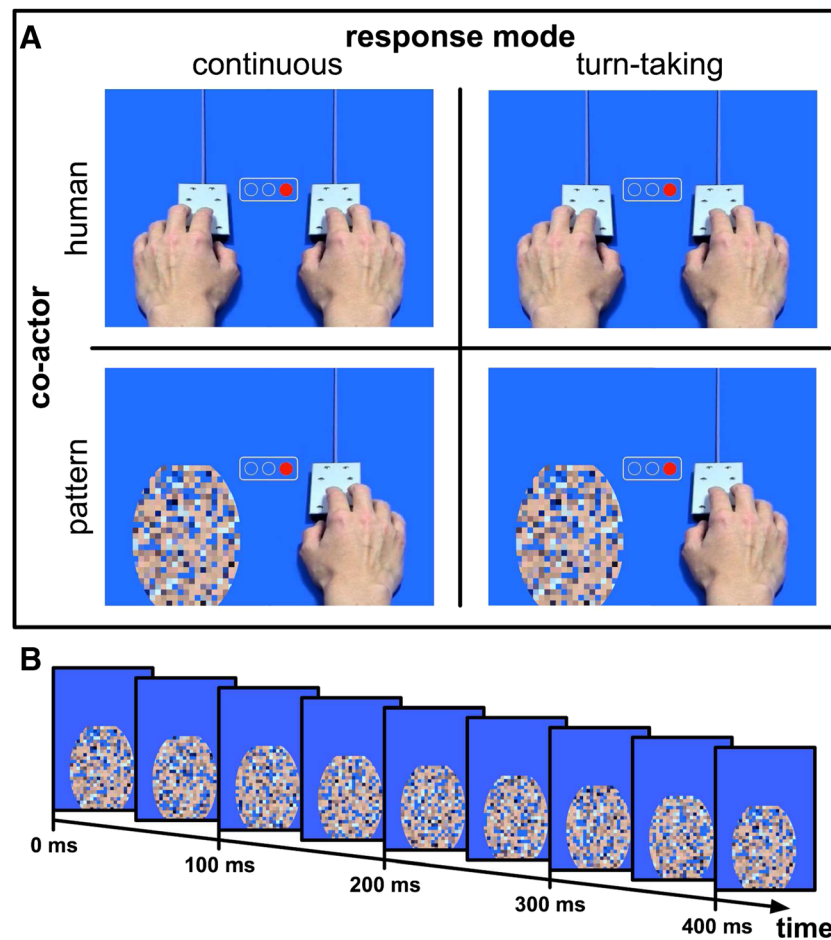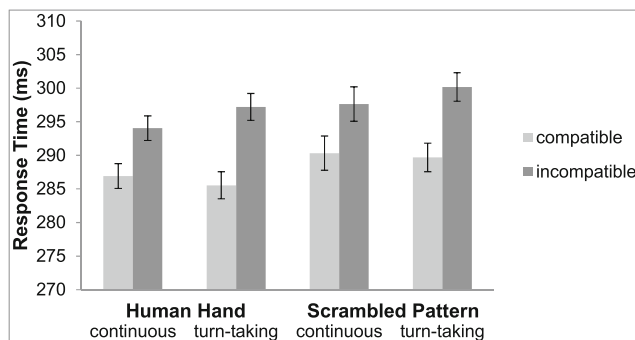
Fig. 3 Experimental design and stimuli used in Experiment 2 (**a**). Response sequence of the non-human co-actor (scrambled pattern) (**b**)

way interaction were not significant, all Fs(1, 31) < 1.30, all ps > .27, all η_p^2 < .04.

Discussion

Using a scrambled pattern (excluding any social features, and producing abstract response events), we found a larger

Fig. 4 Mean response times in Experiment 2 for compatible (*light gray*) and incompatible (*dark gray*) trials for the human hand (*left side*) and the scrambled pattern (*right side*) plotted separately for continuous and turn-taking response modes. *Error bars* represent standard errors of the mean differences (Pfister & Janczyk, 2013)

compatibility effect when co-actors responded in a turn-taking manner than when responding continuously. Again, reliable cSEs were observed for human and non-human co-actors, and cSEs were statistically comparable for both types of co-actors.

The results of our first two experiments suggest that reliable cSEs can be found for non-human virtual agents independent of whether they produce action events in a continuous or in a turn-taking, stimulus-related way. The significant increase of the compatibility effect in the turn taking as compared to the continuous response mode may suggest that the cSE in previous studies (Dolk et al., 2013) may have even been larger, if a turn-taking response mode would have been applied.

Experiment 3

In Experiment 3, we tested whether this pattern of results can also be observed for a non-human co-actor that is even less similar to the virtual effector controlled by the participant. The objects used so far in our experiments were present on the monitor throughout the entire trial and had about the same size as the human hand. In Experiment 3, the non-human

co-actor was not permanently visible on the screen, and covered only the size of the moving finger of the human hand. Participants performed the same joint go/nogo Simon task as in the previous experiments with the following exceptions. As a non-human co-actor, we now used a small scrambled pattern that was not permanently visible on the screen, but was only displayed when a response event was required. The scrambled pattern appeared whenever a green dot was displayed (turn-taking response mode) or appeared and disappeared with a constant interval (continuous response mode). We further tackled another question concerning the human co-actor conditions. In previous virtual versions of the joint go/nogo Simon task (Müller et al., 2011; Tsai et al., 2011), and in our first two experiments, a left virtual human hand was used as a human co-actor that was a mirror image of the right human hand. With these stimuli, the participant might get the impression that the two human hands on the monitor belong to a single person, which might facilitate the integration of both action events. In the human co-actor conditions of Experiment 3, we used pictures of two right hands from different persons as hand stimuli in order to test whether two effectors clearly belonging to two different persons can induce a cSE of about the same size as two effectors, which potentially belong to a single person.

Method

Participants A new set of thirty-two students from the University of Muenster participated in the experiment (10 male, M_{age} = 22.5 years, SD_{age} = 2.8 years, ranging from 19 to 30 years). They received course credit or a financial compensation for their participation. All participants fulfilled the same criteria and were treated in the same way as in the previous experiments.

Stimuli For the human hand stimuli, we took pictures of the right hands of two different persons (see Fig. 5a). As in the previous experiments, the response sequence of both hands consisted of four images. The hand displayed on the left side of the monitor (see Fig. 5b) had a size of 8.3 cm × 13.3 cm. The hand on the right side (see Fig. 5c) had a size of 9.1 cm × 13.4 cm. The index fingers of both hands were located 7.6 cm from the midline of the monitor.

For the non-human co-actor we used a rectangular pattern of 9 × 19 small squares with similar color properties as the index finger of the hand displayed on the left side. The pattern was based on a rectangular picture of the index finger of the left human hand. As in Experiment 2, the picture was divided into 9 × 19 squares of equal size, the average color value was calculated for each of the squares, and afterwards the position of the squares was shuffled. One shuffled image was generated according to this procedure (see Fig. 5a and d). The final picture

was 2.2 cm × 4.7 cm in size, and was displayed at the same position as the left index finger. The response sequence of the non-human co-actor consisted of eight pictures presented in succession for 50 ms each. The second picture of the sequence was the picture of the scrambled square in front of a blue background. All other pictures of the sequence were blank blue screens (see Fig. 5d). Thus different from Experiment 2, the scrambled square was not permanently visible on the screen, but was displayed for 50 ms after a green dot appeared (turn-taking response mode) or was displayed every 400 ms for 50 ms (continuous response mode).

Apparatus, task and procedure Apparatus, task and procedure were the same as in the previous experiments. During the instruction phase, a printed image of the scrambled pattern was shown to participants. As in Experiment 2, participants were not informed that the pattern was based on an image of the left index finger.

Results

Response time analysis We again excluded all go-trials for the participant in which responses were incorrect (0.09 %), or 2.5 standard deviations faster or slower than the mean RT of each participant and condition (2.1 %). The results are shown in Fig. 6.

A significant main effect of compatibility, $F(1, 31) = 36.81$, $p < .001$, $\eta_p^2 = .54$, indicated faster responses in compatible trials (289 ms) than incompatible trials (298 ms). The size of the compatibility effect was modulated by the type of response mode as indicated by a significant interaction of response mode × compatibility, $F(1, 31) = 7.20$, $p = .01$, $\eta_p^2 = .19$, with a larger compatibility effect when the co-actor responded in a turn-taking mode (11 ms) than in a continuous response mode (6 ms). Again, we found no statistical difference of the compatibility effects between the human co-actor (10 ms) and the non-human co-actor (7 ms), $F(1, 31) = 3.00$, $p = .09$, $\eta_p^2 = .09$. Further, the main effect of response mode reached significance, $F(1, 31) = 4.30$, $p = .05$, $\eta_p^2 = .12$, with faster RTs when the co-actor responded continuously (290 ms) than when responding in a turn-taking mode (297 ms). The main factors of co-actor, as well as the interaction of response type × movement and the three-way interaction were not significant, all $Fs(1, 31) < 1.50$, all $ps > .23$, all $\eta_p^2 < .05$.

Planned post hoc t tests revealed significant compatibility effects in all four conditions-human/continuous: 9 ms, $t(31) = 3.56$, $p < .01$; human/turn-taking: 12 ms, $t(31) = 5.11$, $p < .001$; non-human/continuous: 3 ms, $t(31) = 2.47$, $p = .02$; non-human/turn-taking: 11 ms, $t(31) = 4.83$, $p < .001$ (see Fig. 6).

Error analysis The main effect of response mode was significant, $F(1, 31) = 5.74$, $p = .02$, $\eta_p^2 = .16$, indicating fewer errors
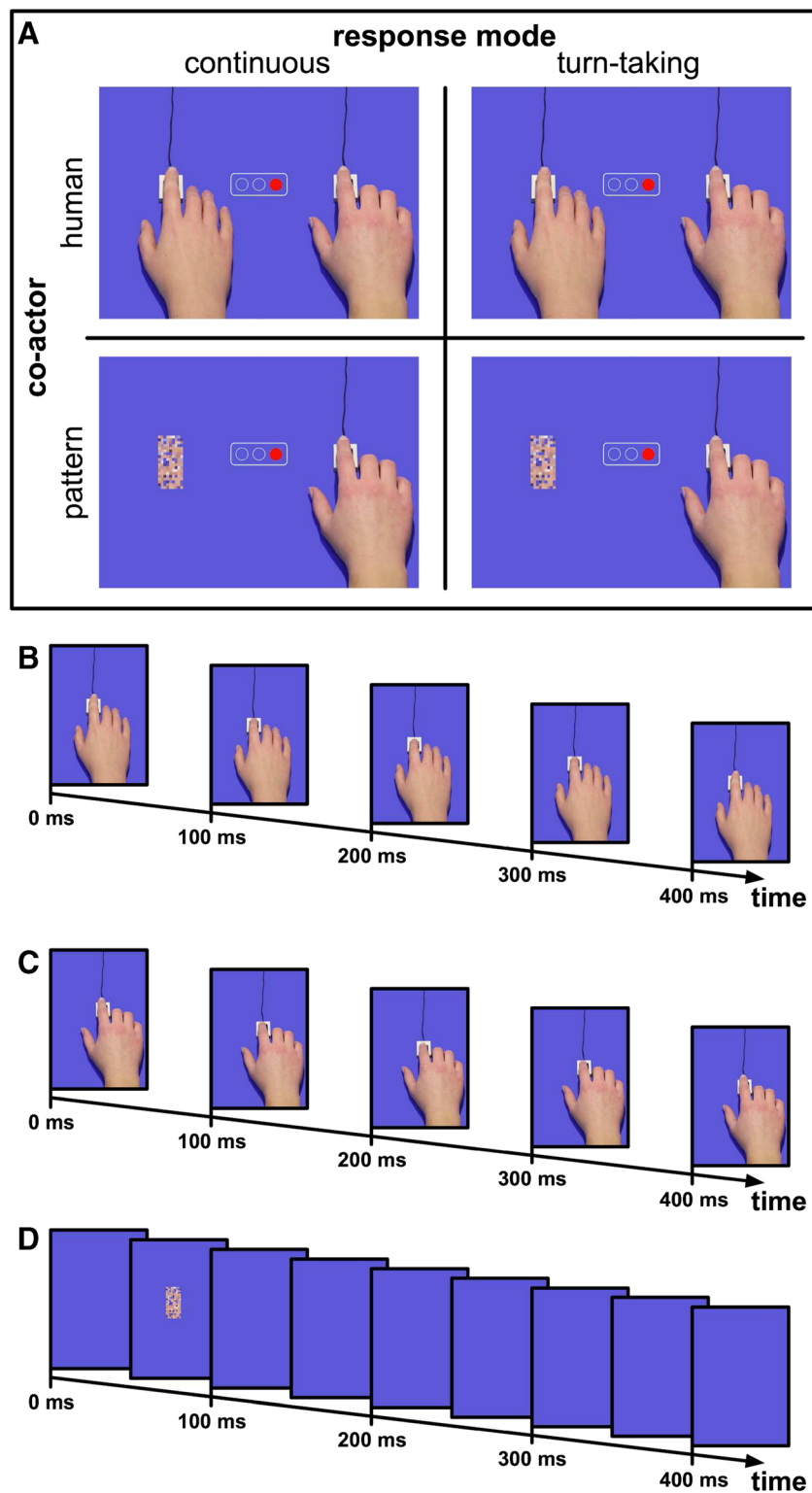
Fig. 5 Experimental design and stimuli used in Experiment 3 (**a**). Response sequence of the human co-actor (*human hand*). (**b**). Response sequence of the human hand participants controlled with their own button presses (**c**). Response sequence of the non-human co-actor (*scrambled pattern*) (**d**)

in the turn-taking conditions (0.3 %) than in the continuous response mode conditions (0.5 %). All other main effects, two-way interactions and the three-way interaction were not significant, all $Fs(1, 31) < 1.84$, all $ps > .18$, all $\eta_p^2 < .06$.
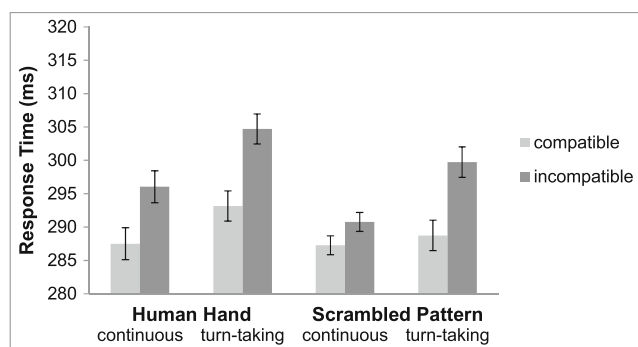
Fig. 6 Mean response times in Experiment 3 for compatible (*light gray*) and incompatible (*dark gray*) trials for the human hand (*left side*) and the scrambled pattern (*right side*) plotted separately for continuous and turn-taking response modes. *Error bars* represent standard errors of the mean differences (Pfister & Janczyk, 2013)

Discussion

Experiment 3 shows that even for an abstract, scrambled pattern that was not permanently visible on the screen a reliable cSE can be found. Like in Experiment 2, the cSE was larger in the turn-taking than in the continuous condition. The cSEs for human and non-human co-actors were again statistically comparable.

The cSEs for the human co-actor conditions (mean cSE = 10 ms) were comparable in size to the cSEs for the human co-actor conditions of Experiments 1 (mean cSE = 10 ms), and Experiment 2 (mean cSE = 11 ms). Hence, performing the joint go/nogo Simon task together with another human hand that clearly belongs to another person induces a similar cSE as performing the task together with another left human hand, which could be perceived as belonging to the same person.

Experiment 4

Across all three experiments, we did not find significant differences in the size of the cSE for human and non-human co-actors, which is not in line with previous studies showing a cSE for a human hand, but no cSE for a wooden hand. As these results shed some doubt on the reliability of previous findings on joint action with wooden hands, we aimed to directly replicate the pattern of results that was usually observed for the wooden hand. A wooden hand might be a special kind of non-human co-actor, as it is of similar biological form as a human hand, but clearly non-human in nature. This is not the case for the objects used by Dolk et al. (2013), which have a different shape than the human hand. Sharing a task with a non-human co-actor that is human-like in appearance but non-human in nature might be a confusing experience and might lead to an active suppression of the confusing element (i.e., the non-human wooden co-actor). When the visual

information on the left side is actively suppressed, actions are no longer spatially coded, and hence no cSE should be observed. This could be the reason why no cSE was observed for wooden hands in previous studies. In Experiment 4, we aimed to test the potential specificity of wooden hand stimuli. Participants performed the virtual version of the joint go/nogo Simon task together with a human hand and a wooden hand, which either responded continuously or in a turn-taking mode. Based on previous studies (Müller et al., 2011; Tsai & Brass, 2007; Tsai et al., 2011), we expected to observe a significant cSE for the human co-actor, but no cSE for the wooden co-actor.

Method

Participants A new set of thirty-two students from the University of Muenster participated in the experiment (5 male, M_{age} = 21.5 years, SD_{age} = 3.8 years, ranging from 18 to 35 years). All participants fulfilled the same criteria and were treated in the same way as participants in Experiment 1.

Stimuli The human hand stimuli were the same as in Experiment 3 (see Fig. 7a). The response sequence of the wooden hand consisted of pictures of a right wooden hand showing the wooden index finger in the same positions as the human hand (see Fig. 7b). The wooden hand was displayed on the left side of the monitor at the same position as the human co-actor and had a size of 8.1 cm × 13.2 cm.

Apparatus, task and procedure Apparatus, task and procedure were the same as in the previous experiments. During the instruction phase the wooden hand that was used to create the stimuli was shown to participants.

Results

Response time analysis We again excluded all go-trials for the participant in which responses were incorrect (0.11 %), or 2.5 standard deviations faster or slower than the mean RT of each participant and condition (2.7 %). The results are shown in Fig. 8.

We observed a significant main effect of compatibility, $F(1, 31) = 11.66$, $p < .01$, $\eta_p^2 = .27$, with faster responses in compatible trials (281 ms) than incompatible trials (287 ms). Compatibility effects were of identical size for the human co-actor (7 ms) and the wooden co-actor (7 ms), as indicated by a non-significant interaction of co-actor x compatibility, $F(1, 31) < .01$, $p = .96$, $\eta_p^2 < .01$. The size of the compatibility effect was, however, significantly modulated by the type of response mode, $F(1, 31) = 5.58$, $p = .03$, $\eta_p^2 = .15$, showing a larger compatibility effect when the co-actor responded in a turn-taking mode (9 ms) than when responding continuously (5 ms). Further, we found a significant main effect of response
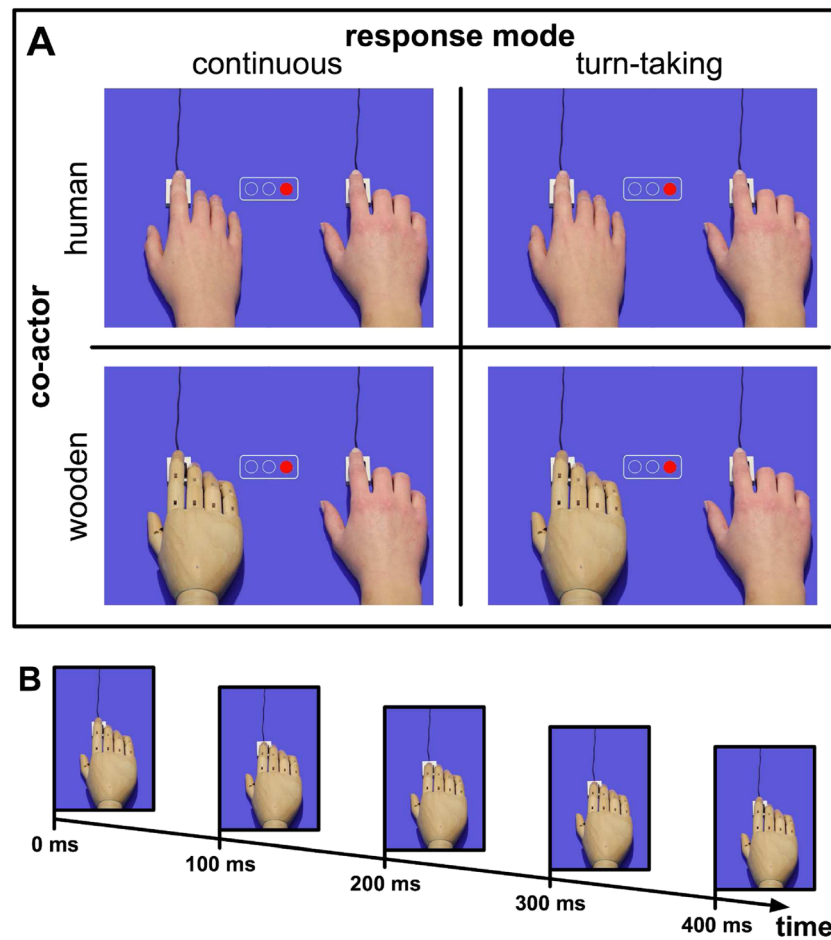
Fig. 7 Experimental design and stimuli used in Experiment 4 (**a**). Response sequence of the non-human co-actor (*wooden hand*) (**b**)

mode, $F(1, 31) = 5.13$, $p = .03$, $\eta_p^2 = .14$, with faster RTs when the co-actor responded continuously (281 ms) than when responding in a turn-taking mode (288 ms). The main factor of co-actor as well as the interaction of response mode × co-actor and the three-way interaction were not significant, all Fs$(1, 31) < .60$, all ps $> .44$, all $\eta_p^2 < .02$.
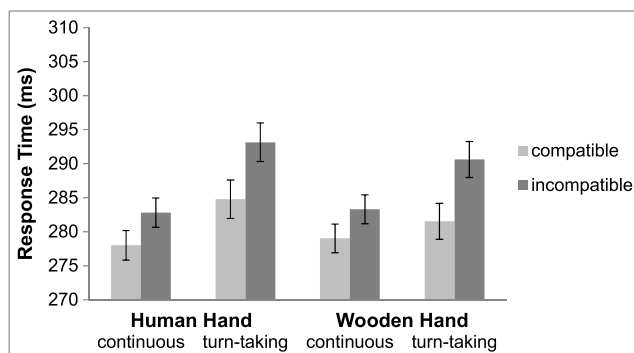
Fig. 8 Mean response times in Experiment 4 for compatible (*light gray*) and incompatible (*dark gray*) trials for the human hand (*left side*) and the wooden hand (*right side*) plotted separately for continuous and turn-taking response modes. *Error bars* represent standard errors of the mean differences (Pfister & Janczyk, 2013)

Planned post hoc t tests revealed significant compatibility effects in the human/continuous (5 ms), $t(31) = 2.22$, $p = .03$, the human/turn-taking (8 ms), $t(31) = 2.96$, $p < .01$, and the non-human/turn-taking condition (9 ms), $t(31) = 3.45$, $p < .01$. The compatibility effect in the non-human/continuous condition slightly missed the significance level (4 ms), $t(31) = 2.02$, $p = .05$ (see Fig. 8).

Error analysis The interaction of response mode × co-actor, $F(1, 31) = 3.89$, $p = .06$, $\eta_p^2 = .11$, all main effects, remaining two-way interactions and the three-way interaction were not significant, all Fs$(1,31) < 2.23$, all ps $> .14$, all $\eta_p^2 < .07$.

Discussion

In line with the results of Experiment 2 and 3, we found significantly larger cSEs when co-actors responded in a turn-taking mode than when responding continuously. We again observed full-blown cSEs for the human and non-human co-actor, which were again of comparable size. This finding is at odds with the findings of previous studies using wooden hands as co-actors (Müller et al., 2011; Tsai & Brass, 2007;

Tsai et al., 2011) in which the cSE was typically absent for the wooden co-actor, but present for the human co-actor.

Our finding of a comparable cSE for human and wooden co-actors suggests that the seemingly special status of the wooden hand among other objects (with a rather human-like appearance, but being clearly non-human in nature) cannot be the cause for the absence of the cSE for wooden hands which was observed in previous studies (Müller et al., 2011; Tsai & Brass, 2007; Tsai et al., 2011).

Data analysis across all experiments

We conducted an additional RT-analysis across all four experiments, in order to clarify whether the presence of a significant interaction of response mode and compatibility is a stable finding across all experiments and types of co-actors used, and whether the absence of the significant interaction of co-actor and compatibility was due to a lack in statistical power. We calculated a repeated measures analysis of variance (ANOVA) across the data of all four experiments ($n = 128$) including co-actor (human, non-human), response mode (continuous, turn taking), and compatibility (compatible, incompatible) as within-subjects factors, and Experiment (1, 2, 3, 4) as a between-subjects factor (see Fig. 9).

The analysis yielded a significant main effect of compatibility, $F(1, 124) = 86.72$, $p < .001$, $\eta_p^2 = .41$, indicating that responses were faster for compatible trials (285 ms) than for incompatible trials (293 ms). The interaction of response mode × compatibility was significant, $F(1, 124) = 16.71$, $p < .001$, $\eta_p^2 = .12$, with a larger compatibility effect for the turn-taking (10 ms) than for the continuous response mode (6 ms). The interaction of co-actor × compatibility was not significant, $F(1, 124) = 2.47$, $p = .12$, $\eta_p^2 = .02$, indicating no statistical difference between compatibility effects for human (9 ms) and

non-human co-actors (7 ms). The main effect of response mode was significant, $F(1, 124) = 12.61$, $p < .001$, $\eta_p^2 = .09$, with slower RTs for the turn-taking response mode (291 ms) as compared to the continuous response mode (287 ms). The main effects of co-actor and experiment, all remaining two-way and three-way interactions as well as the four-way interaction were not significant, all Fs < 2.20, all ps $> .14$, all $\eta_p^2 < .05$.

Across all four experiments, significant compatibility effects were observed for all conditions—human/continuous: 7 ms, $t(127) = 6.03$, $p < .001$; human/turn-taking: 10 ms, $t(127) = 8.66$, $p < .001$; non-human/continuous: 5 ms, $t(127) = 4.58$, $p < .001$; non-human/turn-taking: 9 ms, $t(127) = 8.07$, $p < .001$ (see Fig. 9).

In order to better integrate our findings in the existing literature, Table 1 provides an overview of the sizes of the cSE for non-human co-actors that were observed in our study, and in previous studies. In the table, we also provide information about the sample sizes.

Discussion

The analysis across all four experiments confirmed that the cSE is larger for turn-taking than for continuous responses independent of the type of co-actor. We found no statistically significant difference in the size of the cSE between human and non-human co-actors, suggesting that the absence of a significant interaction of co-actor and compatibility we observed in the separate experiments was not due to a lack of statistical power.

General Discussion

Using a virtual version of the joint go/nogo Simon task, our findings provide evidence for a cSE when sharing a go/nogo Simon task with a non-human co-actor varying from more object-like or abstract (Japanese waving cat, abstract scrambled patterns) to more human-like (wooden hand) form characteristics. There was no difference in the size of the cSE for human and non-human co-actors. However, the way in which the co-actor delivered response events affected the cSE for both, human and non-human co-actors. We found larger cSEs when the co-actor responded in a turn-taking way, which is the usual way in which humans respond in the joint go/nogo Simon task, as compared to continuous responding. A reliable cSE was, however, present under both, turn-taking and continuous response mode conditions, suggesting that event production does not need to be task-related in order for a cSE to occur. Our findings allow for two main conclusions: Previous findings providing evidence of a cSE for event producing objects (Dolk et al., 2011, 2013) can be replicated and
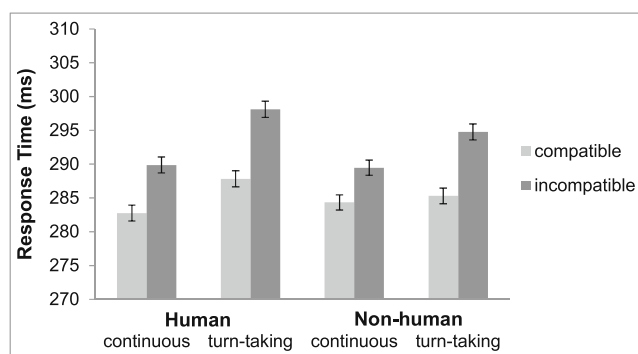
Fig. 9 Mean response times averaged across all four Experiments for compatible (*light gray*) and incompatible (*dark gray*) trials for the human hand (*left side*) and the non-human co-actor (*right side*) plotted separately for continuous and turn-taking response modes. *Error bars* represent standard errors of the mean differences (Pfister & Janczyk, 2013)

Table 1 Overview of Joint Simon effects (cSEs) observed for non-human co-actors or event-producing objects

Study	cSE(ms)	p	n	Response mode	Task setting
Tsai and Brass (2007)					
Dynamic condition	1[a]	ns	20	turn-taking	virtual
Tsai et al. (2011, Exp. 1)					
Different effector condition	2[a]	ns	18	turn-taking	virtual
Müller et al. (2011)					
Biological condition	4	ns	38	turn-taking	virtual
Present study					
Non-human/continuous condition	9	***	128	turn-taking	virtual
Non-human/turn taking condition	5	***	128	continuous	virtual
Stenzel et al. (2012)					
Machine-like robot condition	1	ns	24	turn-taking	real
Dolk et al. (2011, Exp. 3)					
Device present condition	9	*	20	continuous	real
Dolk et al. (2013)					
Metronome present condition	8	*	16	continuous	real
Clock present condition	10	**	16	continuous	real
Cat present condition	19	***	16	continuous	real

In the studies co-actors either responded in a turn-taking or in a continuous way and were either virtual or real. For virtual task settings, only conditions were included in which the participant controlled an effector with a human-like appearance

cSE joint Simon effect, *n* sample size, *ns* nonsignificant

[a] As the size of the cSE was not mentioned in the text, it was estimated on the basis of the graphical description of results

*$p < .05$. **$p < .01$. ***$p < .001$

extended to virtual task setups, and are not restricted to the use of a continuous response mode.

The finding of larger cSEs for turn-taking than continuous responses favors an explanation based on the perceived similarity between own and others' action events, with larger cSEs for perceptually similar events (i.e., when both actors respond in a turn-taking mode) than for dissimilar events (i.e., when both actors respond in different modes). The higher similarity for turn-taking responses might also be due to an enhanced similarity in agency (Csibra & Gergely, 2006; Gergely et al., 2007), which has been shown to modulate the cSE (Stenzel et al., 2014). A higher similarity between both actors leads to a larger action discrimination problem. The action discrimination problem can be resolved by a stronger intentional weighting (Memelink & Hommel, 2013) of discriminable action features like the spatial position of the response. A stronger spatial coding of the responses leads to increased dimensional overlap between stimulus features and response features, which increases the cSE (for a detailed description of the referential coding account see Dolk, Hommel, Colzato, et al., 2014).

We observed significant cSEs under continuous response conditions in which no task rule was present and the co-actor did not perform the complementary part of the go/nogo task. This finding is not in line with the task corepresentation account (Sebanz et al., 2005) assuming that a cSE should only be present under turn-taking response conditions, in which task rules can be used to predict the complementary responses of the co-actor.

At first glance, our findings seem to be at odds with other joint action studies using virtual task setups (Müller et al., 2011; Tsai & Brass, 2007; Tsai et al., 2011) showing an absence of the cSE when interacting with a non-human co-actor. The analysis across all four experiments including 128 participants supported our findings from the single experiments showing no difference in the cSE between human and non-human co-actors (a finding observed in all four experiments), but increased cSEs under turn-taking as compared to continuous response conditions (observed in Experiments 2, 3, and 4). Given the high statistical power of this analysis as compared to previous studies (Tsai & Brass, 2007: $n = 20$; Tsai et al., 2011, Experiment 1: $n = 18$; Müller et al., 2011, biological condition: $n = 38$) this difference in results seems stunning. But why might this be the case? One potential difference between our study and previous studies might concern the instruction of participants. A recent study showed that the presence or absence of a cSE critically depends on ascribing agency to the co-actor, i.e., on perceiving the co-actor as being the initiator of the action effect (Stenzel et al., 2014). When explicitly pointing out during task instructions how the virtual

co-actor is controlled by the computer, agency would be disrupted for this co-actor, and hence the cSE could be diminished. Another way to diminish the cSE in the virtual task setting might be to provide more information to participants portraying the non-human co-actor as clearly inanimate. Instructions about the animacy of the co-actor may be effective in disrupting referential spatial coding of responses by allowing a way to solve the action discrimination problem on a different level than space (Stenzel et al., 2012). For example, responses could be discriminated along the features agency vs. no-agency rather than right vs. left. As a result, spatial recoding of the participant's response would no longer be necessary, and hence the cSE would break down.

The present findings allow a new view on the cSE. Rather than inducing a cSE by providing task instructions about the humanness of a virtual co-actor (Müller et al., 2011; Tsai & Brass, 2007), the cSE seems to be typically present when sharing a task with a human or non-human co-actor. When focusing on conceptual differences between both actors during task instructions, the compatibility effect might decrease, as this allows other ways of resolving the action discrimination problem and hence disrupts spatial response coding.

Our finding that the cSE is modulated by the type of response mode but not by the type of co-actor suggests that the perceived similarity between both actors can also be induced by procedural aspects of the task, like the way in which action events are delivered, and not only by perceptual features of the event (Sellaro, Dolk, Colzato, Liepelt, & Hommel, 2015). These procedural aspects of the task relating to event production had a stronger impact on joint task performance than the physical appearance or humanness of the interaction partner. However, as the cSE is a relatively small effect compared to the standard Simon effect, we should note here that power issues might also play a role for the question of whether a cSE is statistically significant or not (for an overview of the size of the cSE for non-human co-actors across various studies see Table 1).

It might be possible that the cSEs we observed in our virtual task settings might not be caused by the same mechanisms as cSEs in real task settings, in which co-actors are physically present. As our virtual paradigm requires the integration of a nonbody-related distal action effect, the compatibility effects in our study might reflect action-effect compatibility (e.g., Hommel, 1993, 1996). In contrast to virtual task setups, real task settings may also require task corepresentation. However, a vast majority of the knowledge that we have about the social brain is based on virtual task setups using pictures or videos of faces, bodies or actions (see e.g., Press, 2011, for a review). A clear advantage of virtual setups is that experimental factors can be controlled more precisely than in real task settings. For example, real humans produce continuous effects such as breathing or body movements in addition to the relevant response events. As the present study not only shows that

human and non-human co-actors produce similar compatibility effects but also demonstrates that both effects are modulated in the same way by additional factors (here response mode), we believe that the same mechanisms are underlying both compatibility effects. This assumption is also supported by the theoretical account of referential coding, assuming that action coding and control are based on the anticipation of action effects, independent of whether the level of action coding is more proximal (visual, auditory or tactile feedback of the hand as in the joint go/nogo Simon task with two real persons) or more distal (visual feedback of the virtual hand as in the virtual joint Simon task). In future studies, effects of response mode on the cSE should nevertheless be investigated in real task settings where human and non-human co-actors are physically present even though the control of all other continuous effects may be more difficult than in virtual task setups.

Taken together, we consistently found reliable cSEs for virtual human and non-human co-actors. These results support the assumptions of the referential coding account allowing for a cSE for any co-actor (human or non-human) that produces action events which are sufficiently similar to the participant's own action events. While our findings showed no effects of the co-actor's humanness on the cSE, the way in which the co-actor produced action events modulated the size of the cSE with larger cSEs for turn-taking than for continuous responses. This finding can be explained by changes in the perceived similarity between internally generated and externally perceived events and the corresponding need to discriminate between these events during joint action.

References

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. Processes of change in brain and cognitive development. *Attention and Performance, XXI,* 249–274.

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., & Liepelt, R. (2011). How "social" is the social Simon effect? *Frontiers in Psychology, 2,* 84.

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,* 974.

Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2013). The (not so) social Simon effect: A referential coding account. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1248–1260.

Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2014). The joint flanker effect: Less social than previously thought. *Psychonomic Bulletin and Review, 21,* 1224–1230.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30,* 412–431.

Gergely, G., Egyed, K., & Király, I. (2007). On pedagogy. *Developmental Science, 10,* 139–146.

Hommel, B. (1993). Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research, 55,* 270–279.

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research, 59,* 176–186.

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,* 849–878.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility – A model and taxonomy. *Psychological Review, 97,* 253–270.

Liepelt, R. (2014). Interacting hands: The role of attention for the joint Simon effect. *Frontiers in Psychology, 5,* 1462.

Memelink, J., & Hommel, B. (2013). Intentional weighting: A basic principle in cognitive control. *Psychological Research, 77,* 249–259.

Müller, B. C. N., Brass, M., Kühn, S., Tsai, C.-C., Nieuwboer, W., Dijksterhuis, A., & van Baaren, R. B. (2011). When Pinocchio acts like a human, a wooden hand becomes embodied: Action co-representation for non-biological agents. *Neuropsychologia, 49,* 1373–1377.

Pfister, R., & Janczyk, M. (2013). Confidence intervals for two sample means: Calculation, interpretation, and a few simple rules. *Advances in Cognitive Psychology, 9,* 74–80.

Press, C. (2011). Action observation and robotic agents: Learning and anthropomorphism. *Neuroscience & Biobehavioral Reviews, 35,* 1410–1418.

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.

Sellaro, R., Dolk, T., Colzato, L., Liepelt, R., & Hommel, B. (2015). Referential coding does not rely on location features: Evidence for a non-spatial joint Simon effect. *Journal of Experimental Psychology: Human Perception and Performance, 41,* 186–195.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81,* 174–176.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility* (pp. 31–86). New York, NY: North-Holland.

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.

Stenzel, A., Chinellato, E., Tirado Bou, M. A., del Pobil, Á. P., Lappe, M., & Liepelt, R. (2012). When humanoid robots become human-like interaction partners: Co-representation 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,* 595.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28,* 675–735.

Tsai, C.-C., & Brass, M. (2007). Does the human motor system simulate Pinocchio's actions? *Psychological Science, 18,* 1058–1062.

Tsai, J. C. C., Knoblich, G., & Sebanz, N. (2011). On the inclusion of externally controlled actions in action planning. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1407–1419.