Supplementary Material

Driving Simulator and Environment Specifications

The driving simulator was manufactured by DriveSafety and included the front three quarters of the body and the front interior of a sedan. It features a projected LCD instrument cluster controlled by a computer, foot controls, and realistic steering force feedback. The virtual road environment was projected onto three forward screens and one rear screen, each with a resolution of 1024 x 768 pixels, for 120 degrees of forward view and 40 degrees of rear view visible through the side and rearview mirrors. The driving environment contained standard roads, intersections, traffic control devices, other vehicles, and pedestrians; in addition, purely visual elements including buildings, trees, and sky were present. The simulator's sound system produced realistic interior and exterior sounds, and road vibration was simulated through the floorboard.

Coherence Calculation

An advantage of coherence compared to correlation in connectivity analysis is its relative insensitivity to differences in the shape of the hemodynamic response function (HRF) (Sun et al., 2004; Lauritzen et al., 2009), which has been shown to vary between regions in the same individual's brain (Handwerker et al., 2004). Coherence was computed using Welch's method with a 48-point discrete Fourier transform Hanning window and a 24-point overlap between windows (Welch, 1967).

Networks of Interest

We utilized the same centers for the regions of interest in the theory-driven analysis (social pain and mentalizing networks) as in Schmälzle et al. (2017), with the exception of replacing the nodes in the lateral temporal cortex with the temporal poles by choosing the closest regions in the atlas (Power et al., 2011). This change was made in order to be more consistent with the literature,

especially D'Argembeau et al. (2007), from which the coordinates are drawn, and does not substantively change the conclusions reported. The regions are pictured below along with their MNI coordinates.
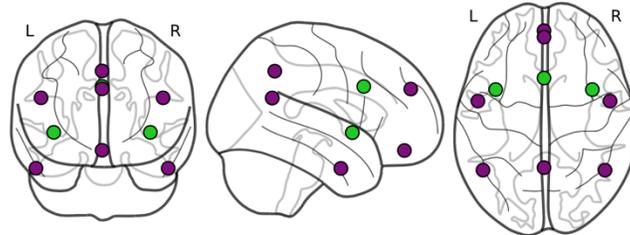

Figure S1. The locations of the social pain and mentalizing networks' regions of interest. The mentalizing network is shown in purple, and the social pain network is green.

Network	Region	Coordinates
Social pain	ACC	(0, 16, 32)
Social pain	left AI	(-38, 7, -4)
Social pain	right AI	(38, 7, -4)
Mentalizing	dorsomedial PFC	(0, 53, 30)
Mentalizing	ventromedial PFC	(0, 48, -18)
Mentalizing	PC	(0, -54, 44)
Mentalizing	left TPJ	(-48, -56, 23)
Mentalizing	right TPJ	(48, -56, 23)
Mentalizing	left TP	(-52, -2, -32)
Mentalizing	right TP	(52, -2, -32)

Table S1. MNI coordinates of the centroids of the regions in the social pain and mentalizing networks of interest.

Global Connectivity and Overfitting

Our method relies on the global connectivity of each region rather than connectivity between pairs of regions to minimize the risk of overfitting a predictive model from connectivity data. One danger

when using connectivity as a predictor is the large number of pairwise connections between regions. Our model considered the global connectivity of 30 regions of interest from a whole-brain parcellation of 264 regions. If we computed connectivity between each pair of regions, the model would contain 7,455 potential features. With only 57 participants in our sample, creating models from thousands of connectivity features in this dataset would likely overfit the training sets and have very poor generalization to the population. Consequently, we focused the analysis of the overall connectivity from our 30 regions to the rest of the brain, as this provides an index of each region's global brain activity. This allows us to address our theoretical question of interest while overcoming the technical challenge of overfitting the model since the resulting model only has 30 theory-guided potential features (instead of 7,455).

Feature Selection in Scikit-Learn

For each training set of 56 participants, we used scikit-learn's SelectKBest algorithm to sequentially select the k best features according to their F scores (Pedregosa et al., 2011). Common wisdom suggests choosing approximately one feature for every 10-15 observations for small sample sizes (Babyak, 2004), corresponding to between 4 and 6 features for our dataset; however, we expanded our search to 10 since our features were theoretically motivated. We then performed leave-one-out cross-validation using ordinary least squares regression so that each participant's conformity to peer influence was predicted using a model trained only on the other 56 participants. In the first analysis with 30 theory-driven regions, the model with highest accuracy in cross-validation occurred with k = 7, so we focused on the 7-feature model. In the second whole-brain analysis with 264 regions, models with 5 and 6 features performed very well, and we opted to select 6 features based on cross-validation scores.

Controlling for Drive Order

Because participants completed the solo and passenger drives in random order, we performed a set of cross-validated analyses to account for the possibility of drive order as a confounding variable. Specifically, we trained a model with 31 potential features: the 30 brain regions from the networks of interest plus drive order. However, including drive order as a predictor did not improve predictivity, so we did not pursue this approach further.

Connectivity Differences in Response to Social Exclusion

Although the mean connectivity to the rest of the brain in seven of the nine regions in the social pain network was higher during exclusion than during inclusion, none was significantly different (Figure S2). Similarly, we observed nonsignificant trends in the mentalizing network: Five of the six regions in the medial PFC were more connected to the whole brain on average during inclusion than during exclusion, whereas each region in the bilateral temporal poles exhibited more global connectivity during exclusion than during inclusion on average. Overall, although we observed weak connectivity differences between Cyberball conditions, with two thirds of the regions more strongly connected to the rest of the brain during exclusion than during inclusion, these differences were not significant or robust. This arose due to substantial heterogeneity across participants in the degree of change from exclusion to inclusion, highlighting the importance of considering individual differences.
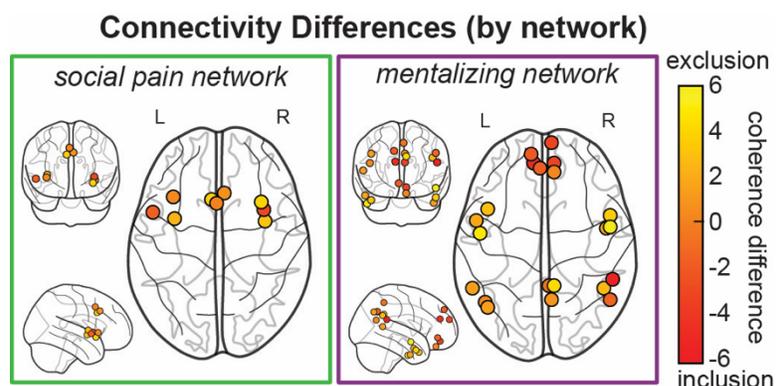
Figure S2. Differences in global connectivity between social exclusion and social inclusion in the social pain network (left) and the mentalizing network (right), averaged across all participants. The nodes with more connectivity to the rest of the brain during exclusion than during inclusion are yellow, and the nodes that are more connected during inclusion than during exclusion are red. Note: none of the differences pictured is statistically significant; results are depicted for descriptive purposes.

Predictivity for Both Risk-Accepting and Risk-Averse Passengers

To verify that global connectivity was predictive of conformity for participants regardless of passenger type, we examined the the relationship between the observed values of conformity and the predicted values of conformity of our model (in cross-validation) for both risk-accepting and risk-averse passengers. The correlation between actual and predicted conformity was statistically significant ($p < 0.005$) for both participants with risky passengers and those with safe passengers; see Figure S3.
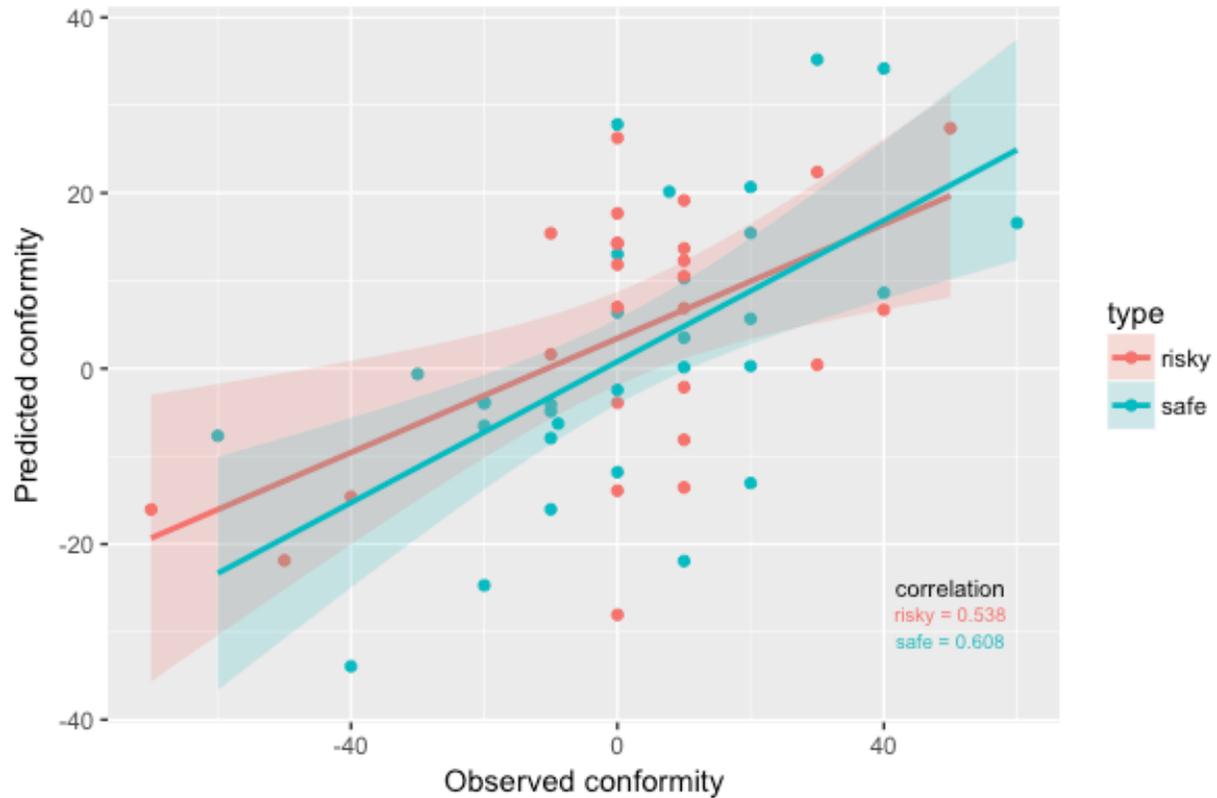
Figure S3. Predicted vs. observed conformity for participants with risky and safe passengers. The predicted conformity (within cross-validation) is plotted against observed conformity for all participants (red = participant with risk-accepting passenger; teal = participant with risk-averse passenger). Lines of best fit are plotted separately by passenger type with 95% confidence intervals shown; the same relationship between global connectivity and conformity is found for participants with both passenger types.

Whole-brain Connectivity Analysis Confirms Importance of Theory-Driven Regions

We replicated our analysis using theory-driven networks using a whole-brain analysis. This analysis was conducted in part to assess the robustness of the first analysis, investigating whether the *a priori* networks of interest would still be identified among the 264 whole-brain regions. Furthermore, it examined whether the out-of-sample R^2 value could be substantially improved

when the connectivity features were not limited to approximately ten percent of the possible regions (i.e., when not limiting to theory-driven ROIs but rather searching across the whole brain).

As shown in Figure S3 (top left), the best prediction was achieved from 6 of the 264 whole-brain regions, and these top 6 regions predicted individual differences in conformity with an out-of-sample R^2 of 0.303 (root mean squared error of 20.40 in cross-validation), which is slightly less but comparable to the theory-driven analysis. Half of these regions overlap with or neighbor the theory-driven social pain and mentalizing regions, including one region in the AI (social pain region, outlined in green in Figure S3) and two regions in the right TPJ (one mentalizing region, outlined in purple in Figure S3, and an adjacent TPJ atlas ROI). Furthermore, the whole-brain analysis identified three additional regions, two in the left motor/premotor cortex and one in the right dorsolateral PFC (outlined in black in Figure S3).

Following the same procedure used in the theory-driven analysis, we performed 10,000 permutations to verify that this result does not arise from chance or overfitting, and the whole-brain model outperformed 99.77% of the shuffled models (p = 0.0024; Ojala and Garriga, 2010). We also confirmed that the difference in connectivity between exclusion and inclusion was more predictive than that of either state alone, with coherence calculated at any frequency band, with social exclusion being more predictive (R^2 = 0.219) than social inclusion (R^2 = 0.161).
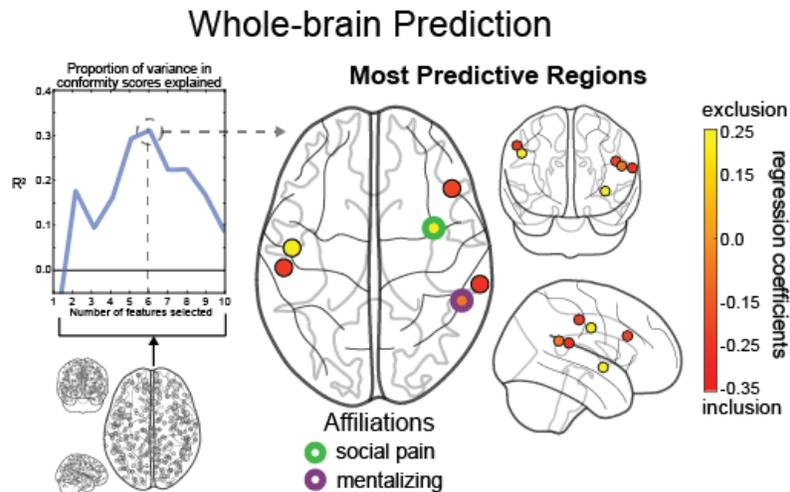
Whole-brain Prediction

Figure S3. The regions from the whole brain atlas whose connectivity to the rest of the brain is most predictive of conformity. Top left: The out-of-sample R^2 values obtained when selecting between 1 and 10 regions as features in our model, with 6 features being most predictive of conformity. Right: The 6 regions selected by our feature-selection algorithm. The region with a green outline is in the social pain network, and the region with a purple outline is in the mentalizing network. The other four regions were not among the 30 regions representing either network of interest. The color of the center of each region indicates its regression coefficient. Regions for which more connectivity during social exclusion than social inclusion is predictive of conformity are yellow, and regions for which more connectivity during social inclusion than social exclusion is predictive of conformity are red.

References

Babyak, M. A. (2004) What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosomatic Medicine, 66, 411–21.

D'Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Balteau, E., Luxen, A., et al. (2007) Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. Journal of Cognitive Neuroscience, 19, 935–44.

Handwerker, D. A., Ollinger, J. M., & D'Esposito, M. (2004) Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. Neuroimage, 21, 1639–51.

Lauritzen, T. Z., D'Esposito, M., Heeger, D. J., & Silver, M. A. (2009) Top-down flow of visual spatial attention signals from parietal to occipital cortex. Journal of Vision, 9, 1-14.

Ojala, M., & Garriga, G. C. (2010) Permutation tests for studying classifier performance. Journal of Machine Learning Research, 11, 1833–63.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011) Scikit-learn: machine learning in Python. Journal of Machine Learning Research, 12, 2825−30.

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., et al. (2011) Functional network organization of the human brain. Neuron, 72, 665–78.

Schmälzle, R., O'Donnell, M. B., Garcia, J. O., Cascio, C. N., Bayer, J., Bassett, D. S., et al. (2017) Individual differences in connectivity dynamics during social exclusion reflect social network structure. Proceedings of the National Academy of Sciences of the United States of America.

Welch, P. D. (1967) The use of a fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Transactions on Audio Electrouacoustics. AU-15, 70-3.