



# Joint attention, social-cognition, and recognition memory in adults

Kwanguk Kim<sup>1</sup> and Peter Mundy<sup>1,2</sup>\*

<sup>1</sup> Department of Psychiatry, MIND Institute, University of California at Davis, Sacramento, CA, USA

<sup>2</sup> School of Education, University of California at Davis, Davis, CA, USA

## Edited by:

Leonhard Schilbach,  
Max-Planck-Institute for Neurological  
Research, Germany

## Reviewed by:

Ron Dotsch, Radboud University  
Nijmegen, Netherlands  
Andrew Bayliss, University of East  
Anglia, UK

## \*Correspondence:

Peter Mundy, School of Education,  
Social Attention and Virtual Reality  
Lab, University of California at Davis,  
202 Cousteau Place, Davis, CA  
95618, USA.  
e-mail: pcmundy@ucdavis.edu

The early emerging capacity for Joint Attention (JA), or socially coordinated visual attention, is thought to be integral to the development of social-cognition in childhood. Recent studies have also begun to suggest that JA affects adult cognition as well, but methodological limitations hamper research on this topic. To address this issue we developed a novel virtual reality paradigm that integrates eye-tracking and virtual avatar technology to measure two types of JA in adults, Initiating Joint Attention (IJA) and Responding to Joint Attention (RJA). Distinguishing these types of JA in research is important because they are thought to reflect unique, as well as common constellations of processes involved in human social-cognition and social learning. We tested the validity of the differentiation of IJA and RJA in our paradigm in two studies of picture recognition memory in undergraduate students. Study 1 indicated that young adults correctly identified more pictures they had previously viewed in an IJA condition (67%) than in a RJA (58%) condition,  $\eta^2 = 0.57$ . Study 2 controlled for IJA and RJA stimulus viewing time differences, and replicated the findings of Study 1. The implications of these results for the validity of the paradigm and research on the affects of JA on adult social-cognition are discussed.

**Keywords:** joint attention, information processing, eye-tracking, virtual avatar, brain-behavior

## INTRODUCTION

Human beings have an exquisitely honed capacity to coordinate their attention with that of other people. This capacity is referred to as joint attention (JA; Bruner, 1995). JA is often defined in terms of socially coordinated visual attention, and operationalized with measures of the ability to follow the gaze of another person to adopt a common point of reference (Responding to Joint attention, RJA), and to use one's own gaze direction or gestures to initiate a common point of reference with another person (Initiating Joint Attention, IJA; Seibert et al., 1982). Theory suggests that, as facility with joint visual attention increases, it becomes internalized as the capacity to coordinate joint mental attention with others. This developmental re-description of overt visual JA in infancy to internal mental JA in the later preschool period provides a vital foundation for human social-cognition and social learning (Tomasello et al., 2005; Mundy et al., 2009).

The potential for a deeper understanding of the role of JA in human cognitive development has emerged recently in the guise of several studies on the impact of JA on information processing in adults. Bockler et al. (2011) reported an innovative study that indicated that the experience of JA with another person enhanced participants' mental spatial rotation problem solving in order to judge the similarity of variously positioned images of right versus left hands. This finding is consistent with the long-standing hypothesis that JA affects and is affected by the self-referenced spatial information processing ability of the individual (Butterworth and Jarrett, 1991). In another study, Frischen and Tipper (2004) have published a seminal paper that indicates the cuing effects of the experience of JA triggers visual orienting and information

processing responses that are significantly more resistant to inhibition of return (IOR) in adults than is observed in a non-social attention cuing task. Linderman et al. (2011) have corroborated this finding in a study that shows that social attention cuing via hand gestures also leads to greater resistance to IOR than non-social cuing. Interesting, indirect comparison of the results of these last two studies suggests that gaze triggered cuing may lead to greater resistance to IOR than hand gesture cuing. Finally, Bayliss et al. (2006) have observed that social-gaze directed cuing to pictures (i.e., emulated RJA) enhanced the subjective positive valence of the pictures for adults compared to a condition involving non-social directional cuing to pictures.

These findings in research with adults are consistent with data for studies of young children. Striano et al. (2006a,b) have reported that active versus passive caregiver JA cuing leads to better short-term picture recognition memory in infants, and associated EEG data indicative of enhanced neural depth of processing. These observations have been replicated and extended by Kopp and Lindenberger (2011) who observed that JA is related to EEG evidence of depth of processing that is associated with long term as well as short-term picture recognition memory in infancy. An additional study reports evidence of more widespread frontal, central, parietal neural network activity, and better depth of processing among toddlers during a word learning task in joint attention, versus non-joint attention conditions (Hirotsani et al., 2009).

## HOW DOES JOINT ATTENTION AFFECT INFORMATION PROCESSING?

The observations of Hirotsani et al. (2009) of widespread frontal, central, and parietal neural network activation during the

experience of JA in toddlers is consistent with the parallel and distributed information processing model of JA (Mundy, 2003; Mundy et al., 2009; Mundy and Jarrold, 2010). Accordingly, JA is enabled by the parallel processing of internal information about one's own visual attention with external information about the visual attention of other people. This type of joint processing of information about *self and other attention* entails the activation of a distributed anterior and posterior cortical attention neural network. With practice during early development the integrated joint processing of *self* attention and *others'* attention becomes automatically engaged in social interactions as a frontal-parietal social-executive function. Evidence for the activation of this distributed system during JA has been documented in adults (Williams et al., 2005; Redcay et al., 2010; Schilbach et al., 2010) as well as young children (e.g., Mundy et al., 2000; Henderson et al., 2002; Grossman and Johnson, 2010). This model of a parallel and distributed social-information processing model system serves to improve the understanding of phenomenon associated with JA development, *including its association with enhanced depth of processing*.

Keysers and Perrett (2006) propose that the relational processing of self-referenced and other referenced information, such as what occurs in JA, may be thought of in terms of Hebbian learning. Hebb (1949) proposed that neural networks that are repeatedly active at the same time become associated, such that specific activity (e.g., re-presentations) in one network triggers activity in the other (Hebb, 1949). Parallel and distributed cognitive theory suggests that depth of encoding is optimized by the simultaneous activation of multiple neural networks during information processing (e.g., Otten et al., 2001; Munakata and McClelland, 2003). These aspects of theory suggest that processing information in the context of JA would be likely to enhance depth of processing and memory by embedding declarative and episodic encoding of shared experience in association with the parallel activation of a distributed neural network engaged in processing of information pertaining to the attention of self and the attention of others.

#### **DIFFERENCES IN IJA VERSUS RJA INFORMATION PROCESSING EFFECTS**

The parallel and distributed processing model of JA also helps to explain the phenomenon of dissociated RJA and IJA development. Although they share common processes RJA and IJA measures are not highly correlated and have unique paths of associations with developmental outcomes in typical children (Brooks and Meltzoff, 2002; Mundy et al., 2007). An uneven pattern of development favoring RJA over IJA has also been observed in children affected by Autism Spectrum Disorders (Mundy et al., 1994) and in comparative studies of JA development in primates (Tomasello and Carpenter, 2005). Research from an information processing perspective reveals that RJA and IJA may be associated with the activation of different frontal and parietal networks, as well as common cortical systems, in adults (Redcay et al., 2010; Schilbach et al., 2010), as well as children (Mundy et al., 2000). The observed neural network differences coincide with functional difference between the more self-referenced (egocentric) spatial and self-motivated (volitional) processes of IJA versus the more allocentric spatial referenced but less self-motivated, involuntary, responsive processes of RJA (Butterworth and Jarrett, 1991; Mundy, 2003;

Schilbach et al., 2010). Hence, parallel and distributed processing model raises the hypothesis that if differences between IJA and RJA reflect substantial differences in the degree to which these types of JA are associated with self-referenced processing then the experience of IJA and RJA may be expected to impact encoding, memory, and learning differently (Mundy and Jarrold, 2010). This is because a long-standing literature suggests that processing information under-self-referenced conditions promotes organization, elaboration, and encoding of information that is generally superior to comparative other referenced conditions (Symons and Johnson, 1997).

Research on JA on adult cognition (e.g., Frischen and Tipper, 2004; Bockler et al., 2011; Linderman et al., 2011) and cognitive neuroscience (e.g., Williams et al., 2005; Redcay et al., 2010; Schilbach et al., 2010) are on the frontier of social-cognitive science. However, the future of such work may benefit from advances in methodology. Aside from two groundbreaking imaging studies (Redcay et al., 2010; Schilbach et al., 2010) cognitive studies of JA in adults have been limited to examinations of phenomenon associated with RJA (e.g., Williams et al., 2005; Frischen et al., 2007). On the other hand, prior research and JA theory strongly suggest it may be illuminating, albeit methodologically more challenging, to compare and contrast RJA and IJA measures in the study of JA and cognitive processing in adults (Redcay et al., 2010; Schilbach et al., 2010).

#### **AIMS OF THE CURRENT STUDY**

To contribute to the advancement of research on social attention coordination in adults this study was designed to test a new JA paradigm that employs Virtual reality (VR) and eye-tracking methods to emulate both IJA and RJA. VR platforms offer the opportunity to develop paradigms for the study of social processes that are at once well controlled, yet ecologically valid (e.g., Kim et al., 2007, 2010). In this study we employed a VR paradigm to test the hypothesis that differences in processes associated with IJA and RJA may impact encoding, memory, and learning differently in adults. This hypothesis was examined by comparing the impact of RJA versus IJA on information processing in an adult picture recognition memory task. A picture memory task was chosen because prior research with children and adult picture recognition memory paradigms are sensitive to self-referenced processing effects (Craik et al., 1999; Henderson et al., 2009), and the effects of JA on stimulus encoding (Bayliss et al., 2006; Striano et al., 2006a,b). Comparative data from these two conditions provided data to address the hypothesis that IJA and RJA reflect discrete as well as common social neurocognitive processes (Mundy et al., 2009; Redcay et al., 2010; Schilbach et al., 2010) and, therefore, may be associated with distinctive effects on information processing (Mundy and Jarrold, 2010), in this case picture encoding and memory.

In addition, a standardized measure of spatial memory was included in this study to determine if individual differences in general mnemonic abilities affected performance on the JA tasks in this study. A *spatial working memory task* was selected for this purpose because research and theory have related spatial processing to JA (Butterworth and Jarrett, 1991; Bockler et al., 2011). However, to our knowledge no studies have been reported that

directly examined the hypothesized association between JA task performance and spatial ability. Too little is currently known about the empirical relations of spatial processing and joint attention allow for more than the *a priori* null hypothesis that encoding under IJA and RJA conditions would be equally associated with spatial working memory in this study. Finally, three types of pictorial stimuli were used; faces, buildings, and abstract patterns to provide evaluation of whether JA tasks had stimulus general or stimulus specific effects on information processing.

## STUDY 1

### MATERIALS AND METHODS

#### Participants

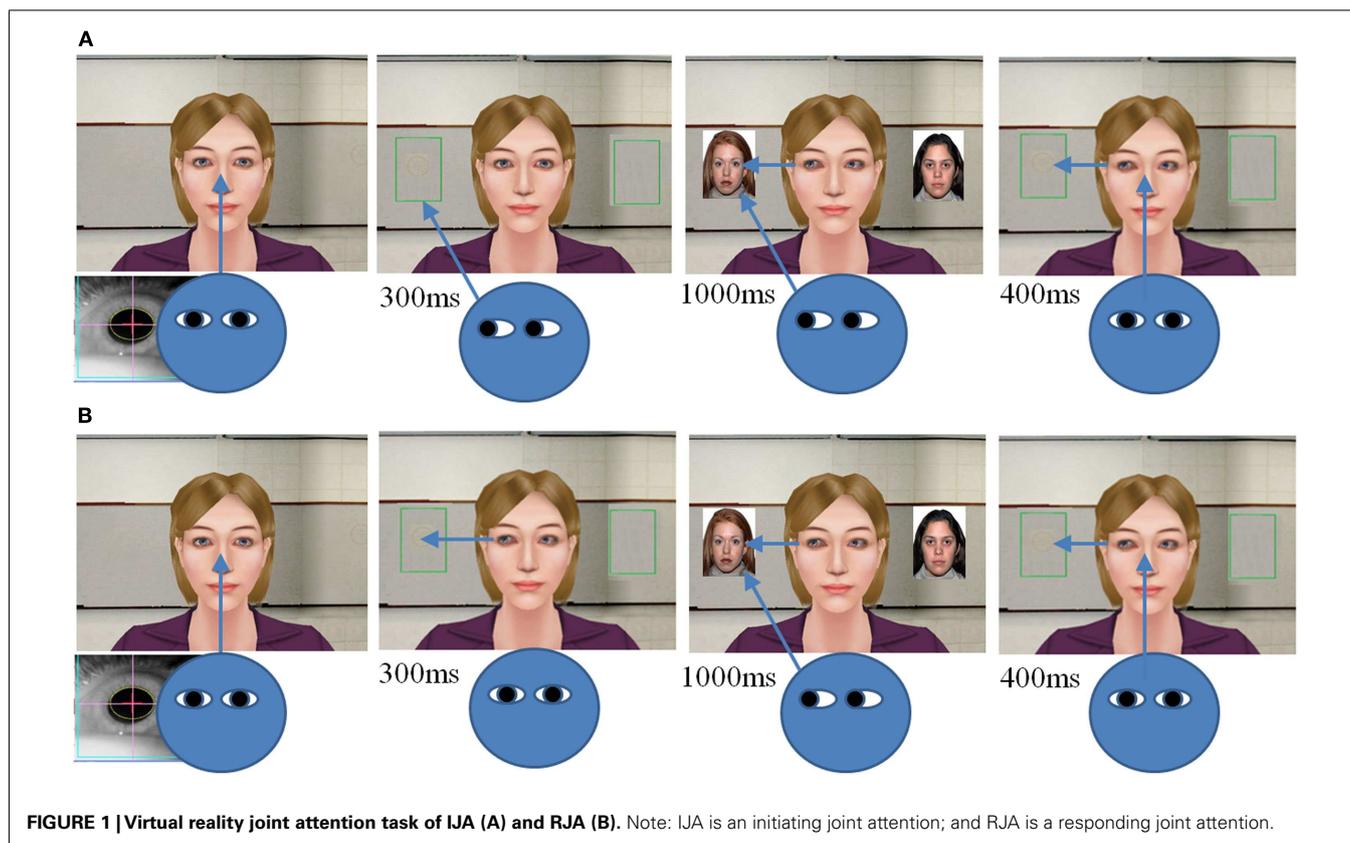
This protocol was approved by the University of California at Davis (UC Davis) Institutional Review Board prior to recruitment. Participants were students at UC Davis recruited through the Department of Psychology Research Participation System. Thirty-three participants consented to participate (76% female;  $n = 25$ ). Participants' self-reported ethnicities included: Asian/Pacific Islander ( $n = 21$ ; 63.6%), Caucasian ( $n = 6$ ; 18.2%), Hispanic-American ( $n = 5$ ; 13.2%), and Other ( $n = 1$ ; 3.0%). No participants reported currently being prescribed any psychiatric medications.

#### Joint attention task

**VR-JA task.** In the present study, we designed a VR paradigm that integrates eye-tracking and virtual avatar technologies to measure human JA. There were two JA conditions in the current study: IJA and RJA conditions. In both conditions, participants were

presented with an image of the upper body and face of a female avatar (**Figure 1**) and participants were asked to study pictures of houses, faces, or abstract designs that appear to the right and left of the avatar. The RJA and IJA tasks differed in how the participants chose or were directed to view pictures that appeared to the right and left of the avatar on each discrete trial. On each trial in the RJA condition, participants were directed to fixate a small "+" that appeared between the avatar's eyes. They were instructed to follow the avatar's gaze shift to the view the left-hand or right-hand picture (**Figure 1A**). The duration of avatar gaze shifts was 300 ms. After 300 ms different pictures then appeared to the left and right sides of the avatar for 1000 ms, and the participant viewed the target that was the focus of the avatars gaze. After 1000 ms the pictures disappeared and participants were directed to return to midline. The avatar's gaze was maintained to the picture location for an additional 400 ms to enable the participants to observe that the avatar had been sharing attention with the participant for the entire trial (**Figure 1B**).

In the IJA condition, participants were instructed to fixate the avatar and then to choose to look to the left or right of the avatar to view a picture. By way of eye-tracker feedback the VR software then triggered the avatar to shift her gaze within 300 ms to the region of interest (ROI) defined by the line of regard of the participants. Thus, in this condition the avatar followed the gaze shift of the participants. Identical to the RJA condition pictures to be studied appeared to the avatar's left and right sides for 1000 ms. After the pictures disappeared participants were requested to immediately return their gaze to midline, but the avatar's gaze remained to the



left or right for 400 ms to insure participants were aware that the avatar had followed their gaze (**Figure 1A**). All participants were presented with four blocks of IJA and RJA conditions, and each block consisted of 12 learning trials (4 trials of face, house, and abstract stimuli). The order of the condition was counterbalanced across participants. During the learning phase, participants were asked to remember as many pictures as possible. The pictures used in each condition (IJA and RJA) are available upon request from the authors. The order of avatar's gazing direction and the order of pictures were counterbalanced, and the inter task interval was jittered from 1 to 3 s.

After each block of IJA or RJA trials, participants were presented with a set of familiar and novel pictures from the same three categories (faces, houses, and abstract designs). Each test phase contained 36 test pictures consisting of: (a) 12 viewed in the JA task with accompanying avatar directed gaze (target pictures), 12 pictures presented that were not designated for viewing (non-target pictures) by the avatar gaze direction in the RJA condition, or the participants in IJA condition (non-target pictures) and 12 pictures that had not been presented as target or non-target pictures (novel pictures). The order of target, non-target, and novel test trials was counterbalanced across participants. Thus, each participant was presented with two blocks of 12 pictures to study in the IJA Condition (24 pictures) and the RJA condition (24 pictures). They were also tested for picture recognition of the 24 target pictures versus 24 non-target pictures and 24 novel pictures.

**Hardware and software.** The VR paradigm for this study was created using a 3D development platform (Vizard 3.0; WorldViz, Santa Barbara, CA, USA). This paradigm developed with this version of Vizard are implemented with a mono-head mounted display system (HMD; Z800 3DVisor, eMagin; **Figure 2**) with a 40° field of view system (OLED displays are 0.59 inch diagonal, and it is equivalent of a 105" screen at 12 feet). During the experiment adjustable head-bands were used to fit the HMD on a participant's head. An infrared eye-tracker (Arrington Research, Inc.) attached to the bottom of the right video screen of the HMD recorded participants right eye movement relative to visual HMD stimulus presentations (spatial resolution: 0.15°; temporal resolution: 60 Hz). The system was controlled by a desktop workstation running Windows XP (Microsoft) equipped with a high-end graphics card (nVidia). A standard keyboard was also used to record participants' responses. A sequence of 16 visual fixations points that covered the visual field of the HMD video monitors was presented to calibration each participant's real-time eye position data. The individual calibration data was saved with a time stamp (Arrington Research, Inc.) for each participant and integrated with the Vizard (WorldViz) to enable participant eye movement to trigger avatar gaze shifts in a specified fashion in the IJA condition. **Figure 2** illustrates a schematic diagram of this program.

The major dependent measures in the VR-JA picture recognition task were the *hit* (target pictures viewed in the learning phase and later correctly recognized in the testing phase) and the *false alarms* (errors of commission where participants incorrectly identified a novel picture in test trials as one viewed in the study trials; Snodgrass and Corwin, 1988). The percentage of correct identification (hit) and the errors of commission (false alarms) across

blocks were compared between IJA and RJA conditions. A third variable was the hit rate for pictures that were presented during study trials, but were not indicated by Avatar gaze (RJA condition), or chosen by participants in the IJA condition.

**Eye-tracking data.** Participant eye-tracking data were analyzed within each block of IJA and RJA study trials. Although, the automated paradigm presented all stimuli for 1000 ms of viewing time, the actual viewing time (stimulus fixation) for each participant could vary. To address this issue, we calculated the total viewing time on each stimulus presentation trail using the eye-tracking data and log-file information of each participant. Matlab 7.1 (MathWorks, Inc., Natick, MA, USA) was used to calculate the trial viewing times, and the ROI for this analysis included the left or right target areas. The eye-tracking data was stored unsmoothed, and if data was out of threshold in  $x$ ,  $y$  points (i.e., participant blink their eyes) it counted as a null value. The *total viewing time across trials in the two IJA and RJA blocks* were compared.

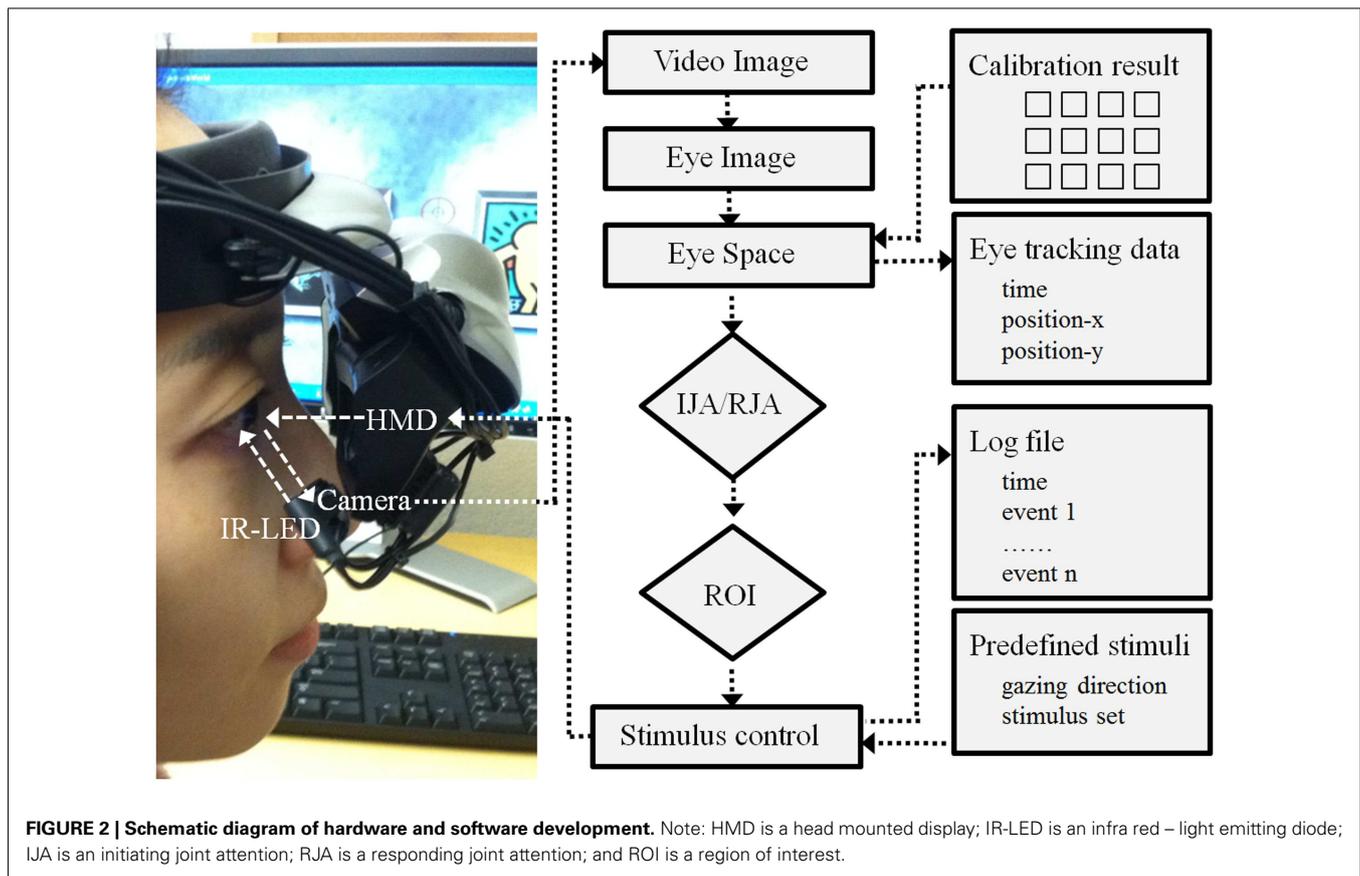
#### **Standardized spatial working memory task**

The *Finger Windows* task from the Wide Range Assessment of Memory and Learning, 2nd Edition (WRAML-2, Sheslow and Adams, 2003) was administered to participants prior to presentation of the VR-JA task. This task provided a measure of each participant's ability to encode and remember spatial-sequential information. This task provided a measure of visual spatial working memory (Sheslow and Adams, 2003). In the *Finger Windows task*, an examiner presented participants with a card depicting an array of holes or windows. The examiner presented a sequence of increasingly difficult trials that involved touching the end of a pencil to a sequence of windows. Participants were asked to place his/her finger in the same sequence (correct order) of windows. The examiner presented a set of sequences guided by the form until participants made errors attempting to replicate three consecutive sequences. The total number of correct sequences modeled constituted the total score for this task. The scores of each task were converted to standard scores following the WRAML-2 manual (Sheslow and Adams, 2003).

#### **Procedure**

Upon arrival at the research site, participating adults were informed about the nature of the research and they provided signed consent to participate according to the university approved IRB protocol for the study. They were then asked to complete a brief questionnaire to gather data on age, gender, ethnicity, medical status (e.g., current medications), and the *Finger Window tasks* of WRAML-2. The experimenter then assisted the participants with their head-placement of the HMD. The infrared camera for eye-tracking was adjusted to participant's eye position and calibrated (**Figure 2**). Participants then completed a practice block of the VR-JA tasks during which they practiced following the avatar's gaze on RJA trials and directing the avatar's gaze with the participants' line of regard on IJA trials. HMD and eye-tracking set-up, as well as VR-JA practice trials required ~10–15 min per participant.

Following the practice trials the participants were presented with two blocks of the experimental JA picture study conditions (IJA and RJA) in a within-subject experimental design. The order



of the IJA and RJA conditions within blocks was counterbalanced across all participants to control for order effects (Figure 3). The participants interacted with the virtual avatar under IJA or RJA condition, and were asked to remember as many pictures as they could during each block. Thirty-six recognition test trials were presented immediately after each block of the four study trials and participants were asked to identify pictures they had seen before in the previous block of study trials. Before each block, including the practice block, the eye-tracker was calibrated for each participant to ensure accurate tracking. During the calibration, participants looked toward 16 predefined points on the HMD screen. After each block, participants rested for ~5 min in order to prevent fatigue effects. After completion of all blocks, participants were debriefed about the purpose of the study. The experiment lasted for 90–120 min.

## RESULTS

The two variables computed for this study were the *percent of correct identifications* (hits) and the *false positive identifications* (false alarms) in the recognition trials. These variables were calculated separately for the JA target pictures and the novel pictures. The percentages of correct identification and false positive identification of pictures were averaged within the two blocks of trials for IJA and RJA conditions. The correct and false positive identification data were transformed according to signal detection theory (Stanislaw and Todorov, 1999; Schulz et al., 2007) such that: (1) A  $\beta$  index was computed to measure response bias or the general

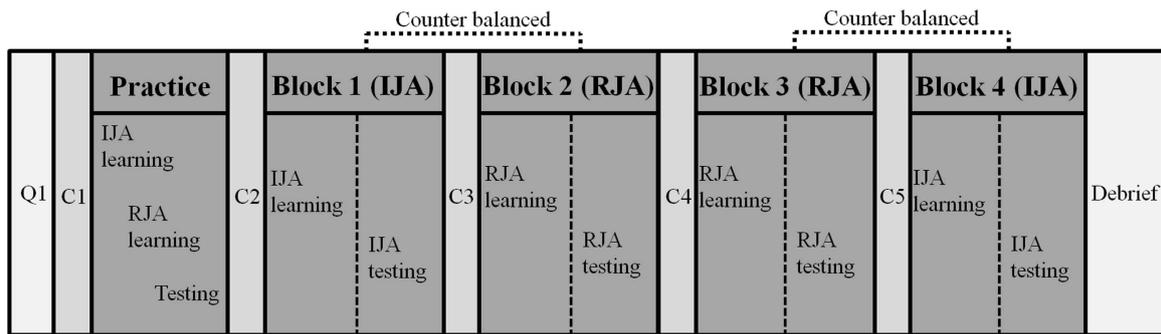
tendency to respond yes or no during IJA or RJA condition picture recognition tests; and (2) A  $d$ -prime ( $d'$ ) index was computed to measure response sensitivity (correct picture recognition) in each JA condition unaffected by response bias. The  $d$ -prime index was the main dependent variable for analysis in this study. Nevertheless, data on  $\beta$  index were also presented for relevant data analyses.

### The effects of IJA and RJA on recognition memory

A two-way (IJA versus RJA stimulus conditions) within ANOVA was conducted for  $d$ -prime. The results indicated that the  $d$ -prime differed significantly between the two conditions,  $F(1, 32) = 43.01$ ,  $p < 0.001$ ,  $\eta^2 = 0.57$ . As shown in the Table 1, participants correctly identified more pictures that they had viewed in the IJA condition ( $M = 66.9$ ;  $SD = 16.9$ ) than in the RJA condition ( $M = 57.8$ ;  $SD = 16.4$ ). A correlation analysis also indicated that there was significant consistency in the pattern of individual differences in  $d$ -prime displayed by participants across the IJA and RJA conditions,  $r(32) = 0.67$ ,  $p < 0.001$ .

A second two-way (IJA versus RJA stimulus conditions) within ANOVA revealed that participants did not display reliable differences on  $\beta$  ( $\beta$ ), or response specificity, across the IJA and the RJA conditions,  $p > 0.54$  (see Table 1).

In each trial in the IJA and RJA conditions pictures were presented that were not the target of shared attention with the avatar, but could be viewed by the participants. Since they were not targeted we expected little evidence of memory for these pictures and



**FIGURE 3 | Experimental procedure.** Note: Q1 included an informed consent form, a personal history form, a picture memory task, a finger window task; C1, C2, C3, C4, and C5 included a calibration of each participant to ensure accurate eye-tracking.

**Table 1 | The effects of IJA and RJA on mean memory recognition memory (SD in parentheses).**

	Average percentage of correct identification	Average percentage of false identification	<i>d</i> -Prime	Beta
IJA	66.9 (16.9)	12.4 (10.0)	1.77 (0.76)	0.37 (0.35)
RJA	57.8 (16.4)	21.2 (13.9)	1.11 (0.65)	0.34 (0.36)

IJA, initiating joint attention; RJA, responding joint attention.

no differences in recognition memory between the IJA and RJA conditions. These assumptions were confirmed. A two-way (IJA versus RJA stimulus conditions) within ANOVA for the *average percentage of non-target identification* revealed very low correct hit rates associated with both conditions and no difference across the two conditions, IJA ( $M = 26.8$ ;  $SD = 18.9$ ) and RJA ( $M = 27.4$ ;  $SD = 18.0$ ),  $p > 0.80$ .

#### The relations between spatial working memory ability and the VR-JA recognition memory

A correlation analysis was conducted to determine if spatial working memory was related to stimulus encoding and recognition memory in the IJA and RJA condition. The results indicated that performance on the *Finger Windows spatial memory task* was significantly correlated with the *d-prime* in the IJA condition,  $r(32) = 0.46$ ,  $p < 0.007$ , but not in the RJA condition,  $r(32) = 0.25$ ,  $p > 0.15$ . However, there was no evidence of a significant difference in the comparison of the magnitude of these correlations.

#### Viewing time variability in the IJA and RJA condition

In Study 1 stimuli were presented for 1000 ms in each trial, but participants did not necessarily view (fixate) stimuli across the entire presentation time. Consequently, actual *viewing time* may have differed across the IJA and RJA conditions. We examined this possibility in a sequence of analyses. The average trial study time in the IJA and RJA conditions was calculated from eye-tracking data for each participant. Two of the 33 participants had missing data on this variable because of data saving errors of the eye-tracker. A two-way (IJA versus RJA stimulus conditions) within ANOVA

for *viewing time* revealed that participants spent longer time to looking at the stimulus in the IJA condition ( $M = 882$ ;  $SD = 111$ ) than in the RJA condition ( $M = 600$ ;  $SD = 85$ ),  $F(1, 30) = 233.4$ ,  $p < 0.001$ ,  $\eta^2 = 0.89$ . Correlation analyses also revealed that individual differences in viewing time were consistent across the IJA and RJA conditions [ $r(30) = 0.481$ ,  $p < 0.006$ ]. Finally, the correlation between *viewing time* and *d-prime for recognition memory* approached a conventional level of significance in the IJA condition:  $r(30) = 0.314$ ,  $p = 0.086$  and exceeded the 0.05 alpha criteria in the RJA condition:  $r(30) = 0.372$ ,  $p = 0.039$ .

#### Gender effects

The gender ratio in the current study was (76% female;  $n = 25$ ). Independent samples *t*-test were conducted to test differences between females and males in all dependent measures in the current study: the *average percentage of correct identification* in IJA and RJA conditions; the *average percentage of false positive identification* in IJA and RJA conditions; the *d-prime* in IJA and RJA conditions; the *viewing time* in IJA and RJA conditions; and the *memory abilities of participants*. No significant differences were found in any of the dependent measures (all  $ps > 0.05$ ).

#### Stimuli effects

Each test phase contained three types of stimuli (face, abstract designs, and house). The possible effects of stimuli were explored in a two-way (IJA and RJA stimulus conditions) by three-way (Face, Abstract, and House stimulus type) within ANOVA was computed for the *d-prime scores*. The results revealed significant main effects for JA Condition,  $F(1, 32) = 40.40$ ,  $p < 0.001$ ,  $\eta^2 = 0.56$ , and Stimulus Type,  $F(2, 64) = 3.33$ ,  $p < 0.05$ ,  $\eta^2 = 0.09$ . The interaction between JA Condition and Stimulus Type was also significant,  $F(2, 64) = 4.96$ ,  $p < 0.01$ ,  $\eta^2 = 0.13$ , such that JA effects were apparent on *abstract* and *house* stimuli  $t(32) = 4.73$ ,  $p < 0.001$ ,  $t(32) = 4.51$ ,  $p < 0.001$  respectively, but not for condition comparisons of the scores for *face* stimuli (see Table 2).

#### DISCUSSION

The results of Study 1 were consistent with the hypothesis that JA may affect information processing in adults, and that IJA and RJA may impact encoding and memory differently. The results were also consistent with hypothesis that spatial information processing

**Table 2 | The percent of hit and false alarm in the three types of stimuli.**

	IJA		RJA	
	Hit	False alarm	Hit	False alarm
Face	75.38 (19.39)	29.80 (18.81)	70.08 (20.48)	29.60 (17.96)
Abstract designs	66.59 (22.70)	11.74 (8.23)	51.14 (22.40)	15.53 (11.17)
House	58.33 (28.92)	14.06 (11.32)	53.03 (23.60)	26.58 (18.94)

IJA, initiating joint attention; RJA, responding joint attention.

may be involved in JA and its impact on information processing. However, the interpretation of the data from Study 1 was complicated by the observation that participants viewed pictures for less time in the RJA than IJA condition. This raised the possibility that differences in stimulus viewing time may have contributed to differences in recognition memory that were observed for the IJA and RJA encoding conditions. To examine this possibility, and to provide data from a second independent sample on JA and recognition memory in adults, a second study was conducted with a modified paradigm that provided improved control of picture viewing time across the IJA and RJA conditions.

## STUDY 2

### MATERIALS AND METHODS

To address the need to control viewing time, the RJA condition of the VR paradigm was modified. In Study 1 pictures were presented 300 ms after the avatar shifted gaze. However, participants varied in their response latency to avatar gaze shifts, and those with longer latencies had less opportunity to view the picture stimuli during the 1000 ms stimulus presentation interval. To control for this source of variability in Study 2 picture presentation was yoked to participant gaze shifts in the RJA condition. Pictures appeared after participants shifted their gaze to the correct part of the stimulus field (Figure 1A, green rectangles) in response to the spatial eye direction cue of the Avatar. On each trial in the RJA condition, participants were again directed to fixate a small “+” that appeared between the avatar’s eyes and were instructed to follow the avatar’s gaze shift to the left or right. After the participants shifted their gaze to follow the avatar’s gaze direction, the pictures then appeared to the left and right sides of the avatar for 1000 ms. After 1000 ms the pictures disappeared and participants were requested to return to midline, and the avatar’s gaze was maintained to the pictures location for 400 ms. This matched the viewing opportunity in the IJA condition where, as in Study 1, pictures were presented for 1000 ms after the participant to look left or right of the avatar on a given trial.

### Participants and experiment design

The participants were a new sample of students at UC Davis recruited through the Department of Psychology Research Participation System. Twenty-six participants consented to participate (58% female;  $n = 15$ ). Participants’ self-reported ethnicities included: Asian/Pacific Islander ( $n = 11$ ; 39.3%), Caucasian ( $n = 10$ ; 35.7%), Hispanic-American ( $n = 4$ ; 14.3%), and Other ( $n = 1$ ; 3.6%). No participants reported currently being prescribed

**Table 3 | The actual viewing time (ms) in Experiment 1 and 2.**

	IJA viewing time	RJA viewing time (ms)
Experiment 1	882 (111)	600 (85)
Experiment 2	900 (110)	940 (83)

IJA, initiating joint attention; RJA, responding joint attention.

any psychiatric medications. Other than the change in the RJA condition methods, the procedures, measures, and experimental design were exactly same as those described for Study 1.

## RESULTS

### Viewing time variability in the IJA and RJA condition

The average trial viewing time in the IJA and RJA conditions was calculated from eye-tracking data for each participant. Out of the 26 participants one had missing data on this variable because of data saving error of the eye-tracker. A two-way (IJA versus RJA condition) ANOVA was conducted to test for differences in the viewing time. Viewing time did not differ across the IJA condition ( $M = 900$ ;  $SD = 110$ ) and the RJA condition ( $M = 940$ ;  $SD = 83$ ),  $p > 0.10$ . As expected, however, the change in RJA condition methods resulted in significantly greater viewing time in the RJA condition in Study 2 than in Study 1,  $t(54) = 15.03$ ,  $p < 0.01$  (see Table 3). There was no difference between the viewing times for the IJA condition in Study 2 versus Study 1 (see Table 3). As was the case in Study 1, correlation analyses again indicated that there was a significant positive correlation between participants actual viewing times in the IJA and RJA conditions,  $r(25) = 0.465$ ,  $p < 0.02$ , in Study 2.

### The effects of IJA and RJA on recognition memory

A two-way (IJA versus RJA condition) within ANOVA was conducted to test for differences in  $d$ -prime between IJA and RJA. Consistent with data from Study 1, the results indicated that the participants displayed higher  $d$ -prime for recognition memory for pictures viewed in the IJA condition rather than the RJA condition,  $F(1, 25) = 7.16$ ,  $p < 0.013$ ,  $\eta^2 = 0.22$  (see Table 4). Also consistent with Study 1, a two-way (IJA versus RJA conditions) ANOVA failed to detect any condition effects for  $\beta$  indicating that participants displayed no differences in specificity, or errors of commission, after viewing pictures in the RJA and IJA conditions,  $p > 0.75$  (Table 4). Thus, after controlling for the possible effect of viewing time analyses of the difference between the effects IJA and RJA conditions on recognition memory was significant in the sample in Study 2. However, the estimate of effect sizes associated with the difference between the IJA and RJA conditions was smaller in Study 2,  $\eta^2 = 0.22$ , than in Study 1,  $\eta^2 = 0.57$ .

Correlation analyses were computed to examine the consistency of individual differences in memory performance across the two JA conditions. A positive correlation between  $d$ -prime in the IJA and RJA conditions was observed that approached a conventional level of significance ( $r = 0.37$ ,  $p \leq 0.08$ , one-tailed analysis). The actual viewing time in IJA condition was correlated with  $d$ -prime recognition memory measure in IJA condition,  $r(25) = 0.54$ ,  $p = 0.005$ , but the actual viewing time in the RJA condition was not correlated

**Table 4 | The effects of IJA and RJA on memory recognition in the Experiment 2.**

	Average percentage of correct identification	Average percentage of false identification	<i>d</i> -Prime	Beta
IJA	68.6 (8.4)	11.5 (10.6)	1.80 (0.55)	0.37 (0.26)
RJA	59.5 (14.7)	16.0 (12.4)	1.39 (0.77)	0.39 (0.30)

IJA, initiating joint attention; RJA, responding joint attention.

with the respective *d*-prime recognition memory measure in that condition,  $r(25) = 0.21$ ,  $p > 0.10$ .

A two-way (IJA versus RJA conditions) within ANOVA for differences in *non-target identification* (pictures presented but not viewed with the avatar) revealed no significant differences between the JA conditions: IJA ( $M = 24.0$ ;  $SD = 18.6$ ) and RJA conditions ( $M = 20.8$ ;  $SD = 16.2$ ),  $p > 0.10$ .

### THE RELATION BETWEEN SPATIAL WORKING MEMORY AND THE VR-JA RECOGNITION MEMORY

A correlation analysis was again conducted to determine if performance on a spatial working memory task was associated with memory in the JA conditions. The results indicated that association between performance on the Finger Windows spatial memory task and *d*-prime in the IJA condition approached a conventional level of significance,  $r(25) = 0.35$ ,  $p < 0.08$ , but this was not the case in the RJA condition,  $r(25) = 0.15$ ,  $p = 0.24$ . The difference between the magnitudes of these correlations was not significant. However, the convergent pattern of a significant correlation between spatial working memory and IJA performance in Study 1 ( $r = 0.46$ ) and a correlation that approached significance, with no evidence of a reliable association with RJA, suggests that spatial working memory may be a more consistent or stronger correlate of IJA than RJA process in adults.

### Gender effects

The sample for Study 2 was more balanced with regard to gender (58–42% female to male ratio) than in Study 1 (75–25% female to male ratio). However, no significant gender effects were observed with any of the dependent measures in Study 2 (all  $ps > 0.10$ ).

### Stimuli effects

Finally, the possible effect of stimulus types in Study 2 was again examined. A 2 (IJA versus RJA condition) by 3 (Face, Abstract, or House stimuli) ANOVA was performed for the *d*-prime scores. This revealed significant main effects for JA condition,  $F(1, 25) = 8.35$ ,  $p < 0.008$ ,  $\eta^2 = 0.25$ , and a marginally significant effect for stimuli,  $F(1, 25) = 2.97$ ,  $p = 0.06$ ,  $\eta^2 = 0.11$ . In addition, the interaction effect between JA and stimuli was significant,  $F(2, 50) = 3.63$ ,  $p < 0.04$ ,  $\eta^2 = 0.13$ . The pattern of data associated with this interaction was the same as the one observed in Study 1. Specifically, IJA and RJA *d*-prime scores differed for abstract and house stimuli,  $t(25) = 3.15$ ,  $p < 0.004$ ,  $t(25) = 3.12$ ,  $p < 0.004$  respectively, but not for the for face stimuli,  $p > 0.90$  (see Table 5).

**Table 5 | The percent of hit and false alarm in the three types of stimuli in the second experiment.**

	IJA		RJA	
	Hit	False alarm	Hit	False alarm
Face	73.56 (15.14)	29.78 (19.13)	74.03 (21.19)	31.60 (21.06)
Abstract designs	68.75 (16.68)	12.50 (11.18)	51.76 (22.40)	13.94 (10.50)
House	63.46 (16.17)	12.16 (13.13)	53.66 (23.43)	17.60 (13.08)

IJA, initiating joint attention; RJA, responding joint attention.

## DISCUSSION

The study of JA has long been associated with research and theory on social-cognition. In developmental science social-cognition has often been defined singularly in terms measures that signify the degree to which an individual can infer the intentions, beliefs, or emotions of another person. However, this is but one operational definition of social-cognition. Another perspective suggests that one vital element of social-cognition may reflect the effects that social attention coordination may have on human information processing (Mundy et al., 2009). This perspective has emerged, in part, from a literature that suggests that JA affects encoding of pictures and words in infancy (Striano et al., 2006a,b; Hirotani et al., 2009; Kopp and Lindenberger, 2011). It is also supported by the results of studies with adults that have reported that JA effects spatial information processing, attribution of stimulus valence, and resistance to IOR in studies of visual orienting (Frischen and Tipper, 2004; Bayliss et al., 2006; Bockler et al., 2011; Linderman et al., 2011).

In our opinion these are ground breaking studies. They open up a new perception of the role of JA in human development. Accordingly, JA is not only a faculty of mind that is vital to developing an understanding of the minds of other people, but also one that may play a vital role in the phylogenetic and ontogenetic advancements of the human faculties for learning and memory. To encourage sustained research in this vein we undertook the development and testing of a paradigm that could be used in controlled studies of the effects of different types of JA on adult cognitive processes. The initial data from this paradigm were promising.

Previous research has indicated that RJA and IJA appear to tap divergent as well as convergent processes in learning and development in young children (e.g., Mundy et al., 2007; Vaughan Van Hecke et al., 2007; Meltzoff and Brooks, 2008), and – RJA and IJA are associated with distinct and common neural networks in adults (Redcay et al., 2010; Schilbach et al., 2010). The two studies reported here on the effects of JA on adult picture encoding provide data that are consistent with this previous pattern of findings. When participants directed the gaze of an avatar-social-partner and shared attention to pictures (IJA condition) their recognition memory for the pictures was enhanced compared to when they followed the gaze of an avatar and shared attention to pictures (RJA condition). Hence, encoding, or depth of processing of the picture stimuli was facilitated by the experience of initiating shared attention, rather than responding to the gaze direction of others, in virtual social interactions. In Study 1 differences in picture viewing

time across the IJA and RJA likely contributed to the effects of viewing condition on recognition memory. However, the revised methods employed in Study II demonstrated that, participants displayed significantly better picture memory (stimulus encoding) in the IJA rather than RJA condition, even when picture viewing time was better controlled and comparable across JA conditions.

The nature of the factors that distinguish information processing during IJA and RJA are not yet clear. One possibility emerged from testing the hypothesis that JA involves, or affects spatial information processing skills (Butterworth and Jarrett, 1991; Bockler et al., 2011). As expected, encoding in the context of JA was associated with an independent standardized measure of visual spatial working memory performance in adults. However, unexpectedly, evidence of this association was observed only relative to encoding in the IJA condition in both Study 1 and 2. The current limits of our knowledge of JA do not allow for a definitive interpretation of these observations. Nevertheless, self-centered (egocentric) spatial information processing has been observed to trigger different episodic information processing relative to other referenced (allocentric) spatial information processing (Gomez et al., 2009). *Post hoc*, it may well be that differences in egocentric and allocentric spatial processing are associated with IJA and RJA respectively, and this distinguishing cognitive characteristic contributed to the differences in IJA and RJA encoding observed in this study. Of course, in these studies the spatial working memory measure was presented prior to the VR paradigm. It is possible then that testing spatial working memory somehow primed encoding in the IJA condition. However, even if this were true, spatial information processing would still appear to have had a selectively stronger effects on IJA versus RJA in this study. The estimates of these effect sizes, though, were modest, 0.46,  $r^2 = 0.21$  and 0.35,  $r^2 = 0.12$ , in Study 1 and 2 respectively. Therefore, it is important to consider other factors that may contribute to the differential impact of the experience of IJA versus RJA on stimulus encoding.

Initiating JA may benefit information processing because it involves greater self-referenced processing than RJA (Mundy and Jarrold, 2010). Self-referenced processing refers to implicit, subjective, and pre-reflective processing and integration of information from one's own body (e.g., heart rate, volitional muscle movement) with perceptual and cognitive activity, such as maintenance of goal related intentions in working or integrating perceptual input with information from long term memory (Northoff et al., 2006). The nature of self-referenced information processing specific to JA may be more precisely described in terms of: (a) proprioception, such as feedback from ocular muscle control and the vestibular system related to the spatial direction of one's own visual attention and head posture (see Butterworth and Jarrett, 1991 for relate discussion), and (b) interoception including, information about arousal and the positive (rewarding), neutral, or negative valence of self perception of the object or event, as well as the valence of sharing attention with a social partner.

Self-referenced processing may facilitate encoding through one of several mechanisms. Craik et al. (1999) suggested that self-reference processing triggers an extensive frontal network involved in re-presentations of one's own identity and this provides a rich matrix of associative encoding opportunities that increase

the likelihood of deep and efficient stimulus encoding. Second, self-referenced spatial processing may involve different networks specific to the role of hippocampus in memory. Self-referenced processing may involve left hippocampal activation, whereas other referenced spatial processing (TJA) may involve right hippocampal activation to a greater extent (Burgess et al., 2002). Thirdly, Gilboa (2004) suggest that self and other referenced processing may be associated with differences in retrieval. Self-referenced processing may yield encoding and memory that "relies on quick intuitive 'feeling of rightness' to monitor the veracity and cohesiveness of retrieved memories" versus other referenced episodic encoding that "require more conscious elaborate monitoring to avoid omissions, commissions, and repetitions" (Gilboa, 2004).

A fourth is that the self-referenced processing of IJA more volitional and intentional in nature than is the more reactive, reflexive, or involuntary processing associated with RJA (Friesen and Kingstone, 1998; Mundy, 2003). Consequently, IJA and RJA may be distinguished by the degree that they are associated with motivation/reward system activation that distinguishes intentional self-generated goal related action and reflexive goal related action (Mundy, 1995). Consistent with this possibility Schilbach et al. (2010) observed that IJA was associated with activation of neural reward circuitry of the ventral striatum (bilaterally) in adults than was RJA. Since neural reward circuit activation plays a role in facilitating encoding (Holroyd and Coles, 2002) ventral striatal activation in may have played a role in facilitating encoding in the IJA condition in this study. In this regard it is important to note that the type of reflexive social-orienting involved in RJA may be more effortful or place greater task related cognitive demands on participants than IJA. This notion is supported by recent research that the developmental of RJA is associated with the effortful control of attention (Vaughan Van Hecke et al., 2012). It may well be then, that the results of this study reflect an RJA executive impediment to rapid encoding of briefly presented stimuli, as much as they reflect an encoding enhancement of IJA.

Of course *hypotheses* also raises the possibility that the effects reported here are not specific to encoding in the context of social attention coordination, but rather reflect a general effect of whether or not participants are free to choose stimuli to view or are directed to view stimuli before encoding. This calls for a future study comparing IJA and RJA encoding effects with analogous no-social volitional versus directed stimulus encoding conditions.

Concrete test of each of these hypotheses could be provided with fMRI studies of the cortical correlates of stimulus encoding in conditions that emulate the experience of IJA and RJA. Indeed, part of the value of the data from this study is that they converge with those of imaging studies of Redcay et al. (2010) and Schilbach et al. (2010) to suggest that the use of virtual emulations of the different types of JA experience (initiating versus responding) provides a valid means for the more precise scientific examination social attention coordination processes, and their cognitive and neurocognitive sequelae in adults. VR applications to the study of the effects of JA on information processing may also be useful in research on forms of developmental pathology characterized by impairments in social attention coordination and social-cognition, such as Autism. While the scientific literature on

Autism is replete with important studies of the role of JA in early development and intervention (i.e., Kasari et al., 2006; Mundy et al., 2009) few research tools have previously been developed to facilitate the study of the role of JA in learning and development in school aged children and adults with ASD.

With regard to the latter it may be important to recognize that JA encoding affects were only apparent in both studies for picture of buildings and abstract patterns, but not for picture of faces. Moreover, participants displayed better face recognition than building or abstract pattern recognition across all conditions in both studies. The later observation was consistent with a literature that indicates that encoding of faces is often supported by neural “expertise” systems involving fusiform networks that are not as consistently activated in encoding non-facial stimuli (e.g., Carey et al., 1992; Farah et al., 1998; De Hann et al., 2002; Gauthier et al., 2003). Presumably this expertise system serves as an executive function that enhances face encoding in many people. The data in Study 1 and 2 suggest that processes that are involved in JA effects on encoding and those involved in facial encoding may be distinguishable. Shared social attention may have its most discernible impact on encoding that require participants to engage in stimulus organization and depth of processing that is not well supported by previously acquired executive or expertise functions. Alternatively, researchers have often assumed that the processes leading to face processing and JA impairments in Autism Spectrum Disorders are highly related if not isomorphic (Schultz, 2005; Mundy et al., 2009). It may well be that the study of the impact of JA on face and non-face stimulus encoding in samples of affected individuals and comparison groups may shed new light on this prominent assumption in research on the nature of Autism.

## REFERENCES

- Bayliss, A., Paul, M., Cannon, P., and Tipper, S. (2006). Gaze cuing and affective judgments of objects. I like what you look at. *Psychon. Bull. Rev.* 13, 1061–1066.
- Bockler, A., Knoblich, G., and Sebanz, N. (2011). Giving a helping hand: effects of joint attention on mental rotation of body parts. *Exp. Brain Res.* 211, 531–545.
- Brooks, R., and Meltzoff, A. (2002). The importance of eyes: how infants interpret adult looking behavior. *Dev. Psychol.* 38, 958–966.
- Bruner, J. S. (1995). “From joint attention to the meeting of minds: an Introduction,” in *Joint Attention: Its Origins and Role in Development*, eds C. Moore and P. J. Dunham (Hillsdale, NJ: Lawrence Erlbaum), 1–14.
- Burgess, N., Maguire, E., and O’Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron* 35, 625–641.
- Butterworth, G., and Jarrett, N. (1991). What minds have in common is space: spatial mechanisms in serving joint visual attention in infancy. *Br. J. Dev. Psychol.* 9, 55–72.
- Carey, S., De Schonen, S., and Ellis, H. (1992). Becoming a face expert. *Phil. Trans. R. Soc. Biol. Sci.* 335, 95–103.
- Craik, F. I. M., Moroz, T. M., Moscovitch, M., Stuss, D. T., Winocur, G., Tulving, E., and Kapur, S. (1999). In search of the self: a positron emission tomography study. *Psychol. Sci.* 10, 26–34.
- De Hann, M., Humphreys, K., and Johnson, M. (2002). Developing a brain specialized for face processing: a converging methods approach. *Dev. Psychobiol.* 40, 200–212.
- Farah, M., Wilson, K., Drain, M., and Tanaka, J. (1998). What is special about face processing? *Psychol. Rev.* 105, 482–498.
- Friesen, C., and Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychon. Bull. Rev.* 5, 490–495.
- Frischen, A., Bayliss, A., and Tipper, S. (2007). Gaze cuing of attention: visual attention, social cognition and individual differences. *Psychol. Bull.* 133, 694–724.
- Frischen, A., and Tipper, S. (2004). Orienting attention via observed shift evokes longer term inhibitory effects: implications for social interactions, attention, and memory. *J. Exp. Psychol. Gen.* 133, 516–533.
- Gauthier, I., Curran, T., Curby, K., and Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nat. Neurosci.* 6, 428–432.
- Gilboa, A. (2004). Autobiographical and episodic memory—and and the same? Evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia* 42, 1336–1349.
- Gomez, A., Rousset, S., and Baciou, M. (2009). Egocentric-updating during navigation facilitates episodic memory retrieval. *Acta Psychol.* 132, 221–227.
- Grossman, T., and Johnson, M. (2010). Selective prefrontal cortex responses to joint attention in early infancy. *Biol. Lett.* 6, 540–543.
- Hebb, D. (1949). *The Organization of Behaviour*. New York, NY: Wiley.
- Henderson, H. A., Zahka, N. E., Kojkowski, N. M., Inge, A. P., Schwartz, C. B., Hileman, C. M., Coman, D. C., and Mundy, P. C. (2009). Self-referenced memory, social cognition, and symptom presentation in autism. *J. Child Psychol. Psychiatry* 50, 853–861.
- Henderson, L., Yoder, P., Yale, M., and McDuffie, A. (2002). Getting the point: electrophysiological correlates of protodeclarative pointing. *Int. J. Dev. Neurosci.* 20, 449–458.
- Hirofani, M., Stets, M., Striano, T., and Friedericic, A. (2009). Joint attention helps infants learn new words: event related potential evidence. *Neuroreport* 20, 600–605.
- Holroyd, C., and Coles, M. (2002). The neural basis of human error processing: reinforcement learning, dopamine and the error related negativity. *Psychol. Rev.* 109, 679–709.
- Kasari, C., Freeman, S., and Paparella, T. (2006). Joint attention and symbolic play in young children with autism: a randomized controlled intervention study. *J. Child Psychol. Psychiatry* 47, 11–620.
- Keysers, C., and Perrett, D. (2006). Demystifying social-cognition: a Hebbian perspective. *Trends Cogn. Sci. (Regul. Ed.)* 8, 501–507.
- Kim, K., Kim, J., Park, D., Jang, H., Ku, J., Kim, C., Kim, I., and Kim, S. (2007). Characteristics of social perception assessed in schizophrenia using virtual reality. *Cyberpsychol. Behav.* 10, 215–219.

## STUDY LIMITATIONS

A major limitation of study is that not enough is known about JA or the differences between IJA and RJA processes, or their impact on information processing, to go beyond *post hoc* explanations in the discussion of the results of this study. So, while the results added to weight of evidence that JA does indeed affect adult information processing (Frischen and Tipper, 2004; Bayliss et al., 2006; Bockler et al., 2011; Linderman et al., 2011), the nature of the specific processes involved remain a vital and open topic for a new generation of research. We have, perhaps, all too blatantly exceeded the limits of the data in discussing four alternative hypothetical factors that may be involved in JA and the differential effects of IJA and RJA on cognitive processing. However, we would hasten to add that we believe that each of these alternatives is readily open to empirical examination in future experimental cognitive or neurocognitive (e.g., imaging) studies using variations of the paradigm described in this paper. We hope the results contribute to new methods, new questions and renewed enthusiasm for inquiry into JA as a major yet poorly understood facility of the human mind, while recognizing the clear limits of the empirical contribution of this initial study in our planned program of research.

## ACKNOWLEDGMENTS

This work was supported by National Institutes of Health (USA) Grant 1R21MH085904 – 01, the University of California at Davis Department of Psychiatry Lisa Capps Endowment for Research on Neurodevelopmental Disorders, and the National Research Foundation of Korea Grant funded by the Korean Government (Ministry of Education, Science and Technology; NRF-2011-357-E00041).

- Kim, K., Kim, S., Cha, K., Park, J., Rosenthal, M. Z., Kim, J., Han, K., Kim, I., and Kim, C. (2010). Development of a computer-based behavioral assessment of checking behavior in obsessive-compulsive disorder (2010). *Compr. Psychiatry* 55, 86–93.
- Kopp, F., and Lindenberger, U. (2011). Effects of joint attention on long-term memory in infants: and event-related potentials study. *Dev. Sci.* 14, 660–672.
- Linderman, O., Pines, N., Rueschemeyer, S., and Bekkering, H. (2011). Grasping the others's attention: the role of animacy in action cuing of joint attention. *Vision Res.* 51, 940–944.
- Meltzoff, A., and Brooks, R. (2008). Self experiences a mechanism for learning about others: a training study in social cognition. *Dev. Psychol.* 44, 1–9.
- Munakata, Y., and McClelland, J. (2003). Connectionist models of development. *Dev. Sci.* 6, 413–429.
- Mundy, P. (1995). Joint attention and social-emotional approach behavior in children with autism. *Dev. Psychopathol.* 7, 63–82.
- Mundy, P. (2003). The neural basis of social impairments in autism: the role of the dorsal medial-frontal cortex and anterior cingulate system. *J. Child Psychol. Psychiatry* 44, 793–809.
- Mundy, P., Block, J., Vaughan Van Hecke, A., Delgado, C., Venezia Parlade, M., and Pomares, Y. (2007). Individual differences and the development of infant joint attention. *Child Dev.* 78, 938–954.
- Mundy, P., Card, J., and Fox, N. (2000). EEG correlates of the development of infant joint attention skills. *Dev. Psychobiol.* 36, 325–338.
- Mundy, P., and Jarrold, W. (2010). Infant joint attention, neural networks and social-cognition. *Neural. Netw.* 23, 985–997.
- Mundy, P., Sigman, M., and Kasari, C. (1994). Joint attention, developmental level, and symptom presentation in children with autism. *Dev. Psychopathol.* 6, 389–401.
- Mundy, P., Sullivan, L., and Mastergeorge, A. (2009). A parallel and distributed processing model of joint attention and autism. *Autism Res.* 2, 2–21.
- Northoff, G., Heinzell, A., Greck, M., Bermpohl, F., Debrowolny, H., and Panksepp, J. (2006). Self referenced processing in our brain – a meta analysis of imaging studies of the self. *Neuroimage* 31, 440–457.
- Otten, L., Henson, R., and Rugg, M. (2001). Depth of processing effects on neural correlates of memory encoding. *Brain* 125, 399–412.
- Redcay, E., Dodell-Feder, D., Pearson, M., Mavros, P., Kellner, M., Gabrieli, J., and Saxe, R. (2010). Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage* 50, 1639–1647.
- Schilbach, L., Wilms, M., Eickhoff, S., Romanzetti, S., Tepest, R., Bente, G., Shah, N. J., Fink, G. R., and Voegeley, K. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* 22, 2702–2715.
- Schultz, R. (2005). Developmental deficits in social perception in autism: the role of the amygdala and the fusiform face area. *Int. J. Dev. Neurosci.* 23, 125–141.
- Schulz, K. P., Fan, J., Magidina, O., Marks, D. J., Hahn, B., and Halperin, J. M. (2007). Does the emotional go/no-go task really measure behavioral inhibition? Convergence with measures of a non-emotional analog. *Arch. Clin. Neuropsychol.* 22, 151–160.
- Seibert, J. M., Hogan, A. E., and Mundy, P. C. (1982). Assessing interactional competencies: the early social-communication scales. *Infant Ment. Health J.* 3, 244–258.
- Sheslow, D., and Adams, W. (2003). *Wide Range Assessment of Memory and Learning Second Edition Administration and Technical Manual*. Lutz, FL: Psychological Assessment Resources.
- Snodgrass, J., and Corwin, J. (1988). Perceptual identification thresholds for 150 fragmented pictures from the Snodgrass and Vanderwart picture set. *Percept. Mot. Skills* 67, 3–36.
- Stanislaw, H., and Todorov, N. (1999). Calculation of signal detection theory. *Behav. Res. Methods Instrum. Comput.* 31, 137–149.
- Striano, T., Chen, X., Cleveland, A., and Bradshaw, S. (2006a). Joint attention social cues influence infant learning. *Eur. J. Dev. Psychol.* 3, 289–299.
- Striano, T., Reid, V., and Hoel, S. (2006b). Neural mechanisms of joint attention in infancy. *Eur. J. Neurosci.* 23, 2819–2823.
- Symons, C. S., and Johnson, B. T. (1997). The self-reference effect in memory: a meta-analysis. *Psychol. Bull.* 121, 371–394.
- Tomasello, M., and Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monogr. Soc. Res. Child Dev.* 70, 1–136.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–690.
- Vaughan Van Hecke, A., Mundy, P., Acra, C. F., Block, J., Delgado, C., Parlade, M., Meyer, J., Neal, R., and Pomares, Y. (2007). Infant joint attention, temperament, and social competence in preschool children. *Child Dev.* 78, 53–69.
- Vaughan Van Hecke, A., Mundy, P., Block, J., Delgado, C., Parlade, M., Polmares, Y., and Hobson, J. (2012). Infant responding to joint attention, executive processes, and self-regulation in preschool children. *Infant Behav. Dev.* 35, 303–311.
- Williams, J., Waiter, G., Perra, O. A., Perrett, D., Murray, A., and Whitten, A. (2005). An fMRI study of joint attention experience. *Neuroimage* 25, 133–140.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 December 2011; accepted: 28 May 2012; published online: 14 June 2012.

Citation: Kim K and Mundy P (2012) Joint attention, social-cognition, and recognition memory in adults. *Front. Hum. Neurosci.* 6:172. doi: 10.3389/fnhum.2012.00172

Copyright © 2012 Kim and Mundy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.