

Combining fMRI and Eye-tracking for the Study of Social Cognition

Kristin Marie Rusch^{1,2,3} 

¹Laboratory for Multimodal Neuroimaging, Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany. ²Department of Neurology and Neurorehabilitation, Hospital zum Heiligen Geist, Academic Teaching Hospital of the Heinrich-Heine-University Düsseldorf, Kempen, Germany. ³Center for Mind, Brain and Behavior (CMBB), University of Marburg and Justus Liebig University Giessen.

Neuroscience Insights
Volume 16: 1–4
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/26331055211065497



ABSTRACT: The study of social cognition with functional magnetic resonance imaging (fMRI) affords the use of complex stimulus material. Visual attention to distinct aspects of these stimuli can result in the involvement of remarkably different neural systems. Usually, the influence of gaze on neural signal is either disregarded or dealt with by controlling gaze of participants through instructions or tasks. However, behavioral restrictions like this limit the study's ecological validity. Thus, it would be preferable if participants freely look at the stimuli while their gaze traces are measured. Yet several impediments hamper a combination of fMRI and eye-tracking. In our recent work on neural Theory of Mind processes in alexithymia, we propose a simple way of integrating dwell time on specific stimulus features into general linear models of fMRI data. By parametrically modeling fixations, we were able to distinguish neural processes associated with specific stimulus features looked at. Here, I discuss opportunities and obstacles of this approach in more detail. My goal is to motivate a wider use of parametric models — usually implemented in common fMRI software packages — to combine fMRI and eye-tracking data.

KEYWORDS: fMRI, eye-tracking, Theory of Mind, Tom, mentalizing, alexithymia, emotional awareness, gaze, fixations, parametric modeling, free viewing, viewing behavior, gaze behavior

RECEIVED: November 12, 2021. **ACCEPTED:** November 22, 2021.

TYPE: Commentary

FUNDING: The author received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Kristin Marie Rusch, University of Marburg, Rudolf-Bultmann-Str. 8, 35039 Marburg, Germany. Email: Kristin.Rusch@med.uni-marburg.de

About 3 decades now, functional Magnetic Resonance Imaging (fMRI) enables the *in vivo* study of cognitive processes in the human brain. To investigate neural function with fMRI, researchers most commonly make use of visual stimuli. Especially in the social sciences, those stimuli are usually complex. Multiple features are displayed together, competing for visual attention. In order to nonetheless guarantee a controlled experimental environment, the participants' gaze is often guided by instruction or task. However, this way of restricting gaze is accompanied by an artificial pronunciation of specific stimulus features which would not necessarily represent the main locus of visual attention in a more naturalistic setting. Further, despite of being confronted with the same task, scan paths of participants may still differ. Finally, not even all experimental tasks permit restrictions of scan paths, in particular when studying higher cognitive functions. Thus, it would be preferable to allow participants to freely look at the experimental stimuli.

While in fMRI experiments natural viewing behavior can be measured by the use of an MR-compatible eye-tracking system, the subsequent analysis of fMRI data in combination with the information from the eye-tracker is not straightforward. One of the biggest challenges is the overcoming of differences in temporal resolution. Whereas major eye movements happen in milliseconds, fMRI data are usually sampled at a rate of less than one scan per second. In our study,¹ however, we propose a simple method to associate fMRI signal with gaze while participants freely explore the experimental stimuli.

The results from our study highlight several benefits of combining measurement and analysis of fMRI and eye-tracking data. First, interindividual differences in scan path can come to light. Second, unwanted effects of different gaze behaviors on neural signal are controlled for. And third, neural mechanisms associated with individual stimulus properties can be distinguished, instead of only assessing effects of the stimulus as a whole (Figure 1). So how can the joint analysis of fMRI and eye-tracking data be implemented? I will exemplify this from a more practical point of view by outlining the approach from our study.

Our goal was to examine differences in neural Theory of Mind (ToM) processing according to emotional awareness (ie, alexithymia). To measure ToM, we used a widely applied complex visual task (ie, the *Reading the Mind in the Eyes Test*, RMET²). In the RMET, participants choose a word out of 4 that best describes the mental state shown in a photograph of eyes. Since the stimulus items are spatially distributed in the display, it is mandatory for task completion that participants examine the stimuli in an unconstrained manner. Being aware of the potential bias in fMRI signal through interindividual gaze differences,³ we collected eye-tracking data to include them in the analysis. More specifically, in addition to modeling the average effect per stimulus, we integrated cumulated dwell time on predefined stimulus features (dwell time on eyes and on words) as parametric modulators into general linear models. The parametric modulators modeled the signal amplitude as a



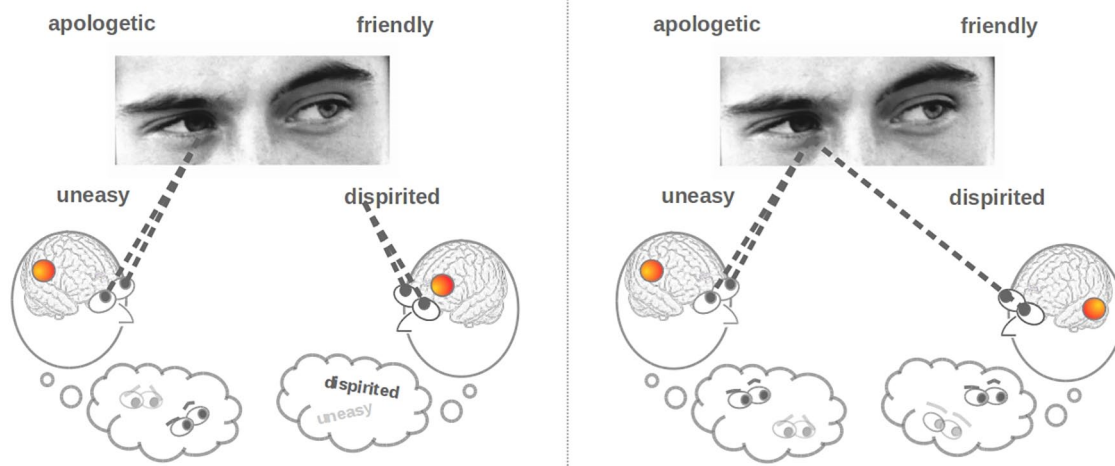


Figure 1. *Left:* Two participants perform a widely applied ToM task, the Reading the Mind in the Eyes Test (RMET).² While freely exploring the stimulus, their viewing preferences differ. Whereas the left one mainly operates on the eyes, the right one focusses more on words to complete the task. Different brain areas are activated (orange circles). Importantly, without controlling for the effects of gaze, we cannot discriminate between neural ToM processes and those neural processes related to different viewing behaviors. *Right:* Regardless of the viewing preferences, in order to complete the RMET both participants need to look at the eyes and the words. If we compare neural activations observed while the participants look at the same stimulus feature, the effects of gaze are controlled for. Here, activational pattern of the participants still differ, meaning that they operate on the same task aspect in a different way.

function of the sum of fixation durations per stimulus. Thus, by cumulating fixations over the presentation time of a stimulus, we were able to bypass the problem of incompatibilities in temporal resolution between fMRI and eye-tracking. For each parametric modulator, this approach results in a topographical map of brain regions in which the neural signal varies with visual attention to specific stimulus features. By orthogonalizing the parametric modulators, we assured their independence from both the regressor modeling the average stimulus effect as well as all associated parametric modulators previously entered into the model.⁴

As we collected eye-tracking data while participants freely explored the RMET stimuli, we were able to detect that participants with different levels of alexithymia show different viewing patterns. In particular, high alexithymics looked less at eyes. In order to control for potential effects of the revealed gaze differences on neural signal, we coded dwell time on eyes and on words as parametric modulators. Hence, we assured that interindividual differences in neural signal were not simply the result of different viewing behaviors. Further, we were able to dissociate the neural processes linked with attention to eyes expressing mental states from those associated with gaze at words describing mental states. That is, we modeled modality-specific ToM mechanisms. By comparing the topographical maps resulting from the parametric modulation analyses, we found that during mentalizing different regions belonging to the core ToM network process different modalities. While looking at ToM-related words was associated with signal variations in the medial prefrontal cortex (MPFC), alexithymia-dependent effects of dwell time on eyes revealed in the anterior temporo-parietal junction (TPJa). Based on these results, we conclude that the MPFC contributes to language-based, explicit ToM processing. In contrast, only in high alexithymics the TPJa seems to be involved in mentalizing. In light of former

studies demonstrating that high alexithymics have difficulties with mentalizing and the results from our study showing that high alexithymics look less at eyes (while completing the ToM task equally well as low alexithymics), we propose that in high alexithymics the involvement of the TPJa during mentalizing may deteriorate ToM performance which in turn may prompt implicit eye avoidance.

Thus, by only modeling the average effect as usually done in standard fMRI analysis, we would not have been able to draw conclusions as detailed as outlined here. In the following, I will delineate the rationale behind our approach in more detail.

Why Gaze Matters

In addition to temporal problems just addressed, several reasons may account for the few approaches that have been made toward a simultaneous measurement and analysis of fMRI and eye-tracking data so far. For example, it has been suggested for a long time that gaze is externally guided by task and the salience of stimulus features.⁵ Thus, researchers using task-based fMRI may not have been motivated to scrutinize interindividual differences in viewing behavior. This is also reflected in the widespread use of complex visual paradigms for the study of social cognition, often without caring about the multiple options for participants of where to look at. But even for the investigation of individuals with known atypical gaze behavior, as for example people with depression, autism spectrum disorder, or social anxiety, potential effects of gaze on neural signals are often completely ignored. However, recent research advises us to be more cautious about the deployment of visual attention and its impact on neural systems. Accumulating findings point toward a significant influence of the individual's characteristics on visual scene exploration. Gaze traces differ according to skills, traits and prior experiences.^{1,6} Preferences for particular scan paths on social stimuli can even partly be

explained by heritable factors.⁷ After all, viewing behavior seems to be determined by an interaction of genetic and environmental factors. Thus, task and stimulus properties undoubtedly shape attentional traces but not as independent from systematic differences due to characteristics of the individual as long suggested.

How do interindividual differences in viewing behavior influence neural processing? We must acknowledge that we know very little about the impact of systematic gaze differences on the signal in the social brain. To my knowledge, only a handful of studies dealt with this topic (eg, Gamer and Buchel⁸). Only recently, using a simple and clever study design, Hadjikhani et al³ demonstrated an influence of different attentional foci on the fMRI signal. They showed that guiding gaze toward the eye region compared with freely viewing faces was associated with different neural processes not only in early sensory regions but also in higher level brain areas like the prefrontal cortex. Thus, the effects of different scan paths clearly extend beyond sensory areas. Most importantly, by comparing the average signal during free viewing with the average signal during viewing of the same faces but with a fixation cross directing gaze toward the eyes, the authors demonstrated differences in neural patterns associated with 2 frequent manipulations used in fMRI studies.

Ultimately we do not really know how to interpret the average neural signal to a complex stimulus. Neural processes associated with the most frequently attended stimulus aspect may be detected. Signals related to meta-cognitive processes might prevail. However, signals related to competing visual foci might be lost in noise. An example from the literature on social cognition that illustrates these thoughts may be the study of the ToM concept.

Much work has been done to understand the neural mechanisms of ToM. But depending on the behavioral manipulation or experimental stimulus used to investigate ToM processes, the studies resulted in several heterogenic maps of brain regions proposed to be involved in ToM. Only few regions were activated regardless of the task, among them the MPFC and the TPJ. In their inspiring article, Schaafsma et al⁹ therefore called for a reformulation of the ToM concept. They proposed to break down the concept by analyzing processes related to single (task) components in order to subsequently reassemble a better validated and informed idea of ToM.

The combination of fMRI and eye-tracking as laid out here is an advance in exactly this direction. The simultaneous assessment of neural processes associated with distinguished task aspects is not limited to the study of ToM but can also be applied to other domains. Alternative approaches for a unified use of fMRI and eye-tracking were proposed.^{10,11} However, some obstacles need to be overcome in any case.

Challenges and Future Directions

In our study, we applied the proposed method of parametrically modulating gaze to a task in which the stimuli are presented

for a multiple of the time of repetition (TR). In fMRI, this is typically the case for epoch models. Neural processes associated with a single saccade or fixation can be modeled as an event, but due to the temporal resolution of standard fMRI it is difficult to distinguish between 2 succeeding events close in time. However, advances toward faster fMRI sequences for the combination of fMRI and eye-tracking have been made.¹² Nevertheless, the conditions for measuring eye traces in the MR scanner are suboptimal. Complicating factors are, for example, the distance between the eye-tracker and the participant, the lighting conditions, or the scanner coil blocking the view, especially for lateral gaze. Drooping eyelashes or eyelids further aggravate continuous measurement. Spatial imprecisions due to drift and the addressed suboptimal measuring conditions should be already taken into account when designing the experimental stimuli. For example, to be distinguished stimulus features should be presented with sufficient spatial distance.

Unfortunately, all this usually leads to the exclusion of a large proportion of participants. Hence, in order to improve measuring conditions, future studies need to determine those obstacles in a quantitative manner, as, for example, in Peitek et al¹³. Interesting approaches predicting eye movement patterns from fMRI scans – ie completely independent from an eye-tracker – were recently introduced.^{14,15} Although the proposed parametric modulation of gaze is straightforward, I suggest proof-of-principle studies, on the one hand, to indicate whether the models capture the targeted processes and, on the other hand, to examine the strength and reliability of revealed effects. A validation study may, for example, examine whether visual attention to a face embedded in a complex stimulus environment engages core face regions.

In conclusion, the proposed way of combining fMRI and eye-tracking data is an easily implemented and promising method for gaining more sophisticated insights into the neural mechanisms of social cognition. Ignoring likely bias through gaze instead may conceal valuable information.

Acknowledgement

I thank Andreas Jansen and Hannes Rusch for valuable comments.

Author Contributions

KMR wrote the article.

ORCID iD

Kristin Marie Rusch  <https://orcid.org/0000-0002-9492-0259>

REFERENCES

1. Zimmermann KM, Schmidt KD, Gronow F, Sommer J, Leweke F, Jansen A. Seeing things differently: gaze shapes neural signal during mentalizing according to emotional awareness. *Neuroimage*. 2021;238:118223.

2. Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I. The “reading the mind in the eyes” test revised version: a study with normal adults, and adults with asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry*. 2001;42:241-251.
3. Hadjikhani N, Zurcher NR, Lassalle A, Hippolyte L, Ward N, Johnels JÅ. The effect of constraining eye-contact during dynamic emotional face perception: an fMRI study. *Soc Cogn Affect Neurosci*. 2017;12:1197-1207.
4. Mumford JA, Poline J-B, Poldrack RA. Orthogonalization of regressors in fMRI models. *PLoS One*. 2015;10:e0126255.
5. Borji A, Sihite DN, Itti L. Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study. *IEEE Trans Image Process*. 2013;22:55-69.
6. de Haas B, Iakovidis AL, Schwarzkopf DS, Gegenfurtner KR. Individual differences in visual salience vary along semantic dimensions. *Proc Natl Acad Sci U S A*. 2019;116:11687-11692.
7. Constantino JN, Kennon-McGill S, Weichselbaum C, et al. Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature*. 2017;547:340-344.
8. Gamer M, Buchel C. Amygdala activation predicts gaze toward fearful eyes. *J Neurosci*. 2009;29:9123-9126.
9. Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R. Deconstructing and reconstructing theory of mind. *Trends Cogn Sci*. 2015;19:65-72.
10. Marsman JBC, Renken R, Velichkovsky BM, Hooymans JMM, Cornelissen FW. Fixation based event-related fmri analysis: using eye fixations as events in functional magnetic resonance imaging to reveal cortical processing during the free exploration of visual images. *Hum Brain Mapp*. 2012;33:307-318.
11. Henderson JM, Choi W. Neural correlates of fixation duration during real-world scene viewing: evidence from fixation-related (FIRE) fMRI. *J Cogn Neurosci*. 2015;27:1137-1145.
12. Korosteleva A, Ushakov V, Malakhov D, Velichkovsky B. Event-related fMRI analysis based on the eye tracking and the use of ultrafast sequences. Paper presented at: Advances in Intelligent Systems and Computing; August 1-3, 2018; Moscow, Russia; Springer:107-112.
13. Peitek N, Siegmund J, Parnin C, Apel S, Brechmann A. Toward conjoint analysis of simultaneous eye-tracking and fMRI data for program-comprehension studies. Paper presented at: Proceedings of the Workshop on Eye Movements in Programming—EMIP '18; June 14-17, 2018, Warsaw, Poland; ACM Press:1-5.
14. O'Connell TP, Chun MM. Predicting eye movement patterns from fMRI responses to natural scenes. *Nat Commun*. 2018;9:5159.
15. Son J, Ai L, Lim R, et al. Evaluating fMRI-Based estimation of eye gaze during naturalistic viewing. *Cereb Cortex*. 2020;30:1171-1184.