

# **Eyes on Emotion: Dynamic Gaze Allocation During Emotion Perception from Speech-like Stimuli**

**Minke J. de Boer**<sup>1,2,3,\*</sup>, **Deniz Başkent**<sup>1,2</sup> and **Frans W. Cornelissen**<sup>1,3</sup>

<sup>1</sup> Research School of Behavioural and Cognitive Neurosciences (BCN), University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

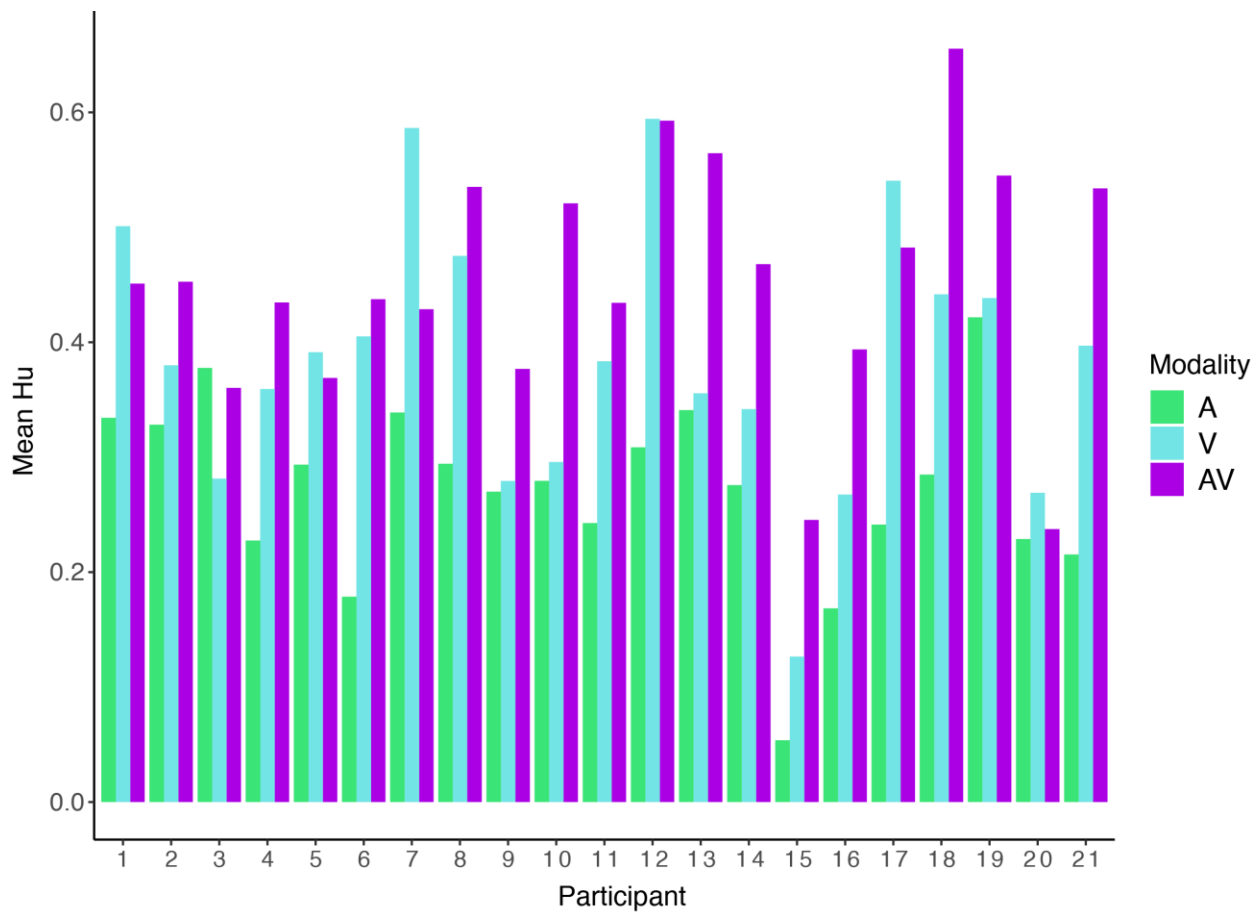
<sup>2</sup> Department of Otorhinolaryngology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

<sup>3</sup> Laboratory for Experimental Ophthalmology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

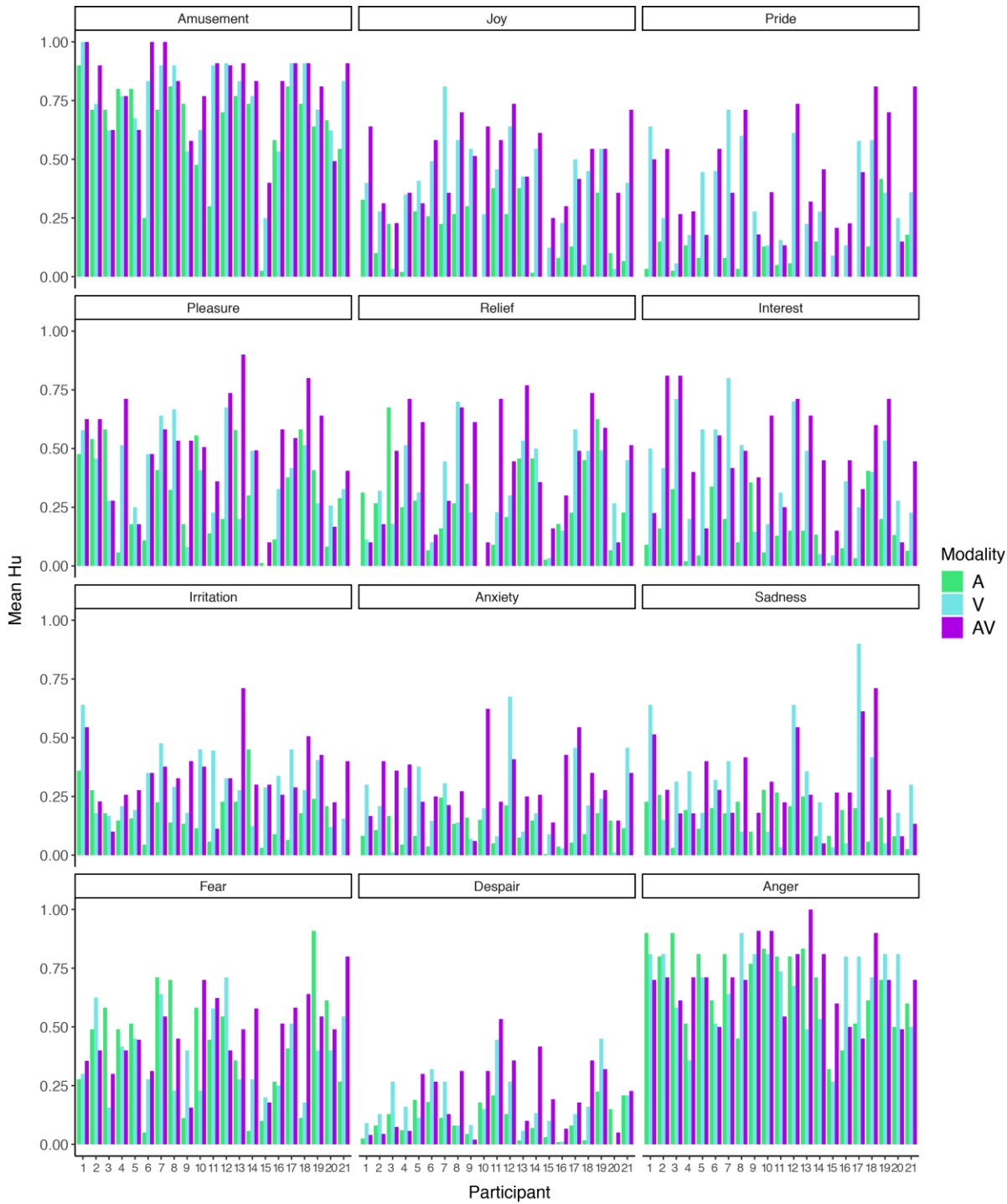
Received 20 November 2019; accepted 29 may 2020

\* To whom correspondence should be addressed. E-mail: minke.de.boer@rug.nl

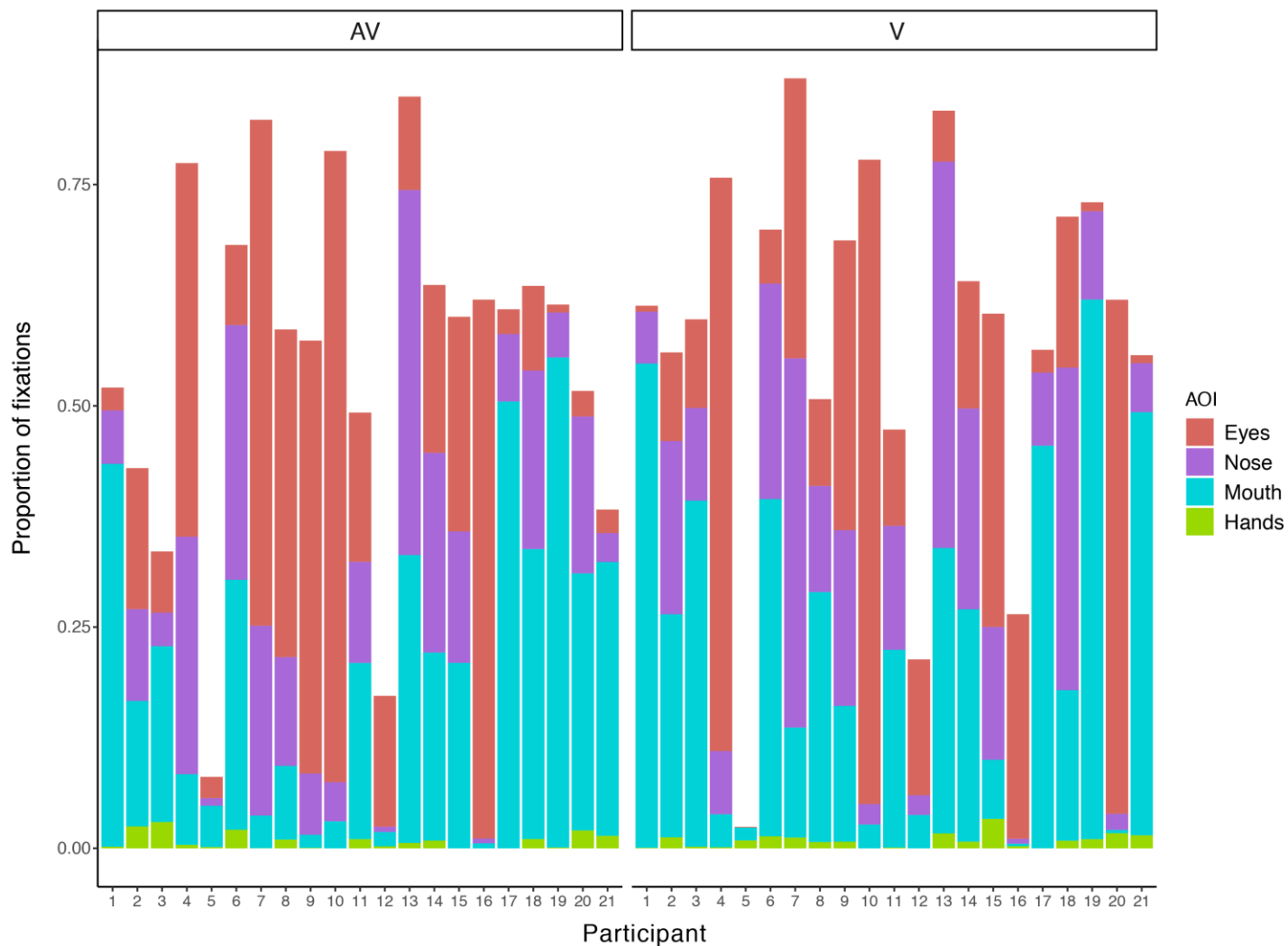
## **Supplementary Material**



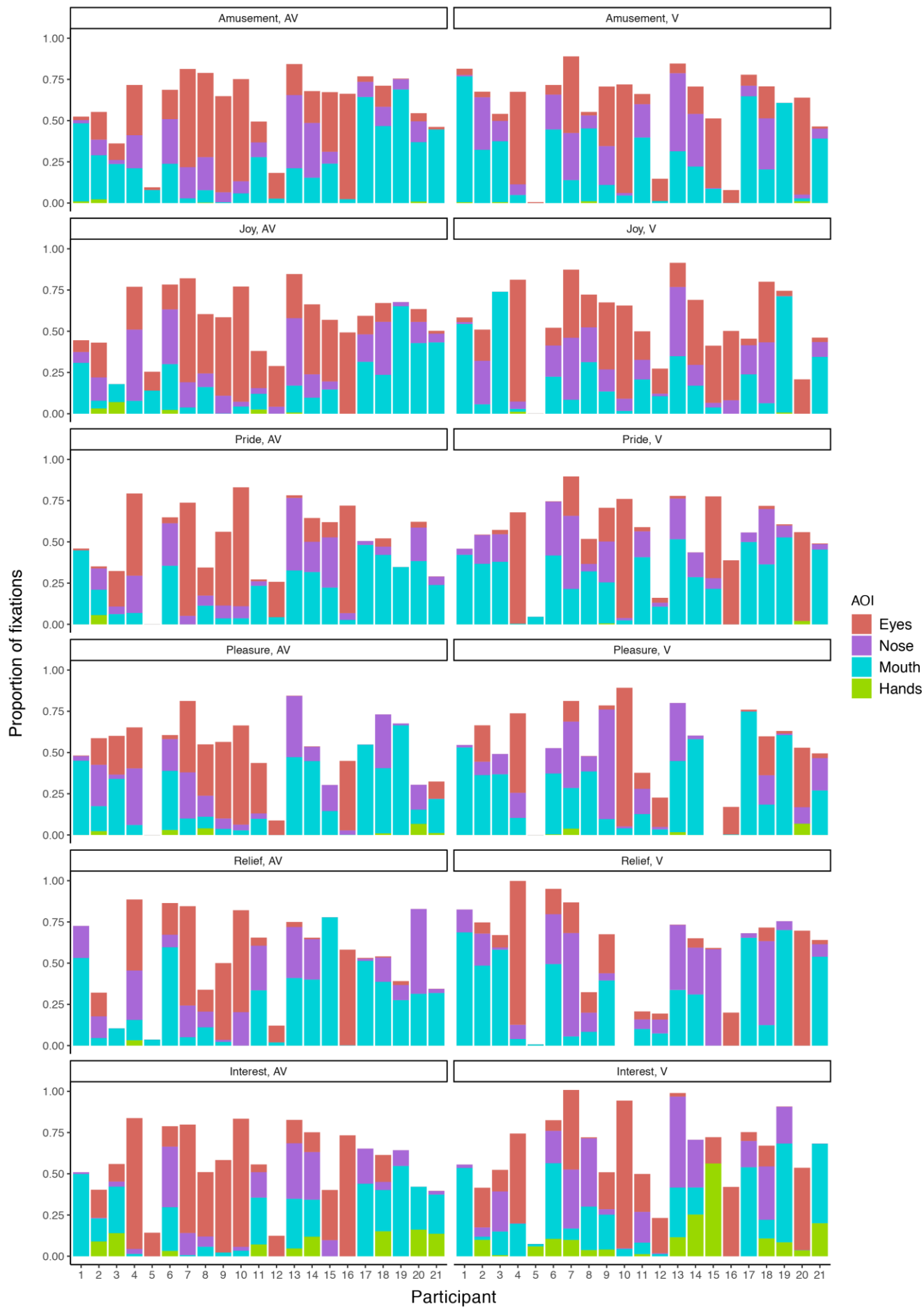
**Figure S1.** Task performance for each modality and participant, shown in unbiased hit rates, averaged over trials. Each bar shows an individual’s mean accuracy for a specific modality. Auditory-only performance modality is indicated by green bars, visual-only performance is indicated by blue bars, and AV performance is indicated by purple bars.

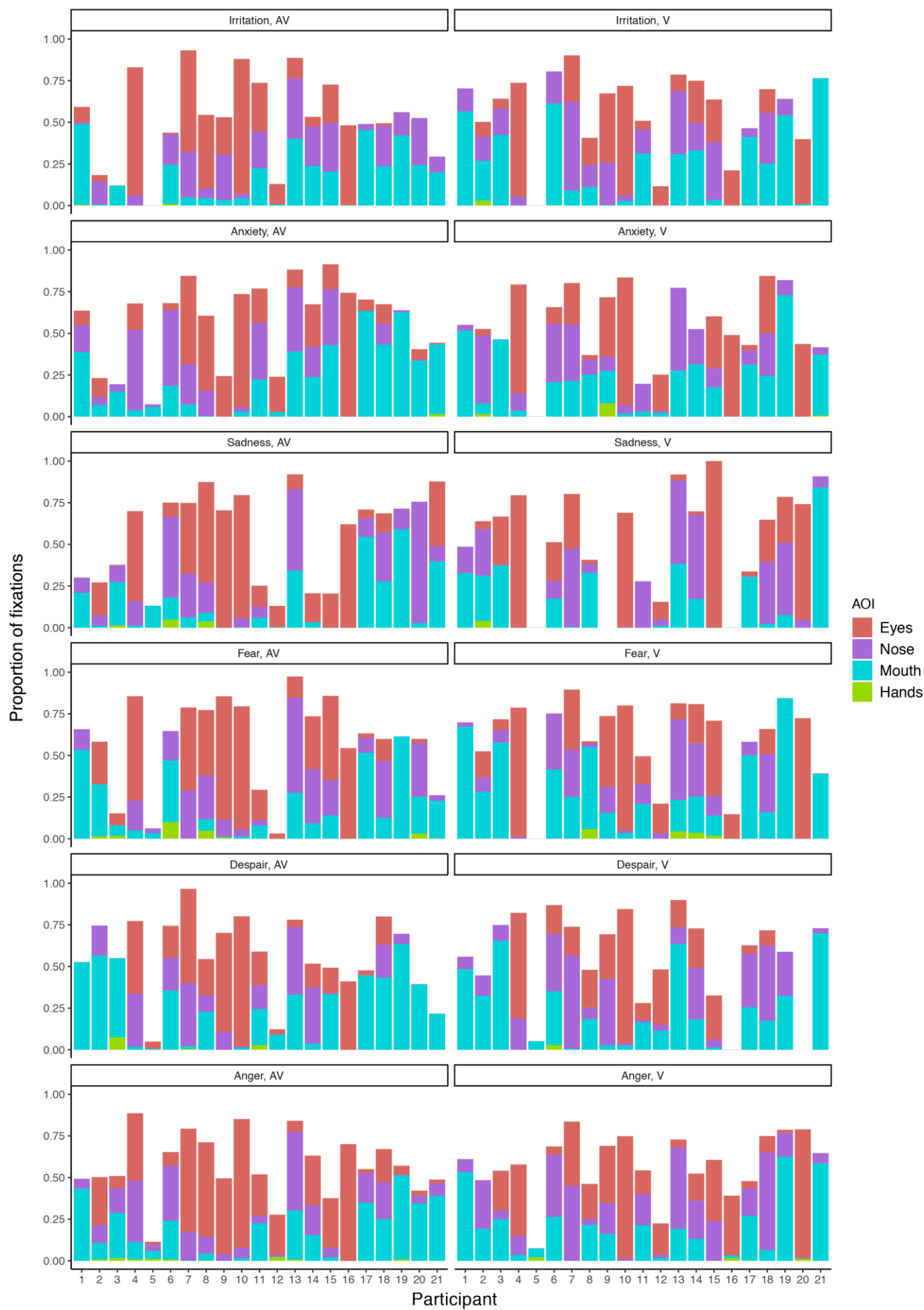


**Figure S2.** Task performance for each modality, emotion, and participant, shown in accuracy percent correct scores, averaged over trials. Each panel shows the result for one of the emotions. Each bar shows an individual's mean accuracy for a specific modality. Auditory-only performance modality is indicated by green bars, visual-only performance is indicated by blue bars, and AV performance is indicated by purple bars.

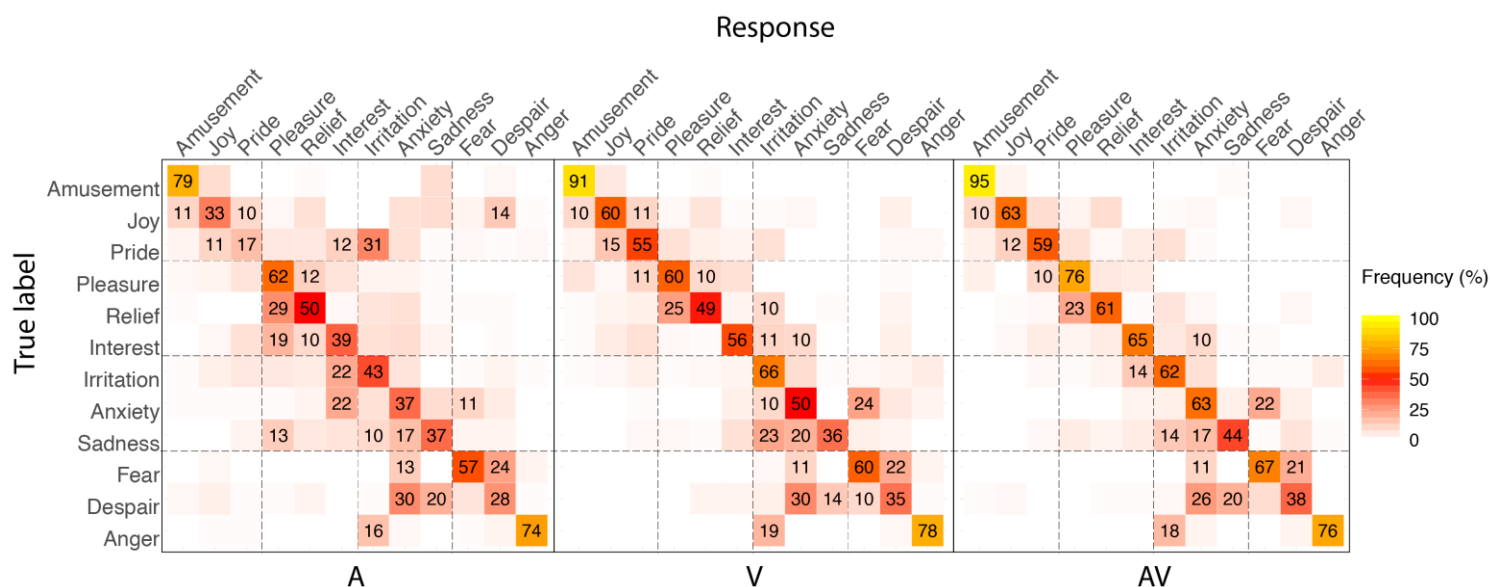


**Figure S3.** Average fixation proportions for each modality and participant, averaged over trials. Only fixation proportions for correct trials are shown. Each panel shows the result for one of the modalities. Each bar shows an individual's mean fixation proportions on the different areas of interest (AOIs). Fixations on the eyes are shown in red, fixations on the nose are shown in purple, fixations on the mouth are shown in blue, and fixations on the hands are shown in green.





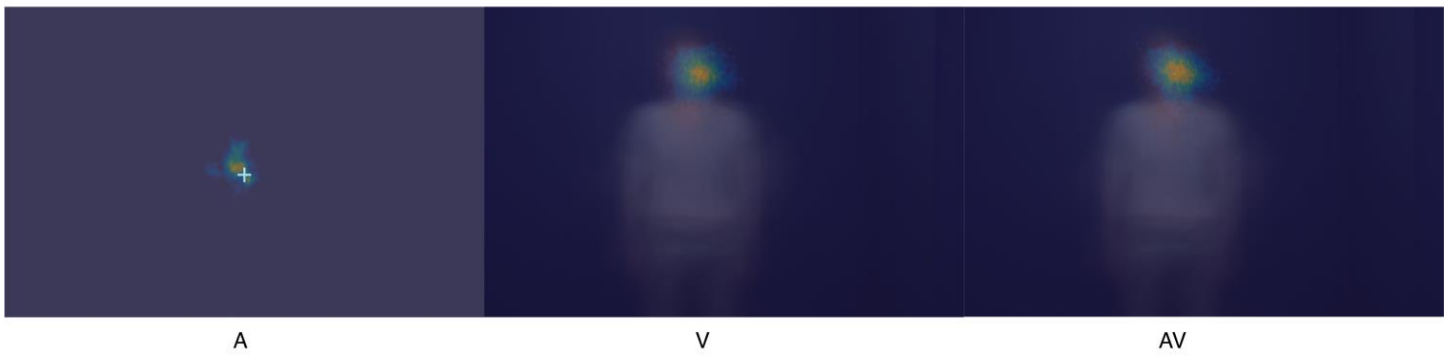
**Figure S4.** Average fixation proportions for each modality, emotion, and participant, averaged over trials. Only fixation proportions for correct trials are shown. Each panel shows the result for one of the modality–emotion combinations. Each bar shows an individual’s mean fixation proportions on the different areas of interest (AOIs). Fixations on the eyes are shown in red, fixations on the nose are shown in purple, fixations on the mouth are shown in blue, and fixations on the hands are shown in green.



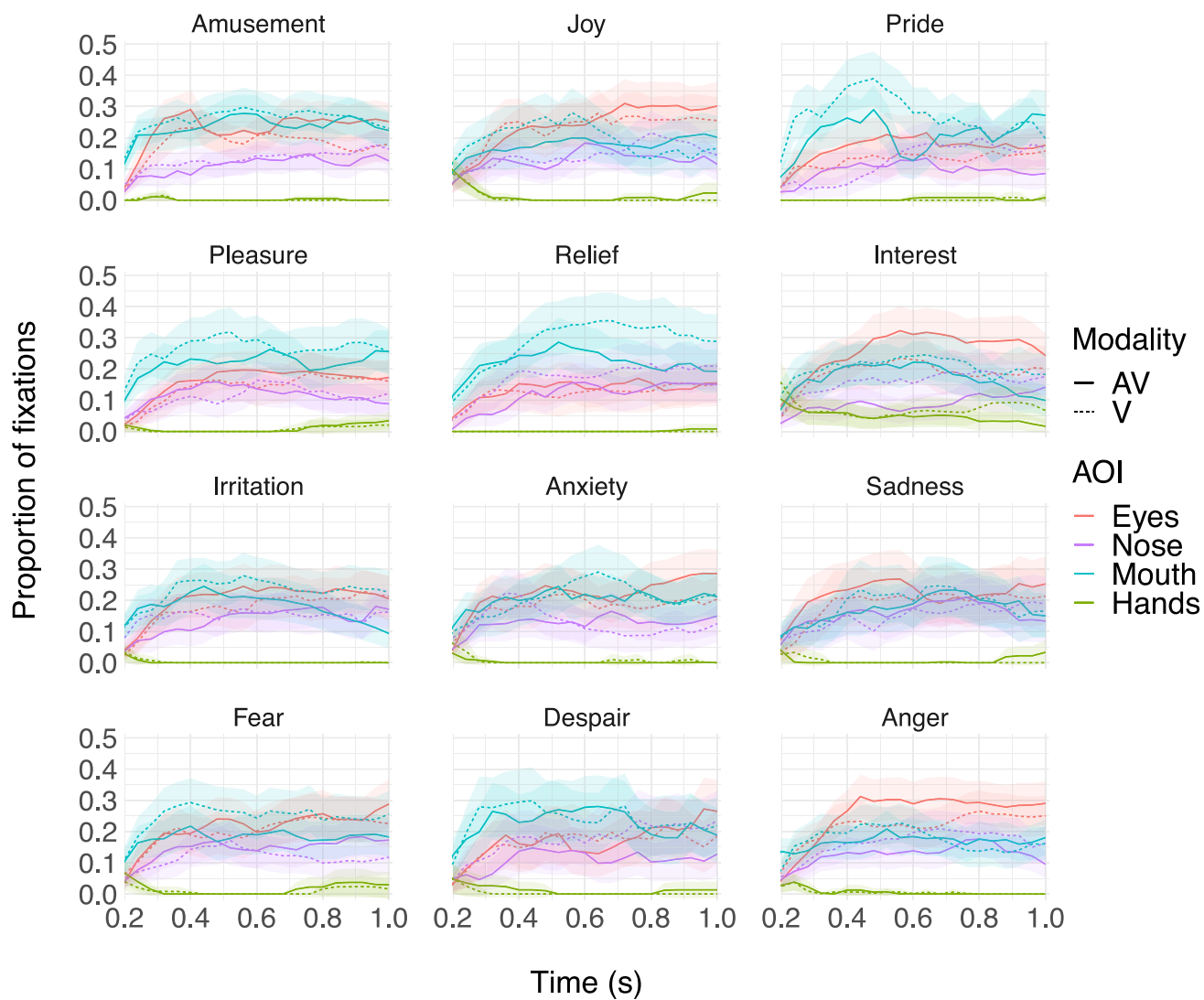
**Figure S5.** Confusion matrices for all modalities. Each row represents the true label for an emotion (i.e., the correct answer), while each column represents the given response. Values on the diagonal indicate correct responses, while values off the diagonal indicate incorrect responses. Values are frequency percentages. Only values > 10% are given a value label, below 10% only the color is indicative of the frequency.



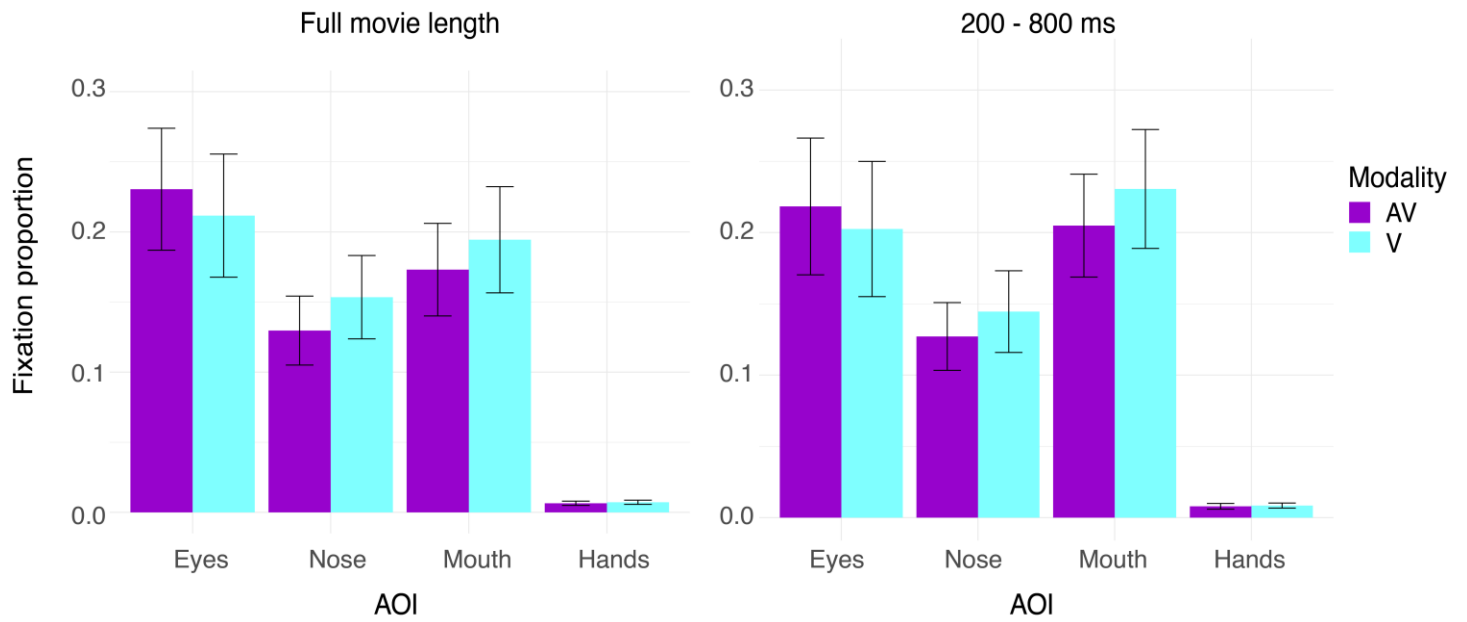
Heatmaps projected on an average of presented stimuli



**Figure S6.** Fixation heat maps for all modalities, averaged over participants and trials. Colors indicate how often a certain area was fixated (yellow: often, dark blue: hardly ever). The heat maps are projected on top of an image averaging the presented stimuli in each modality (i.e., audiovisual and visual-only: an average of all movies, audio-only: a fixation cross).



**Figure S7.** Fixation patterns for correct trials on all areas of interest (AOIs) across the analyzed time course, averaged over participants. The panels show fixation proportions for different emotions. Shaded areas around each line denote the standard error of the mean (SEM).



**Figure S8.** Fixation proportions for correct trials on all areas of interest (AOIs: eyes, nose, mouth, and hands) averaged over: the full movie length (left) and the analyzed time window of 200–800 ms (right), both averaged over stimuli and participants. Error bars denote the standard error of the mean (SEM).