Joint Perception: Gaze and the Presence of Others

Daniel C. Richardson (dcr@eyethink.org)
School of Psychology, Reading University,
Whiteknights Road, Reading, RG6 6AL, UK

Merrit A. Hoover (mhoover@ucsc.edu)
Arezou Ghane (aghane2002@yahoo.com)
Department of Psychology, University of California, Santa Cruz
1056 High Street, Santa Cruz, CA 95064 USA

Abstract

We document a new phenomenon: participants' eye movements are influenced by the belief that they are either looking at pictures alone, or that a person next door is looking at the same pictures. The pictures were in sets of four. One had a positive and one a negative valence, and others were neutral. On a trial by trial basis, eye movements to the negative versus the positive pictures were modulated by participants' belief that they were looking alone or jointly. A second experiment manipulated the beliefs participants had about the person in the next cubicle, in an attempt to tease apart possible explanations for this effect of 'joint perception'. In each case, a minimal sense of cooperation with another appears to produce distinct cognitive effects, in manner similar to that found in studies of 'joint action'. We conclude that there may be a pervasive effect of social context upon cognitive and perceptual processes.

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

Introduction

Imagine you are at a large, international conference. The parallel sessions are simultaneously offering little of interest, and so you wander into the empty hall where the posters are on display. You begin to study one, and after a few moments, become immersed. Then you hear a shuffle over your shoulder. Someone else is looking at the poster too. You pass over the distraction and return to your study. How is your perception of the poster now changed by the knowledge that someone else is looking at it? Does your attention increase? Do you look more critically or more generously? The other person could be a senior researcher, famed for scathing remarks, or a young graduate student. Does your impression of the other person change the influence of their presence on your perception?

By their presence - most forcibly by looking into your eyes - other people compel you to realize that you are an object for them, Sartre (1948) argues. The arresting sensation of direct eye contact has certainly been investigated by philosophers (Stack & Plant, 1982) and psychologists (Argyle, 1967), who find it closely connected to the interpersonal forces of dominance and attraction. Here we are interested in another aspect of 'the other'. Sartre (1948) wrote that in the presence of another, "Your world is suddenly haunted by the Other's values, over which you have no control." Is there a systematic effect of social context upon perceptual processes?

Social context

In the standard information processing accounts of cognition, the theoretical model is of a cognitive systems that thinks and acts in isolation. Accordingly, laboratory experiments quarantine subjects away from a social context. While this approach has been greatly productive, for example, in mapping out the behavioural and anatomical components of cognition, it risks excluding critical phenomena from the laboratory. Notably, elsewhere, neuroscientists are discovering that the brain is highly tuned to social information (Cacioppo, Visser & Pickett, 2005), yet such variables are not often part of cognitive theories. Thus, while the tools of cognitive psychology are being used to study social phenomena to great effect, the potential impact of social context upon cognitive and perceptual processing is rarely explored.

Though the standard cognitive model marginalizes social context, there have been notable exceptions. Studies of situated cognition (Barsalou, Breazeal & Smith, 2007; Robbins & Ayded, in press) show that cognition 'in the wild' is intimately linked not only to representations of the external world, but also to the cognitive processes of others. For example, Hutchins (1995) observed the ways that navy navigators distribute cognitive processes between themselves by using external tools and representations, such as maps and notations.

Recently, experimental methods are starting to reveal the mechanisms involved in such joint activity. Experimenters have taken some classic experiments in the information processing model of cognitive psychology, and placed them in a rudimentary social context (Sebanz, Bekkering & Knoblich, 2006). For example, in a traditional stimulus-response compatibility task, participants make a judgment about one stimulus property (colour) and ignore another stimulus property (location). If there is an incompatibility between the irrelevant location property and the response (left or right finger movement), then reaction times increase, as the irrelevant property activates the incompatible response representation (Simon, 1969). Sebanz and colleagues divided this task between two people. They sat next to each other, and each person responded to one colour: in effect, each acting as one of the fingers of a participant in Simon's (1969) experiment. Although each individual had only one response to execute (and hence no need to represent the incompatible response), they still showed slower responses in the incompatible trials. There was no incompatibility effect when performing the same single

response task alone. When acting jointly, participants represented their partners' actions as if they were their own. Further evidence suggests that representing the actions of another occurs 'quasi-automatically' during joint action, even when it leads to a decrement in performance.

As well as representing the actions of others, there is recent evidence that we monitor their potential responses by keeping track of a rich set of social variables. Crosby, Monin & Richardson (in press) found that people look to those around them when interpreting potentially offensive behaviour. Eye movements of participants were tracked as they watched a video of four people giving their views on university admissions policy. All four discussants were visible at all times. When one person voiced strong, potentially offensive remarks against affirmative action, the amount of time participants looked at a Black discussant increased. Importantly, these looks did not occur when participants viewed an identical video, but were led to believe that the Black individual could not hear the comments because his headphones were turned off. Therefore, participants do not simply direct attention to members of minority groups whenever affirmative action is mentioned. Instead, they maintain an awareness of others' knowledge state, social identity and possible reaction to particular attitudes.

There is mounting evidence that (when it is provided) participants ground simple cognitive tasks within a social context. Inspired by the simplicity of the joint action experiments, our goal was a take a straightforward case of perception, eye movements to pictures of differing affect, and to examine the effect of introducing a minimal social context. Experiment 1 was designed simply to determine whether social context alters the way participants view these images.

Experiment 1

The social context in our experiment began when participants entered the laboratory in pairs. They were led into separate cubicles next door to each other, and believed (correctly) that they were going to be run together in an experiment. Each saw sets of pictures presented on screen for a short time. In the 'joint perception' trials, they heard an experimenter's voice say, 'You will both be looking at the same set of pictures,' and in the 'lone perception' trials they heard, 'You will be looking at pictures. Your partner will be looking at symbols'. There were also filler trials in which the participants looked symbols (but were told their partner was looking at pictures). In this way, the social context was varied trial by trial throughout the experiment, and consisted of no more than a belief about what the person next door could see on their screen.

The perceptual task was to look at sets of four pictures that appeared simultaneously onscreen. On each trial, participants saw two 'neutral' fillers, typically, anodyne objects or landscapes,one pleasant image, and one unpleasant image. The pictures were collected according the the subjective opinion of the experimenters. While it is certain that this may have produced an uneven stimuli set, the categories are divergent enough that it is implausible, for example, that what one person would call an unpleasant image, another would call pleasant. The full set of positive and negative images are shown in Figure 1.

Valance was chosen as a stimulus dimension because what is termed negative and positive images are often loaded with social values. For example, there are taboos against unclean and violent things, on the one hand; and more Calvinist taboos against staring covetously at pleasant things, such as overly indulgent cakes or scantily clad models on the other. To the extent to which positive and negative images interact with social mores, they seemed good bait for effects of social context.

Our simple hypothesis was that the minimal social context will interact with the perceptual task. Biases towards negative and positive images will shift when the participants believe they are looking alone or jointly with another. Our prediction was that a bias towards positive stimuli will increase when the participants believe they are looking alone.

Method

Participants 27 undergraduates from the University of California, Santa Cruz took part in the experiment in exchange for course credit. We did not collect data from 6 due to equipment problems and failures to calibrate because of corrected vision. Participants were strangers who had not met prior to the experiment.

Apparatus Participants were run, two at a time, in the Eye Think lab's speech and gaze tracking system. Each participant sat in a cubicle in a reclining chair, looking up at an arm mounted 19" LCD screen approximately 60cm away. A Bobax3000 remote eye tracker was mounted at the base

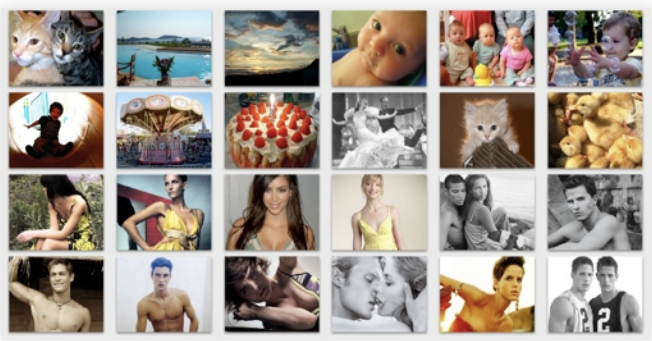

Figure 1A. Positive images

Figure 1B. Negative images

of the display, consisting of a camera focused on the participant's eye and a set of LED illuminators. The participant wore a headset, through which they could hear the stimuli and speak to the experimenter. The eye trackers passed image data to intel iMacs positioned outside the cubicle. The iMacs calculated gaze position for each participant approximately 30 times a second, presented stimuli and recorded regions of interest that were being fixated. Data was also streamed to an experimenter's computer, which saved an audio-video record of what the participant saw, heard and said during the experiment, superimposed with their gaze position.

Design There were 32 trials presented in a random order. Half of these were experimental trials. Participants saw either a set of pictures or a set of symbols, and were (truthfully) told that their partner was either looking at the same stimuli, or stimuli of a different type. The picture sets consisted of two neutral images, one positive and one negative (Figure 1), and were placed in random locations on a two by two grid. Symbol sets were taken at random from a set of geometric patterns found in various font sets. The other half of the trials were distractor trials, in which one of the participants saw a single picture with three symbols, and the other saw four pictures. These trials, and the choice of abstract, Zener card symbols were intended to resemble ESP experiments and draw attention away from the experimental contrasts of interest.

Procedure Participants waited to begin the experiment in a common waiting area. The experimenter briefly introduced herself and the participants to each other, and then they were led to adjacent cubicles. Each underwent a short calibration session, typically lasting no more than a minute. Once they were ready, the experimenter warned them over the intercom that the experiment was to begin. Each trial began with a pre-recorded voice describing the upcoming stimuli (pictures/symbols) and social context. Participants were given no explicit instructions about the images. In the key experimental contrast, they were told either 'You will both be looking at the same pictures' or 'You will be looking at pictures. Your partner will be looking at symbols'. The visual stimuli then appeared and remained onscreen for eight seconds. After a 1000ms ISI, the next trial commenced. Computers in the two cubicles communicated so that the trials were presented in synchrony throughout the session. Upon completion, participants were debriefed and informed of the true experimental hypothesis.

Results

Positive and negative images were attended to differently according to the social context. In the alone condition, negative images were fixated for 2436ms (SD=770) and positive images for 1853ms (SD = 676), while in the joint condition, negative images were fixated for 2158ms (SD=872) and positive images for 2370ms (SD=s =961). A 2 (picture valance: negative or positive) x 2 (social context: joint or alone) ANOVA revealed supported this significant interaction ($F(1,20)=5.52$, $p<.05$), with no main effects (All Fs <1). As Figure 2 shows, when they believed

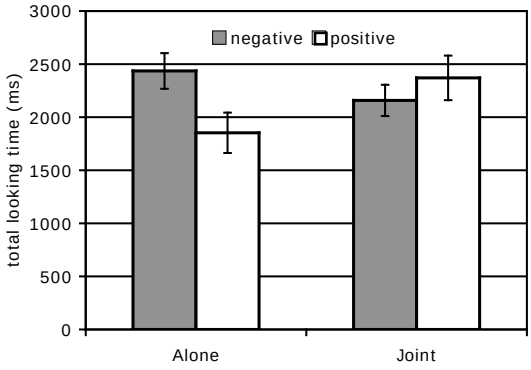

Figure 2. Mean looking times, Experiment 1.

that they were looking alone, participants looked more at the negative stimuli, but looked more at the positive stimuli when they thought their partner could see the same thing. Figure 3 reveals the timecourse of these gaze patterns. In both conditions, there was an early peak in gazes to the negative stimuli. In the lone condition this preference persisted, but in the joint condition it gave way to a preference for positive images.

Discussion

Participants' gaze patterns to positive and negative stimuli were influenced by whether they thought that a person in the next cubical was looking at the same pictures, or an unrelated set of symbols. Indeed, this social context could flip on a trial by trial basis, yet participants kept track of it and modulated their looking. The joint perception hypothesis was supported: even a minimal sense of social context can change a perceptual process. Our prediction of the effects of joint perception fared less well:when looking alone, participants' gaze was biased towards negative stimuli. Why does social context affect gaze, and why in this direction? We put forward two main working hypotheses, which are neither exhaustive nor mutually exclusive.

Focal Image Hypothesis When we enter into any joint activity, coordination is all important. Communication is a prime example of coordination (Clark, 1996) and when we talk, we implicitly agree upon names for novel objects (Clark & Brennan, 1991), align our spatial reference frames (Schober, 1993) and use each others' syntactic structures (Branigan, Pickering & Cleland, 2000). Our accents become more similar (Giles, Coupland & Coupland, 1992), we sway our bodies in synchrony (Shockley, Santana & Fowler, 2003; Condon & Ogdon, 1971) and even scratch our noses together (Chartrand & Bargh, 1999). This coordination at multiple levels facilitates many functions of communication, such as promoting goodwill, establishing reference and setting up common ground. Richardson and colleagues found that two people looking at a common scene, such as a painting, closely coordinated their gaze patterns. This coordination increased with the degree of shared knowledge possessed by the participants (Richardson, Dale & Kirkham, 2007), and the level of gaze coordination was causally linked to the success of language comprehension (Richardson & Dale, 2005).

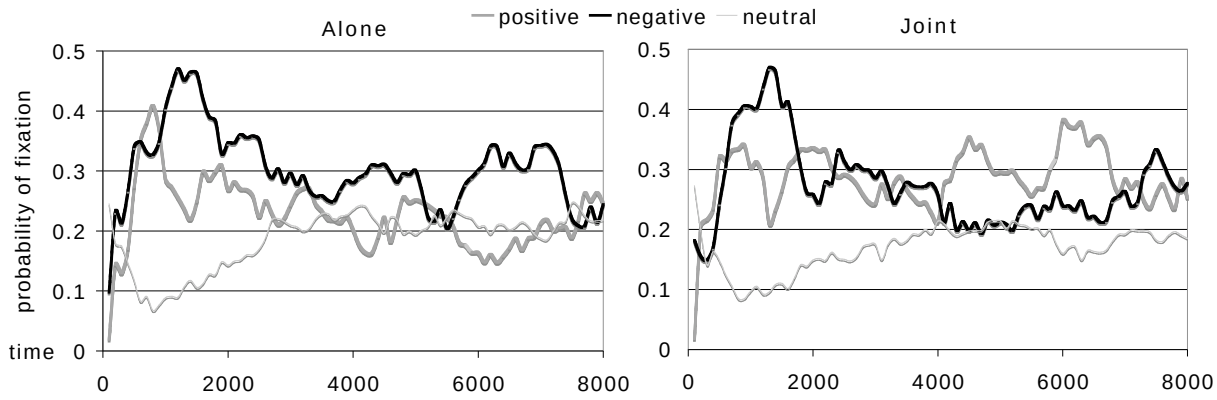

Figure 3. Timecourse of gaze preferences, Experiment 1.

We hypothesized that participants in our experiment implicitly entered into a joint activity. Perhaps anticipating a future discussion of the stimuli between themselves, they attempted to coordinate gaze patterns with their partner, when they believed they were looking jointly. In other words, in the joint condition they looked at the pictures they thought another person would prefer: the pleasant ones.

The term 'focal image' comes from Schelling (1960) who found that people were very good at guessing what images others find salient. Schelling realized that everyday cases of verbal reference are often ambiguous. We say, 'Hand me the fork,' in the presence of many such items, yet listeners unproblematically infer the same referent. For example, when presented with a page full of items, such as watches from a catalog, participants agreed with each other which one was most likely to be referred to as 'the watch' (Clark, Schreuder & Buttrick, 1983).

The focal image hypothesis assumes that the 'baseline' preference is to look at the unpleasant images. The focal image hypothesis posits that this preference is overridden in the joint case, when participants imagine that other people would like to see nice images, and look accordingly. So the hypothesis asserts that people individually prefer the macabre, but (incorrectly) think that other people have more genial tastes.

Threat Calculus Hypothesis Walking down a dark alley, do you jump quicker at noise behind you when you are alone, or when you are walking with others? Our intuition is that of sensible necessity, when you are alone you attend more to stimuli that could be potentially threatening. This response is so ingrained that even the minimal solitude of our 'lone perception' condition, participants increase their attention to negative (and hence threatening) images.

This hypothesis may seem to ascribe too much paranoia to our participants, since it suggests they are continually aware of potential dangers. However, they do live a society with a daily colour coded 'threat level'. There is also strong experimental evidence that we have a heightened sensitivity to negative stimuli. Threatening subliminal words are better detected (Dijksterhuis & Aarts, 2003) and demand more attentional resources than positive words (Pratto & John, 1991). At the timescale of milliseconds, we seem to sense threat within stimuli. At the timescale of minutes, with more complex stimuli, our understanding of threat seems more malleable. In particular, it can be influenced by those around us. In Darley and Latane's (1968) studies of bystander intervention, the presence of other people made an individual less likely to help in a threatening or difficult situation, such as helping a stranger who has collapsed on a street. Partly this was because the individual would rely on others to interpret the situation as an emergency, and partly the others would 'diffuse responsibility' for taking action.

Task demand hypotheses There are many other factors that could be contributing to our effect. Perhaps participants suspect that following presentation there will be some sort of memory test. This stirs their competitive spirit or fear of failure, and on joint trials they aim to look more evenly at all the pictures (rather than simply focusing on the negative one) so that their memory performance does not suffer a poor comparison with their partners'. Or perhaps they imagine that their own gaze patterns will be later compared with their partners', and so do not want to appear, by comparison, to be unduly attracted to unpleasant images. These hypotheses have not been directly tested yet, but note that to some degree they all suggest that the participants are changing their behaviour because they believe that their partner is looking at particular images. In this sense, they presuppose the focal image hypothesis.

Experiment 2

The hypotheses we lay out are not directly at odds with each other. The focal image hypothesis attempts to describe the special case of joint looking, assuming that lone looking is the baseline. Conversely, the threat calculus hypothesis assumes that lone looking is something to be explained. In experiment 2 we devised a test of the focal image hypothesis, but for these reasons, the results cannot rule out threat calculus as a possibility. The experiment aimed to reveal any calculations participants make about the type of images their partner may prefer. We manipulated participants' beliefs about their partner, who in this case was an experimental confederate in one of two guises as an art major. In one case, he or she wore a tie dye t-shirt and liked "taking pictures of everything I find beautiful and pleasing". In the other case, he or she wore a death metal t-shirt and liked 'taking pictures of everything I find bizarre and

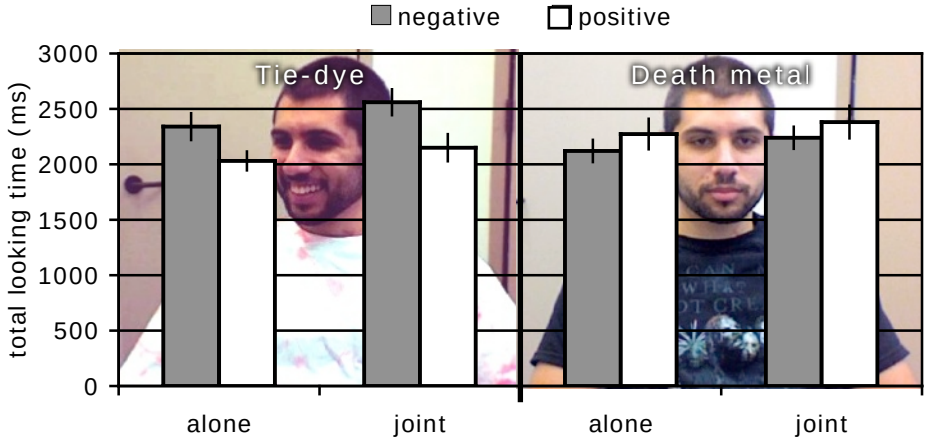

Figure 4. Mean looking times, Experiment 2.

shocking'. Figure 4 shows examples of our male research assistant in his two guises.

Following the focal image hypothesis, our prediction was that participants would not look differently in the alone condition, but would follow their partners' assumed preferences in the joint looking condition, looking more at the positive pictures with the tie-dye confederate, and more at the negative pictures with the death metal confederate.

Method

Participants 44 undergraduatesfrom UC Santa Cruz participated for course credit. 4 could not be run because of calibration errors.

Apparatus The apparatus was identical to Experiment 1.

Design Stimuli were the same as Experiment 1. The distractor trials were dropped, since debriefing following Experiment 1 suggested that participants were not suspicious about the joint/alone manipulation, and were content to assume it was an experiment investigating stimuli preferences. This allowed us to double the number of experimental trials.

Procedure There were no differences to Experiment 1 except for the fact that participants were run with a confederate who was a male or a female research assistant. They wore one of two shirts, shown in figure 4, but otherwise behaved in a consistent and unremarkable fashion. Prior to calibration, with the excuse of breaking the ice, the experimenter asked the participants to say a few things about themselves. The confederate was asked to go first and said, "My name is [Jackie / Daniel]. I'm from Kresge and I'm an art major. I mostly do photography. I'm all about taking pictures of everything I find [bizarre and shocking / beautiful and pleasing]."

Results The identity of the confederate influenced participants' looking behaviour to positive and negative images. We hypothesized no effect of confederate in the alone condition but an interaction in the joint condition, and so carried out separate planned comparisons of the two social contexts. In the alone condition, a 2 confederate (tie-dye vs death metal) x 2 valence (positive vs negative) ANOVA revealed no significant interaction ($F(1,38)=3.46$) and no main effects (all Fs <1). A similar ANOVA on the joint condition found a significant interaction between confederate and valence ($F(1,38)=4.09$, $p<.05$), with no main effects (all Fs <1).

The patterns of these means can be seen in Figure 4. Although, as hypothesized, there was an interaction of confederate, to our surprise it had the opposite effect that we predicted. When looking jointly with someone who professed to like pleasant things, participants spent roughly 25% longer on the negative pictures (Tukey's $p<.05$).

Discussion

Evidence for the focal image hypothesis was not forthcoming. In the important joint condition, participants did not spend more time looking at the images their partner would prefer. Indeed, when paired with the tie-dye confederate, participants looked more at the pictures that were anything but 'beautiful and pleasing'. We draw two conclusions from this result, or rather, see two directions for further research.

Firstly, we do not think we can discard the focal image hypothesis yet. The stimulus categories of positive and negative valence were inherited from the first experiment, but one could argue that they do not naturally lend themselves to individuals' preference. A better test would be to vary picture content: a confederate could confess to being a sports fan or an architecture student. In the joint condition the focal image hypothesis would predict looks to buildings or sporting activities to increase. If no increase was found in the alone condition, a general priming effect of talking about such preferences could be ruled out.

We have one reason why the focal image hypothesis failed in this experiment. But we have not explained the positive result of a preference for negative images in the presence of a tie-dye confederate. One possibility, and our second direction for future research, concerns the threat calculus hypothesis.

We imagined that participants modulate their attention to threatening stimuli in the light of their social context. Being alone is inherently riskier, and looking alone, the need for vigilance is greater. Looking jointly, there is another set of eyes to watch out for threatening things, and the help of

another if required, and so the need to attend to negative stimuli diminishes. However, this 'threat calculus' rests on the assumption that ones partner does indeed attend to negative stimuli, and would indeed be useful in a threatening situation. Perhaps in Experiment 2, participants judged that the tie-dye wearing aesthete would not fulfill these roles, and their threat calculus led them to increase attention to negative stimuli when looking jointly. Although the pattern of our results is consistent with the threat calculus hypothesis, we cannot say that that Experiment 2 adduces direct support. For that, we need either to untangle stimuli that are threatening versus simply negative, or to experimentally manipulate the degree to which the other participant is seen as ameliorating threat.

Conclusion

Believing that another person, out of sight, is this moment looking at the same images as you are alters how you look at them. Your beliefs about that person also influence how you look. We have established, therefore, two aspects of the new phenomenon of joint perception: the presence and the type of social context can interact with perceptual processes. Surprisingly, perhaps, the condition of looking alone or looking jointly was not established once at the start of the session, and its effects left to accrue throughout the experiment; the social context flipped back and forth, trial-by-trial. The causal route of this behaviour has not yet been made clear, whether it is monitoring threat, the intention to coordinate, or some other factor. We think that, at the least, this research reveals the powerful and persuasive effect of social context upon cognitive and perceptual processes.

Acknowledgments

We would like to thank Herb Clark, Natasha Kirkham, and Michael Spivey for helpful comments and Norine Doherty, Jacquelyn Espino, Victor Hernandez, Daniel Janulaitis and Jonas Nagel for data collection and insightful discussion.

References

Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge: Cambridge University Press.

Barsalou, L.W., Breazeal, C., & Smith, L.B. (2007). Cognition as coordinated non-cognition. *Cognitive Processing*, 8, 79-91.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic coordination in dialogue. *Cognition*, 75, B13-B25.

Cacioppo, J.T., Visser, P.S.& Pickett C.L.(eds.), *Social neuroscience: People thinking about thinking people*. Cambridge, MA: The MIT press.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76, 893-910.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: APA.

Clark, H.H., Schreuder, R. & Buttrick, S., (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior,* 22, 245-258.

Condon, W., & Ogston, W. (1971). Speech and body motion synchrony of the speaker– hearer. In D. Horton & J. Jenkins (Eds.), *The perception of language* (pp. 150 – 184). Columbus, OH: Charles E. Merrill.

Crosby, J.R., Monin, B. & Richardson, D.C. (in press). Where Do We Look During Potentially Offensive Behavior? *Psychological Science*

Darley, J.M. & Latane, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8, 377-383.

Dijksterhuis, A., & Aarts, H. (2003). On wildebeests and humans: The preferential detection of negative stimuli. *Psychological Science,* 14, 14–18.

Giles, H., Coupland, N., & Coupland, J. (1992). Accommodation theory: Communication, context and consequences. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation* (pp. 1-68). Cambridge: Cambridge University Press.

Hutchins, E., 1995, *Cognition in the Wild.* Cambridge, MA: MIT Press.

Pratto, F., & John, O.P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology,* 61, 380–391

Richardson, D.C & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science,* 29, 1045–1060.

Richardson, D.C, Dale, R. & Kirkham, N.Z., (2007) The art of conversation is coordination: common ground and the coupling of eye movements during dialogue, *Psychological Science,* 18(5), 407-413

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

Sartre, J-P. (1948), *Being and Nothingness*, tr. Hazel E. Barnes, New York: Philosophical Library.

Schelling, T. C. *The Strategy of Conflict*, Cambridge, Mass.: Harvard University Press.

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

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, 10, 70-76

Shockley, K., Santana, M.V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception & Performance,* 29(2), 326-332.

Simon, H. (1969). Reactions towards the source of the stimulation. *Journal of Experimental Psychology,* 81.

Stack, G.J., & Plant, R.W. (1982). The Phenomenon of "The Look", *Philosophy and Phenomenological Research*, 42, 3, 359-373.