NeuroImage 84 (2014) 648–656

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Revealing the neural networks associated with processing of natural social interaction and the related effects of actor-orientation and face-visibility

Manish Saggar *, Elizabeth Walter Shelly, Jean-Francois Lepage, Fumiko Hoeft, Allan L. Reiss

Center for Interdisciplinary Brain Sciences Research, Department of Psychiatry and Behavioral Sciences, Stanford University, USA

ARTICLE INFO

Article history:
Accepted 19 September 2013
Available online 29 September 2013

Keywords:
Social cognition
Event-related fMRI
Non-verbal communication
Actor-orientation
Face-visibility

ABSTRACT

Understanding the intentions and desires of those around us is vital for adapting to a dynamic social environment. In this paper, a novel event-related functional Magnetic Resonance Imaging (fMRI) paradigm with dynamic and natural stimuli (2 s video clips) was developed to directly examine the neural networks associated with processing of gestures with social intent as compared to nonsocial intent. When comparing social to nonsocial gestures, increased activation in both the mentalizing (or theory of mind) and amygdala networks was found. As a secondary aim, a factor of actor-orientation was included in the paradigm to examine how the neural mechanisms differ with respect to personal engagement during a social interaction versus passively observing an interaction. Activity in the lateral occipital cortex and precentral gyrus was found sensitive to actor-orientation during social interactions. Lastly, by manipulating face-visibility we tested whether facial information alone is the primary driver of neural activation differences observed between social and nonsocial gestures. We discovered that activity in the posterior superior temporal sulcus (pSTS) and fusiform gyrus (FFG) was partially driven by observing facial expressions during social gestures. Altogether, using multiple factors associated with processing of natural social interaction, we conceptually advance our understanding of how social stimuli is processed in the brain and discuss the application of this paradigm to clinical populations where atypical social cognition is manifested as a key symptom.

Introduction

Although Darwinian theory of evolution states that only the fittest will survive, cooperation within the same- and between-species is very common (Axelrod and Hamilton, 1981; Henrich et al., 2003). In particular, human societies outrank all other species based on our large-scale cooperation, even among genetically unrelated individuals (Fehr and Fischbacher, 2004; Nowak, 2012). The ability to efficiently communicate and interact is perhaps what makes humans such remarkable outliers relative to the rest of the animal kingdom. For a successful social interaction, we use a complex array of cues including facial expressions, gaze, gestures, speech, intonation and cadence to gauge the intentions of others around us. In comparison to the lower-level visual system, where the processing of stimuli is relatively linear and sequential (Gold et al., 2012; Shapley, 2009), processing of social stimuli is highly complex and non-linear, mainly due the inherent complexities of the stimuli itself. Thus, delineating the neural correlates of natural social interactions can be a daunting task. However, it is crucial that we understand how our brains process social cues, so that clinical populations where atypical social cognition is evident can be better understood and treated.

To understand and reveal the neural correlates of basic social interactions, this paper focuses on non-verbal communications, which can be used for obtaining and transmitting socially relevant information by performing different facial expressions, gaze movements, and hand gestures (Montgomery and Haxby, 2008). Several fMRI studies have investigated these communications in healthy adults (Conty and Grezes, 2012; Gallagher and Frith, 2004; Knutson et al., 2008; Lotze et al., 2006; Montgomery and Haxby, 2008; Montgomery et al., 2007; Morris et al., 2005; Schilbach et al., 2006; Villarreal et al., 2008). Although initial studies were limited to using static pictures of faces or bodies as stimuli (Montgomery and Haxby, 2008), several recent studies have explored new ways of presenting stimuli to better capture the neural correlates of realistic nonverbal communications. For example, researchers have started to appreciate the value of using short video clips to present dynamic and realistic stimuli to participants (Gallagher and Frith, 2004; Knutson et al., 2008; Lotze et al., 2006; Montgomery et al., 2007; Villarreal et al., 2008). Video clips are preferred, as compared to the static pictures, because the dynamic images tend to elicit higher activation in the brain regions responsible for processing and producing affective content (e.g., amygdala) as well as explicit body movements (e.g.,

* Corresponding author.
E-mail address: saggar@stanford.edu (M. Saggar).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.09.046

pSTS), thereby providing more information regarding social stimulus processing than static pictures (Grèzes et al., 2007). Other studies have used virtual immersion paradigms, e.g., Morris and colleagues contrasted trials where a virtual character was present with trials where just a framed picture was shown (Morris et al., 2005). Similarly, Schilbach and colleagues used virtual characters to differentiate between the neural networks associated with processing of one-on-one social interaction and processing of passively viewed social interaction between others (Schilbach et al., 2006). Using the virtual immersion paradigms, it is experimentally feasible to develop and study complex social interaction scenarios. However, it is unclear whether the neural networks engaged during a virtual interaction overlap (and to what degree) with the networks engaged during a *natural* interaction.

Given the complex nature of a social stimulus and its processing, several previous studies have attempted to fragment the stimuli into different categories and examine the associated neural networks in isolation. For example, researchers have examined the neural correlates of nonverbal interaction using just facial expressions (Montgomery and Haxby, 2008) or isolated hand movements (Villarreal et al., 2008). Further, recent studies have also attempted to differentiate and partition gestural stimuli into different categories, e.g., transitive versus intransitive (Villarreal et al., 2008), communicative versus object-based (Morris et al., 2005), or expressive versus body-referred (Lotze et al., 2006). Although these studies provide crucial information regarding how our brains process social cues in isolation, it is unclear what brain regions are engaged explicitly for processing cues involved in the "social" interaction gestalt. More importantly, because the social cognition system can be safely assumed to be more than just the sum of its parts, *fragmenting* social stimuli into different components and studying them in isolation could potentially obscure a more accurate and holistic understanding of this system.

To address some of the limitations in the extant literature and to investigate the neural basis of processing nonverbal interaction without finely fragmenting its components, we developed the Dynamic Social Gesture (DSG) task. This task uses dynamic and natural gesture stimuli (video clips of 2 s each) in an event related fMRI design. The DSG task is comprised of three factors (Fig. 1). First, sociability – short clips of actors performing interactive (e.g., "a friendly wave") or non-interactive gestures (e.g., "reaching for a cup"). Thus, in this paper, a gesture is deemed as social if it is *intended* to elicit a response from the participant and non-social otherwise. Very few studies to date have *directly* examined the neural networks associated with sociability (Mainieri et al., 2013; Morris et al., 2005). Of these studies, Morris and colleagues used a virtual immersion paradigm, where the social trials were defined as those in which a virtual character was present and the nonsocial trials were defined by the absence of the virtual character (and instead a framed picture of human face was shown). Furthermore, the virtual character always performed non-interactive gestures (e.g., scratching its face).

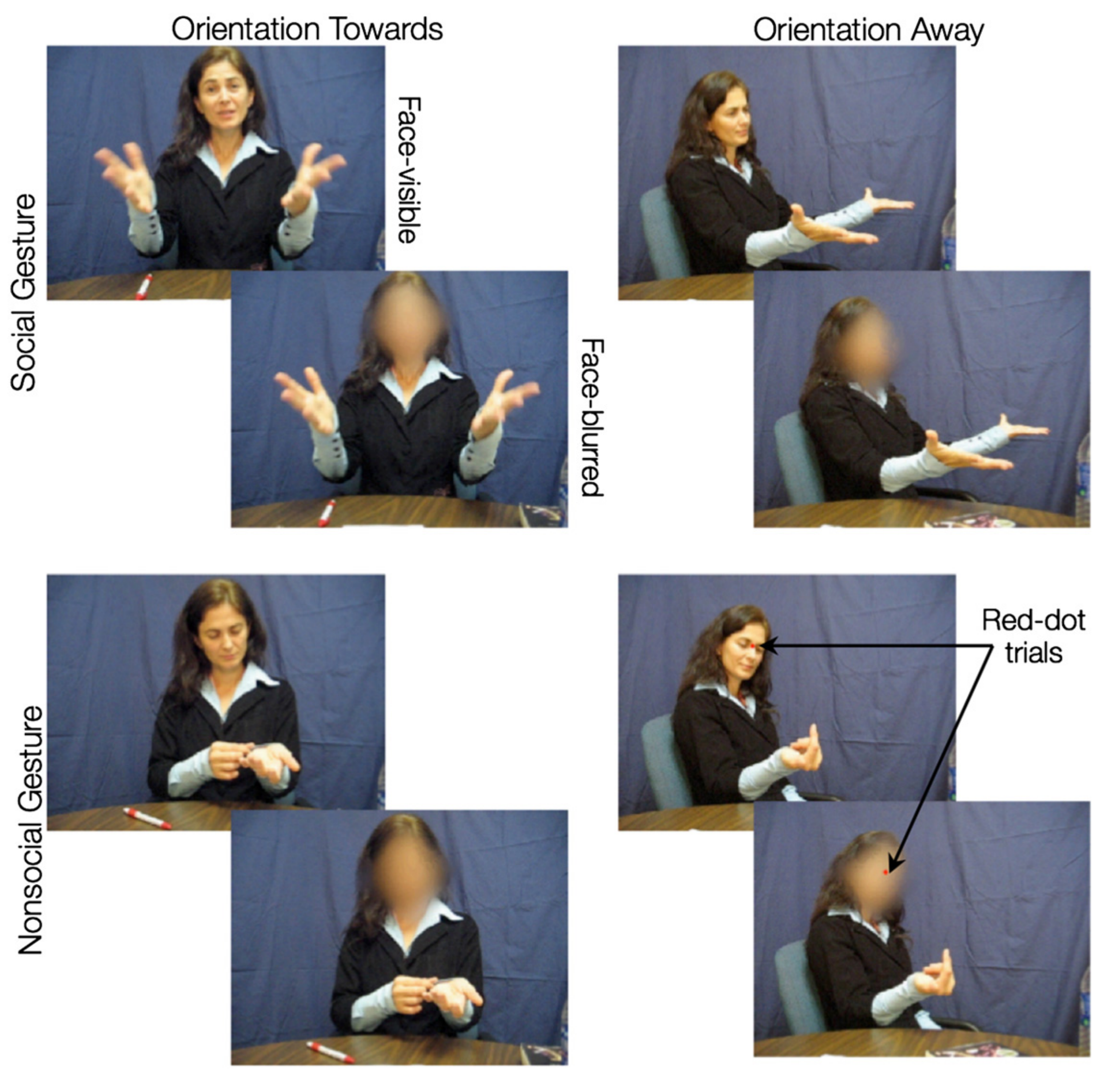

Fig. 1. DSG task contains three factors: sociability, actor-orientation, and face-visibility. Here we show sample snapshots from our dataset of 200 dynamic social gesture stimuli. Two gestures are shown here — implore (social) and looking at arms (nonsocial). To make sure that participants were attending to the stimuli, half of the gestures had a red dot that appeared on the actor's face (between the eyes) 1 s after the stimulus-onset. As a cover task, participants were instructed to press a button as soon as they detected the red dot.