I See What You See: Eye Movements in Real-World Scenes Are Affected by Perceived Direction of Gaze

Monica S. Castelhano[1], Mareike Wieth[2], and John M. Henderson[3]

[1] Queen's University, Kingston ON K7L 3N6, Canada
[2] Albion College, Albion MI 49224, USA
[3] University of Edinburgh, Edinburgh EH8 9YL, UK

Abstract. In this chapter, we report an investigation the influence of the saliency of another person's direction of gaze on an observer's eye movements through real-world scenes. Participants' eye movements were recorded while they viewed a sequence of scene photographs that told a story. A subset of the scenes contained an actor. The actor's face was highly likely to be fixated, and when it was, the observer's next saccade was more likely to be toward the object that was the focus of the actor's gaze than in any other direction. Furthermore, when eye movement patterns did not show an immediate saccade to the focused object, observers were nonetheless more likely to fixate the focused object than a control object within close temporal proximity of fixation on the face. We conclude that during real-world scene perception, observers are sensitive to another's direction of gaze and use it to help guide their own eye movements.

1 Introduction

During the exploration of a visual scene, our eyes change fixation position about three times a second via saccadic eye movements. This active process of gathering visual information is necessary because high visual acuity at the center of fixation falls off rapidly and continuously into a low-resolution surround. Eye movement control is the process of directing fixations through a scene in real time to serve ongoing perceptual, cognitive, and behavioral activities [22]. Early studies demonstrated that from the first fixation on a scene, areas that are interesting and informative are fixated, whereas uniform, uninformative areas are not [35,37]. Specific scene areas are selected for further scrutiny using two main sources of information: stimulus-based information generated from the image, and top-down, memory-based information generated from visual and cognitive systems [20]. For example, stimulus-based information that is salient such as discontinuities in luminance and color are known to affect how fixations are deployed over a scene [27,45]. In addition, fixations are often biased toward particular regions of interest due to top-down factors such as task demands, attentional sets, context effects, and knowledge of scene schemas [8,15,21,23,52,54].

L. Paletta and E. Rome (Eds.): WAPCV 2007, LNAI 4840, pp. 251–262, 2007.

Another potential source of top-down information concerning what is interesting and informative in a scene is the focus of another person's attention. The direction that another person is looking is known to elicit a reflexive response in our own gaze [14,38,47], but whether this social cue is used to direct attention during the exploration of a novel visual scene remains unknown. Although not adequately addressed in the literature, the role of gaze perception in the allocation of attention within real-world scenes remains an important issue in both theories of attentional control and scene perception [28,34,50,51]. The present study investigates how perception of another's gaze influences the control of eye movements during real-world scene perception.

1.1 Gaze Perception in Children and Infants

Gaze perception has traditionally been studied in the context of the development of theory of mind [5,7,9,48] and for its influence on social interactions [6,29,30,53]. Gaze detection and perception in infants has been proposed to be one of the initial building blocks toward an understanding of intentionality, "mindreading", and theory of mind [5,11,14,46]. Infants have been shown to fixate on the eyes of a face more than any other facial feature [39,41], and have also been shown to use the gaze of others (as well as body and head movements and orientation) to redirect their attention appropriately [9,12,48].

1.2 Gaze Perception in Machines

The role of developmental learning in robots has allowed researchers to rethink how learning takes place and joint attention has played an important role in this [36,40,42,43]. Researchers found that joint attention accelerates robot learning [42] and can emerge as a process without the need of direct task evaluation from an external observer (thus, more closely mimicking infant learning) [40]. Although the mechanisms behind how infants and robots learn joint attention may differ [49], each points to the importance of joint attention in extracting relevant and important information from the surrounding environment.

1.3 Gaze Perception in Adults

In adults, the importance of gaze perception has been highlighted in interpreting the intentions of others. For example, direct gaze is linked both to negative and positive intentions, while averted gaze is linked to avoidance or taken as a sign to end the current interaction [3,29,30]. Moreover, gaze perception has been explored in the context of classical visual attention paradigms. These studies demonstrate that as adults, humans are very efficient at detecting direction of gaze [53] and tend to direct their own attention automatically (both with and without eye movements) in the direction of the perceived gaze of another. The latter effect is observed regardless of the relationship of the perceived gaze to the task at hand, and even when it is detrimental to performance [13,16,17,33,38,47]. Additionally, the deployment of attention due to gaze perception (vs. other symbolic, central cues such as arrows) has proven to have a stronger cueing effect and leads to a more efficient deployment of attention [18,31]. The idea that gaze cues may play a special and differing role in attentional control was further supported by the finding that this type of deployment has different underlying cortical structures and may be a result of differing cognitive processes aimed at specifically integrating social information [25].

2 Experiment

In the present study, we explored whether a person's face is salient within a scene in the presence of other salient objects, and furthermore whether viewed gaze direction is another important factor controlling an observer's eye movements during scene exploration. It is commonly assumed that faces in scenes attract fixations [15,54], but there is little direct empirical support for this belief. The current experiment will allow us to investigate the question of how salient an actor's face is compared to other objects in a scene.

We are also interested in how the gaze cue is interpreted and whether it has an effect on the subsequent allocation of attention in real-world scenes. On the one hand, past studies have shown that to encode the details from a scene, fixations must land on or near those details [44,24,26]. Based on these findings, we would expect that an actor's face to be fixated if that actor's direction of gaze is to influence an observer's attention. On the other hand, we also know that observers can readily pick up on gaze information from luminance differences between the iris and sclera [1,2] as well as body and head positions [4,10,19,32], and it is also possible that attention could be directed to the location receiving attention by another person without explicitly fixating the other person's gaze beforehand. We will look at whether there is a direct link between fixating the gaze and region of the actor's focus or whether fixating the eyes is necessary for the influence of gaze to take effect.

More importantly, we will examine whether an object that is the focus of attention of an actor does in fact receive ”priority” when exploring the various contents of a real-world scene. It is not yet clear whether the direction of an actor's gaze in a scene will influence the attention of an observer at all. Studies of gaze perception suggest that observers will most likely use the direction of gaze of another as an attentional cue [13,16,17,33,38,47] and reflexively orient in the direction of perceived gaze. However, another possibility is that in interpreting a scene, the reflexive tendency seen with gaze perception in simpler stimuli can be over-ridden, allowing attention to be directed to other aspects of the scene. In the present study, we used eye movement measures to assess the influence of the direction of gaze of a depicted actor on an observer's eye movement behavior.

In the current study, participants viewed a sequence of scenes presented like a slide show that portrayed the story of a janitor (the actor) cleaning an office. Of the scenes presented, only a subset had the actor gazing at a particular object within the scene. For these critical scenes, when actor was looking at a particular object (the focused object), we examined how that fact affected where the observer looked in the scene.

2.1 Methods

Participants. Fifty-three undergraduates from Michigan State University participated in the study for credit in an introductory psychology course. All subjects had normal or corrected to normal vision.

Apparatus and Stimuli. The stimuli consisted of 54 full-colored photographs depicting the story of a janitor cleaning an office (see Figure 1). The scenes were presented on an NEC Multisync XE 15-in. (38.1 cm) color monitor driven by a Hercules Dynamite Pro super video graphics adapter card. The scenes subtended 22 x 29 of visual angle viewed from 55 cm away. Participants' head movements were restricted using a bitebar and headrest, and eye movements were recorded by an ISCAN RK-416HS eye movement monitor operating at a sampling rate of 120 Hz.

Procedure. Participants were first calibrated on the eyetracker, and were recalibrated when necessary. A sequence of 54 scenes was then presented in the form of a slide show that told a story. Each scene was presented for 5 s. Participants were instructed to view the sequence of scenes to understand the story, paying attention to the details.

2.2 Results

Twenty-seven critical scenes containing the actor were analyzed. An example of a critical scene is shown in Figure 1, and includes the scan pattern of one participant. For each critical scene five regions were defined: the actor's face, the focused object (the object that the actor was looking at), a non-focused control object (an object about the same size as the focused object but that the actor was not looking at), and two additional control regions constructed by reflecting the face and focused object regions around the horizontal midline of the scene (thus, including the same total area as these regions, and controlling for their distance from the scene center). We examined two questions: First, would the actor's face attract fixations? Second, would the actor's direction of gaze affect the observer's eye movements in the scene?

To examine whether faces attract attention in scenes, we generated four measures for each of the 27 critical scenes containing the actor: (1) The latency from onset of the scene to the initial fixation on the actor's face, compared to the latency from onset of the scene to the initial fixation on the control regions, (2) The total fixation time on the actor's face compared to the control regions, (3) The number of new gazes on the actor's face, defined as the number of times that the observer's fixation moved from another area of the scene to the actor's face (given that the face was fixated) compared to the control regions, and (4) The percentage of trials that the face was looked at in comparison to the control regions.

Figure 2 shows the means for the four measures for each of the defined regions of interest. We found that for every measure, the face was viewed significantly

Fig. 1. Example of one scene containing the actor, with one participant's eye movements over the scene depicted in white; lines represent saccades and circles represent fixations

longer and sooner than the control regions: (1) Latency to Face Fixation (face vs. non-focused object: $t(52) = -16.68$, $p < .01$; face vs. face control: $t(52) = -2.66$, $p < .01$); (2) Total Fixation Time (face vs. non-focused object: $t(52) = 7.67$, $p < .01$; face vs. face control: $t(52) = 6.56$, $p < .01$); (3) Number of New Gazes on the face (face vs. non-focused object: $t(52) = 8.45$, $p < .01$; face vs. face control: $t(52) = 5.64$, $p < .01$); (4) Percentage of Trials Fixated (face vs. non-focused object: $t(52) = 16.81$, $p < .01$; face vs. face control: $t(52) = 13.44$, $p < .01$). However, as is shown in Figure 2, fixations to the focused object occurred at an even shorter latency than those to the face. It is the analysis of the focused object that we turn to next, followed by further examination of the relationship between examination of the face (gaze) and focused object.

To examine the question of whether the actor's direction of gaze affect the eye movements of an observer, we used four measures like those used to investigate face fixation: (1) The latency from onset of the scene to the initial fixation on the focused object, compared to the latency from onset of the scene to the initial fixation on the control regions, (2) The total fixation time on the focused object, compared to the control regions, (3) The number of new gazes on the focused object, defined as the number of times that the observer's fixation moved from another region of the scene to the focused object (given that the object was fixated) compared to the control regions, and (4) The percentage of trials that the focused object was looked at in comparison to the control regions.

First we examined fixation patterns on the focused object in comparison to the object controls and found that for all measures, the focused object was attended more than any other region: (1) Latency To Fixation (focused object vs. non-focused object: $t(52) = -20.48$, $p < .01$; focused object vs. control object: $t(52) = -12.40$, $p < .01$); (2) Total Fixation Time (focused object vs. non-focused object: $t(52) = 14.67$, $p < .01$; focused object vs. control object: $t(52) = 13.72$, $p < .01$); (3) Number of New Gazes on the focused object (focused object vs. non-focused object: $t(52) = 16.76$, $p < .01$; focused object vs. control

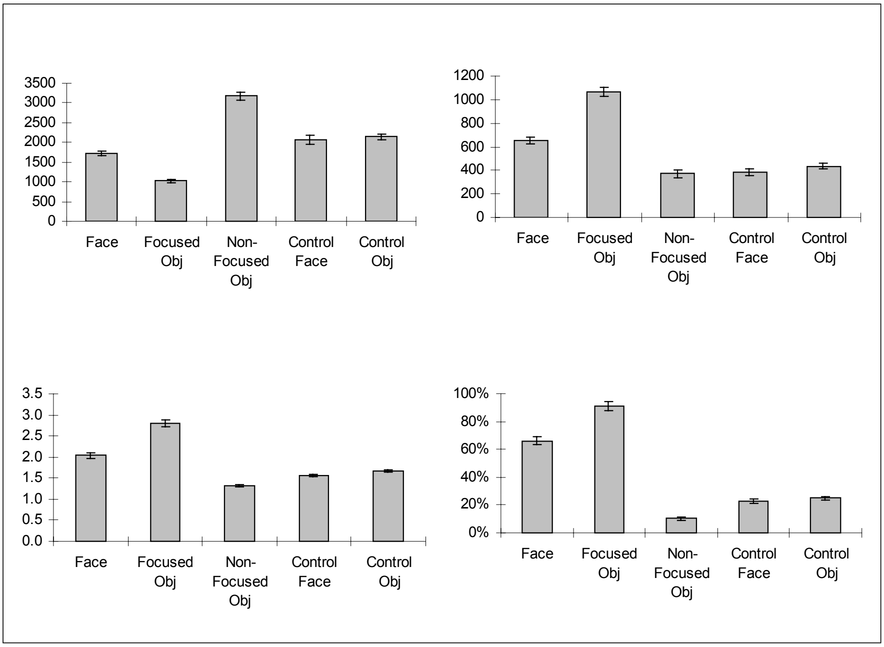

Fig. 2. Eye movement measures by region. (a) Latency to first fixation, (b) Total fixation time, (c) Average number of times gaze moved to each region, and (d) Percentage of trials each region fixated. Error bars represent the standard error of the mean.

object: $t(52) = 12.74$, $p < .01$); (4) Percentage of Trials Fixated (focused object vs. non-focused object: $t(52) = 22.74$, $p < .01$; focused object vs. control object: $t(52) = 19.10$, $p < .01$).

Secondly, we also compared the latency of fixating the focused object to the latency to the face area and found that as was the case for the focused object versus control regions, the focused object was fixated sooner than the face, $t(52) = -10.32$, $p < .01$, and as is expected based on the cuing paradigm, observers did spend more time examining the focused object than the face as shown by the total fixation time ($t(52) = 12.17$, $p < .01$), number of new gazes ($t(52) = 9.66$, $p < .01$), and percentage of trials that these regions were fixated ($t(52) = 7.96$, $p < .01$). Thus, it seems that in some cases, the gaze of the person was perceived without directly looking at the actor's face or gaze. As was mentioned, there are a number of cues (head position and body orientation) that can cue which way the actor's gaze was pointed. However, we also know that with scene stimuli (as opposed to simple cuing paradigms) that there may be a number of candidate objects that can be interpreted as the focused object. Thus, we would expect that despite the other cues to the direction of the actor's gaze, that there would be a tight link between fixations on the gaze and on the focused object. In the following analyses, we further explore this link.

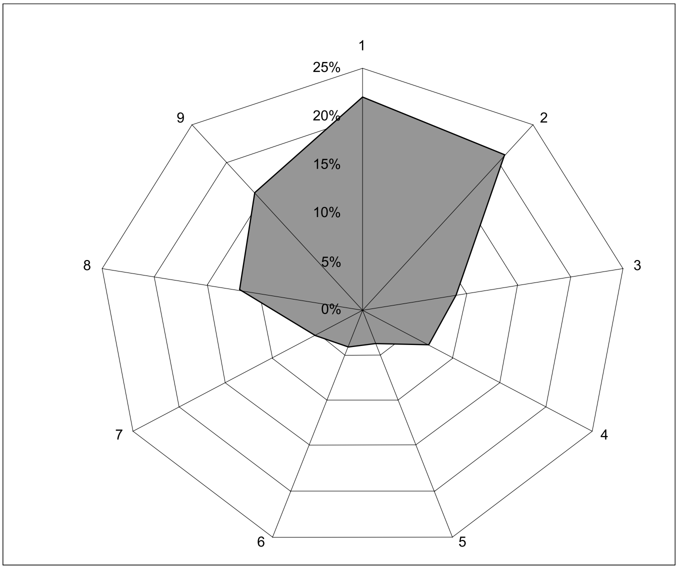

Fig. 3. (a) Direction of saccades away from the actor's face, in polar coordinates. Each point indicates the mean percentage of saccades falling within a 40 degree arc centered on each direction. Direction 1 is the direction from the face to the focused object, with consecutive directions plotted clockwise in the picture plane.

To provide a more detailed examination of the observer's eye movement behavior related to the direction of the depicted actor's gaze, we measured the direction of each saccade that immediately followed fixation on the actor's face. The results for this analysis are depicted in Figure 3. In this analysis, regions were defined around each fixation point on the face in each scene, with Region 1 always representing the direction in which the focused object was located with respect to that point, and the other directions defined clockwise in the picture plane from that direction. Each saccade from the actor's face was then assigned to one of these nine discrete 40o regions. The value for each region represents the percentage of saccades that fell into it, when subjects were saccading away from the face.

A one-way ANOVA revealed that the pattern of saccade proportions was different across the regions ($F(8, 468) = 71.92$, $p < 0.01$). A Dunnett's pairwise multiple comparison t-test (comparing each region to Region 1) indicated that the largest proportion of saccades fell into Region 1, with the exception of Region 2 (all comparisons with Region 1 were significant, with the exception of Region 2). These results strongly indicate that observers were most likely to move their own eyes in the general direction of the depicted actor's gaze after fixating the actor's face than in another direction.

To investigate the dynamics of eye movements between the focused object and the face, we calculated the percentage of times that the focused object was fixated soon after fixating the face (within 4 fixations). Given that the face was fixated, the focused object was fixated within 4 fixations on 35% of the trials. This value compares to fixating the non-focused object on 8% of the trials ($t = 10.91$, $p < .01$) and fixating the control region on 4% of the trials ($t = 8.97$, $p < .01$) following fixation on the face.

There was also a strong tendency for participants to move back and forth from the focused object to the face, as shown by the fact that given the focused object was fixated, the face was fixated within the next 4 fixations on 43% of the trials. These eye movement patterns illustrate that fixation tended to move between the face and the focused object. As discussed above, due to the complexity of scenes and multiple object candidates that could be the focus of the actor's gaze, it seemed that viewers were adopting a checking strategy to make sure that they had found the object that the actor was focused on. This check required moving from the object back to the face to see if they had correctly determined the actor's gaze direction.

3 Discussion

The purpose of the present study was to explore whether a viewed person's gaze direction is an important factor in determining the eye movements of an observer during real-world scene exploration. We also sought to determine whether a viewed person's face would attract fixations in the presence of other salient objects within a real-world scene as an enabling condition for asking our primary question. Given the strong social cues that another person's gaze provides, we predicted that despite the increased complexity of scenes, an actor's gaze would direct an observer's fixation to the region of focus. To investigate these questions, we presented participants with a series of scene photographs in the form of a slide show that told a story. We asked whether a depicted actor's face would tend to be fixated, and if so, whether following fixations would be affected by the direction of the actor's gaze.

There were two main sets of results. First, observers showed a strong tendency to fixate the actor's face when it was present in the scene. More specifically, the face of the depicted actor was more likely to be fixated than control regions. Furthermore, when the face was fixated, it was fixated sooner, for longer duration, and more often than control regions. These results support the observation that has been made by other investigators in qualitative analyses of eye movements during scene viewing: Observers tend to look at faces when they are present in scenes (e.g., [8,15,54]). Unlike prior studies, however, the present results provide initial quantitative support for this conclusion.

Second, the main finding of the present study was that the eye movements of an observer were influenced by the direction of gaze of the viewed actor. We found that the object that was the focus of the actor's gaze was fixated on a larger percentage of trials than any other defined region, including the actor's face. Furthermore, given that it was fixated, the focused object was fixated with shorter latency, more often, and for a longer duration, than any other scene region including the actor's face. It is interesting that it was fixated sooner than the actor's face because it suggests that observers were picking up on other gaze cues (such as head and body position) to determine the focus of the actor's gaze. These results suggest that the focus of another person's gaze has a strong influence on an observer's own gaze behavior. In addition, when the gaze of the actor was directly fixated, participants were more likely to look in the direction of the actor's gaze than in any other direction. This is the first demonstration that an observer's eye movements are influenced by the direction of gaze of an actor in an image of a real-world scene.

An important current topic in eye movement control is the degree to which eye movements in real-world scenes are determined by stimulus properties versus by top-down factors such as the meaning of the scene and the task of the viewer [20,22]. Current saliency models of eye movement control in scenes place the emphasis on stimulus properties such as discontinuities in luminance and color [27]. While these types of models do a reasonable job of predicting eye movements under some conditions, particularly when the images are not meaningful and/or the viewing task is not active [45], the present study suggests that when the observer is engaged in an active viewing task, eye movements are driven by the need to interpret the meaning of the scene. In the present case, scene interpretation was strongly tied to looking at a human actor's face, determining the actor's own direction of gaze, and following that gaze to understand what that actor was looking at. Therefore, it seems that social cues provided in the form of perceiving an actor's gaze plays a large role in the direction of an observer's gaze, even within complex, real-world scenes that may have other competing regions of interest.

Acknowledgments. This work was supported by grants from the National Science Foundation (BCS-0094433 and ECS-9873531). Data analysis and manuscript preparation were supported by the Army Research Office (W911NF-04-1-0078); the opinions expressed in this article are those of the authors and do not necessarily represent the views of the Department of the Army or any other governmental organization. Reference to or citations of trade or corporate names does not constitute explicit or implied endorsement of those entities or their products by the author or the Department of the Army). We thank Bryan Corpus for his help with data collection and Jeremy Athy for his role as the janitor. Please address correspondence to MSC, Department of Psychology, Queen's University, Kingston ON, K7L 3N6, Canada or to JMH, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK.

References

1. Ando, S.: Luminance-induced shift in the apparent direction of gaze. Perception. 31, 657–674 (2002)
2. Ando, S.: Luminance-induced shift in the apparent direction of gaze. Perception. 33, 1173–1184 (2004)
3. Argyle, M., Cook, M.: Gaze and mutual gaze. Cambridge University Press, Cambridge, England (1976)

4. Bayliss, A.P., di Pellegrino, G., Tipper, S.P.: Orienting of attention via observed eye gaze is head-centred. Cognition 94, B1–B10 (2004)
5. Baron-Cohen, S.: Mindblindness: An essay on autism and theory of mind. MIT Press, Cambridge, MA (1995)
6. Barresi, J., Moore, C.: Intentional relations and social understanding. Behavioral and Brain Sciences 19, 107–154 (1996)
7. Brooks, R., Meltzoff, A.N.: The importance of eyes: How infants interpret adult looking behavior. Developmental Psychology 38, 958–966 (2002)
8. Buswell, G.T.: How people look at pictures. The University of Chicago press, Chicago, IL (1935)
9. Butler, S.C., Caron, A.J., Brooks, R.: Infant understanding of the referential nature of looking. Journal of Cognition and Development 1, 359–377 (2000)
10. Cline, M.G.: The perception of where a person is looking. American Journal of Psychology 80, 41–50 (1967)
11. Corkum, V., Moore, C.: The origins of joint visual attention in infants. In: Moore, C., Dunham, P.J. (eds.) Joint attention: Its origins and role in development, pp. 61–83. Lawrence Erlbaum, Hillsdale, NJ (1998)
12. D'Entremont, B., Hains, S., Muir, D.: A demonstration of gaze following in 3- to 6-month olds. Infant Behavior and Development 20, 569–572 (1997)
13. Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., Baron-Cohen, S.: Gaze perception triggers visuospatial orienting by adults in a reflexive manner. Visual Cognition 6, 509–540 (1999)
14. Emery, N.J.: The eyes have it: the neuroethology, function and evolution of social gaze. Neuroscience and Biobehavioral Reviews 24, 581–604 (2000)
15. Friedman, A.: Framing pictures: the role of knowledge in automatized encoding and memory for gist. Journal of Experimental Psychology: General 108, 316–355 (1979)
16. Friesen, C.K., Kingstone, A.: The eyes have it!: Reflexive orienting is triggered by nonpredictive gaze. Psychonomic Bulletin and Review 5, 490–495 (1998)
17. Friesen, C.K., Kingstone, A.: Abrupt onsets and gaze direction cues trigger independent reflexive attentional effects. Cognition 87, B1–B10 (2003)
18. Friesen, C.K., Ristic, J., Kingstone, A.: Attentional effects of counterpredictive gaze and arrow cues. Journal of Experimental Psychology-Human Perception and Performance 30(2), 319–329 (2004)
19. Gibson, J.J., Pick, A.D.: Perception of another person's looking behavior. American Journal of Psychology 76, 386–394 (1963)
20. Henderson, J.M.: Human gaze control in real-world scene perception. Trends in Cognitive Sciences 7, 498–504 (2003)
21. Henderson, J.M., Brockmole, J.R., Castelhano, M.S., Mack, M.: Image salience versus cognitive control of eye movements in real-world scenes: Evidence from visual search. In: van Gompel, R., Fischer, M., Murray, W., Hill, R. (eds.) Eye movement research: Insights into mind and brain, pp. 537–562. Elsevier, Oxford (2007)
22. Henderson, J.M., Ferreira, F.: Scene perception for psycholinguists. In: Henderson, J.M., Ferreira, F. (eds.) The interface of language, vision, and action: Eye movements and the visual world, pp. 1–58. Psychology Press, New York (2004)
23. Henderson, J.M., Weeks Jr., P.A., Hollingworth, A.: Effects of semantic consistency on eye movements during scene viewing. Journal of Experimental Psychology: Human Perception and Performance 25, 210–228 (1999)

24. Henderson, J.M., Williams, C.C., Castelhano, M.S., Falk, R.J.: Eye movements and picture processing during recognition. Perception & Psychophysics 65, 725–734 (2003)
25. Hietanen, J.K., Nummenmaa, L., Nyman, M.J., Parkkola, R., Hamalainen, H.: Automatic attention orienting by social and symbolic cues activates different neural networks: An fMR1 study. Neuroimage 33(1), 406–413 (2006)
26. Hollingworth, A., Henderson, J.M.: Accurate visual memory for previously attended objects in natural scenes. Journal of Experimental Psychology: Human Perception and Performance 28, 113–136 (2002)
27. Itti, L., Koch, C.: Computational modeling of visual attention. Nature Reviews: Neuroscience 2, 194–203 (2001)
28. Kingstone, A., Smilek, D., Ristic, J., Friesen, C.K., Eastwood, J.D.: Attention, Researchers! It is time to take a look at the real world. Current Direction in Psychological Science 12(5), 176–184 (2003)
29. Kellerman, J., Lewis, J., Laird, J.D.: Looking and loving: The effects of mutual gaze on feelings of romantic love. Journal of Research in Personality 23, 145–161 (1989)
30. Kleinke, C.L.: Gaze and eye contact: A research review. Psychological Bulletin 100, 78–100 (1986)
31. Langdon, R., Smith, P.: Spatial cueing by social versus nonsocial directional signals. Visual Cognition 12(8), 1497–1527 (2005)
32. Langton, S.R.H.: The mutual influence of gaze and head orientation of the analysis of social attention direction. The Quarterly Journal of Experimental Psychology A 53, 825–845 (2000)
33. Langton, S.R.H., Bruce, V.: Reflexive social orienting. Visual Cognition 6, 541–567 (1999)
34. Langton, S.R.H., O'Donnell, C., Riby, D.M, Ballantyne, C.J.: Gaze cues influence the allocation of attention in natural scene viewing. Quarterly Journal of Experimental Psychology 59 (12), 2056–2064 (2006)
35. Loftus, G., Mackworth, N.: Cognitive determinants of fixation location during picture viewing. Journal of Experimental Psychology: Human Perception and Performance 4(4), 565–572 (1978)
36. Lungarella, M., Metta, G., Pfeifer, R., Sandini, G.: Developmental robotics: a survey. Connection Science 15, 151–190 (2003)
37. Mackworth, N.H., Morandi, A.J.: The gaze selects informative details within pictures. Perception & Psychophysics 2, 547–552 (1967)
38. Mansfield, E.M., Farroni, T., Johnson, M.H.: Does gaze perception facilitate overt orienting? Visual Cognition 10, 7–14 (2000)
39. Maurer, D.: Infants' perception of facedness. In: Field, T., Fox, M. (eds.) Social Perception in Infants, Ablex, Norwood, NJ (1995)
40. Morita, A., Yoshikawa, Y., Hosoda, K., Asada, M.: Joint attention with strangers based on generalization through joint attention with caregivers. In: IROS. Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 4, pp. 3744–3749 (2004)
41. Morton, J., Johnson, M.: CONSPEC and CONLEARN: A two-process theory of infant face recognition. Psychological Review 98, 164–181 (1991)
42. Nagai, Y., Asada, M., Hosoda, K.: A developmental approach accelerates learning of joint attention. In: ICDL 2002. Proceedings of the 2nd International Conference on Development and Learning, pp. 277–282 (2002)

43. Nagai, Y., Hosoda, K., Asada, M.: Joint attention emerges through bootstrap learning. In: Proceedings of the IEEE/RSJ International Conference on Interlligent Robots and Systems, pp. 168–173 (2003)
44. Nelson, W.W., Loftus, G.R.: The functional visual field during picture viewing. Journal of Experimental Psychology: Human Learning and Memory 6(4), 391–399 (1980)
45. Parkhurst, D., Law, K., Niebur, E.: Modeling the role of salience in the allocation of overt visual attention. Vision Research 42, 107–123 (2002)
46. Perrett, D., Emery, N.J.: Understanding the intentions of others from visual signals: Neuropsychological evidence. Cahiers de Psychologie Cognitive 13, 683–694 (1994)
47. Ricciardelli, P., Briocolo, E., Aglioti, S.M., Chelazzi, L.: My eyes want to look where your eyes are looking: Exploring the tendency to imitate another individual's gaze. Neuroreport 13, 2259–2264 (2002)
48. Scaife, M., Bruner, J.S.: The capacity for joint visual attention in the infant. Nature 253, 265–266 (1975)
49. Scassellati, B.: Theory of mind for a humanoid robot. Autonomous Robots 8(1), 13–24 (2002)
50. Smilek, D., Birmingham, E., Cameron, D., Bischof, W., Kingstone, A.: Cognitive ethology and exploring attention in real-world scenes. Brain Research 1080, 101–119 (2006)
51. Smilek, D., Eastwood, J.D., Reynolds, M.G., Kingstone, A.: Metacognitive errors in change detection: Missing the gap between lab and life. Consciousness and Cognition 16(1), 52–57 (2007)
52. Torralba, A., Oliva, A., Castelhano, M.S., Henderson, J.M.: Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. Psychological Review 113, 766–786 (2006)
53. von Grünau, M., Anston, C.: The detection of gaze direction: A stare-in-the-crowd effect. Perception 24, 1297–1313 (1995)
54. Yarbus, I.A.: Eye movements and vision. Plenum Press, New York (1967)