Gaze cueing in naturalistic scenes under top-down modulation - A conceptual replication

Jonas Großekathöfer1, Kristina Suchotzki1, & Matthias Gamer1

1 Julius-Maximilian University, Würzburg

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

Correspondence concerning this article should be addressed to Jonas Großekathöfer,

Marcusstraße 9-11, 97070 Würzburg. E-mail: [jonas.grossekathoefer@uni-wuerzburg.de](mailto:jonas.grossekathoefer@uni-wuerzburg.de)

# Abstract

Humans as social beings rely on information provided by conspecifics. One important signal in social communication is eye gaze. The current study (*n*=90) sought to replicate and extend previous findings of attentional guidance by eye gaze in complex everyday scenes. In line with previous studies, longer, more and earlier fixations for objects referenced by gaze were observed in free viewing conditions. To investigate how robust this prioritization is against top-down modulation, half of the observers received a memory task that required scanning the whole scene instead of exclusively focusing on *referenced* objects. Interestingly, similar gaze cueing effects occurred in this condition. Moreover, the human beings depicted in the scene received a large amount of attention even though they were irrelevant to the current task. These results indicate that the mere presence of other human beings as well as their gaze orientation have a strong impact on attentional exploration.

*Keywords:* ?keywords?

Word count: ?wordcount

Gaze cueing in naturalistic scenes under top-down modulation - A conceptual replication

# Introduction

Humans in their social environment rely on the information conspecifics provide. This does not only hold for reading explicit signals, as in conversations, but also for implicit signals, as in gazes. Specifically, if an individual looks into a certain direction, this information is often read spontaneously by an observer and redirects his or her attention towards the referred object or location. Such guidance of someone else’s attention is also called gaze cuing, whereas the process of following someone’s gaze is often referred to as joint attention.

The most frequently used paradigm to investigate such attentional shifts is the so-called gaze-cueing paradigm (Driver et al., 1999; Friesen & Kingstone, 1998; Langton & Bruce, 2000). Studies using this paradigm show that perceived gaze cues like faces/eyes looking at a certain location (?) lead to reflexive attentional shifts, which can result in a processing benefit for the looked at locations (for a very recent study see Langton, McIntyre, Hancock, & Leder, 2017). Even though gaze cues are crucial for joint attention, the standard gaze-cueing paradigm is criticized for lacking ecological validity, because (among others) these studies use isolated heads (Friesen & Kingstone, 1998; Langton & Bruce, 2000) or even cartoon heads (Driver et al., 1999; Ristic & Kingstone, 2005) as gaze cues (for an overview of this criticism see: Risko, Laidlaw, Freeth, Foulsham, & Kingstone, 2012). For example, in a recent study Hayward, Voorhies, Morris, Capozzi, and Ristic (2017) did not find reliable links between classical gaze cueing tasks and real social engage.

The relevance of ecologically (?) valid stimuli is a core aspect of social attention research (Itier & Batty, 2009). Social attention research investigates attentional consequences of social interactions and focuses often on similarities and differences of different types of social stimuli (Risko et al., 2012), from real human interactions (Freeth, Foulsham, & Kingstone, 2013; for example: Laidlaw, Foulsham, Kuhn, & Kingstone, 2011) to highly controlled laboratory settings with isolated faces (for example: Langton et al., 2017).

*To account for these issues in gaze cuing paradigms, Zwickel and Võ (2010) conducted a joint attention study with a full person (instead of an isolated face) as directional cue. In this study, the authors used a free viewing instruction, meaning that participants had no explicit task to fulfil but were simply instructed to freely explore the stimuli. Zwickel and Võ (2010) argue, that the lack of a specific task puts gaze following to a stricter test, as ….. This is explicitly in contrast to a study of Castelhano, Wieth, and Henderson (2007) who asked participants to understand the story depicted in their stimuli (?). Additionally, without a task represents naturalistic viewing behavior and therefore adds more ecological validity to classical gaze cuing paradigms. They presented their participants multiple scenes, one after the other (?), each for several seconds. Each scene contained either a person or a loudspeaker, and two objects. In each scene the person or the loudspeaker were directed towards one of the two objects (subsequently called the referenced object). Results of the study showed that participants fixated the referenced object remarkably earlier, more often and longer than the not referenced object. However, the prioritization of the object occurred only when the person referenced the object. By showing that leaving saccades from the head (but not from the loudspeaker) landed most often onto the referenced object the results give direct support for the relation between cue type and object role. The attentional focus of the person in the scene guided attentional distribution of the observer. Interestingly, the same was not true for the loudspeaker. The referenced objects were not just focused because they might have been salient by themselves (due to e.g. positioning) or they were referenced by another object, but became more salient only if they were references by a person. Thereby, in their study Zwickel and Võ (2010) provide evidence that joint attention is a direct consequence of gaze cues, happens spontaneously and has high relevance even in more realistic situations.* It remains unclear, however, how robust these joint attention effects are. So, our conceptual replication of Zwickel and Võ’s work aims at replicating and extending the findings of Zwickel and Võ’s (2010).

To extend this line of research and investigate the robustness of the findings of Zwickel & Vo (2010), the influence of top-down modulations on joint and social attention in naturalistic scenes is investigatedshowed that social attention is influenced by multiple factors like social status (Foulsham, Cheng, Tracy, Henrich, & Kingstone, 2010) or expectations of … (Perez-Osorio, Müller, Wiese, & Wykowska, 2015). These studies have in common that they manipulate viewing behavior of the observer by manipulating the stimuli. For example, Foulsham et al. (2010) build the stimulus set from stimuli that were previously rated for social status and confirmed the predicted shift in attention with eye-tracking measures.

In the present study, however, viewing behavior should not be manipulated via the stimuli but via explicit instructions given to observers. All participants first completed a viewing task in which the they saw… and then a memory task in which they …. In the explicit encoding group, participants received the explicit information that the viewing task will be followed by a memory test. In contrast, in the free viewing group, participants did not receive this instruction before the viewing task but only right before the memory task.. The instruction manipulation is thought to induce a more explicit and systematic encoding of the presented scenes, specifically towards all objects in the explicit encoding group.

Additionally, the consequences for latter cognitive processes, here memory effects, are examined for the different objects and groups (?). Observers from the free viewing group who did not receive any instruction should provide an unaffected viewing behavior, just as Zwickel and Võ (2010) demonstrated. Both groups had to recall as many as things as possible from the scenes in a free-recall memory test. With this manipulation it is thought to demonstrate top-down influence on social attention aspects, but also its influence on joint attention. As Zwickel and Võ (2010) showed, a person in a scene influences the viewing behavior of observers spontaneously and without further instruction or manipulation.

In the given study, the motivation for the manipulation of the instructions for the memory test was twofold. First, it is thought to test robustness of the social and joint attention effects and second joint attentional effects on memory can be observed. The influence of the manipulation is expected to reduce spontaneous viewing behavior and foster a more systematic processing of the scene. It is important to note, that it is not expected that the influence of the presence of the person vanishes completely for social and joint attentional effects due to top-down modulation. It is supposed to be weaker, because it does not represent the optimal strategy for the observer. Therefore, the voluntary part of the scene processing might be overwritten. However, as social and joint attentional effects are expected to be (partially) independent of volition and occur spontaneously and reflexive the effects should still be visible regardless for the social attention part and also for the joint attention part of this study. Specifically, prioritization effects for the head should be smaller but still visible in the explicit encoding group. For the joint attention measures a decreased prioritization for the referenced object is expected for the explicit encoding group is decreased compared with the not referenced object. An optimal strategy for observer in the explicit encoding group would be to ignore completely the person in the scene, resulting in smaller differences in attentional prioritization between the two object roles. Additionally, memory effects sensitive for the role of the object were examined, to answer the question whether more attentional resources on an object pays off with enhanced memorability. This is mainly interesting for the free viewing group with the unbiased viewing behavior, as it is expected that both objects bind comparable amount of attention due to the more systematic processing in the explicit encoding group.

Zwickel and Võ (2010) used computer-generated stimuli. Here, stimuli that are more naturalistic and a bigger sample size is used. Consequently, real photographs are chosen over computer rendered scenes. As often, by being more naturalistic experimental control is reduced. The consequences are hold to a minimum by producing the stimuli the same way. In particular, each scene was photographed four times, with the individual looking twice to the left and right to each object. Although, as much as possible was controlled for in the photographs, real photographs contain small unavoidable changes. Whereas Zwickel and Võ (2010) rotate the figure in the computer rendered scene, they have complete control of all the changes, e.g. angel of the body or facial expression. Here, another photography was taken. As consequence, for example, the body orientation within the four balanced scenes cannot be perfectly controlled and might differ slightly between photographs within scenes. With higher power and stimuli that are more naturalistic this study will underline Zwickel and Võ (2010) findings. However, it is presumed that the effect showed by Zwickel and Võ (2010) is genuine for social stimuli. Therefore, no non-social condition comparable to the loudspeaker-condition was tested.

The foundation of this study is to examine top-down modulation on gaze cueing and to replicate the findings from Zwickel & Võ (2010).

*H1*. Consequently, a prioritization for the *referenced* objects is predicted. In line with Zwickel and Võ (2010), this prioritization is measured in multiple ways. First, there should be an early fixation bias towards the *referenced* object. During presentation time, the total time that the *referenced* object is fixated should be prolonged, with more fixations as well, compared to the not-referenced object. For saccades it is expected, that saccades leaving from the head are is more likely to move to the *referenced* object, in contrast to the not-*referenced* object.

*H2*. New, compared to Zwickel and Võ’s (2010) study, is the instruction condition. For gaze following it is expected that the prioritization of the *referenced* object will decrease for subjects in the explicit encoding group, due to the instruction of the memory test and a more explicit and systematic processing of the scene. This means smaller differences between referenced and not-referenced objects all in the fixation measurements are expected. For leaving saccades it is also expected that the under H1 expected difference between referenced and not-referenced objects decreases.

*H3*. For the follow-up memory test it is assumed, that observer with announced memory test (in the *explicit encoding group*) will recall more items than observes in the … group (?), because they process the scenes more thorough. Additionally it is expected, that in the free-recall condition *referenced* objects are better recalled than not-*referenced* objects.

*H4*. Additionally to the joint attentional effects stated in H1, the basic effects of social attention are expected. This means, that the head will be prioritized over the body. This prioritization can be measured, in multiple ways. Again, it is expected, that first fixation fall earlier on the head, that it is fixated longer and more often than the body. Additionally, fixations occur earlier on the head, than on the *referenced* object, because the gaze-cue needs to be processed in advance.

*H5*. Due to the instruction of the memory test, and the assumed change in processing the scene, the viewing behavior should change for social stimuli. Due to explicit encoding, it is expected that natural viewing behavior that is known to prioritize the head is reduced and therefore the head loses some of its salience. As a consequence, the head should be prioritized stronger in the free viewing group compared to the explicit encoding group, resulting in longer fixations, faster fixations and more fixations.

# Methods

## Participants

In this study 94 observers (65 female and 29 male) between 18 and 55 years (*M* = 24.73, *SD* = 5.04) participated voluntarily. All observer had normal or corrected vision and were recruited at the University of Würzburg’s online subject pool or by blackboard. For participation observers received study credit or 5€. One observer was excluded due to missing data.

## Stimuli and Apparatus

The experimental stimuli consisted of pictures of 26 different scenes with a single individual in the center looking at one of two objects of interests in the right or left half of the picture. The direction of the gaze and the place of the objects were balanced, creating 104 unique naturalistic scenes in the end. For each participant, a set was randomly generated from this pool containing one version of each scene, resulting in 26 trials Eye movements were tracked with the corneal-reflexion method and were recorded with an EyeLink1000 tower system, sampling at 1000 Hz. The eye tracker and stimuli were controlled by the software Presentation® (Neurobehavioral Systems).

## Design and Procedure

The experimental design was a 2 x 2 mixed observer design. First, as a two-level factor the instruction was manipulated between subjects (free viewing vs. explicit encoding). Additionally, as a two-level within subjects factor the role of the object was manipulated (referenced vs. not-referenced).

First, observers were asked to give full informed consent. Then the eye-tracker was calibrated for the observer. According to the instruction manipulation, only half of the observers were told that there was a follow-up memory test. All observers were then told to look at the scenes as they would look at photographs. The presentation order of the scenes was randomized and the roles of the objects and their location were completely balanced and controlled. In each trial, the scene was presented for 10 seconds. The inter trial interval was randomized between 2 and 4 seconds. After the last trial participants filled in questionnaires (demographics, autism-questionnaire (short), and Inventar soziale kompetenz). To prevent primacy and recency effects the questionnaire session was also used as a puffer for the following memory task. Afterward filling in the questionnaires, all observers were asked to recall as many items as possible in a free recall memory-test. Observers received credit or payment afterwards.

## Data analysis

For the analysis, standard configuration of SR Research’s EyeLink DataViewer software was used to categorize eye movements into saccades and fixations. Saccades were defined as eye movements exceeding a velocity threshold of 30°/sec or an acceleration threshold of 8.000°/sec². Fixations were defined as time periods between saccades. Regions-of-Interest (ROI) were hand-drawn around the relevant objects and the face and body of the individual in the center. These ROIs were color-coded for cued and uncued objects and for the head and body of the individual on the pictures to determine gaze locations.

R (3.3.1, R Core Team, 2016) and the R-packages afex (0.18.0, Singmann, Bolker, Westfall, & Aust, 2017), bindrcpp (0.2, K. Müller, 2017), car (2.1.6, Fox & Weisberg, 2011), dplyr (0.7.4, Wickham, Francois, Henry, & Müller, 2017), estimability (1.2, R. Lenth, 2016), forcats (0.2.0, Wickham, 2017a), ggplot2 (2.2.1, Wickham, 2009), ggpubr (0.1.6, Kassambara, 2017), lme4 (1.1.14, Bates, Mächler, Bolker, & Walker, 2015), lsmeans (2.27.61, R. V. Lenth, 2016), magrittr (1.5, Bache & Wickham, 2014), Matrix (1.2.12, Bates & Maechler, 2017), papaja (0.1.0.9492, Aust & Barth, 2017), purrr (0.2.4, Henry & Wickham, 2017), readr (1.1.1, Wickham, Hester, & Francois, 2017), stringr (1.2.0, Wickham, 2017b), tibble (1.3.4, K. Müller & Wickham, 2017), tidyr (0.7.2, Wickham & Henry, 2017), and tidyverse (1.2.1, Wickham, 2017c) are used for all analyses.

# Results

## Joint attention

As a measure of prioritization, fixation duration, fixation number, fixation latency, as well as leaving saccades are used. For all measures, higher values indicate prioritization.

### Fixations.

***Fixation duration.*** The duration of all fixations is cumulated for each object. Specifically, the cumulative time (in ms) a fixation rested on an object was divided by the total time spent fixating the rest of the scene. By that, a relative measures for fixation duration was gained. Observer show a bias towards the *referenced* object, which was fixated longer (see Table/Figure). As predicted the 2 x 2 ANOVA on fixation duration revealed a significant main effect for reference, *F*(1*,*91) = 6*.*86, *p* = *.*010, *ηG*2 = *.*017, with a longer fixation duration for … compared to ….. It also revealed a main effect for condition *F*(1*,*91) = 18*.*02, *p < .*001, ηG2 = *.*133, with a longer fixation duration for … compared to …. Against prediction there was no significant two-way interaction, *F*(1*,*91) = 2*.*47, *p* = *.*120, *ηG*2 = *.*006.

***Fixation latency.*** An additional measurement of prioritization is fixation latency. This describes for each object the mean time out of all first fixations per object each trial. This measure shows that objects are fixated earlier when they are *referenced*, but also when the observer was in the explicit encoding condition. The 2 x 2 ANOVA on fixation latency revealed a significant main effect for reference *F*(1*,*91) = 43*.*38, *p < .*001, *ηG*2 = *.*140, with an earlier fixation of … compared to …, and a significant main effect of condition, *F*(1*,*91) = 33*.*33, *p < .*001, *ηG*2 = *.*194, , with an earlier fixation of … compared to …,. The interaction shows a trend towards significance, *F*(1*,*91) = 3*.*37, *p* = *.*070, *ηG*2 = *.*013, suggesting a smaller difference for reference for the *explicit encoding group* compared to …..

***Fixation number.*** As a third measurement of prioritization, fixation number, as the count of fixations per object, divided by the total number of fixations during scene presentation, was calculated. The 2 x 2 ANOVA on fixation number showed … It shows, that objects were fixated more often when *referenced*, *F*(1*,*91) = 10*.*18, *p* = *.*002, *ηG*2 = *.*019. Objects were also more often fixated in the explicit encoding condition, and again against prediction there was no interaction, *F*(1*,*91) = 2*.*09, *p* = *.*152, *ηG*2 = *.*004.

### Saccades.

***Leaving saccades.*** The second hypothesis predicts that there are more saccades leaving from the head to the *referenced* object. Furthermore it is claimed, that an interaction between instruction and reference results in increasing differences in the free viewing group. The first claim finds statistical support from an Anova. There is a significant main effect for reference, *F*(1*,*91) = 39*.*87, *p < .*001, *ηG*2 = *.*151, and for group, *F*(1*,*91) = 25*.*14, *p < .*001, *ηG*2 = *.*141. No confirmation was found for the predicted interaction between reference and instruction *F*(1*,*91) = 1*.*41, *p* = *.*239, *ηG*2 = *.*006.

### Memory.

The free-recall memory test showed, that observers who received the instruction of the memory test before the free viewing remembered more items than observers from the other group. An Anova showed its statistical significance, *F*(1,92) = 33.23, *p* < .001, ηG2 = .234. However, against prediction, referencing an object did not influence memory performance, F(1,92) = 0.43, p = .516, ηG2 = .001. There was also no interaction, F(1,92) = 0.02, p = .878, ηG2 = .000, such that referencing an object would increase memorability only for one group, as it is stated in the hypothesis for the free viewing group.

## Social attention

A similar pattern to the objects can be seen when comparing the head with the body region. Again, there were the same measurements to account for prioritization. For the social attention part, Anovas are conducted for the same dependent variables indicating prioritization, but the head is compared against the body region.

**Face vs body fixation (H4).** The first measurement shows, that head regions were fixated longer than the body (Table X), *F*(1*,*91) = 175*.*08, *p < .*001, *ηG*2 = *.*467, but there is no difference between the instruction-groups, *F*(1*,*91) = 2*.*70, *p* = *.*104, *ηG*2 = *.*016, and no interaction between the two factors, *F*(1*,*91) = 3*.*25, *p* = *.*075, *ηG*2 = *.*016.

The same pattern can be seen for fixation number, with a strong effect for region, with more fixations on the head, *F*(1*,*91) = 144*.*50, *p < .*001, *ηG*2 = *.*415 and no difference between the groups, *F*(1*,*91) = 1*.*25, *p* = *.*266, *ηG*2 = *.*008, and no interaction *F*(1*,*91) = 2*.*03, *p* = *.*158, *ηG*2 = *.*010.

Latencies of fixation differed also remarkably between head and body, but only for region *F*(1*,*91) = 125*.*00, *p < .*001, *ηG*2 = *.*446. There was neither an effect for instruction, *F*(1*,*91) = 0*.*84, *p* = *.*361, *ηG*2 = *.*004 nor was there an interaction, *F*(1*,*91) = 1*.*29, *p* = *.*259, *ηG*2 = *.*008.

**Free viewing vs explicit encoding for face fixations (H5).** A post Hoc t-test between head fixations for each group provides information whether heads are differently processed dependent on the instruction observers get. It reveals, that only fixation latency differed significantly between groups, ∆*M* = 241*.*85, 95% CI [*−*457*.*22, *−*26*.*48], *t*(79*.*99) = *−*2*.*23, *p* = *.*028, with faster fixations for the free viewing group (*M* = , *SD* =) compared to the …group (*M* = , *SD* =). For fixation duration there was only a trend for difference between groups, ∆*M* = *−*0*.*03, 95% CI [0*.*00, 0*.*07], *t*(88*.*82) = 1*.*84, *p* = *.*069. In addition, fixation number did not differ, ∆*M* = *−*0*.*02, 95% CI [*−*0*.*01, 0*.*05], *t*(88*.*80) = 1*.*41, *p* = *.*161.

# Discussion

The main attempt to replicate Zwickel and Võ (2010)’s findings on attentional prioritization from the mere presence of a person in a scene on gaze behavior was successful. As in Zwickel and Võ (2010) and others studies (for example Birmingham, Bischof, & Kingstone, 2008), a strong prioritization can be seen for the head of the person in the scene. Accordingly, the head was fixated first, before any other ROI. Besides that, the heads were also fixated more often, meaning that during scene presentation observers kept looking at the head. Consequently, no ROI was fixated longer in total than the head. Additionally, also in line with Zwickel and Võ (2010) *referenced* objects were preferred over the object not being *referenced* in a given trial. The preference for the *referenced* object is consistent over all measures. First, the object that was referencedwas fixated remarkably earlier. Furthermore, objects were also fixated longer with more frequent fixations when they were *referenced*. Therefore, the gaze cue by the person does not only lead to more thorough processing during the whole time of scene presentation, but also guides attentional resources very early. All these findings together indicate a strong prioritization of objects referenced through the gaze cue of the person in the scene. Moreover, the consistent prioritization of the head and the *referenced* object indirectly suggest a link between these two regions. To investigate a direct relation leaving saccades from the head are examined as well. It is shown that saccades leaving the head are more likely to end on the *referenced* object. This underlines a direct link between the head and the *referenced* object. So far, the given results replicate all aspects Zwickel and Võ (2010) demonstrated for the presence of a person compared to a loudspeaker in their study. However, this replication included a top-down modulation, used photographic scenes and had considerably more power than the original study. Although it is unclear how the photographs changed the results compared to the computer rendered scenes, the higher power lead to more precise estimation of the effect sizes. Note, however, that this was not a direct but a conceptual replication, showing that Zwickel and Võ (2010) findings can be generalized from computer rendered to photographic scenes. Besides replicating, this study aimed also at extending the line of research, generalizing it to new stimuli and at testing the robustness of the pattern in gaze behavior as it was shown here against top-down modulation, specifically an instructed task prior the viewing of the pictures. Unsurprisingly the task scores show, that observer, who knew about the free-recall task in advance perform better in recalling items. More interestingly, viewing behavior was different depending on instruction group. Specifically, observer pay in general more attention to objects when they received the instruction prior to the viewing part, and at the same time, the head loses some of its natural salience. Both findings are consistent over all measures. These findings suggest that observers with instruction had a more systematic gaze pattern, where they preferred all task relevant objects. **However, no interaction between top-down modulation and reference of an object can be found.**

Consistent for social attention as well as joint attention, prioritization in all measurements remained stable also in the memory task instruction group. The prioritization shown for the head and for the *referenced* object was unaffected by the given top-down modulation. The attentional guidance of social and joint attentional processes can be seen, even when observers investigate the scene more systematic. This evidence provides support for the automaticity and reflexivness of joint and social attentional effects, even when the observer voluntarily aims at scanning the scene systematically. Nevertheless, observer with instructed free-recall task performed better in the subsequent memory test (?). However, the contribution of the attentional reallocation remains unclear. In particular, **whether an object was cued or not by the person, did not influence its probability being recalled**. This is in contrast to the increased attention referenced objects receive. That is partially inconsistent with the hypothesis, which stated correct that scene processing changes, but the actual link is missing that the change in processing for solving the task can be explained from the attentional reallocation of attention. Even as it is assumed, these findings do not indicate a link between spent attentional resources and performance for recall. Overall, the results provide additional support to previous findings that attention is shifted reflexively where other persons are looking at (e.g. Ristic & Kingstone, 2005, Hayward et al. 2017). This evidence, which was previously extended to free viewing of static naturalistic scenes by Zwickel and Võ (2010), is shown to be robust against top-down modulation. Even when attentional allocation changes due to a more systematic viewing pattern, social and joint attentional shifts are still affected by the mere presence of a person, comparable as if the observers perform unbiased free viewing. All in all, our results indicate that the mere presence of other human beings as well as their gaze orientation have a strong impact on attentional exploration. The observed attentional guidance of the gaze was so robust to resist even top-down modulation. It is concluded that attentional guidance by social attentional and joint attentional trigger is very robust.

# References

Aust, F., & Barth, M. (2017). papaja: Create APA manuscripts with R Markdown. Retrieved from <https://github.com/crsh/papaja>

Bache, S. M., & Wickham, H. (2014). Magrittr: A forward-pipe operator for r. Retrieved from <https://CRAN.R-project.org/package=magrittr>

Bates, D., & Maechler, M. (2017). Matrix: Sparse and dense matrix classes and methods. Retrieved from <https://CRAN.R-project.org/package=Matrix>

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48. doi:[10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Gaze selection in complex social scenes. Visual Cognition, 16(2-3), 341–355. doi:[10.1080/13506280701434532](https://doi.org/10.1080/13506280701434532)

Castelhano, M. S., Wieth, M., & Henderson, J. M. (2007). I see what you see: Eye movements in real-world scenes are affected by perceived direction of gaze. In International workshop on attention in cognitive systems (Vol. 3, pp. 251–262). Springer; Association for Research in Vision; Ophthalmology (ARVO). doi:[10.1167/3.9.307](https://doi.org/10.1167/3.9.307)

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. Visual Cognition, 6(5), 509–540. doi:[10.1080/135062899394920](https://doi.org/10.1080/135062899394920)

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. Cognition, 117(3), 319–331. doi:[10.1016/j.cognition.2010.09.003](https://doi.org/10.1016/j.cognition.2010.09.003)

Fox, J., & Weisberg, S. (2011). An R companion to applied regression (Second.). Thousand Oaks CA: Sage. Retrieved from <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>

Freeth, M., Foulsham, T., & Kingstone, A. (2013). What affects social attention? Social presence, eye contact and autistic traits. PLoS ONE, 8(1), e53286. doi:[10.1371/journal.pone.0053286](https://doi.org/10.1371/journal.pone.0053286)

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. Psychonomic Bulletin & Review, 5(3), 490–495. doi:[10.3758/BF03208827](https://doi.org/10.3758/BF03208827)

Hayward, D. A., Voorhies, W., Morris, J. L., Capozzi, F., & Ristic, J. (2017). Staring reality in the face: A comparison of social attention across laboratory and real world measures suggests little common ground. Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 71(3), 212–225. doi:[10.1037/cep0000117](https://doi.org/10.1037/cep0000117)

Henry, L., & Wickham, H. (2017). Purrr: Functional programming tools. Retrieved from <https://CRAN.R-project.org/package=purrr>

Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. Neuroscience & Biobehavioral Reviews, 33(6), 843–863. doi:[10.1016/j.neubiorev.2009.02.004](https://doi.org/10.1016/j.neubiorev.2009.02.004)

Kassambara, A. (2017). Ggpubr: ’Ggplot2’ based publication ready plots. Retrieved from <https://CRAN.R-project.org/package=ggpubr>

Laidlaw, K. E. W., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. Proceedings of the National Academy of Sciences, 108(14), 5548–5553. doi:[10.1073/pnas.1017022108](https://doi.org/10.1073/pnas.1017022108)

Langton, S. R. H., & Bruce, V. (2000). You must see the point: Automatic processing of cues to the direction of social attention. Journal of Experimental Psychology: Human Perception and Performance, 26(2), 747–757. doi:[10.1037/0096-1523.26.2.747](https://doi.org/10.1037/0096-1523.26.2.747)

Langton, S. R. H., McIntyre, A. H., Hancock, P. J. B., & Leder, H. (2017). Saccades and smooth pursuit eye movements trigger equivalent gaze-cued orienting effects. The Quarterly Journal of Experimental Psychology, (just-accepted), 1–30. doi:[10.1080/17470218.2017.1362703](https://doi.org/10.1080/17470218.2017.1362703)

Lenth, R. (2016). Estimability: Tools for assessing estimability of linear predictions. Retrieved from <https://CRAN.R-project.org/package=estimability>

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. Journal of Statistical Software, 69(1), 1–33. doi:[10.18637/jss.v069.i01](https://doi.org/10.18637/jss.v069.i01)

Müller, K. (2017). Bindrcpp: An ’rcpp’ interface to active bindings. Retrieved from <https://CRAN.R-project.org/package=bindrcpp>

Müller, K., & Wickham, H. (2017). Tibble: Simple data frames. Retrieved from <https://CRAN.R-project.org/package=tibble>

Perez-Osorio, J., Müller, H. J., Wiese, E., & Wykowska, A. (2015). Gaze following is modulated by expectations regarding others’ action goals. PLoS ONE, 10(11), e0143614. doi:[10.1371/journal.pone.0143614](https://doi.org/10.1371/journal.pone.0143614)

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Risko, E. F., Laidlaw, K. E. W., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: Toward an empirical approach to concerns about ecological validity. Frontiers in Human Neuroscience, 6(May), 1–11. doi:[10.3389/fnhum.2012.00143](https://doi.org/10.3389/fnhum.2012.00143)

Ristic, J., & Kingstone, A. (2005). Taking control of reflexive social attention. Cognition, 94(3), B55–B65. doi:[10.1016/j.cognition.2004.04.005](https://doi.org/10.1016/j.cognition.2004.04.005)

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2017). Afex: Analysis of factorial experiments. Retrieved from <https://CRAN.R-project.org/package=afex>

Wickham, H. (2009). Ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. Retrieved from [http://ggplot2.org](http://ggplot2.org/)

Wickham, H. (2017a). Forcats: Tools for working with categorical variables (factors). Retrieved from <https://CRAN.R-project.org/package=forcats>

Wickham, H. (2017b). Stringr: Simple, consistent wrappers for common string operations. Retrieved from <https://CRAN.R-project.org/package=stringr>

Wickham, H. (2017c). Tidyverse: Easily install and load ’tidyverse’ packages. Retrieved from <https://CRAN.R-project.org/package=tidyverse>

Wickham, H., & Henry, L. (2017). Tidyr: Easily tidy data with ’spread()’ and ’gather()’ functions. Retrieved from <https://CRAN.R-project.org/package=tidyr>

Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). *Dplyr: A grammar of data manipulation*. Retrieved from <https://CRAN.R-project.org/package=dplyr>

Wickham, H., Hester, J., & Francois, R. (2017). *Readr: Read rectangular text data*.

## Retrieved from <https://CRAN.R-project.org/package=readr>

Zwickel, J., & Võ, M. L.-H. (2010). How the presence of persons biases eye movements.

*Psychonomic Bulletin & Review*, *17*(2), 257–262. doi:[10.3758/PBR.17.2.257](https://doi.org/10.3758/PBR.17.2.257)