Supplementary Information

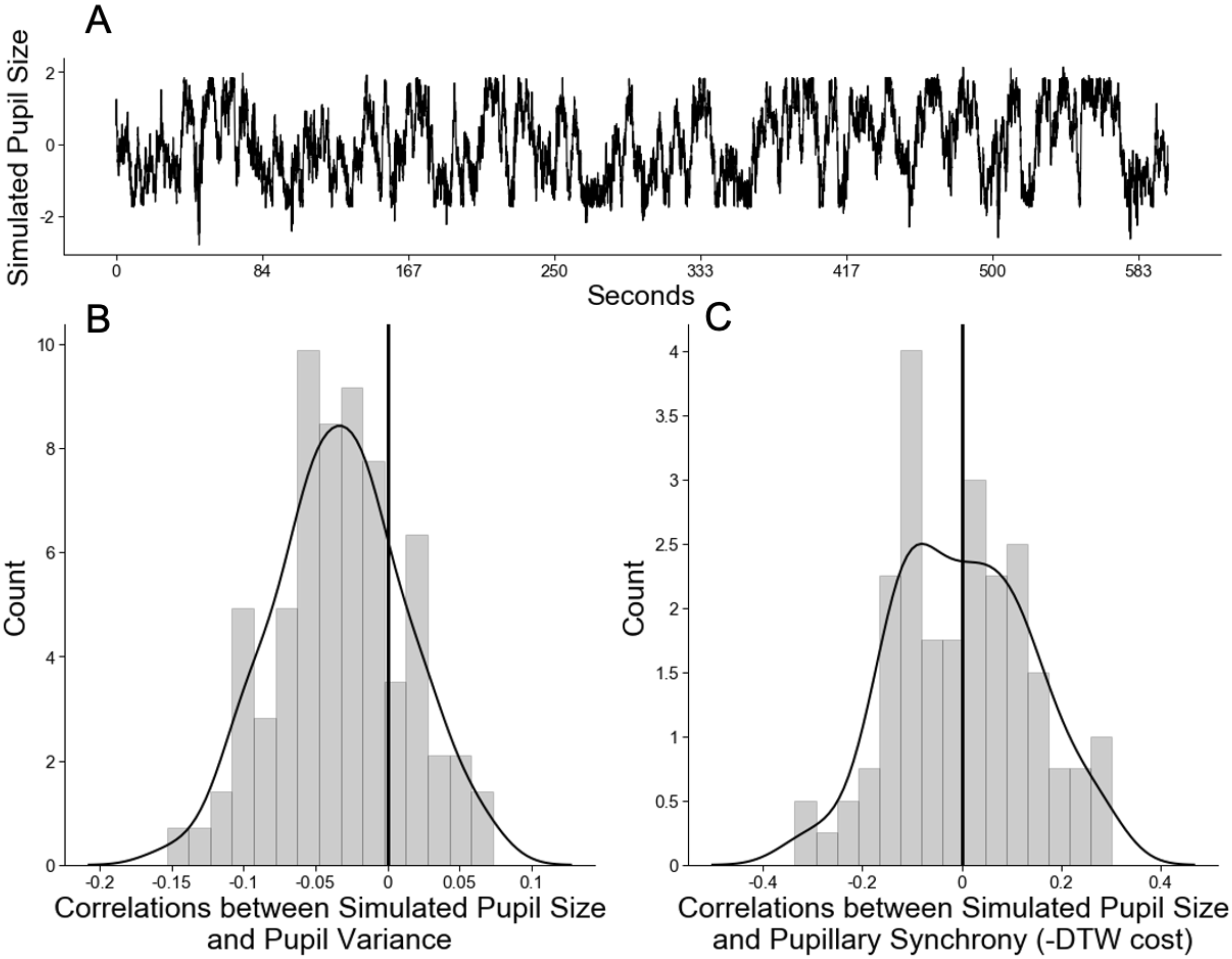
Eye contact marks the rise and fall of shared attention in conversation

Sophie Wohltjen & Thalia Wheatley

The relationship between pupil size and pupillary synchrony as a function of variance

The Multi-level Vector Auto-regression (mlVAR) analysis revealed a positive contemporaneous relationship with pupil size and pupillary synchrony, such that the pupil dilations of conversation partners tended to become more synchronous the larger their pupil size. This result could potentially be explained by less variance at larger pupil sizes artificially inflating synchrony scores. To investigate this, we divided the pupillary time series for each conversation partner into one-second windows. For each of these windows, we measured 1) mean pupil size and 2) amount of variance in the pupillary signal. We then correlated pupil size with pupil variance in these windows for each participant. A one-sample t-test revealed that the distribution of the resulting pearson's R values was significantly negatively skewed from zero (M R value = -0.26, SD = 0.22; t(93) = -11.29, p < 0.001: i.e., greater variance at smaller pupil sizes).

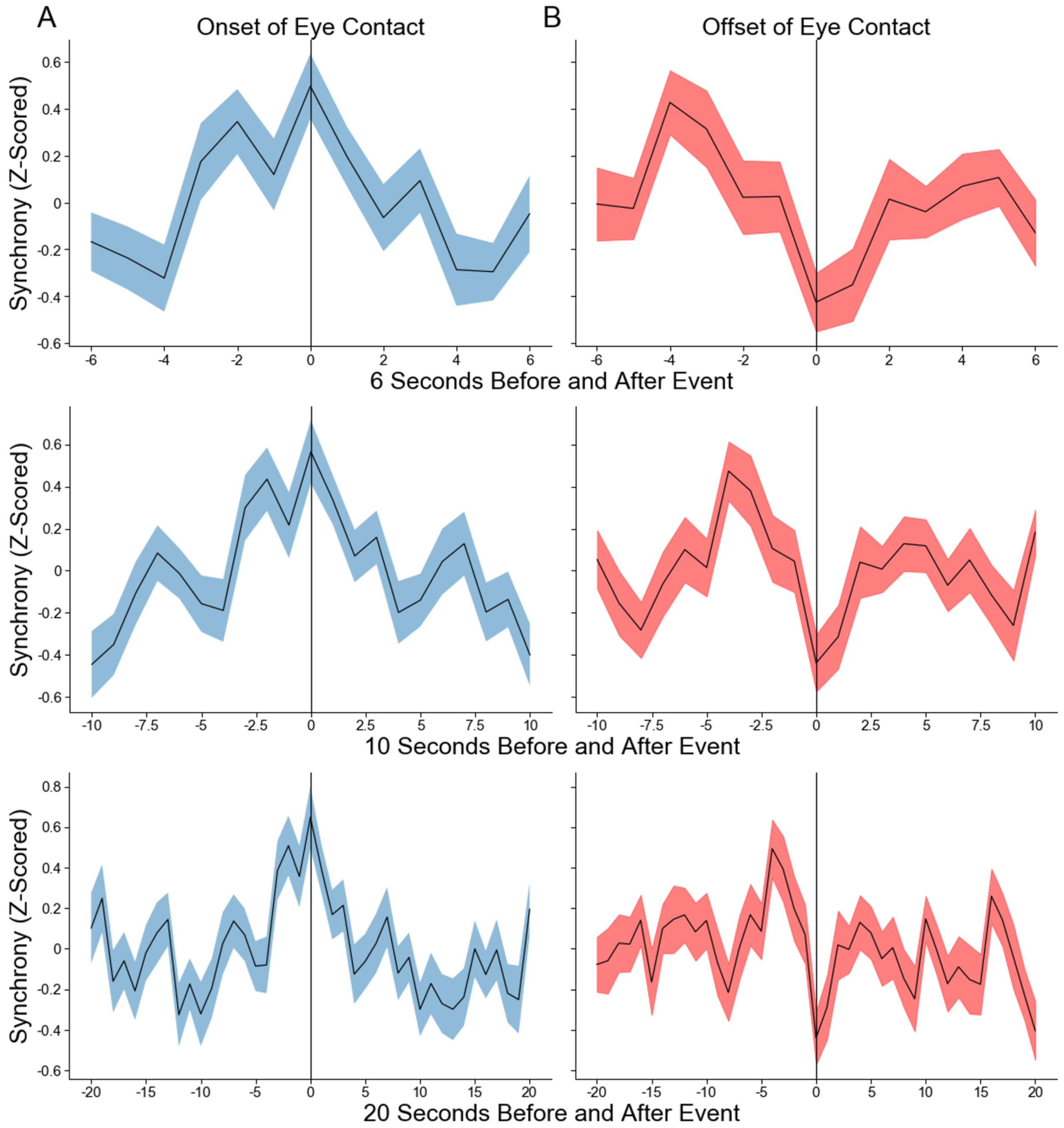
To test whether this difference in variance across pupil size accounted for the observed positive relationship between pupil size and pupillary synchrony, we created 94 simulated "pupillary" time series (equal to the 94 participants in our study) using a random walk model that sampled points within the same size range and standard deviation of our true pupillary data (M size = 0.005; SD = 0.15; size range = -3.75 – 1.86; see supplementary figure 1A). We then randomly added noise to these time series based on the probability in the real data of a given amount of variance (variance ranged from 0 to 2.47) at a given pupil size. With this addition of noise, we confirmed that the simulated data had the same significant negative relationship between pupil size and pupil variance as the real pupil data (M R value = -0.03, SD = 0.04; t(93) = -7.26, p < 0.001; supplementary figure 1B). We then tested whether this simulated data showed the same relationship between pupil size and inverse DTW cost (synchrony) as we observed with the real pupil data. There was no such relationship between pupil size and synchrony for the simulated data (M R value 0.002, SD=0.14; t(93) = 0.16, p = 0.87; supplementary figure 1C), suggesting that amount of variance alone cannot fully explain the relationship observed between pupil size and pupillary synchrony in our participants.



Supplementary Figure 1. **A)** Example of a simulated pupillary time series using a random walk model with added noise based on pupil size for a single subject; **B)** Distribution of pearson's R values obtained by correlating simulated pupil size with simulated pupil variance, with a black vertical line drawn at zero. This distribution was significantly negatively skewed from zero ($p < 0.001$). **C)** Distribution of pearson's R values obtained by correlating simulated pupil size with simulated pupillary synchrony. This distribution was not significantly skewed from zero ($p = 0.87$), suggesting that differences in variance do not solely explain the relationship between pupillary synchrony and pupil size in our true data.

Calculating the fluctuations of synchrony around eye contact over a range of timescales

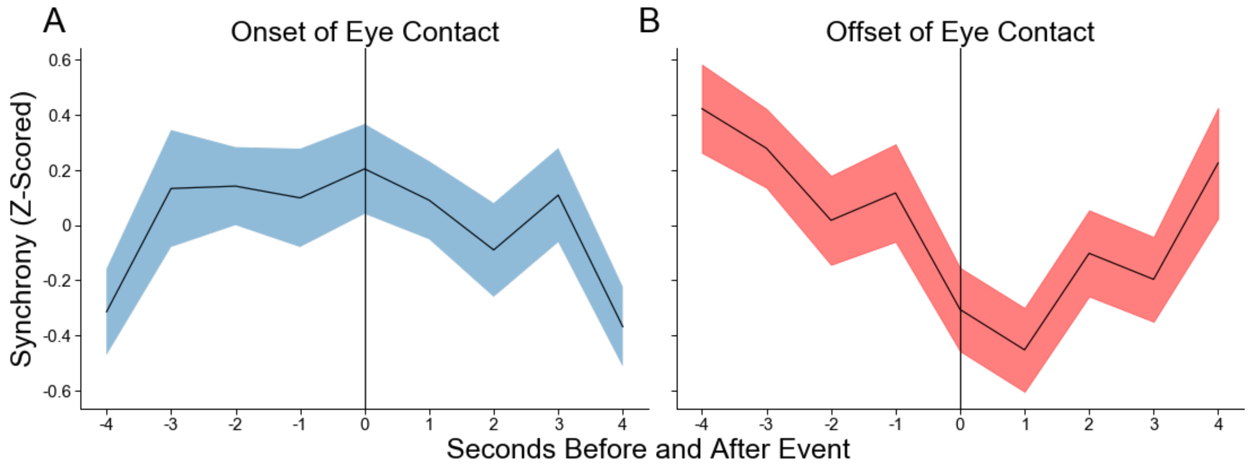
In the main text, we report the nine seconds surrounding the onsets and offsets of eye contact. This choice was motivated by first plotting the fluctuations of synchrony around eye contact at a range of time series lengths, and noting that regardless of length, a fluctuation of around nine seconds surrounding eye contact emerged. Supplementary figure 2 illustrates this effect, depicting synchrony time series surrounding eye contact at three different time series lengths.



Supplementary Figure 2. Synchrony curves surrounding the **A)** onset and **B)** offset of eye contact at a range of timescales. Regardless of the length of time used to compute each synchrony curve, we found that a nine-second parabolic fluctuation around the onset and offset of eye contact emerged.

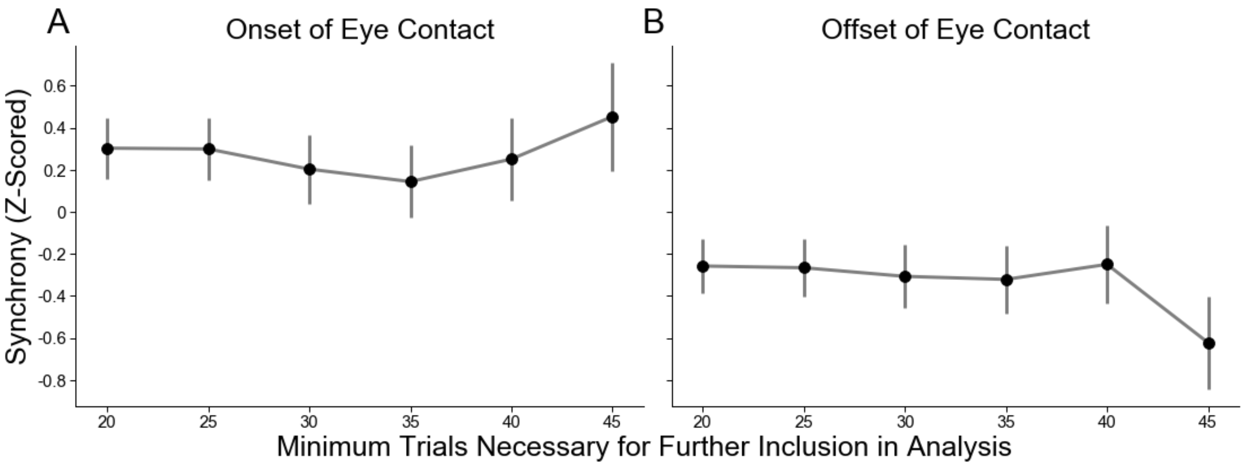
Controlling for naturalistic data issues in event-related analysis

Each dyad had a unique time series of eye contact corresponding to their natural behavior during their conversation. Allowing eye contact "events" to vary freely provided two potential challenges for event-related analyses. First, multiple instances of eye contact occurring closely in time could cause issues when, for example, dyads made eye contact, broke it, and then made it again a second later. In our original analysis, these two instances of eye contact would have corresponding synchrony time series that overlapped with one another, but would both be included in the analysis. Second, dyads could vary widely in how much eye contact they made, with some dyads making very little, thereby potentially reducing the reliability of their average pupillary fluctuation around eye contact. To investigate the potential impact of these two data issues on our original result, we conducted an analysis that included only instances of eye contact that were spaced at least four seconds apart and accepted only dyads who had at least 30 of these more widely-spaced moments of eye contact. Doing so did not change any of the results reported in the main text that used all the data: a quadratic contrast remained the best fit for dyads' eye contact onset synchrony curves ($\beta = -0.47$,t = -2.88, CI = -0.8 - -0.15, p = 0.004) and for dyads' eye contact offset synchrony curves ($\beta = 0.59$, t = 3.65, CI = 0.27 - 0.91, p < 0.001; supplementary figure 3).



Supplementary Figure 3. Analysis plots limited to data in which eye contact events were spaced >4 seconds. **A)** Time series of synchrony in the four seconds leading up to and following the onset and **B)** offset of eye contact.

We chose 30 trials as our minimum because it was the most stringent threshold that still retained at least two-thirds of dyads. However, to ensure our result was not due to an arbitrary threshold choice, we computed synchrony curves at a range of trial minimums (20-45) and found that, regardless of the threshold, the results were consistent (for onset, β ranged from -0.84 to -0.41, t ranged from -3.57 to -2.88, p ranged from 0.0004 to 0.02; for offset, β ranged from 0.42 to 1.3, t ranged from 2.02 to 4.89, p ranged from < 0.0001 to 0.04). Synchrony values for the full range of minimum trial thresholds at the onset and offset of eye contact are plotted in supplementary figure 4.



Supplementary Figure 4. **A)** Line graph depicting average synchrony values at the onset and **B)** offset of eye contact at a range of minimum trial thresholds. All error bars plotted depict standard error. Our original result —that eye contact peaks at the onset of eye contact and declines until the offset of eye contact — was robust to these additional controls.

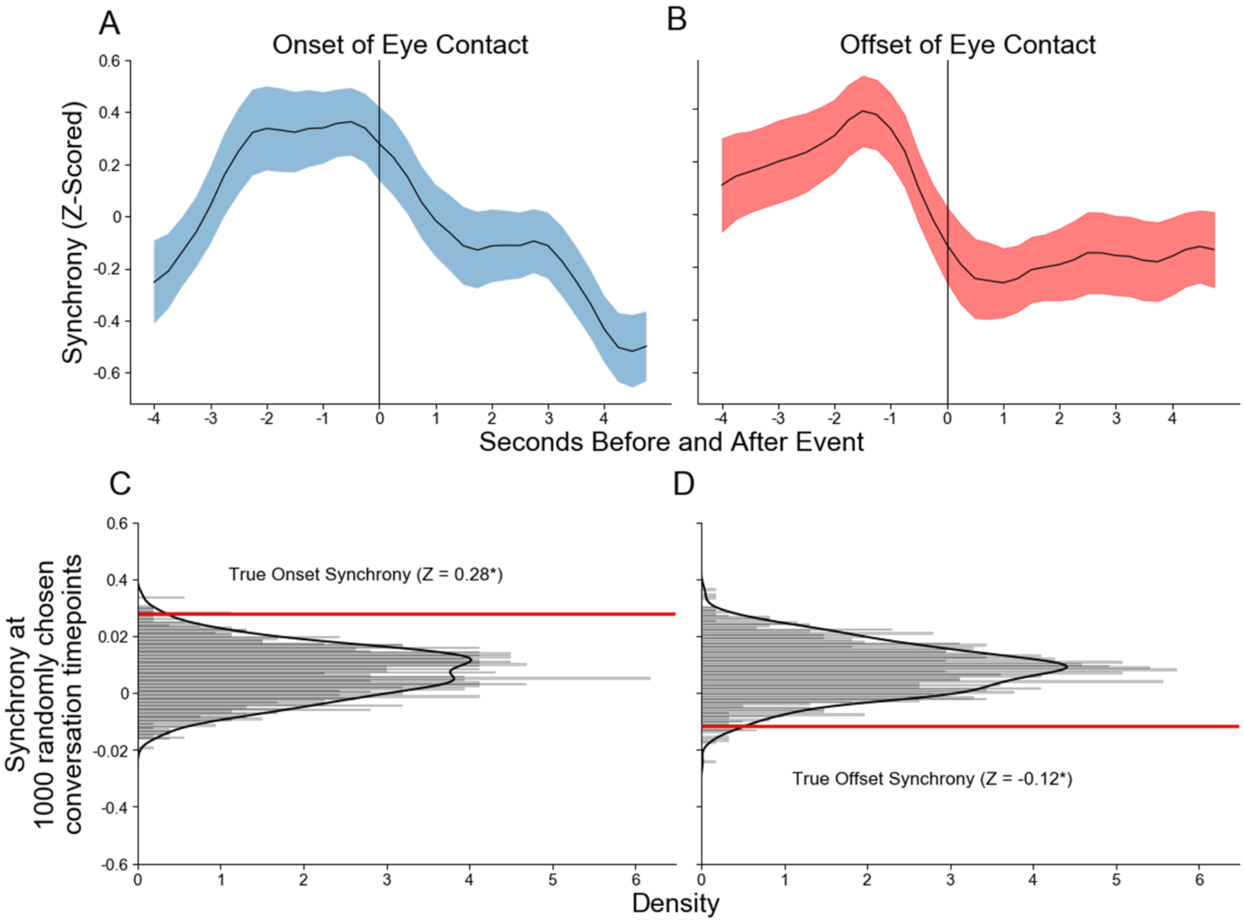
Examining pupillary synchrony including instances of eye contact lasting less than one second

The 1hz sampling rate used to compute pupillary synchrony necessarily excluded a large number of brief (<1 second) eye contact events from our analysis (M number of instances excluded = 117.28, SD = 49.72). We tested whether these brief instances of eye contact also produced fluctuations of pupillary synchrony similar to the longer (>1 second) instances in the primary analysis.

Capturing brief instances of eye contact required an increase in the sampling rate of pupillary synchrony. In order to ensure that any change in our original pupillary synchrony curves was the result of the new inclusion of brief instances of eye contact and not the change in our pupillary synchrony sampling window, we calculated pupillary synchrony in one-second increments, but used a 250ms rolling window (250ms is the shortest duration of an eye fixation[1]). In this way, the DTW algorithm still assessed the optimal warping path over a larger window (1 minute), but recomputed that path every 250ms making it additionally sensitive to fluctuations associated with brief eye contact. Because of the rolling window, there was no longer one specific time point that captured the onset of an instance of eye contact, there were five. To try to get a more accurate "onset" or “offset” using the rolling window, we assigned five synchrony curves – corresponding to the five samples in which the onset or offset was measured – to each instance of eye contact.

Consistent with the results reported in the main text, the model predicting the onset of eye contact was significant ($f(35,1656) = 3.78$, $R^2 = 0.05$, $p < 0.001$), with a quadratic contrast showing the best fit for dyads’ pupillary synchrony curves ($\beta = -1.11, t = -7.81$, CI = -1.39 – -0.83, $p < 0.001$). The model predicting the offset of eye contact was also significant ($f(35,1656) = 2.31$, $R^2 = 0.03$, $p < 0.001$), with a linear contrast showing the best fit for dyads’ pupillary synchrony curves ($\beta = -0.97, t = -6.75$, CI = -1.25 – -0.69, $p < 0.001$; supplementary figure 5A and 5B).

.



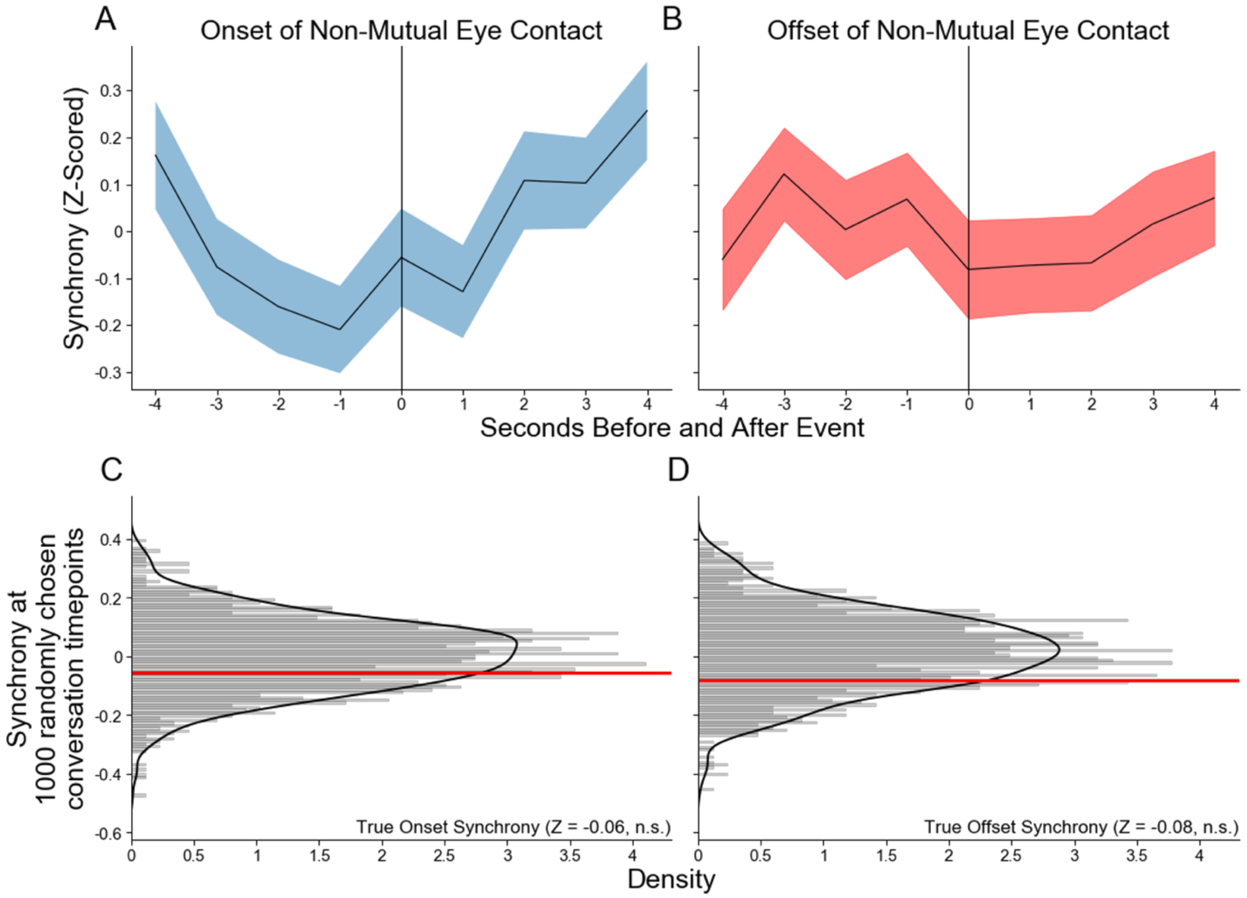
Supplementary Figure 5. Event-related analysis for moments of eye contact lasting 250ms or longer, with pupillary synchrony calculated in a rolling window of one-second segments with 250ms of overlap. **A)** Time series of pupillary synchrony in the four seconds leading up to and following the onset and **B)** offset of eye contact. **C)** Results of two permutation tests comparing the onset and **D)** offset of eye contact to 1000, randomly chosen, nine-second moments in each conversation, per participant. The distributions depicted above were created by taking the pupillary synchrony value at the "onset" and "offset" point (position 5) of each randomly chosen moment in the conversation. The true pupillary synchrony values for the onset and offset of eye contact are represented by the red horizontal lines in each figure. As in the original analysis, pupillary synchrony at the onset and offset of eye contact is significantly higher and lower, respectively, than would be expected by chance.

To see whether pupillary synchrony at the onset and offset of eye contact was higher or lower than would be expected at any other point in the conversation, we again compared true eye contact onsets and offsets to a randomly sampled (using python’s “random” package), equal number of moments in the conversation where eye contact was not made, creating “pseudo onsets” and “pseudo offsets.” (see results in main text for full description of this analysis). We found that synchrony at the onset and offset of eye contact, including brief eye contact, was significantly higher and lower, respectively, than would be expected at any other point in the conversation (onset Z = 0.28, p=0.013; offset Z = -0.12, p = 0.018; figure 5C and 5D).

Investigating the fluctuations of pupillary synchrony around non-mutual eye contact

Our main finding was that pupillary synchrony peaked at the onset of eye contact and then immediately began to decline until reaching its nadir at the offset of eye contact. We tested whether pupillary synchrony also fluctuates around non-mutual eye contact -— when one individual gazes at their partner’s eyes, but their partner does not reciprocate. To perform this analysis, we computed DTW (as an inverse measure of pupillary synchrony) around each instance of non-mutual eye contact, following the method described in the main text (see methods: event-related analysis).

Specifically, we computed a linear model predicting the average pupillary synchrony curve per dyad and specifying planned contrasts for how synchrony might vary over time. The model for the offset of non-mutual eye contact was not significant (f(8,837) = 0.36, R^2 = 0.003, p = 0.9; see figure 6B). The model for the onset of non-mutual eye contact was significant (f(8,837) = 2.15, R^2 = 0.01, p = 0.02; see supplementary figure 6A). However, this fit was in the opposite direction of what we observed with mutual eye contact and did not survive permutation testing. Permutation testing involved randomly sampling an equal number of moments in the conversation where eye contact was not made, creating “pseudo onsets” and “pseudo offsets” (see results in main text for full description of this analysis). Pupillary synchrony at the onset or offset of non-mutual eye contact was not significantly higher or lower than would be expected at any other point in the conversation (onset Z = -0.06, p=0.69; offset Z = -0.08, p = 0.77; figure 6C and 6D). This result suggests that the fluctuations of synchrony around mutual eye contact are unique to mutual eye contact. One person looking at their partner’s eyes is not enough to create reliable changes in pupillary synchrony. We note that because non-mutual eye gaze involves a gaze shift, this analysis further suggests that gaze shifts alone cannot account for the relationship observed between mutual eye contact and pupillary synchrony.



Supplementary Figure 6. Event-related analysis for moments of *non-mutual* eye contact. **A)** Timeseries of synchrony in the four seconds leading up to and following the onset and **B)** offset of non-mutual eye contact. **C)** Results of two permutation tests comparing the onset and **D)** offset of non-mutual eye contact to 1000, randomly chosen, nine-second moments in each conversation, per participant. The distributions depicted above were created by taking the pupillary synchrony value at the “onset” and “offset” point (position 5) of each randomly chosen moment in the conversation. The true pupillary synchrony values for the onset and offset of eye contact are represented by the red horizontal lines in each figure. Pupillary synchrony at the onset and offset of non-mutual eye contact is not significantly higher or lower than would be expected by chance.

Permutation testing to verify the relationship between eye contact, pupillary synchrony, and conversational engagement

Eye contact and pupillary synchrony were significantly and positively predictive of dyads' mean reported conversational engagement. Further, eye contact marginally moderated these effects such that, when dyads *were not* making eye contact, synchrony and engagement were positively related, but when dyads *were* making eye contact, this positive relationship was reversed. However, these main effects and interactions were small (eye contact main effect $\beta = 0.028$; synchrony main effect $\beta = 0.012$; interaction effect $\beta = -0.017$). To further test the reliability of these effects, we permuted our data within and between subjects 5000 times and compared our true effects to null distributions created by those permutations. For these tests, within-subjects permutations consisted of shuffling the eye contact and synchrony time series for each dyad and computing the relationship between these shuffled time series and dyads' mean engagement. Between-subjects permutations consisted of shuffling fully intact eye contact and synchrony time series between different dyads (e.g. synchrony from dyad A and eye contact from dyad B is assigned to dyad C) and computing the relationship with engagement for these pseudo-dyads.

Eye Contact Main Effect

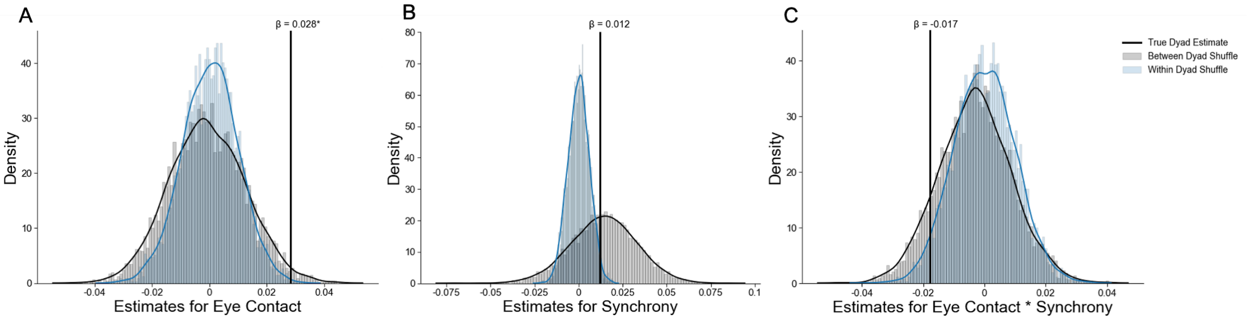
The relationship between eye contact and reported engagement (true $\beta = 0.028$, $p = 0.006$) was robust to both within-subjects ($p = 0.002$) and between-subjects ($p = 0.02$) permutation tests (figure 7A). This suggests that eye contact is a robust predictor of conversational engagement, above and beyond random variation in the eye contact time series and in a way that is specific to the individual structure of a dyad's conversation. The finding that this effect is robust to a between-subjects permutation test suggests that dyads are not merely adhering to global conversation norms when employing eye contact, but making it as needed to increase engagement in their unique conversations.

Pupillary Synchrony Main Effect

The relationship between pupillary synchrony and reported engagement (true $\beta = 0.012$, $p = 0.05$) was robust to a within-subjects permutation test ($p = 0.02$), suggesting that pupillary synchrony predicts conversational engagement above and beyond random variation in the synchrony time series. However, this effect was not robust to a between-subjects permutation test (p = 0.56; figure 7B). The lack of a between-subjects effect was due, at least in part, to a linear trend across dyads to become both more engaged and more synchronous over the course of a conversation. This suggests that, as dyads conversed, they increased shared attention which corresponded to a similar increase in reported engagement.

Interaction between Eye Contact and Pupillary Synchrony on Engagement

The marginal interaction between pupillary synchrony and eye contact on reported engagement (true β = -0.017, p = 0.08) was robust to a within-subjects permutation test (p = 0.04), suggesting that this interaction, while marginal, predicted conversational engagement above and beyond random variation in synchrony and eye contact time series. However, this effect was not robust to a between-subjects permutation test (p = 0.12; figure 7C) due, at least in part, to pupillary synchrony and reported engagement being correlated across all dyads. Thus, although pupillary synchrony and engagement are inversely correlated *during* eye contact, outside of these moments, pupillary synchrony and reported engagement increase together over the course of the conversation.



Supplementary Figure 7. Results of both within- and between-subjects permutation tests testing **A)** the positive relationship between eye contact and reported engagement, **B)** the positive relationship between pupillary synchrony and reported engagement, and **C)** the interaction between eye contact and synchrony on reported engagement against null-distributions created by shuffling eye contact and synchrony time series 5000 times both within- and between- dyads. The relationship between eye contact and engagement was robust to both within- and between- subjects permutation tests. The relationship between pupillary synchrony and engagement and the interaction between eye contact and pupillary synchrony on engagement were robust to within-subjects permutation tests. However, we found that global increases in both shared attention and mean reported engagement over the course of dyads' conversations produced non-significant between-subjects permutation tests for both the relationship between pupillary synchrony and engagement and the interaction between synchrony and eye contact on engagement.

References

1. Rogers, S.L., Speelman, C.P., Guidetti, O., & Longmuir, M. (2018). Using dual eye-tracking to uncover personal gaze patterns during social interaction. *Scientific Reports, 8,* 1-9. DOI:10.1038/s41598-018-22726-7