

The role of transients in action observation

Geoff G. Cole¹, Timothy N. Welsh², & Paul A. Skarratt³

1. Centre for Brain Science, University of Essex.
2. Faculty of Kinesiology & Physical Education, University of Toronto.
3. Department of Psychology, University of Hull.

Running head: Transients and social IOR

Address for correspondence

Geoff G. Cole
Centre for Brain Science
University of Essex
Wivenhoe Park
Colchester
CO4 3 SQ
E-mail: ggcole@essex.ac.uk
Tel: 01206 872331
Fax: 01206 873801

Abstract.

A large number of studies have now described the various ways in which the observation of another person's dynamic movement can influence the speed with which the observer is able to prepare a motor action themselves. The typical results are most often explained with reference to theories that link perception and action. Such theories argue that the cognitive structures associated with each share common representations. Consequently, action preparation and action observation are often said to be functionally equivalent. However, the dominance of these theories in explaining action observation effects has masked the potential contribution from processes associated with the detection of low-level 'transients' resulting from observing a body movement, such as motion and sound. In the present review, we describe work undertaken in one particular action observation phenomenon ('social inhibition of return') and show that the transient account provides the best explanation of the effect. We argue that future work should consider attention capture and orienting as a potential contributing factor to action observation effects more broadly.

Introduction.

In 2005, Welsh and colleagues described a phenomenon in which participants took relatively more time to prepare and initiate a reaching action to the same spatial location as a co-actor's recent response than to an alternate unresponded-to location. This pattern of response initiation times (reaction times or RTs) has come to be known as the *social inhibition of return* effect (social IOR) because the task and pattern of RTs mimics the inhibition of return phenomenon observed in paradigms in which individual perform alone (see Klein, 2000 for review). A large number of experiments have now been published that have employed this joint action paradigm, all of which have helped to determine the effect's various characteristics and parameters (e.g., Atkinson, Skarratt, Simpson, & Cole, 2014; Cole, Skarratt, & Billing, 2012; Cole, Wright, Doneva, & Skarratt., 2015; Cole, Atkinson, D'Souza, Welsh, & Skarratt, 2018; Doneva & Cole, 2014; Hayes, Hansen, & Elliott, 2010; Janczyk, Welsh, & Dolk, 2016; Lyons, Weeks, & Elliott, 2013; Ondobaka, de Lange, Newman-Norlund, Wiemers, & Bekkering, 2012; Skarratt, Cole, & Kingstone, 2010; Welsh et al., 2005; Welsh et al., 2007; Welsh, Manzone, & McDougall, 2014; Welsh, McDougall, & Weeks, 2009; Welsh, Ray, Weeks, Dewey, & Elliott, 2009). In the present article, we review the literature with a view to identifying why so-called "social IOR" occurs. We will consider the central question of whether the effect is most likely to be due to relatively high level mechanisms associated with the representation of observed actions or, alternatively, whether lower level processes associated with the detection of visual transients provides a better explanation.

Our principal argument will be that this particular effect can be seen as a test case for the understanding of joint action phenomena more broadly. We suggest that the joint action literature has not adequately considered the possible influence that visual transients have on the effects under examination. When a person makes a dynamic movement, it generates a multitude of low level audiovisual signals, each one of which is individually able to 'capture' and summon attention towards it (Skarratt et al., 2010). These include changes to chrominance and luminance, motion

signals generated by the onset and/or continuation of motion, as well as any sound associated with the initiation and termination of the movement (e.g., lifting or pressing a response button or screen). The fact that these constituent features of an action have rarely been considered can be seen in a special issue on joint action published by *Experimental Brain Research* in 2011.

Understanding the Mechanisms of Joint Action reported 30 empirical articles in which only two made reference to body movement capturing attention (Cook & Bird, 2011; Hollander, Jung, & Prinz, 2011). Indeed, the word attention is rarely used within the action observation field. When it is, it is usually employed to describe 'joint attention' in which two people attend to the same stimulus (e.g., Bockler, Koblich, & Sebanz, 2011; for review see Frischen, Bayliss, & Tipper, 2007). There is, however, good reason to suspect that transients play at least a mediating role in action observation. Using the paradigm joint-action experiment, Dolk, Hommel, Prinz, and Liepelt (2013) showed that the 'joint Simon effect' (Sebanz, Knoblich, & Prinz, 2003) can be induced by any stimulus that is salient enough to attract attention to the location where the co-actor would normally make their response. Thus, it is the contention here that visual transients and the attentional processes engaged thereafter may play a primary role in action observation and associated joint action effects, and in social IOR in particular.

Social IOR

The classic cognitive psychology paradigm involves individuals performing tasks alone and the vast majority of what we know about human cognition has of course come from this basic method. However, in recent years psychologists have begun to examine a range of phenomena when individuals act in conjunction with another person (for reviews, see Cole, Skarratt, & Kuhn, 2016; Galantucci, & Sebanz, 2009; Skarratt, Kuhn & Cole, 2012). At the forefront of this movement is research on action observation (e.g., Sebanz & Knoblich, 2009). Interest in this area has been partly motivated by the notion that humans are social animals and spend much of their time interacting with others and sharing cognitive tasks. The resultant work has very much been placed within the

context of action recognition and understanding. Action-perception models typically argue that the mechanisms associated with action and those associated with perception share cognitive representations (e.g., Hommel, 2009; Jeannerod & Frak, 1999; Knoblich & Sebanz, 2006; Prinz, 1997). That is, the same visuomotor processes mediate both performing a given action and interpreting the action in others. One corollary of this, then, is that shared action-perception representations can influence one other and this has indeed been shown in numerous studies in which individuals act alone. For instance, a rightward shift of an observer's attention can facilitate a right-handed button press (Cole & Kuhn, 2010). Because models of action and perception often argue that these systems are tightly coupled and are functionally equivalent, it also follows that observation of another person's dynamic movement can influence one's own action planning. That is, because the effects (sensory and perceptual consequences) of the action are tightly coupled to the action codes that bring about those effects, the observation of someone else generating those effects in the environment activates the effect codes in the observer which concomitantly activates the associated action codes in the observer. Thus, the hypothesized tight coupling between perception and action codes can explain why the motor system becomes active while observing the actions of another individual.

In the basic social IOR paradigm, two participants, or co-actors, sit facing each other at a small table and typically have their right hand holding down a 'home' button (see Figure 1). Each person takes turns to reach out to targets that appear on the workspace, usually via a touch screen monitor, placed flat, or via response button boxes. In the original Welsh et al. (2005) study, co-actors A and B were required to reach to the illumination of a target light that could appear on either the left or right, and make two consecutive responses in an AABBAABB.... sequence. The critical analysis compared RTs for reaching responses to targets that appeared at the same location as the previous target ('repeated') or at the different location ('opposite'). Because co-actors made two responses, Welsh et al. were able to analyse the time it took to repeat their own movement compared to making a different movement. Furthermore, because co-actors alternated responses, this also

allowed analysis of responses to the same location as a co-actor's previous response, again relative to the opposite location. To put this another way, the experiment examined RT to repeat one's own action, compared to a different one, and RT to repeat another person's action or to execute a different one. As well as measuring the time between target onset and the target being touched, Welsh et al. also measured the time between the hand leaving the home button and touching the target, i.e., movement time.

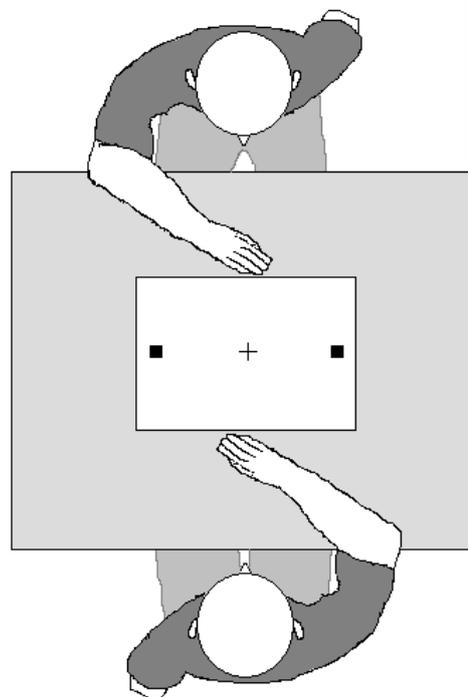


Figure 1. The social IOR paradigm. Co-actors take turn to reach out and touch one of the two targets. Results show that participants are slower to initiate a reach when it is directed to a location where their partner just reached to on the previous turn. Thus, if the person depicted at the bottom reaches to his left, the person at the top will then be slower to reach to his right.

There were two central findings of interest. Results showed that RT to repeat one's own movement was relatively longer compared with executing a different movement. This 'within-person' inhibitory effect, as Welsh et al. called it, is perhaps unsurprising. Psychology has a long history of assessing the (presumed) inhibitory processes that are activated when a range of

behaviours are performed. For example, the likelihood that a dyslexic individual will omit writing a particular letter is partly dependent upon whether the same letter recently appeared. The probability of dropping a letter decreases, in a linear manner, as a function of where in a word it appeared previously (McKay, 1987). The argument is that the initial presentation induces inhibition. A further example, and one more familiar to visual cognition researchers, is the inhibition that occurs subsequent to attention being shifted to an object or location (Posner & Cohen, 1984; see Klein, 2000, for a review).

Perhaps most interestingly, and central to the present article, is the effect which emerged when Welsh et al. analysed the first of the two successive responses participants made in the sequence. That is, the reaching action performed immediately following a co-actor's response. Results showed that RTs associated with reaches to the same location as the co-actor's most recent response were longer than RTs to the different location; referred to by Welsh et al. as 'between-person' inhibition. Thus, if co-actor A had just reached to their right, co-actor B would be take longer to initiate a reach to their own left (i.e., to the same target location) than if they reached to their own right (i.e., to the opposite target location). One other aspect of the data worth noting is that no differences were observed in relation to movement time. Indeed, no subsequent published study that has measured movement times have found any effects (e.g., Welsh et al., 2005; Skarratt et al., 2010) suggesting that the execution of a movement by co-actor B may not be affected by the preceding movement of co-actor A. It is worth noting that this conclusion (i.e., regarding movement execution) is drawn with some caution because detailed kinematic analyses of the executed movements in these tasks have yet to be performed. Thus, future research may want to explore the characteristics of the executed movements in this social task to determine if the trajectories of these movements are affected in manner consistent those in studies in which individuals complete cue-target tasks on their own (e.g., Neyedli & Welsh, 2012). Nonetheless, the extant data suggest that between-person effects are primarily pre-motoric in nature.

Three accounts of social IOR

As is common in visual cognition, the initial studies that employed the arm reaching paradigm were primarily concerned with establishing the empirical parameters of the phenomenon itself, whereas later work began exploring its theoretical implications across an increasingly wide literature. Since the original observation, the phenomenon has been examined in relation to, for instance, other behavioural characteristics (e.g., gambling behaviour, Lyons, Weeks, & Elliot, 2013); whether the inhibition arises from free-choice and self-initiated movements rather than prescribed target responses (Cole, Wright, Doneva, & Skarratt, 2015); whether it is sensitive to the action goals of a co-actor (Cole, Atkinson, D'Souza, Welsh, & Skarratt, 2018; Janczyk, Welsh, & Donk, 2016) and to the identity of a co-actor (Doneva, Atkinson, Skarratt, & Cole, 2017); and whether it occurs in non-typically developed individuals such as those with autism (Welsh, et al., 2009) and those with anorexia nervosa (Dalmaso et al., 2016). Indeed, many authors, ourselves included, presumed that the effect arises from essentially the same inhibitory processes first reported by Posner and Cohen (1984), and thus reflected in the name *social inhibition of return* adopted by Skarratt et al. (2010).

Nevertheless, three main explanations for the phenomenon have now been put forward. In simple terms, the 'co-representation' account (Welsh et al., 2005) posits that the effect is due to a combination of attention cuing and action representation; the 'movement-congruency' account (Ondobaka, et al., 2012) suggests that it is an imitation effect; and the 'transient' account (Cole et al., 2012) argues that even when taking into account the importance of inference and agency, the phenomenon is still better understood as an IOR-like effect induced by lower-level motion transients. Differences and similarities are shared amongst the three theories. For instance, the co-representation and transient explanations both include inhibition as the central process, whereas the movement congruency explanation does not. The co-representation and movement congruency accounts both incorporate action observation mechanisms, whereas the transient account is solely

concerned with capture by motion. Furthermore, response location is not important to the movement congruency explanation but is to the co-representation account.

The co-representation account (Welsh et al., 2005).

Recall that the basic effect shows that an individual is relatively slow when reaching to the same location as a co-actor's previous response. In their original article, Welsh et al. (2005) suggested two central processes that generate this effect. In one, it was said that when co-actor A reaches to location x , the perceptual system of co-actor B represents this observed action as if it had been performed by the observer themselves. That is, consistent with the work on the action observation system and co-representation in joint action (e.g., Sebanz & Knoblich, 2007), the observation of a goal-directed action engages the response codes that mediate the initiation of the same action in an observer. As a result of this response co-activation and representation, any subsequent effect which usually results when a person performs an arm reach when acting alone should therefore also occur in a passive individual who merely observes the action taking place. As we described above, inhibition is involved in many cognitive and neural processes including motor behaviour. This is also known to be the case in the specific example of reaching out one's arm; as shown in the original Welsh et al. article where participants took more time to initiate their own arm reach to a repeated than to a different location. This phenomenon was also replicated by Cowper-Smith, Eskes and Westwood (2013; see also Neyedli & Welsh, 2012) in participants acting alone, i.e., outside of the social IOR paradigm. Thus, the argument is that the consequence of merely seeing an action is that the observer will be relatively slow to perform the same action immediately after. This aspect of the Welsh et al. explanation therefore relies on models that link perception and action (e.g., Prinz, 1997). Welsh et al. also suggested that the mirror neuron system (di Pellegrino, Fadiga, Gallese, & Rizzolatti, 1992; Rizzolatti & Craighero, 2004) acts as a mediating mechanism. This system is known to be active when an individual perceives and performs a similar action. As Welsh et al.

(2007, p. 955) stated, “We hypothesize that the activation of the mirror neuron system during the observation of the response mimicked the activity association with the actual response”.

The second component of the Welsh et al. explanation concerns attention orienting, specifically IOR. In their seminal article, Posner and Cohen (1984) showed that RTs to detect a stimulus are relatively long if the target appears in a location where the observer’s attention had been recently oriented. In the social IOR paradigm therefore, co-actor A’s reaching movement to a location can act as pre-cue, orienting co-actor B’s attention to the location and subsequently lengthening RTs as a direct result of inhibition emerging there. The inhibition that is thought to be occurring in the Welsh et al. account is thus assumed to be partly due to inhibition elicited by attentional orienting and partly due to motor movement inhibition (if one can indeed draw a distinction between the two processes; see Rizzolatti, Riggio, & Sheliga, 1994; Welsh & Weeks, 2010).

The movement congruency account (Ondobaka et al., 2012).

It is well established that humans have a tendency to mimic the behaviour of others (e.g., Chartrand, 1999). This tendency can occur in a wide variety of ways from the adoption of a foreign accent through to the development of institutionalised practices in an organisation. Not only do humans tend to imitate others in this general sense, observing another person perform *a specific dynamic movement* can facilitate the same movement in an observer. This mimicry has been extensively examined with the so-called *direct matching* paradigm. Indeed, the imitation account of social IOR effectively argues that the standard procedure is actually a direct matching experiment. In the basic direct matching paradigm (e.g., Brass, Bekkering, Wohlschlagel, & Prinz, 2000; Liepelt, von Cramon, & Brass, 2008), a photographic image of a hand is presented oriented such that it is in the same position as the participant’s own response hand. One of the fingers on the stimulus hand lifts which is immediately followed by a target that is superimposed onto the hand. The participant is required to discriminate the target and respond by lifting one of their own fingers. The critical

manipulation is the compatibility between the finger raised on the image hand and the finger that needs to be raised by the participant. For example, a 'compatible' trial is one in which the index finger is raised on the stimulus hand and the participant needs to raise their own index finger. An 'incompatible' trial is one in which the middle finger is raised on the stimulus hand but the participant needs to raise their index finger. RTs are typically shorter on compatible trials than on incompatible trials. As with most other action observation studies, results from this paradigm are usually interpreted with reference to mechanisms that link perception and action. Essentially, observation of the finger movement primes an observer to imitate the same movement. Furthermore, an arm movement direct matching effect was reported by Kilner, Paulignan, and Blakemore (2003).

With respect to the social IOR paradigm, the movement congruency explanation argues that when co-actor A reaches out to, say, their left, this action primes co-actor B to reach out to their own left, i.e., an identical movement within an egocentric framework. This compatibility and same response priming will manifest as relatively short RTs when reaching to the location opposite to where a co-actor just reached to. Or to put another way, RTs will be longer when reaching to the same location because when reaching to the same location a different dynamic movement occurs. The movement congruency account thus makes the exact same prediction as the co-representation account but with a different mechanism (facilitation instead of inhibition).

The transient account (Cole et al., 2012).

An abundance of work has examined the mechanisms and processes associated with the processing of transient events and attention capture. Beginning with the seminal work of Yantis and Jonides (1984), research began to examine which properties of a visual scene are particularly effective at marshalling cognitive resources (e.g., Folk, Remington, & Johnston, 1992; Jonides & Yantis, 1988; Yantis & Gibson, 1994; Yantis & Hillstrom, 1994; Yantis & Johnson, 1990). A central finding is that a relatively large luminance change is one such stimulus (i.e., Posner, 1980). There are

clear adaptive reasons as to why luminance change will receive preferential processing, and virtually all mammalian visual systems include a pathway dedicated to the (very rapid) processing of luminance change information (i.e., magnocellular channel; see Steinman, Steinman, & Lehmkuhle, 1997). This notion is explicit in the behavioural urgency hypothesis (Franconeri & Simons, 2003), which attempts to predict what class of stimuli ought to capture attention (e.g., looming more than receding stimuli; Lin, Franconeri, & Enns, 2008; Skarratt, Gellatly, Cole, Pilling, & Hulleman, 2014). Increased sensitivity to visual transients is an extremely efficient evolutionary strategy for detecting potentially dangerous object/animal appearances because virtually all such events will be accompanied by a luminance change. Indeed, much of the attention capture work concerned the issue of whether any such event can attract attention in the absence of changes in luminance (e.g., Cole, Kentridge, & Heywood, 2005; Gellatly & Cole, 2000). Furthermore, there will be little or no cost in shifting attention to a non-threatening visual event; better to make a false positive than a false negative. There is therefore good reason to assume that virtually all transient events have the potential to attract attention. These events include another individual moving a body part.

The transient account of social IOR thus stipulates that the reaching movement of co-actor A triggers two responses in co-actor B: an attentional orienting response to the reached-to location, as well as the activation of the mechanisms that lead to the emergence of IOR. Target onset and the arm movement may be said to be functionally equivalent to the luminance cue in the classic IOR procedure. That is, a peripheral transient event that triggers an automatic orienting response in an observer. The return of the arm back to the home button can also be said to be functionally equivalent to the central cue that immediately follows the peripheral cue in the basic IOR procedure. This cue has the effect of withdrawing attention from the peripheral location. IOR is then measured and defined as the relative slow speed with which targets presented at the cued peripheral location can be processed. In this account of social IOR, *any* event that has the capacity to attract attention can induce IOR, whether it be a door opening, the appearance of an internet pop-up, a friend waving, a Posner cue, or an arm reach in the social IOR task.

Why does social IOR occur?

The work undertaken so far can be characterised as addressing the following, sometimes related, questions. What occurs when visual transients are minimised, equated across conditions, or even abolished? How is the phenomenon affected when co-actors respond in very different ways? How long does it last? What happens when the effects of transients are pitted directly against the effects of action? Do its basic characteristics behave like attention or action effects? Is it sensitive to relatively high level manipulations such as intended goal? It is worth noting that not all experiments have been necessarily intended to assess these questions but their designs often do allow conclusions to be made. Furthermore, the first two questions stated above are central to the transient and movement congruency accounts, respectively. Moreover, it is relatively easy to examine these two accounts because they unambiguously predict a particular pattern of data. This is less so for the co-representation account because it includes aspects of the other two theories.

The influence of visual transients.

Welsh et al. (2005) first raised the possibility that attention capture by low-level transients could explain the basic social IOR finding. They therefore undertook a control experiment designed to rule out capture by one such stimulus. Recall that in the original paradigm, participants reach to the location of a light that illuminates (onsets) abruptly. Even when it is not their turn to respond, this stimulus is fully visible to the co-actor and inevitably attracts their attention. It may therefore be this target stimulus alone that induces IOR. In the control experiment of Welsh et al., participants wore goggles with liquid crystal lenses that became opaque 20 ms before the onset of their partner's target and remained so until 10 ms after the target had been extinguished. Thus, the "observing" participant saw the response of the acting partner, but not the sudden onset target stimulus that drove the response. Results showed that IOR still occurred, ostensibly ruling out the possibility that the transient associated with target onset induces the effect. This experiment did not however control for the other visual transients that occur in the basic procedure, as co-actors still saw their

partner's entire peripheral arm movement. In a follow-up article therefore, Welsh et al. (2007; Experiment 1) used the goggles procedure not only to occlude target onset but all of the arm reach (but not arm return). Again, social IOR was observed. This finding was taken as evidence that attentional capture from motion of the limb towards the target does not explain the phenomenon. Findings from these partial-vision experiments have also been used as evidence for the involvement of the mirror neuron system. One of the experimental techniques employed in mirror neuron research has been to occlude part/most of the observed action (Umiltà et al., 2001). Studies using this approach show that the mirror neuron system is activated under such conditions. This suggests that it is not the specifics of the movement that is represented, but the goal of the action. The Welsh et al. goggles procedure did not however occlude the arm returning, itself a large visual transient. Thus, the motion signal generated by this arm return could have been enough to induce the effect.

Of course, not every visual transient associated with a partner's response can be occluded; no possible social IOR effect could occur if a participant does not see or know where their co-actor just reached to. Skarratt, et al. (2010) therefore took the approach of keeping visual transients to a bare minimum by obscuring all visual information except the most central portion of the response. That is, the observing participant witnessed only their co-actor's hand beginning to leave, and return to, the home position. This occlusion of the peripheral segments of the movement was achieved by the positioning of two large physical barriers between participants, one to the left and the other to the right, and so restricting all visible information to a central aperture along the vertical midline. Results again showed that social IOR was observed. However, obscuring all peripheral information with the use of physical barriers, or goggles, only controls for the effects of *peripheral* transients. The visibility of the central portion could still be enough to generate attention cueing and thus inhibition. This suggestion is based on the well-established fact that attention-capturing cues do not just orient attention to the specific cued location, regions and objects in close proximity to the cue also receive facilitated processing; an effect central to the gradient model of attentional allocation (La Berge & Brown, 1989). An important aspect of this basic finding with respect to the present work, is the fact

that the vertical meridian plays a central role in this process. Cues have been shown to have greater facilitatory effects on targets presented within the same hemifield relative to the opposite hemifield (see Henderson & Macquistan, 1993). As one might expect following an attention shift, within-hemifield attention capture leads to within-hemifield inhibition. For instance, Bennett and Pratt (2001) cued one of four possible locations (two in each hemisphere) and presented targets at one of 441 possible positions (each 1 degree apart). Results showed that IOR spreads within hemisphere. Similarly, whilst employing a visual search task and keeping orienting distance constant, Pierce, Cruse, and Green (2017) showed that IOR was greater when attention was reoriented to a new location within the same hemifield relative to a new location within the opposite hemifield. The central point with respect to social IOR is that cueing a hemifield with a salient motion transient, even if not particularly peripheral (i.e., near to fixation), is likely to induce IOR for the peripheral target location. Furthermore, even cues that are more centrally located than those mentioned above, and ones that include minimal motion transient (i.e., central gaze cues), have now been shown to induce IOR under some circumstances (see Frischen & Tipper, 2004; Okamoto-Barth & Kawai, 2006). The general point here is that, despite suggestions to the contrary by many authors in the field (e.g., Lyons, et al., 2013), obscuring all but a very small portion of the available transients associated with a partner's response does not rule out the possibility that a motion signal induces social IOR. One does have to add however that no researchers have yet systematically examined whether restricting the visibility of an arm reach induces within-hemifield inhibition. The absence of such an effect would prevent a straightforward application of a classical IOR account. As none of the lower level visual information is localised to the specific inhibited regions, the slowing effect may have occurred in restricted vision experiments because participants *inferred* their partner's response (but see Welsh, et al., 2014, and below, for an assessment of an inference effect). This effect would not be expected given the purely mechanistic IOR account described above.

Varying the size of the transient

One approach to assessing whether visual transients drive social IOR, is to ask whether the behavioural characteristics of the effect are similar to the characteristics of attention. A number of studies have examined the relationship between the salience or size of an attention cue and its ability to shift attention. For example, Fuller, Park, and Carrasco (2009; see also Kean & Lambert, 2003) showed that rather than being 'all-or-nothing', the magnitude of an attention cueing effect is related to the salience of the cue. Later work (Diaz-Tula, Morimoto, & Vanvaud, 2015) additionally reported that one's ability to *withdraw* attention from a location is also related to the salience (i.e., brightness contrast) of the stimulus at the attended location. We can therefore ask whether relatively large motion signals induced by the observed arm movement in social IOR creates relatively large orienting effects.

Of the fourteen published experiments that have happened to include changes in motion signal size across conditions, seven have found a significantly larger social IOR effect when the signal was larger. For example, in one of the restricted-vision conditions of Welsh et al. (2007; see above), co-actors saw only 50 ms of the arm movements in one block of trials and the whole arm movement (i.e., reach and return) in another block. Results showed a larger effect in the full vision condition relative to the partial vision condition. In Manzone, Cole, Skarratt, and Welsh (2017), one of the co-actors was instructed to reach out to the target, as in the standard procedure, whilst their partner was required to press either a left or right hand button using their left and right hand. This meant that one co-actor saw large visual transients when their partner responded (i.e., an arm reach), whilst the other co-actor saw a small transient (i.e., a button press). Results showed that social IOR was significantly smaller in the latter condition. Interestingly, this effect was itself reduced under conditions of partial vision, again via the use of goggles. This pattern of effects is predicted by the transient account since the difference in transient signal between an arm reach and button press is reduced under conditions of partial vision. Note however that this smaller effect is also consistent with the action co-representation account because different responses were being observed and

executed – without a co-representation of the same response, simulation of the response should not occur and the subsequent social IOR should not emerge.

Some of these data therefore support the transient explanation, where the size of the signal appears to modulate the magnitude of social IOR, as the account predicts. However, this does not always occur. Seven of the fourteen experiments that manipulated size of transient did not find a difference in the size of social IOR.

Of all the experiments that have manipulated transient size, the work of Hayes et al. (2010) is perhaps the most pertinent. The experiment is particularly relevant because on half of their trials, a co-actor would reach to the target location whilst on the other half they would reach to the other (i.e., opposite) side. This meant that sometimes a participant saw a large transient *all on one side of the display* (i.e., target onset plus arm movement), whilst other times the transient signals occurred on *both sides* (i.e., target onset on one side and the arm movement on the other). This contrasts with the other experiments described above, in which the transient-size manipulations were less pronounced since they all occurred on one side of the display. Furthermore, the experiment of Hayes et al. manipulated where participants reached across blocks such that either both co-actors reached to the target, both reached to the opposite side, one reached to the target and the other opposite, or vice versa. The authors presented means for each of the four blocks. These means show that in the two blocks where a partner reached to the target location, i.e., where the participant saw a large transient on one side of the display, the social IOR effect was 26 ms and 15 ms. In the two blocks where a partner reached to the side opposite to where the target appeared, i.e., where the participant saw a transient signal on both sides, the effect was 4 ms and 10 ms. Although one has to be cautious in this interpretation because the authors did not analyse these particular means for statistical significance, these results do suggest that the location and size of the visual transients had additive effects on IOR. When these all occur on one side, the magnitude of the effect is large relative to when transients appear on both sides, as the transient account predicts. Note that there

was also an overall effect of the arm movement even when transients occurred on both sides. It is tempting to argue that this is due to action observation processes; processes above and beyond attention capture by the transient signals. However, the arm movement signal is inevitably larger than the target onset signal. This itself could explain the larger effect induced by the arm compared to the target.

One particularly interesting aspect of the Hayes, et al. study is the way in which the target event and arm movement were considered as being processed by different systems. The authors stated that they aimed to dissociate the effects of “stimulus alerting events” (i.e., target onset) from that of observed actions. For example, “Our empirical objective was to determine the relative importance of the alerting signal and observing another performer on between-person inhibitory effects” (p. 297). Similarly, the authors made reference to the “stimulus *and* observed movement” (italics added), and when discussing their findings, they stated that “observation of another person’s movements mediated and, in some cases, completely negated the impact of a stimulus event” (p. 304). For these authors then, it appears that seeing a large arm movement occurring immediately in front of an observer is not a ‘stimulus event’, or at the very least is less of a stimulus event than is the target onset, despite the fact that the latter is inevitably much smaller. This, we argue, is indicative of joint action work in general. Action observation is considered as only stimulating perception-action mechanisms whereas processes that arise from the detection of transients are not considered as important, if considered at all. Of course the authors might contend that they did not mean that arm movements are not a stimulus capable of capturing attention, but the manner in which (or the channel through which) these body/movement related stimuli and sudden onset stimuli are processed are different.

Mimicking the arm movement transient

In an effort to control and/or examine the role of visual transients, a number of authors have included a (usually blocked) condition in which the response by a co-actor is replaced by a luminance

cue that appears at the location where the co-actor would have responded. In other words, a stimulus that mimics the response of a partner. This approach has been undertaken either when a co-actor is absent or when they are present but does not respond. A variant of this method is to present a stimulus that traverses the workspace between the co-actor's home button and response location. As Welsh et al. (2007) stated, the rationale is that "if between-person IOR is the result of motion alone, then movement of the [stimulus] towards one of the targets will evoke IOR". The authors (Experiment 3) replaced the arm movement with a triangle that moved part way across the display (similar to the restricted vision of the hand movement condition in Experiment 2). Results showed that such a stimulus did not induce IOR, thus supporting a non-transient account of the phenomenon. Indeed, Lyons, et al. (2013) took this effect, and the partial-vision goggles experiments, as evidence that social IOR is due to the action observation system: "Together these effects lead us to suggest that between-person IOR is due to the mirror neuron system," (p. 5). Replacing a co-actor's reach with a luminance cue has now been employed on three other occasions, two of which showed an inhibitory effect. For example, Doneva and Cole (2014; see also Cole et al., 2018) replaced the arm reach with an animated black square that extended from the home button to the target location (i.e., became a rectangle). Such results do therefore suggest that the arm reach acts as an effective attention cue. Skarratt et al. (2010) however showed that the presence of social IOR (under restricted viewing) occurs only when one performs with a real person and not with an automaton. The automaton was a pre-recorded partner who was projected life-size onto a screen opposite the participant. This projection therefore included much of the transient information that occurred in the real biological partner condition, yet did not give rise to the same inhibitory effect. Although this finding suggests that transients may not necessarily cue attention in the context of social IOR, one has to be cautious when interpreting null results, especially when the experiment currently stands as the only one in which social IOR has been examined in relation to simulated rather than real biological movement.

There is, however, an inherent problem with any experiment that examines the transient hypothesis, or attempts to control it, by mimicking social IOR data with attention capturing cues. Such experiments effectively become a replication of a Posner and Cohen (1984) cuing experiment; a procedure well known to induce attention orienting. Any failure to show a cueing effect when an arm reach is replaced by a luminance cue may effectively be a failure to induce a cueing effect. The corollary to this is that any effect found is simply a replication of a cueing experiment. In other words, showing (or not) a cueing effect with a salient cue may not actually tell us anything about the mechanisms that give rise to social IOR. Of course, there is the alternative possibility that social and non-social stimulus features are processed in separate and independent channels (for example see Böckler, van der Wel, & Welsh, 2014).

Co-actors perform different actions.

Recall that Ondobaka et al. (2012) argued that when a person reaches out to, say, their right hand side this will facilitate the same egocentric movement in an observer. Thus, in the basic social IOR procedure, responses will take longer to initiate when reaching to the opposite location than when reaching to the same location because this action does not mimic the movement just observed. One method that will allow a direct assessment of the movement congruency account is where co-actors perform different kinematic movements. This is because no imitation can take place when there is no congruency of actions.

Atkinson et al. (2014) employed a design with two main conditions in which the movements performed by co-actors were either the same as each other (i.e., reaching), with only the stimuli at the reached-to location changing, or a second in which movements were different and the stimuli remained the same. In the latter case, one co-actor reached whilst their partner merely pointed to the target location. The authors reasoned that if social IOR is due to action observation mechanisms, the basic effect should not be observed when co-actors perform different movements because such movements are always incongruent (i.e., one reaches, the other points). In contrast, no modulation

of the basic effect should be observed when actions are the same. Results showed that when co-actors performed the same movements, social IOR was modulated (by properties of the stimuli). In contrast, when co-actors performed different movements, social IOR was not modulated. That is, the effect was no bigger, or smaller, if both participants had been reaching compared to if one had been reaching and the other pointing. This does not therefore support the movement congruency account since no imitation, i.e., movement congruency, can occur when a participant is required to perform a different action than their partner.

One can argue that predicting no modulation of social IOR when co-actors perform the same action, as Atkinson et al. did, is somewhat conservative given that many cognitive phenomena can be manipulated by additional factors. A more liberal hypothesis therefore is to state that an action observation effect should be present, even if modulated, when co-actors perform the same actions. However, in a further experiment reported by Atkinson et al., the authors found no social IOR effect at all even when co-actors performed identical actions but to slightly different locations. In this experiment, there were two left and two right response locations (see Figure 2). That is, the left/right response positions were not shared; each co-actor had their own response positions. Both the movement congruency and co-representation accounts predict that social IOR should occur under this condition because co-actors perform the same action- although responses were to different locations, recall that the Welsh et al. explanation includes action observation as a component. However, as stated, no social IOR effect occurred in this situation. Doneva and Cole (2014) also examined the hypothesis that social IOR should occur only when co-actors can imitate each other. The authors replicated the basic procedure and included an additional block in which a (confederate) co-actor sat in an elevated position and responded with her feet. Results showed that social IOR was observed in both the standard (i.e., confederate arm reach) and feet response conditions. As with the results of the reaching/pointing experiment of Atkinson et al., it is difficult to

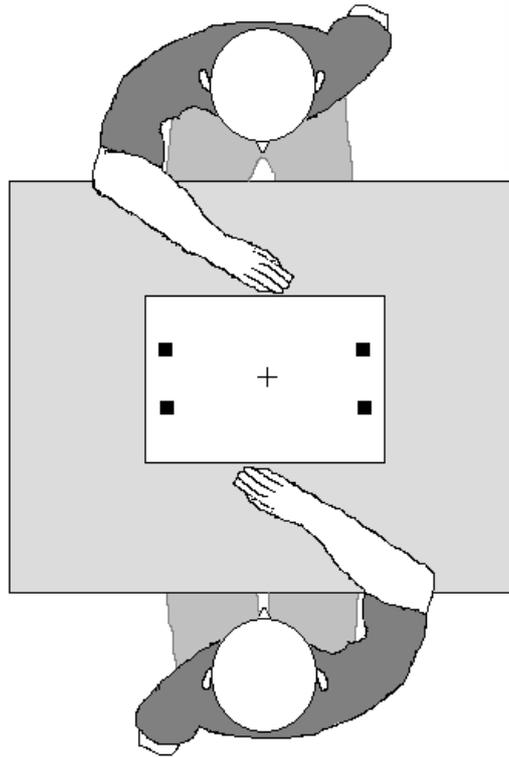


Figure 2. Response location arrangement used in Atkinson et al. (2014).

argue that these results are due to an imitation mechanism that represent the same action that is perceived and performed. One could argue however, that imitation processes were still operating in both experiments. For instance, even when one co-actor uses her foot to reach to her right, a co-actor who then reaches to his right with his hand/arm is still in some sense performing a congruent movement; he is also reaching to his right. This of course revolves around the issue of what is an appropriate definition of imitation.

The Mazzone et al. (2017) experiment described previously represents the most recent assessment of whether social IOR occurs when co-actors perform different tasks. Recall that in their variant of the basic procedure, one co-actor reached and the other co-actor used their left and right hands to press a left and right button in response to a target that appeared on the left or right. As with Atkinson et al. (2014), and Doneva and Cole (2014), results showed that social IOR was evoked irrespective of the action that one's co-actor performed when the full movements and target onset

were observed. When only a portion of the aiming movement was observed, however, no social IOR was observed in the person performing the key press.

As well as refuting the imitation account, these results support the transient account; as long as a person shifts another person's attention to a location, irrespective of how this is done, inhibition will be induced at that location. This is perhaps most explicit in a variant of the social IOR procedure undertaken by Doneva et al. (2014.) In this set up, *all* actions of one co-actor were occluded; only their shoulders and head could be seen. Rather than reaching out, this co-actor operated one end of a wooden arrow behind a screen. The arrow was pivoted in the middle and its front protruded through the screen and would point to the target. It was thus conceptually similar to the finger pointing procedure of Atkinson et al. (2014) but, unlike that paradigm, no body actions were seen at all. Results showed that the same inhibitory effect was observed. That is, RTs were relatively slow to reach to a location that had just been pointed to by the wooden arrow. Again, as long as a location is cued, irrespective of how, a reach to that location will be slowed, as predicted by the transient hypothesis.

The time course of social IOR

Another approach to determining why social IOR occurs is to ask how long the effect lasts. That is, at what point after seeing an action will its effect on the observer cease? The rationale is that if action representation mechanisms are responsible then the phenomenon should persist for some duration. This is because, when an action is viewed, time will inevitably elapse between the observed and performed action. Work on motor area activity following action observation suggests, as one would predict, that the effects last for many seconds. For instance, Lestou, Pollick, and Kourtzi, (2008) found peak activity in motor areas up to six seconds following action observation. Similarly, Gazzola, Rizzolatti, Wicker, and Keysers (2007) found this figure to be nine seconds. Furthermore, mirror neuron activity is thought to be associated with delayed imitation in which the performed action is prevented for a few seconds (Kruger et al., 2014), several minutes (Rogers,

Young, Cook, Giolzetti, & Ozonoff, 2008), and even days (Meltzoff & Moore, 1994). Action representation accounts of social IOR therefore makes the firm prediction that the effect should persist for a number of seconds. With respect to the duration of standard IOR, it appears to last for 3 seconds at most (but see Tipper, Grison, & Kessler, 2003). Although this is also within the range of seconds, it is still shorter than what one would expect, and shown, with action observation processes.

Two studies have so far examined the time course of social IOR, with the basic method being the manipulation of the time interval between consecutive targets. Skarratt et al. (2010) employed target onset asynchronies of 1200 and 2400 ms, and found that the effect only occurred at the shorter interval. Doneva et al. (2016) employed intervals of 1000, 2200, 2400, and 4600 ms, and similarly found that the effect only occurred at the shortest duration. These two experiments suggest that social IOR is not a long phenomenon, expiring at some point before 2200 ms. Thus, the time course of the phenomenon does not support the notion that it is due to action representation mechanisms associated with imitation. This does not therefore support the two explanations that include action representation as a component. One does have to note that the imitation account places greater emphasis on action observation mechanisms that does the co-representation account. Recall that the latter also includes the mechanisms of IOR as a major component. Thus, one might say that although this time course analysis refutes the imitation account, it is less equivocal in terms of the co-representation account.

Is social IOR sensitive to goals?

A number of authors have argued that if social IOR is mediated by the action observation/planning system then it should be sensitive to the goals of co-actors. This goal-sensitivity should emerge because the action observation system is known to be responsive to the goals of an action rather than the dynamic movement that achieves the goal (e.g., Bekkering, Wohlschläger, & Gattis, 2000; Gattis, Bekkering, & Wohlschläger, 2002; Rizzolatti, Fabbri-Destro &

Cattaneo, 2008). Indeed, Longo, Kosobud, and Bertenthal (2008) suggested that the “default response” of the action observation system is to represent the goals of the movement. Furthermore, goals have been shown to influence action planning in the direct matching paradigm of Wohlschläger and Bekkering (2002), in which participants view and respond with a finger movement (see also Bouquet, Shipley, Capa, & Marshall, 2011).

Fourteen social IOR experiments, across five articles (Cole et al., 2012; Cole et al., 2018; Janczyk et al., 2016; Ondobaka et al., 2012; Ondobaka, Newman-Norlund, de Lange, & Bekkering, 2013; see also Hayes et al., 2010), have now been published that manipulate goal compatibility. In the standard procedure, co-actors always have a single goal in that they are required to reach to and touch the target. When, in contrast, goals are manipulated, co-actors perform one of two tasks on each trial. Importantly, the goal is either compatible with their partner’s previous response or incompatible. For example, in the experiment of Cole et al. (2012; Experiment 3), a pencil was placed at both target locations and co-actors were required to either use it to write a digit on an adjacent pad or use its opposite end to erase a digit. This task was blocked so that on one block both co-actors wrote, on another block both erased, on a further block one wrote whilst the other erased, and in the last block, vice versa. Results showed the usual social IOR effect but one that was not modulated by what task the pair performed. That is, there was no interaction between response location and goal compatibility suggesting that the magnitude of the social IOR across the two conditions was not different. This was also the case in two additional experiments using different tasks (e.g., pick up a cup and either pantomime a drinking action or throwing liquid over one’s shoulder). Ondobaka et al. (2012; see also Ondobaka, et al., 2013) in contrast did report a goals effect on social IOR. In their experiment a single playing card appeared at each of the target locations and participants were required to reach and touch the highest card on some trials and the lowest on others. Again, co-actors either had the same goal (e.g., both reach to the highest) or a different goal (e.g., one reaches high, the other low). Unlike Cole et al., results showed that a social IOR effect only occurred when both had the same goal.

A further assessment of the goals issue was undertaken by Cole et al. (2018). The authors made the point that goals can be defined, and thus experimentally operationalised, in a number of different ways. Indeed, it was suggested that the conception of a 'goal' is more experimenter-dependent than reflecting a distinct cognitive process. For example, one can argue that participants always have the same goal, that of responding to the appearance of a target, or even completing an experiment for a researcher. Cole et al. thus defined and operationalised goals in a number of ways (e.g., action at the end-point location) and attempted to show a goals effect in a series of seven experiments. No effect of goals on social IOR was observed. One experiment even included a collaborative task in which co-actors were given the joint goal of completing a dot-to-dot drawing together by picking up a pencil at the cued location. Additionally, the authors could not replicate the goals effect reported by Ondobaka et al. (2012) when closely replicating their method. Using a variant of the social IOR goals procedure, Janczyk et al. (2016) have also reported the absence of a goals effect.

Overall, the empirical work does not support the notion that the mechanisms responsible for social IOR also represent the goals of the action. This in turn suggests that the phenomenon is not due to the system that is known to represent goals, i.e., action observation and planning mechanisms.

Welsh, McDougall, and Weeks (2009).

Virtually all published experiments that enable an examination of why social IOR occurs do so by assessing one, and sometimes two, of the three main theories we have set out. The study reported by Welsh et al. (2009), in contrast, is unique in that it happens to test each of the three accounts in a single experiment. Although the authors did not make this point (recall that the Ondobaka et al. explanation was not published until 2012), each theory is effectively pitted directly against each other. Furthermore, the experiment's design possibly enables an examination of the *relative contribution* that attention capture, and resultant inhibition at a location, makes to the

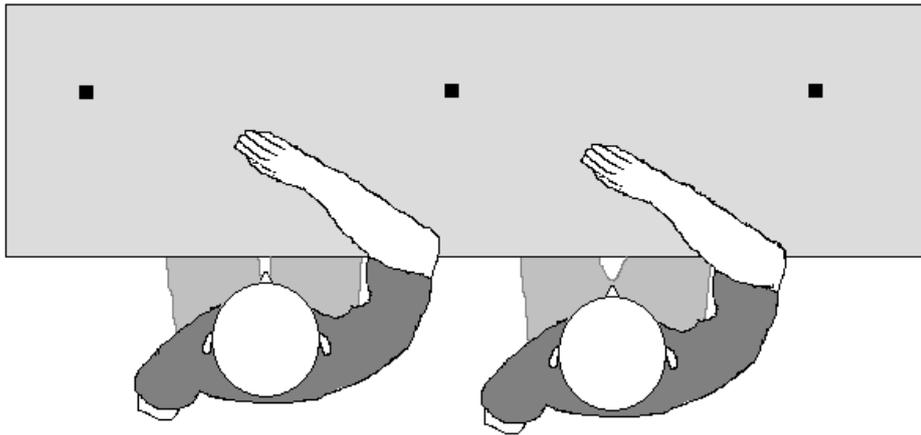


Figure 3. Sitting positions and target locations employed by Welsh et al. (2009).

effect compared with the inhibition of the same action. Recall that the Cole et al. account is solely based on the former, and the Welsh et al. account includes both the former and latter.

Unlike the standard procedure in which co-actors sit opposite each other across a table, Welsh et al. (2009) had participants sitting next to each other facing the same way (see Figure 3). That is, one sat on the left, the other to their right. Located on the table were three possible response locations. These locations were equally spaced and the middle location could be a target position for both co-actors. This spatial arrangement meant that a participant's response (immediately following their partner's response) could be classified in one of three ways: 1) to a different location with a different action, 2) to a different location with the same action, 3) to the same location with a different action. For instance, if the co-actor sitting on the left had just responded to her left and the co-actor sitting on the right then reaches to his right, this latter response would be to a different location using an action that had a different direction (i.e., reaching to the right). Similarly, if the left co-actor had just responded to her left and the right co-actor then reaches to his left, to touch the central location, this latter response would be to a different location but using an action that had the same direction. Finally, if the left co-actor just responded to her right, to touch the central location, and the right co-actor then reaches to his left to touch that same

location, this latter response would be to the same location using an action that had a different direction (i.e., she reaches right, he reaches left).

Recall that when participants sit opposite each other, as in the standard paradigm, both the movement congruency and the two inhibition accounts make the same prediction. This however is not the case when participants sit adjacently facing the same way. Because it concerns direct matching (i.e., movement congruency), the Ondobaka et al. account makes the firm prediction that when a participant has just reached out to her right, this should facilitate the same movement in their co-actor on the next trial. Or to put another way, when a participant has just reached out to the right, their co-actor should be relatively slow when required to reach out to their left. Welsh et al. however found that RTs were significantly shorter in the latter condition, that is, participants were faster to perform a different action. Given the unambiguity of what the movement congruency account predicts, these data clearly refute this explanation; a direct matching effect does not occur. Indeed, these results support the co-representation account which states that co-actors inhibit the same movement as the one just observed.

As stated, the Welsh et al. (2009) experiment also pits the transient and co-representation accounts directly against each other. This is because on some trials participants performed the same action as their co-actor's previous response but to a different location whilst on other trials participants performed a different action but to the same location. These two components are therefore balanced across the two conditions. If the action component of social IOR and the location component exert their effects in equal measure, these two conditions should produce no significant difference in RT. Any difference, in contrast, would reveal which of the two components contributes the most. Results showed that RTs were longer when co-actors responded to the same location compared to when they performed the same action. In other words, location-based inhibition was greater than action inhibition. In sum, these data not only support the transient account, they also support the co-representation explanation because both an action and location are inhibited.

However, in terms of relative contribution to social IOR, these results show that attention orienting (to a location) plays a greater role than action observation processes.

Although the results of Welsh et al. (2009) suggest that action observation mechanisms play a significant role in the social IOR effect, the means for the three conditions of interest can also be explained solely by the transient hypothesis. A well-established effect within attention capture/inhibition work is that RTs for targets that occur at a cued location are longer than RTs for targets presented in the *opposite hemifield* (e.g., Berlucchi, Tassinari, Marzi, & Stefano, 1989; Collie, Maruff, Yucel, Danckert, & Currie, 2000; Tassinari, Aglioti, Chelazzi, Marzi, & Berlucchi, 1987). Indeed, this is almost the paradigm definition of IOR, since the uncued location in the classic IOR paradigm is usually in the opposite field. With respect to the experiment of Welsh et al. (2009), in the conditions where a participant performs a different movement to a different location, the cue (i.e., arm movement) is always in the opposite hemifield to that of the target. This should result in relatively short RTs. In contrast, where a participant performs the same movement to a different location, the cue appears in the same hemifield as the target on half of the trials and in the opposite hemi-field on the other half. Because the cue-target hemifield positions are now equated, this should result in significantly longer RTs relative to the condition just described; as was found by Welsh et al. Finally, in the conditions where a participant performs a different movement to the same location, the cue always appears in the same hemifield as the target, indeed the same place. This scenario should result in the longest RTs of all; exactly what Welsh et al. found. In fact, the Welsh et al. experiment could be conceived as a classic Posner and Cohen cueing paradigm in which the (hemifield) location of the cue and target are systematically manipulated such that they generate cued, uncued, and control (i.e., both cued and uncued) conditions.

In sum, although the Welsh et al. (2009) experiment suggests that both action and location inhibition occurs in social IOR, the latter may provide the most parsimonious account of the phenomenon. Clearly, additional experimentation will be needed to tease these issues apart.

Correlations of social IOR with non-social IOR

A number of studies have examined the degree to which within-participant IOR relates to social IOR. The explicitly stated rationale has been that if social IOR is indeed due to IOR mechanisms there should be a good correlation between the two. Three studies have now employed this rationale. One might add that a significant association between within- and between-participant IOR is a necessary condition of the transient hypothesis. In the Welsh et al. (2009) experiment, in which both co-actors faced the same way (see above), results showed high within- and between-participant IOR correlations for both repeated movement endpoint trials ($r = .67$) and repeated movement direction trials ($r = .73$). Doneva et al. (2015) also found a medium-to-large correlation between the degree to which a central wooden arrow (and peripheral cues) could induce IOR and social IOR ($r = .49$). Only one published study has examined the correlation between IOR induced by the classic Posner and Cohen paradigm and social IOR. Atkinson et al. (2014) observed a correlation of $r = .37$. These analyses do therefore suggest that social IOR is indeed an IOR phenomenon, and IOR is closely related to attention capture. Of course, one could make the additional argument that the above correlation values show that not all the variance is explained by IOR. In other words, additional processes must explain the rest of the variance; these could include individual differences in action observation processes.

Foraging facilitation, habituation, and the gambler's fallacy. Our review has concentrated on the three main accounts of social IOR. Furthermore, these are explanations that the available empirical data can address. There are however other theories that have been posited. As with the vast majority of cognitive psychology, the three main accounts of social IOR are concerned with immediate mechanisms and processes. A 'distal' evolutionary-based explanation has however also been posited. This explanation can be seen as accompanying the transient and co-representation accounts but, if true, would not support the movement congruency account. Following Klien and Macinnes (1999), standard (i.e., non-social) IOR is often said to be a foraging facilitator. This account

suggests that natural selection has favoured the behaviour because it facilitates visual search. Specifically, the mechanisms that have developed through evolution work to inhibit the orienting of attention and saccades to locations that have just been searched/attended. This 'functional' explanation has also been applied to social IOR by many authors (e.g., Janczyk et al., 2016). Indeed, it can be argued that the inhibition effect should be particularly pronounced when visual search occurs with cooperating pairs of individuals given that our Pleistocene ancestors are likely to have searched in groups (see Binford, 1986).

A number of predictions follow from the foraging facilitator hypothesis, and a few published social IOR studies do allow an assessment of the theory. A central prediction is that simply *knowing* where a co-actor has reached to (i.e., searched) should induce the effect. When individuals are searching an environment it is clearly beneficial to know where another person has recently searched; a system that requires a person to actually see where another person has searched is not particularly efficient. Only one published report has examined whether knowing where another person has searched induces social IOR. Welsh, et al. (2014) obscured all visible partner responses via the use of goggles, but co-actors received auditory information (a high or low tone in one study, and "blue" or "green" indicating the colour of the target in a second study) that informed them as to where their co-actor had just reached to. Results showed no social IOR effect in this scenario. This does not therefore support the foraging facilitator hypothesis. We will add that one has to remember what it actually means for IOR, social or non-social, to have evolved via Darwinian selection - A single person was born during the Pleistocene period who had a genetic propensity for being approximately 25 ms slower to respond to stimuli appearing at a location where their attention just was; this propensity provided the individual with an advantage over individuals and that this advantage enabled them to survive and to have more children who in turn passed on this genetic propensity.

Finally, it may also be that habituation and/or an erroneous reasoning process, known as the gambler's fallacy, may also play a role in social IOR. Because the former has occasionally been suggested as an explanation for the basic (i.e., non-social) IOR effect (see Dukewich, 2009), it follows that the same explanation may account for social IOR. Habituation is often described as a mechanism that redistributes processing away from repetitive events (e.g., Stephenson & Siddle, 1983) and as Dukewich (2009) pointed out, this is similar to Posner, Rafal, Choate, and Vaughan's (1985) early account of IOR (i.e., "novelty seeking"). However, little work has been undertaken on whether this explains IOR and none on whether it explains, or plays any role in social IOR. One prediction that does follow from the habituation account is that the size of the social IOR effect should be related to the number of repeated target presentations at the same location. That is, the greater the number of repeats the greater the habituation. Behaviourally however, this explanation of social IOR will look no different to another possible explanation of social IOR, that is, the gambler's fallacy. Gambler's fallacy is the erroneous misconception that the probability of a truly random event occurring decreases if it occurred recently; for example, believing that a coin toss will result in a Head if there has just been a run of Tails. There is some evidence that IOR is related to the gambler's fallacy. Lyons et al. (2003) showed that observers who demonstrated particularly pronounced (arm reaching) IOR were more likely to switch choices on a betting paradigm (following a win). Indeed, if one considers the basic social IOR paradigm, participants may well believe that the target location is more likely to switch (than stay) following a run of targets being presented at the same location. However, Cole et al. (2015) found that when given a free choice of response location (i.e., the participant decides which side to reach), co-actors are less likely to reach to the location just reached to by their partner. This suggests that neither the gambler's fallacy nor habituation explains social IOR. Future research may want to more directly target these alternative accounts.

Conclusions

In this article we have reviewed work on the social IOR joint action phenomenon with the aim of assessing why the effect occurs. In particular, we have examined whether it is likely to be due to action observation mechanisms or due to the detection of motion and/or other visual transients. The movement congruency account (Ondobaka et al. 2012) can be confidently discounted. The effect still occurs when co-actors perform different actions, and its time course is likely to be too short to allow imitation to occur. Furthermore, the correlations between standard individual IOR and social IOR are also evidence against this account. Moreover, when co-actors sit side-by-side and face the same way, a congruent movement is inhibited rather than facilitated. Because it involves both action inhibition and location inhibition, the co-representation account (Welsh et al., 2005) is perhaps the most difficult to assess, or at least distinguish from the transient account, since the former requires an experiment that shows both components operating simultaneously. Despite this, the study of Welsh et al. (2009) provides evidence that both types of inhibition occur in social IOR. Our review has however shown that transient processing, and subsequent IOR, is the most parsimonious account of the phenomenon (Cole et al., 2012). As long as an observer's attention is shifted to a location, IOR associated with stimuli at that location will be generated. This account can also explain the data observed by Welsh et al. (2009) that appear to show both action and location inhibition operate in social IOR. However, even in that experiment, the influence of transients is larger than the influence of action observation.

More broadly, we suggest that work on action observation and joint action should begin to consider the fact that when an observer sees another person perform a body movement, this is a transient event that is likely to shift the observer's attention. This orienting may well contribute to, or even explain in its entirety, the phenomenon under examination. For example, Constable et al. (2017) showed that a motor contagion-like effect only emerged when the observer followed the limb of a co-actor with their eyes. When the participant centrally-fixated and did not follow the limb, no motor contagion-like deviations in movement execution emerged suggesting that, at minimum, the focus of attention on the moving limb and associated eye movements enhanced the

motor contagion effect in the limb movements. Attention to an object or person can also act as a mediating mechanism that leads to a further process, as was shown by Dolk et al. (2013) with respect to the joint Simon effect and 'reference codes', i.e., the spatial coding of one's action relative to other attention capturing events (see Hommel, 1996). Attention may therefore assist in the initiation of action observation representations. Thus, transient motion and action observation need not be mutually exclusive. Indeed, the combining of attention orienting and action observation processes is explicit in the 'gaze imitation hypothesis' (see Atkinson, Simpson & Cole, 2018; Mansfield, Farroni, & Johnson, 2003). This hypothesis posits that observing eye gaze generates an oculomotor program in the observer, which subsequently induces the same gaze behaviour. If attention orienting and action observation mechanisms do both play a role in joint action effects, future work may want to examine the relative contribution that each makes. A related line of work could assess whether body movements are particularly effective at attracting attention when compared to other motion signals.

Of course, it is not the case that every body action we are exposed to attracts our attention. For instance, a person adjusting their spectacles is not likely to be relevant to an observer and there is a long and ongoing debate as to whether transient events attract attention in a purely bottom up manner (e.g., Folk, Remington, & Johnson, 1992; Theeuwes, Atchley, & Kramer, 2000). One does also have to acknowledge that an attention shift need not necessarily lead to awareness of the event (see Lamme, 2003 for discussion). However, when one considers the impoverished nature of the stimuli in the standard action observation experiment, it is likely that the salient body movement will shift attention. One also has to note that attention shifts induced by biological motion, or any motion, do not always lead to inhibition, as has often been shown with gaze following.

References

Atkinson, M., Skarratt, P.A., Simpson, A., & Cole, G. G (2014). Is social IOR due to action co-

representation? *Acta Psychologica*, 150, 85-93.

Atkinson, M., Simpson, A., & Cole, G. G. (2018). Visual attention and action: How cueing, direct mapping and social interactions drive orienting. *Psychonomic Bulletin and Review*, 25, 1585-1605.

Bekkering, H., Wohlschläger, A., & Gattis, G. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 53(A), 153-164.

Bennett, P. J., & Pratt, J. (2001). The spatial distribution of inhibition of return. *Psychological Science*, 12, 76-80.

Berlucchi, G., Tassinari, G., Marzi, C.A., & Di Stefano, M. (1989). Spatial distribution of the inhibitory effect of peripheral non-informative cues on simple reaction time to non-fixated visual targets. *Neuropsychologia*, 27, 201-221.

Binford, L. (1986). Human ancestors: Changing views of their behavior. *Journal of Anthropological Archaeology*, 3 235-57.

Böckler, A., van der Wel, R.P., & Welsh, T.N. (2014). Catching eyes: Effects of social and non-social cues on attention capture. *Psychological Science*, 25, 720-727.

Bouquet, C. A., Shipley, T. F., Capa, R. L., & Marshall, P. J. (2011). Motor contagion: Goal-directed actions are more contagious than non-goal-directed actions. *Experimental Psychology*, 58, 71-78.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and

- executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, 44, 124-143.
- Cole, G. G., Skarratt, P.A., & Billing R-C. (2012). Do action goals mediate social inhibition of return? *Psychological Research*, 76, 736-746.
- Cole, G. G., Wright, D., Doneva, S. & Skarratt, P. A. (2015). When your decisions are not (quite) your own: Action observation influences free choices. *PLoS ONE* 10(5): e0127766.
- Cole, G. G., & Kuhn, G. (2010). What the experimenter's prime tells the observer's brain. *Attention, Perception & Psychophysics*, 72, 1367-1376.
- Cole, G. G., Skarratt, P. A. & Kuhn, G. (2016). Real-person interaction in visual attention research. *European Psychologist*, 21, 141-149.
- Cole, G. G., Atkinson, M., D'Souza, A., Welsh, T. & Skarratt, P. A. (2018). Are goal states represented during kinematic imitation? *Journal of Experimental Psychology: Human Perception and Performance*, 44, 226-242.
- Cole, G. G., Kentridge, R. W., & Heywood, C. A. (2005). Object onset and parvocellular guidance of attentional allocation. *Psychological Science*, 16, 270-274.
- Collie, A., Maruff, P., Yucel, M., Currie, J., & Danckert, J. (2000). Spatiotemporal distribution of facilitation and inhibition of return arising from the reflexive orienting of covert attention. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1733–1745.
- Constable, M.D., de Grosbois, J., Lung, T., Tremblay, L., Pratt, J. & Welsh, T.N. (2017). Eye movements may cause motor contagion effects. *Psychonomic Bulletin & Review*, 24, 835-841.

- Cook, J., & Bird, G. (2011). Social attitudes differentially modulate imitation in adolescents and adults. *Experimental Brain Research*, *211*, 601–612.
- Cowper-Smith, C. D., Eskes, G., & Westwood, D. (2013). Motor inhibition of return can affect prepared reaching movements. *Neuroscience letters*, *541*, 83–86.
- Dalmaso, M., Castelli, L., Scatturin, P., Carli, L., Todisco, P., Palomba, D., & Galfano, G. (2016). Altered social attention in anorexia nervosa during real social interaction. *Scientific Reports*, *6*, 1–11.
- Diaz-Tula, A., Morimoto, C. H., & Ranvaud, R. D. (2015). A mathematical model of saccadic reaction time as a function of the fixation point brightness gain. *Attention, Perception, & Psychophysics* *77*, 2153-2165
- Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2013). The (not so) social simon effect: A referential coding account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1248-1260.
- Doneva, S., & Cole, G. G. (2014). The role of attention in a joint-action effect. *PLOS One*, *9*, 3, e91336.
- Doneva, S., Atkinson, M. Skarratt, P. & Cole, G. G. (2017). Action or attention in social inhibition of return? *Psychological Research*, *81*, 143-54.
- Dukewich, K. R. (2009). Reconceptualizing inhibition of return as habituation of the orienting Response. *Psychonomic Bulletin and Review*, *16*, 238-251.
- di Pellegrino, G., Fadiga, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176-180.

- Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*, 161-177.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1030-1044.
- Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, *65*, 999-1010.
- Frischen, A., & Tipper, S. P. (2006). Long-term gaze cueing effects: Evidence for retrieval of prior states of attention from memory. *Visual Cognition*, *14*, 351-364.
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze-cueing of attention: Visual attention, social cognition and individual differences. *Psychological Bulletin*, *133*(4), 694-724.
- Fuller, S., Park, Y., & Carrasco, M. (2009). Cue contrast modulates the effects of exogenous attention on appearance. *Vision Research*, *49*, 1825–1837.
- Galantucci, B., & Sebanz, N. (2009). Joint action: Current perspectives. *Topics in Cognitive Science*, *1*, 255–259.
- Gattis, M., Bekkering, H., & Wohlschläger, A. (2002). Goal-directed imitation. In A. N. Meltzoff & W. Prinz (Eds.), *The imitative mind* (pp. 183–205). Cambridge, England: Cambridge University Press.
- Gazzola, V. V., Rizzolatti, G. G., Wicker, B. B., & Keysers, C. C. (2007). The anthropomorphic brain:

- The mirror neuron system responds to human and robotic actions. *Neuroimage*, 35, 1674–1684.
- Gellatly, A.R.H., & Cole, G. G. (2000). Accuracy of target detection in new object and old object displays. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 889–899.
- Hayes, S. J., Hansen, S., & Elliot, D. (2010). Between-person effects on attention and action: Joe and Fred revisited. *Psychological Research*, 74, 302–312.
- Henderson, J. M., & Macquistan, A. D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics*, 53, 221–230.
- Hollander, A., Jung, C., and Prinz, W. (2011). Covert motor activity on NoGo trials in a task sharing paradigm: evidence from the lateralized readiness potential. *Experimental Brain Research*, 211, 345–356.
- Hommel, B. (1996). No prevalence of right-left over above-below spatial codes. *Perception & Psychophysics*, 58, 102–110.
- Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychological Research*, 73, 512–526.
- Jeannerod, M., & Frak, V. (1999). Mental imaging of motor activity in humans. *Current Opinion in Neurobiology*, 9, 735–739.
- Janczyk, M., Welsh, T. M., & Dolk, T. (2016). A role for goals for social inhibition of return? *The Quarterly Journal of Experimental Psychology*, 69, 2402–2418.

- Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception and Psychophysics*, *43*, 346-354.
- Kean, M. & Lambert, A. (2003). The Influence of a salience distinction between bilateral cues on the latency of target-detection saccades. *British Journal of Psychology*, *94*, 373-388.
- Kilner, J., Paulignan, Y., & Blakemore, S. (2003). An interference effect of observed biological movement on action. *Current Biology*, *13* (6), 522–525.
- Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a Foraging facilitator in visual search. *Psychological Science*, *10*, 346–352.
- Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, *7*, 12-18.
- Lestou, V., Pollick, F., & Kourtzi, Z. (2008). Neural substrates for action understanding at different description levels in the human brain. *Journal of Cognitive Neuroscience*, *20*, 324–341.
- Liepelt, R., von Cramon, D. Y., & Brass, M. (2008). What is matched in direct matching? Intention attribution modulates motor priming. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 578-591.
- Lin, J. Y., Franconeri, S. L., & Enns, J. T. (2008). Objects on a collision path with the observer demand attention. *Psychological Science*, *19*, 686-692.
- Longo, M. R., Kosobud, A., & Bertenthal, B. I. (2008). Automatic imitation of biomechanically possible and impossible actions: Effects of priming movements versus goals. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 489–501.

Lyons, J., Weeks, D. J., & Elliott, D. (2013). The Gambler's Fallacy: A Basic Inhibitory Process?

Frontiers in Psychology, 4, 72.

MacKay, D.G. (1987). Self-inhibition, perceptual feedback and error detection. In D. G. MacKay, (ed.)

(1987). *The organization of perception and action: A theory for language and other cognitive skills* (1-254). Berlin: Springer-Verlag.

Mansfield, E., Farroni, T., & Johnson, M. (2003). Does gaze perception facilitate overt orienting?

Visual Cognition, 10, 7–14.

Manzone, J., Cole, G., G., Skarratt, P.A., & Welsh, T.N. (2017). Response-specific effects in a joint

action task: social inhibition of return effects do not emerge when observed and executed actions are different. *Psychological Research, 81*, 1059-1071.

Meltzoff, A. N., & Moore, M. K. (1994). Imitation, memory, and the representation of persons. *Infant*

Behavior & Development, 17, 83–99.

Neyedli, H., & Welsh T.N. (2012). The processes of facilitation and inhibition in a cue-target

paradigm: Insight from movement trajectory deviations. *Acta Psychologica, 139*, 159-165.

Okamoto-Barth, S. and N. Kawai (2006). The role of attention in the facilitation effect and another

“inhibition of return.” *Cognition, 101*, B42–B50.

Ondobaka, S., de Lange, F. P., Newman-Norlund, R. D., Wiemers, M., & Bekkering, H. (2012).

Interplay between action and movement intentions during social interaction. *Psychological Science, 23*, 30-35.

Ondobaka, S., Newman-Norlund, R.D., de Lange F.P., Bekkering, H. (2013). Action recognition

- depends on observer's level of action control and social personality traits. *PLoS ONE*, *8*, 11, e81392.
- Pierce, A. M., Crouse, M. D., and Green, J. J. (2017). Evidence for an attentional component of inhibition of return in visual search. *Psychophysiology*, *54*, 1676-1685.
- Posner, M. I. (1980) Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3-25.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X* (pp. 531-556). Hillsdale, NJ: Erlbaum.
- Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, *2*, 211-228.
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129-154.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
- Rizzolatti, G., Fabbri-Destro, M., & Cattaneo, L. (2008). Mirror neurons and their clinical relevance, *Nature Clinical Practice Neurology*, *5*, 24-34.
- Rizzolatti, G., Riggio, L., & Sheliga, B.M. (1994). Space and selective attention. In Umiltà, C., Moscovitch, M. (Eds.), *Attention and performance XV*. Cambridge, MA: MIT Press.
- Rogers, S. J., Young, G. S., Cook, I., Giolzetti, A., & Ozonoff, S. (2008). Deferred and immediate imitation in regressive and early onset autism. *Journal of Child Psychology and Psychiatry*, *9*, 449–457.

- Sebanz, N. & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science, 1*, 353-367.
- Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition, 88*, 11-21.
- Skarratt, P. A., Cole, G. G., Kingstone, A. (2010). Social inhibition of return. *Acta Psychologica, 134*, 48-54.
- Skarratt, P. A., Gellatly, A. R. H., Cole, G. G., Pilling, M. & Hulleman, J. (2014). Looming motion primes the visuomotor system. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 566-579.
- Skarratt, P. A., Kuhn, G., & Cole, G. G. (2012). Visual cognition during real social interaction. *Frontiers in Human Neuroscience, 6*, 196.
- Steinman, B. A., Steinman, S. B., & Lehmkuhle, S. (1995). Visual attention mechanisms show a center-surround organization. *Vision Research, 35*, 1859-1869.
- Stephenson, D., & Siddle, D. (1983). Theories of habituation. In D. Siddle (Ed.), *Orienting and habituation: Perspectives in human research* (pp. 183-236). Chichester, U.K.: Wiley.
- Tassinari, G., Aglioti, S., Chelazzi, L., Marzi, C.A., & Berlucchi, G. (1987). Distribution in the visual field of the costs of voluntarily allocated attention and the inhibitory after effects of covert orienting. *Neuropsychologia, 25*, 55-72.
- Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics, 49*, 83-90.

- Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). On the time course of top-down and bottom-up control of visual attention. In S. Monsell & J. Driver (Eds.). *Attention and performance XVII: Control of Cognitive processes* (pp. 105–125). Cambridge, MA: MIT Press.
- Tipper, S. P., Grison, S., & Kessler, K. (2003). Long-term inhibition of return of attention. *Psychological Science, 14*, 19-25.
- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing. A neurophysiological study. *Neuron, 31*, 155–165.
- Welsh, T.N., Manzone, J., & McDougall, L.M. (2014). Knowledge of response location alone is not sufficient to generate social inhibition of return. *Acta Psychologica, 153*, 153-159.
- Welsh, T. N., Elliot, D., Anson, J. G., Dhillon, V., Weeks, D. J., Lyons, J. L., & Chua, R. (2005). Does Joe influence Fred's actions? Inhibition of return across different nervous systems. *Neuroscience Letters, 385*, 99-104.
- Welsh, T. N., Lyons, J., Weeks, D. J., Anson, J. G., Chua, R., Mendoza, J., & Elliot, D. (2007). Within- and between-person inhibition of return: Observation is as good as performance. *Psychonomic Bulletin & Review, 14*, 960-956.
- Welsh, T. N., McDougall, L. M., & Weeks D. J. (2009). The performance and observation of action shape future behaviour. *Brain and Cognition 71*, 64–71.
- Welsh, T. N., Ray, M. C., Weeks, D. J., Dewey, D., & Elliott, D. (2009). Does Joe influence Fred's action? Not if Fred has autism spectrum disorder. *Brain Research, 1248*, 141-148.
- Welsh, T.N., & Weeks, D.J. (2010). Visual selective attention & action. In D. Elliott and M.A. Khan

(Eds.), *Vision and Goal-Directed Movement: Neurobehavioral Perspectives*. (pp.39-58).

Human Kinetics, Champaign Illinois.

Wohlschläger, A., & Bekkering, H. (2002). Is human imitation based on a mirror-neurone system?

Some behavioural evidence. *Experimental Brain Research*, *143*, 335–341.

Yantis, S., & Gibson, B. S. (1994). Object continuity in apparent motion and attention. *Canadian*

Journal of Psychology, *48*, 182-204.

Yantis, S., & Hillstrom, A.P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant

visual objects. *Journal of Experimental Psychology: Human Perception and Performance*, *20*,

95-107.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual

search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 601-621.