

# **Real person interaction in visual attention research**

Geoff G. Cole<sup>1</sup>, Paul A. Skarratt<sup>2</sup> and Gustav Kuhn<sup>3</sup>

1. Centre for Brain Science, University of Essex.
2. Department of Psychology, University of Hull.
3. Department of Psychology, Goldsmiths, University of London.

**Running head: Ecological validity and visual cognition**

**Key words:** Visual cognition– Cognitive ethology – Social attention - Ecological validity

**Submitted:** 12/09/2014

**Word count:** 5419

**Address for correspondence**

Gustav Kuhn

Department of Psychology

New Cross

London SE14 6NW

**Abstract**

An important development in cognitive psychology in the past decade has been the examination of visual attention during real social interaction. This contrasts traditional laboratory studies of attention, including ‘social attention’, in which observers perform tasks alone. In this review we show that although the lone-observer method has been central to attention research, real-person interaction paradigms have not only uncovered the processes that occur during ‘joint attention’ but have also revealed attentional processes previously thought not to occur. Furthermore, the examination of some visual attention processes almost invariably requires the use of real-person paradigms. Whilst we do not argue for an increase in ‘ecological validity’ for its own sake, we do suggest that research using real person interaction has greatly benefited the development of visual attention theories.

## **Introduction**

During the 1930s a movement, developed by a number of naturalists, began in Europe. *Ethology* was concerned with the biological basis of nonhuman animal behaviour and examined mechanisms responsible for innate ‘routines’ and ‘patterns’. A major aspect of

ethology was an emphasis on the importance of the method used to understand behaviour. The proponents of the new field, which emerged from Austria, Germany, and the Netherlands, argued that the study of animal behaviour should consider the natural environment in which behaviour occurs. This notion explicitly contrasted with the stimulus-response methods undertaken by the American behaviourists in which models of behaviour were generated from laboratory based studies of animals, usually rats and pigeons. Although many fundamental aspects of behaviour were derived from the behaviourist approach, the field was often criticized for its lack of ecological validity. Whilst behaviourism waned in the 1950s and 60s, as a result of the so-called cognitive revolution, ethology grew. This culminated in the joint awarding of a Nobel prize for three of its leading figures: Konrad Lorenz and Niko Tinbergen for largely developing ethology and Karl von Frisch for discovering the ‘dance language’ of the bee. Ethology was enormously influential and has provided the basis for many areas of behavioural science such as developmental psychology, behavioural ecology (i.e., gene selection) and evolutionary psychology.

Within the last 10 years a larger consideration of the environment in which behaviour occurs has begun to influence another field of psychology, namely visual attention. Attention is the process by which some aspects of the environment are prioritised for further processing at the expense of other aspects. A number of metaphors have been suggested in an attempt to best characterise the way in which this occurs including the attentional spotlight (Broadbent, 1982; Posner, Snyder, & Davidson, 1980) and zoom-lens model (Eriksen & St. James, 1986). As with many other cognitive phenomena (e.g., memory), visual cognition has been primarily based on experiments in which a single observer performs a required task. Furthermore, this lone-observer method has helped to reveal many fundamental aspects of visual cognition (e.g., Duncan, 1984; Eriksen & Eriksen, 1974; Posner, 1980; Raymond, Shapiro, & Arnell, 1992; Simons & Rensink, 2005; Tipper, 1985; Watson & Humphreys, 1997). However, a number of authors have argued that the single observer method fails to take into account the fact that humans are social animals and usually spend some part of each day interacting with others. The charge, in other words, is that attention researchers have neglected one important

aspect of the ecology in which attention operates; when a person is engaged in social interaction.

In this paper we discuss the potential importance of real-person interaction paradigms in terms of studying visual attention. However, importantly, and unlike many others before us, we do not advocate the use of real-world situations and ‘ecological validity’ from a principled point of view. Indeed, most would agree that as behavioural scientists our primary concern should be with processes and mechanisms rather than where they are studied. Instead, we make the point that some fundamental processes associated with some well-established attention phenomena may need to be revised based on the findings from real-person interaction studies, and we present inhibition of return as a test case for this. We then review other studies showing how real and depicted representations of social stimuli have produced fundamentally differing results. We also argue that the understanding of the processes associated with one particular broad phenomenon (i.e., ‘joint action’) almost invariably requires the use of real-person paradigms.

### **Visual attention during real social interaction**

The importance of examining cognitive processes during social interaction is, of course, concerned with the broader question of ecological validity in science; nor is it a new one, predating both ethology and psychology. In the early days of the Enlightenment, a group calling themselves the ‘everyday chemistry movement’ argued that chemistry should consider its everyday settings in addition to the laboratory (see Banaji & Crowder, 1989). Similarly during the 1980s proponents of ‘everyday memory’ highlighted the applicability of memory research and posed questions relevant to practical settings (e.g., Gruneberg, Morris, & Sykes, 1978). An influential ecological approach was also applied to visual perception by Gibson (e.g., Gibson, 1979). He emphasised the role of motion in perception because most percepts are often generated during self-locomotion, where an individual is interacting with the environment. More recently, a number of authors have also pointed out that visual perception tends to occur in conjunction with action, and paradigms have been developed that examine this (see Hommel, 2009; Jeannerod, 1999; Milner & Goodale, 1995; Prinz, 1997). For

instance, Welsh and Pratt (2006) showed that the type of planned action a person makes to a visual target will determine the degree to which it is likely to attract attention. The authors showed that when a standard keyboard press is required, a task-irrelevant offsetting stimulus will reduce the effectiveness with which a new target object captures attention, but not when a full reaching response is made. Thus, vision-for-action researchers have argued that vision is better understood when its functional significance in the real world is taken into account (see also Milner & Goodale, 1995). However, the area that most explicitly advocates real-world processes is the developing field of *cognitive ethology*. Kingstone, Smilek, Birmingham, Cameron and Bischof (2005), as central proponents of the field, argue that traditional attention research is based on two assumptions; that attentional processes are stable across a variety of situations and the explanatory power of a process can be maximised by reducing all variability in a situation except for the process being examined. The authors argue that the assumption of stability means that scientists do not feel obliged to examine whether the processes they assess in the laboratory translate to the real world. Kingstone, Smilek, and Eastwood (2008) additionally argued that phenomena should ideally be studied as they occur in the real-world environment *before* examining them under the controlled situation of the laboratory. As Chisolm, Chapman, Amm, Bischof, and Smilek (2014) stated, “By starting at the natural level, one’s subsequent investigations are grounded in cognition and performance as it occurs in real life, and hence, through comparison between life and lab, one can determine which lab-based findings are likely to scale up to a natural environment and which findings are specific to a controlled research environment.” (Chisolm, et al. 2014).

That attention can be modulated by social factors has been acknowledged for many decades. The early eye tracking studies of Yarbus (1967) showed that a person located in a visual scene is particularly well attended. Furthermore, the field of ‘social attention’ has become a major area within attention and is based on the premise that attention often takes place in a social setting. Although social attention by definition emphasises social context, experimental paradigms have not typically assessed attention during real social interactions. Experiments usually present participants with schematic representations of social stimuli. For instance, an abundance of studies have now demonstrated the importance of eyes in face

perception (e.g., Birmingham, Bischof, & Kingstone, 2008b, 2009; Fletcher-Watson, Findlay, Leekam, & Benson, 2008; Kuhn & Land, 2006; Kuhn, Tatler, & Cole, 2009; 2002; Walker-Smith, Gale, & Findlay, 1977). However, the vast majority of such studies present faces in isolation with as little context as possible. Consequently, the face is likely to be already attended. Thus, these findings may only demonstrate eye preference because they are the most complex or salient aspect of an impoverished pre-selected stimulus. Some authors have attempted to improve the ecological validity of social stimuli by using photographs or video clips of real world interactions. For instance, Kuhn, et al. (2009) examined visual attention by showing a magician performing a magic trick and measuring where in the scene participants were most likely to look. Similarly, Foulsham, Cheng, Tracy, Henrich, and Kingstone (2010) assessed eye movements when participants were asked to watch videos of students engaging in conversation. Although these studies present more realistic social stimuli, their use of filmed rather than real social interactions must still question whether the findings accurately measure the cognitive processes underlying real world social behaviour.

In contrast to purely *depicted* representations of social scenes, a number of authors have begun to examine cognition, and more specifically attention, whilst individuals interact with other people. Examining visual cognition when others are present is not in itself a new phenomenon. Solomon Asch's work on conformity is perhaps the best known example. In one of his classic experiments, participants were required to match a sample line with one of three test lines with respect to length. The central results showed that judgements were influenced by what other (confederate) individuals thought, often concurring with their selection even when the sample and test lines were markedly different. Although one might assume that the *perception* of the lines themselves was not changed, recent neuroimaging work has provided evidence that social influence can modulate perception. Berns, Chappelow, Zink, Pagnoni, Martin-Skurski, and Richards (2005) imaged observers whilst they performed a mental rotation task during which they saw the responses of other 'participants' (i.e., confederates). The central results showed that incorrect responses resulting from the influence of the group was associated with functional changes in occipital-

parietal networks responsible for perception. The authors thus argued that perception of the stimuli had been modulated by social influence.

One of the basic tasks concerning attention and its oculomotor mechanisms is visual search, in which an individual searches an environment for a target item (e.g., set of keys). Brennan, Chen, Dickinson, Neider and Zelinsky (2008) made the point that many everyday situations involve performing this type of task with another individual, such as when a child and adult look through a picture book together. A number of studies have begun to examine the efficiency of search when an individual has knowledge of where a co-searcher is looking (known as *gaze transfer*). In their experiment, Brennan et al. (2008) asked pairs of participants to search displays for a specific letter that would appear amongst distractors. Cursors allowed each participant to see where their partner was gazing at any moment during the search. Brennan et al. found that that search was almost twice as efficient when a participant was able to use another person's search behaviour, in contrast to searching alone. In a follow-up study, Neider, Chen, Dickinson, Brennan and Zelinsky (2010) compared gaze and speech transfer when two participants jointly searched for a gunman in a virtual scene and were required to agree on the target's location in order to complete a trial. Results showed that RTs were shorter for one participant to locate targets once it had been located by their co-actor.

Gaze transfer effects have also been shown in applied settings. Litchfield, Ball, Donovan, Manning, and Crawford (2010) assessed whether co-searching improved identification of pulmonary nodules in chest x-ray inspection amongst novice radiographers. Results showed this was indeed the case: participants were more efficient in identifying nodules when shown the search behaviour of another individual. Furthermore, many authors now argue that knowledge of an expert's search behaviour can be useful in training novices (Nalanagula, Greenstein, & Gramopadhye, 2006; Sadasivian, Greenstein, Gramopadhye, & Duchowski, 2005). Other examples of gaze transfer include increased fault detection rates during aircraft (Sadasivian et al., 2005) and circuitry board inspection (Nalanagula et al., 2006), and better puzzle problem solving (Velichkovsky, 1995). Moreover, gaze transfer has

been shown in a paradigm where transfer occurs by direct observation of another person's gaze (seated opposite) rather than via gaze cursors (e.g., Macdonald & Tatler, 2013). Outstanding questions that follow from gaze transfer research are whether it improves search efficiency because it represents the collaborative search of another agent, or simply because it provides a spatial cue that can be used to guide search behaviour. Evidence for the former would require search improvements beyond that conferred by the pseudorandom movements of a nonagent cursor. Relatedly, the mediating mechanisms are not yet understood. For instance, gaze cursors may only exert their influence through attentional capture processes (e.g., Cole & Kuhn, 2010a; Yantis & Jonides, 1984) or may involve Theory of Mind processes. Finally, these studies do differ in terms of the degree to which participants are physically interacting with the gazer. For instance, some studies have participants in the same room but do not directly interact (e.g., Muller, Helmert, Pannasch, & Velichkovsky, 2013). Despite these outstanding questions, the above studies provide clear examples of how real-person interaction paradigms have benefitted visual attention work. Furthermore, the gaze transfer method also demonstrates that real-person paradigms do not just improve ecological validity, they also lead to the development of cooperation models.

### **Inhibition of return**

The studies reviewed in the previous section help to reveal aspects of visual attention that occur when we interact with others. Indeed, these real-person interaction paradigms are explicitly designed to examine the mechanisms associated with social attention. However, work carried out using real social interaction may also reveal that some visual attention principles need to be revised as a consequence. A case in point concerns one of the most widely researched visual attention phenomena, inhibition of return (IOR). IOR, first described by Posner and Cohen (1984), refers to the inhibition generated when visuomotor resources are repeatedly directed to a spatial location on successive occasions. In other words, humans are slower to process a stimulus when it appears at a location recently attended. An abundance of work has now examined its various characteristics and the mechanisms giving rise to the effect. For instance, IOR impairs the detection and

identification of objects (e.g., Pratt & Abrams, 1999) for a period of approximately 3000 ms (Samuel & Kat, 2003), and is evident for various response modes including eye saccades and limb movements (Abrams & Dobkin, 1994; Klein & McInnes, 1999). IOR is also known to ‘follow’ objects as they move around the environment as well as being tagged to fixed spatial locations (Abrams & Dobkin, 1994; Tipper, Driver, & Weaver, 1991). Finally, it is cross-modal, having been shown to occur within and between the visual, auditory and tactile domains (Mondor, Breau, & Milliken, 1998; Poliakoff, Spence, O’Boyle, McGlone, & Cody, 2002).

One of the key findings from the IOR literature was that the basic phenomenon was thought to occur only as a result of peripheral ‘exogenous’ attention-capturing events. Shifts of visual attention are often dichotomised as being either exogenous (‘bottom-up’) or ‘endogenous’ (‘top down’). Whereas the former refers to shifts induced by peripheral and task-irrelevant events that ‘capture’ attention automatically, the latter refers to shifts that are under volitional control. Endogenous attention is mostly associated with tasks in which an observer is presented with a central (e.g., arrow or eye gaze) stimulus that indicates the likely position of an upcoming target. (e.g., Friesen, & Kingstone, 1998). It is a shift of attention in response to this type of cue that consistently failed to elicit IOR. Thus, if an arrow cue points to the right, IOR will not be generated on the right. This contrasts with the robust effect seen when attention is shifted by an exogenous peripheral cue. As Weger, Abrams, Law, and Pratt (2008) stated, the absence of IOR following central cues assumed “axiomatic status” from the early days of IOR work. Indeed, some authors used IOR as an indicator that an exogenous attentional shift, as opposed to endogenous, had occurred (Pratt, Hillis, & Gold, 2001; Theeuwes & Godijn, 2002). The few studies that have reported IOR induced by centrally presented cues (Frischen, Smilek, et al., 2007; Frischen & Tipper, 2004; Okamoto-Barth & Kawai, 2006) have done so with specific manipulations of the basic paradigm. For example, by using particularly long intervals between the initial attentional shift and the subsequent target. Furthermore, the fact that only a few studies have shown IOR induced by central cues may suggest that many failed to find such an effect and consequently did not submit the work for publication. However, it is becoming clear from the dozens of experiments published to

date that IOR can be generated robustly and more generically with a central cue if the cue is a real individual that interacts with the participant, an effect termed ‘between-participant’ IOR (Welsh et al., 2005; 2007) or ‘social’ IOR (Skarratt, Cole, & Kingstone, 2010; Skarratt, Cole, & Kuhn, 2012).

In the basic social IOR paradigm, two individuals sit opposite each other across a table. Built within the table is a flat screen monitor used to present stimuli and record responses. Participants take turns (i.e., alternate) to reach out and touch a target that can appear on either the left or right of the workspace. Results show that reaction time to respond to a target is slowed if a participant reaches to the same location to which their co-actor has just reached. For instance, if Participant A is required to respond to the target on their right, Participant B will now be slower to reach to their left (i.e., the same spatial location). The effect does not arise during the arm movement itself, as once the motor programming has taken place the time to execute the response remains the same in all the critical comparisons. Rather, it is the perceptual representation of the target location and/or the preparation of a motoric response to it that is inhibited. Although some authors have posited alternative theories to explain the basic effect (e.g., action congruency mechanisms, where observing a co-actor reach to their right primes an observer to reach to their own right; Ondobaka, Lange, Newman-Norlund., Wiemers, & Bekkering, 2012), most studies suggest the effect is due to mechanisms that give rise to IOR. Specifically, when a co-actor reaches to their left, a corresponding visuomotor response to the same spatial location is elicited in the observer. If on the next trial the observer is required to reach to that location (i.e., on their right), IOR generated by the stimulus will slow their responses.

This effect is placed very much within the context of ‘joint action’ work (see next section), in which a participant’s behaviour is influenced by the actions of another individual. As with the basic (i.e., lone participant) IOR phenomenon, many of its attributes and characteristics have now been uncovered. Within the context of the IOR and central cueing, the important finding is that this social IOR effect can be induced when only the first part of a partner’s arm movement can be seen and acts, to that end, as a central cue. For instance,

Welsh et al. (2007) had participants wear shutter goggles that occluded a participant's view of their co-actors target appearance and most of their arm reach, ensuring only the initial 50ms of their response remained visible. Similarly, Skarratt, et al. (2010) occluded all peripheral stimuli with the use of physical barriers placed between co-actors that restricted the view of each participant to a central aperture. Importantly therefore, this restricted information acted similarly to a central cue stimulus used to orient attention in classical IOR studies. However, in disagreement with the consensus view that IOR does not typically occur in response to central cues, IOR is reliably elicited when the cue is the action of a real person sitting opposite. In sum, an effect long thought not to occur, such that it had reached "axiomatic status" (e.g., Weger et al., 2008), is robustly induced in a real-person interaction paradigm. This illustrates the importance of real social situations for theories of visual attention.

### **Direct comparison of real–person and no real–person paradigms**

There are other cases in which attention processes have been shown to be different when real-person interaction paradigms are employed. For instance, people will often avoid the eye gaze of another person (Ellsworth, Carlsmith, & Henson, 1972) but not the depicted image of a stranger (Henderson, Williams, & Falk, 2005; Itier, et al., 2007). Similarly, pedestrians are more likely to look where another person is looking when walking behind them, as opposed to walking in front (i.e., approaching them; Gallup, Chong, & Couzin, 2012). Laidlaw, Foulsham, Kuhn, and Kingstone (2011) directly compared eye movements during interaction with a real (confederate) person or interaction with the same person that appeared via a video screen. The confederate posed as another research participant waiting in a room. Perhaps unsurprisingly, results showed that whereas participants often gazed towards the confederate presented on the video screen, they rarely did so in person. However, results also showed that participants' level of sociality, as measured by the Autism Quotient Inventory (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) correlated with the amount of time spent gazing at the live confederate. This more subtle effect was not observed in the video condition. Similar conclusions concerning the difference between real and artificial social interactions have been drawn from studies examining eye movements in

response to social cues in autism (e.g., Freeth, Foulsham, & Kingstone, 2013; Nation & Penny, 2008). Skarratt et al. (2010) also directly compared the effects of using a real human versus a depicted human on the social IOR phenomenon described above. In the depicted situation a participant performed the alternate reaching task in conjunction with an avatar who was a life-sized image of a human participant projected onto a screen opposite. Interestingly, results showed that the social IOR phenomenon was found only when participants performed with a real person and not with the projected co-actor. These findings further illustrate the differences that can be observed between real and depicted representations of social interaction.

As suggested in an earlier section, other people's eyes can provide an observer with a variety of valuable information, including where their attention is. This is most often examined with the gaze cueing paradigm in which participants are shown a central face whose eyes and/or head are directed to the left or right (Langton, & Bruce, 1999). A target then appears either on the left or right. Results typically show that stimuli are processed more rapidly when presented in the gazed-at position. This is taken as evidence that an attentional shift has been elicited by the gaze direction. It is often implicitly assumed amongst social attention workers (or explicitly, see Wu, Bischof, & Kingstone, 2014) that other people's gaze acts to signal information present in the environment. Indeed, many authors have also suggested that attentional orienting in response to gaze following may be adaptive, as it could facilitate the detection of potentially threatening stimuli. This has lead to the suggestion that gaze has a special status in terms of attentional cueing. However, one of the central and somewhat surprising findings from the abundance of gaze cueing work is that central arrow cues are equally effective in shifting attention as are eye gaze cues, which challenges the notion that social cues have a special status in attentional orienting. However, a recent study suggests that the standard gaze cueing method has underestimated the size of the gaze following effect. Cole, Smith and Atkinson (2015) repeated the basic procedure but used a real person as the cue (see also, Lachat, et al., 2012). As expected, the central results showed that the stimulus person shifted an observer's attention to the gazed-at location. Importantly, the effect size was approximately three times larger than that elicited using standard

schematic stimuli. Thus the findings suggest that eye gaze is indeed a particularly effective attention cue when a real person is used to orient an observer's attention.

The potential importance of using a real person in the gaze cueing paradigm is also illustrated with respect to a recent debate concerning gaze following and Theory of Mind. Some authors have argued that Theory of Mind mechanisms are activated when an observer's attention is shifted by another person's gaze. Nuku and Bekkering (2008), for instance, argue that the observer infers that the agent is physically able to attend to the target." (Page 340). Indeed, a person's gaze suggests that something is being perceived at the gazed location. As Calder et al. (2002) point out, gaze "implies that the person may have some intention or goal towards this particular object. In other words, gaze engages the mechanisms involved in the attribution of intentions and goals to others..." (Page 1130). However, Langton (2009) has argued that the important attribution may be whether the gazer's perceptual mechanisms are working (or not) rather than their mental state. Furthermore, non-agent objects (e.g., a glove) that incorporate a pair of eyes but evidently have no mental state are still effective in shifting attention to the looked-at direction (e.g., Quadflieg, Mason, & Macrae, 2004). Langton (2009) also points out that the typical gaze cueing procedure presents an isolated face that is not actually looking at anything. The mentalising account contrasts the most common account of gaze cueing in which the effect is said to be due to reflective/bottom-up processes mediated by mechanics of eye deviation perception (e.g., Baron-Cohen, 1995; Driver, Davies, Kidd, Maxwell, Ricciardelli, & Baron-cohen, 1999). Accordingly, attention is said to be shifted from the eyes because they deviate towards the looked-at location. In the context of the present discussion, one can argue that a paradigm that uses a real person as the inducing stimulus is going to be particularly beneficial to a debate concerning mental state attribution. One can assume that mechanisms associated with Theory of Mind will be more easily invoked when interacting with a real person as opposed to a schematic representation.

### **Theories of action co-representation**

Real person interaction studies have been instrumental in developing theories concerned with the co-representation of action and perception. Although a small number of studies have

employed the classic lone observer method (e.g., Liepelt, von Cramon, & Brass, 2008) there is an almost tacit assumption that the mechanisms and processes are best examined when two real individuals interact. Perceptuomotor models argue that rather than being separate, action and perception share cognitive representations (e.g., Hommel, 2009; Jeannerod & Frak, 1999; Knoblich & Sebanz, 2006; Prinz, 1997). It follows that one can influence the other, an effect that has been shown in a number of studies in which observers act alone. For instance, a leftward shift of attention will facilitate a left-handed response (e.g., Cole & Kuhn, 2010b). Furthermore, orientation judgements about everyday objects are facilitated when the way the object is typically used is compatible with the required response (e.g., Ellis & Tucker, 2000). A central prediction made by perceptuomotor models is that the observation of an action performed by *another individual* activates the motor system of the observer. Thus, observing a manual response can influence one's own action, thus giving rise to movement congruency effects. For instance, Liepelt, et al. (2008) presented photographs of a hand that had a target number placed in the middle of the image. Participants were asked to discriminate the target and make a response by raising either their index or middle finger. The critical manipulation was that the hand on the photograph also had either its index or middle finger raised. Results showed that when the target required the index finger to be raised responses were faster if the depicted hand also had the index finger raised. The same effect occurred for the middle finger. Thus a response is facilitated when there is congruency between an observed action and the response required. However, although such effects can be shown with depicted images, theories that link action and perception within a common representation have greatly benefited from real-person interaction paradigms. Indeed, 'joint action' experiments, in which two individuals jointly share a task, have become the major tool in this field.

The most notable example is the Social Simon effect (Sebanz, Knoblich & Prinz, 2003). In the basic Simon task (Simon, 1969), an individual is required to discriminate a target stimulus with a left or right response. The important manipulation is that the target can appear either on the left or right side of the display. Although target location is not task relevant results show that responses are slower when the target location and response location are not compatible. For instance, when the target is on the right and a left response is required. This

effect is abolished when participants perform half of the task in which they are only required to respond to one target with one hand (as in a ‘Go-No-Go’ task). However, when this single response task is performed with another individual sitting adjacent who is required to respond to the presentation of the other target, the effect is reinstated. The Social Simon effect has motivated numerous joint-action studies, the vast majority of which use a real person as the co-actor. For instance, Hommel, Colzato, and van den Wildenberg, (2009, but see Dolk, Hommel, Prinz, & Liepelt, 2013) showed that the effect is sensitive to the valence of the co-actor.

### **Issues of ecological validity?**

Real-person interaction paradigms are often associated with issues of ecological validity and social cognition work has been criticised for its lack of ecological validity (e.g., Birmingham, Bischof, & Kingstone, 2009). However, what level of ecological validity is sufficient when it is considered? One could argue that paradigms that use images of real faces are more ecologically valid than those using schematic faces, but some would argue that real faces need to be embedded within a richer context, such as within a scene (Birmingham, Bischof, & Kingstone, 2008). Movie clips depict the rapid information changes that can underlie human interactions more accurately than do static images, but both may lack the context that may be critical to observing behaviour in the natural environment. Indeed, even studies that attempt to reproduce authentic social interactions, such as those involving watching a magician perform magic tricks (Kuhn & Tatler, 2005), or sitting in a waiting room (Laidlaw, et al., 2011) are artificial. In those studies, the very act of measuring individuals’ eye movements with their full awareness means they are likely to act in a non-natural way (but see Wu, Bischof, & Kingstone, 2013, who used hidden cameras to record eye movements surreptitiously). Ecological validity may be considered a continuum and the potential for real social interactions lies at some point on this continuum. For certain cognitive processes, such as the use of eye to signal information, there is a strong need for real social interaction, but other processes may be less sensitive. It is only by studying

behaviour at these different levels that we can fully understand the true nature of social cognition.

One can also argue that ecological validity is often not a good idea at all. The goal of psychology, and indeed science in general, is to generate general models that can explain many phenomena. This is often contrasted with paradigm-driven approaches concerned with specific phenomena, and specific phenomena is precisely what ecological validity gives the experimenter. Furthermore, simplifying mental processes and studying them in isolation under well-controlled conditions clearly has its advantages. For example, in their seminal work, Heider and Simmel (1944) used simple geometric stimuli to study rather abstract forms of social cognition, and demonstrated that people imbue mentalistic characteristics on simple geometric shapes. These truly non-ecologically valid stimuli have been immensely useful for studying a wide range of social cognitive processes, since it allows researchers to isolate some of the key component. Moreover, some behavioural processes only manifest themselves under non-ecologically valid conditions. For example, brain damaged patients often have very specific cognitive impairments, and can develop strategies that allow them to adapt in the real world. Prosopagnosic patient MS (see Cole, Heywood, Kentridge, Fairholme, & Cowey, 2003) is not able to discriminate between a male and female face on any standard face recognition test, in which hair is typically removed. However, he is considerably better at such discrimination in the real world where he uses hair style as a guide.

Rather than assuming that all experimental designs need to be ecologically valid it may be more useful to consider the specific situations under which ecological validity matters, and the level of abstraction that is acceptable. Research on gaze cueing has involved a wide range of experimental set-ups and stimuli. Some have used schematic faces (Friesen & Kingstone, 1998), others images of real faces (Driver et al., 1999) and some have even used real people (Cole, et al. 2015). However, what is the appropriate level of abstraction? This is an important question because undertaking experiments with a real person as a cue is much more complex than simply presenting images of faces on a computer screen. The level of abstraction depends on the overall research question. For example, Teufel, Alexis, Clayton,

and Davis, (2010) investigated the extent to which gaze cueing is modulated by mental states (see above), such as whether the person can see the target or not. The experimenters used a standard cueing task, but rather than presenting pictures, they showed videos of a real person. Crucially, however, the researchers were interested in how the mental state of seeing influenced gaze cueing, and participants were therefore required to believe that the person depicted in the video was real. The researchers developed an elaborate deceptive procedure that convinced participants that the person was indeed real. According to the authors it was essential that participants believed the person was real, and nothing short of a video clip would have been sufficient. For the purpose of this research question, the level of abstraction was sufficient, and clearly much higher in ecological validity than the standard gaze cueing paradigm. However, as there was no opportunity for a real social interaction, the results may not tell us that much about social cognition within real settings.

### **Concluding remark**

The lone-observer method has, for good reason, been central to research on visual cognition and has helped uncover some of the fundamental aspects of attention, and cognition more generally. In this review we have highlighted how the presence of real people can influence a range of attention processes. Clearly, these findings have important methodological as well as theoretical implications. Methodologically, they give weight to the argument that social attention research should consider the impact that the use of depictions of people in the absence of a real person have on cognitive processes. However, we would not advocate that improvements in ecologically valid are of particular importance; generating models of behaviour is the central concern irrespective of the situation. Rather, we argue, and have shown, that real-person interaction paradigms can lead to an increase in our understanding of attentional processes. One challenge for future experimenters is knowing when a real-person paradigm is required. We posit that such paradigms are only needed when examining issues concerned with social cognition.

### **References**

- Abrams, R.A., & Dobkin, R.S. (1994). Inhibition of return: Effects of attentional cuing on eye movement latencies. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 467-477.
- Banaji, M. R., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist*, 44, 1185-1193.
- Baron-Cohen S. (1995). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In: Moore C, Dunham PJ, editors. *Joint attention: Its origins and role in development*. Hillsdale, NJ: Erlbaum; 41–59.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism Spectrum Quotient (AQ) : Evidence from Asperger Syndrome/High Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, 31, 5-17.
- Berns, G.S., Chappelow, J., Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., & Richards, J. (2005). Neurobiological correlates of social conformity and independence during mental rotation. *Biological Psychiatry*, 58, 245-253.
- Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Social attention and real world scenes: the roles of action, competition, and social content. *Quart J Exp Psych*, 61(7), 986-999.
- Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vis Res*, 49, 2992–3000.
- Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, 106, 1465–1477.
- Broadbent, D.E. (1982). Task combination and selective intake of information. *Acta Psychologica*, 50, 253–290.
- Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., & Young, A.W. (2002). Reading the mind from eye gaze. *Neuropsychologia*, 40, 1129–1138.

- Cole, G. G., Heywood, C. A., Kentridge, R. W., Fairholm, I. & Cowey, A. (2003). Attentional capture by colour and motion in cerebral achromatopsia. *Neuropsychologia*, 41, 1837-1846.
- Cole, G. G., & Kuhn, G. (2010a). Attentional capture by object appearance and disappearance. *Quarterly Journal of Experimental Psychology*, 63, 147-159.
- Cole, G. G., & Kuhn, G. (2010b). What the experimenter's prime tells the observer's brain. *Attention, Perception & Psychophysics*, 72, 1367-1376.
- Cole, G. G., Smith, D, & Atkinson, M. (2015). Mental state attribution and the gaze cueing effect. *Attention, Perception & Psychophysics*, 77, 1105-1115.
- 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.
- Driver, J., Davies, M., Kidd, P., Maxwell, M., Ricciardelli, P., & Baron-cohen, S. (1999). Gaze perception triggers reflective visuospatial orienting. *Vis Cogn*, 6(5), 509-540.
- Duncan, J. (1984). Selective attention and the organization of visual information. *J Exp Psych Gen*, 113, 501-517.
- Ellis, R. & Tucker, M. (2000). Micro-affordance: the potentiation of components of action by seen objects. *British Journal of Psychology*, 91, 451–471.
- Ellsworth, P. C., Carlsmith, J. M., & Henson, A. (1972). The stare as a stimulus to flight in human subjects: A series of field experiments. *J Pers Soc Psychol*, 21, 302-311.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non- search task. *Perc Psych*, 16, 143–149.
- Eriksen, C. W., & St. James, J. D. (1986). Visual-Attention within and around the Field of Focal Attention - a Zoom Lens Model. *Perc Psych*, 40, 225-240.
- Fletcher-Watson, S., Findlay, J. M., Leekam, S. R., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception*, 37, 571-583.

- Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: effects of social status and speaking. *Cognition*, 117, 319-331.
- Freeth M, Foulsham, T. & Kingstone, A. (2013). What affects social attention? Social presence, eye contact and autistic traits. *PLoS One*, 8(1), e53286.
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psych Bull & Rev*, 5, 490–495.
- Frischen, A., Smilek, D., Eastwood, J. D., & Tipper, S. P. (2007). Inhibition of return in response to gaze cues: The roles of time course and fixation cue. *Vis Cogn*, 15, 881-895.
- Frischen, A., & Tipper, S. P. (2004). Orienting attention via observed gaze shift evokes longer term inhibitory effects: Implications for social interactions, attention, and memory. *J Exp Psych Gen*, 133, 516-533.
- Gallup, A.C., Chong, A. & Couzin, I.D. (2012) The directional flow of visual information transfer between pedestrians. *Biology Letters* 8(4), 520-522.
- Gibson, J.J. (1979). The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.
- Gruneberg. M., Morris, P. & Sykes, R. (Eds) (1978) Practical Aspects of Memory. Academic Press. London.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57, 243-259. doi: 10.2307/1416950
- Henderson, J.M., & Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, 33, 98-106.
- Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psych Res*, 73, 512-526.
- Hommel, B., Colzato, L.S., & van den Wildenberg, W.P.M. (2009). How social are task representations? *Psychological Science*, 20, 794-798.
- Itier, R. J., Villate, C., & Ryan, J. D. (2007). Eyes always attract attention but gaze orienting is task-dependent: Evidence from eye movement monitoring. *Neuropsychologia*, 45, 1019-1028.

- Jeannerod, M. (1999). To act or not to act: Perspectives on the representation of actions. *Quarterly Journal of Experimental Psychology*, 52A, 1-29.
- Jeannerod, M., & Frak, V. (1999). Mental imaging of motor activity in humans. *Current opinion in neurobiology*, 9(6), 735-739.
- Klein, R. M., & McInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science*, 10, 790-801.
- Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, 15(3), 99-104.
- Kuhn, G., & Land, M. F. (2006). There's more to magic than meets the eye. *Curr Bio*, 16, R950-R951.
- Kuhn, G., & Tatler, B. W. (2005). Magic and fixation: Now you don't see it, now you do. *Percept*, 34, 1155-1161.
- Kuhn, G., Tatler, B. W., & Cole, G. G. (2009). You look where I look! Effect of gaze cues on overt and covert attention in misdirection. *Vis Cogn*, 17, 925-944.
- Lachat F., Conty L., Hugeville L., George N. (2012). Gaze cueing effect in real face-to-face situation. *Journal of Nonverbal Behaviour*, 36, 177-190.
- Laidlaw, K. E., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *Proc Natl Acad Sci U S A*, 108(14), 5548-5553.
- Langton, S., & Bruce, V. (1999). Reflexive orienting to social attention signals. *Visual Cognition*, 6, 541–568.
- Langton, S. R. H. (2009). Gaze perception: is seeing influenced by believing? *Cur Bio*, 19, R651–R652.
- 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.
- Litchfield, D. , Ball, L. J. , Donovan, T. , Manning, D. J. , & Crawford, T. (2010). Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. *Journal of Experimental Psychology Applied*, 16(3), 251-262.

- Macdonald R. G., Tatler, B. W. (2013). Do as eye say: Gaze cueing and language in a real-world social interaction. *Journal of Vision*, 13 (4), 6, 1-12
- Milner, A. D., & Goodale, M. A. (1995). The visual brain in action. Oxford: Oxford Press.
- Mondor, T. A., Breau, L. M., & Milliken, B. (1998). Inhibitory processes in auditory selective attention: Evidence of location-based and frequency-based inhibition of return. *Perception & Psychophysics*, 60, 296-302.
- Müller, R., Helmert, J.R., Pannasch, S., & Velichkovsky B.M. (2013). Gaze transfer in remote cooperation: Is it always helpful to see what your partner is attending to? *Quarterly Journal of Experimental Psychology*, 66(7), 1302-1316.
- Nalanagula, D., Greenstein, J. S., & Gramopadhye, A. K. (2006). Evaluation of the effect of feedforward training displays of search strategy on visual search performance. *International Journal of Industrial Ergonomics*, 36, 289-300.
- Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: is it normal? Is it automatic? Is it social? *Dev Psych*, 20, 79-97.
- Neider, M. B., Chen, X., Dickinson, C. A., Brennan, S. E. & Zelinsky, G. J. (2010). Coordinating spatial referencing using shared gaze. *Psychonomic Bulletin & Review*, 17(5), 718-724.
- Nuku, P., & Bekkering, H. (2008). Joint attention: Inferring what others perceive (and don't perceive). *Consc Cogn*, 17, 339-349.
- Okamoto-Barth, S., & Kawai, N. (2006). The role of attention in the facilitation effect and another "inhibition of return". *Cognition*, 101, B42-50.
- 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.
- Poliakoff, E., Spence, C., O'Boyle, D. J., McGlone, F. P., & Cody, F. W. J. (2002). Tactile inhibition of return: non-ocular response inhibition and mode of response. *Experimental Brain Research*, 146, 54-59.

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., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *J Exp Psych Gen*, 109, 160-174.

Pratt, J., & Abrams, R. A. (1999). Inhibition of return in discrimination tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 229 - 242.

Pratt, J., Hillis, J., & Gold, J.(2001). The effect of the physical characteristics of cues and targets on facilitation and inhibition. *Psychonomic Bulletin & Review*, 8, 489-495.

Prinz, W. (1997). Perception and action planning. *European journal of cognitive psychology*, 9, 129-154.

Quadflieg, S., Mason, M. F., & Macrae, C. N., (2004) The owl and the pussycat: gaze cues and visuospatial orienting. *Psych Bull Rev*. 11, 826–831.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception & Performance*, 18, 849-860.

Risko, E. F., Laidlaw, K. E., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in human neuroscience*, 6 .

Sadasivan, S., Greenstein, J. S., Gramopadhye, A. K., & Duchowski, A. T., (2005). Use of eye movements as feedforward training for a synthetic aircraft inspection task, In Proceedings of the ACM CHI 2005 Conference, Portland, OR, 141-149.

Samuel, A. G., & Kat, D. (2003). Inhibition of return: A graphical meta-analysis of its time course and an empirical test of its temporal and spatial properties. *Psychonomic*

*Bulletin and Review*, 10, 897–906.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36(04), 393-414.

Sebanz, N., Knoblich, G., Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition*, 88, 11-21.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, 81, 174-176.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, 9, 6-20.

Skarratt, P. A., Cole, G. G., & Kuhn, G. (2012). Visual cognition during real social interaction. *Frontiers in Human Neuroscience*, 6.

Skarratt, P. A., Cole, G. G., & Kingstone, A. (2010). Social inhibition of return. *Acta Psych*, 134, 48-54.

Teufel, C., Alexis, D. M., Clayton, N. S., & Davis, G. (2010). Mental-state attribution drives rapid, reflexive gaze following. *Att Percept Psych*, 72, 695-705.

Theeuwes, J., & Godijn, R. (2002). Attentional and oculomotor capture. In C. Folk & B. Gibson. (Eds.), Attraction, distraction and action: Multiple perspectives on attentional capture (pp. 121149). New York: Elsevier Science.

Tipper, S.P. (1985). The negative priming effect: Inhibitory priming by ignored objects.

*Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 37A, 571–590.

Tipper, S. P., Weaver, B., Cameron, S., Brehaut, J. C., & Bastedo, J. (1991). Inhibitory mechanisms of attention in identification and localization tasks: Time course and

disruption. *Journal of Experimental Psychology. Learning Memory And Cognition*, 17(4), 681–692.

Velichkovsky, B. M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition*, 3, 199-224.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception*, 6, 313-326.

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psych Rev*, 104, 90-122.

Weger, U., Abrams, R. A., Law, M. B., & Pratt, J. (2008). Attending to objects: Endogenous cues can produce inhibition of return. *Visual Cognition*, 16, 659-674.

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. *Neurosci Letters*, 385, 99-104.

Welsh, T. N., & Pratt, J. (2006). Inhibition of return in cue-target and target-target tasks. *Exp Brain Res*, 174, 167-175.

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.

Wu, D. W.-L., Bischof, W. F., & Kingstone, A. (2013). Looking while eating: The importance of social context to social attention. *Scientific Reports*, 3, 2356.

Wu, D. W. L., Bischof, W. F., & Kingstone, A. (2014). Natural gaze signaling in a social context. *Evolution and Human Behavior*, 35(3), 211-218.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.

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

visual search. *J Exp Psych Hum Percept Perf*, 10, 601-621.