Joint perception: gaze and social context

Daniel C. Richardson[1]*, Chris N. H. Street[1], Joanne Y. M. Tan[1], Natasha Z. Kirkham[2], Merrit A. Hoover[3] and Arezou Ghane Cavanaugh[4]

[1] Cognitive, Perceptual and Brain Sciences, University College London, London, UK
[2] Centre for Brain and Cognitive Development, University of London, Birkbeck, London, UK
[3] School of Medicine, Oregon Health and Science University, Portland, OR, USA
[4] Department of Psychology, University of California, Riverside, CA, USA

We found that the way people looked at images was influenced by their belief that others were looking too. If participants believed that an unseen other person was also looking at what they could see, it shifted the balance of their gaze between negative and positive images. The direction of this shift depended upon whether participants thought that later they would be compared against the other person or would be collaborating with them. Changes in the social context influenced both gaze and memory processes, and were not due just to participants' belief that they are looking at the same images, but also to the belief that they are doing the same task. We believe that the phenomenon of joint perception reveals the pervasive and subtle effect of social context upon cognitive and perceptual processes.

Keywords: vision, joint action, eye movements, social cognition, situated cognition

INTRODUCTION

Social context, the real or imagined presence of other people (Allport, 1954), is a ubiquitous psychological force. Cognition is enveloped by social context (Smith and Semin, 2004; Smith and Conrey, 2009). Yet the effects of social context upon cognition often fall between the cracks of social and cognitive psychology. In cognitive and perceptual laboratories, we typically place participants in an experimental quarantine, away from the confounds of social influence. As a consequence, we have many elegant demonstrations of the different behavioral and neurological responses to social versus non-social stimuli (e.g., Cacioppo et al., 2005; Birmingham et al., 2008; Senju and Johnson, 2009), but little idea of how these and other stimuli are processed in a social versus a non-social context. Increasingly, as this special edition shows, researchers are advocating that neuroscience and cognitive psychology directly address issues of social context and interaction (Schilbach, 2010; Obhi and Sebanz, 2011; Shibata et al., 2011).

In this paper, we asked what is the difference between perceiving something by your self and perceiving it at the same time as another person? When a student hovers over your shoulder while you read their paper, does it influence your evaluation? When someone sits down on the sofa while you are watching TV, does their presence intrude upon your experience of the show? What if you are watching a show alone, but know that a friend across town is also tuned in? We term this phenomenon joint perception: the changes that happen when people believe that they are experiencing something at the same time as another person. To isolate these effects from the demands of social interaction, we minimized the social content of joint perception. Participants could not see, hear or interact with each other. We presented images to participants, tracked their gaze, and manipulated—on a trial by trial basis— whether or not they believed that an unseen other person was looking at the same sets of images.

It is hard to discern, from the literature, whether such a minimal social context will have any influence visual attention, as the presence of a social context is often intertwined with social interaction. For example, language use requires a high level of social interaction. When two people talk, their eye movements can be highly sensitive to what they think each other knows and sees (Horton and Keysar, 1996; Bromme et al., 2001; Nadig and Sedivy, 2002; Hanna et al., 2003; Metzing and Brennan, 2003; Brown-Schmidt et al., 2008; Richardson et al., 2009). In contrast, other researchers have argued that people can be striking egocentric (Keysar et al., 2000; Keysar, 2007) in the way they deploy their gaze during language processing.

At a lower level of social context, there are experiments in which two people do not speak to each other but are engaged in the same task. For example, in a traditional stimulus-response compatibility task, participants make a judgment about one stimulus property (color) and ignore another stimulus property (location). If there is an incompatibility between the irrelevant property and the response (e.g., the stimulus is on an opposite side of the screen to the response button) then reaction times increase (Simon, 1969). Sebanz et al. (2003) divided such a task between two people. The participants sat next to each other, and each person responded to one color: in effect, each acting as one of the fingers of a participant in Simon's (1969) experiment. Though each person had only one response to execute, they showed an incompatibility effect when acting together. There was no incompatibility effect when performing the same single response task alone. When engaged in a task together, participants represent their partners' actions as if they were their own.

People will also attune to stimuli that an experimenter identifies as shared: if they are told that someone else is looking at a stimulus, that increases its salience (Shteynberg, 2010). More subtly, people configure their attentional state to that of others.

Their ability to attend to global or local features in a Navon figure (Navon, 1977) is influenced by the knowledge that a co-actor is attending to local or global features (Böckler et al., 2012). Infants follow the gaze of others (Senju and Johnson, 2009), and if their attention is drawn to an event by another's gaze (compared to a non-social cue such as an arrow) they learn more about that event (Wu and Kirkham, 2010; Wu et al., 2011).

In short, people are highly responsive to where others are looking, if they are given that information. In the these experiments, we address a more rudimentary issue: if people simply know that others are looking, but not where, how do they change their gaze patterns?

At this lowest level of social context, the eye movement literature is largely agnostic. From early eye movement research it has been shown that differences in expertise (Buswell, 1935) and cognitive process (Yarbus, 1967; Just and Carpenter, 1976) exert a top-down effect on gaze. But social context itself has not been a variable of concern in eye movement research (e.g., Henderson, 2003), in the way that it has been studied elsewhere (Zajonc, 1965).

The studies we present contrast with those which explicitly give participants a task to perform with another (Sebanz et al., 2003), which explicitly tell participants what another person is attending to (Brennan et al., 2008; Shteynberg, 2010; Böckler et al., 2012) and experiments in which participants communicate with each other (Richardson et al., 2007, 2009; Dale et al., 2011). We presented participants with a set of normed images, knowing that that they would be biased to attend to some over others. Instruction, interaction and cooperation with another person were absent, and we focused on changes in perception that were brought about just by the knowledge that the images were experienced with another person or not. By focusing on this minimal social context, we can explore the shifts in perceptual processes that occur in response to the presence of others, prior to communication, joint action or cooperation taking place.

EXPERIMENT 1

Pairs of participants who did not know each other, or interact during the experiment, sat in opposite corners of the lab. We presented them with sets of four images, on screen for eight seconds. On different trials, they each believed that the other participant was looking at the same images, or that the other was looking at a set of unrelated symbols (**Figure 1**). The four pictures were taken from a normed database (Lang et al., 2005). In each set, there was one picture with a negative valence (e.g., crying child), one with a positive valence (e.g., a smiling couple) and two neutral images with no strong valence (e.g., a person reading). Negative images are considered more potent than equivalently-valenced

FIGURE 1 | Trial schematic.
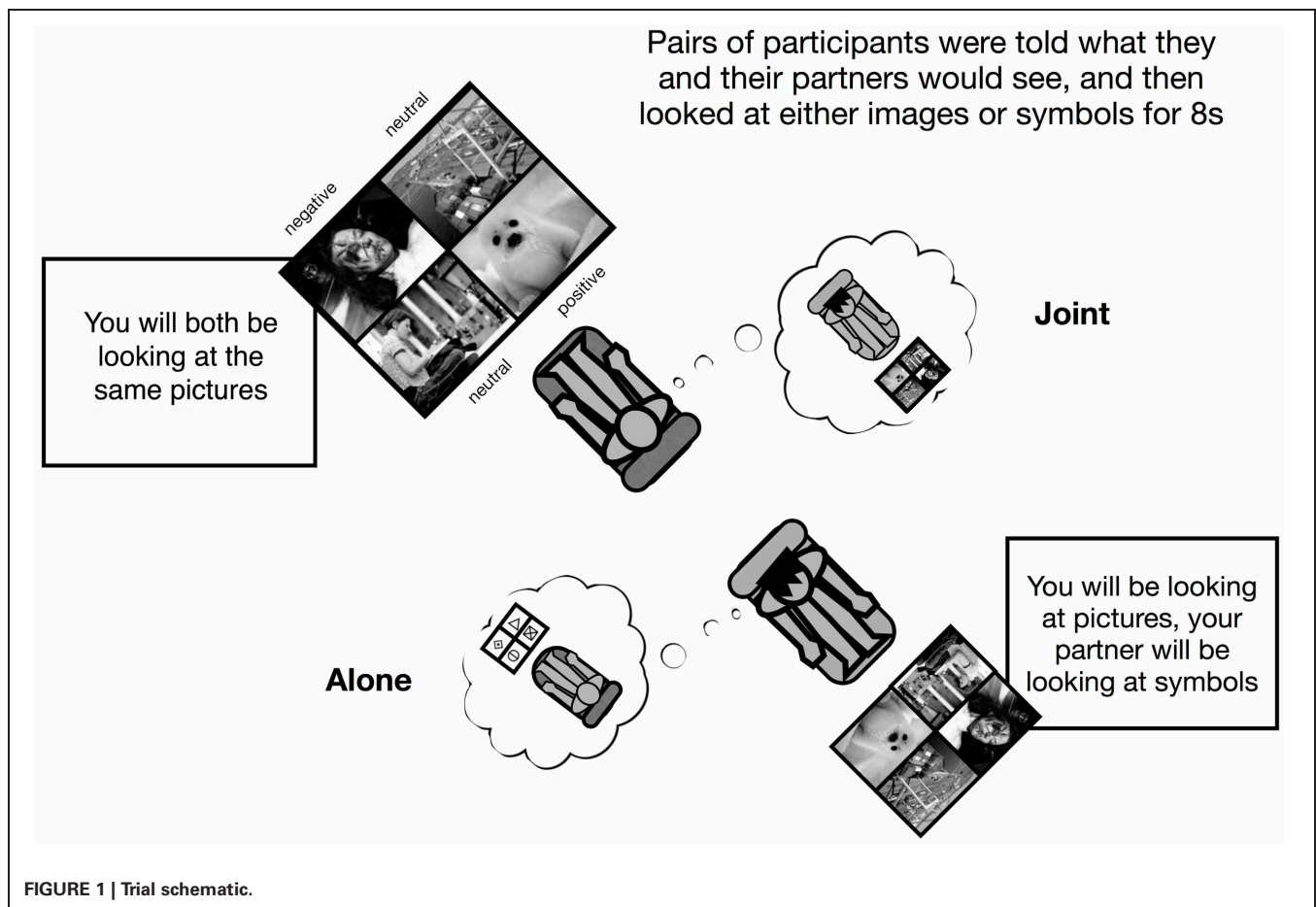
positive images (for reviews, see Baumeister et al., 2001; Rozin and Royzman, 2001). We anticipated, therefore, that the negative stimuli were more likely to receive participants' attention in line with previous work (Smith et al., 2003; Norris et al., 2004; Hajcak and Olvet, 2008). We tested the hypothesis that this attentional bias would be influenced by the minimal social context of the participants' belief that they were looking at the pictures jointly or alone.

METHODS

Participants

There were 20 undergraduates from University College London who took part in the experiment in exchange for course credit. The participants were randomly paired and did not interact. We did not collect data from two due to equipment problems and failures to calibrate. Although we ran pairs of participants in the lab, each participant's data were analysed independently as they could not see each other or interact. At debriefing, participants did not give any indication that they realized we would be comparing their gaze patterns during the joint and alone conditions.

Apparatus

Participants were positioned in opposite corners of a 5 m² room. They could not see each other or each others' displays. Each participant sat in a reclining chair looking up at an arm mounted 19″ LCD screen approximately 60 cm away. A custom built remote eye tracker was mounted at the base of each display. The participants wore headsets, through which they could hear the stimuli and speak to the experimenter. Two iMacs calculated gaze position for each participant approximately 100 times a second, presented stimuli and recorded fixation position parsed into regions of interest. The experimenter's computer saved an audio-video record of what the participants saw, heard and said during the experiment, superimposed with their gaze positions.

Design and procedure

We presented participants with 64 trials in a random order. **Figure 1** provides a schematic. At the start of each trial a prerecorded voice and text message informed participants about the type of images they were about to see, and what their partner would see. Half the time participants saw a set of four pictures, and half the time they saw a set of four symbols. Counterbalanced with the image type, participants were either (truthfully) told that their partner would be looking at the same or a different image type. In each picture trial, two were chosen randomly from a set of neutral images, one from a set of positive, and one from a set of negative images. The sets were created by selecting from Lang et al.'s (2005) database of normed images according to their valency ratings, to produce non-overlapping, equally spaced categories: neutral (valence from 4.8 to 5.2, $M = 5$), positive (7.6–8.3, $M = 8$), and negative (1.6–2.4, $M = 2$). The pictures were of real world scenes as might be seen in a newspaper. The symbol sets, which served only as filler items in this design, were taken at random from a set of geometric patterns found in various font sets. The images were displayed onscreen for 8 s. Following a blank screen for 1 s, the next trial began.

Results and discussion

We calculated the total looking times to positive and negative images across each trial, as shown in **Figure 2**. These times were different when participants were looking alone versus jointly. A 2 (picture valence: negative or positive) × 2 (social context: joint or alone) ANOVA showed a significant interaction [$F_{(1, 17)} = 9.96$, $p = 0.006$, $\eta^2 = 0.37$], and a significant difference between valence conditions only in the joint condition (Tukey's HSD, $p < 0.01$). When they believed that their partner was looking at the same stimuli, participants looked more at the negative images. There was no significant difference when they believed they were looking alone. There was a main effect of picture valence [$F_{(1, 17)} = 5.24$, $p = 0.04$, $\eta^2 = 0.24$] but not of social context alone ($F < 1$).

Participants in this experiment could not see or interact with each other, and had no knowledge of each others' gaze or attentional focus. They were not instructed to perform a task with each other or coordinate their activity in any way. They simply viewed pictures by themselves, with or without the experimenter's assurance that an unseen partner could see the same thing. Yet surprisingly their eye movements were systematically shifted by this minimal social context on a trial-by-trial basis. It was not
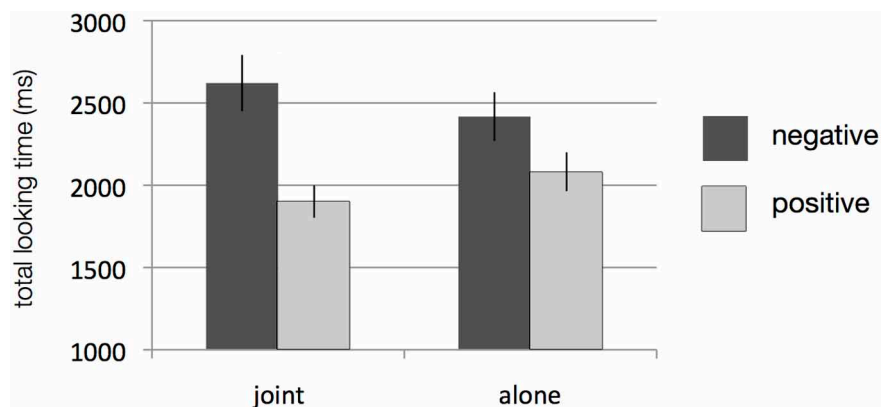

FIGURE 2 | Results form Experiment 1.

simply that shared images received greater attention, as found by Shteynberg (2010). Indeed, in our experiment there was no main effect on looking times overall. More specifically than has been shown before, we found that when set images were believed to be shared there was a shift in participants' distribution of attention.

EXPERIMENT 2

We have demonstrated that eye movements are influenced by beliefs about social context. One could argue, however, that eye movements are indicative of lower level perceptual processing alone, and that in cognitive terms they are epiphenomenal. Although there are theoretical and empirical arguments against this view (Spivey et al., 2009), we wanted to investigate whether minimal social context differences were also reflected in a measure of cognitive performance: recognition memory. In this version of the paradigm, eye movement measures were not taken but, following presentation blocks, participants' memory for the images was tested. We hypothesized that minimal social context which affected attention in Experiment 1 would be sufficient to affect memory here.

METHODS

The experiment was identical to Experiment 1 apart from the following details.

Participants

There were 36 undergraduates from University College London who took part in the experiment in exchange for course credit. We did not use data from eight because, at debriefing, the participants indicated some awareness of our hypotheses.

Design and procedure

All participants were run simultaneously in separate cubicles of a computer lab. At the start of the experiment, an instruction screen told them that they would be collaborating with a partner on a memory task, and that the computer had randomly paired them with another participant in the group. They saw a fake text message from the other participant greeting them, and were invited to respond with a short message. In fact, the participants were not paired with anyone and had no interaction with each other.

There were two identical blocks. In the presentation phase of each, participants saw eight trials that were identical to those shown in Experiment 1: half were picture presentations, and half were symbols. On half the trials participants were told that they were looking the same images as their partners, and on the other half that they were looking at different images. Following that, there were 32 test trials, which consisted of a single picture presented until the participants made a yes or no response to indicate whether they had seen it before. On half the occasions, the picture had been previously presented and was either one of the negative or one of the positive images.

Results

Accuracy recognizing pictures that had been seen before was 85%, and did not differ between experimental conditions. Following standard work in visual memory (Sternberg, 1969) and, more specifically, work on the social tuning of memory (Shteynberg, 2010), we used reaction times as a more sensitive measure of memory performance. A 2 (valence) × 2 (social context) ANOVA found a significant interaction [$F_{(1, 27)} = 6.98$, $p = 0.014$, $\eta^2 = 0.21$]. In the joint looking condition, the negative images ($M = 758$ ms, $SD = 114$) were recognized faster than the positive ($M = 794$ ms, $SD = 120$). Conversely, in the alone condition, positive images ($M = 785$ ms, $SD = 113$) were recognized faster than the negative ($M = 828$ ms, $SD = 155$). There was a main effect of social context [$F_{(1, 27)} = 8.01$, $p = 0.009$, $\eta^2 = 0.23$], but not of valence ($F < 1$).

Discussion

Looking at something together affects more than eye movements. The images that received more visual attention in previous experiments, according to their valence and the social context, were also remembered more efficiently in this study. This result echoes Shteynberg's (2010) finding that when participants believe other people are examining the same stimuli as they are, those images become more "psychologically prominent." But in contrast, here and in Experiment 1, participants were not told which of the images the other person was looking at. They simply knew *that* another person was looking, and this minimal social context influenced which particular images attracted more attention and proved easier to recall. In the following experiment, we investigated exactly what it was about the "minimal social context" that brought about this attentional shift. In other words, what counts as "looking together"?

EXPERIMENT 3

There are at least two ways to interpret "looking together," which up until now we have treated as a single idea. On the one hand, looking together could mean just experiencing a set of images at the same time. On the other hand, it could mean examining the same images, but also having the same goal, attitude or intention towards them.

Our joint perception paradigm was based on work in joint action (Sebanz et al., 2003; Galantucci and Sebanz, 2009). Joint action effects do not occur if the participant is simply sat next to another person (Tsai et al., 2006), or if that person's button pressing actions are not intentional (their finger is moved by a mechanical device). Also, if the participant is acting jointly, but with a computer program (Tsai et al., 2008) or a marionette's wooden hand (Tsai and Brass, 2007) there is no stimulus-response incompatibility effect. Participants only form representations of another when that person's genuine, intentional actions are engaged in the same task (Atmaca et al., 2011).

In the current experiment, we began to investigate whether the same sort of conditions circumscribing joint action also determine joint perception. Unlike those described above, in this experiment the participants always believed that they were examining the same images. What changed, trial-by-trial, was the task that they were doing, and the task that they believed their partner was doing. Sometimes they or their partner were memorizing the pictures, sometimes they were scanning them for the presence of a small X. We predicted that joint perception effects would be strongest when participants believed that they were not just passively sharing an experience, but also engaged in the same task.

METHODS

The experiment was identical to Experiment 1, apart from the details below.

Participants

There were 32 University College London students who participated for course credit. Data from four participants were unusable due to equipment calibration problems.

Design

The instruction screen defined two tasks for the participants. In a memory task, they had to remember the pictures for a later test (which never actually took place). In the search task, they had to look for a translucent X superimposed on one image, and press the mouse button that they held in one hand if they detected it. They were informed that both their own task and their partner's task could change from trial to trial, but both of them would always see the same pictures.

At the start of each trial, participants were told their task for the upcoming presentation. A large icon at the top of the screen represented the task (visual search or memory), and a smaller icon below showed their partner's task (shown in **Figure 3**). They also heard a voice say "You will be [memorizing/searching]. Your partner will be [memorizing/searching]".

There were 40 trials. In half the participant was told to memorize the stimuli and in half to search for an X. Participants' own task was crossed with the task they were told their partner was doing. Half the time they were told that their partner performed the same task, and half a different task. On eight trials (spread evenly across conditions), an X appeared at a random location on one of the images.

Results and discussion

Participants showed a robust preference for negative images over positive images only when they believed that they and their partner had been assigned the same task. We calculated the total amount of time spent looking at the critical negative and positive images on trials where there was no X (we did not analyse the 20% of trials when there was an X present, as X and participants' responses to it would interfere with how they allocated their attention to each image). A 2 (valence) × 2 (own task: memory/search) × 2 (other's task: same/different) ANOVA was performed, and the means for each cell are displayed in **Figure 3**. There was a significant two way interaction between valence and other's task [$F_{(1, 27)} = 10.08$, $p = 0.004$, $\eta^2 = 0.41$]. Post hoc tests show that the difference between positive and negative images was significant when the participants believed they were doing the same task (Tukey's HSD $p = 0.01$), but did not reach significance when they were doing a different task. There was also a main effect of valence [$F_{(1, 27)} = 19.19$, $p = 0.0001$, $\eta^2 = 0.27$], but all other main effects and interactions were non significant (all $Fs < 1$).

The effect of joint perception does not occur simply when participants believe that another person is experiencing the same stimuli. It is necessary for them to believe that the other, unseen person is engaged in the same task as themselves. This task could be to memorize the pictures, which presumably would require processing the meaning of an image, or the task could just be to search for a visual feature, which requires only superficial processing. Regardless, the effect of joint perception arises whenever these tasks are believed to be done together.

GENERAL DISCUSSION

Social context exerts a pervasive effect on perception. Even a minimal social context, when the difference between looking alone and
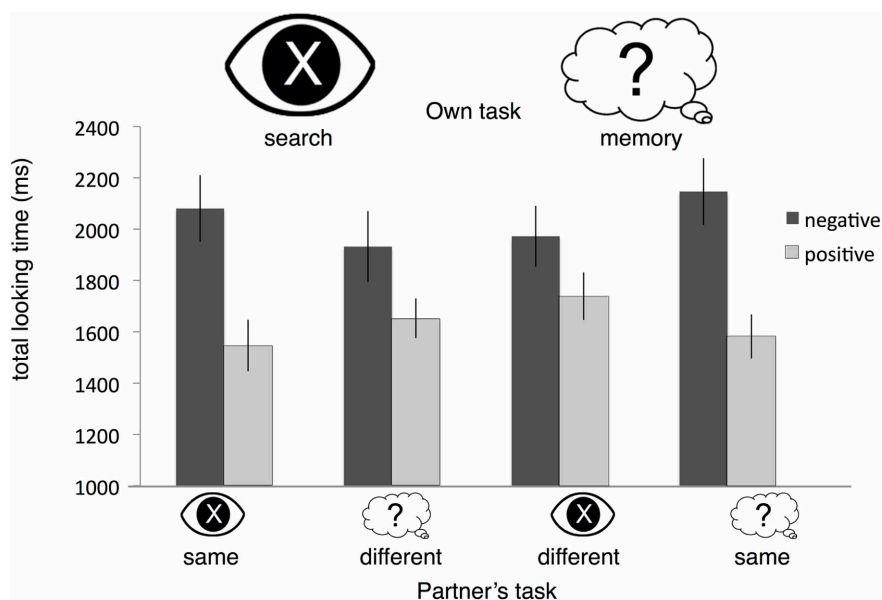
FIGURE 3 | Results from Experiment 3. Looking times showed a significant interaction between valance and whether or not the participant's partner was belived to be doing the same or a different task.

looking jointly is as small as possible, produces distinct behavioral and cognitive effects. Shared exposure is not sufficient to produce these effects alone: participants must also believe that they are engaged in the same task when processing the shared stimuli.

This result is distinct from other findings in area between social and cognitive psychology. There are many interesting studies of joint action (e.g., Obhi and Sebanz, 2011), but our experiments are different because participants are not instructed to coordinate their behavior or act together. There are many interesting studies on joint attention and how people use information about each other's attentional state (Brennan et al., 2008; Shteynberg, 2010; Böckler et al., 2012), but our experiments are different because participants are given no knowledge of where the other is looking. And finally, there are many studies of attentional coordination during social interaction and language use (e.g., Richardson et al., 2007), but in our experiments there is no interaction between people at all. Nevertheless, despite the very minimal nature of this minimal social context, it produces a systematic shift in participants' attention.

In these first experiments, we have tried to understand the conditions under which joint perception influences attention. But we have not yet addressed the direction of these effects. Why is it that sharing images in our paradigm led to increased attention specifically to the negative pictures? Here we discuss four alternatives: social context modulates the strength of the negativity bias specifically, or it modulates attention and alertness more broadly; social context increases the degree to which there is alignment with emotions, or alignment with saliency.

It has been argued that the negativity bias exists because of a learnt or evolved priority to detect threats in the environment (Baumeister et al., 2001; Rozin and Royzman, 2001). If social context was associated with an increase in perceived threat or anxiety, then it would follow that joint perception could increase the negativity bias specifically. This is possible, but it seems unlikely that our participants would have felt increased threat from each other. All participants were first year undergraduate students at UCL, and so were members of similar or overlapping social groups. Even if they did feel some anxiety in each others' presence, it is not clear why that threat would change trial-by-trial according to the stimuli they believed each other could see. However, to fully discount this possibility, we would need to experimentally manipulate the anxiety felt by participants, perhaps by changing their in/out group relationship.

The second possibility is that the social context of joint perception increases some broad cognitive factor such as alertness, in the way that the presence of others can cause social facilitation (Zajonc, 1965). It has been shown, for example, that when participants are engaged in a dialogue, it can increase alertness and counter the effects of sleep deprivation (Bard et al., 1996). Perhaps the lower level of social context used in this experiment, and modulated trial-by-trial, also increased alertness. This increased engagement would presumably benefit the negative images first of all, since there is a pre-existing bias towards them. However, under this account, it remains a puzzle why there would be no corresponding increase in looks to positive items at all. One would expect a main effect of social context on look times to these two items (compared to the neutral items), but throughout our experiments we found an interaction between social context and valence.

A third possibility draws on work in social psychology showing that social interaction leads to emotional alignment. When people interact, they are motivated to form a "shared reality" (Hardin and Higgins, 1996): a speaker will adapt the content of their message to align with the beliefs and emotions of their audience (reviewed by Echterhoff et al., 2009). Similarly, when people collaborate in groups, they tend to align with the group emotion (Hatfield et al., 1993; Wageman, 1995; Barsade, 2002). Since individuals are attuned to negative stimuli, it is conceivable that in a group, this shared negativity bias would be amplified as people seek to align with each other. Over repeated experiences, perhaps this social alignment towards negative stimuli becomes ingrained. In this light, our joint perception phenomenon could be seen as a form of minimal, imagined cooperation that is sufficient to evoke a learnt alignment towards negative images.

The final alternative is that the joint perception effect is not driven by emotion, per se, but by salience. This account draws on observations of language use and the rich joint activity of social interaction. Language is remarkably ambiguous. "Please take a chair," could refer to a variety of actions with a variety of chairs in a room. Conversations do not grind to a halt however, because people are very good at resolving ambiguous references by drawing on knowledge about the context and assumptions that they have in common (Schelling, 1960). For example, when presented with a page full of items, such as watches from a catalogue, participants agreed with each other which one was most likely to be referred to as "the watch" (Clark et al., 1983).

When we enter into any conversation, such coordination is all important (Clark, 1996), and can be seen at many levels of behavior. When we talk, we use the same names for novel objects (Clark and Brennan, 1991), align our spatial reference frames (Schober, 1993), use each others' syntactic structures (Branigan et al., 2000), sway our bodies in synchrony (Condon and Ogston, 1971; Shockley et al., 2003) and even scratch our noses together (Chartrand and Bargh, 1999). When we are talking and looking at the same images, we also coordinate our gaze patterns with each other (Richardson and Dale, 2005), taking into account the knowledge (Richardson et al., 2007) and the visual context (Richardson et al., 2009) that we share. In short, language engenders a rich, multileveled coordination between speakers (Shockley et al., 2009; Louwerse et al., in press).

Perhaps the instruction stating that images were being viewed together was enough to turn on some of these mechanisms of coordination, even in the absence of any actual communication between participants. When images were believed to be shared, participants sought out those which they imagined would be more salient for their partners. Since saliency is driven by the valence of the images in our set, paying more attention to the most salient means paying more attention to the negative image. In this way, it can be argued that the shifts brought about by joint perception are the precursors to the more richly interactive forms of joint activity studied in other fields.

Our experiments echo a point that social psychologists have made from the outset. The presence and actions of others can

have a powerful effect on an individual's motivations, goals and judgments (Triplett, 1898; Sherif, 1935; Lewin, 1936; Festinger, 1950; Asch, 1951; Allport, 1954; Heider, 1958; Zajonc et al., 1969). Beliefs and judgments are not formed in cognitive isolation, but always in the context of the thoughts and opinions of those around us (Smith and Semin, 2004). Here we have shown that these lessons from social psychology can be applied to a simple perceptual process in a minimal social context. Merely the belief that stimuli are attended to alone or with another is enough to activate coordinative behaviours that are the basis of joint action, communication and social interaction. The pervasive effects of social context have theoretical implications for how we view cognition (Robbins and Aydede, 2009), adding to calls to consider social interaction at its heart (Smith and Semin, 2004; Barsalou et al., 2007).

ACKNOWLEDGMENTS

REFERENCES

Allport, G. W. (1954). *The Nature of Prejudice*. Cambridge, MA: Perseus Books.

Asch, S. E. (1951). "Effects of group pressure upon the modification and distortion of judgement," in *Groups, Leadership and Men*, ed H. Guetzkow (Pittsburgh, PA: Carnegie Press), 177–190.

Atmaca, S., Sebanz, N., and Knoblich, G. (2011). The joint flanker effect: sharing tasks with real and imagined co-actors. *Exp. Brain Res.* 211, 371–385.

Bard, E. G., Sotillo, C., Anderson, A. H., Thompson, H. S., and Taylor, M. M. (1996). The DCIEM map task corpus: spontaneous dialogue under sleep deprivation and drug treatment. *Speech Commun.* 20, 71–84.

Barsade, S. G. (2002). The ripple effect: emotional contagion and its influence on group behavior. *Adm. Sci. Q.* 47, 644–675.

Barsalou, L. W., Breazeal, C., and Smith, L. B. (2007). Cognition as coordinated non-cognition. *Cogn. Process.* 8, 79–91.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., and Vohs, K. D. (2001). Bad is stronger than good. *Rev. Gen. Psychol.* 5, 323–370.

Birmingham, E., Bischof, W. F., and Kingstone, A. (2008). Social attention and real world scenes: the roles of action, competition, and social content. *Q. J. Exp. Psychol.* 61, 986–998.

Böckler, A., Knoblich, G., and Sebanz, N. (2012). Effects of a coactor's focus of attention on task performance. *J. Exp. Psychol. Hum. Percept. Perform.* doi: 10.1037/a0027523. [Epub ahead of print].

Branigan, H. P., Pickering, M. J., and Cleland, A. A. (2000). Syntactic coordination in dialogue. *Cognition* 75, 13–25.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., and Zelinsky, G. J. (2008). Coordinating cognition: the costs and benefits of shared gaze during collaborative search. *Cognition* 106, 1465–1477.

Bromme, R., Rambow, R., and Nückles, M. (2001). Expertise and estimating what other people know: the influence of professional experience and type of knowledge. *J. Exp. Psychol. Appl.* 7, 317–330.

Brown-Schmidt, S., Gunlogson, C., and Tanenhaus, M. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition* 107, 1122–1134.

Buswell, G. (1935). *How People Look at Pictures: A Study of the Psychology and Perception in Art*. Chicago, IL: University of Chicago Press.

Cacioppo, J. T., Visser, P. S., and Pickett, C. L. (Eds.) (2005). *Social Neuroscience: People Thinking About Thinking People*. Cambridge, MA: The MIT press.

Chartrand, T. L., and Bargh, J. A. (1999). The chameleon effect: the perception-behavior link and social interaction. *J. Pers. Soc. Psychol.* 76, 893–910.

Clark, H. H. (1996). *Being there: Putting Brain, Body, and the World Together Again*. Cambridge, MA: MIT Press.

Clark, H. H., and Brennan, S. E. (1991). "Grounding in communication," in *Perspectives on Socially Shared Cognition*, eds L. B. Resnick, J. Levine, and S. D. Teasley (Washington, DC: APA), 127–149. Reprinted in "Groupware and computer-supported cooperative work," in *Assisting Human-Human Collaboration*, ed R. M. Baecker (San Mateo, CA: Morgan Kaufman Publishers Inc.), 222–233.

Clark, H. H., Schreuder, R., and Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *J. Verbal Learn. Verbal Behav.* 22, 245–258.

Condon, W., and Ogston, W. (1971). "Speech and body motion synchrony of the speaker-hearer," in *The Perception of Language*, eds D. Horton and J. Jenkins (Columbus, OH: Charles E. Merrill), 150–184.

Dale, R., Kirkham, N. Z., and Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Front. Psychol.* 2:355. doi: 10.3389/fpsyg.2011.00355

Echterhoff, G., Higgins, E. T., and Levine, J. M. (2009). Shared reality experiencing commonality with others' inner states about the world. *Perspect. Psychol. Sci.* 4, 496–521.

Festinger, L. (1950). Informal social communication. *Psychol. Rev.* 57, 271–282.

Galantucci, B., and Sebanz, N. (2009). Joint action: current perspectives. *Top. Cogn. Sci.* 1, 255–259.

Hajcak, G., and Olvet, D. M. (2008). The persistence of attention to emotion: brain potentials during and after picture presentation. *Emotion* 8, 250–255.

Hanna, J., Tanenhaus, M., and Trueswell, J. (2003). The effects of common ground and perspective on domains of referential interpretation. *J. Mem. Lang.* 49, 43–61.

Hardin, C. D., and Higgins, E. T. (1996). "Shared reality: How social verification makes the subjective objective," in *Handbook of Motivation and Cognition: The Interpersonal Context*, Vol. 3, eds E. T. Higgins and R. M. Sorrentino (New York, NY: Guilford), 28–84.

Hatfield, E., Cacioppo, J. T., and Rapson, R. L. (1993). Emotional contagion. *Curr. Dir. Psychol. Sci.* 2, 96–99.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York, NY: John Wiley and Sons.

Henderson, J. (2003). Human gaze control in real-world scene perception. *Trends Cogn. Sci.* 7, 498–504.

Horton, W., and Keysar, B. (1996). When do speakers take into account common ground? *Cognition* 59, 91–117.

Just, M., and Carpenter, P. (1976). Eye fixations and cognitive processes. *Cogn. Psychol.* 8, 441–480.

Keysar, B. (2007). Communication and miscommunication: the role of egocentric processes. *Intercult. Pragmatics* 4, 71–84.

Keysar, B., Barr, D. J., Balin, J. A., and Brauner, J. S. (2000). Taking perspective in conversation: the role of mutual knowledge in comprehension. *Psychol. Sci.* 11, 32–38.

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2005). *International Affective Picture system (IAPS): Digitized Photographs, Instruction Manual, and Affective Ratings (Tech. Rep. A-6)*. Gainesville, FL: University of Florida, Center for Research in Psychophysiology.

Lewin, K. (1936). *Principles of Topological Psychology*. New York, NY: McGraw Hill.

Louwerse, M. M., Dale, R. A., Bard, E. G. and Jeuniaux, P. (in press). Behavior matching in multimodal communication is synchronized. *Cogn. Sci.*

Metzing, C., and Brennan, S. (2003). When conceptual pacts are broken: partner-specific effects on the comprehension of referring expressions. *J. Mem. Lang.* 49, 201–213.

Nadig, A., and Sedivy, J. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychol. Sci.* 13, 329–336.

Navon, D. (1977). Forest before trees: the precedence of global features in visual perception. *Cogn. Psychol.* 9, 353–383.

Norris, C. J., Chen, E. E., Zhu, D. C., Small, S. L., and Cacioppo, J. T. (2004). The interaction of social and emotional processes in the brain. *J. Cogn. Neurosci.* 16, 1818–1829.

Obhi, S. S., and Sebanz, N. (2011). Moving together: toward understanding the mechanisms of joint action. *Exp. Brain Res.* 211, 329–336.

Richardson, D. C., and Dale, R. (2005). Looking to understand:

the coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cogn. Sci.* 29, 1045–1060.

Richardson, D. C., Dale, R., and Kirkham, N. Z. (2007). The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. *Psychol. Sci.* 18, 407–413.

Richardson, D. C., Dale, R., and Tomlinson, J. M. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cogn. Sci.* 33, 1468–1482.

Robbins, P., and Aydede, M. (Eds) (2009). *The Cambridge Handbook of Situated Cognition.* Cambridge, UK: Cambridge University Press.

Rozin, P., and Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Pers. Soc. Psychol. Rev.* 5, 296–320.

Schelling, T. C. (1960). *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

Schilbach, L. (2010). A second-person approach to other minds. *Nat. Rev. Neurosci.*

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition* 47, 1–24.

Sebanz, N., Knoblich, G., and Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition* 88, B11–B21.

Senju, A., and Johnson, M. H. (2009). Atypical eye contact in autism: models, mechanisms and development. *Neurosci. Biobehav. Rev.* 33, 1204–1214.

Sherif, M. (1935). A study of some social factors in perception. *Arch. Psychol.* 27, 17–22.

Shibata, H., Inui, T., and Ogawa, K. (2011). Understanding interpersonal action coordination: an fMRI study. *Exp. Brain Res.* 211, 569–579.

Shockley, K., Richardson, D. C., and Dale, R. (2009). Conversation and coordination structures. *Top. Cogn. Sci.* 1, 305–319.

Shockley, K., Santana, M. V., and Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 326–332.

Shteynberg, G. (2010). A silent emergence of culture: the social tuning effect. *J. Pers. Soc. Psychol.* 99, 683–689.

Simon, J. R. (1969). Reactions toward the source of the stimulation. *J. Exp. Psychol.* 81, 174–176.

Smith, E. R., and Conrey, F. R. (2009). "The social context of cognition," in *Cambridge Handbook of Situated Cognition,* eds P. Robbins and M. Aydede (Cambridge, UK: Cambridge University Press), 454–466.

Smith, E. R., and Semin, G. R. (2004). Socially situated cognition: cognition in its social context. *Adv. Exp. Soc. Psychol.* 36, 53–117.

Smith, N. K., Cacioppo, J. T., Larsen, J. T., and Chartrand, T. L. (2003). May i have your attention, please: electrocortical responses to positive and negative stimuli. *Neuropsychologia* 41, 171–183.

Spivey, M. J., Richardson, D. C., and Dale, R. (2009). "The movement of eye and hand as a window into language and cognition," in *Oxford Handbook of Human Action,* eds E. Morsella, J. A. Bargh, and P. M. Gollwitzer (New York, NY: Oxford University Press), 225–249.

Sternberg, S. (1969). Memory-scanning: mental processes revealed by reaction-time experiments. *Am. Sci.* 57, 421–457.

Triplett, N. (1898). The dynamogenic factors in pacemaking and competition. *Am. J. Psychol.* 9, 507–533.

Tsai, C. C., and Brass, M. (2007). Does the human motor system simulate Pinocchio's actions? co-acting with a human hand versus a wooden hand in a dyadic interaction. *Psychol. Sci.* 18, 1058–1062.

Tsai, C. C., Kuo, W. J., Hung, D. L., and Tzeng, O. J.-L. (2008). Action co-representation is tuned to other humans. *J. Cogn. Neurosci.* 20, 2015–2024.

Tsai, C. C., Kuo, W. J., Jing, J. T., Hung, D. L., and Tzeng, O. J. L. (2006). A common coding framework in self-other interaction: evidence from joint action task. *Exp. Brain Res.* 175, 353–362.

Wageman, R. (1995). Interdependence and group effectiveness. *Adm. Sci. Q.* 40, 145–180.

Wu, R., and Kirkham, N. Z. (2010). No two cues are alike: depth of learning during infancy is dependent on what orients attention. *J. Exp. Child Psychol.* 107, 118–136.

Wu, R., Gopnik, A., Richardson, D. C., and Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Dev. Psychol.* 47, 1220–1229.

Yarbus, A. (1967). *Eye Movements and Vision.* New York, NY: Plenum Press.

Zajonc, R. B. (1965). Social facilitation. *Science* 149, 269–274.

Zajonc, R. B., Heingartner, A., and Herman, E. M. (1969). Social enhancement and impairment of performance in the cockroach. *J. Pers. Soc. Psychol.* 13, 83–92.