



HAL
open science

Timing of grip and goal activation during action perception: a priming study

Jérémy Decroix, Solène Kalénine

► **To cite this version:**

Jérémy Decroix, Solène Kalénine. Timing of grip and goal activation during action perception: a priming study. *Experimental Brain Research*, 2018, 236 (8), pp.2411-2426. 10.1007/s00221-018-5309-0. hal-02109222

HAL Id: hal-02109222

<https://hal.science/hal-02109222v1>

Submitted on 5 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Title :** Timing of grip and goal activation during action perception : A priming study

2 **Author names :** Jérémy Decroix & Solène Kalénine

3 **Affiliations :** Univ. Lille, CNRS, CHU Lille, UMR 9193 - SCALab - Sciences Cognitives et Sciences
4 Affectives, F-59000 Lille, France

5 **Authors' adresse :**

6 Laboratoire SCALab UMR CNRS 9193
7 Université de Lille, SHS
8 Batiment A2 - Niveau forum +2
9 Domaine universitaire du Pont de Bois, BP 149
10 59653 Villeneuve d'Ascq Cedex, France
11

12 **Corresponding author:** Solène Kalénine

13 Mail : solene.kalenine@univ-lille3.fr

14 Phone : +33 3 20 41 73 25

15

16 **Abstract**

17 Several models of action recognition acknowledge the involvement of distinct grip and goal representations in the
18 processing of others' actions. Yet their functional role and temporal organisation are still debated. The present
19 priming study aimed at evaluating the relative timing of grip and goal activation during the processing of
20 photographs of object-directed actions. Action could be correct or incorrect owing to grip and/or goal violations.
21 Twenty-eight (Experiment 1) and twenty-five (Experiment 2) healthy adults judged the correctness of target
22 actions according to object typical use. Target pictures were primed by action pictures sharing the same grip or
23 same goal, both the same grip and same goal or none. Primes were presented for 66 or 300ms in Experiment 1 and
24 for 120 or 220ms in Experiment 2. In Experiment 1, facilitative priming effects were observed for goal and grip
25 similarity after 300 ms primes but only for goal after 66ms primes. In Experiment 2, facilitative priming effects
26 were found for both goal and grip similarity from 120 ms of prime processing. In addition, results from a control
27 condition in Experiment 2 indicated that mere object priming could partially account for goal similarity priming
28 effects, suggesting that object identity may help the observer to make predictions about possible action goals.
29 Findings demonstrate an early and first activation of goal representations, as compared to grip representations, in
30 action decoding, consistent with predictive accounts of action understanding. Future studies should determine to
31 what extent the timing of grip and goal activation is context-sensitive.

32

33 **Key-words :** Action understanding; Goals; Grip; Action semantics; Priming

34

35 **Accepted in Experimental Brain Research:**

36 Decroix, J. & Kalénine, S. (2018). Timing of grip and goal activation during action perception: A priming study.

37 *Experimental Brain Research*, 236, 2411-2426. doi: [10.1007/s00221-018-5309-0](https://doi.org/10.1007/s00221-018-5309-0)

1 **Introduction**

2 When carrying out actions, humans not only execute movements but aim at achieving goals (Jacob and
3 Jeannerod 2005; Kilner 2011). Yet action goals may be conceptualized at different levels of generality, and action
4 planning and control may occur at multiple levels in the action goal hierarchy. For example, one may reach and
5 grasp a cup (direct goal) to drink or to move it away (distant goal). There is considerable evidence supporting the
6 existence of action planning at these different levels. For instance, performing the same reach and grasp action
7 (direct goal), varies depending on the next action step (distant goals, Ansuini, Santello, Massaccesi, & Castiello,
8 2005; Quesque, Lewkowicz, Delevoeye-Turrell, & Coello, 2013). In addition, the neural mechanisms underlying
9 direct and distant goals can be dissociated during action planning (van Schie & Bekkering, 2007). Overall, the
10 action planning literature highlights the importance of distant goals (such as action intentions or outcomes) and
11 direct goals (such as kinematics parameters, e.g. reach trajectory, grip configuration; Grafton & Hamilton, 2007)
12 in guiding object-directed actions (Thill et al. 2013; Cooper et al. 2014; van Elk et al. 2014).

13 In line with the claim of close action-perception relationships (Barsalou 2008; Gentsch et al. 2016), the
14 ability to understand observed actions have been proposed to involve a similar structure as the one identified during
15 action planning (Wolpert et al. 2003; Grafton and Hamilton 2007). Several findings indicate that humans are
16 specifically tuned to interpret others' movements as intended and goal-directed (Hrkać et al. 2014; Novack et al.
17 2016). Consistent with this idea, fMRI and TMS studies consistently report brain dissociations between the regions
18 involved in decoding action kinematic parameters (reach, grip, effector) and those associated with processing goals
19 (Grafton and Hamilton 2007; Cattaneo et al. 2010; Jacquet and Avenanti 2015; Wurm and Lingnau 2015). Yet the
20 functional role and the temporal organization of kinematics ("how") and goal ("what for") representations during
21 decoding of observed actions remain a matter of debate.

22 On the one hand, "bottom-up" approaches consider action decoding as bottom-up propagation from the
23 processing of kinematic parameters to the activation of goal representations (Zentgraf et al. 2011; Rizzolatti and
24 Fogassi 2014). Consistent with these views, several studies have demonstrated that there is sufficient invariant
25 information in movements kinematics to discriminate between different action goals (e.g., grasp to pour vs. grasp
26 to drink) and that this information is indeed used by the observers to anticipate the goal of the observed action
27 (Ansuini et al. 2014; Cavallo et al. 2016). On the other hand, "top-down" approaches suggest that contextual
28 information is first used to compute a prediction about the forthcoming action, the actual kinematics being then
29 used to evaluate this prediction (Kilner et al. 2007; Bach et al. 2014). Visual context is indeed known to influence
30 action decoding (Wurm and Schubotz 2012, 2016) and recent evidence suggests that prior knowledge about others'
31 goals directly biases processing of kinematic parameters (Hudson et al. 2016a, b).

32 Together, bottom-up and top-down approaches are difficult to disentangle based on empirical evidence.
33 Processing of action kinematics contributes to subsequent decisions about action intentions, but it does not inform
34 about whether movement analysis is driven by the prior activation of a representation of the possible action goal.
35 Similarly, although contextual information influences the processing of observed actions, it remains to clarify
36 whether predictions about possible action goals guide action decoding in the absence of externally-driven
37 expectancies (visual scene, prior knowledge about other's intentions). The present study aims at separating bottom-
38 up and top-down approaches of action understanding based on temporal data, as suggested by Catmur (2015).
39 Indeed, whereas bottom-up approaches imply that representations of action kinematics would be first activated,
40 top-down approaches suggest that goal representations would be first recruited.

1 Assessing the timing of grip and goal activations during action perceptual processing requires to
2 experimentally manipulate information about action kinematics and action goal independently. A few behavioural
3 experiments have dissociated kinematic and goal information during the processing of visual actions (van Elk et
4 al. 2008; Kalénine et al. 2013). Kalénine et al. (2013) found that healthy participants were slower to determine that
5 two videos displayed different actions when the two actions differed only in terms of kinematic parameters (e.g.
6 applying detergent with circular versus straight wipe), or in terms of goals (e.g., applying versus removing
7 detergent), as compared to when the two actions differed in terms of both kinematic and goal. Van Elk et al. (2008)
8 showed that response times to determine that the goal of the action displayed in a picture was correct (“Is the object
9 held at the correct goal location?” e.g., cup near the ear versus near the mouth) were slower when the grip
10 component of the action was incorrect. Conversely, response times to decide whether the grip was correct (i.e., “Is
11 the object grasped with an appropriate grip?” e.g., pencil grasped with power clench versus precision grip) were
12 slower when the goal was incorrect. Moreover, interference from incorrect goals was greater than from incorrect
13 grips. Overall, both studies support the hypothesis that action representations at the kinematic and goal levels play
14 distinct but complementary roles in the processing of observed actions. However, although they provide empirical
15 evidence about the relative “weight” of the different action dimensions in decisions about perceived actions,
16 previous studies do not directly inform about the timing of processing of grip and goal information during action
17 decoding.

18 Using a priming paradigm involving action pictures containing action violations, the present study goes
19 one step further by investigating the relative timing of grip and goal activation during action recognition. Target
20 actions were judged correct or incorrect after being primed by action pictures sharing the same grip and/or the
21 same goal. The manipulation of different prime durations allowed to directly assess the timing of activation of grip
22 and goal action representations. Prime durations were designed according to EEG studies on perceptual processing
23 of object-directed actions. In these studies, brain activity of participants is recorded while they are determining the
24 goal of an observed action. Overall, decoding others’ action goals impacts brain activity in different time windows,
25 the earliest starting around 60 ms (Ortigue et al. 2009) and the latest emerging from 300 ms (Ortigue et al. 2009;
26 van Elk et al. 2012; Avanzini et al. 2013) of visual action processing. Thus, 60 and 300 ms were chosen as prime
27 duration boundaries. The rationale was that short prime durations should only allow the processing of the first
28 steps of action decoding. Depending of which is processed first (kinematic parameters for bottom-up approaches;
29 goal for top-down approaches), selective grip and goal priming should be observed at short prime durations, with
30 faster judgments for action pairs sharing either the same grip or the same goal. At long prime durations, both grip
31 and goal representations should be activated and elicit action priming effects. [In addition, effects of partial grip](#)
32 [violations \(e.g., grip incorrect but goal correct, see below stimuli section\) or partial goal violations \(e.g., grip](#)
33 [correct but goal incorrect, see below stimuli section\) should be visible: the co-activation of incongruent grip and](#)
34 [goal information at long prime durations should interfere with target action processing.](#)

1 **EXPERIMENT 1**

2 **Materials and methods**

3 **Participants**

4 Thirty-one healthy participants took part in the study¹. Three of them were left-handed or ambidextrous
5 according to the Edinburgh handedness inventory (EHI; Oldfield, 1971) and were excluded. The final sample
6 included 28 participants (mean age 22, age range 18-36, 23 females). All were right-handed (mean EHI= 87%,
7 from 45% to 100%), reported normal or corrected-to-normal vision and provided written informed consent. They
8 were not paid but could receive extra course credit for their participation. The protocol was approved by the Ethical
9 Committee of the University and was in accordance with the declaration of Helsinki (1964, revised in 2013).

10 **Stimuli**

11 Twenty manipulable objects were selected (Appendix 1). Four 1024*683 pixel coloured photographs of
12 each reference object were taken, all involving the same right-handed actress interacting with the object in one of
13 four conditions. The object-directed action could be correct or not by crossing the correctness of the grip used by
14 the actress and the correctness of the action goal (Figure 1). Thus, there were four types of action stimuli: actions
15 that were fully correct with no violation and actions that were incorrect owing to goal violations, grip violations,
16 or both grip and goal violations. Correct grips were defined as the typical grasp-to-use grip of an object. Thus, an
17 incorrect grip corresponded to an atypical (although not impossible) grip considering the typical use of the object.
18 Similarly, correct goals were defined as the actor's goal to use the object in order to fulfil its typical/primary
19 function. Thus, a goal was considered incorrect when the object could not be used for its typical function.
20 Importantly, the incorrect grip did not prevent the correct goal to be achieved. Thus, grip and goal were
21 manipulated independently. Note that contrary to van Elk et al. (2008) who manipulated goal representation in
22 terms of goal location (e.g. bring cup to the mouth versus to the ear), we manipulate the orientation of the object
23 to make its functional use possible or not, and then focus on the functional goal of the action (e.g., bring upright
24 cup versus up-side down cup to the mouth). [Moreover, the recipient object on which the object involved in the
25 action may act upon \(e.g. plant for water can\) was not included in the picture to minimize expectancies that may
26 be driven from other sources of information than the use of the tool.](#)

27 **Experimental design**

28 Photographs were implemented in a priming paradigm. For each reference object, action photographs
29 without violations ("correct" actions) and action photographs with both grip and goal violations (all "incorrect"
30 actions) were used as targets for the "yes" (correct) and "no" (incorrect) judgment task. For each target, all four
31 types of photographs were used as primes. Thus, each reference object was associated with 8 pairs of photographs
32 divided in 4 types of action pairs. The prime could display the same grip and same goal (Grip and goal similar), a
33 different grip but the same goal (Goal similar only), the same grip but a different goal (Grip similar only) or a
34 different grip and a different goal (All different) as the target.

¹ A sample size of about 30 participants was chosen to ensure sufficient statistical power (.80) for anticipated moderate effect sizes (Cohen d = .50 for the critical paired comparisons).

1 There was a total of 2 prime-target grip similarity (grip similar, grip different) x 2 prime-target goal
 2 similarity (goal similar, goal different) x 2 response types (yes=correct target action, no=incorrect target action) x
 3 20 objects = 160 prime-target pairs. The design of the different priming conditions is presented on Figure 1.

4 Prime duration was manipulated in 2 conditions, a short (66 ms) and a long (300 ms) prime duration. The
 5 160 pairs of photographs were then presented twice, once with a prime duration of 66ms and once with a prime
 6 duration of 300ms, leading to 320 experimental trials. Trials were divided into 4 blocks of 80 trials, each block
 7 containing an equal number of trials of the 4 prime-target pairs x 2 response types x 2 prime duration conditions.

8 **Control measure of perceptual similarity**

9 An objective index of perceptual similarity between prime and target was computed using the FSIM
 10 algorithm (Zhang et al. 2011) to assess potential differences in terms of low-level visual features between prime-
 11 target pairs in the different conditions. According to this measure, pairs in the ‘grip similar only’ condition
 12 presented a higher degree of perceptual similarity than pairs in the ‘goal similar only’ and all different conditions
 13 ($p = .001$), the two having the same degree of perceptual similarity ($p = .09$), see also Figure 1 and Appendix 2 for
 14 details. Note that the perceptual similarity index in the grip and goal similar condition had no variance as the pair
 15 was composed of the same exact picture. Perceptual similarity scores were taken into account in a complementary
 16 analysis (below).

17 [INSERT FIG 1]

18 **Procedure**

19 Participants were seated at 100cm from the screen so that the action in the picture would appear within 5
 20 degrees of visual angle. Their task was to judge whether the action presented in the picture was correct or not
 21 (forced choice). Instructions were displayed on the screen and carefully explained by the experimenter to ensure
 22 that participants understood the correct/incorrect distinction.² Twelve representative practice trials with feedback
 23 were provided with 3 objects that were not included in the actual experiment. The experimental session was the
 24 same as the practice session but without feedback. The experiment was conducted with E-Prime V2.0.10.353
 25 software (Psychology Software Tools, Pittsburgh, PA).

26 Each trial began with a 2500 ms fixation cross, then a prime was displayed during 66 or 300ms, followed
 27 by a 66 ms grey screen mask, and finally the target until participants’ response. Participants answered “yes”
 28 (correct) or “no” (incorrect) using their left and right hands on left and right extreme keys of the response box.
 29 Response mapping (yes/no, left/right) was counterbalanced between participants. Reaction times (RT) and correct
 30 responses (accuracy) were recorded. The four blocks and the 80 experimental trials within each block were
 31 randomly presented. Short breaks were proposed between blocks. The experiment lasted about 1 hour.

32
 33 **Data Analysis**

² “On each trial, you will see two successive pictures showing an actor using an object. The first picture will
 always be briefly presented. You will have to judge the second photography. You will have to determine, as fast
 and as accurately as possible, if the presented action is correct or not according to the typical use of the object.
 The use of an object is atypical when the object is used for another purpose or in another manner as the typical
 one. You will start with a training in which you will have feedback.”

1 Correct RT were analysed as a function of grip similarity between prime and target (grip similar, grip
2 different), goal similarity between prime and target (goal similar, goal different), prime duration (66ms, 300ms)
3 and response type (yes, no); the three factors being repeated between participants and items. Response type was
4 not a factor of interest but was introduced in the model since effect of yes/no response have been frequently
5 reported. Effects of grip and goal similarities were expected to be modulated by prime duration. According to
6 bottom-up approaches, grip information should be processed first and effect of grip similarity should be more and
7 perhaps exclusively visible at 66 ms, as compared to effect of goal similarity. Alternatively, according to top-down
8 approaches, the goal should be activated first and effect of goal similarity should be more and perhaps exclusively
9 obvious at 66 ms, as compared to effect of grip similarity. For long latencies, observers should have activated both
10 goal and grip information about the action regardless of whether observed actions are processed top-down or
11 bottom up. Thus, both effects of grip and goal priming effects were expected at 300ms.

12 In brief, we expected an interaction between prime duration, grip priming and goal priming. The
13 modulation of grip and goal priming by prime duration would reflect the different timing of activation of grip and
14 goal information within the 66-300 ms time window. The interaction between grip and goal priming effects would
15 reflect the detection of the incongruency between goal and grip information within the prime and would confirm
16 that both grip and goal representations have been co-activated.

17

18 **Results**

19 **Data preparation for reaction times analysis**

20 All analyses were carried out with R version 3.3.2 (R Core Team 2017) and R Studio version 1.0.136.
21 Overall, participants made few errors (mean = 2%, range .3% - 6%). Since the task was relatively easy to perform
22 and errors were not of primary interest, errors were solely analysed in order to verify the absence of any trade-off
23 effect. A chi-square test for independence indicated that errors were equally distributed between conditions ($\chi^2(3)$
24 = .014, $p = .99$). Errors were then replaced by median RT of the remaining trials in each condition. RT superior to
25 1500ms and inferior to 200ms were considered outliers and then replaced by median RT of the remaining trials in
26 each condition. Overall, 2.84 % of the data were replaced. Final data are presented on Figure 2.

27 [INSERT FIGURE 2]

28 **Mixed-effect model analysis of correct action recognition RT**

29 A mixed-effect model comparison approach was used to analyse correct RT. This approach allows taking
30 into account both items and participants as sources of variation and considering their possible interaction with the
31 factors of interest (e.g., a given participant or a given item, may be more sensitive to grip similarity than another).
32 Analyses were carried out with the lmer function of the lme4 1.1-12 package (Bates et al. 2015b).

33 The full model included grip similarity (GRIP), goal similarity (GOAL), prime duration (DURATION),
34 response type (RESP) and related first-order, second-order, and third order interactions as fixed effects. It also
35 included random intercepts and random slopes for GRIP, GOAL and RESP for participants and random intercepts
36 for DURATION and RESP for items (1). The choice of this random effect structure resulted from a compromise

1 between the maximum random effect structure (Barr et al. 2013) and the optimal random structure supported by
2 the data (Bates, Kliegl, Vasishth, & Baayen, 2015; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017; see
3 Appendix 3A for determination of the full model structure).

4 Augmented and reduced models (i.e., with and without the fixed effects of interest) were compared using -
5 2LogLikelihood Ratio test distributed like χ^2 with degrees of freedom corresponding to the number of parameters
6 added. Parameter estimates of the models are provided with their corresponding t-values, when relevant. Pseudo
7 R^2 effect-sizes were computed with the r.squaredLR function of the MuMIn 1.15.6 package (Barton 2016). Note
8 however that pseudo R^2 for mixed-models cannot be interpreted using classical standards for effect size.

9 **Effect of grip and goal similarity priming**

10 The model comparison did not show any significant 4-way interaction between GRIP, GOAL, DURATION,
11 and RESP ($\chi^2(1) = .8861$, $p = .346$, $R^2 = .10^{10^{-4}}$). As expected, the 3-way interaction between GRIP, GOAL, and
12 DURATION was significant ($\chi^2(1) = 120.57$, $p = .001$, $R^2 = 133.67^{10^{-4}}$). We thus tested the interaction model
13 between GRIP and GOAL for each DURATION. At 66 ms, there was no interaction between GRIP and GOAL
14 ($\chi^2(1) = .11$, $p = .740$, $R^2 = .25^{10^{-4}}$). There was a main effect of GOAL ($\chi^2(1) = 8.06$, $p = .004$, $R^2 = 17.98^{10^{-4}}$;
15 Estimate goal similar-goal different = - 8.53 ms, SE = 3 ms, $t = -2.84$), but no significant main effect of GRIP
16 ($\chi^2(1) = 1.74$, $p = .187$, $R^2 = 3.89^{10^{-4}}$; Estimate grip similar-grip different = - 3.97 ms, SE = 3 ms, $t = -1.32$)
17 reflecting goal priming but no grip priming at 66 ms. In contrast, at 300 ms, the interaction model between GRIP
18 and GOAL was significant ($\chi^2(1) = 241.11$, $p = .001$, $R^2 = 523.96^{10^{-4}}$). Parameter estimates of the model showed
19 a main effect of GOAL (Estimate goal similar-goal different = - 15.71 ms, SE = 3.53 ms, $t = -4.44$) and a main
20 effect of GRIP (Estimate grip similar-grip different = - 31.51 ms, SE = 2.90 ms, $t = -10.86$), reflecting both goal
21 and grip priming at 300 ms. In addition, the interaction estimate indicated that action primes with partial grip or
22 goal violations slowed down target judgements by 62 ms in comparison with all similar or all different primes
23 (estimate = -61.78, SE= 3.92, $t = -15.74$).

24 **Complementary analysis: influence of perceptual similarity**

25 Spearman's rank correlations were computed between the goal and grip similarity priming effects
26 obtained in each response condition and the corresponding perceptual similarity indices. There were no significant
27 correlations (all $p > .66$, see Appendix 2).

1 Discussion

2 This experiment used a priming paradigm involving object-direct action photographs with grip and/or
3 goal violations to assess the relative timing of grip and goal activation during the processing of others' actions.
4 The action goal was defined in functional terms and corresponded to the use of the object according to its typical
5 function (e.g. water bottle: to drink). Action primes were expected to influence response times to determine
6 whether target actions were correct or not when a) the prime shared grip or goal information with the target and
7 when b) prime duration was sufficient to activate the common representation. We found that action primes with
8 similar grip and similar goal both facilitated target action processing when presented for 300 ms. Conversely, when
9 action primes were presented for 66 ms, only those sharing the same goal facilitated target action processing.
10 Moreover, primes containing incongruent grip and goal information (partial violations) were more slowly
11 processed than primes with no violation when presented for 300 ms.

12 Priming effects for 300ms prime duration are in line with previous experiments (van Elk et al. 2008;
13 Kalénine et al. 2013) with the co-activation of both grip and goal representations during action decoding. Here, as
14 long as participants had sufficient time to process the prime (300 ms), both grip and goal information influenced
15 target action processing. Interestingly, the magnitude of grip similarity priming was greater than that of goal
16 similarity priming. In our study, participants did not have to explicitly process the different action dimensions
17 manipulated in the primes. This suggests that whereas grip information may be more difficult to process when
18 judged explicitly (cf. Van Elk et al., 2008), it strongly impacts action judgements when implicitly processed.
19 Moreover, our priming paradigm highlights a cost on target processing when primes contained incongruent grip
20 and goal information (partial action violations). This result mirrors the slower action decisions observed on action
21 stimuli containing partial violations in previous studies (Kalénine et al., 2013; van Elk et al., 2008). In contrast to
22 previous studies, participants were not explicitly instructed about the action dimensions manipulated. Thus, the
23 cost entailed by partial action violations indicates that once activated, grip and goal representations automatically
24 interact with one another during action perceptual processing. This extends the claim of the involvement of distinct
25 but interrelated kinematic parameters and goal representations in perception.

26 When primes were presented for 66ms, only goal similarity priming was able to facilitate target action
27 judgments. Absence of early grip priming may be interpreted in the light of previous studies that found that
28 sensorimotor simulation of specific kinematic parameters performed by others may not arise before 200 ms (Naish,
29 Houston-Price, Bremner, & Holmes, 2014). Although modulations of corticospinal excitability by observed
30 actions have been found as early as 90ms (e.g., Lepage et al. 2010; Cavallo et al. 2014 see Naish et al., 2014 for
31 review), they would first correspond to a general preactivation of the motor system before reflecting the mapping
32 of the specific kinematic parameters perceived onto the observer's own motor system. Therefore, even though the
33 grip may have been processed during 66ms of prime processing, this processing may have not been sufficient to
34 influence target action judgments.

35 In contrast, early activation of goal-related information has been previously observed. In particular, EEG
36 studies showed that discriminating between the different possible action goals of an observed action starts
37 influencing brain activity after 60 ms of visual processing (e.g., Ortigue, Thompson, Parasuraman, & Grafton,
38 2009). Yet these early responses have often been associated to perceptual processes (Catmur, 2015) rather than
39 early activation of goal-related information. Earlier and stronger facilitative priming effects may then be expected
40 between pairs of action pictures that are more perceptually similar. However, this possibility was ruled out using

1 the FSIM measure of low-level perceptual similarity. Grip similar pairs were more perceptually similar than goal
2 similar pairs and there was no correlation between the amplitude of grip and goal similarity priming effects and
3 perceptual similarity differences between prime-target pairs. Thus, it is unlikely that the early facilitative goal
4 priming effect merely reflects the greater perceptual similarity between action pairs in this condition. In contrast,
5 goal priming results are consistent with an early activation of goal representations. Critically, goal activation was
6 not only visible early but first. Thus, the relative timing of goal and grip activation supports a top-down processing
7 of observed actions in our task. This suggests that participants first activate a prediction about the possible action
8 goal of the actor that will guide the subsequent processing of the kinematic parameters of the action, which
9 occurred between 60 and 300 ms.

10 To further specify the time-course of grip and goal activation, and in particular determine the minimal
11 time required to activate both grip and goal representations from action visual stimuli, a second Experiment was
12 carried out. Intermediate prime durations were selected between 66 and 300 ms. A first duration of 120 ms was
13 chosen, which corresponds to the late boundary of first time window of action understanding identified at the
14 neurophysiological level (e.g., Ortigue et al., 2009). At 120 ms, facilitative effects of goal similarity were expected,
15 which would confirm the earliness of goal activation. A second prime duration of 220ms was selected to fit the
16 timing at which specific motor responses have been reported during action observation (Naish et al., 2014). Effects
17 of grip similarity should emerge between 120 and 220 ms. Experiment 2 followed the same rationale and paradigm
18 as Experiment, asides from two additions:

- 19 a) The strength of the perceptual mask used between prime and target was enhanced for a greater control of low-
20 level perceptual priming effects.
- 21 b) A neutral prime condition presenting the object without action was added to further evaluate the role of object
22 information in priming effects involving object-related actions.

1 **EXPERIMENT 2**

2 **Materials and methods**

3 **Participants**

4 Thirty healthy participants took part in the study. Five of them were classified as left-handed or
5 ambidextrous according to the EHI (Oldfield 1971), and were excluded from the study. The final sample included
6 25 participants (mean age 21, age range 18-30, 16 females). All were right-handed (mean EHI= 84 %, from 43 %
7 to 100 %), reported normal or corrected-to-normal vision and provided written informed consent. They were not paid
8 but could receive extra course credit for their participation. The protocol was approved by the Ethical Committee
9 of the University and was in accordance with the declaration of Helsinki (1964, revised in 2013).

10 **Stimuli**

11 Experiment 2 used the same prime-target action pairs as Experiment 1, divided in 4 conditions: 'Grip and
12 goal similar', 'Goal similar only', 'Grip similar only', and 'All different'. Furthermore, a control condition was
13 added in which the target was coupled with a neutral action-free prime in which the object was displayed in front
14 of the actress with her hands resting on the table (Neutral). This control condition was designed to evaluate to what
15 extent mere object priming could account for the expected effects of goal and/or grip similarity priming (see
16 complementary analysis below). Overall, there was a total of 5 prime-target pairs x 2 response types x 20 objects
17 = 200 prime-target pairs (see Figure 1, and Appendix 1). Trials were divided into 4 blocks of 100 trials, each block
18 containing an equal number of trials of the 5 prime-target pairs x 2 response types x 2 prime duration conditions.

19 **Procedure**

20 Experiment 2 evaluated two intermediate prime durations of 120 ms and 220 ms. A few additional
21 changes were applied to the procedure. The grey mask used in the first experiment was replaced by a black and
22 white pixelated mask and the target picture was spatially shifted from the prime picture by 0.5 degree of visual
23 angle. These subtle changes have been made to the procedure in order to provide a stronger control of low-level
24 visual effects such as visual persistence and low-level spatial overlap. Apart from those modifications, the
25 procedure and the rationale of Experiment 2 were the same as Experiment 1. The four blocks and the 100
26 experimental trials within each block were randomly presented. Short breaks were proposed between blocks. The
27 experiment lasted about 1 hour.

28 **Data analysis**

29 As in Experiment 1, RTs were analysed as a function of grip similarity between prime and target (grip
30 similar, grip different), goal similarity between prime and target (goal similar, goal different), prime duration
31 (120ms, 220ms) and response type (yes, no); the three factors being repeated between participants and items. As
32 goal similarity priming was observed from 66 ms of prime duration in Experiment 1, it was expected at both 120
33 ms and 220 ms of prime duration in Experiment 2. In contrast, as grip similarity priming was evidenced only for
34 300 ms prime duration in Experiment 1, grip decoding was supposed to occur sometime between 66 and 300 ms
35 and grip priming effects could possibly not be visible at one or both of these intermediate prime durations.

1 As in Experiment 1, we expected an interaction between prime duration, grip priming and goal priming.
2 The modulation of grip and goal priming by prime duration would reflect the different timing of activation of grip
3 and goal information within the 120-220 ms time window. Again, the interaction between grip and goal priming
4 effects would reflect the detection of the incongruency between goal and grip information within the prime and
5 would confirm that both grip and goal representations have been co-activated.

6 Results

7 Data preparation for reaction times analysis

8 Again, error rate was relatively low (mean = 3%, from 0.4% to 12%). A chi-square test for independence
9 revealed that errors were equally distributed between conditions ($\chi^2(4) = .0126, p > .99$). Errors were replaced by
10 median RT of the remaining items in each condition. RT superior to 1500 ms and inferior to 200 ms were
11 considered outliers and then replaced by median RT of the remaining items in each condition. Overall, 3.1 % of
12 the data were replaced. Final data are reported on Figure 2.

14 Mixed-effect model analysis of correct action recognition RT

15 The same mixed-effect model approach was used to analyse correct RT in Experiment 2. The full model
16 included grip similarity (GRIP), goal similarity (GOAL), prime duration (SOA) and response type (RESP) and
17 related first-order, second-order and third order interactions as fixed effects. It also included random intercepts
18 and random slopes for RESP for participants and intercept and random slopes for DURATION and RESP for items
19 (2). The final model structure is justified in Appendix 3B.

20 Effect of grip and goal similarity priming

21 Comparison of augmented and reduced models highlighted a 4-way interaction between GRIP, GOAL,
22 DURATION and RESP ($\chi^2(1) = 3.82, p = .051, R^2 = 4.78 \cdot 10^{-4}$). We therefore tested the interaction between GRIP,
23 GOAL, and DURATION for each response type separately.

24 For YES responses, the GRIP x GOAL x DURATION interaction was not significant ($\chi^2(1) = 1.34, p = .246,$
25 $R^2 = 3.36 \cdot 10^{-4}$). The GRIP x GOAL interaction was thus evaluated regardless of prime duration for yes responses.
26 The GRIP x GOAL interaction model was significant ($\chi^2(1) = 34.04, p = .001, R^2 = 84.74 \cdot 10^{-4}$). Parameter estimates
27 of this model showed a main effect of GOAL (Estimate goal similar-goal different = - 33.83 ms, SE = 2.96 ms, t
28 = -11.43) and a main effect of GRIP (Estimate grip similar-grip different = - 20.44 ms, SE = 2.96 ms, t = -6.91),
29 reflecting both goal and grip priming at intermediate prime durations. In addition, the interaction estimate indicated
30 that action primes with partial grip or goal violations slowed down target judgements by 32 ms in comparison with
31 primes that were all similar or all different (estimate = -32.00, SE= 4.07, t = -7.87).

32 For NO responses, the GRIP x GOAL x DURATION interaction was significant ($\chi^2(1) = 16.21, p = .001, R^2$
33 $= 39.97 \cdot 10^{-4}$). The GRIP x GOAL interaction model was significant at 120 ms ($\chi^2(1) = 6.86, p = .008, R^2 = 34.24$
34 $\cdot 10^{-4}$). Parameter estimates of this model did not show any main effect of GOAL or GRIP (Estimate goal similar-
35 goal different = -0.38 ms, SE = 4.25 ms, t = -0.09; Estimate grip similar-grip different = - 3.78 ms, SE = 4.25 ms,
36 t = -0.89). Yet the interaction estimate indicated that action primes with partial grip or goal violations slowed down

1 target judgements by 16 ms in comparison with primes that were all similar or all different (estimate = -15.75, SE=
2 6.01, $t = -2.62$). The GRIP x GOAL interaction model was also significant at 220 ms ($\chi^2(1) = 76.66$, $p = .001$, R^2
3 = $376.04^{10^{-4}}$). Parameter estimates of the model demonstrated a main effect of GOAL (Estimate goal similar-goal
4 different = -13.41 ms, SE = 3.86 ms, $t = -3.47$) and a main effect of grip (Estimate grip similar-grip different = -
5 17.79 ms, SE = 3.86 ms, $t = -4.61$). The interaction estimate showed a cost of 48 ms on target judgements for
6 action primes with partial grip or goal violations compared to primes that were all similar or all different (estimate
7 = - 48.26, SE= 5.46, $t = -8.84$).

8 Together, results of Experiment 2 showed effects of goal and grip similarity from 120 ms of prime processing.
9 Interactions with response type and prime duration suggest a subtle additional modulation of grip and goal
10 similarity effects with prime duration (120, 220 ms) in the case of no-responses. We will go back to these nuances
11 in the Discussion of Experiment 2.

12 [INSERT TABLE 1]

13 **Complementary analysis: Effect of object priming**

14 If object identity is used to derive information about the goal or grip of the action, then the pattern of goal
15 and grip similarity priming effects should follow the pattern of object priming effects. To evaluate this hypothesis,
16 we replaced the conditions in which grip or goal overlapped between prime and target with the neutral priming
17 condition. In other words, neutral prime-target pairs (object similar) were compared with prime-target pairs that
18 differed on grip (grip dissimilar conditions) on the hand, and with prime-target pairs that differed on goal (goal
19 dissimilar conditions, see Table 1). This comparison was tested in situations where we observed an effect of grip
20 similarity and goal similarity in the main analysis of Experiment 2.

21 For YES responses, object similarity did not show any advantage over grip dissimilarity ($\chi^2(1) = 1.02$, p
22 = $.312$, $R^2 = 2.55^{10^{-4}}$). This stands in contrast with the grip similarity priming observed in the main analysis and
23 indicates that object repetition is not responsible of the grip similarity priming effect. Object similarity did show a
24 14 ms advantage over goal dissimilarity, however ($\chi^2(1) = 10.43$, $p = .002$, $R^2 = 26.03^{10^{-4}}$; estimate = -13.83, SE
25 = 4.28, $t = -3.23$), suggesting that object similarity priming could at least partially account for the effect of goal
26 similarity in the main analysis.

27 For NO responses, object similarity was systematically less facilitative than grip similarity (at 120 ms:
28 $\chi^2(1) = 12.06$, $p = .001$, $R^2 = 60.13^{10^{-4}}$; estimate = +13.32, SE = 3.83, $t = 3.48$; at 220 ms: $\chi^2(1) = 23.95$, $p = .001$,
29 $R^2 = 119.05^{10^{-4}}$; estimate = +19.32, SE = 3.93, $t = 4.91$) and goal similarity (at 120 ms: $\chi^2(1) = 19.94$, $p = .001$,
30 $R^2 = 84.32^{10^{-4}}$; estimate = +21.24, SE = 5.15, $t = 4.12$; at 220 ms: $\chi^2(1) = 28.73$, $p = .001$, $R^2 = 143.00^{10^{-4}}$;
31 estimate = +30.42, SE = 5.65, $t = 5.38$). Facilitative effects of grip or goal similarity for no responses were thus
32 not due to object similarity priming.

33 **Complementary analysis: influence of perceptual similarity**

34 As in the Experiment 1, Spearman's rank correlations were computed between the goal and grip similarity
35 priming effects obtained in each response condition and the corresponding perceptual similarity indices. There
36 were no significant correlations (all $p > .29$, see Appendix 2).

1 **Discussion**

2 The second experiment demonstrates the activation of both grip and goal representations from 120 ms of
3 processing of the action photographs. We used the same priming paradigm as in Experiment 1 with different prime
4 durations (120 and 220 ms) and with the addition of a neutral control prime condition (actor and object present
5 without any performed action) to assess the role of object information in the action priming effects observed.
6 Overall, we found that the recognition of correct actions was facilitated when action primes shared the same action
7 goal or the same grip as the target. This facilitation was visible from 120 ms of prime processing for “yes”
8 responses and the magnitude of the action priming effect was more important for goal than for grip similarity. For
9 “no” responses, goal and grip similarity priming effects were only significant at 220 ms.

10 The effect of response type in Experiment 2, although not predicted, is not at odds with the literature on action
11 processing (e.g., Yoon, Humphreys, & Riddoch, 2010). Van Elk et al. (2008) also found an effect of response type,
12 with an impact of action violations on action decisions only when the dimension of the action evaluated
13 (correctness of grip or correctness of goal) was correct. This may suggest that regardless of the action dimension
14 considered, correct representations have a greater weight than incorrect representations in the comprehension of
15 observed actions. In our experiment, goal and grip similarity priming effects were not completely absent for “no”
16 responses, but the emergence of both goal and grip priming effects was slightly delayed and becomes visible from
17 220 ms of prime duration, in contrast with priming effects for “yes” responses evidenced regardless of prime
18 duration from 120 ms. Although a better understanding of the reasons underlying this delay is needed, the effect
19 of response type on action priming effects does not impact the main result of Experiment 2, namely the appearance
20 of grip similarity priming from 120 ms of prime processing, and the consistency of the whole data pattern between
21 Experiments 1 and 2. At extremely short and long prime durations (Experiment 1), the effect of response type was
22 not significant since at 66 ms prime duration was far too short to trigger grip similarity priming and at 300 ms grip
23 similarity priming was already well established. Response type affected the prime durations at which grip
24 similarity priming became obvious, between 120 and 220 ms (Experiment 2).

25 The main finding of Experiment 2, particularly clearly highlighted for “yes” responses, shows both grip and
26 goal similarity priming effects from 120 ms of prime processing, with a greater magnitude of goal priming effects.
27 Note that moreover, as in Experiment 1, a cost for primes conveying incongruent grip and goal information (i.e.,
28 primes with partial violations) was observed from 120 ms regardless of response type, as reflected by significant
29 interactions between grip and goal similarity. Since both grip and goal representations have to be activated in order
30 to induce a cost when incongruent, this effect reinforces the idea that grip similarity priming emerges as early as
31 120 ms. Again, we can be fairly confident that differences in low-level perceptual similarity between prime-target
32 pairs cannot account for the pattern of priming effects. The stronger masking procedure used in Experiment 2
33 minimized the possible influence of low-level similarity priming and, more importantly, there was no correlation
34 between the amplitude of grip and goal priming effects and perceptual similarity indices in the different conditions
35 (Appendix 2).

36 Finally, the complementary analysis involving the neutral prime condition indicated that prime-target pairs sharing
37 mere object similarity showed an advantage over prime-target pairs with dissimilar goals, but not dissimilar grips
38 (in the “yes” response condition). This suggests that the advantage of prime-target pairs sharing goal similarity
39 over prime-target pairs with dissimilar goals (i.e., the goal priming effect in the main analysis) could be due to
40 mere object priming. However, the magnitude of the goal priming effect (34 ms) was twice as big as the object

1 priming effect (14 ms), indicating that object priming may only partially account for the goal similarity priming
2 effects observed for “yes” responses. One interpretation of this additional result is that observers may use some
3 information about object identity to activate a prediction about the goal of the action. We will go back to this
4 interpretation in the General Discussion.

5

6 **General discussion**

7 Whereas the involvement of grip and goal representations during the processing of others’ actions is well
8 accepted, their specific role and temporal organization remain debated. Bottom-up approaches propose that action
9 decoding is first guided by the processing of kinematic parameters, whereas top-down approaches suggest that
10 predictions about others’ goal precede the use of kinematic information. The present study evaluates these
11 hypotheses by assessing the timing of grip and goal activation during correctness judgments of photographs of
12 object-directed actions. The selective activation of grip or goal representations was identified using action primes
13 sharing the same grip or the same goal as target actions. Information about activation time course was captured by
14 varying prime duration (66 ms and 300 ms in Experiment 1; 120 ms and 220 ms in Experiment 2). Goal similarity
15 priming was visible for all prime durations whereas grip priming emerged from 120 ms of prime duration.
16 Interestingly, action primes conveying incongruent grip and goal information interfered with action recognition
17 when there was enough time for grip and goal representations to be co-activated, namely from 120 ms. These
18 results indicate that goal representations are activated first and provide additional evidence for distinct but
19 complementary roles of grip and goal representations during the processing of others’ actions.

20 Top-down approaches (Kilner 2011) have been supported by recent studies highlighting the involvement
21 of contextual information in action understanding (Iacoboni et al. 2005; Giglio et al. 2013; Amoruso and Urgesi
22 2016; Wurm and Schubotz 2016; Schenke et al. 2016). In such studies, contextual information is provided before
23 the presentation of the kinematic parameters of an action. They show that congruent context helps determining
24 action goals from observed kinematics. This demonstrates that contextual information can shape a priori
25 predictions of the observer about the goal of the action, which will influence the processing of kinematic
26 parameters. However, these studies do not allow determining at what moment of action processing a representation
27 of the action goal is activated in the absence of prior contextual information providing externally-driven
28 expectancies. In our protocol, information about the goal and the kinematics of object-directed actions is presented
29 simultaneously in action photographs. The hand grip on the object provides information about the kinematics of
30 the action (e.g. grasp water bottle) and the orientation of the object informs about the goal of the action (e.g. drink).
31 Thus, timing differences between goal and grip priming effects reflect differences in the time course of activation
32 of representations of action goals (e.g., representation of the object typical function) and kinematic parameters (i.e.
33 representation of the functional hand posture for object use) from the perceived object-related action. Our study
34 then makes a unique contribution by demonstrating the early and first activation of goal representations.

35 The photographs of object use actions involved the hand effector on the tool. One may argue that the
36 object is a context per se from which goal expectancies could be driven. Goals expectancies driven from the object
37 may thus lead to faster activation of goal representations. Although we consider the tool object an integral part of
38 object-directed actions, we acknowledge that prediction about the action goal can be driven from object identity
39 (e.g. water bottle: to drink). Yet this is also true for predictions regarding the grip parameter (e.g. water bottle:

1 whole hand clench). Consistent with the idea, recent theoretical assumptions assign a central role to object
2 information in both execution and perception of object-directed actions (Bach et al. 2014; van Elk et al. 2014). In
3 Experiment 2, mere object similarity priming showed a similar advantage as goal similarity priming, indicating
4 that processing of object identity may have led to the activation of a set of potential goals corresponding to object
5 typical use. Importantly, however, processing of object identity cannot fully account for the differences between
6 goal and grip priming effects. First, the magnitude of goal similarity priming was twice as big as the magnitude of
7 object priming. Second, grip similarity priming, in contrast, was not similar to object priming even if the object
8 could easily activate grip representations. Together, the pattern of results demonstrate that object-directed actions
9 tend to be processed top-down in accordance to predictive accounts of action understanding. The observer uses
10 the available information to make first a prediction about the goal of the action (rather than a prediction about the
11 grip of the action) that will be then confronted to the processing of the kinematic parameters.

12 Finally, the late activation of grip representations in our experiments adds to the inconsistent and widely
13 discussed literature on the role of kinematic information in the processing of observed actions. Some evidence
14 clearly indicates that processing of kinematic parameters fuels decisions about action goals (Ansuini et al. 2014;
15 Quesque and Coello 2015; Cavallo et al. 2016). Still, despite the poor contextual information provided, kinematic
16 processing may still be guided by predictions about possible action goals, especially in studies requiring decisions
17 between a limited subset of possible goals. In Lewkowicz et al. (2015) for example, video clips displayed an actor's
18 arm reaching and moving a dowel in order to move it away or take it for himself. Observers may first pre-activate
19 the only two possible goals (move away versus take it) and then use kinematics information to verify each
20 hypothesis. Actions would be processed in a predictive manner. Finally, the use of kinematic information in action
21 understanding may be flexible, context- and task-dependent. Some experiments have found that on-line kinematic
22 information is not necessary or not always used in the processing of observed actions (Longo et al. 2008; Neal and
23 Kilner 2010; Naish et al. 2013; Geangu et al. 2015). In parallel, reliance on kinematic information is increased
24 when the object is absent (Thioux and Keysers 2015), when the action is difficult to recognize (Nicholson et al.
25 2017) or when the recognition of the action requires the processing of fine-grained kinematic information (Tidoni
26 et al. 2013). Thus, a strong possibility is that actions performed by others can be flexibly decoded through top-
27 down or bottom-up mechanisms, depending of contextual variables (e.g., object presence, visual conditions,
28 contextual scene, social factors, or even kinematic information provided by others' actions see for example Manera
29 et al. 2011) and task demands.

30 **Limitations to the study**

31 There are a few limitations to this study that deserve consideration. First, participant samples were not
32 gender balanced and although there was no obvious a priori or a posteriori evidence of gender effects, the
33 possibility of gender modulation of action processing cannot be ruled out. Second, the set of object stimuli was
34 relatively heterogeneous and included both tools (objects related to specific usage) and non-tools (objects with
35 many possible usages). It is thus difficult to determine from the present study whether object status has an impact
36 on action decoding. Third, we used static action stimuli for a strict control of grip and goal visual information
37 input. This leaves the issue of the generalization of the present findings to naturalistic, dynamic action perception
38 open. Last but not least, although grip and goal information were manipulated independently in the present design,
39 some asymmetry in the relation between the two pieces of information remains: an incorrect grip can still be used

1 to achieve a correct action, as defined as the typical use of the object (e.g., water bottle: grasp-to-drink), whereas
2 an incorrect goal (e.g., water bottle up-side down) prevents the achievement of this action. This is not a strict
3 asymmetry, however, since other possible and meaningful actions may be achieved in the incorrect goal conditions,
4 even if less probable and/or less associated to object function (e.g., grasp-to-move actions when transporting
5 objects are possible regardless of object orientation). In order to evaluate whether goal information remains first
6 processed when it is less predictive of action correctness, future research could attempt to cross two alternatives
7 (and ideally equally likely) grips and goals for a given object (e.g. pour versus drink from cup).

8 **Conclusion**

9 To conclude, the present study provides evidence for the activation of distinct goal and kinematic
10 parameters representations within the first 120 ms of processing of observed actions, with an early and first
11 activation of goal representations. Results also suggest that early goal representations may have been at least
12 partially derived from information related to object identity. These results are in line with top-down predictive
13 approaches of action understanding in which a prediction about the action goal is first computed and then compared
14 with the processing of kinematic parameters. Future research should examine to what extent the timing of goal and
15 grip activation is sensitive to context and task demands. Context-dependent changes in grip and goal temporal
16 dynamics would support the hypothesis of flexible bottom-up and top-down mechanisms in action decoding.

1 **Acknowledgments**

2 This work was funded by the French National Research Agency (ANR-16-CE28-0003 and ANR-11-
3 EQPX-0023) and benefited from a regional fellowship (Hauts-de-France) to J. Decroix.

4 **References**

- 5 Amoruso L, Urgesi C (2016) Familiarity modulates motor activation while other species' actions are observed:
6 A magnetic stimulation study. *Eur J Neurosci* 43:765–772. doi: 10.1111/ejn.13154
- 7 Ansuini C, Cavallo A, Bertone C, Becchio C (2014) The visible face of intention: Why kinematics matters. *Front*
8 *Psychol* 5:1–6. doi: 10.3389/fpsyg.2014.00815
- 9 Avanzini P, Fabbri-Destro M, Campi C, et al (2013) Spatiotemporal dynamics in understanding hand-object
10 interactions. *Proc Natl Acad Sci U S A* 110:15878–85. doi: 10.1073/pnas.1314420110
- 11 Bach P, Nicholson T, Hudson M (2014) The affordance-matching hypothesis: how objects guide action
12 understanding and prediction. *Front Hum Neurosci* 8:254. doi: 10.3389/fnhum.2014.00254
- 13 Barr DJ, Levy R, Scheepers C, Tily HJ (2013) Random effects structure for confirmatory hypothesis testing:
14 Keep it maximal. *J Mem Lang* 68:255–278. doi: 10.1016/j.jml.2012.11.001
- 15 Barsalou LW (2008) Grounded Cognition. *Annu Rev Psychol* 59:617–645. doi:
16 10.1146/annurev.psych.59.103006.093639
- 17 Barton K (2016) MuMIn: Multi-Model Inference.
- 18 Bates D, Kliegl R, Vasishth S, Baayen H (2015a) Parsimonious mixed models. *arXiv Prepr arXiv150604967* 1–
19 27. doi: arXiv:1506.04967
- 20 Bates D, Mächler M, Bolker B, Walker S (2015b) Fitting Linear Mixed-Effects Models Using **lme4**. *J Stat*
21 *Softw*. doi: 10.18637/jss.v067.i01
- 22 Catmur C (2015) Understanding intentions from actions: Direct perception, inference, and the roles of mirror and
23 mentalizing systems. *Conscious Cogn* 1–8. doi: 10.1016/j.concog.2015.03.012
- 24 Cattaneo L, Sandrini M, Schwarzbach J (2010) State-dependent TMS reveals a hierarchical representation of
25 observed acts in the temporal, parietal, and premotor cortices. *Cereb Cortex* 20:2252–2258. doi:
26 10.1093/cercor/bhp291
- 27 Cavallo A, Heyes C, Becchio C, et al (2014) Timecourse of mirror and counter-mirror effects measured with
28 transcranial magnetic stimulation. *Soc Cogn Affect Neurosci* 9:1082–1088. doi: 10.1093/scan/nst085
- 29 Cavallo A, Koul A, Ansuini C, et al (2016) Decoding intentions from movement kinematics. *Sci Rep* 6:37036.
30 doi: 10.1038/srep37036
- 31 Geangu E, Senna I, Croci E, Turati C (2015) The effect of biomechanical properties of motion on infants'
32 perception of goal-directed grasping actions. *J Exp Child Psychol* 129:55–67. doi:
33 10.1016/j.jecp.2014.08.005

- 1 Gentsch A, Weber A, Synofzik M, et al (2016) Towards a common framework of grounded action cognition:
2 Relating motor control, perception and cognition. *Cognition* 146:81–89.
- 3 Giglio ACA, Minati L, Boggio PS (2013) Throwing the banana away and keeping the peel: Neuroelectric
4 responses to unexpected but physically feasible action endings. *Brain Res* 1532:56–62. doi:
5 10.1016/j.brainres.2013.08.017
- 6 Grafton ST, Hamilton AFDC (2007) Evidence for a distributed hierarchy of action representation in the brain.
7 *Hum Mov Sci* 26:590–616. doi: 10.1016/j.humov.2007.05.009
- 8 Hudson M, Nicholson T, Ellis R, Bach P (2016a) I see what you say: Prior knowledge of other’s goals
9 automatically biases the perception of their actions. *Cognition* 146:245–250. doi:
10 10.1016/j.cognition.2015.09.021
- 11 Hudson M, Nicholson T, Simpson WA, et al (2016b) One step ahead: The perceived kinematics of others’
12 actions are biased toward expected goals. *J Exp Psychol Gen* 145:1–7. doi: 10.1037/xge0000126
- 13 Iacoboni M, Molnar-Szakacs I, Gallese V, et al (2005) Grasping the intentions of others with one’s own mirror
14 neuron system. *PLoS Biol* 3:0529–0535. doi: 10.1371/journal.pbio.0030079
- 15 Jacquet PO, Avenanti A (2015) Perturbing the action observation network during perception and categorization
16 of actions’ goals and grips: State-dependency and virtual lesion TMS effects. *Cereb Cortex* 25:598–608.
17 doi: 10.1093/cercor/bht242
- 18 Kalénine S, Shapiro AD, Buxbaum LJ (2013) Dissociations of action means and outcome processing in left-
19 hemisphere stroke. *Neuropsychologia* 51:1224–1233. doi: 10.1016/j.neuropsychologia.2013.03.017
- 20 Kilner JM (2011) More than one pathway to action understanding. *Trends Cogn Sci* 15:352–357. doi:
21 10.1016/j.tics.2011.06.005
- 22 Kilner JM, Friston KJ, Frith CD (2007) Predictive coding: An account of the mirror neuron system. *Cogn*
23 *Process* 8:159–166. doi: 10.1007/s10339-007-0170-2
- 24 Kristjansson A (2008) “I know what you did on the last trial” - a selective review of research on priming in
25 visual search. *Front Biosci* 13:1171. doi: 10.2741/2753
- 26 Lepage JF, Tremblay S, Théoret H (2010) Early non-specific modulation of corticospinal excitability during
27 action observation. *Eur J Neurosci* 31:931–937. doi: 10.1111/j.1460-9568.2010.07121.x
- 28 Longo MR, Kosobud A, Bertenthal BI (2008) Automatic imitation of biomechanically possible and impossible
29 actions: effects of priming movements versus goals. *J Exp Psychol Hum Percept Perform* 34:489–501. doi:
30 10.1037/0096-1523.34.2.489
- 31 Manera V, Becchio C, Schouten B, et al (2011) Communicative interactions improve visual detection of
32 biological motion. *PLoS One*. doi: 10.1371/journal.pone.0014594
- 33 Matuschek H, Kliegl R, Vasishth S, et al (2017) Balancing Type I error and power in linear mixed models. *J*
34 *Mem Lang* 94:305–315. doi: 10.1016/j.jml.2017.01.001

- 1 Naish KR, Houston-Price C, Bremner AJ, Holmes NP (2014) Effects of action observation on corticospinal
2 excitability: Muscle specificity, direction, and timing of the mirror response. *Neuropsychologia* 64:331–
3 348. doi: 10.1016/j.neuropsychologia.2014.09.034
- 4 Naish KR, Reader AT, Houston-Price C, et al (2013) To eat or not to eat? Kinematics and muscle activity of
5 reach-to-grasp movements are influenced by the action goal, but observers do not detect these differences.
6 *Exp Brain Res* 225:261–275. doi: 10.1007/s00221-012-3367-2
- 7 Neal A, Kilner JM (2010) What is simulated in the action observation network when we observe actions? *Eur J*
8 *Neurosci* 32:1765–1770. doi: 10.1111/j.1460-9568.2010.07435.x
- 9 Nicholson T, Roser M, Bach P (2017) Understanding the goals of everyday instrumental actions is primarily
10 linked to object, not motor-kinematic, information: Evidence from fMRI. *PLoS One* 12:1–21. doi:
11 10.1371/journal.pone.0169700
- 12 Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*
13 9:97–113. doi: 10.1016/0028-3932(71)90067-4
- 14 Ortigue S, Thompson JC, Parasuraman R, Grafton ST (2009) Spatio-temporal dynamics of human intention
15 understanding in temporo-parietal cortex: A combined EEG/fMRI repetition suppression paradigm. *PLoS*
16 *One*. doi: 10.1371/journal.pone.0006962
- 17 Quesque F, Coello Y (2015) Perceiving what you intend to do from what you do: evidence for embodiment in
18 social interactions. *Socioaffective Neurosci Psychol* 5:28602. doi: 10.3402/snp.v5.28602
- 19 R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical
20 Computing, Vienna
- 21 Rizzolatti G, Fogassi L (2014) The mirror mechanism: recent findings and perspectives. *Philos Trans R Soc*
22 *Lond B Biol Sci* 369:20130420. doi: 10.1098/rstb.2013.0420
- 23 Schenke KC, Wyer NA, Bach P (2016) The Things You Do: Internal Models of Others' Expected Behaviour
24 Guide Action Observation. *PLoS One* 11:e0158910. doi: 10.1371/journal.pone.0158910
- 25 Thioux M, Keysers C (2015) Object visibility alters the relative contribution of ventral visual stream and mirror
26 neuron system to goal anticipation during action observation. *Neuroimage* 105:380–394. doi:
27 10.1016/j.neuroimage.2014.10.035
- 28 Tidoni E, Borgomaneri S, di Pellegrino G, Avenanti A (2013) Action simulation plays a critical role in deceptive
29 action recognition. *J Neurosci* 33:611–23. doi: 10.1523/JNEUROSCI.2228-11.2013
- 30 van Elk M, Bousardt R, Bekkering H, van Schie HT (2012) Using goal- and grip-related information for
31 understanding the correctness of other's actions: An ERP study. *PLoS One* 7:1–8. doi:
32 10.1371/journal.pone.0036450
- 33 van Elk M, van Schie H, Bekkering H (2014) Action semantics: A unifying conceptual framework for the
34 selective use of multimodal and modality-specific object knowledge. *Phys Life Rev* 11:220–250. doi:

1 10.1016/j.pprev.2013.11.005

2 van Elk M, Van Schie HT, Bekkering H (2008) Conceptual knowledge for understanding other's actions is
3 organized primarily around action goals. *Exp Brain Res* 189:99–107. doi: 10.1007/s00221-008-1408-7

4 Wolpert D, Doya K, Kawato M (2003) A unifying computational framework for motor control and social
5 interaction. *Philos Trans R Soc Lond B Biol Sci* 358:593–602. doi: 10.1098/rstb.2002.1238

6 Wurm MF, Lingnau A (2015) Decoding Actions at Different Levels of Abstraction. *J Neurosci* 35:7727–7735.
7 doi: 10.1523/JNEUROSCI.0188-15.

8 Wurm MF, Schubotz RI (2012) NeuroImage Squeezing lemons in the bathroom : Contextual information
9 modulates action recognition. *Neuroimage* 59:1551–1559. doi: 10.1016/j.neuroimage.2011.08.038

10 Wurm MF, Schubotz RI (2016) What's she doing in the kitchen? Context helps when actions are hard to
11 recognize. *Psychon Bull Rev* 1–7. doi: 10.3758/s13423-016-1108-4

12 Yoon EY, Humphreys GW, Riddoch MJ (2010) The paired-object affordance effect. *J Exp Psychol Hum Percept*
13 *Perform* 36:812–824. doi: 10.1037/a0017175

14 Zentgraf K, Munzert J, Bischoff M, Newman-Norlund RD (2011) Simulation during observation of human
15 actions - Theories, empirical studies, applications. *Vision Res* 51:827–835. doi:
16 10.1016/j.visres.2011.01.007

17 Zhang L, Zhang L, Mou X, Zhang D (2011) FSIM: A feature similarity index for image quality assessment.
18 *IEEE Trans Image Process* 8:2378–2386.

19

1 APPENDIX 1: List of objects used in Experiments 1 and 2

Carafe
Coffee cup
Fork
Hairbrush
Hairdryer
Hammer
Knife
Liquid soap
Magnifying glass
Pencil
Phone
Screwdriver
Tea spoon
Teapot
Toothbrush
Torch
Cream Tube
Water bottle
Watering can
Wine glass

2

1 APPENDIX 2

2 Priming paradigms are known to be influenced by low-level visual features (e.g., Kristjansson, 2008). Thus, our
3 results may be driven by perceptual differences between conditions and not by the activation of different levels of
4 action representation. For example, earlier and stronger priming effects may be expected between pairs of action
5 pictures that are more similar perceptually. We used the FSIM algorithm developed by Zhang, Zhang, Mou &
6 Zhang (2011) to assess image similarity based on low-level visual features. The closer to 1 the index is the more
7 pictures are similar perceptually. First, an index of perceptual similarity was computed between each type of prime
8 type (correct grip but incorrect goal, incorrect grip but correct goal, incorrect grip and incorrect goal, neutral action-
9 free) and target (correct action target, incorrect action target) and then attributed to each type of prime-target pairs
10 (Grip similar only, goal similar only, all different, neutral). No indices were computed for grip and goal similar
11 pairs, as prime and target were the same exact picture.

12 Perceptual similarity was modelled as a function of prime-target pairs (Grip similar only, goal similar only, all
13 different, neutral) and RESP (yes, no) as fixed effects and items as random intercepts. Model comparison did not
14 show any interaction between RESP and type of pair ($\chi^2(3) = 2.5356, p = 0.4689$) and no main effect of RESP
15 ($\chi^2(1) = 1.7477, p = 0.1862$). We did observe however a main effect of type of pair ($\chi^2(3) = 213.11, p < 0.001$).
16 Compared to ‘all different’ pairs, perceptual similarity was higher in ‘similar grip only’ pairs (estimate = 0.055,
17 SE = 0.004, $t = 12.36, p < 0.001$), but lower in neutral pairs (estimate = -0.044, SE = 0.004, $t = -9.96, p < 0.001$).
18 ‘All different’ and ‘goal similar only’ pairs were not different from each other (estimate = 0.007, SE = 0.004, $t =$
19 1.67, $p = 0.09$).

20 To assess the relation between grip and goal priming effects and the perceptual similarity indices in the different
21 goal and grip similarity conditions, Spearman’s rank correlations were computed. Table 2 summarizes the results
22 of the analysis.

23 [INSERT TABLE 2]

1 APPENDIX 3A: Determination of the random effect structure in mixed-effect models of Experiment 1

2

3 Following the recommendation of Barr et al. (2013), we first attempted to fit a model with the maximum random
4 structure, including, for both subjects and items, random intercepts, random slopes for GRIP, GOAL,
5 DURATION, and RESP and random slopes for the 4 possible interactions between GRIP, GOAL, DURATION,
6 and RESP. As expected, the model was overparametrized in regards to the data available and failed to converge to
7 a stable solution (see Bates, Kliegl, Vasisht & Baayen, 2015; Matuschek, Kliegl, Vasisth, Baayen & Bates, 2017).
8 Then following the latest recommendations of Bates et al. (2015) we determined the optimal random structure
9 supported by the data. First higher-order random slopes (random slopes reflecting interactions) were removed.
10 Then we looked for possible further reduction of the remaining random effect structure by conducting a principal
11 component analysis on the random terms of the model using the rePCA function from the RePsychLing package
12 version 0.0.4 developed by Bates et al. (2015). The analysis identified 5 components, three of them being sufficient
13 to explain 100% of variance. The random slope for DURATION was the least representative factor of these three
14 components for subject and was thus removed from the random effect structure of the model. For items, the least
15 representative factor was GRIP and GOAL, the two factors were then removed from the random effect structure
16 of the model. Consequently, the final model included random intercepts, GRIP, GOAL and RESP as random slopes
17 for subject, and random intercepts, DURATION and RESP as random slopes for items.

18 The following full model was finally considered in Experiment 1 (Bold indicates fixed effects of interest)

19 RT = **GRIP**+ **GOAL** + DURATION + RESP
20 + GRIP*DURATION + GOAL*DURATION
21 + **GRIP*GOAL** + GRIP*RESP + GOAL*RESP + DURATION*RESP
22 + **GRIP*GOAL*DURATION** + **GRIP*GOAL*RESP** + GRIP*DURATION*RESP + GOAL*DURATION*RESP
23 + **GRIP*GOAL*DURATION*RESP**
24 + (GRIP+GOAL+RESP | Participant) + (DURATION +RESP| Item)

25 APPENDIX 3B: Determination of random effect structure in mixed effect models of Experiment 2

26 As for Experiment 1, the maximum random structure was not suited for the data of Experiment 2. Moreover, the
27 full model did not converge with GRIP, GOAL, DURATION, and RESP as random slopes for participants and
28 items. A first principal component analysis on the random terms of the model indicated that the GRIP, GOAL and
29 DURATION for subjects were the least representative factors of the components explaining 100% of variance and
30 were removed from the analysis. For items, the least representative factors explaining 100% of variance were
31 GRIP and GOAL and were then removed from the analysis. Thus, the final model included random intercepts and
32 random slopes for RESP for subjects, and random intercepts, RESP and DURATION as random slopes for items.

33 The following full model was finally considered in Experiment 2 (Bold indicates fixed effects of interest)

34 RT = **GRIP**+ **GOAL** + DURATION + RESP
35 + GRIP*DURATION + GOAL*DURATION
36 + **GRIP*GOAL** + GRIP*RESP + GOAL*RESP + DURATION*RESP
37 + **GRIP*GOAL*DURATION** + **GRIP*GOAL*RESP** + GRIP*DURATION*RESP + GOAL*DURATION*RESP
38 + **GRIP*GOAL*DURATION*RESP**
39 + (RESP | Participant) + (DURATION +RESP| Item)

1 **Table Captions**

2 **Table 1** Mean correct reaction times and standard deviations according to prime-target pairs, prime duration (66,
3 300 for Experiment 1; 120, 220ms for Experiment 2) and response type in Experiment 1 and 2.

4 **Table 2** Spearman's rank correlations between goal and grip similarity priming effects and the corresponding
5 perceptual similarity difference scores

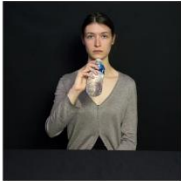
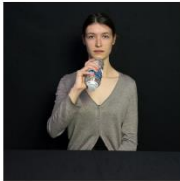
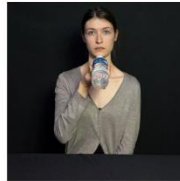
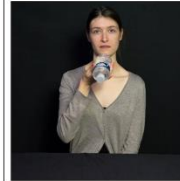
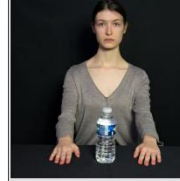

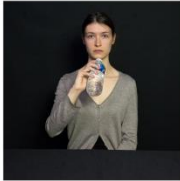

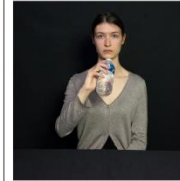
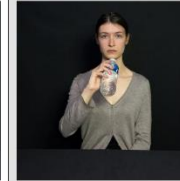



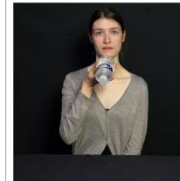


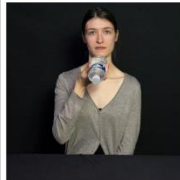
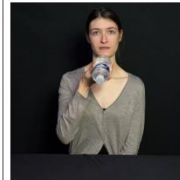
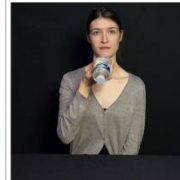
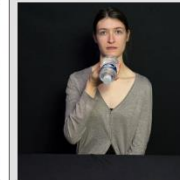
6

1 **Figure Captions**

2 **Fig. 1** Example of action photographs containing grip and/or goal violations for a given stimulus. In this example,
3 the actress grasps the upright (full and closed) water bottle with a typical clench grip, a correct action regarding
4 the typical function of the object (drink). In the photograph presenting a goal violation, the actress grasps the bottle
5 with the same typical clench grip but the bottle is upside down, preventing her from drinking. In the photograph
6 presenting a grip violation, the actress grasps the bottle with an atypical grip from below that does not prevent her
7 from drinking. The photograph combining grip and goal violations displays the actress grasping the upside-down
8 bottle with the atypical grip. Action photographs are then divided into eight prime-target pairs according grip
9 similarity, goal similarity and response type. An additional neutral prime-target pair was included in Experiment
10 2. The FSIM index provides a measure of low-level perceptual similarity between prime and target.

11 **Fig. 2** Individual mean priming effects in milliseconds according to grip similarity, goal similarity, prime duration
12 and response type. Prime durations of 66 and 300 ms were tested in Experiment 1, and prime durations of 120 and
13 220 ms were evaluated in Experiment 2. Boxes represent the interquartile range. Black horizontal lines represent
14 the medians of the distribution. Computed with the pirateplot function of R (R version 3.3.2 (R Core Team 2017)
15 and R Studio version 1.0.136.).

16

Response Type	YES				
Prime-target pairs	Grip and goal similar	Grip similar only	Goal similar only	All different	Neutral (Experiment 2)
Prime					
Target					
Grip similarity Priming effect	Shared	Shared	Different	Different	NA
Goal similarity priming effect	Shared	Different	Shared	Different	NA
FSIM perceptual index mean(SD)	1	0,91 (+/-0,03)	0,87 (+/-0,04)	0,86 (+/-0,04)	0,82 (+/-0,03)
Response Type	NO				
Prime-target pairs	All different	Goal similar only	Grip similar only	Grip and goal similar	Neutral (Experiment 2)
Prime					
Target					
Grip similarity Priming effect	Different	Different	Shared	Shared	NA
Goal similarity priming effect	Different	Shared	Different	Shared	NA
FSIM perceptual index mean(SD)	0,86 (+/-0,04)	0,87 (+/-0,04)	0,91 (+/-0,03)	1	0,81 (+/-0,03)

1

2

