The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/134344

Please be advised that this information was generated on 2019-09-09 and may be subject to change.
Anticipatory Planning Reveals Segmentation of Cortical Motor Output During Action Observation

Loes Janssen¹, 2, 3, Bert Steenbergen² and Richard G. Carson⁴, 5*

¹Donners Institute for Brain, Cognition and Behaviour, ²Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, 6500 HE, The Netherlands, ³Department of Clinical Epidemiology, VieCuri Medical Centre, Venlo, 5912 BL, The Netherlands, ⁴Trinity College Dublin, Trinity College Institute of Neuroscience and School of Psychology, Dublin 2, Ireland and ⁵School of Psychology, Queen’s University Belfast, Belfast, BT7 1NN, Northern Ireland

Address correspondence to Prof. Richard G. Carson, Trinity College Institute of Neuroscience and School of Psychology, Lloyd Building, Trinity College Dublin, Dublin 2, Ireland. Email: richard.carson@tcd.ie

It has been argued that the variation in brain activity that occurs when observing another person reflects a representation of actions that is indivisible, and which plays out in full once the intent of the actor can be discerned. We used transcranial magnetic stimulation to probe the excitability of corticospinal projections to 2 intrinsic hand muscles while motions to reach and grasp an object were observed. A symbolic cue either faithfully indicated the required final orientation of the object and thus the nature of the grasp that was required, or was in conflict with the movement subsequently displayed. When the cue was veridical, modulation of excitability was in accordance with the functional role of the muscles in the action observed. If however the cue had indicated that the alternative grasp would be required, modulation of output to first dorsal interosseus was consistent with the action specified, rather than the action observed—until the terminal phase of the motion sequence during which the object was seen lifted. Modulation of corticospinal output during observation is thus segmented—it progresses initially in accordance with the action anticipated, and if discrepancies are revealed by visual input, coincides thereafter with that of the action seen.

Keywords: human, motor cortex, reaching–grasping movement, transcranial magnetic stimulation, visuomotor

Introduction

It is well established that observation of action is associated with changes in the excitability of corticospinal projections to muscles that would be engaged in replication of the movements being observed (Fadiga et al. 1995, 2005). These variations are functionally and temporally specific—the amplitude of responses to transcranial magnetic stimulation (TMS) recorded from the first dorsal interosseus (FDI) muscle of a passive observer is modulated in phase with the grip aperture exhibited by an actor seen to approach and lift an object (Gangitano et al. 2001). Furthermore, they may be anticipatory in nature, thus providing for predictive coding of the behavior of others (Kilner et al. 2004; Borroni et al. 2005; Aglioti et al. 2008; Urgesi et al. 2010; Alaerts et al. 2012; Sartori et al. 2012).

The phenomenon has been attributed (Rizzolatti and Fadiga 1998) to a mirror system akin to that inferred on the basis of single-cell recordings from frontal and parietal areas of nonhuman primates (Rizzolatti and Craighero 2004). There is however a notable point of divergence. Animal studies suggest that successive elements of the observed movement sequence are encoded separately (Fadiga et al. 2000; Fadiga and Craighero 2003; Craighero et al. 2007). For example, in describing the behavior of neurons in the ventral premotor cortex (area F5) that were active during the observation of specific actions, Gallese et al. (1996) noted that these could be classified on the basis of their level of discharge during the various phases of the compound movement that was viewed. It has been further demonstrated that subsets of neurons in this brain region respond when the final phase of an evolving action is implied, whereas others are responsive only when it is portrayed explicitly (Umiltà et al. 2001).

In contrast, on the basis of evidence derived from noninvasive cortical stimulation, it has been argued that human observers represent action sequences as indivisible ensembles (Gangitano et al. 2004; Fadiga et al. 2005; Hauk et al. 2008; Cattaneo and Rizzolatti 2009). Gangitano et al. (2004) reported that the variations in corticospinal excitability exhibited typically during observation of a reach to grasp movement, are not obtained when the action does not proceed as expected (delayed maximum aperture of the fingers or premature closure of the hand). The authors concluded that “resonant motor plan(s),” formed initially upon viewing the start of a movement sequence, are neither modulated nor substituted when the observed action does not proceed accordingly. It has been inferred that instead they are suppressed if the visual characteristics of the observed movement fail to match those that are anticipated (Gangitano et al. 2004).

In exploring observation-to-execution mappings expressed via the human motor system, it has been noted that when the intrinsic properties (e.g., the ostensible level to which a vessel is filled) of an object, and the kinematic characteristics of its manipulation (i.e., when seen lifted) are in conflict, corticospinal output is modulated in accordance with kinematic rather than object-related cues (Alaerts et al. 2010). At first glance, this might suggest that the engagement of the cortical motor network strictly mirrors only the motor output characteristics of the observed actions (Cavallo et al. 2012). There is however evidence that these processes are also subject to cognitive penetration. When semantic information (e.g., a “light” or “heavy” label) conflicts with the physical characteristics of the object of an observed action, modulation of corticospinal output is reduced, but not reversed (Senot et al. 2011). It has thus been proposed that—as when the observed action fails to match that which is anticipated (Gangitano et al. 2004), when semantic and movement-related cues are in conflict, corticospinal output is gated rather than regulated in accordance with symbolic information (Senot et al. 2011).

We reasoned that prior symbolic cues—if action oriented, could in principle be used to establish whether human
observers represent action sequences as indivisible ensembles or if successive elements of a movement sequence are encoded separately. In contrast to approaches in which variants of a single action (e.g., a precision grasp) have been used (e.g., Gangitano et al. 2004; Senot et al. 2011), the potential segmentation of corticospinal excitability was considered in the context of alternative actions that in their implementation required different patterns of muscle engagement. This was achieved by utilizing the phenomenon of anticipatory movement planning—when preparing to grasp an object, account is taken not only of its current configuration, but also of the demands associated with the goal of the prospective action (Johnson-Frey et al. 2004). Critically, this planning occurs not only when movements are executed, but also when observed (e.g., Flanagan and Johansson 2003).

Participants watched videos of actors performing object manipulations using either a precision or whole-hand grasp. Prior to movement initiation, a symbolic cue indicated the final goal of the action. On this basis, the required grasp could be inferred. In some instances (incongruent condition), the cue was in conflict with the action that was subsequently displayed. During each trial, motor-evoked potentials (MEPs) were evoked in 2 hand muscles differentially engaged in precision and whole-hand grasps, at 1 of 6 intervals defined relative to the unfolding action sequence.

Critically, mapping a specific cue to a distinct action provided the means of establishing the extent to which (in the incongruent cue condition) the modulation of corticospinal excitability unfolded in accordance with one alternative (i.e., the cued action) or the other alternative (i.e., the observed action). The specific power of this design lies in the facility to not only detect a deviation from the modulation associated with the action that was cued (as in single action designs), but also to establish (through the use of inferential tests of equivalence) whether there was reversion to the pattern of modulation associated with the alternative action (i.e., that being observed). Our findings revealed that the variation of corticospinal output that occurs during passive observation progresses initially in accordance with the action that is anticipated, and if discrepancies are revealed by visual input, coincides thereafter with that of the action seen. The elements of observed movement sequences are thus encoded separately.

Materials and Methods

Participants
Thirteen healthy right-handed adults (6 males) aged 18–38 years gave written informed consent to the procedures, which were approved by the local ethical committee and conducted in accordance with the Declaration of Helsinki.

Materials
Twelve videos displayed an actor reaching for and grasping an object, which was then inserted into 1 of 2 holes in a wooden block (Fig. 1). The object was a black bar (diameter 14 mm, length 70 mm) attached at one end to the center of a white disk (diameter 108 mm, thickness 12 mm). It could be grasped easily with either a precision grasp—on the black bar, or whole-hand grasp—on the white disc. The block had a small hole in which the black bar fitted precisely, and a large hole in which the white disc fitted precisely. Comfortable insertion of the object into the small hole required a whole-hand grasp of the white disc. Insertion of the object into the large hole required a precision grasp of the black bar. An arrow between the 2 holes could point either to the small hole (whole-hand grasp required), or to the large hole (indicating a precision grasp), or between the holes (no indication of the required action).

Videos commenced with the object in a holder, the block positioned to the left, and the actor’s prone (right) hand to the right (Fig. 1). The block was oriented with the small hole on the left and the large hole on the right, or vice versa. After 1000 ms, a yellow ring (diameter 100 mm) around the ring appeared for 2000 ms to highlight the goal cue. The hand started moving 3000 ms following the disappearance of the ring and subsequently reached and grasped the object either by holding the black bar (precision grasp) or the white disc (whole-hand grasp).

In the majority of trials (67%), the grasp was in accordance with the goal cue (congruent). In some (22%) trials, however, the grasp conflicted with the cue (incongruent). We also included "no-cue" trials (11%) in which the arrow provided no indication of the required action—the actor used either a precision or a whole-hand grasp to insert the object in the corresponding hole.

The 12 different videos comprised combinations of 3 cues (congruent, incongruent, no-cue), 2 grasps (precision, whole-hand), and 2 block orientations (small hole left, small hole right). All were of 10 480 ms duration, followed by the presentation of a black screen for 2000 ms.

Electromyography and TMS
The electromyographic (EMG) activity of the FDI and abductor digitii minimi (ADM) of the right hand was recorded using bipolar surface electrodes. EMG signals were amplified, filtered (30 Hz–1 kHz), and digitized (16 bit) at a sampling rate of 5000 Hz.

Figure 1. The sequence of events is illustrated for a composite trial. The bar at the top of the figure represents the video timeline. The individual frames (shown for a no-cue condition) correspond to the times (phases) at which TMS was delivered, that is, in separate trials. Also depicted schematically are the instructional cues in which the arrow pointed to the small hole—indicating the requirement for a whole-hand grasp, and in which the arrow pointed to the large hole—indicating the requirement for a precision grasp.
The FDI and ADM muscles perform functional roles in relation to prehensile movements that are differentiated by the nature of the grasp that must be formed. As the FDI acts as a principal agonist (abducts and flexes the index finger) in precision grasps requiring index-thumb opposition, it is engaged to a greater degree than in whole-hand grasps—unless a clockwise rotation must be imparted upon the object (Long 1981). The ADM assumes a significant functional role in grasping large objects (such as the disc used in the present study) with outstretched fingers, and is engaged accordingly. While it is not typically activated to a significant degree in a prototypical pinch opposition between the thumb and index finger, there was an additional component to the precision grasp required in the present study. When the object was lifted with a grasp of the black bar, it was necessary to also extend (and fan) digits 4 and 5—including the little finger, in order to avoid contact with the support stand (Fig. 2). As there is a tendon attachment from ADM to the extensor expansion, the muscle contributes to proximal and distal interphalangeal extension, and is duly engaged (Marzke et al. 1998).

TMS was delivered to the left primary motor cortex (M1) by a Magstim 200 stimulator (Magstim, Carmarthenshire), using a (30-mm internal loop diameter; 85-mm external loop diameter) figure of eight coil. The optimal position to obtain a MEP in the right ADM was located and reproduced by the following means. A spot was first marked on the scalp 6 cm lateral and 2 cm anterior to the vertex (i.e., left M1 hand area), and the coil oriented to induce posterior to anterior current flow in relation to the presumed orientation of the motor strip. The coil was then moved systematically around this initial location in both anterior–posterior and medial–lateral directions, until the largest response was obtained. A line defining the intersection of the 2 loops of the stimulating coil was marked directly on the scalp. An additional marking aligned with the long axis of the coil handle—which bisected the first line, was also made. During the experiment, these markings were used to maintain constant the position and orientation of the coil. The lowest stimulation intensity at which MEPs with peak-to-peak amplitude of at least 50 µV could be recorded in 3 of 5 trials was taken as resting motor threshold (RMT). Stimulation at this intensity also evoked a MEP in right FDI. The level of stimulation used subsequently was 120% RMT. At the beginning and at the end of the session, 3 sets of 10 control MEPs were obtained.

The rationale in using a stimulus intensity of 120% RMT was to generate MEPs (i.e., in the control condition) with amplitudes in the middle of the range that can be obtained for the muscles in question. On the basis of our own experience (e.g., Carroll et al. 2001) and that of others (e.g., Devanne et al. 1997), we were confident that by using this (i.e., fixed) stimulation intensity in the context of action observation, there was scope for the amplitude of the MEPs to increase or decrease. It was thus unlikely that either floor or ceiling effects would be present.

Procedure
The participants were first familiarized with the object manipulation task. They were instructed to object the right hand and insert it into the hole in the block—as cued by the arrow (using a precision or whole-hand grasp as appropriate). The participant was then asked to close their eyes. The experimenter changed the orientation of the block. Following a verbal go-signal, the participant opened their eyes and performed the task again. This was repeated until the participant completed 20 successive trials using the correct grasp. In requiring that the participants perform the task themselves, we were guided by recent findings (e.g., Olsson and Nyberg 2011) that it is necessary to have experienced a motor act in order that corresponding brain regions are recruited when the action is seen being performed. It has also been reported that cueing in the context of action observation is effective only when there has been preceding experience linking a given cue to a specific action performed (Petroni et al. 2010).

The participants were instructed to attentively observe the videos that were displayed on a computer screen positioned 1 m ahead. In a randomly defined subset of trials (18 of 288), they were asked to imitate the movement they had just observed. The purpose of the imitation trials was to ensure that the participants remained engaged. As they were unable to anticipate whether on any given trial they would be required to imitate the movement they had just observed, we reasoned that this arrangement would ensure that the level of attention allocated on every observation trial would be comparable. As there were no errors in imitation, we are confident that the inclusion of this requirement achieved the desired outcome. In most respects, the procedure was similar to that employed during initial practice. There was one key exception. The interleaved test movements were executed using the left hand rather than the right hand. It was anticipated that any action associated changes in the excitability of corticospinal projections to the muscles of the hand performing the interleaved movement, would have minimal impact upon the excitability of the projections to the right hand that were assessed during the observation trials. Short breaks were scheduled after each of the 18 imitation trials. Thus on average, breaks were taken after sets of 16 observation trials.

During each observation trial, TMS was delivered in 1 of 6 phases (delays): termination of the goal cue (3000 ms); 100 ms before movement onset (4900 ms); just after movement onset but grasp as yet unrevealed (5280 ms); maximal aperture of the precision grasp (5680 ms); maximal aperture of the whole-hand grasp (6080 ms); during lifting of the object prior to transport (7280 ms). In order to ensure precision of timing for stimulus delivery, a discrete pulse was encoded on an audio channel of the video file. The rising edge of this pulse was used to trigger the TMS. To limit the total duration, and maintain a high congruent/incongruent trial ratio—thus ensuring the integrity of any associated internal models (Schiffer et al. 2012), we administered TMS in all 6 phases during congruent trials, in 4 phases during incongruent trials, and in 2 phases during no-cue trials. As the videos for the congruent and incongruent trials were identical during the movement...
preparation phases (Fig. 1), we elected not to obtain MEPs at 3000 and 4900 ms in the latter condition. As the no-cue trials acted as a control, principally with respect to the preparation phase, we administered TMS only at 3000 (termination of the goal cue) and 4900 ms (100 ms before movement onset). The 24 congruent combinations (4 videos × 6 delays) were repeated 8 times (total 192), the 16 incongruent combinations (4 videos × 4 delays) 4 times (total 64), and the 8 no-cue combinations (4 videos × 2 delays) 4 times (total 32). The order was random.

Data Reduction and Analyses
The amplitude of MEPs recorded in right FDI and ADM, and the background EMG (root mean squared) 100 ms prior to each stimulus, was used to derive the sums of squares and means squares relevant to each comparison. Corresponding effect sizes (f) were also obtained (Cohen 1969). In addition, an inferential test of equivalence (Wellek and Michalis 2010), was calculated for each comparison. As the equivalence tests are random.

In a supplementary experiment, we established the patterns of activity in FDI and ADM during execution of the 2 variants of the task. For this purpose, we engaged a further 5 participants and asked them to each perform 10 instances of the full grasp action and 10 instances of the precision grasp action. The EMG recordings were normalized in time with respect to the onset and offset of the movement derived from a video record. The amplitude of each (enveloped (50 Hz) and rectified) EMG record was normalized with respect to the maximum value obtained on that individual trial.

As inspection of Figure 2 reveals, in the precision grasp variant of the task, for ADM there is an initial interval of low amplitude activity as the posture of the hand is modified, followed by a period of muscle quiescence coinciding with the time of maximum aperture in the whole-hand grasp condition. Thereafter, activity in ADM increased, and was maintained at a relatively stable level through the lift phase. This is in accordance with the requirement that the little finger be extended to avoid contact with the support stand, as the object is lifted. Activity in the FDI muscle was less conspicuously biphasic. In all participants, the maximum level of engagement of FDI occurred just in advance of the moment the object was lifted. In the whole-hand grasp condition (Fig. 3), activity in the ADM muscle reached a maximum following the time of maximum hand aperture, and declined thereafter through the lift phase. In FDI, the maximum level of activity in this condition was registered just in advance of the moment the object was lifted. Notably, the overall level of activity and period of modulation recorded in FDI was markedly greater in the precision than in the whole-hand variant of the task. For the ADM muscle, at least in the period before the object was lifted, the contrast was less conspicuous. There was remarkable consistency across the participants in the degree to which these patterns were expressed. These EMG data provide one context in which to assess the variations in excitability of corticospinal projections to FDI and ADM obtained during observation of the same actions in the main experiment (as described below).

### Table 1
Mean (in mV) ± SD (calculated across 13 participants) responses to TMS evoked in first dorsal interosseus (FDI) at various phases of the action being observed

<table>
<thead>
<tr>
<th>Time of stimulation (ms)</th>
<th>3000</th>
<th>4900</th>
<th>5280</th>
<th>5680</th>
<th>6080</th>
<th>7280</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cue Precision</td>
<td>2.24 ± 1.01</td>
<td>2.21 ± 0.96</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Whole-hand</td>
<td>2.42 ± 1.22</td>
<td>2.29 ± 0.95</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Congruent cue Precision</td>
<td>2.36 ± 1.05</td>
<td>2.25 ± 1.08</td>
<td>2.32 ± 0.94</td>
<td>2.17 ± 0.83</td>
<td>2.24 ± 0.83</td>
<td>2.42 ± 0.91</td>
</tr>
<tr>
<td>Whole-hand</td>
<td>2.37 ± 1.18</td>
<td>2.15 ± 0.77</td>
<td>2.16 ± 0.83</td>
<td>1.84 ± 0.62</td>
<td>2.05 ± 0.89</td>
<td>2.17 ± 0.87</td>
</tr>
<tr>
<td>Incongruent cue Precision</td>
<td>—</td>
<td>—</td>
<td>2.18 ± 0.86</td>
<td>1.94 ± 0.79</td>
<td>2.11 ± 0.88</td>
<td>2.43 ± 0.91</td>
</tr>
<tr>
<td>Whole-hand</td>
<td>—</td>
<td>—</td>
<td>2.22 ± 1.16</td>
<td>2.23 ± 1.12</td>
<td>2.24 ± 0.91</td>
<td>2.20 ± 1.06</td>
</tr>
</tbody>
</table>

Each phase was defined by the time of stimulation relative to onset of the trial. Note that for the Incongruent Cue conditions, the row labels indicate the action that was displayed.

### Table 2
Mean (in mV) ± SD (calculated across 13 participants) responses to TMS evoked in abductor digit minimi (ADM) at various phases of the action being observed

<table>
<thead>
<tr>
<th>Time of stimulation (ms)</th>
<th>3000</th>
<th>4900</th>
<th>5280</th>
<th>5680</th>
<th>6080</th>
<th>7280</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cue Precision</td>
<td>0.61 ± 0.33</td>
<td>0.61 ± 0.44</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Whole-hand</td>
<td>0.60 ± 0.36</td>
<td>0.62 ± 0.39</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Congruent cue Precision</td>
<td>0.65 ± 0.37</td>
<td>0.58 ± 0.33</td>
<td>0.62 ± 0.57</td>
<td>0.57 ± 0.43</td>
<td>0.54 ± 0.37</td>
<td>0.61 ± 0.35</td>
</tr>
<tr>
<td>Whole-hand</td>
<td>0.67 ± 0.47</td>
<td>0.58 ± 0.30</td>
<td>0.61 ± 0.33</td>
<td>0.63 ± 0.36</td>
<td>0.69 ± 0.39</td>
<td>0.67 ± 0.43</td>
</tr>
<tr>
<td>Incongruent cue Precision</td>
<td>—</td>
<td>—</td>
<td>0.64 ± 0.46</td>
<td>0.57 ± 0.29</td>
<td>0.57 ± 0.33</td>
<td>0.60 ± 0.40</td>
</tr>
<tr>
<td>Whole-hand</td>
<td>—</td>
<td>—</td>
<td>0.54 ± 0.24</td>
<td>0.59 ± 0.34</td>
<td>0.67 ± 0.55</td>
<td>0.68 ± 0.54</td>
</tr>
</tbody>
</table>

Each phase was defined by the time of stimulation relative to onset of the trial. Note that for the Incongruent Cue conditions, the row labels indicate the action that was displayed.
Table 3
Outcomes of inferential tests of equivalence applied to the normalized amplitudes of motor-evoked potentials (MEPs) recorded in the first dorsal interosseus (FDI) in the incongruent condition whereby a whole-hand grasp was cued and a whole-hand grasp was observed subsequently

<table>
<thead>
<tr>
<th>Equivalence ($I_{12}$)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cue: whole hand; Grasp: precision vs. Cue: whole hand; Grasp: whole hand</td>
<td>Post onset: 0.16, $P = 0.03$</td>
</tr>
<tr>
<td>Max-precision</td>
<td>0.34, $P = 0.05$</td>
</tr>
<tr>
<td>Max-whole-hand</td>
<td>0.07, $P = 0.01$</td>
</tr>
<tr>
<td>Lift</td>
<td>1.46, 0.36</td>
</tr>
</tbody>
</table>

| Cue: whole hand; Grasp: precision vs. Cue: precision; Grasp: precision | Post onset: 1.38, $P = 0.33$ |
| Max-precision | 1.91, 0.53 |
| Max-whole-hand | 1.42, 0.35 |
| Lift | 0.17, 0.03 |

Table 4
Outcomes of inferential tests of equivalence applied to the normalized amplitudes of motor-evoked potentials (MEPs) recorded in the first dorsal interosseus (FDI) in the incongruent condition whereby a precision grasp was cued and a whole-hand grasp was observed subsequently

<table>
<thead>
<tr>
<th>Equivalence ($I_{12}$)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cue: precision; Grasp: whole hand vs. Cue: whole hand; Grasp: whole hand</td>
<td>Post onset: 0.43, $P = 0.07$</td>
</tr>
<tr>
<td>Max-precision</td>
<td>1.89, 0.52</td>
</tr>
<tr>
<td>Max-whole-hand</td>
<td>1.43, 0.35</td>
</tr>
<tr>
<td>Lift</td>
<td>0.19, 0.03</td>
</tr>
</tbody>
</table>

| Cue: precision; Grasp: whole hand vs. Cue: precision; Grasp: precision | Post onset: 1.62, 0.42 |
| Max-precision | 0.21, 0.03 |
| Max-whole-hand | 0.01, 0.00 |
| Lift | 2.50, 0.72 |

### Results

**No Cue**

We first examined situations in which no advance information was provided concerning the action that might follow. In these trials, MEPs were evoked at the termination of the (no) cue phase, and 100 ms prior to the onset of movement. It was verified that for FDI and ADM respectively, the amplitudes of the MEPs elicited in the 2 phases did not vary systematically from one another ($F < 1$, $P > 0.20$). On this basis, means were obtained for each muscle (Fig. 4). These provided references with respect to which the values obtained in the congruent and incongruent conditions could be assessed.

It is important to note that the responses obtained following the provision of a noninformative (but presumably alerting) cue before the onset of the action then observed—were of greater ($\approx 140\%$) magnitude than those obtained in the controls undertaken prior to and following the observation trials. This reflects an anticipatory increase in corticospinal excitability in the general context of cued action observation (Lago and Fernandez Del Olmo 2011). In relation to the effects of prior symbolic cues indicating the nature of the action that would follow (described in the sections below), these therefore reflect muscle-specific and action goal-related variations of corticospinal excitability superimposed upon the elevation attributable to the generic requirements of the task.

**Congruent Cue**

When the observer saw a whole-hand grasp cued and executed (Fig. 4A, triangles: dashed green line), MEPs elicited in FDI diminished in amplitude following the onset of the actor’s movement and reached their lowest point prior to the maximum aperture of the grasp. When a precision grasp was specified (Fig. 4A, diamonds: dashed blue line), a similar pattern of modulation across phases was obtained; however, the excitability of projections to this muscle was elevated to a greater degree. The difference between the conditions was most pronounced at the maximum aperture of the precision grasp ($F_{1, 108} = 5.91$, $P < 0.05$, $f = 0.23$) — highlighted by filled symbols in Figure 4A.

The pattern of variation observed for ADM was distinct from that of FDI, and consistent with the functional role of the muscle in the actions observed. When a whole-hand grasp was cued and demonstrated (Fig. 4B, triangles: dashed green line), MEPs elicited in ADM were most elevated during maximum aperture. Whereas, when a precision grasp was cued and then seen performed (Fig. 4A, diamonds: dashed blue line), ADM MEPs diminished in amplitude during this phase (Fig. 4B). Thus, the greatest difference between conditions coincided with the maximum aperture of the whole-hand grasp ($F_{1, 108} = 11.09$, $P < 0.01$, $f = 0.32$) — highlighted by filled symbols in Figure 4B.

**Incongruent Cue**

In light of the functionally specific, variations in corticospinal excitability present when the advance information and following action were congruent, we sought to determine the pattern of modulation exhibited when they were incongruent. If successive elements of an observed movement sequence were encoded separately, it would be anticipated that equivalence tests would indicate that corticospinal output was initially in concordance with the action cued, and subsequently (i.e., from one point in time onward) equivalent to that associated with the action that is seen.
When a whole-hand grasp was cued and a precision grasp demonstrated subsequently (Fig. 4A, circles: solid purple line), the amplitude of MEPs elicited in FDI matched initially those obtained for the whole-hand grasp (i.e., correctly cued) (Fig. 4A, triangles: dashed green line). The MEP amplitudes were equated immediately following the onset of movement, through to the phase during which the maximum aperture of a whole-hand grasp would have been anticipated ($P < 0.05$). This is highlighted by the green dashed ovals on Figure 4A indicating values equivalent to those of the cue: whole-hand; grasp: whole-hand condition (the outcomes of the corresponding inferential analyses are given in the upper portion of Table 3, first 3 rows). The sustained impact of the erroneous advance information was such that: at the time the maximum aperture of the precision grip was observed, the MEP amplitudes remained markedly lower than those obtained when a precision grasp was both cued and observed ($F_1, 108 = 4.34$, $P < 0.05$, $f = 0.20$), thus matching in this additional respect the condition in which the observer saw a whole-hand grasp cued and executed. During the final phase of the movement in which the actor was seen to lift the object using a precision grasp however, the MEPs were elevated in amplitude, such that they were equivalent ($P < 0.05$) to those recorded in the precision grasp (i.e., correctly cued) condition. This is highlighted by the blue dashed oval on Figure 4A indicating a value equivalent to that of the cue: precision; grasp: precision condition; green ovals indicate a value equivalent to that of the cue: whole-hand; grasp: whole-hand condition.

A corresponding transition was evident if the advance information indicated (erroneously) that a precision grasp would follow (Fig. 4A, squares: solid red line). FDI MEPs were of equivalent amplitude ($P < 0.05$) to those exhibited when a
precision grasp was both cued and executed (Fig. 4A, diamonds: dashed blue line): during phases in which the maximum aperture of a precision grasp would have been anticipated, and the maximum aperture of the whole-hand grasp was observed. This is highlighted by the blue dashed ovals on Figure 4A indicating values equivalent to those of the cue: precision; grasp: precision condition (the outcomes of the corresponding inferential analyses are given in the lower portion of Table 4, rows 2 and 3). In the former case, MEP amplitudes were reliably greater than when a whole-hand grasp was both cued and observed ($F_{1, 108} = 4.79, P < 0.05, f = 0.21$), thus matching in this additional respect the condition in which the observer saw a precision grasp cued and executed. During the final lift phase of the depicted (whole-hand grasp) action, the MEP amplitudes decreased to a level that could not be discriminated ($P < 0.05$) from those obtained when the (whole-hand) cue and grasp were congruent (Fig. 4A). This is highlighted by the green dashed oval on Figure 4A indicating a value equivalent to that of the cue: whole-hand; grasp: whole-hand condition (the outcome of the corresponding inferential test is given in the upper portion of Table 4, fourth row).

With respect to the FDI muscle therefore, the outcomes of the inferential analyses corroborated the observation that corticospinal output was initially in concordance with the action anticipated, and subsequently equivalent to that associated with the action that is seen. This was the case both in circumstances in which a whole-hand grasp was cued and a precision grasp then demonstrated, and when a precision grasp was cued and a whole-hand grasp shown subsequently. These data thus provide strong support for the hypothesis that successive elements of an observed movement sequence are encoded separately.

An incorrect cue having been presented, variations in the excitability of projections to ADM were attenuated relative to when the nature of the action had been pre-empted faithfully (Fig. 4B). Thus, if a whole-hand grasp was cued and a precision grasp demonstrated (Fig. 4B, circles: solid purple line), the diminution of MEP amplitude through the max-precision and max-whole-hand phases were less marked, than that obtained when a precision grasp was cued and observed (Fig. 4B, diamonds: dashed blue line). The attenuation was most pronounced during the phase in which the maximum aperture of a whole-hand grasp would have been anticipated ($f = 0.21$ vs. 0.32). The values obtained in the 2 conditions did however tend towards convergence (i.e., for the incongruent condition—correspondence with the action observed) during the final (lift) phase of the sequence.

A consistent profile of attenuation—in this case relative to the condition in which whole-hand grasp was cued and observed, was expressed when a precision grasp was cued and a whole-hand grasp demonstrated. Thus, the potentiation of MEP amplitude observed in the cue: precision; grasp: whole-hand condition (Fig. 4B, squares: solid red line) was less marked through the max-precision and max-whole-hand phases than that obtained when a whole-hand grasp was cued and observed (Fig. 4B, triangles: dashed green line). This effect was no longer present during the final (lift) phase of the sequence.

In general therefore, the influence of the incongruent cue appeared to be manifested (i.e., in modulation of MEP amplitude) earlier for the ADM muscle than for FDI. Inspection of the EMG profiles obtained for the real execution of the grasps indicates that for both the precision and whole-hand variants, activation of ADM occurred prior to that of FDI (Figs 2 and 3). Thus, precocious changes in the excitability of corticospinal projection to ADM are in accordance with the relative timing of muscle engagement that is required to generate these movements. Nonetheless, temporal segmentation of corticospinal output in the incongruent conditions was expressed more prominently for FDI than for ADM.

Discussion

We used TMS to probe the excitability of corticospinal projections to intrinsic hand muscles while motions to reach and grasp an object were observed. Each action was preceded by a symbolic cue that either faithfully indicated the required final position of the object, and thus, the nature of the grasp that was required, or was in conflict with the movement subsequently displayed. When the cue was veridical, the responses recorded from FDI and ADM were modulated in accordance with their functional role in the action that was observed (Gangitano et al. 2001). If however the symbolic cue had indicated that the alternative grasp would be required, a quite distinct pattern of modulation was obtained. When a precision grasp was cued and a whole-hand grasp seen executed subsequently, or a whole-hand grasp cued and a precision grasp then displayed, the modulation of corticospinal output to FDI was consistent with the action specified, rather than the action observed—until the terminal phase of the motion sequence during which the object was seen lifted. The excitability of corticospinal projections to the ADM muscle was also modified by advance information. When there was incongruence between the action specified and the action observed, the modulation of corticospinal output to ADM was attenuated relative to that obtained when the cue faithfully represented the required final position of the object.

When the ostensible intrinsic properties of an object (e.g., purported weight) and its motion are incongruent, the variation in corticospinal output exhibited by an observer is commensurate with the kinematics displayed rather than the forces implied (Alaerts et al. 2010). If, however, there is a mismatch between a semantic cue (an object labeled “heavy” or “light”) and a weight-determined kinematic profile, modulation of corticospinal excitability is diminished, and a general inhibition of motor output pertains (Senot et al. 2011). Similarly, if the anticipated and observed kinematics are at odds, there occurs abolition of the modulation associated with the anticipated action, rather than substitution by a profile commensurate with the observed motion. On this basis, it has been proposed that resonant motor plans, primed as observation commences, either proceed to completion or are suppressed as inconsistencies are revealed subsequently by visual input (Gangitano et al. 2004; Fadiga et al. 2005; Hauk et al. 2008; Cattaneo and Rizzolatti 2009).

Unlike most previous investigations, we studied 2 alternative grasp actions (precision or whole-hand). The overall level of activity and degree of modulation recorded in FDI during execution was in accordance with its functional contribution, and was thus markedly greater in the precision than in the whole-hand variant of the task. While the level of ADM activity present during execution of the 2 task variants was not dissimilar, the temporal pattern of modulation was clearly differentiated. In the whole-hand grasp, the ADM recruitment contributed to the enclosure of the object, whereas when the
precision grasp was required, ADM activity served primarily to extend the little finger to avoid contact with the support stand.

The observation of each grasp action also gave rise to a distinct profile of corticospinal excitability. As in some instances the prior symbolic cue misrepresented the grasp that would then be depicted, this design offered the means to determine whether, as the action unfolded, modulation of corticospinal output proceeded in conformity with the grasp that was anticipated, or was altered in accordance with the grasp that was observed. Tests of equivalence provided the strong inferential basis upon which to substantiate this determination. In contrast to the conclusion presented elsewhere—that motor plans once primed proceed through to completion, our findings reveal that modulation of corticospinal output during passive observation may progress first in accordance with the action that is anticipated, and if discrepancies are revealed by visual input, coincide thereafter with the action that is seen. This phenomenon was expressed with particular clarity in relation to the amplitude of motor potentials evoked in the FDI muscle. Indeed, a single mean value diverged from the general pattern (Fig. 4A). And yet for the ADM also, there was little evidence of a pattern of modulation primed by the prior symbolic cue proceeding independently of the observed action. Nor was there an overall suppression of corticospinal output. Rather, there was a transition to the profile of corticospinal excitability designated by the motion seen, albeit at reduced gain.

Single-cell recordings from non-human primates have revealed clusters of neurons in premotor regions that encode distinct sequential elements of distal limb movements (Rizzolatti et al. 1988; Rizzolatti and Fadiga 1998). Groups of cells in ventral premotor cortex that discharge not only during the execution of hand actions but also during their observation can be differentiated on this basis (Gallese et al. 1996). Indeed, subsets of neurons in this region remain engaged when the final parts of an action sequence are not shown directly, but must be inferred from the preceding elements (Umiltà et al. 2001). Prior to the present study, there has been little direct evidence that a similar segmentation of action sequence is expressed via the human motor system during observation of others.

The manner in which an object is grasped is dependent on the purpose of the action. The preference for a comfortable joint-angle configuration at the end of the movement, even when this necessitates an awkward posture when an object is first approached, has been termed the end-state comfort effect (Rosenbaum et al. 1992). Thus, when planning how to grasp an object, account is taken of the prospective demands associated with the goal of the action (Johnson-Frey et al. 2004). Such anticipatory planning implies that the purpose of a movement is represented prior to encoding of the manner in which it will be given effect.

The symbolic cue used in the present study indicated to the observer the purpose of the action, and thus (indirectly) the ultimate joint-angle configuration required for its realization. The phases of movement giving rise to this outcome could be inferred only on the basis of anticipatory planning. Circumstances in which the cue is accurate cannot provide the means upon which to determine whether these phases are represented sequentially by the observer or as an ensemble. If however the symbolic cue specifies one of a set of possible outcomes, and an alternative is then shown, the modulation of corticospinal output that occurs as the action unfolds provides a basis for this determination. In this regard, the outcomes of the present study reveal that humans exhibit temporal segmentation in representing and anticipating the behavior of others.

Congruent spiking modulation of single units and commensurate variations in local field potentials have been observed in the primary motor cortex of macaques during action and action observation (Tkach et al. 2007). The changes in corticospinal excitability registered via TMS when humans perform similar tasks may reflect corresponding processes. That which remains to be resolved is the causal status of fluctuations in functional connectivity among elements of a wider brain network that also includes the inferior frontal cortex (IFC; ventral premotor cortex and inferior frontal gyrus), parietal and temporal cortices, in mediating both concurrent and anticipatory responses to the seen action of others.

In a recent study in which repetitive TMS (rTMS) was used to transiently disrupt regions within IFC, Avenanti et al. (2012) reported abolition of the increases in corticospinal excitability otherwise obtained when the actions of observed others are implied by static pictures. In contrast, when rTMS was delivered to the superior temporal sulcus (STS)—putatively a relay linking “lower order” visual and frontoparietal regions (e.g., Rizzolatti and Luppino 2001), the modulation of corticospinal excitability was more pronounced. These data were interpreted as support for the supposition that a reciprocal control system is engaged in such tasks, whereby activity registered initially in STS influences visuomotor (frontoparietal) processing, which in turn alters information processing in temporal regions. Of particular interest in the present context is the more general hypothesis that information flow from premotor to parietal and middle temporal cortices is determined by the degree to which forward models (Kawato 1999) of sensory consequences can be imputed from the observed actions of others (Schippers and Keysers 2011). As far as we are aware there has not yet been an extension of this line of reasoning to encompass the impact of such task-dependent variations in functional connectivity on the state of circuits in the primary motor cortex. There is however empirical evidence to suggest that activity in presupplementary motor area (pre-SMA) is elevated if prediction is required following transient occlusion of an observed action (Stadler et al. 2011). To the extent that anticipatory and concurrent responses to the viewed actions of others are mediated by distinct patterns of functional connectivity in brain circuits that project to primary motor cortex, a means of mediating the temporal segmentation of corticospinal excitability observed in the present study is provided, at least in principle.

There were differences between projections to the FDI and ADM muscles in the degree to which this temporal segmentation was expressed. In light of the contingent relation between experience of a motor act and the brain activity that arises during observation of that act (e.g., Petroni et al. 2010; Olsson and Nyberg 2011), a parsimonious account of this feature of the data may be derived from our observation that the differential modulation of EMG activity that occurs in the real execution of the precision and whole-hand grasps is considerably more pronounced for FDI than for ADM (Figs 2 and 3). Nonetheless, there are other factors that may be implicated. The features of a cortical circuit mediating visually guided grasping, that comprises the anterior intraparietal area (AIP), ventral premotor cortex (PMv), and M1 are now well documented (Castiello 2005). There is a high degree of functional specificity in the interactions between PMv and M1. Conditioning
TMS applied to PMv has larger impact upon projections from M1 to FDI in the context of a precision grip, and a greater influence on projections from M1 to ADM when a whole-hand grasp is adopted (Davare et al. 2010). The AIP plays a part in mediating the integration of object specifications derived from visual input as the shape of the hand is specified during grasping. AIP activity is however much weaker during a whole-hand grasp than a precision grasp (Begliomini et al. 2007). Corticometoneuronal projections to FDI exhibit higher gain than those to ADM (Ziemann et al. 2004a, 2004b). Thus, whether considered in relation to their descending projections, or the task contexts in which they assume a primary functional role (precision vs. whole-hand grip), the representations of FDI and ADM are differentiated within the cortical circuits that regulate visually guided grasping (see also Ni et al. 2006). A corresponding distinction expressed during the observation of action is therefore to be anticipated.

Is it conceivable that the effects observed in the incongruent trials, particularly at the beginning of each trial, were attributable to an influence of the symbolic cue that transcended the anticipatory modulation of corticospinal excitability that was expressed once the movement was seen to commence? MEPs were elevated (relative to controls undertaken prior to and following the observation trials) in the no-cue condition in which the direction of the arrow provided no information concerning the nature of the forthcoming action. A conspicuous feature of the data presented in Tables 1 and 2 is the similarity of these values to those obtained in the cued conditions at termination of the goal cue (following a 2000 ms presentation) and 1900 ms thereafter (100 ms before movement onset). Indeed, in none of these instances did a value obtained for the cued trials (precision or whole-hand) for either FDI or ADM lie outside the 95% confidence interval defined for the corresponding no-cue trials. It appears therefore that the impact of a symbolic cue on anticipatory movement planning becomes manifest only upon initiation of the action that is witnessed.

Beyond the fundamental knowledge generated by the present study, consideration might also be given to the potential practical application of the findings in rehabilitation, for example following stroke. It has been proposed previously (e.g., Small et al. 2012) that action observation and imitation may prove beneficial in this context. The current outcomes, and indeed those concerning the patterns of functional brain connectivity that mediate concurrent and anticipatory responses to the seen action of others, suggest that interventions based on action observation (e.g., Ertelt et al. 2007) might be tailored—through manipulation of specific task demands (e.g., the requirements for anticipation), to the constellation of cerebral damage sustained by individual stroke survivors.

Notes
Richard Carson thanks Atlantic Philanthropies for their generous support of his research through their funding of the NEIL (Neuroenhancement for Independent Lives) program at Trinity College Institute of Neuroscience. Conflict of Interest: None declared.

References


