Eye Contact Is A Marker of Alignment in Conversation

SOPHIE WOHLTJEN AND THALIA WHEATLEY, Dartmouth College

INTRODUCTION

When we interact with one another, we initiate a cascade of social cues and decisions that are modulated by the words, facial expressions, body language, and actions of our interaction partner. These adaptations that we make during interaction serve to align us, create a shared understanding, with the result (ideally) of successful communication (Hasson & Frith, 2016). One of the most ubiquitous signals of successful communication is eye contact. Couples make more of it than strangers (Rubin, 1970), and it is notably diminished in autism, a disorder characterized by social deficits (Chita-Tegmark, 2016). A wealth of research exists on the strategic ways we use eye contact to communicate (Argyle & Dean, 1965; Kendon, 1967; Rutter et al., 1978), but it remains unclear whether eye contact might be a useful marker of the mutual adaptation and alignment that occurs during interaction. In the present paper I use eye tracking to measure two individuals' pupil dilations, as a continuous index of their attention, while they are engaged in natural conversation. Our data suggest that eye contact is a signal of shared attention between two individuals as they engage in successful communication.

METHOD

This study is based on previous work showing that when people pay attention to one another, their pupil dilations, indexing that attention, synchronize (Kang & Wheatley, 2017). In the current study, a set of two participants wear eye-tracking glasses and are seated across from one another. They are instructed to have a conversation, about anything they would like, for ten minutes. Afterward, they watch a video of their conversation while continuously rating how engaged they were, and report their overall feelings about the conversation in a questionnaire.

Participants

94 subjects comprising 47 dyads (mean age: 20.64; 51 females) participated in the study.

Eye-Tracking Data Collection

Pupil dilation data was collected continuously while dyads engaged in conversation. Pupil diameter was recorded from both eyes at 60Hz using either SensoMotoric Instruments (SMI) wearable eye-tracking glasses (SensoMotoric Instruments, Teltow, Berlin) or Pupil Labs wearable eye-tracking glasses (Pupil Labs, Berlin, Germany). Dyads were seated across from one another at a distance of approximately 3 feet in a luminance-controlled testing room. Lux values were recorded at multiple points in the room to ensure that luminance in the room did not exceed the 150 lux necessary to elicit a luminance-induced pupil dilation (Maqsood, 2017).

Ratings of Engagement

After conversation, participants watched a video recording of their conversation and continuously reported how engaged they were at each moment by moving a slider bar (PsychoPy software, Peirce, 2007). Participants' continuous engagement ratings were recorded at 10hz.

Eye Contact Annotations

We annotated moments of eye contact during the conversation using ELAN annotation software (Brugman et al., 2004). Annotations were obtained by watching videos recorded from a camera on the eye-tracking glasses located near the participant's nasion. These videos depicted each participant from the viewpoint of their partner's own eyes, so moments of eye contact were defined as moments when the participant in the video looked straight at the camera. Two independent annotations were obtained for each participant, and moments in which the two independent annotations agreed that eye contact was being made were moments that were accepted into the final set of eye contact annotations. Each participant's annotations were then compared with the eye contact annotations of their conversation partner to find moments of mutual gaze.

Preprocessing and Quality Control
Raw pupil dilation time series were preprocessed for further analysis by first performing linear interpolation over eye-blinks and other dropout in signal, removing subjects who required more than 25% of their data interpolated. To remove spikes, the data was then median filtered (5th order) and low-pass filtered at 10 hz and detrended to reduce drift.

Dynamic Time Warping
Dynamic Time Warping (DTW) is an algorithm used to compare signals that may be offset in time (Berndt & Clifford, 1994). We used DTW to measure synchrony between 1) dyads' pupils during conversation and 2) dyads' continuous ratings of engagement recorded after the conversation. The DTW algorithm divides two signals into a user-defined number of segments, each representing some window of time. Then, DTW calculates the cosine similarity of these chunks. DTW makes three comparisons for each chunk: the same chunk on both signals, and that chunk on one time series with the following chunk on the other. The cosine similarity values associated with these three segment pairs are compared. The segment pair that yields the smallest cosine similarity value is deemed to be the time at which the two signals best align, and both signals are adjusted according to that cosine similarity comparison. Each adjustment incurs a penalty, or a "cost" of realignment. The sum of these penalties yields an overall cost value, which represents the overall effort involved in warping one signal onto the other. Higher cost values will indicate greater dissimilarity between two patterns.

RESULT

A linear mixed effects analysis of the relationship between the number of instances of eye contact, the sum of time spent making eye contact, dyads' engagement similarity (computed using DTW) and dyads' pupillary synchrony was performed in R using the lme4 package (Bates et al., 2014). For this analysis, conversations were separated into two-minute bins in order to get a wide sample of values for how subjects chose to make eye contact. Number of instances of eye contact, sum of time spent making eye contact, and engagement similarity were entered into the model as fixed effects. We included random intercepts for dyads, bins, and the type of eye tracker that was used. All variables computed using DTW were log-transformed to correct for homoscedasticity.

There was a significant main effect of reported engagement similarity on pupillary synchrony ($t(419.2) = 2.952$, $p=0.003$), suggesting that when subjects were similarly engaged, their pupils were more synchronous. This provided initial evidence that pupillary synchrony was indexing true alignment of attention between conversation partners. However, if both participants in a dyad reported being *disengaged* at the same time, this would still produce a high engagement similarity score. To test whether the synchrony of dyads' engagement ratings were primarily driven by mutual disengagement, we ran another linear mixed model predicting subjects engagement similarity with their mean engagement ratings. If moments of engagement similarity were primarily driven by moments of mutual disengagement, we would expect to see these two values negatively related, such that a lower mean engagement resulted in more engagement similarity. We found that the opposite was true ($t(163.54) = 4.163$, $p < 0.001$), with higher mean engagement predicting higher engagement similarity. This suggests that our original effect was primarily the result of subjects being mutually engaged and thus more synchronous with one another.

There was also a significant main effect of the number of instances of eye contact on pupillary synchrony ($t(447.2) = 2.686$, $p=0.007$), indicating that the more times subjects made eye contact, the more synchronous they were. However, there was no effect of the sum of time spent making eye contact on pupillary synchrony, suggesting that it is the *frequency of making* eye contact, rather than the total time spent making it, that is most indicative of shared attention.

Reverse Correlation
Reverse correlation has been used in neuroimaging studies to reveal what properties of a stimulus yield common fluctuations in hemodynamic response across individuals (Hasson et al., 2004; Spiers & Maguire, 2007). We adapted these methods to the current design to create time series of pupillary synchrony between dyads and reverse correlate moments of pupillary synchrony with eye contact. We predicted that times of greatest pupillary synchrony would correspond with moments when dyads were making proportionally more eye contact. Pupillary synchrony was calculated by parsing each pupillary time series into three-second epochs and comparing those epochs within each dyad using DTW. Eye contact was calculated by parsing a binary time series of eye contact, comprised of samples where eye contact either was or was not being made, into three-second epochs and taking the average of the eye contact in these epochs, resulting in a proportion of eye contact made per three-second window.

We performed a correlation between eye contact time series and pupillary synchrony time series for every dyad, resulting in a distribution of 47 correlations. This distribution was significantly positively skewed away from 0 (Mean $R= 0.044$, $t(47) = 2.62$, $p=0.01$), suggesting that subjects were more synchronous during moments when a larger proportion of eye contact was being made. To more robustly test the probability that this distribution showed a significant positive relationship between eye contact and synchrony, we compared our true mean to a null distribution created by comparing pseudo-pairs of eye contact and pupillary synchrony (i.e. the eye contact time series from one dyad with the synchrony time series from a different dyad). The true mean correlation fell entirely outside this null distribution, providing further evidence that the relationship we saw was not due to chance.

Multi-level Vector Auto-regression
To investigate the directionality of the relationship between eye contact, pupil size, and pupillary synchrony, we used the R package mlVAR (Epskamp et al., 2017), which is used for the analysis of multivariate time series using a network framework. This method allows for the estimation of temporal, contemporaneous, and between subjects networks of relationships between time series variables, where variables are "nodes" in the network, and the partial correlations between those variables are "edges". The temporal network is estimated by calculating the regression coefficients obtained by predicting one variable at time t from another variable at time $t-1$. Once the temporal network has been estimated, the contemporaneous network is calculated by taking partial correlations of the residuals from the temporal network. Finally, the between subjects network is calculated by taking partial correlations of mean-centered predictors and adding person-means as level 2 predictors.

We obtained temporal, contemporaneous, and between-subjects networks for the relationship between pupil size, pupillary synchrony, and eye contact. We found a significant between-subjects relationship for eye contact and pupillary synchrony (partial $r = 0.23, p<0.05$), suggesting that dyads who made significantly more eye contact during their conversations also showed more pupillary synchrony. We also found contemporaneous network relationships between pupil size and eye contact (partial $r = 0.05, p<0.05$) and pupil size and synchrony (partial $r = 0.22$, $p<0.05$), suggesting that in moments when our pupils dilate, we both make more eye contact and are more synchronous with our conversation partner. Finally, we found temporal network relationships suggesting that both pupil size (partial $r = 0.07, p<0.05$) and synchrony (partial $r = 0.04$, $p<0.05$) in one moment lead to eye contact in the next. This temporal network result suggests that, **rather than being a strategy used to promote synchrony during conversation, eye contact may be a behavioral readout of the alignment that is already occurring.**

CONCLUSION

In this work, we have provided the first evidence that eye contact may serve as an indication of pupillary synchrony, and thus shared attention, in conversation. We first linked pupillary synchrony to a behavioral rating of engagement, establishing that synchrony is in fact a marker of two individuals aligning their levels of engagement with one another over the course of the conversation. We subsequently linked pupillary synchrony to eye contact, both in terms of the number of instances of eye contact and in the overall proportion of eye contact made over the course of the conversation. Further, we found evidence that this link may be temporally directed: rather than causing synchrony, eye contact appears to be a consequence of shared attention.

REFERENCES

Argyle, M., & Dean, J. (1965). EYE-CONTACT, DISTANCE AND AFFILIATION. *Sociometry*, *28*, 289–304.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. In *arXiv [stat.CO]*. arXiv. http://arxiv.org/abs/1406.5823

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *KDD Workshop*, *10*, 359–370.

Brugman, H., Russel, A., & Nijmegen, X. (2004). Annotating Multi-media/Multi-modal Resources with ELAN. *LREC*. https://www.researchgate.net/profile/Hennie_Brugman/publication/238473992_Annotating_Multimedia_Multi-modal_resources_with_ELAN/links/53fd90320cf22f21c2f80afd.pdf

Cheong, J. H., Brooks, S., & Chang, L. J. (2017). *FaceSync: Open source framework for recording facial expressions with head-mounted cameras*. https://psyarxiv.com/p5293/

Chita-Tegmark, M. (2016). Social attention in ASD: A review and meta-analysis of eye-tracking studies. In *Research in Developmental Disabilities* (Vol. 48, pp. 79–93). https://doi.org/10.1016/j.ridd.2015.10.011

Epskamp, S., Deserno, M. K., & Bringmann, L. F. (2017). mlVAR: multi-level vector autoregression. *R Package Version 0. 4*.

Hasson, U., & Frith, C. D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *371*(1693). https://doi.org/10.1098/rstb.2015.0366

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*(5664), 1634–1640.

Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology. General*, *146*(4), 569–576.

Kendon, A. (1967). Some functions of gaze-direction in social interaction. In *Acta Psychologica* (Vol. 26, pp. 22–63). https://doi.org/10.1016/0001-6918(67)90005-4

Maqsood, F. (2017). Effects of varying light conditions and refractive error on pupil size. *Cogent Medicine*, *4*(1). https://doi.org/10.1080/2331205x.2017.1338824

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13.

Rubin, Z. (1970). Measurement of romantic love. *Journal of Personality and Social Psychology*, *16*(2), 265–273.

Rutter, D. R., Stephenson, G. M., Ayling, K., & White, P. A. (1978). The timing of Looks in dyadic conversation. *The British Journal of Social and Clinical Psychology*, *17*(1), 17–21.

Spiers, H. J., & Maguire, E. A. (2007). Decoding human brain activity during real-world experiences. *Trends in Cognitive Sciences*, *11*(8), 356–365.