Context-Dependent Action Understanding: Automatic Integration of Contextual Information in Action Processing

Authors: Yang Yisong, Lin Jing, He Xiaoyan, Yin Jun, Yin Jun

Date: 2020-02-21T00:00:00+00:00

Abstract

Regarding the mechanisms of action understanding, simulation theory posits that the brain spontaneously simulates others' actions, with identical actions leading to identical understanding, whereas theory theory holds that people reason about others'actions based on the principle of rationality, such that the same action occurring in different contexts yields different understanding. However, the kinematic features of the action stimuli currently employed covary with the context in which they occur, making it difficult to distinguish whether action understanding supports simulation theory or theory theory. Through two experiments, animation techniques were employed to generate chasing actions in both constrained and unconstrained contexts, using EEG mu suppression as an index of action processing to test the aforementioned two perspectives. Specifically, in the constrained context, obstacles were present, requiring the chaser to alter its movement direction to bypass the obstacles and gradually approach the target from behind; whereas in the unconstrained context, no obstacles were present, yet the chaser maintained the same movement pattern as in the constrained scenario. The results revealed that when the chasing action occurred in a constrained context, where a clear action goal could be inferred based on the principle of rationality, the degree of mu suppression was greater than in the unconstrained context (Experiment 1); whereas when only the chaser moved, i.e., the approach target was uncertain, although the physical differences between constrained and unconstrained contexts remained identical to Experiment 1, the difference in mu suppression between conditions disappeared (Experiment 2). This mu suppression was not a generalization of attention-related occipital alpha activity. These findings suggest that contextual information in which actions occur influences people' s understanding of actions, implying that actions are understood through a reasoning process, thereby supporting the theory theory perspective.

Contextual Modulation of Action Interpretation: Automatic Integration of Situational Contexts During Action Understanding

YANG Yisong[1], LIN Jing[1], HE Xiaoyan[1], YIN Jun[1]* [1]Department of Psychology and Institute of Psychology, Ningbo University, Ningbo, 315211, China

Abstract

Regarding the mechanisms of action understanding, simulation theory posits that the brain spontaneously simulates others' actions, yielding identical understanding for identical actions, whereas theory theory proposes that people reason about others' actions based on rationality principles, such that identical actions occurring in different contexts produce different interpretations. However, existing action stimuli confound kinematic characteristics with contextual information, making it difficult to distinguish whether action understanding supports simulation theory or theory theory. Through two experiments, we employed animation technology to generate chasing actions in constrained versus unconstrained contexts, using EEG mu suppression (an index of action processing) to test these competing accounts. In the constrained context, obstacles required the chaser to alter direction to bypass them and gradually approach the target from behind; in the unconstrained context, no obstacles were present, yet the chaser maintained the same movement pattern as in the constrained condition. Results revealed that when chasing actions occurred in a constrained context—where a clear action goal could be inferred based on rationality principles—mu suppression was greater than in the unconstrained context (Experiment 1). However, when only the chaser moved (i.e., the approached target was ambiguous), the difference in mu suppression between conditions disappeared despite identical physical differences between contexts (Experiment 2). This mu suppression effect did not generalize to occipital alpha activity associated with attention. These findings indicate that contextual information influences action understanding, supporting an inferential process consistent with theory theory.

Keywords: action understanding; chasing action; context; simulation theory; theory theory; mu suppression

1 Introduction

Social interaction constitutes a vital component of human life (Miller, 2001). To interact smoothly with others, individuals must process others' actions to comprehend and identify their goals and intentions, thereby effectively predicting others' behavior and planning their own (Hauser & Wood, 2010; Wellman, Cross, & Watson, 2001). For instance, in soccer matches, players must continuously process opponents' dribbling actions to determine whether their purpose is passing or shooting. Impairment of this ability severely impacts daily life, as evidenced by autism spectrum disorder patients who exhibit deficits in processing and recognizing action intentions (Carter Leno et al., 2019; Klin, Jones, Schultz, & Volkmar, 2003). However, action processing presents a challenge: observers can only access physical-level kinematic information, yet must interpret the underlying goals and intentions—a process characterized by high uncertainty (Ullman, Baker, Macindoe, Evans, Goodman, & Tenenbaum, 2009; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016). For example, when Xiao Li continuously approaches Xiao Wang while walking, this could be interpreted either as Xiao Li chasing Xiao Wang or simply as Xiao Li hurrying and coincidentally moving closer to Xiao Wang. Consequently, how people understand actions has remained a central challenge in cognitive science, with implications for elucidating human mental systems and building socially capable robots (Blakemore & Decety, 2001).

Regarding the mechanisms of action understanding, one influential theoretical framework is simulation theory, also known as the direct matching hypothesis (Gallese & Goldman, 1998; Rizzolatti & Craighero, 2004; Ye, 2016). This theory proposes that individuals automatically match others' actions with their own motor representations through spontaneous internal simulation, such that action kinematics determine action understanding. Primary supporting evidence comes from the discovery of mirror neurons/systems (Rizzolatti, Fogassi, & Gallese, 2001). Di Pellegrino et al. (1992) found that when macaques observed others' actions, activation patterns in premotor cortex matched those produced when the monkeys executed the same actions themselves. Subsequent research identified mirror neurons as the neural basis for action processing, implementing action understanding through this mirroring simulation mechanism (Fogassi et al., 2005; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). Similar neural substrates —termed the mirror neuron system—were later identified in humans, comprising primarily the inferior parietal lobule and inferior frontal gyrus (or premotor cortex) (Iacoboni et al., 1999; Rizzolatti & Craighero, 2004). Researchers also discovered that during social interaction, people automatically mimic others' postures, expressions, and speech—the "chameleon effect" (Chartrand & Bargh, 1999). Moreover, observing others' finger or arm movements that are inconsistent with one' s own actions interferes with action execution (Dijkerman & Smit, 2007; Kilner, Paulignan, & Blakemore, 2003). These findings demonstrate that observing others' actions activates one' s own corresponding motor representations, automatically generating mimicry. Critically, although 3-month-old infants generally struggle to understand action goals, providing them with sufficient toy interaction opportunities (i.e., making the toy a target) or chances to perform grasping actions enables them to acquire action understanding abilities (Sommerville, Woodward, & Needham, 2005). Simulation theory' s matching process implies that observing others' actions activates one' s own corresponding motor representations, with more actions to match and simulate requiring greater cognitive resources—a prediction supported by recent findings that mirror system activation is stronger when processing multiple agents' actions compared to a single agent (Cracco, De Coster, Andres, & Brass, 2016). Collectively, these studies suggest that simulation theory' s direct matching process constitutes a key mechanism for action understanding. Since direct matching determines action understanding, actions with identical kinematic characteristics should yield identical understanding outcomes, whereas actions with different kinematics should produce different processing results (Csibra, 2007; Rizzolatti & Craighero, 2004).

However, researchers have challenged simulation theory because it struggles to explain certain activation patterns in mirror neurons or systems (Brass, Schmitt, Spengler, & Gergely, 2007; Csibra, 2007). Gallese et al. (1996) found that macaque mirror neurons did not activate when observing experimenters pretending to grasp objects that did not exist—contradicting simulation theory' s prediction, as pretend grasping and actual grasping share essentially identical kinematics. Human studies similarly revealed that for actions with nearly identical kinematics, mirror system activation was stronger when grasping targets were present versus absent (Koski et al., 2002). Thus, for actions with identical kinematics, understanding differed depending on whether they were transitive or intransitive. Furthermore, for grasping actions interpreted as different intentions based on contextual information, mirror system activation also varied: grasping a cup could be understood as drinking when filled with coffee, but as cleaning when empty with food residue present (Iacoboni et al., 2005). These results indicate that direct matching of action characteristics does not determine action understanding outcomes; rather, understanding is influenced by information beyond the action itself. Consequently, researchers proposed that people understand behavior not by matching observed actions with their own motor representations, but by reasoning to ascertain why the behavior was performed (i.e., its goal or intention), leading to theory theory of action understanding (Gergely & Csibra, 2003; Jacob & Jeannerod, 2005; Ye, 2016). This perspective holds that action understanding is based on rationality principles, assuming that actors execute minimally costly actions to achieve goals within current situations. Through inference, observers identify the most probable goal or intention for the observed action, with this process influenced by extra-action information. In other words, identical actions occurring in different contexts yield different understanding outcomes. This theory has garnered support from developmental psychology, cognitive neuroscience, and artificial intelligence (Brass et al., 2007; Gergely & Csibra, 2003; Jara-Ettinger et al., 2016). For example, when agent A approaches agent B via the shortest path (consistent with rationality principles), 10-month-old infants perceive A' s goal as approaching B, but when A approaches B via jumping movements, infants consider A' s purpose ambiguous (Csibra et al., 2003). In behavioral imitation, children also consider the context in which actions occur rather than purely mimicking others' behaviors (Gergely, Bekkering, & Király, 2002); this holds true for adults as well, as imitation of finger movements is modulated by whether the action is intentional or mechanically forced (Liepelt, Cramon, & Brass, 2008). In summary, theory theory' s inferential process can explain more phenomena in action understanding.

Nevertheless, simulation theory proponents argue that theory theory evidence can still be explained by simulation theory (Cavallo et al., 2016; Koul et al., 2018). First, transitive and intransitive actions do not share identical kinematics, so internal simulation processes differ, leading to different mirror system activation. Second, since observed actions are recorded from real humans, actions toward different targets involve different muscle control, potentially creating subtle kinematic differences that could be implicitly detected visually and affect action simulation. For example, grasping actions for drinking versus cleaning may exhibit subtle kinematic distinctions. Finally, for infants observing straight versus jumping actions, the movement patterns themselves differ, so matching yields different understanding outcomes. In essence, the debate between these theories centers on whether kinematic characteristics determine action understanding. Simulation theory claims that kinematics determine understanding, with any differences in understanding outcomes attributable to kinematic differences, independent of context. Theory theory proponents argue that because inferential processes apply rationality principles, even actions with identical kinematics produce different understanding when contextual constraints determine the rationality of goal approach (see reviews by Gergely & Csibra, 2003; Jacob & Jeannerod, 2005). Since current experimental stimuli confound kinematic characteristics with contextual information, it remains difficult to distinguish whether action understanding aligns with simulation theory or theory theory. Therefore, alternative methods are needed, such as strictly controlling kinematics while manipulating action contexts (Pomiechowska & Csibra, 2017), to further investigate action understanding mechanisms.

To manipulate action contexts while strictly controlling kinematic characteristics, we adopted Heider and Simmel’ s (1944) method of generating action patterns through animation technology. This approach simplifies agents into geometric shapes that can simulate everyday behavioral scenarios while effectively controlling extraneous variables, and has been widely applied in social cognition research (Powell & Spelke, 2013; Yin et al., 2016). We selected chasing actions, which played a crucial role in human evolution (Barrett, Todd, Miller, & Blythe, 2005), as our target for action understanding. The experiment included two contexts, each with two moving agents where one agent moved behind the other, creating a chasing pattern (see Figure 1 [Figure 1: see original paper]). Specifically, in the constrained context, agent A changed direction to bypass obstacles and gradually approached agent B from behind. In the unconstrained context, obstacles blocking agent A’ s shortest path to agent B were removed, yet agent A still exhibited the same movement pattern as in the constrained context, approaching agent B’ (which had the same movement pattern as agent B). Thus, agents’ kinematics were identical across contexts, differing only in whether obstacles prevented agent A from taking the shortest path to agent B. From a simulation theory perspective, since kinematics were identical across contexts, the action simulation process should be equivalent, predicting no difference in action processing outcomes. However, from a theory theory perspective, although kinematics were identical, in the constrained context agent A' s direction change could be understood as bypassing obstacles to most efficiently approach agent B—an action consistent with rationality principles. In the unconstrained context, agent A' approaching agent B' violated rationality principles (i.e., not taking the shortest path), yielding different understanding outcomes, with at least weaker goal-directedness toward agent B' . In short, simulation theory predicts identical understanding across contexts, whereas theory theory predicts different understanding, specifically that agent A is more likely to be perceived as targeting agent B in the constrained versus unconstrained context.

To test these predictions, we employed EEG technology, measuring mu rhythm (8-13 Hz) at sensorimotor cortex sites (C3 and C4 electrodes). The neural basis of this index lies in the mirror neuron system and even mentalizing systems (e.g., temporoparietal junction), with activation levels (manifested as mu suppression or desynchronization) varying according to action understanding and considered indicative of action processing (Fox et al., 2016; Muthukumaraswamy, Johnson, & McNair, 2004; Pomiechowska & Csibra, 2017; Ulloa & Pineda, 2007). Notably, mu suppression is greater when processing actions with clear goals versus ambiguous goals (Muthukumaraswamy et al., 2004). If action understanding follows simulation theory predictions, mu activation should be equivalent across contexts due to identical kinematics. If action understanding operates via theory theory mechanisms, mu suppression should be greater in the constrained context where rationality principles enable construction of a goal-directed relationship between the two agents.

2 Experiment 1

This experiment used chasing actions to manipulate action rationality through contextual characteristics, aiming to test whether action understanding follows simulation theory or theory theory mechanisms. We selected parietal mu rhythm as our index. Since the mu band shares the same frequency range (8-13 Hz) as occipital alpha associated with attentional mechanisms (Klimesch, Sauseng, & Hanslmayr, 2007; Perry, Stein, & Bentin, 2011), we also analyzed occipital alpha to test whether mu effects generalized from attention-related alpha activity.

2.1 Participants

Eighteen university students were recruited through advertisements to participate voluntarily. This sample size referenced previous mu studies investigating action processing, which used 16-18 participants and reported medium effect sizes (f between 0.25 and 0.40; Duan et al., 2018; Natraj et al., 2013; Pomiechowska & Csibra, 2017). Based on our repeated-measures design, we set alpha at 0.05, power at 0.80, and effect size at the medium level revealed in previous research (f 0.30; Duan et al., 2018; Pomiechowska & Csibra, 2017). G*Power 3.0.10 calculations indicated a planned sample size of approximately 18. Among the 18 participants, one was excluded due to poor EEG data quality from electrode detachment during recording, and two failed attention check criteria (including these three participants did not change the result pattern). The final sample comprised 15 valid participants (8 male, 7 female) aged 18–27 years (M = 22.67, SE = 0.77), with normal or corrected vision, no color blindness, and no physical or mental illness. Participants received 50 RMB as compensation. To maintain consistent sample sizes, we also recruited 15 valid participants for the subsequent experiment. The study was approved by the Ningbo University Psychology Research Ethics Committee, and all participants provided informed consent after being fully informed about experimental requirements.

2.2 Apparatus and Stimuli

The experiment used a 19-inch CRT monitor with 800×600 resolution and 100 Hz refresh rate; participants viewed the screen from approximately 60 cm. All stimulus presentation and experimental procedures were controlled via MATLAB' s Psychophysics Toolbox (Brainard, 1997).

Experimental stimuli were computer animations created with 3D modeling software Blender 2.78a, displayed at 11.1°×8.3° visual angle. Following Southgate and Csibra (2009), we created dynamic chasing actions (agent A chasing agent B; static schematic shown in Figure 1) and manipulated context by setting obstacles in the chasing scenario. All action videos lasted 3 seconds. Specifically, the larger agent A (2°×2°, the chaser) pursued the smaller agent B (1°×1°, the target) from behind at approximately 16°/s while the target moved toward the lower right screen area at about 14°/s. Agent shapes differed and were randomly selected from 36 different combinations (9 distinct shapes paired combinatorially) across trials. In the constrained context, obstacles blocked the chaser' s shortest path (i.e., straight line) to the target, forcing the chaser to detour around them, though the smaller target could pass through gaps between obstacles. In the unconstrained context, obstacles blocking the chaser' s shortest path were removed (relocated elsewhere), yet both chaser and target maintained the same movement patterns as in the constrained context. Thus, agents' kinematics were identical across the two animation types, with only the context differing.

In addition to these two animation types, we constructed filler animations requiring participants to count their occurrences to maintain attention on the screen. These fillers were identical to the experimental animations except that the chaser disappeared continuously for 1 second between 0.5-2.5 seconds during the chase.

Figure 1 Schematic of actions in constrained (a) and unconstrained (b) contexts

Given that action familiarity might affect mental simulation, we recruited 30 participants (18 male, 12 female, M = 22.20, SE = 0.48) before the formal experiment to rate familiarity with the two video types (rating order counterbalanced across participants) on a 7-point scale from 1 ("not familiar at all") to 7 ("very familiar"). Results showed no significant difference in familiarity ratings between constrained (M = 3.63, SE = 0.31) and unconstrained contexts

(M = 3.47, SE = 0.32), t(29) = 0.47, p = 0.643, Cohen' s d = 0.09. Additionally, to confirm subjectively that participants indeed understood the two animations differently, these participants rated agreement with the statement "To what extent do you agree that the smaller cartoon character is the target that the larger cartoon character is approaching?" on a 7-point scale from 1 ("strongly disagree") to 7 ("strongly agree"). Results revealed that participants were more likely to perceive the smaller agent as the larger agent' s target in the constrained context (M = 5.93, SE = 0.20) than in the unconstrained context (M = 4.83, SE = 0.35), t(29) = 3.20, p = 0.003, Cohen' s d = 0.59. This indicates that participants' understanding of identical actions differed across contexts, confirming effective manipulation.

2.3 Procedure

Each trial' s procedure is shown in Figure 2 [Figure 2: see original paper]. First, a dynamic fixation point appeared at screen center for 480 ms, followed by a blank screen for 300-400 ms. The first frame of the animation then appeared as a cue that the animation would begin. After 300 ms, the animation played for 3 seconds. The inter-trial interval was 900-1100 ms.

Figure 2 Experimental procedure

The 36 trial animations all featured different agent combinations, with identical combinations across conditions. Thus, aside from context, all other physical information was consistent between conditions. Each block included 8-12 filler animations. Participants were instructed to watch the animations carefully and count filler occurrences, reporting the count after each block. If reported filler counts were inaccurate, participants were considered not to have attended properly. Based on this attention check, two participants failed to report filler counts accurately and were excluded.

2.4 Data Recording

EEG was recorded using a NeuroScan Synamps 2 system (Compumedics NeuroScan Inc.) with a 64-channel Ag/AgCl electrode cap following the international 10-20 system. The left mastoid served as reference, with ground electrode located between FPZ and FZ. Vertical EOG was recorded from two electrodes above and below the left orbit, and horizontal EOG from two electrodes 1.5 cm lateral to the eyes. EEG signals were amplified with a gain of 500, sampled at 500 Hz, and band-pass filtered at 0.05-100 Hz. Electrode-scalp impedance was maintained below 5 kΩ. Given that baseline selection can affect mu sensitivity (Tangwiriyasakul, Verhagen, van Putten, & Rutten, 2013), we also recorded 1-minute baseline EEG data while participants viewed a computer screen before and after the formal experiment to obtain baseline EEG without action understanding.

2.5 Data Analysis

Data were analyzed using MATLAB' s EEGLAB and Fieldtrip toolboxes. Preprocessing in EEGLAB involved converting left mastoid reference to averaged mastoids, filtering data at 0.1-100 Hz (fir1), removing EOG electrodes, and performing independent component analysis (ICA). The ADJUST plugin in EEGLAB identified and removed ocular and noise components. The chasing action presentation period—from movement onset (animation start) to 3000 ms post-movement (animation end)—was then segmented. Fieldtrip was used for spectral analysis. Since the action understanding process occurred throughout the entire chasing animation without clear temporal characteristics, we followed previous studies (Brown, Wiersema, Pourtois, & Brüne, 2013; Perry et al., 2011) and used Fast Fourier Transform (FFT) to analyze 8–13 Hz power across the entire 3-second action understanding period. Each segmented data segment (i.e., 3000 ms animation duration) underwent FFT with 10% Hamming windowing to obtain power values (V^2) for each frequency band. For each trial, baseline correction was applied by subtracting power values from the 600–300 ms pre-animation baseline period (i.e., blank screen). Finally, data were averaged per participant and condition to obtain power values for different frequency bands.

Consistent with previous research, our mu band was 8–13 Hz (Fox et al., 2016; Muthukumaraswamy, Johnson, & McNair, 2004; Ulloa & Pineda, 2007). Numerous studies indicate that action-related mu rhythms primarily appear in parietal motor regions, with EEG studies commonly selecting parietal C3 and C4 electrodes (Duan et al., 2018; Pomiechowska & Csibra, 2017); therefore, we analyzed mu activation patterns at these two electrodes. To test whether mu effects generalized from occipital alpha activity, we similarly analyzed alpha at occipital O1 and O2 electrodes following previous research (Klimesch et al., 2007; Perry et al., 2011). To reveal mu specificity in indicating action understanding, we also analyzed beta band (14–40 Hz) at C3 and C4 electrodes, consistent with prior studies (Hobson & Bishop, 2016).

Additionally, given baseline selection's potential impact on mu sensitivity (Tangwiriyasakul et al., 2013), we segmented the pre- and post-experiment baseline EEG data into 3000 ms epochs (to match experimental analysis parameters), yielding 40 segments. These were also analyzed with FFT using 10% Hamming windows, with average power values across frequency bands serving as baseline values for subsequent analysis. For each participant, mu power during chasing action presentation was subtracted from this baseline difference value as the dependent variable. This analysis yielded the same pattern as using the 600–300 ms pre-animation baseline. Since previous research suggests baseline correction for each trial' s spectral power is more appropriate (Hobson & Bishop, 2016), we primarily report results using the 600–300 ms pre-animation baseline.

2.6 Results and Discussion

Parietal mu band power across conditions is shown in Table 1 and Figure 3a [Figure 3: see original paper]. Topographic maps of 8-13 Hz power across conditions and difference maps between conditions are shown in Figure 4 [Figure 4: see original paper]. Mu band power values were significantly lower than baseline (i.e., 0) across conditions, ts $>$ 3.23, ps $<$ 0.006, Cohen' s d $>$ 0.83, demonstrating mu suppression characteristics. A 2 (electrode: C3 vs. C4) × 2 (context: constrained vs. unconstrained) repeated-measures ANOVA on mu power revealed only a significant main effect of context, F(1, 14) = 12.51, p = 0.003, ^{2}p = 0.47, indicating significantly greater mu suppression (higher power values) in the constrained context (M = -1.955 V^2, SE = 0.567) than in the unconstrained context (M = -1.913 V^2, SE = 0.564). Neither the electrode main effect, F(1, 14) = 0.05, p = 0.835, ^{2}p $<$ 0.01, nor the electrode × context interaction, F(1, 14) = 0.08, p = 0.780, ^{2}p $<$ 0.01, were significant.

Table 1 EEG power values (V^2) for different measures in constrained and unconstrained contexts in Experiments 1 and 2

Occipital alpha band power across conditions is shown in Table 1 and Figure 3b. A 2 (electrode: O1 vs. O2) × 2 (context: constrained vs. unconstrained) repeated-measures ANOVA on occipital alpha power revealed no significant effects: electrode main effect, F(1, 14) = 1.56, p = 0.232, ^{2}p = 0.10; context main effect, F(1, 14) = 2.27, p = 0.154, ^{2}p = 0.14; and interaction, F(1, 14) = 0.01, p = 0.951, ^{2}p $<$ 0.01. These results suggest that mu power differences across animation types do not generalize from occipital alpha activity.

Figure 3 Parietal (a) and occipital (b) 8-13 Hz power values across conditions. Note: *p $<$ 0.05; error bars represent within-subject standard errors (Cousineau & O' Brien, 2014)

For the beta band (Table 1), a 2 (electrode: C3 vs. C4) × 2 (context: constrained vs. unconstrained) repeated-measures ANOVA revealed no significant effects: electrode main effect, F(1, 14) = 0.45, p = 0.514, ^{2}p = 0.03; context main effect, F(1, 14) = 2.75, p = 0.119, ^{2}p = 0.16; and interaction, F(1, 14) = 0.87, p = 0.368, ^{2}p = 0.06. This indicates that action understanding primarily activates the mu band.

To ensure our electrode selection did not drive results, we performed cluster-based permutation tests (Maris & Oostenveld, 2007) on 8-13 Hz power across all electrodes, using Monte Carlo simulation with 5000 permutations and requiring significant differences between constrained and unconstrained contexts at a minimum of two adjacent electrodes. Results confirmed significantly greater mu suppression in the constrained versus unconstrained context at C3 and C4 electrodes, with no differences at O1 and O2. Thus, ANOVA results cannot be attributed to electrode selection specificity.

These findings demonstrate that identical chasing actions elicited different mu activation depending on context, and that this effect did not generalize from alpha activity, suggesting that context influences action understanding. Specifically, mu suppression was significantly greater in the constrained context—where the chaser' s action could be clearly understood as goal-directed—compared to the unconstrained context, consistent with theory theory predictions.

3 Experiment 2

Experiment 1' s results aligned with theory theory predictions, showing greater mu suppression in the constrained context. However, in the constrained context of Experiment 1, the chaser was sometimes occluded by obstacles, resulting in different physical visibility durations compared to the unconstrained context. From a simulation theory perspective, these physical differences might produce different action simulation degrees, leading to inconsistent mu activation. Experiment 2 tested this possibility.

3.1 Method

The chaser maintained identical kinematics to Experiment 1, and we used the same contextual manipulation, but the target remained stationary at the movement endpoint. This design preserved the same physical differences between conditions as Experiment 1, but based on inferential reasoning, the chaser' s behavioral intention remained ambiguous before reaching the stationary target, making the rationality of its prior path unclear. Consequently, the degree to which the stationary object was perceived as an approach target should be equivalent across constrained and unconstrained contexts (Brass et al., 2007; Marsh et al., 2014; Southgate & Csibra, 2009). Therefore, if simulation processes underlie action understanding, the contextual modulation effect from Experiment 1 should persist when the target is stationary. However, if action understanding operates via theory theory' s inferential process, the contextual modulation effect should disappear, as reasoning outcomes would be similar across contexts.

Sixteen university students voluntarily participated, with one excluded for failing attention checks (including this participant did not change the result pattern). The final sample comprised 15 valid participants (7 male, 8 female) aged 19-26 years (M = 22.33, SE = 0.65), with normal or corrected vision, no color blindness, and no physical or mental illness.

The experimental design and procedure matched Experiment 1, except the target remained stationary at the movement endpoint. Data acquisition and analysis were identical to Experiment 1. As in Experiment 1, we recruited 30 participants (18 male, 12 female, M = 22.20, SE = 0.48) before the formal experiment to rate animation familiarity and action goal perception. Familiarity ratings did not differ between constrained (M = 3.67, SE = 0.35) and unconstrained contexts (M = 3.57, SE = 0.33), t(29) = 0.34, p = 0.735, Cohen' s d = 0.06. Moreover, participants perceived the stationary agent as the moving agent' s target equally across constrained (M = 4.87, SE = 0.35) and unconstrained contexts (M = 4.93, SE = 0.35), t(29) = 0.13, p = 0.896, Cohen' s d = 0.02. This indicates equivalent understanding of identical actions across contexts, confirming effective manipulation.

3.2 Results

Mu band power across conditions is shown in Table 1 and Figure 3a. Topographic maps of 8-13 Hz power and difference maps between conditions are shown in Figure 4. Mu power at midline electrodes was significantly lower than baseline, ts > 3.28, ps < 0.005, Cohen' s d > 0.84, again demonstrating mu suppression. A 2 (electrode: C3 vs. C4) × 2 (context: constrained vs. unconstrained) repeated-measures ANOVA on mu power revealed no significant effects: electrode main effect, $F(1, 14) = 0.07$, $p = 0.791$, $^2p < 0.01$; context main effect, $F(1, 14) = 0.13$, $p = 0.720$, $^2p < 0.01$; and interaction, $F(1, 14) = 0.13$, $p = 0.723$, $^2p < 0.01$.

For occipital alpha band power (Table 1 and Figure 3b), a 2 (electrode: O1 vs. O2) × 2 (context: constrained vs. unconstrained) repeated-measures ANOVA revealed no significant effects: electrode main effect, $F(1, 14) = 1.55$, $p = 0.234$, $^2p = 0.10$; context main effect, $F(1, 14) = 0.01$, $p = 0.925$, $^2p < 0.01$; and interaction, $F(1, 14) = 0.45$, $p = 0.514$, $^2p = 0.03$. Cluster-based permutation tests identical to Experiment 1 revealed no electrodes showing significant 8-13 Hz power differences between conditions.

To verify that Experiment 1' s results indeed stemmed from action understanding differences, we combined C3 and C4 data and conducted a 2 (experiment: Experiment 1 vs. Experiment 2; between-subjects) × 2 (context: constrained vs. unconstrained; within-subjects) ANOVA on parietal mu power. Results showed no significant main effects of experiment, $F(1, 28) < 0.01$, $p = 0.999$, $^2p < 0.01$, or context, $F(1, 14) = 2.45$, $p = 0.129$, $^2p = 0.08$, but a significant experiment × context interaction, $F(1, 14) = 4.77$, $p = 0.037$, $^2p = 0.15$. This suggests that experimental manipulation modulated parietal mu power differently across contexts, further supporting that Experiment 1' s results reflected action processing differences.

For the beta band (Table 1), a 2 (electrode: C3 vs. C4) × 2 (context: constrained vs. unconstrained) repeated-measures ANOVA revealed no significant effects: electrode main effect, $F(1, 14) = 0.12$, $p = 0.730$, $^2p < 0.01$; context main effect, $F(1, 14) = 0.40$, $p = 0.540$, $^2p = 0.03$; and interaction, $F(1, 14) = 0.51$, $p = 0.485$, $^2p = 0.04$.

These results demonstrate that despite preserving the same physical differences as Experiment 1, making the target ambiguous eliminated mu power differences between conditions. This indicates that Experiment 1' s results cannot be explained by different physical differences inducing different internal simulations, but rather by different inferential outcomes across constrained versus unconstrained contexts.

4 Discussion

Using animation technology to generate chasing actions and manipulating action contexts, we conducted two experiments measuring EEG mu suppression to investigate whether action understanding follows simulation theory or theory theory mechanisms. Results showed that when chasing actions occurred in constrained contexts—where clear action goals could be inferred based on rationality principles—mu suppression was greater than in unconstrained contexts (Experiment 1). However, when only the chaser moved (i.e., the approached target was ambiguous), mu suppression differences between conditions disappeared despite identical physical differences (Experiment 2). These findings indicate that contextual information influences action understanding, supporting theory theory.

Physical differences and familiarity with action contexts cannot explain these results. First, Experiment 1’ s constrained and unconstrained contexts shared identical kinematics, and Experiment 2 controlled for physical differences that contexts might introduce. Yet when contextual constraints made actions rational, mu suppression was higher than when actions violated rationality principles. Therefore, mu suppression differences indeed reflected different action understanding outcomes. Second, participants’ familiarity ratings for animations across contexts were equivalent, ruling out familiarity as an explanation for mu suppression differences. Third, different contexts did not modulate occipital alpha activation, excluding the possibility that mu suppression differences resulted from attention-related alpha effects. Finally, referencing previous mu suppression studies on action understanding (Fox et al., 2016; Pomiechowska & Csibra, 2017), the effect of goal clarity on this component is approximately 0.04 V^2, comparable to our study (0.05 V^2), with effect sizes larger than previous research (Cohen’ s d = 0.3-0.6; Experiment 1’ s effect size converted to Cohen’ s d = 1.88). This analysis demonstrates the sensitivity and stability of this index in indicating action processing. In summary, when both chaser and target moved, mu suppression differences across constrained and unconstrained contexts indeed reflected different action understanding outcomes, supporting theory theory.

The influence of contextual information on mu suppression aligns with previous research. For instance, mu suppression differs when observing grasping actions in drinking versus cleaning contexts (Amoruso & Urgesi, 2016); observing handle-grasping elicits greater mu suppression than observing other cup-grasping locations (Natraj et al., 2013); and observing object-directed grasping versus pointing actions with similar kinematics elicits stronger mu suppression, with this difference disappearing when actions occur in communicative contexts (preceded by communicative signals) (Pomiechowska & Csibra, 2017). However, these studies also involved physical differences across contexts, making their findings compatible with both simulation theory and theory theory predictions. Our study strictly controlled kinematics while producing different inferential outcomes across contexts with identical simulation processes, allowing us to distinguish between the two accounts. Our results support theory theory, holding important implications for action understanding mechanisms.

Action understanding involves inferring potentially infinite goals or intentions behind specific actions. For example, grasping a cup could be interpreted as drinking or moving it aside. Action understanding represents a classic inverse problem: observed actions are outcomes requiring interpretation of underlying causes—goals or intentions (Csibra, 2007; Csibra & Gergely, 2007). Simulation-based matching to one' s own experience struggles to provide clear action interpretations. Solving inverse problems requires constraints, and theory theory' s rationality-based inferential process—assuming actors use minimal-cost, economical actions to achieve goals—provides cognitive constraints for action understanding (Csibra & Gergely, 2007). Thus, people can compute action costs in relation to context to infer the most probable goal. In computational terms, this can be modeled via Bayesian inference: p(goal|action) = p(action|goal) × p(goal) / p(action), where p(goal|action) represents the inferred goal probability given the current action, p(action) and p(goal) represent the prior probabilities of the action and goal, and p(action|goal) can be computed based on rationality principles—the probability of producing the current action given a particular goal (Csibra, 2007). This modeling approach has successfully predicted action goals and aligned with human action understanding results (Baker, Saxe, & Tenenbaum, 2009).

Inferential action understanding explains how people comprehend novel actions. For example, when encountering robots for the first time, although their form and kinematics differ completely from previously observed human behaviors, observers attribute goals and intentions to robots taking shortest paths to objects, perceiving object acquisition as the robot' s purpose (Kamewari et al., 2005). Moreover, inferential action understanding may enable infants' action interpretation abilities. Research shows that 3-month-old infants can accurately understand action goals (Luo, 2011), and 15-month-olds can understand previously unseen pretend actions (Onishi, Baillargeon, & Leslie, 2007), such as pretending to drink from an empty cup. However, inferential processes do not exclude experience' s role. Experience provides additional cognitive constraints beyond rationality principles. For instance, furtive tailing behavior is not economical, yet people can draw on accumulated experience to infer that the tailer wants to approach without being detected, thus rationalizing this non-economical action. Therefore, as experience accumulates, people develop richer rationalizations for seemingly irrational behaviors and can more accurately interpret others' actions.

Our animation-based approach enabled effective manipulation of kinematics and action contexts while strictly controlling extraneous variables, providing a methodological model for future action understanding research. However, this method abstracts human actions to some degree, and the ecological validity of our conclusions requires testing with actual human actions.

5 Conclusion

Using chasing actions as stimuli and mu suppression as an index of action processing, two experiments investigated the cognitive mechanisms of action understanding. The findings indicate that contextual information influences action understanding, suggesting that people understand actions through inferential processes and supporting theory theory in action processing.

References

Amoruso, L., & Urgesi, C. (2016). Contextual modulation of motor resonance during the observation of everyday actions. *Neuroimage*, 134, 74–84.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, 113(3), 340–349.

Barrett, H. C., Todd, P. M., Miller, G. F., & Blythe, P. W. (2005). Accurate judgments of intention from motion cues alone: A cross-cultural study. *Evolution and Human Behavior*, 26(4), 313–331.

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, 2(8), 561–567.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436.

Brass, M., Schmitt, R. M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: Inferential processes versus action simulation. *Current Biology*, 17(24), 2117–2121.

Brown, E. C., Wiersema, J. R., Pourtois, G., & Brüne, M. (2013). Modulation of motor cortex activity when observing rewarding and punishing actions. *Neuropsychologia*, 51(1), 52–58.

Carter Leno, V., Vitoratou, S., Kent, R., Charman, T., Chandler, S., Jones, C. R., ⋯& Simonoff, E. (2019). Exploring the neurocognitive correlates of challenging behaviours in young people with autism spectrum disorder. *Autism*, 23(5), 1152–1164.

Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports*, 6, 37036.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6), 893–910.

Cracco, E., De Coster, L., Andres, M., & Brass, M. (2016). Mirroring multiple agents: Motor resonance during action observation is modulated by the number of agents. *Social Cognitive and Affective Neuroscience*, 11(9), 1495–1502.

Csibra, G. (2007). Action mirroring and action interpretation: An alternative account. In P. Haggard, Y. Rosetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition. Attention and performance XXII* (pp. 435-459). Oxford: Oxford University Press.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, 27, 111-133.

Csibra, G., & Gergely, G. (2007). "Obsessed with goals" : Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, 124, 60-78.

Cousineau, D., & O' Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods*, 46(4), 1149-1151.

Dijkerman, H. C., & Smit, M. C. (2007). Interference of grasping observation during prehension, a behavioural study. *Experimental Brain Research*, 176, 387-396.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91(1), 176-180.

Duan, J., Yang, Z., He, X., Shao, M., & Yin, J. (2018). Automatic attribution of social coordination information to chasing scenes: Evidence from mu suppression. *Experimental Brain Research*, 236(1), 117-127.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, 308(5722), 662-667.

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N., Vanderwert, R. E., ···Van Ijzendoorn, M. H. (2016). Assessing human mirror activity with EEG Mu rhythm: A meta-analysis. *Psychological Bulletin*, 142(3), 291-313.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2), 593-609.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind reading. *Trends in Cognitive Sciences*, 2(12), 493-501.

Gergely, G., Bekkering, H., & Király, I. (2002). Developmental psychology: Rational imitation in preverbal infants. *Nature*, 415(6873), 755.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, 7, 287-292.

Hauser, M., & Wood, J. (2010). Evolving the capacity to understand actions, intentions, and goals. *Annual Review of Psychology*, 61, 303-324.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57, 243-259.

Hobson, H. M., & Bishop, D. V. (2016). Mu suppression–a good measure of the human mirror neuron system? *Cortex*, 82, 290-310.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one' s own mirror neuron system. *PLoS Biology*, 3(3), e79.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286, 2526-2528.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: A critique. *Trends in Cognitive Sciences*, 9, 21-25.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, 20(8), 589-604.

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development*, 20(2), 303-320.

Kilner, J., Paulignan, Y., & Blakemore, S. (2003). An interference effect of observed biological movement on action. *Current Biology*, 13(6), 522-525.

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition-timing hypothesis. *Brain Research Reviews*, 53(1), 63-88.

Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: Lessons from autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1430), 345-360.

Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M. C., Mazziotta, J. C., & Iacoboni, M. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cerebral Cortex*, 12(8), 847-855.

Koul, A., Cavallo, A., Cauda, F., Costa, T., Diano, M., Pontil, M., & Becchio, C. (2018). Action observation areas represent intentions from subtle kinematic features. *Cerebral Cortex*, 28(7), 2647-2654.

Liepelt, R., Cramon, D., & Brass, M. (2008). What is matched in direct matching? Intention attribution modulates motor priming. *Journal of Experimental Psychology: Human Perception and Performance*, 34(3), 578-591.

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science*, 14(2), 453-460.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177-190.

Marsh, L. E., Mullett, T. L., Ropar, D., & Hamilton, A. F. D. C. (2014). Responses to irrational actions in action observation and mentalising networks of the human brain. *NeuroImage*, 103, 81-90.

Muthukumaraswamy, S. D., Johnson, B. W., & McNair, N. A. (2004). Mu rhythm modulation during observation of an object-directed grasp. *Cognitive Brain Research*, 19(2), 195-201.

Miller, S. (2001). *Social action: A teleological account.* Cambridge: Cambridge University Press.

Natraj, N., Poole, V., Mizelle, J. C., Flumini, A., Borghi, A. M., & Wheaton, L. A. (2013). Context and hand posture modulate the neural dynamics of tool-object perception. *Neuropsychologia*, 51(3), 506-519.

Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica*, 124, 106-128.

Perry, A., Stein, L., & Bentin, S. (2011). Motor and attentional mechanisms involved in social interaction—Evidence from mu and alpha EEG suppression. *Neuroimage*, 58(3), 895-904.

Pomiechowska, B., & Csibra, G. (2017). Motor activation during action perception depends on action interpretation. *Neuropsychologia*, 105, 84-91.

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, 110(41), E3965-E3972.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169-192.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Review Neuroscience*, 2, 661-670.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, 96(1), B1-B11.

Southgate, V., & Csibra, G. (2009). Inferring the outcome of an ongoing novel action at 13 months. *Developmental Psychology*, 45(6), 1794-1798.

Tangwiriyasakul, C., Verhagen, R., van Putten, M. J., & Rutten, W. L. (2013). Importance of baseline in event-related desynchronization during a combination task of motor imagery and motor observation. *Journal of Neural Engineering*, 10(2), 026009.

Ullman, T., Baker, C. L., Macindoe, O., Evans, O., Goodman, N. D., & Tenenbaum, J. B. (2009). Help or hinder: Bayesian models of social goal inference. In Y. Bengio et al. (Eds.), *Advances in neural information processing systems 22.* Vancouver: NIPS Foundation.

Ulloa, E. R., & Pineda, J. A. (2007). Recognition of point-light biological motion: Mu rhythms and mirror neuron activity. *Behavioural Brain Research*, 183(2), 188-194.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655-684.

Ye, H. (2016). The significances of mirror neurons. *Acta Psychologica Sinica*, 48(4), 444-456.

Yin, J., Xu, H., Ding, X., Liang, J., Shui, R., & Shen, M. (2016). Social constraints from an observer' s perspective: Coordinated actions make an agent' s position more predictable. *Cognition*, 151, 10-17.

Note: Figure translations are in progress. See original paper for figures.