Registered Report

Anticipating actions and corticospinal excitability: A preregistered motor TMS experiment

Oliver Genschow a,*, Lara Bardi b and Marcel Brass b

a University of Cologne, Germany
b Ghent University, Belgium

Article info
Article history:
Protocol received: 16 August 2016
Protocol accepted: 12 May 2017
Received 17 January 2018
Reviewed 7 February 2018
Accepted 28 April 2018
Action Editor Chris Chambers
Published online 25 May 2018

Keywords:
Anticipation
Prediction
Corticospinal excitability
Motor TMS
Imitation

Abstract
Past research on action observation and imitation suggests that observing a movement activates a corresponding motor representation in the observer. However, recent research suggests that individuals may not only reflexively simulate the observed behavior but also simulate and engage in anticipated action without another person actually engaging in it. For example, it has been demonstrated that observing a triggering event (i.e., nose wrinkling) that potentially leads to the anticipation of an action (i.e., nose scratching) increases the likelihood that the observer will perform that action. In the present research, we applied motor Transcranial Magnetic Stimulation (TMS) to investigate such anticipated social action effects at the neurophysiological level within a trial-by-trial measure. While a pilot study suggests that observing nose wrinkling elicits stronger motor evoked potentials (MEPs) in participants' biceps muscles than observing control events, this effect could not be fully replicated in a preregistered study. Although a post hoc meta-analysis across both studies supports the general hypothesis, these results need to be taken cautiously.

Implications of the results reported in the manuscript are discussed.

© 2018 Elsevier Ltd. All rights reserved.

Individuals tend to automatically imitate a wide range of different behaviors (for a meta-analysis, see Cracco, Bardi, Desmet, Genschow, Rigoni, & De Coster, 2018), such as facial expressions (Dimberg, 1982), language (Cappella & Planalp, 1981; Giles & Powsland, 1975; Webb, 1969, 1972), emotions (for an overview, see Hess & Fischer, 2013), postures (LaFrance, 1982), gestures (Bernier, 1988; Cracco, Genschow, Radkova, & Brass, 2018), or simple movements (Brass, Bekkering, Wohlschlager, & Prinz, 2000; Genschow & Florack, 2014; Genschow & Schindler, 2016; Genschow, Florack, & Wänke, 2013; Genschow et al., 2017)—to name just a few examples.

Classical perception-action theories in social psychology (for an overview, see Chartrand & Dalton, 2009) and cognitive psychology (for an overview, see Heyes, 2011) suggest that such imitative phenomena are based on shared representations of observed and executed actions. Ideomotor theory (Greenwald, 1970; Prinz, 1990, 1997), for example, explains imitative response tendencies by stating that the observation of an action primes and thus facilitates the execution of a compatible action, because observed and executed actions activate the same sensory representations.

* Corresponding author. University of Cologne, Social Cognition Center Cologne, Richard-Strauss Str. 2, 50931, Köln, Germany.
E-mail address: oliver.genschow@uni-koeln.de (O. Genschow).
https://doi.org/10.1016/j.cortex.2018.04.014
0010-9452/© 2018 Elsevier Ltd. All rights reserved.
At the neurophysiological level, imitation has predominantly been explained in relation to the mirror neuron system (e.g., Bienen, Roebroeck, Goebel, & Sakai, 2009; Brass & Heyes, 2005; Catmur, Walsh, & Heyes, 2009; Cook & Bird, 2011). The mirror neuron system refers to a network of motor areas in the frontal and parietal cortex that do not only respond to action execution but also to action observation (Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010). Although the mirror neuron system was initially documented in the monkey brain (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996), there is now converging evidence that a similar system exists in humans as well (Molenberghs, Cunnington, & Mattingley, 2012). Particular support for the idea that the motor system is active during action execution as well as action observation was provided by different neurophysiological experiments (for a review, see Caspers, Zilles, Laird, & Eickhoff, 2010), including motor Transcranial Magnetic Stimulation (TMS; Fadiga, Craighero, & Olivier, 2005; Naish, Houston-Price, Brenner, & Holmes, 2014; Urgesi, Candidi, Fabbro, Romani, & Aglioti, 2006a). For example, Fadiga, Fogassi, Pavesi, and Rizzolatti (1995) stimulated the primary motor cortex of human subjects with TMS during the observation of hand movements and measured Motor Evoked Potentials (MEPs) from subjects’ hand muscles. The researchers found during action observation an increase in the MEP amplitude in participants’ hand muscles that would be used to execute the observed movements (see also Urgesi et al., 2006a).

In sum, the above-reviewed literature indicates that observing a whole action sequence in someone else directly activates the corresponding action plans in the observer. However, would already the anticipation of a future action be sufficient to activate corresponding motor actions too? Although there is not yet a clear answer to this question, there is reason to believe that this might, indeed, be the case. For instance, recent theoretical models (e.g., Kilner, Friston, & Frith, 2007; Wilson & Knoblich, 2005) assume that during action observation, individuals tend to constantly simulate other persons’ ongoing actions and infer its behavioral outcomes in order to prepare one’s own actions (see also Lamme, Fischer, & Decety, 2007). Kilner et al. (2007) propose a computational approach, which assumes that the mirror neuron system (Rizzolatti, Fogassi, & Gallese, 2001) infers the most likely intention of an observed action by minimizing the prediction error at all levels of the cortical hierarchy involved in action observation. Similarly, Wilson and Knoblich (2005) propose a so-called emulator that internally simulates others’ action execution. This simulation process then provides immediate information about the ongoing course of the observed action as well as its probable immediate future.

Evidence for such a physiological simulation process comes from a series of neurophysiological experiments (e.g., Cardellicchio, Sinigaglia, & Costantini, 2013; Kilner, Vargas, Duval, Blakemore, & Sirigu, 2004; Pierno et al., 2006; Umiltà et al., 2001; Urgesi et al., 2010). For instance, in a seminal study, Umiltà et al. (2001) measured mirror neuron activations in macaque monkeys while the monkeys observed a fully visible action directed towards an object or the same action with its endpoint being hidden. The results show that the majority of mirror neurons became active during the whole action presentation, but also when the final action towards the object was hidden. This result was taken as evidence for the hypothesis that motor representations of an action performed by others can be internally generated in the observer even when the complete visual description of the action is lacking.

In a related study, Kilner et al. (2004) measured the Readiness Potential (RP)—an electrophysiological marker of motor preparation—while human participants observed different video clips of another person. In half of the video clips, participants observed a hand movement grasping an object. In the other half of the videos the hand remained stationary. At the beginning of each clip, a color cue indicated whether the hand would move or remain in the same position. When the onset time of the upcoming arm movement was predictable, a rise of the RP was observed before the actual movement’s onset. This result suggests that the mere knowledge of an upcoming movement is sufficient to activate one’s own motor system. In a similar study, Urgesi et al. (2010) presented participants snapshots of hand movements while applying motor TMS. The snapshots either depicted the starting, the middle or the end phase of a movement. The researchers found that observing a movement’s start phase and middle phase engendered significantly higher motor facilitation than observing the final phase.

In sum, neurophysiological studies indicate that the motor system is active when individuals observe an action. Moreover, recent research suggests that individuals also simulate anticipated actions activating the motor system as well. However, despite first support for anticipative mechanisms, it is important to note that in most of the previous experiments participants were aware of the next following action. That is, participants knew what kind of movement would follow, because they had seen the model executing the movement in previous trials, or a cue announced the movement.

With reference to “anticipated action” we went one step further and recently demonstrated that when observing another person, individuals actually engage in the action that the other person might show in the near future although the other person never engages in this action and no cue announces the action (Genscheow & Brass, 2015; Genscheow, Klomfar, d’Haene, & Brass, 2018). For instance, in two experiments Genscheow and Brass (2015) tested whether the observation of an event that could potentially trigger another person’s action is sufficient to produce the anticipated action in the observer. Importantly, participants were never exposed to the actual movement. In a first experiment participants watched two 10 min lasting videos of a female model who was reading a story. In one video, the model was constantly scratching actions and in the other video her hair was constantly falling into her face. While watching the two videos, participants were videotaped. Afterwards, we coded how often participants engaged in anticipated actions related to nose wrinkling (e.g., nose scratching) and hair falling (e.g., hair stroking). The results gave first evidence for our hypothesized anticipated action effect: when watching the nose wrinkling video, participants engaged in more nose scratching actions than hair stroking actions and vice versa for watching the hair falling video. In a second experiment we tested whether anticipated action is based on the inference of the model’s desire to act. The results demonstrated that
participants engaged less often in anticipated action when we experimentally reduced the model's desire to act.

Despite first evidence for anticipated action and the assumption that this effect is due to an inferred desire to act, our behavioral studies do not reveal its specific characteristics. In fact, the experiments used a setup in which participants observed a number of triggering events (e.g., nose wrinkling) over a long period of time. To measure anticipated action, we simply counted how often participants carried out the target actions over a period of 10 min. While we can infer from this methodological setup that observing the triggering event increases the likelihood of engagement in the anticipated action, this approach has several disadvantages. First, because within such paradigms, researchers observe participants over a certain time period and then calculate a sum score of executed actions, it is not possible to measure the degree to which each model's event directly triggers action tendencies in the observer. Second, this methodological approach depends on subjective ratings of coders and might, thus, lack accuracy and reliability. An approach that would allow for a more accurate and reliable way to measure to which degree each single event leads to a response in the participant is to measure anticipative processes on a trial-by-trial basis (for a more detailed discussion, see Genschow, Van Den Bossche, Cracco, Bardi, Rigoni, & Brass, 2017).

Moreover, with reference to local or stimulus enhancement, our previous studies leave open an alternative explanation that is related to an ongoing debate in the literature on non-human primate imitation (e.g., Tomasello & Call, 1997). That is, it might be that observing someone repeatedly wrinkling his or her nose slowly guides participants' attention towards the nose, activating the motor system and increasing the activity in those muscles that would allow for executing movements towards this body part. Based on such an explanation, it would not be the inferred desire to act that triggers anticipated action, but rather the directed attention towards a specific body part.

In the present research, we aimed at applying an accurate and reliable measure that assess on a trial-by-trial basis the degree of anticipated action on a neurophysiological level. In addition, we aimed at testing to which degree attention plays a role in anticipated action. To this end, we modified our stimulus material and applied motor TMS—a method that allows for accurately and reliably measuring the activation of the motor system on a trial-by-trial basis (cf. Clark, Tremblay, & Ste-Marie, 2004). Single-pulse TMS over the primary motor cortex (M1) can induce (motor evoked potential MEP) in a peripheral muscle and the amplitude of evoked MEPs indexes the excitability of the cortico-spinal tract. Here, a trial-by-trial motor TMS procedure was carried out. That is, we measured MEPs from participants' biceps muscles while they observed video clips of a person engaging in a triggering event (i.e., nose wrinkling) or a control event (i.e., frowning). The biceps muscle was chosen as being the muscle that would be involved when participants would lift their arm in order to scratch their nose. If observing the triggering event directly activates the motor system in the observer, we should observe larger MEPs in the biceps when participants observe the nose-wrinkling event as compared to the frowning event. Moreover, if the observation of a triggering event directly activates the motor system, one should expect an increase in participants' MEPs already at an early point in time. To get further information on the temporal unfolding of anticipated action, we thus applied motor TMS at three different points in time: at the beginning, in the middle and at the end of each video clip.

To test the possibility that anticipated action is based on an attentional effect, we added a condition in which a feather tickled the model's nose without the model displaying any nose wrinkling. Therefore, participants' attention was directed towards the nose without the model displaying any itching sensation. If the attention explanation is correct, one would expect similarly high MEPs in the feather condition as in the condition where the model is engaging in nose wrinkling without any external stimulation. If, however, the inferred-desire hypothesis accounts, one would expect smaller effects in the feather condition as compared to the wrinkling condition, because in the feather condition the model does not engage in any nose wrinkling action. Participants may, thus, infer that the model's desire to scratch the nose is not very strong.

In order to test our research questions, we first conducted a pilot study. Based on these results we specified our hypotheses and then ran a power analysis in order to estimate the number of participants needed for a highly powered experiment. With this number in mind we then ran a preregistered experiment.

1. **Pilot study**

1.1. **Method**

1.1.1. **Ethics statement**

The ethics committee at Ghent University, Belgium, approved the procedure and method of the study.

1.1.2. **Data availability**

The data file of the study is available from the Open Science Framework database. The URL necessary to access our data is: https://osf.io/5pkys/?view_only=a5defabba53d4f56b07d1caca1db699.

1.1.3. **Participants and design**

We recruited participants via the participant pool of the Department of Experimental Psychology at Ghent University in exchange for 20 €. Nineteen participants accepted our invitation. Based on initial problems with measuring MEPs at participants' biceps muscle, we could not detect a reliable EMG signal for four participants. Moreover, we excluded one participant due to too few remaining trials (N < 10) per condition after cleaning the data. The final sample, thus, contained 14 right-handed male participants with an age ranging from 19 to 42 (M = 24.29; SD = 6.78). We tested only male participants due to restrictions by the local university's ethics committee. We applied a 2 (itching: yes vs no) × 2 (displayed actions: facial actions vs feather actions) × 3 (timing: start vs middle vs end) within-subject design. While the factor timing was applied within each trial, the factors triggering event and applied induction varied randomly across trials.
1.1.4. Observation task
In the observation task participants observed 4 different video clips of a male model each lasting 10 sec. In the experimental condition two videos were shown that suggested nose itching. In one condition the model constantly wrinkled his nose, whereas in the other condition a feather was constantly tickling his nose. Critically, in the latter condition the model did not show any reaction—that is, he did not wrinkle his nose. In the control condition we presented two videos unrelated to nose itching. In one video the model was constantly frowning and in the other condition the shadow of a feather was constantly moving around the model’s nose. It is important to note that any movement within the videos (i.e., wrinkling, frowning, tickling) was taking place in a constant way without any breaks. The whole task contained 10 experimental blocks. In every block, each video was repeated two times in random order. In total, each video was thus presented 20 times. Participants were instructed to watch the videos and to imagine what the model in the videos may feel. Before each video, a fixation cross was displayed for 1000 msec.

1.1.5. TMS stimulation and MEP recording
Electromyographical (EMG) activity was recorded with the ActiveTwo system (BioSemi). Sintered 11 × 17 mm active Ag–AgCl electrodes were placed over the right biceps brachii. This muscle contributes to flexing the forearm towards the head. The active electrode was placed on the right biceps and the reference electrode closer to the elbow joint (ventral surface). The two electrodes were about 3 cm apart. The ground electrode was placed on the right shoulder bone, near the neck. EMG signal was amplified (internal gain scaling), digitized at 2 kHz, high-pass filtered at 3 Hz, and stored on a PC for offline analysis. TMS pulses were delivered by a biphasic magnetic stimulator (Rapid2; Magstim). A 70 mm figure of eight coil was held tangentially to the skull with the handle pointing backward and laterally at a 45° angle to the sagittal plane. The optimal position of the coil was defined as the position in which TMS evoked the largest MEP in the contralateral biceps brachii and was determined for each subject separately. The stimulation intensity was determined based on the resting motor threshold (rMT) of the biceps muscle, which is defined as the intensity that evokes a MEP larger than 50 μV in 50% of the cases (Rossini et al., 1994). Participants were equipped with a swimming cap on which the optimal position for stimulation was marked so that the experimenter could easily track the correct position of the coil during the experiment. During the experiment, a mechanical arm held the TMS coil. Experimenters continuously monitored the coil position during the sessions and between each experimental block the position was adjusted when necessary. Stimulation intensity during the recording session was set to 110% of the rMT. Average intensity was 72.67% (range 58%–84%) of the maximal stimulator output.

Similar to previous studies that measured MEPs at the biceps (e.g., Zijdewind, Butler, Gandevia, & Taylor, 2006), participants’ right elbow was flexed with their forearms in a 45° angle throughout the whole experiment (see Fig. 1 for an illustration). The left arm was placed on the left leg. For each observed video, three single TMS pulses were delivered. The time of the three pulse deliveries were randomly varied with a minimal interval of 2 sec between each pulse. The first pulse was delivered not earlier than 2.5 sec after video onset and the last pulse was delivered not later than 9.5 sec after the video onset (See Fig. 2 for an illustration).

1.1.6. Data preparation
The EMG data were analyzed following a procedure similar to that adopted in previous studies (e.g., Bardi, Bundt, Notebaert, & Brass, 2015; Catmur, Walsh, & Heyes, 2007; Cavallo, Heyes, Becchio, Bird, & Catmur, 2014). That is, for each TMS trial, the peak-to-peak amplitude of the MEP was calculated. This was done by extracting the epochs starting 500 msec before and after the actual event (i.e., the TMS pulse) from the recorded data. We calculated the (root mean square RMS) value of the background EMG activity for the 500 msec period preceeding the TMS pulse. When this value was above 100 μV, data for the trial were rejected (2.5% of all trials). This was done to prevent contamination of the MEP measurements by significant fluctuations in background EMG.

The peak-to-peak amplitude of each trial was calculated for the 20–50 msec window following an event (i.e., the typical time range at which a MEP occurs). For each subject, the peak-to-peak amplitude of the MEPs was computed using MATLAB software. Individual trials with amplitude of at least 50 μV (71.0% of all trials) were averaged per participant and condition. Finally, we discarded all MEPs that were three SDs below or above participants’ average MEP (15.4%). To prepare data for analyses MEPs were z-transformed. Only participants with at least 10 valid trials per condition were included in the data analyses.

1.2. Results
To test the presence of a muscle-specific anticipated action effect, we entered the data into a 2 (itching: yes vs no) × 2 (displayed actions: facial actions vs feather actions) × 3 (timing: start vs middle vs end) repeated-measures analysis of variance (ANOVA). The ANOVA
yielded no significant main effects, $F(1, 13) < 2.07$, $p > .173$. The two-way interaction between itching and timing as well as the two-way interaction between displayed actions and timing was not significant either, $F(1, 13) < .31$, $p > .73$. However, more important for our hypothesis was the significant interaction between itching and displayed actions, $F(1, 13) = 5.58$, $p = .034$, $\eta^2 = .30$ (See Fig. 3) suggesting that the nose wrinkling video elicited the highest MEPs compared to all other conditions. A Planned contrast analysis confirms this interpretation. That is, participants’ z-transformed MEPs in the nose wrinkling condition ($M = .07, SD = .11$) were higher as compared to all other conditions ($M_{\text{frowning}} = -.04, SD_{\text{frowning}} = .13; M_{\text{feather tickling}} = -.03, SD_{\text{feather tickling}} = .14; M_{\text{feather shadow}} = -.01, SD_{\text{feather shadow}} = .09$), $F(1, 13) = 6.77$, $p = .022$, $\eta^2 = .34$. The three-way interaction between itching, displayed action, and timing was not significant, $F(1, 13) = .46$, $p = .64$, suggesting that irrespective of the time at which TMS pulses were delivered, the pattern remained the same.

Moreover, the significant interaction between itching and displayed action and the carried out contrast analysis suggest that anticipated action is not based on an attentional effect. If attention would play a crucial role in anticipated action, one would have expected similarly high MEPs in the nose wrinkling condition as compared to the conditions, in which attention was directed towards the model’s nose (i.e., feather tickling and feather shadow condition). Contrary to the attention hypothesis, however, the results show that only the nose wrinkling condition elicited a rise in participants’ MEPs. Besides speaking against the attentional hypothesis, this finding supports the inferred-desire hypothesis putted forward in previous research on anticipated action (cf., Genschow & Brass, 2015). In the feather condition the model did not engage in any nose wrinkling action although his nose was clearly tickled. Therefore, participants may have inferred that the model’s desire to scratch the nose was not very strong. This reduced inferred desire may then have inhibited the rise in the MEP.

### 2. Preregistered experiment

Despite first evidence for anticipated action on a neurophysiological level, there are a couple of open questions that need to be further addressed in order to strengthen our claims. First, although our results suggest that participants inferred a reduced desire to act in the model when his nose was tickled by a feather, we do not have any direct evidence for this claim. Thus, we aimed at testing participants’ inferred desire in a separate pretest.

Second, the setup of the pilot study leaves open whether anticipated action is muscle-specific. Alternatively, it could be that anticipating another person’s action activates the motor system in general and not only the muscles that are involved in the anticipated action. Naish et al. (2014) suggest that in order to ensure that effects are muscle-specific, EMG recordings should be taken during action execution as well as observation. Moreover, adding a condition in which the
anticipated action does not involve the measured muscle, but an unrelated muscle (e.g., mouth) would help demonstrating that the effects are indeed muscle-specific.

Third, the sample of the pilot study was with N = 14 rather low. Moreover, post-hoc power analysis indicates that our study was with a power of 1 – β = .51 not fully powered. In light of the current debate about the crisis of confidence in psychological research (e.g., Open Science Collaboration, 2015) it is crucial to replicate previous findings with highly powered studies in order to draw better conclusions (e.g., Lakens & Evers, 2014). Thus, in order to test muscle-specificity and to replicate the findings of the pilot study with a highly powered study we aimed at running a preregistered experiment. Based on the results of the pilot study we formulated the following hypothesis:

H1: As compared to the feather actions, facial actions lead to higher MEPs when observing the itching videos than when watching the non-itching videos.

2.1. Pretest

In order to test whether participants indeed infer a reduced desire to act in the model when his nose was tickled by a feather we conducted an online study in which 27 participants (23 female) with a mean age of 26.19 (SD = 7.46) ranging from 19 to 46 took part. Participants watched all videos that had been used in the pilot study. After watching each video they indicated their agreement with the following two statements “I think the man would like to scratch his nose”, “I think the man had a strong desire to scratch his nose” on 7-point scales (1 = is not at all the case; 2 = is entirely the case). In order to prepare data for analysis we averaged the two responses for each video. The results demonstrate that participants thought that the model has a stronger desire to scratch his nose when he was wrinkling his nose (M = 5.81, SD = 1.03) as compared to when a feather was tickling his nose (M = 2.96, SD = 1.90), t (26) = 6.96, p < .0001, dz = 1.34, as compared to when he was frowning (M = 1.98, SD = 1.29), t (26) = 13.70, p < .0001, dz = 2.64, and as compared to when the shadow of the feather was moving around his nose (M = 1.70, SD = 1.06), t (26) = 14.47, p < .0001, dz = 2.78. These results are in line with our conclusion of the pilot study indicating that participants perceive the model’s desire to scratch his nose as reduced when a feather is tickling his nose without him showing any reactions.

2.2. Method

2.2.1. Ethics statement

The ethics committee of Ghent University, Belgium, approved the procedure and method of the study.

2.2.2. Data availability and data collection

The data file of the study as well as the laboratory log indicating that the data collection took place after in principle acceptance of the manuscript is available from the Open Science Framework database. The URL necessary to access data and laboratory plan is: https://osf.io/5pkys/?view_only=a5defabba53d4f56b07d1caca1dab699.

2.2.3. Participants and design

In order to estimate the sample size, we conducted a power analysis based on the crucial interaction effect between itching and displayed actions using G’Power 3.1 ( Faul, Erdfelder, Lang, & Buchner, 2007). First, we computed for both displayed actions (i.e., facial actions and feather actions) the difference between the itching and the non-itching condition in order to produce two compound scores. Note, the difference between these two scores reflects the interaction effect found in the pilot study. The two compound scores were then subjected to a power analysis for paired samples t-test. Based on an effect size of dz = .63, a power of 1 – β = .95, an alpha error probability of α = .05 (two-tailed) G’Power estimated an optimal sample size of N = 35 participants.

Similar as in the pilot study, participants were recruited via the participant pool of the Department of Experimental Psychology at Ghent University for a payment of 25 €. Due to modified regulations of the ethical guidelines, in contrast to the pilot study, we were allowed to test female participants as well. In order to compensate for possible dropouts due to the preregistered selection criteria, we invited 45 participants into our TMS lab. We could not reliably identify the motor hot spot for two participants. Moreover, we excluded eight participants due to too few remaining trials in the experimental conditions (i.e., fewer than 10 trials in the experimental trials) after cleaning the data. The final sample, thus, contained 35 participants (25 female) with an age ranging from 18 to 29 (M = 22.09; SD = 3.08). The design of the experiment consisted of a 2 (itching: yes vs no) × 2 (displayed actions: facial actions vs feather actions) within-subject design.

2.2.4. Procedure

The procedure was similar as the one of the pilot study. However, we added some further control conditions in order to perform reality checks and to test whether anticipated action is muscle-specific. Table 1 gives an overview of all assessed conditions.

First, we included neutral trials before and after each block. That is, we presented before the experimental blocks two times a 10 sec lasting blank screen with a fixation cross on it. The total amount of those neutral trials was, thus, 22. For each screen, participants were instructed to just watch the screen while we delivered three single TMS pulses. Similar as the experimental videos, the time of the three pulse deliveries randomly varied with a minimal interval of 2 sec between each pulse. The first pulse was delivered not earlier than 2.5 sec after screen onset and the last pulse was delivered not later than 9.5 sec after the screen onset.

Second, intermixed with these neutral trials, we presented participants two times a video in which they anticipated a movement unrelated to nose scratching in order to be able to test whether anticipated action is muscle-specific. More specifically, we presented a video in which a spoon was directed towards the model’s mouth in order to trigger an anticipated mouth opening movement. The video had the same length as the other videos (i.e., 10 sec) and TMS pulses were delivered at the same three points in time as in the other videos. Over the whole experiment the video was presented 20 times.
Third, in the end of the experiment, we included a condition in which participants actually performed the anticipated action. In 20 trials, participants were instructed to move their right arm towards their nose. At the start of each trial a fixation cross appeared. Afterwards, a cue appeared (i.e., a blue “X”). At this point participants were instructed to prepare the movement (i.e., to slowly start moving the arm upwards). After a variable time interval (500–1500 msec), a green colored “X” (i.e., imperative stimulus) was presented for 3000 msec. Participants were instructed to respond to the imperative stimulus with lifting their arm towards the nose as fast as possible. As soon as they touched their nose, they returned to the resting position. Between the cue and the imperative stimulus (i.e., during response preparation), we delivered one TMS pulse and measured EMG activity in participants’ biceps muscle. We chose to establish this condition in the end of the experiment, because executing the target action before or within the experimental blocks might have primed the anticipated action interfering with the hypothesized anticipated action effect.

In line with the pilot study, stimulation intensity of the TMS pulses was set to 110% of the rMT. Average intensity was 70.06% (range 44%–90%) of the maximal stimulator output.

2.2.5. Data preparation
Data preparation was carried out similar to the pilot study. That is, with respect to excluding participants from analysis, we applied the same filter that we used in the pilot study. Specifically, we did not include participants for whom we could not reliably identify the motor hot spot (N = 2) and participants for who in at least one experimental condition less than 10 trials with valid MEPs were detected (N = 8). With respect to data exclusion, we excluded trials in which the (root mean square RMS) value of the background EMG activity was higher than 100 μV (2.5%), in which the MEP was smaller than 50 μV (.7%), and in which an individual’s MEP was three SD’s below or above his or her average MEP (1.3%).

Possible variations in corticospinal excitability related to TMS per se was controlled by subtracting the mean MEP in the neutral conditions from each mean of the experimental conditions (for a similar procedure see Sartori, Bucchioni, & Castiello, 2013).

3. Results

3.1. Preregistered analyses

3.1.1. Reality check
In a first analysis, we applied a reality check of our measure. That is, we tested whether during our experiment MEPs were actually measured from the right target muscle (muscle that would be used to perform the anticipated action, i.e., arm movement in order to scratch the nose). In doing so, we compared participants’ mean MEPs in the condition in which they prepared the target movement with the mean MEPs in the neutral condition (in which they did not have to execute any action) with a t-test for dependent samples. The results demonstrate that participants’ z-transformed MEPs in the condition in which they prepared the target movement (M = 1.27, SD = .86) were higher as compared to the neutral condition (M = −.05, SD = .15), t(34) = 8.64, p < .0001, dz = 1.45. This indicates that we indeed measured from the muscle that would be used to perform the anticipated action.

3.1.2. Anticipated action
An overview of all mean values is given in Table 2. As no effect of timing was found in the pilot study, we did not have a specific hypothesis for this factor and left it therefore out of the preregistered analyses. To test for the presence of an anticipated action effect, we entered the baseline corrected data into a 2 (itching: yes vs no) × 2 (displayed actions: facial actions vs feather actions) repeated-measures analysis of variance (ANOVA). Neither the main effect for itching, nor the two-way interaction between itching and displayed actions was significant, F(1, 34) < .38, p > .54. However, the ANOVA yielded a significant main effect for displayed action, F(1, 34) = 7.71, p = .009, ηp² = .19, indicating stronger MEPs (baseline corrected and z-transformed) when observing facial actions (M = .03, SD = .21) than when observing feather actions (M = −.02, SD = .20).

In line with the pilot study we ran a contrast analysis in order to test whether the nose wrinkling condition led to higher MEPs (baseline corrected and z-transformed) as compared to all other video conditions. Although the effect went descriptively in the same direction as in the pilot study, the test did not reach significance, F(1, 34) = 2.32, p = .137.

Muscle specificity. In a third analysis we tested whether there was a difference between the two anticipated action conditions. That is, we compared the nose wrinkling condition with the condition in which a spoon was directed to the model’s mouth with a t-test for dependent samples. The t-test did not reveal a significant difference for MEPs (baseline corrected and z-transformed) between the nose wrinkling condition (M = .03, SD = .24) and the spoon condition (M = .003, SD = .18), t(34) = .67, p = .51.

3.2. Post hoc analyses
The preregistered hypothesis tests did not support our predictions. Therefore, in order to further the understanding of our results, we ran additional post hoc analyses in an exploratory fashion. First, we applied Bayesian statistics using JASP, an open source statistical package (Love et al., 2015), in order to
assert the strength of evidence in favor of H0 of the two central hypotheses tests—that is, the interaction between itching and displayed actions as well as the contrast between the nose wrinkling condition and all other video conditions. We report $BF_{01}$, which gives the ratio with which the null hypothesis is favored over the alternative hypothesis (i.e., a larger $BF_{01}$ argues in favor of the null hypothesis; see Dienes, 2014, for an overview). We set the Cauchy prior width based on the standardized effect sizes of our pilot data ($r = .30$ for the interaction and $r = .33$ for the contrast). This approach yielded $BF_{01} = 1.71$ for the interaction and $BF_{01} = .68$ for the contrast. A Bayes factor of 1 is conventionally considered as no evidence and a Bayes factor between 1 and 3 is considered as anecdotal evidence (Andraszewicz et al., 2015; Jeffreys, 1961). Therefore, these analyses do not provide strong evidence for the null hypothesis. However, the analyses were based on standardized effect sizes. As the SD are larger in the preregistered experiment than in the pilot study, the standardized effect sizes of the preregistered study are reduced. Thus, one may argue that analyzing the data based on the raw, rather than standardized effect sizes, is more informative. Therefore, we ran a second Bayesian analysis in which we took advantage of the online tool provided by Dienes et al (cf. Dienes, Coulton, & Heather, 2018). This approach yielded $BF = .44$ for the interaction and $BF = 1.27$ for the contrast. A $BF < 1/3$ indicates evidence for the null hypothesis and $BF > 3$ indicates evidence for the alternative hypothesis rather the null hypothesis. And a $BF$ between 1/3 and 3 indicates that the data are insensitive. Thus, the Bayesian analyses based on the raw effect sizes indicate that the preregistered experiment does neither support the null nor the alternative hypothesis.

Second, we tested whether the nose wrinkling condition differed from any of the other experimental conditions (i.e., frowning, feather tickling, feather shadow) with multiple $t$-tests. Fig. 4 illustrates the results. The results indicate that participants’ MEPs (baseline corrected and z-transformed) were significantly higher when watching the nose wrinkling videos ($M = .03, SD = .24$) than when watching the videos in which the feather tickled the model’s nose, ($M = -.03, SD = .22$), $t(34) = 2.22, p = .033, dz = .37$. All other comparisons between the nose wrinkling condition and the other conditions did not reach significance, $t(34) < 1.54, p > .13$.

Third, we investigated whether the timing of the TMS stimulation might have influenced our results. That is, we ran a 2 (itching: yes vs no) $\times$ 2 (displayed actions: facial actions vs feather actions) $\times$ 3 (timing: start vs middle vs end) repeated-measures analysis of variance (ANOVA). Neither the two-way interaction between timing and itching, nor the two-way interaction between timing and displayed actions, nor the three-way interaction between timing, itching and displayed action was significant, $F < .69, p > .51$.

Fourth, we tested whether participants’ gender might have influenced our results. However, the interaction between itching and observed action was neither significant for female participants, $F(1, 24) = 1.61, p = .22$, nor for male participants, $F(1, 9) = 1.81, p = .21$. Likewise, a contrast analysis testing whether the nose wrinkling condition led to higher MEPs (baseline corrected and z-transformed) as compared to all other video conditions was neither significant for female participants, $F(1, 24) = 2.79, p = .11$, nor for male participants, $F(1, 9) = .002, p = .97$. Moreover, when running a 2 (itching: yes vs no) $\times$ 2 (displayed actions: facial actions vs feather actions) $\times$ 2 (gender: male vs female) ANOVA, gender did not interact with any of the other factors, $F(1, 34) < 1.49, p > .12$.

Fifth, we tested whether the TMS intensity determined at the individual level influenced our results by running a 2 (itching: yes vs no) $\times$ 2 (displayed actions: facial actions vs feather actions) ANOVA with stimulation intensity as covariate. However, when controlling for TMS intensity, neither the interaction, nor the contrast between the nose wrinkling condition and all other conditions was significant, $F(1, 33) < 2.69, p > .11$. Likewise, the interaction between itching, displayed action and intensity was not significant either, $F(1, 33) = 2.46, p = .13$.

Sixth, we tested whether we can replicate the results obtained in the pilot study by running the analyses on the MEPs without correcting the data for the baseline. However, neither the interaction between itching and displayed action nor the contrast analysis testing whether the nose wrinkling condition led to higher MEPs (baseline corrected and z-transformed) as compared to all other video conditions was significant, $F(1, 34) < 2.32, p > .13$.

### 3.3. Post hoc analyses across both studies

Although the preregistered study does not replicate the findings of the pilot study, the data reveal a similar pattern at the descriptive level. This raises the question whether the predicted effect can be detected meta-analytically. Thus, we ran

---

**Table 2 – Overview over mean MEPs (baseline corrected and z-transformed) in the preregistered experiment.**

<table>
<thead>
<tr>
<th></th>
<th>Nose wrinkling</th>
<th>Feather tickling</th>
<th>Frowning</th>
<th>Shadow of feather</th>
<th>Spoon to mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>.032</td>
<td>-.028</td>
<td>.027</td>
<td>-.009</td>
<td>.003</td>
</tr>
</tbody>
</table>

---

**Fig. 4 – Z-transformed and baseline corrected MEPs in the preregistered study. Error bars represent standard errors of the means.**
post hoc mini meta-analyses across the pilot study and the preregistered study with the tool introduced by Goh, Hall, and Rosenthal (2016). We first computed for each study the effect size of the contrast between the nose wrinkling condition and all other conditions. In addition, we computed for each study the effect size of the interaction between itching and displayed actions. Afterwards, we transformed each effect size into Fisher’s Z and then ran the mini meta-analyses weighted by the number of subjects participated in each experiment (cf. Goh et al., 2016). The analyses indicate that participants across both studies had stronger MEPs when observing the nose wrinkling video as compared to all other videos $Z = 1.35$, $p < .001$. Likewise, the interaction between itching and displayed action was significant, $Z = .93$, $p < .001$.

4. General discussion

Past research has shown that individuals engage in anticipated action. That is, individuals anticipate what another person’s next action would be and then engage in that action themselves (Genschow & Brass, 2015; Genschow et al., 2018). In the present research, we assessed on a trial-by-trial basis the degree of anticipated action on a neurophysiological level. To this end we applied in a pilot study and in a preregistered study motor TMS. The results of the pilot study give support for anticipated action on a neurophysiological level: Observing nose wrinkling in someone else increased participants’ MEPs in the biceps as compared to observing control events. While the pilot study finds support for anticipated action on a neurophysiological level, the preregistered study could not replicate these findings although the pattern of results was in the same direction.

In order to test whether the pattern of the preregistered study speak in favor of the null hypothesis we ran additional Bayesian statistics. These analyses neither support the null hypothesis nor the alternative hypothesis. One may find it counterintuitive that the preregistered experiment revealed inconclusive Bayes factors despite being powered to .95. However, due to the possibility of the winner’s curse leading to overestimation of the true effect size, and the more general fact that high power alone is insufficient to guarantee robust evidence, such a result is not surprising. Indeed, Dienes and Mclatchie (2018) show that despite high power, non-significant results may not entail evidence for the null hypothesis. This can be the case, because power is not a measure of evidence and does not necessarily guarantee evidence.

Thus, to test the presence of anticipated action across both studies we carried out a post hoc mini meta-analysis and found support for the general idea of anticipated action on a neurophysiological level. However, this result needs to be taken with caution and shortcomings need to be taken into consideration when interpreting the meta-analysis. That is, the meta-analytical effect is strongly influenced by the effect size of the pilot study, which was rather underpowered. Moreover, even if we take the post hoc meta-analysis as evidence for anticipated action on a neurophysiological level, it is not clear whether this effect is muscle specific. That is, it could be that anticipating another person’s action activates the motor system in general and not only the muscles that are involved in the anticipated action. In the preregistered study, we tested for muscle specificity. However, as we do not fully replicate the overall pattern, the interpretation of this test is rather difficult. Finally, the effect size across both studies can be considered as rather small.

Yet another interesting finding of the preregistered study is the main effect of observed action: Observing facial actions, such as nose wrinkling and frowning increased participants’ MEPs in comparison to observing no facial actions (i.e., feather movements and shadow of feather movements on the face). Past research has shown that corticospinal activity during action observation is effector-specific (e.g., Romani, Cesari, Urgesi, Facchini, & Aglioti, 2005; Strafella & Paus, 2000; Urgesi et al., 2006a; Urgesi, Moro, Candidi, & Aglioti, 2006b). However, our finding suggests that corticospinal activity may translate to other body parts as well. Interestingly, van Schaik, Sacheli, Bekkering, Toni, and Aglioti (2017) recently found similar results suggesting that the observation of specific actions leads to a general activation of the motor system. In order to explain such general activation patterns, Naish et al. (2014) propose that action observation elicits an early, non-specific corticospinal activation (at around 90 msec from action onset), followed by a later activation of specific muscles involved in the observed action (from around 200 msec). In our preregistered experiment we delivered TMS pulses not earlier than 2.5 sec after the observation of an action initiation. Nevertheless, we still find a general activation of the motor system. Our results, thus, extend Naish and colleagues’ model by suggesting that even at a late state general activation of the motor system can be observed. However, it has to be noted that our finding was not predicted and therefore calls for a more thorough investigation. Future research may aim at investigating whether our effects can be replicated and at determining the time course of this general motor activation.

Taken together, our preregistered study found a main effect for observed action, but did not replicate the result of the pilot study. Although a post hoc mini-meta-analysis supports the hypothesis of anticipated action, the results need to be taken cautiously and critically discussed. That is, our results raise two major questions that call for a fine graded discussion.

4.1. Why did the preregistered study not replicate the pilot study?

First, we have to consider the possibility that the null hypothesis is in fact true. That is, it could be that anticipated action on a neurophysiological level does either not exist or that it cannot be detected within the methodological setup we applied in the preregistered experiment. As part of our post hoc analyses we applied Bayesian statistics in order to test for $H_0$. However, the results provide rather anecdotal evidence, which might be due to lack of statistical power. Thus, for future research we recommend assessing a larger sample size in order to draw stronger conclusions.

Second, a difference between the pilot study and the preregistered experiment that could account for the conflicting results may lay in the assessed participants: While we tested male participants only in the pilot study, we tested male as well as female participants in the preregistered study. Given that the model in our studies was male and in-group members...
are more strongly represented in the motor system than outgroup members (Bourgeois & Hess, 2008; Cheng & Chartrand, 2003; Genschow & Schindler, 2016; Lakin, Chartrand, & Arkin, 2008; Mondillon, Niedenthal, Gil, & Droit-Volet, 2007; Yabar, Johnston, Miles, & Peace, 2006), one could argue that male, but not female participants show an anticipated action effect in the preregistered study. However, when testing the influence of gender in an exploratory post hoc analysis, we could not detect any difference between male and female participants in the preregistered study. Therefore, we regard it as rather unlikely that gender differences explain the different effects.

Third, reasons may lay in the methodological setup of the preregistered study. Although the method of the preregistered study was very similar to the one of the pilot study, there were some differences that may account for the failed replication. First, we had to make new video materials and we used a different model. Second, the preregistered study contained a different trial structure. That is, between experimental blocks, participants observed videos in which a spoon approached the model’s mouth triggering anticipated mouth opening actions. It could be that such an action was still activated during the experimental blocks interfering with anticipated nose touching action. Finally, one could argue that the baseline correction in the preregistered experiment masked potential effects. However, when we did not correct the MEPs for the baseline, post hoc analyses yielded no significant effects either.

Fourth, it might be that we underestimated the statistical power needed to find an effect of anticipated action on a neurophysiological level. Based on an effect size of dz = .63 in the pilot study we estimated the number of participants needed for the preregistered study. However, studies with small samples often overestimate the effect size of a found effect (e.g., Button et al., 2013). As the sample size of the pilot study was with N = 14 rather small, we might have overestimated the effect size of the true effect. Thus, the 35 participants within the preregistered study were not enough to detect the true effect.

4.2. Why is the effect rather small?

If existent, the effect size of anticipated action in the present paradigm is rather small. There might be different reasons why this is the case. First, a reason may lay in the applied trial structure. While we have used in our behavioral studies (Genschow & Brass, 2015; Genschow et al., 2018) a blocked design in which only one triggering event (e.g., nose wrinkling) per block was observed, all experimental conditions in the present research were presented intermixed within each block. It could be that in the present research observing the triggering event (i.e., nose wrinkling) primed the target action (i.e., nose touching) and prolonged for the following trials. That is, the effect induced by one condition might have induced carryover effects affecting the other conditions reducing the difference between conditions. As a result, the anticipated action effect decreased. In contrast, in the behavioral studies (Genschow & Brass, 2015; Genschow et al., 2018) observing the triggering event repeatedly after each other could have increased the anticipated action effect. If this interpretation indeed accounts, it might be that anticipated action is not a specifically transient phenomenon, but rather an effect that builds up through repeated observation of triggering events. Future research may aim at further testing this interpretation more rigorously.

Second, the setting of the present studies was rather unnatural as compared to the behavioral studies and may have, thus, reduced our effects. That is, we instructed participants to keep their head still and to hold their elbow flexed with their forearms in a 45° angle throughout the whole experiment. It is reasonable to assume that directing participants’ attention to hold the position counteracted the predicted effects. Therefore, for future research a more natural position is advisable.

Third, it might well be that the effect is rather small, because it does not exist. In order to shed light onto this possibility we applied post hoc Bayesian tests in order to test for H0. The results provide anecdotal evidence for H0 indicating the possibility of an effect size of zero. However, given the rather weak support for H0, for future research, we recommend the assessment of larger samples or sequential hypothesis testing (cf. Schönbrodt, Wagenmakers, Zehetleitner, & Perugini, 2017) in order to draw precise conclusions.

5. Summary

Past research has shown that individuals engage in anticipated action (Genschow & Brass, 2015; Genschow et al., 2018). In the present research, we applied motor TMS and tested anticipated action on a neurophysiological level in a pilot study and in a preregistered study. While the pilot study supports our hypothesis, the preregistered study cannot replicate these findings although the pattern of the results was in the same direction. Although a post hoc mini meta-analysis across both studies supports the idea of anticipated action on a neurophysiological level, these results need to be taken with caution. For future motor TMS research on similar topics we recommend higher powered studies, blocked designs and the applications of natural arm positions.

Acknowledgment

This work was supported by a grant from the Swiss National Science Foundation (grant number PZ00P1_168007).

References


