THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2008, 61 (7), 986–998

Psychology Press
Taylor & Francis Group

Social attention and real-world scenes: The roles of action, competition and social content

Elina Birmingham
University of British Columbia, Vancouver, British Columbia, Canada

Walter F. Bischof
University of Alberta, Edmondton, Alberta, Canada

Alan Kingstone
University of British Columbia, Vancouver, British Columbia, Canada

The present study examined how social attention is influenced by social content and the presence of items that are available for attention. We monitored observers' eye movements while they freely viewed real-world social scenes containing either 1 or 3 people situated among a variety of objects. Building from the work of Yarbus (1965/1967) we hypothesized that observers would demonstrate a preferential bias to fixate the eyes of the people in the scene, although other items would also receive attention. In addition, we hypothesized that fixations to the eyes would increase as the social content (i.e., number of people) increased. Both hypotheses were supported by the data, and we also found that the level of activity in the scene influenced attention to eyes when social content was high. The present results provide support for the notion that the eyes are selected by others in order to extract social information. Our study also suggests a simple and surreptitious methodology for studying social attention to real-world stimuli in a range of populations, such as those with autism spectrum disorders.

Over the last decade there has been an explosion of research interest in what has become known as "social attention". This research has generally focused on understanding how one's attention is affected by the presence of other individuals, epitomized by studies showing that infants and adults alike will attend automatically to where someone else is looking (Friesen & Kingstone, 1998; Hood, Willen, & Driver, 1998; see Langton, Watt, & Bruce, 2000, for a review). In the typical laboratory investigation of this sort, a participant is first shown a picture of a real or schematic face with the eyes looking sideways toward the left or right. Shortly thereafter a response target is presented either at the gazed-at location or at the non-gazed-at location. One normally finds that response time (RT) to detect a target is shorter when the target appears at the gazed-at location than when it appears at the non-gazed-at location. Importantly, this effect emerges rapidly and occurs even when gaze direction does not predict where a target is going to appear.

Although these data fit nicely with the intuition that attention is shifted to where people are looking

Correspondence should be addressed to Elina Birmingham, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC, Canada V6T 1Z4. E-mail: ebirmingham2@yahoo.ca

http://www.psypress.com/qjep
DOI:10.1080/17470210701410375

because one cares about where other people are attending, recent research suggests that this original interpretation (e.g., Friesen & Kingstone, 1998; Langton et al., 2000) may have overstated its case. It has been shown that other biological social-communicative cues produce an attention effect that is very similar (Friesen, Ristic, & Kingstone, 2004), if not identical (Ristic, Friesen, & Kingstone, 2002; Tipples, 2002), to what is found for gaze direction. Some of these cues, like head direction (Langton, 2000) and pointing fingers (Watanabe, 2002), obviously pertain to people, and thus their effects can be readily accommodated by the notion that a range of biological social cues that indicate where other people are attending will trigger a shift in one's attention (Kingstone, Smilek, Ristic, Friesen, & Eastwood, 2003). However, it is now clear that almost any cue with a directional component, from arrows (Ristic & Kingstone, 2005; Tipples, 2002) to numbers (Fischer, Castel, Dodd, & Pratt, 2003; Ristic, Wright, & Kingstone, 2006) will produce a shift in one's attention. These latter cues are not directly associated with people, suggesting that biological social-communicative stimuli may not hold any special status when it comes to producing shifts of attention in cueing studies. What is critical is simply that the cues are directional in nature.

Does it follow from this conclusion that biological social-communicative cues, like the eyes, are never given preferential status by the attentional system? No. After all, there is a wealth of data in the research literature indicating that when a picture of a face is presented to participants, they look preferentially (70–80% of the time) at the eyes of that face (Henderson, Falk, Minut, Dyer, & Mahadevan, 2000; Pelphrey et al., 2002; Walker-Smith, Gale, & Findlay, 1977). This would suggest that the attentional system has a preferential bias for the eyes. And yet, if we have learned anything at all from the cueing studies reviewed above, it is that eyes cannot be considered "special" unless one has compared them against nonface stimuli.

With this point in mind, it is sobering to note that the vast majority of the face-scanning studies have mainly presented participants with only a face to look at. Indeed, not only are all other kinds of background information routinely stripped away, but faces normally are presented in a "passport-type" format, with all body features below the neck missing. Thus, it is possible that people look preferentially at the eyes of an isolated face simply because eyes are the most interesting stimuli in a relatively impoverished display. Therefore the special status of eyes relative to other biological social items (e.g., body, arms, legs) and nonsocial items is less clear-cut than one might initially think.

Indeed, looking at the literature with this thought in mind, one cannot help but be struck by how little there is in the way of research data concerning how people look at scenes when a face is observed as it is normally perceived in the real world—that is, among other body parts, faces, and nonsocial items. On this score past investigations are remarkably silent, save for one most noteworthy exception—the seminal work by the Russian physiologist Alfred Yarbus (1965/1967). Yarbus recorded the eye movements of subjects looking at pictures, continuing the earlier work of Buswell (1935). Yarbus found that when observers are shown the picture of a face presented in isolation, observers tend to look at the eyes of the face, whether it is human or another animal (see Figure 1a). As noted above, this preference for the eyes of an isolated face is now a well-established finding. What has been overlooked, however, is that Yarbus also found that if the face is not presented in isolation, but accompanied by its associated body parts such as arms, torso, and legs, the preferential scanning of the eyes tends to disappear both for humans and for other animals (Figures 1b).[1]

[1] We overlaid the scan paths from Yarbus (1965/1967) onto their corresponding pictures. However, the scan-path image was often not the same size as the picture image, and so an exact match was not possible. As a result, it is actually very difficult to know where the clusters of fixations landed within Yarbus's original images, again making it unclear whether eyes received preferential scanning in all (or any) of his pictures.

Figure 1. *Scan paths of an individual face (A), a face accompanied by the rest of the body (B), and a social scene (C). From* Eye Movements and Vision*, by A. L. Yarbus, 1965 (translated by B. Haigh, 1967), New York: Plenum Press. Copyright 1965 by Springer Science and Business Media (pp. 174, 180, and 189). Adapted with permission.*

Thus Yarbus's data suggest that there may not be a preference for scanning the eyes when the face of a person is presented along with its body. Following a similar line of reasoning, the preference for eyes might also be expected to decline if there were other objects competing for attention, such as those typically found within a complex scene. The present investigation put precisely this hypothesis to the test.

Competition from other objects may not, however, be the entire story. We noticed that when Yarbus (1965/1967) showed his participants a painting by Repin that depicted several people in a room (see Figure 1c), participants now tended to look at the faces of the people in the scene. This suggested to us that increasing the social content of a scene, by adding more people to it, may increase participants' interest in the eyes of the characters. Unfortunately, the resolution of Yarbus's eye monitor did not discriminate clearly between observers' scanning of the eyes in the scene from the other available facial features.

And most critically, there is the concern that all of the observers tested by Yarbus were well acquainted with Repin's picture, a concern raised by Yarbus himself: "This evidently accounts for the generally considerable similarity between all the records.. . .Undoubtedly, observers familiar both with the picture and the epoch represented in it would examine the picture differently from people seeing it for the first time and unfamiliar with the epoch it represents" (Yarbus, 1965/1967, p. 192). This concern is reinforced by Yarbus's subsequent demonstration that if people are simply asked different questions about the picture, their fixation pattern can vary dramatically and systematically from the free-viewing conditions. In other words, the fixations on the heads and eyes in the Repin picture may have reflected a shared knowledge of the picture being viewed and the problem it posed for the knowledgeable observer at the time of perception. These considerations, coupled with the fact that Yarbus's observations stem from the use of a single visual scene, leads one to the conclusion that there is a need to examine further the influence of social content on how attention is allocated, with special interest paid to whether it impacts the allocation of attention to the eyes.

In sum, the present study had two main goals. First, we wanted to determine whether, with naïve observers, in free-viewing conditions, and with other items readily available for viewing, there is a preference for scanning the eyes of a single individual in a scene. Second, we wanted to discover whether adding more people to a scene will increase the degree that eyes are scanned.

It is worth noting that, when manipulating the number of people within different scenes, one is immediately faced with the problem of what the people in the scenes should be doing. Because we had no way of knowing how action in a scene would impact scanning patterns (e.g., Repin's painting is ambiguous on this score) we controlled and manipulated this factor by having people photographed doing nothing (e.g., just sitting on their own; *inactive scenes*), doing something (e.g., sitting reading a book on their own; *active scenes*), or, in the case when there were multiple people in a scene, doing something separately (e.g., sitting together but reading individually; *active scenes*) or doing something together (e.g., sharing a book; *interactive scenes*). These last two conditions, in which multiple people in a room are either acting separately or interacting, represent a subtle yet potentially important difference in social content, and thus in keeping with the aims of the present study we wished to examine whether the eyes would be scanned differently in scenes containing social action (several people doing something separately) compared to scenes with social *inter*action (several people doing something together). Examples of these scene types are presented in Figure 2a. Specific experimental details are presented below.

Method

Participants

A total of 20 undergraduate students from the University of British Columbia participated in this experiment. All had normal or corrected-to-normal vision and were naïve to the purpose of the experiment. Each participant received course credit for participation in a 1-hour session.

Apparatus

Eye movements were monitored using an Eyelink II tracking system. The online saccade detector of the eye tracker was set to detect saccades with an amplitude of at least 0.5°, using an acceleration threshold of $9{,}500^\circ/s^2$ and a velocity threshold of 30°/s.

Stimuli

Full-colour images were taken with a digital camera in different rooms in the Psychology building. Image size was 36.5 × 27.5 cm corresponding to 40.1° × 30.8° at the viewing distance of 50 cm, and image resolution was 800 × 600 pixels. A total of 40 scenes were used in the present experiment. Each scene contained either 1 or 3 persons. In the 1-person scenes the individual was either doing something (*active*) or doing nothing (*inactive*). Similarly, in the 3-people scenes, people either did nothing (*inactive*), or did something on their own (*active*), or did something together

Figure 2. *A. Examples of the four scene types. From top to bottom: 1-person active, 1-person inactive, 3-people active, 3-people inactive. B. Corresponding regions of interest used in analysis (eyes, head, body, foreground objects, background objects). C. Corresponding plots of fixations for all participants. (To view this figure in colour please see the article in the online issue.)*

(*interactive*). All scenes were comparable in terms of their basic layout: Each room had a table, chairs, objects, and background items (e.g., see Figure 2).

Procedure

Participants were seated in a brightly lit room and were placed in a chin rest so that they sat approximately 50 cm from the display computer screen. Participants were told that they would be shown several images, each one appearing for 15 s, and that they were to simply look at these images.

Before beginning the experiment, a calibration procedure was conducted. Participants were instructed to fixate a central black dot and to follow this dot as it appeared randomly at nine different places on the screen. This calibration was then validated, a procedure that calculates the difference between the calibrated gaze position and target position and corrects for this error in future gaze position computations. After successful calibration and validation, the scene trials began.

At the beginning of each trial, a fixation point was displayed in the centre of the computer screen in order to correct for drift in gaze position. Participants were instructed to fixate this point and then press the spacebar to start a trial. One of 40 pictures was then shown in the centre of the screen. Each picture was chosen at random and without replacement. The picture remained visible until 15 s had passed, after which the picture was replaced with the drift correction screen. This process repeated until all pictures had been viewed.

Results

Data handling

For each image, an outline was drawn around each region of interest (e.g., "eyes"), and each region's pixel coordinates and area were recorded. We defined the following regions in this manner: eyes, heads (excluding eyes), body (including arms, torso and legs), foreground objects (e.g., tables, chairs, objects on the table), and background objects (e.g., walls, shelves, items on the walls). Figure 2b illustrates these regions.

To determine what regions were of most interest to observers we computed *fixation proportions* by dividing the number of fixations for a region by the total number of fixations over the whole display. These data were area-normalized by dividing the proportion score for each region by its area (Smilek, Birmingham, Cameron, Bischof, & Kingstone, 2006).

To determine whether observers' interest in the regions changed over time we computed the *cumulative fixation proportions* for the regions in 1-s time steps for the 15 s of display duration. These data were area-normalized by dividing the proportion score for each region by its area.

To determine whether fixation time differed among the regions we computed the *duration proportions* for each region. These data were area-normalized by dividing the time score for each region by its area.

Fixation proportions

Before analysing all the data we examined whether participants viewed the "3-people active" scenes differently from the "3-people interactive" scenes. These two scene types were scanned similarly (i.e., there were no significant differences in fixation proportions between comparable regions, all Fs < 1), so we combined the data from these two scene types to create a single data set of "3-people active". We then submitted all the fixation proportion data to a $2 \times 2 \times 5$ within-subjects analysis of variance (ANOVA) with people (1 person vs. 3 people), activity (inactive vs. active) and region (eyes, head, body, foreground, background) as factors.

Figure 3 shows these data for eyes, heads, and other regions. Looking at this figure it is immediately evident that the eyes were fixated far more than any other region, as reflected by a main effect of region, $F(4, 76) = 639.86$, $p < .001$. This strong preference for the eyes was true for both 1-person scenes (.60 average fixation proportion) and 3-person scenes (.64 average fixation proportion). Furthermore, it is clear that while eyes received more fixations than all other regions, heads were also fixated quite frequently, more so than bodies, foreground objects, and background. In addition, there were no differences between bodies, foreground objects, and background. These observations were confirmed with post hoc Tukey–Kramer pairwise comparisons ($p < .05$).

A significant Region × People × Activity interaction, $F(4, 76) = 14.88$, $p < .001$, reflected, however, that when there was activity in the scene, there were more fixations on the eyes when 3 people were in a scene than when 1 person was in the scene. This observation was confirmed by a post hoc pairwise comparison (Tukey–Kramer $p < .05$) of 3-people (.66 fixation proportions) versus 1-person active scenes (.56 fixation

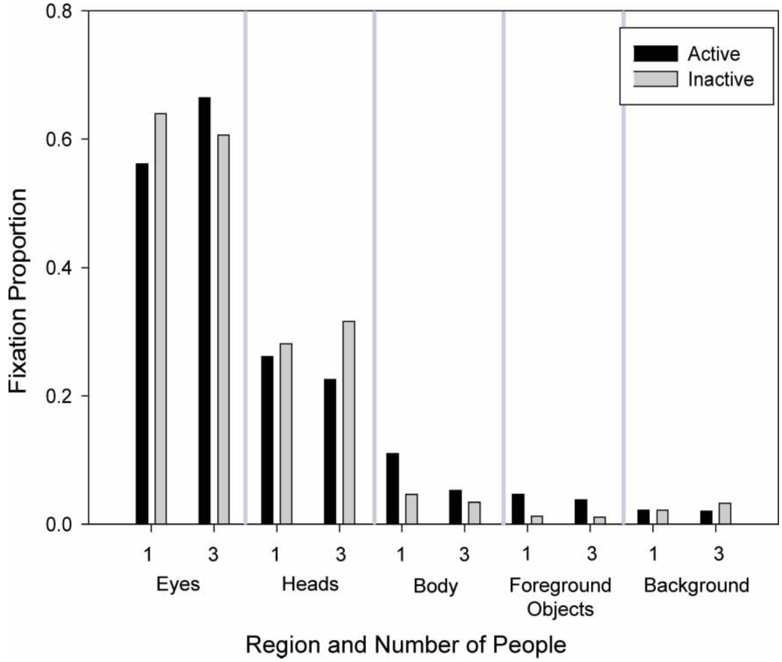

Figure 3. *Fixation proportion data for eyes, head, body, foreground, and background plotted as a function of people, activity, and region. Fixations to eyes were enhanced by increasing social content (i.e., 3-people scenes vs. 1-person scenes) when the scenes contained activity (active scenes).*

proportions). In contrast, when there was no activity in a scene, there were no more fixations on the eyes for 3 people scenes than 1 person scenes (Tukey–Kramer $p > .05$).[2] Further pairwise comparisons ($p < .05$) revealed that observers fixated the eyes more in the 3-active scenes than in the 3-inactive scenes, whereas the opposite effect of activity occurred for 1-person scenes (1-person inactive > 1-person active).

Cumulative fixation proportions

We were also interested in how scanning preferences changed over time. An ANOVA was performed on the cumulative fixation proportion data with people, region, activity, and interval as factors, the last broken into 1-s intervals (0–1 s, 1–2 s, . . .14–15 s).

Figure 4 shows these data for eyes, heads, and other regions. All of the effects from the fixation proportion analysis were again significant in the analysis of cumulative probabilities. For example, there was a main effect of region, reflecting the fact that again eyes were fixated by far the most, $F(4, 7200) = 12{,}406.46$, $p < .001$ (post hoc comparisons revealed that this preference for eyes was significant, $p < .05$, even at the first 1-s interval).[3] Looking at Figure 4 one sees that many of the regions were refixated over time, resulting in an

[2] We were concerned that the overall interaction between people and activity and the bias it has on scanning the eye region might be an artifact of averaging across a few specific actions. Importantly, however, across eight of the nine active scenes (ranging from simply reading a book to threatening to punch each other) there was a larger preference for the eyes in the active 3-people scenes than in the active 1-person scenes, with a binomial test revealing that the probability of eight or more rooms showing this magnitude difference was $p < .02$.

[3] While eyes were fixated the most within the first 1-s interval, the very first fixation was made most often to the head than to any other region, $F(4, 76) = 24.56$, $p < .001$. This movement to the head would appear to be a "signature" of first acquiring the eye region.

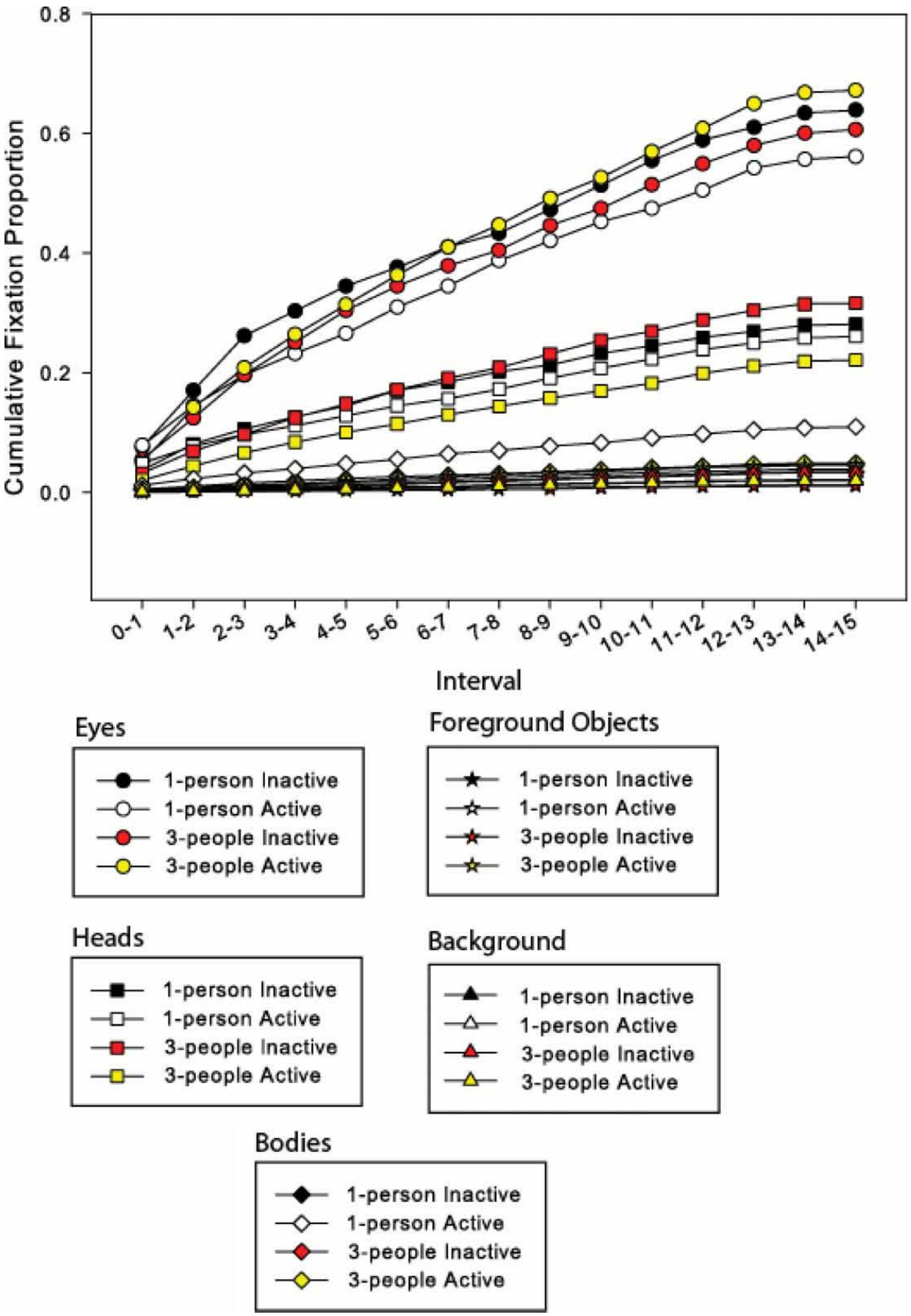

Figure 4. *Cumulative fixation proportions for 1-person scenes and 3-people scenes. Data are plotted as a function of region, activity level, and viewing interval. (To view this figure in colour please see the article in the online issue.)*

overall increase in the cumulative fixation proportions and hence a main effect of interval, $F(14, 7200) = 483.58$, $p < .0001$. This increase in cumulative fixation proportions over time was most pronounced for eyes, reflected by a Region × Interval interaction, $F(56, 7200) = 153.55$, $p < .0001$.

Recall that the key finding from the overall fixation proportion data was that there was an overall preference for the eyes in 1- and 3-person scenes, with scene activity enhancing the preference for eyes in the 3-people scenes. Our cumulative analysis revealed that this preference for eyes in the active 3-people scenes did not begin to emerge until after 6 s, at which point there was a significant difference between fixations on the eyes in 3-people active (yellow circles) relative to 1-person active scenes (white circles). This difference persisted until the end of viewing. This observation was confirmed by a People × Activity × Interval interaction for cumulative fixations on eyes and subsequent post hoc pairwise comparisons ($p < .05$).

Duration proportions

Duration proportions are shown in Figure 5. These data closely parallel the fixation proportion data. For instance, it is clear that the eyes were fixated for the longest out of all the regions, followed by heads, bodies, background, and foreground objects. This was reflected in a highly significant main effect of region, $F(4, 76) = 552.23$; $p < .0001$. Post hoc pairwise comparisons ($p < .05$) revealed significantly longer durations for eyes than for heads, bodies, background, and foreground objects. In addition, heads were fixated for longer than bodies, background, and foreground objects ($p < .05$). In addition to the effect of region, there was also a People × Activity × Region interaction, $F(4, 76) = 17.42$, $p < .0001$. This latter, higher order interaction indicated a similar pattern to that in the fixation proportion analysis. That is, when there was activity in the scene, fixation time was longer on the eyes when 3 people were in a scene than

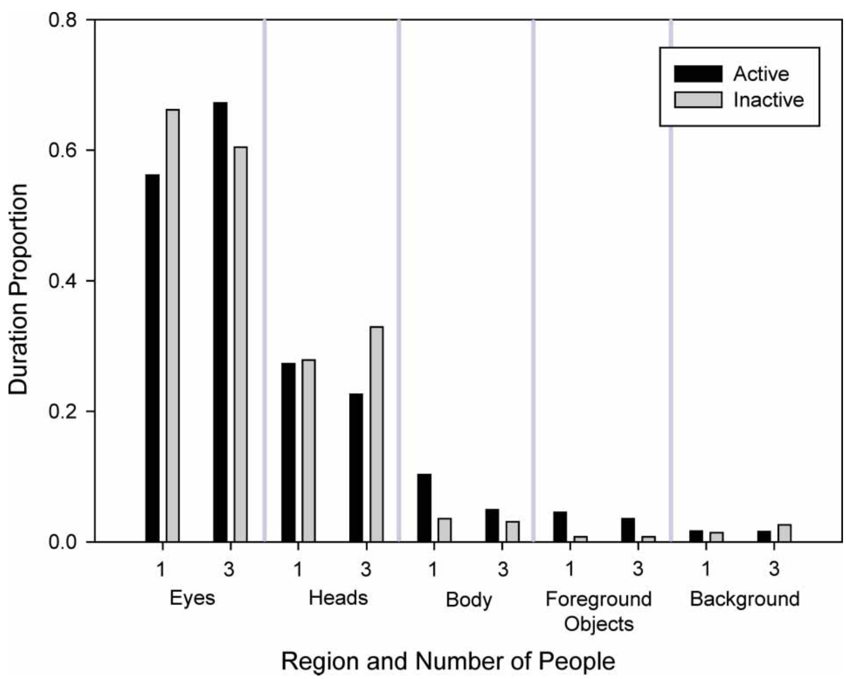

Figure 5. *Duration proportion data plotted as a function of people, activity, and region. Fixation time on eyes were enhanced by increasing social content (i.e., 3-people scenes vs. 1-person scenes) when the scenes contained activity (active scenes).*

when 1 person was in the scene. This observation was confirmed by a post hoc pairwise comparison ($p < .05$). In contrast, when there was no activity in a scene, fixation durations were equal for 3-people scenes and 1-person scenes, $p > .05$. As in the fixation proportion analysis, further pairwise comparisons ($p < .05$) revealed that observers fixated the eyes longer in the 3-active scenes than in the 3-inactive scenes, whereas the opposite effect of activity occurred for 1-person scenes (1-person inactive > 1-person active).

GENERAL DISCUSSION

The present study had two main goals. First, we wanted to determine whether, with naïve observers, in free-viewing conditions, and with other items readily available for viewing, there is a preference for scanning the eyes of a single individual in a scene. While it was well established in the literature that people look at the eyes of a face when the face is the only item available for scanning, it was not at all clear that this finding would hold when a person was presented with a face along with its body and other items in a scene. Indeed we had noted that the seminal work of Yarbus provided evidence to suggest that scanning of the eyes might not receive preference relative to other body parts and/or objects if they were made available for viewing. The results of our study were unequivocal on this issue. People prefer to look at the eyes of one person in a scene, even when there are other items available (average fixation proportions .60 for eyes vs. .40 for elsewhere). Thus, the preferential bias for the eyes of a person persists in real-world scenes containing other body parts and objects. However, it is also clear that eyes did not entirely dominate observers' attention in these scenes, as fixations did frequent (.40) other body parts and objects in the scenes. For instance, the region with the next highest fixation proportion, and duration proportion, was the head region, which was fixated more frequently and for longer than the body region, foreground objects, and background objects (but less so than eyes). Thus, it is clear that observers showed a particular interest in the faces of people in the scenes, focusing especially on their eyes.

The present study also demonstrated that the preferential bias for eyes emerged remarkably early. Our analysis of the cumulative fixation proportions showed that after first fixating the head, the eyes received more fixations than other regions within the first second of viewing. Moreover, observers were more likely to revisit the eye region again and again while viewing a scene, resulting in an enhanced preference for eyes as viewing time was extended.

The second main goal of our study was to discover whether increasing the social content of a scene, by adding more people to it, would increase the extent that the eyes are scanned. We had noted that the eye-scanning data of the Repin painting (Yarbus, 1965/1967) provided indirect support for our proposal that people will look more to the eyes as the social content of a scene is increased. Convergent with this proposal is the recent finding that people look more to the eyes as the need to extract the social information of a scene increases, for instance, in order to infer the attentional states of the people depicted within a scene (Smilek et al., 2006).

The data from the present study shed new light on this issue. First, we found that increasing the social content of a scene does drive people to look more, and longer, at the eyes, but only when the people in the scene are actively doing something. Second, our cumulative probability data revealed that this impact of social content and activity does not first emerge until after 6 s of viewing time, suggesting that this interaction reflects a rather complex level of scene analysis by the observer. It is our speculation that when the social content of a scene is relatively low—that is, when there is only one person in the scene—action draws attention away from the eyes because eye information is not critical to understanding the action. However, when the social content of a scene is relatively high—that is, when there are multiple people within a scene—action draws attention toward the eyes because eye information is critical to understanding the social meaning of the action.

It is clear that this speculation requires future investigation. Nevertheless, it is noteworthy that our data do suggest a subtle, yet powerful, way to examine the observers' sensitivity to changes in social content. For instance, it would be interesting to examine how people with autism scan scenes with one versus many people and how their exploration of these scenes is affected by the action within it. Previous studies have shown that individuals with autism are less likely to spontaneously orient to social stimuli (e.g., people) in their natural environments (Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998; Osterling & Dawson, 1994; Osterling, Dawson, & Munson, 2002; Swettenham et al., 1998) and demonstrate abnormal face processing compared to typically developing individuals (e.g., Behrmann, Thomas, & Humphreys, 2006; Joseph & Tanaka, 2002; Langdell, 1978). Eye movement studies have shown that these social impairments are further characterized by a specific avoidance of eyes (Klin, Jones, Schultz, Volkmar, & Cohen, 2002; Pelphrey et al., 2002). One explanation for these findings is that individuals with autism have a heightened negative emotional response to the eyes of others, and that they avoid the eyes in order to reduce this overarousal (Dalton et al., 2005). This hypothesis would predict that individuals with autism are adverse to an increase in social content. If so, then they might tend to look away from the eyes as people are added to a scene, and their activity increases.

An intriguing possibility is that because eye information may be critical to understanding the social meaning of action when there are three people in a scene, observers might also be likely to make more eye movement transitions between the eyes of the people in 3-people active scenes.[4] That is, perhaps in trying to understand the nature of the social activity occurring in 3-active scenes, observers would have made more eye movements between the eyes of the different people in the scene to look for states of shared or mutual attention. If this was true, and observers made more eyes–eyes transitions in these scenes than in 3-people inactive scenes, it would bolster our conclusion that eye information is critical to understanding the meaning of social action. However, an analysis of eyes–eyes transition frequencies revealed no such differences ($p > .05$). In fact, if anything, there was a nonsignificant trend toward fewer eyes–eyes transitions for 3-active scenes ($M = 0.17$) than 3-inactive scenes ($M = 0.23$). Thus, it may be the case that fixation frequencies and fixation durations are more sensitive to the effects of social content and activity that are transition frequencies. Alternatively, it may be that differences in transition frequencies occurred, but that they were masked by the variation in social activity occurring across the scenes (e.g., variations in mutual versus shared gaze, etc.). Future studies will be required to investigate this possibility further.

Collectively the data from the present study provide support for our interpretation of Yarbus's original work, from which we had hypothesized that scanning of the eyes would be sensitive to the presentation of competing objects and variation in social context. Our study goes well beyond this initial work, however, in systematically testing and confirming that these are important factors to consider when studying how people scan natural scenes. Moreover our work has implications for the social-attention literature. Whereas several studies have shown a preferential bias for eyes within highly simplistic displays (e.g., cut-out of a face against a blank background), to our knowledge the present study provides the first demonstration that this bias is expressed for complex real-world scenes. Our study also provides intriguing evidence that as scene complexity continues to increase—for example, by adding action and social content to a scene—the preference for the eyes will continue to be enhanced.

Finally, our work has implications for recent scene-scanning studies, insofar as we drive home the fact that the social content of a scene is critical to where people look within a scene. To date most

[4] We thank an anonymous reviewer for this suggestion.

studies of this nature have used scenes that do not contain people (de Graef, 1992; Henderson & Hollingworth, 1999; Henderson, Weeks, & Hollingworth, 1999; Rayner & Pollatsek, 1992; Underwood & Foulsham, 2006). Thus, an important question is whether the results from these previous studies are restricted to scenes without people. Our present work suggests that they may well be—that is, people prefer to look at people (for support for this idea see the work of Fletcher-Watson, Findlay, Leekam, & Benson, 2007). It is important to note that we are not claiming that this preference for people—especially the eyes of people—is absolute. It is well established that eye movements change according to task demands. That said, it remains an open question how changes in task demands will affect people's profound tendency to look at the eyes of others.

Original manuscript received 7 December 2006
Accepted revision received 28 March 2007
First published online 23 July 2007

REFERENCES

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing it differently: Visual processing in autism. *Trends in Cognitive Sciences*, *10*, 258–264.

Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*, *8*, 519–526.

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, *28*, 479–485.

de Graef, P. (1992). Scene-context effects and models of real-world perception. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 243–259). New York: Springer-Verlag.

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes shifts in spatial shifts of attention. *Nature Neuroscience*, *6*, 555–556.

Fletcher-Watson, S., Findlay, J. M., Leekam, S. R., & Benson, V. (2007). *Rapid detection of person information in a naturalistic scene.* Manuscript submitted for publication.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*, 490–495.

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 319–329.

Henderson, J. M., Falk, R. Minut, S., Dyer, F. C., & Mahadevan, S. (2000). *Gaze control for face learning and recognition by humans and machines* (Tech. Rep. No. 4, pp. 1–14). East Lansing, MI: Michigan State University, Eye Movement Laboratory.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*, 243–271.

Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during scene viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 210–228.

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*, 131–134.

Joseph, R. M., & Tanaka, J. (2002). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry*, *43*, 1–14.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention researchers! It is time to take a look at the real world. *Current Directions in Psychological Science*, *12*, 176–180.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Defining and quantifying the social phenotype in autism. *American Journal of Psychiatry*, *159*, 895–908.

Langdell, T. (1978). Recognition of faces: An approach to the study of autism. *Journal of Psychology and Psychiatry*, *19*, 255–268.

Langton, S. R. H. (2000). The mutual influence of gaze and head orientation in the analysis of social attention direction. *Quarterly Journal of Experimental Psychology*, *53A*, 825–845.

Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*, 50–59.

Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders*, *24*, 247–257.

Osterling, J., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology*, *14*, 239–251.

Pelphrey, K. A., Sasson, N. J., Reznick, S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, *32*, 249–261.

Rayner, K., & Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal of Psychology*, *46*, 342–376.

Ristic, J., Friesen, C. K., & Kinsgtone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review*, *9*, 507–513.

Ristic, J., & Kingstone, A. (2005). Taking control of reflexive social attention. *Cognition*, *94*, B55–B65.

Ristic, J., Wright, A., & Kingstone, A. (2006). The number line effect reflects top-down control. *Psychonomic Bulletin & Review*, *13*, 743–749.

Smilek, D., Birmingham, E., Cameron, D., Bischof, W. F., & Kingstone, A. (2006). Cognitive ethology and exploring attention in real world scenes. *Brain Research*, *1080*, 101–119.

Swettenham, J., Baron-Cohen, S., Charman, T., Cox, A., Baird, G., Drew, A., et al. (1998). The frequency and distribution of spontaneous attention shifts between social and nonsocial stimuli in autistic, typically developing, and nonautistic developmentally delayed infants. *Journal of Child Psychology and Psychiatry*, *39*, 747–753.

Tipples, J. (2002). Eye gaze is not unique: Automatic orienting in response to uninformative arrows. *Psychonomic Bulletin & Review*, *9*, 314–318.

Underwood, G., & Foulsham, T. (2006) . Visual saliency and semantic congruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*, *59*, 1931–1949.

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

Watanabe, K. (2002). Reflexive attentional shift caused by indexical pointing gestures [Abstract]. *Journal of Vision*, *2*(7), 435a.

Yarbus, A. L. (1967). *Eye movements and vision* (B. Haigh, Trans.). New York: Plenum Press. (Original work published 1965.)