Purity Homophily in Social Networks

Morteza Dehghani[1,2,3], Kate Johnson[1], Joe Hoover[1,2], Eyal Sagi[4], Justin Garten[2,3],

Niki Jitendra Parmar[2,3], Stephen Vaisey[5,6], Rumen Iliev[7], and Jesse Graham[1]

[1]Psychology Department, University of Southern California
[2]Brain and Creativity Institute, University of Southern California
[3]Computer Science Department, University of Southern California
[4]Kellogg School of Business, Northwestern University
[5]Sociology Department, Duke University
[6]Duke Network Analysis Center, Duke University
[7]Ford School of Public Policy, University of Michigan

Author Note

Acknowledgements. We would like to thank Antonio Damasio, Doug Medin, Scott Atran, Steve Read, Simone Schnall, Jon Haidt, Jeremy Frimer, Nina Strohminger, and Peter Carnevale for their comments. This research was supported by AFOSR YIP to MD and NSF IBSS Grant #1520031 to MD, SV, and JG.

Correspondence concerning this article should be addressed to Morteza Dehghani, mdehghan@usc.edu, 3620 S. McClintock Ave, Los Angeles, CA 90089-1061.

Word count (excluding title, references, author affiliations, acknowledgments, figures and figure legends, but including the abstract): 5912

Abstract

Does sharing moral values encourage people to connect and form communities? The importance of moral homophily (love of same) has been recognized by social scientists, but the types of moral similarities that drive this phenomenon are still unknown. Using both large-scale, observational social-media analyses and behavioral lab experiments, we investigated which types of moral similarities influence tie formations. Analysis of a corpus of over 700,000 tweets revealed that the distance between two people in a social-network can be predicted based on differences in the moral purity content – but not other moral content – of their messages. We replicated this finding by experimentally manipulating perceived moral difference (Study 2) and similarity (Study 3) in the lab and demonstrating that purity differences play a significant role in social distancing. These results indicate that social network processes reflect moral selection, and both online and offline differences in moral purity concerns are particularly predictive of social distance. Our research is an attempt to study morality indirectly using an observational big-data study complemented with two confirmatory behavioral experiments carried out using traditional social-psychology methodology.

Keywords: Homophily, Morality, Computational Social Science, Social Networks

Purity Homophily in Social Networks

Social scientists have long recognized the importance of homophily (love of the same) for social bonds – the idea that "birds of a feather flock together" (Byrne, 1961; Lazarsfeld, Merton, & Others, 1954; McPherson, Smith-Lovin, & Cook, 2001). However, most of this research has emphasized how individuals associate or bond with similar others based on demographics such as age, gender, or socioeconomic status (Kossinets & Watts, 2009). More recently, scholars have identified moral values as another possible source of homophily (Vaisey & Lizardo, 2010). People prefer more social and physical distance from others who disagree with them on moralized social issues (Skitka, Bauman, & Sargis, 2005), and also prefer to live in communities with ideologically similar others (Motyl, Iyer, Oishi, Trawalter, & Nosek, 2014). One factor that might drive this tendency toward moral homogeneity is the function of moral cognition as a dynamic coordination device that facilitates third-party convergence on moral judgments (DeScioli & Kurzban, 2009, 2013). From this view, moral cognition can only coordinate third-party judgments to the extent that those third-parties have similar moral values and, accordingly, moral heterogeneity increases the risk of personal costs caused by making a minority moral judgment.

However, despite the growing evidence for moral homophily, little is known about what types of moral similarities matter in processes of moral homophily and social network evolution. In the current research, we approach this question through the framework of Moral Foundations Theory (MFT; Graham et al., 2011; Haidt & Joseph, 2004), which identifies multiple categories of moral values that vary in the degree to which individuals and groups endorse them: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and purity/degradation. We predicted that moral purity concerns carry more social weight than other concerns in determining distance on networks. Physical and spiritual purity concerns are linked with disgust and contamination sensitivities (Horberg, Oveis, Keltner, & Cohen, 2009; Lee & Schwarz, 2010; Preston & Ritter, 2012; Rozin, Haidt, & Fincher, 2009) and thereby tend to amplify perceptions of moral wrongness. Purity violations, compared to harm violations, also tend to be explained in terms of person-based attributes, compared to situation-based attributes (Chakroff & Young, 2015). If purity issues are more likely than other moral issues to lead to dispositional inferences about others, they might also have stronger effects on social network tie formation and dissolution as these behaviors are at least partially dependent on dispositional evaluations of others. We therefore predicted that moral foundation concerns would predict social distance between individuals and that *purity* concerns would be the strongest predictor of distance.

Prior research has been limited by its reliance on coarse measures of both morality and network structure (Graham, 2014). One platform of increasing importance for expressing moral concerns and ideals is social media. The public and persistent nature of social media presents an unprecedented opportunity to assess moral behavior “in the wild,” and to understand moral diversity in social networks and the dynamic relationship between network structure and moral content.

Our investigation relies on a mix of large-scale observational social media analysis and two behavioral lab experiments. In our first study, we used machine learning and natural language processing to measure moral concerns as expressed in real-world contexts, and to investigate how they relate to distance on a large social graph. We then replicated these findings with two behavioral experiments carried out using traditional social-psychology methodology.

Study 1

In studying morality in an environment like Twitter (or Facebook), it is vital to acknowledge the network structure surrounding each user, in addition to the content of their messages. Users are only exposed to tweets from the other users they follow, and a user's place in the network essentially determines the type of content they receive. In the current study, we investigate whether individuals' moral concerns can be used to predict the distance between them in their social network.

Method

Twitter[1] Data Collection. We used the Tweepy API[2] to access the public Twitter stream, which provides random samples of the data flowing through the network. We collected tweets related to the 2013 US government shutdown. We specifically chose this issue as it served to

[1] Twitter is an online social networking site that allows users to send short messages to each other. These messages are called "tweets", and are 140 characters or shorter. Users can read the tweets of the users they follow, and can "retweet" them (i.e., share other people's public messages). Currently, Twitter is the most popular social networking site that allows access to its database. There recently has been an upsurge in using Twitter data in psychological research (e.g. Eichstaedt et al., 2015; Barberá et al. 2015).

[2] http://docs.tweepy.org/en/latest/api.html

highlight the deep ideological differences between political groups in the US, and we anticipated it would receive wide online coverage. We started collecting data on the first day of the shutdown (October 1st, 2013). We searched the public stream for a list of hashtags and pages that were collected independently and agreed upon before we began collecting data. We stopped data collection on October 24th, about a week after the end of the government shutdown. We collected the following information about every tweet: the date and time the tweet was published, the ID of the user who published the tweet, and the content of the tweet itself. Following the period in which we collected the tweets, we gathered information about the network structure within the corpus using the Tweepy API. Specifically, we collected the list of followers and friends for every user in the corpus and used this information to map the network structure.

Overall, after removing non-English Tweets and duplicates, we formed a corpus of 731,332 tweets from 220,251 users. We were able to collect network structure from 188,467 users. Within the corpus of these tweets with network information, 46% of the tweets (339,816) were retweets, and were not used in the analysis due to the possibility of introducing confounds. We categorized tweets as retweets based on the following two criteria: 1. They were marked by Twitter as “retweets” 2. A duplicate of them existed in the corpus from a prior point in time.

Measuring Moral Rhetoric in Twitter. The most common approaches in capturing moral thought and behavior involve directly asking participants to make moral judgments about various issues or observing moral behaviors in lab-based psychological studies. Our work preserves these classic goals and methodologies while leveraging advances in computer science to approach them in a new way. Specifically, in this study we attempt to capture morality more indirectly by observing the naturally-occurring “moral residue” left behind in the texts of social discourse.

Figure 1 outlines the algorithm we used, which is described in detail in the supplemental materials. Our measure relies on word co-occurrence patterns as used by Latent Semantic Analysis (LSA) (Deerwester, Dumais, Landauer, & Furnas, 1990; Landauer & Dumais, 1997). In LSA words are represented as vectors in a semantic vector space derived from a matrix of word co-occurrence frequencies. One important property of this space is that the distance between two words is inversely related to the probability that they will co-occur in the text. A common measure used for the distance is the angle between the vectors representing the words where, for normalized vectors, the cosine of the angle is equivalent to the correlation between the vectors. Because of the way words are used in language, these patterns of co-occurrence are not random and words that relate to similar topics tend to occur together more frequently than unrelated words (e.g., ‘moon’ and ‘earth’ tend to occur with each other more frequently than either tends to occur with ‘gun’).

Another crucial property is that, because this space is linear, vectors can be aggregated in a relatively simple fashion. That is, it is possible to generate vectors that represent the content of short spans of text by using vector addition on the vectors representing the individual words contained in the span (Landauer, McNamara, Dennis, & Kintsch, 2013). We used this to compute an aggregate vector for every tweet in our corpus. Subsequently, we computed the average angle between a tweet and a vector representing a set of terms associated with a particular moral concern. These terms were identified based on the Moral Foundations Dictionary (Graham, Haidt, & Nosek, 2009). In this context, the 'loading' of a tweet for a particular category (e.g. purity), refers to the degree of overlap between the semantic vector representing the tweet, and the vector representing the category of interest.

Our method builds on standard word count techniques widely used in social sciences (e.g. Linguistic Inquiry and Word Count software; Pennebaker, Booth, & Francis, 2007). Specifically, rather than relying on how often a moral term appears in a corpus, we measure the degree of semantic similarity between the terms of interest and tweets. Effectively, LSA-based techniques provide estimations of contextual substitutability of words. For example, if 'rape' is a word of interest, our method determines that 'molestation', 'abuse', 'filthy', 'sex', etc. happen in the same context as 'rape' and provide an observable approximation of its semantic content. Therefore, in our analysis, not only words that are part of the Moral Foundations Dictionary, but words that have been determined to provide an estimation of the semantic content of the dictionary words, are weighted toward the moral loading of a tweet. An additional difference between our method and the standard word-count techniques is that our method is not limited to a fixed dictionary. Even though, the analysis starts with a set of seed words (in this case, words from the Moral Foundations Dictionary), it is the context of the text being analyzed that ultimately determines the vectors representing the categories.

This approach has already been applied to morally-loaded topics such as terrorism and abortion (Sagi & Dehghani, 2013). This method is flexible enough and has been applied to a wide variety of corpora, such as blog entries, political transcripts, and newspaper articles. Moreover, because previous applications of this approach have focused on analyzing 30-word snippets extracted from the texts based on keyword identification, we believe that it is suitable to the analysis of short texts, such as tweets.

To further investigate the effect of non-moral processes on distance, we chose the five psychological processes categories from the Linguistic Inquiry and Word Count Software (Social processes, Affective processes, Cognitive processes, Biological processes, and Relativity; Pennebaker et al., 2007) and used our method to calculate the loadings of each of the tweets with regards to these categories. These categories were used as control, and for "benchmarking" the predictive power of purity difference against several relevant non-moral categories. It is worth noting again that our analysis is based on the degree of overlap between the semantic vectors representing the categories and the vectors representing the tweets, and not on the individual word level. Therefore, the benchmarking analysis is not affected by the differences in sizes of the dictionaries used.

Community Detection. In order to identify the various communities in our data, we formed a network based on 'follower' information. That is, we connected two users on the social graph, using an edge, if one user was the follower of the other. This resulted in 14,904,481 edges (links between users) connecting the 188,467 nodes (users) in the graph. Given the size of the network, we made the simplification that the graph is undirected. This assumption allowed us to use a community detection algorithm that is linear in the number edges. Specifically, we used a community detection algorithm by (Clauset, Newman, & Moore, 2004) which works especially well for large graphs (supplemental materials), but only works on undirected graphs. We used the implementation of this algorithm available in the R iGraph (Csardi & Nepusz, 2006) package. This method discovered two main large communities in the network, which include 93.6% of the users. Manual verification of the affiliations of members (following major Democratic or Republican politicians) of these two communities revealed that they represented these two major political groups in the US.

Distance. We calculated distance between two users on the social graph using the length of the shortest path between them. The shortest path is calculated using breadth-first search in the R iGraph package. Two users have a distance of one when there is an edge connecting them (one of them directly ‘follows’ the other), and a distance of two if there is a user in between them (i.e. the two users have a friend in common, but they don’t directly follow each other). For our analysis we calculated the shortest-path distance between 10,000,000 random sets of users. Approximately, 99.9% of the users in the network are of distance 1 to 5 from each other. Therefore, we only considered distances that fall within this range. Distances of more than 5 are due to missing network structures in our dataset. For within community analysis, we restricted our analysis to distances 4 or less, as within community distance of 5 was quite rare in our database (less than 0.9% of the within community sample).

Analysis. To test the hypothesis that social distancing is associated more strongly with purity concerns than with other moral concerns, we first conducted pairwise comparisons of mean moral difference scores between user pairs for each moral concern at each level of network distance. We also replicated these pairwise comparisons for purity concerns within Liberal and Conservative network clusters in order to investigate the possibility that the hypothesized effect of purity on social distance holds only across political affiliations. Per our hypothesis, we expected to see a linear increase in purity difference with each increase in network distance, however, importantly, we did not expect to see a comparable effect for the other moral concerns. Further, we predicted that the linear effect of purity would be comparable within both Liberal and Conservative clusters. To conduct this analysis, we estimated generalized linear mixed effects models (Bolker et al., 2009; Raudenbush & Bryk, 2002) in which moral difference scores were the dependent variables; social distance was both the grouping factor (N = 5) and the independent variable; and users were included as random effects. While complex, this model permitted us to directly test our hypothesis while also accounting for random effects caused by repeatedly sampling moral loadings from the same users (for example, user A might follow users B, C, and D, and therefore three distance measures would be partially based on user A's moral concerns). We then used Tukey's Honest Significant Difference (Tukey, 1949), a conservative single step multiple comparison procedure, to compare users' mean moral difference scores for each moral concern at each level of distance (i.e. users that are directly connected to each other, those that have a distance of two, etc.). We then repeated these steps for clusters of Liberal and

Conservative users with purity difference as the dependent variable in order to investigate whether the effect of purity on network distance is stable across political groups.

In our second and primary analysis, we used a Support Vector Machine (SVM) regression function with a radial base kernel to evaluate the degree to which moral difference scores for each domain predict social distance. SVMs, which are a supervised machine learning technique first introduced by Vapnik (Vapnik, 2000), are conceptually similar to classical Ordinary Least Squares (OLS) regression in that they estimate the relationship between independent and dependent variables. However, SVMs are more robust to overfitting, compared to OLS regression, and are therefore able to provide a more accurate estimate of model error, particularly for large datasets. To test the hypothesis that purity concerns are the best predictor of social distance, compared to other moral concerns, we compared the root-mean-squared errors (RMSE), a common metric of model fit (Chai & Draxler, 2014; Moriasi et al., 2007), generated from random permutation tests. More specifically, for each regression model we randomly sampled 10,000 data points, 100 different times from the dataset, and for each sample-set performed 10-fold cross validation. We then compared the total RMSEs, across the 100 independent samples, between each moral concern. Finally, to assess the degree to which purity concern difference predicts social distance above and beyond the other moral concern differences, we regressed social distance on all five moral concern difference measures in an OLS multiple regression model.

Results

There was positive correlation between the distance between users and the differences in the purity loadings of their tweets, r (9990240) = 0.137, p < 0.001 (*Figure 2*). The generalized mixed effects model predicting purity difference scores from network distance revealed revealed a main effect of distance F(4, 9990232) = 27445, p < 0.001. General linear models with Tukey's HSD, with the same factors as above, revealed a significant increase in purity difference at each level of distance, such that the purity difference between users increased at every distance level (*Figure 3*): level two (M = 0.0323, SD = 0.025) versus level one (M = 0.0278, SD = 0.022), level three (M = 0.0371, SD = 0.030) versus level two, level four (M = 0.0445, SD = 0.035) versus level three, and level five (M = 0.0471, SD = 0.0465) versus level four (*Table 1*). In other words, users with similar levels of purity concerns had shorter connections between them and as the degree of differences between purity concerns increased so did the distance between the users. While the differences between these estimated means might seem small, it is important to keep in mind both the high dimensionality of the data and the precision of the estimation; as the *Cohen's d*'s indicate, these are substantial effects. Further, as predicted, this monotonic increase was not observed with any other MFT concerns (see *Tables 1a* and *2a* in supplemental material for descriptive statistics and pairwise comparisons). Moreover, examining the relationship between distance and purity difference *within* the two main large communities in the network revealed that this positive linear relationship, between distance and difference in purity loadings holds even within clusters (*Table 1*). This pattern indicates that the observed relationship cannot be explained by general differences between liberals and conservatives (see supplemental material).

Next, we performed Support Vector Machines (Vapnik, 2000) regression to investigate whether the distance between two users can be predicted based on the difference between the moral loadings of their tweets (method section). Specifically, across 100 iterations we randomly sampled 10,000 data points and estimated SVM models with 10-fold cross validation for each moral concern. After each iteration, we extracted the current model's RMSE, which ultimately yielded 100 RMSEs for each moral concern model. We then conducted *t*-tests comparing the distribution of RMSEs extracted from the purity models to those extracted from the other models. These results demonstrate that the difference in purity loadings was the most accurate predictor of distance compared to the loadings of other concerns (*Table 2*). Next, we entered each moral difference measure simultaneously in a multiple regression model estimated with the entire dataset. This model provided strong convergent support for the hypothesis that differences in purity concerns not only have a stronger association with social distance than distances in other moral concerns, but also that this association remains robust even after accounting for the effects of moral concern differences in other domains. More specifically, when controlling for all other moral difference measures, the effect of purity difference ($\beta = 0.13$, SE = 0.007) was substantially larger than the effect of any other moral difference measure (all $|\beta| < 0.06$, all SEs < 0.014; *Table* 3).

Finally, performing Support Vector Regression, with the same settings as above, we observe that difference in purity loadings is a significantly more accurate predictor of distance compared to differences in the loadings of the LIWC Psychological categories: purity vs. Social Processes: t (198) = -8.396, $p < 0.001$, d = -0.594; purity vs. Relativity: t (198) = -6.305, $p <$ 0.001, d = -0.446; purity vs. Cognitive Processes: t (198) = -10.665, $p < 0.001$, d = -0.755; purity vs. Biological Processes: t (198) = -6.1162 $p < 0.001$, d = -0.433; purity vs. Affective Processes: t (198) = -9.556, $p < 0.001$, d = -0.676 (see *Table 3a* in supplemental materials for additional statistics)

Study 2

The results of our first study revealed the existence of purity homophily in social networks. Specifically, we demonstrated that within our corpus there was an increase in social distance as a function of increase in difference in purity loadings. Study 2 was conducted to replicate the effects found in the Twitter data by experimentally manipulating perceived differences in moral concerns to test the effects of perceived differences in moral purity concerns on physical and social distancing preferences, and to determine if these effects were stronger than the effects of perceived differences in other moral concerns. This study was preregistered on the Open Science Framework (https://osf.io/237fk/).

Method

Participants. Three hundred participants (sixty per condition) were recruited for a study titled Judgments and Interactions using Amazon Mechanical Turk (Buhrmester, Kwang, & Gosling, 2011; Casler, Bickel, & Hackett, 2013) (*M* age = 32.66; 60.9% female). Each participant received $0.50 for their participation in the study.

Procedure. Participants were told that the study would ask them to “answer questions about how you perceive different scenarios and rate interaction partner” on the MTurk description page. After providing consent, participants first answered a 20-item questionnaire with the following instructions: “Below you’ll be presented with a variety of situations and be asked to say whether certain behaviors in those situations would be morally wrong. Please use the following scale from 1 to 7, to indicate the degree to which you judge the behavior to be wrong (if at all).” The 20 items were separated into five blocks representing each of the five moral foundations, with four moral judgments in each block. All items from each foundation were presented together on a single page, and the order that the participants completed the foundations was randomized, as was the order of items within each foundation.

The items used for each foundation were selected from a larger pool of items (Clifford et al. 2015), based on extensive pretesting to match foundations on average perceived wrongness and arousal. Example items are “You see a woman clearly avoiding sitting next to an obese woman on the bus” (care/harm); “You see a boy skipping to the front of the line because his friend is an employee” (fairness/cheating); “You see an American telling foreigners that the US is an evil force in the world” (loyalty/betrayal); “You see a group of teenagers joking loudly and goofing off during church services" (authority/subversion); and "You see two first cousins getting married to each other in an elaborate wedding" (purity/degradation; see supplemental materials for full list of items).

Next, participants were given information about their score compatibility with a fictional participant who had also completed the scale. All participants were told that their scores were highly similar to the other person's scores for four out of the five moral domains. Participants were randomly assigned to one of five moral foundation feedback conditions in which they were told that their scores were significantly different (in terms of percent similarity) than their partner's scores for either care, fairness, loyalty, authority, or purity.

Next, participants completed two items designed to assess physical and social distancing. Specifically, participants answered "If you were sitting on a bench with this person, how close to them would you be willing to sit?" on a 6-point scale from 1 (*as near as possible*) to 6 (*as far away as possible;* $M = 3.74$, $SD = 0.89$) and "According to my first feelings (reactions), I would willingly admit a person with these values into the following classifications" using the Bogardus social distance 7-point scale from 1 (*as close relatives by marriage*) to 7 (*would exclude from my country*; $M = 3.27$, $SD = 1.41$). These two items were then normalized using z-scores and added together to create a partner distancing score ($\alpha = .68$).

Finally participants answered demographic questions (age, gender, ethnicity, education, and political ideology) and an attention check item which asked them to identify on which moral foundation they and their partner differed.

Results

Thirty-two participants did not complete the survey and were not included in analyses. If an IP address was recorded more than once in our study, we only included the first set of data recorded with that IP address to minimize the effect of people taking the study multiple times; six responses were removed from analyses for this reason. Thirty-four participants were removed from analyses for failing the attention check item which asked them to identify the moral foundation that they and their partner differed on, leaving a total sample size of 235.

We conducted a one-way, between subjects ANOVA to test the effect of moral foundation feedback condition on z-scored partner distancing (overall $M = 0.02$, $SD = 1.73$). As expected, there was a significant effect of condition on partner distance, F (4, 230) = 3.38, $p < .05$, observed power = .84. Post hoc analyses using Fisher's Least Significant Difference (LSD) protected t-test criterion for significance indicated that participants in the purity foundation feedback condition preferred greater distance from the described partner (*z-score* $M = 0.87$, $SD = 2.09$) compared to all 4 other foundation feedback conditions, (*Figure* 4; Harm *z-score* $M = -0.12$, $SD = 1.46$, $p < .01$, 95% CI [0.36,1.79]; Fairness z-score $M = -2.60$, $SD = 1.67$, $p < .001$, 95% CI [0.51,1.93]; Loyalty *z-score* $M = -0.22$, $SD = 1.53$, $p < .01$, 95% CI [0.41,1.85]; Authority *z-score* $M = 0.03$, $SD = 1.66$; $p < .05$, 95% CI [0.16,1.60]). Participants who read that they responded differently than their partner on purity concerns reported wanting to sit further from their partner on a bench, and were less likely to want someone like their partner in close social circles, compared to participants who read that they responded differently from their partner on concerns from any other domain. All other comparisons between conditions were not significant.

Although previous research has shown demographic differences in purity judgments (Haidt, Koller, & Dias, 1993), there were no significant effects of age, ideology or education on partner distancing. Average purity domain scores predicted partner distancing across conditions, $t = 2.37$, $p < .05$, but average scores from the other 4 domains were not significant predictors of distancing. The effect of condition on partner distance remained significant after controlling for average domain scores and demographic variables, $F\ (4, 225) = 2.80$, $p < .05$, observed power = .76.

Study 3

In Studies 1 and 2 we found evidence supporting our hypothesis that purity concerns play a larger role in social distance dynamics than other moral concerns. Within a large network of Twitter users, Study 1 found that differences between users' purity concerns were more predictive of their network distance compared to their differences for other moral concerns and that the effect of purity differences was comparable between clusters of Liberal and Conservative users. In Study 2, we replicated this effect, finding that purity difference had a stronger positive effect on both self-reported social and physical distance, compared to the effects of the other moral foundations. In Study 3, we sought to replicate the effect of purity on social and physical distance as well as to further investigate several alternative hypotheses that could explain the results obtained in Studies 1 and 2. Specifically, in Study 2 we manipulated perceived moral *difference* by informing participants about their degree of moral difference from their partner. Because our target phenomenon is moral homophily (which may not be simply the opposite of moral heterophobia), in Study 3 we used the same paradigm as in Study 2, but we directly manipulated perceived moral *similarity*, as opposed to moral *difference*. We also measured both participants' political affiliation and religiosity and their perceptions of their partner's political affiliation and religiosity. This permitted us to investigate the possibility that perceptions of differences in purity concerns lead to inferences about others' political and religious positions and it is these inferences, rather than perceived differences in purity concerns, that drive social distance effects. One additional alternative explanation of the results obtained in Study 2 is that participants saw the behaviors depicted in the purity scenarios as particularly unusual and that the effect of purity on distance was primarily a novelty effect. To account for this possibility, scenarios that had been pre-rated for frequency of occurrence (Clifford, Iyengar, Cabeza, & Sinnott-Armstrong, 2015) and wrongness were selected so that the purity scenarios had an average perceived frequency and wrongness that fell between the average frequencies of the other foundations.

Method

Participants. Six hundred adult participants from the United States (*M* age = 31.65; 61.6% female) were recruited on Amazon Mechanical Turk.

Procedure. The general paradigm was the same as the paradigm used in Study 2, however there were some key differences. As in Study 2, participants were assigned to one of five conditions (*n*=120/condition). Across all conditions participants read twenty moral scenarios – four for each moral foundation – and indicated the degree to which they felt the action depicted in each scenario was morally wrong on a 7-point scale (1 = Not at all wrong, 7 = Extremely immoral). Importantly, these scenarios were selected from a pool of scenarios that have been rated according to their perceived frequency of occurrence and wrongness. We selected scenarios so that the mean frequency of occurrence for the purity scenarios fell roughly in the middle of the frequency means for the other foundations and perceived wrongness was balanced across foundations (See *Tables 4a* and *5a* in supplementary material for complete list of scenarios with wrongness and frequency ratings).

As in Study 2, participants were then provided with moral compatibility information about a potential future partner. However, in the current, study participants were only given partner information about one moral foundation. Specifically, all participants were told that they were 92% similar to their potential partner in one moral domain and no additional information about the other moral domains was provided. Importantly, rather than manipulating perceived difference, as in Study 2, in the current study we directly manipulated moral similarity.

Next, participants responded to the same physical and social distancing items used in Study 2. The data obtained through these items were then normalized and aggregated to create a partner distancing score (α = .51). Participants then were asked how religious they thought their partner was (1 = Not at all religious, 6 = Very religious) and how Liberal or Conservative they thought their partner was (1 = Very Liberal, 7 = Very Conservative).

Finally participants answered demographic questions (age, gender, ethnicity, education, and political ideology) and an attention check item which asked them to identify on which moral foundation they and their partner were similar.

Results.

Sixteen participants were dropped due to failing the attention check and twenty-seven duplicate IP addresses were removed. Accordingly, our primary analyses were conducted with N=557; however our results did not change even with the excluded participants included.

We conducted a one-way, between subjects ANOVA to test the effect of moral similarity on z-scored partner distancing. As expected, there was a significant effect of condition on partner distance, $F(4, 552) = 3.48$, $p < .05$, $\eta^2 = 0.025$. In order to compare the effects of each condition on partner distance, a simple regression model with z-scored partner distance as the dependent variable and condition as a 5-level independent variable was estimated[3]. Because the purity condition was coded as ‘0,’ the model intercept represents the mean partner distance for the purity condition and the other regression coefficients represent the estimated mean difference between the purity condition and each of the other conditions. As expected, participants in the purity condition indicated stronger partner approach (b = .22, 95% CI [0.09, 0.51], SE = .08) than did participants in any of the other condition (*Figure 5*): harm condition, b = -.32, 95% CI [-

[3] To estimate this model, condition was represented as four dummy coded variables with the purity condition as the reference category.

0.53, -0.10], *SE* = 0.10, *t*(552) = -2.92, $p < .05$; fairness condition, *b* = -.21, 95% CI [-0.43, 0.00, *SE* = 0.11, *t*(552) = -1.96, *p* = .05; Authority condition, *b* = -.35, 95% CI [-0.56, -0.13, *SE* = 0.11, *t*(552) = -3.10, $p < .05$; Loyalty condition, *b* = -.34, 95% CI [-0.56, -0.12], *SE* = 0.11, *t*(552) = -3.05, $p < .05$ (See *Table 6a* in supplemental material for distance means across conditions). Participants who read that their potential partner was similar to them on purity concerns reported desiring closer physical and social proximity, compared to participants who read that they were similar to their partner on concerns from other domains.

In order to rule out the possibility that perceptions of purity similarity lead to inferences of political and or religiosity similarity, two additional one-way, between-subjects ANOVAs with perceived political similarity and perceived religiosity similarity scores as the dependent variables and condition as the independent variable were estimated. Political and religiosity similarity scores were calculated by subtracting participants' reports of their own political affiliation and religiosity from ratings they assigned to their partner on these dimensions[4]. If perceived purity similarity has a differential effect on perceptions of others' political affiliation and religiosity, then participants assigned to the purity condition should have similarity scores closer to 0 than participants assigned to the other conditions. However, if perceived purity does not have a unique effect on perceived political and religiosity similarity, there should be no difference among similarity perceptions across conditions. As expected, our results support the latter hypothesis. Specifically, there were no significant differences in perceived political

[4] In order to calculate political similarity scores 135 participants who did not respond along the 1-7 liberal-conservative continuum were excluded from analysis (total *N* = 393).

similarity ($F(4, 417) = 1.23$, $p = 0.30$) or religiosity similarity between conditions ($F(4, 552) = 1.12$, $p = 0.35$). This indicates that participants did not make differential inferences about political or religious similarity in the purity condition. These results suggest that it cannot be the case that perceived purity similarity leads to greater inferences of political and or religious similarity and that it is similarity on these latter dimensions that drives purity homophily.

General Discussion

Our results indicate that purity homophily plays a significant role in the formation of connections in social networks. The observational social media study found that purity differences strongly predict social distance. The behavioral experiments confirmed this finding and demonstrated that differences in purity play a more significant role in social distancing than other moral concerns. It is notable that purity predicted social distance not only more than non-moral psychological factors (Study 1 and 3), but more than other types of moral concern too (Studies 1, 2 and 3). This highlights the role of moral purity concerns (and expressions of such concerns) not only for physical contamination and avoidance (Horberg et al., 2009; Lee & Schwarz, 2010; Preston & Ritter, 2012; Rozin et al., 2009), but for social contamination and avoidance as well. Given the large political differences in their endorsement (Graham et al., 2009), it is likely that moral purity concerns play an important role in increasing ideological migration and segregation in the U.S. (Motyl et al., 2014). Prior research has shown that conservative moralizing of topics like religion and sexuality (which are strongly associated with purity) are an important reason for disaffiliation from religious groups (Hout & Fischer, 2002, 2014). A further suggestion of possible purity-related homophily is demonstrated in a Pew Research study (“Political Polarization in the American Public - Pew Research ..,” n.d.), reporting that liberals and conservatives tend to avoid living close to members of the other group, and would be unhappy if their immediate family married someone from the opposing group. Purity may therefore play a strong role in moral homophily because it is frequently used as a basis for political and religious division in our society. The present research cannot determine whether more universal processes or more historically specific processes (or some combination of the two) are at work, but the current results strongly suggest that purity homophily plays an important role in structuring sociability. Overall, the connection between purity differences and the evolution of social networks is a robust one worthy of further study.

Finally, our work is an example in which “Big Data” observational findings are complemented with experimentally-manipulated behaviors in the lab. We believe such triangulation between unobtrusively observing large-scale online behaviors, experimental confirmation of the mechanism in the lab, and theory adjudication can provide many further insights about human cognition.

Authors' Contribution

Dehghani was involved in data collection, analysis and writing. Johnson was involved in data collection and analysis of Studies 2 and 3. Hoover was involved in analysis of Studies 1, 2, 3 and writing. Sagi, Garten & Parmar were involved in data analysis for Study 1. Vaisey, Iliev & Graham were involved in writing the paper.

References

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting From Left to Right Is Online Political Communication More Than an Echo Chamber? *Psychological science*, 0956797615594620.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127–135.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *6*(1), 3–5.

Byrne, D. (1961). Interpersonal attraction and attitude similarity. *Journal of Abnormal and Social Psychology*, *62*, 713–715.

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156–2160.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250.

Chakroff, A., & Young, L. (2015). Harmful situations, impure people: an attribution asymmetry across moral domains. *Cognition*, *136*, 30–37.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, *70*(6 Pt 2), 066111.

Clifford, S., Iyengar, V., Cabeza, R., & Sinnott-Armstrong, W. (2015). Moral foundations vignettes: a standardized stimulus database of scenarios based on moral foundations theory. *Behavior Research Methods*. http://doi.org/10.3758/s13428-014-0551-2

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9.

Deerwester, S., Dumais, S., Landauer, T., & Furnas, G. (1990). Indexing by latent semantic analysis. *Journal of the Association for Information Science and Technology*.

DeScioli, P., & Kurzban, R. (2009). Mysteries of morality. *Cognition*, *112*(2), 281–299.

DeScioli, P., & Kurzban, R. (2013). A solution to the mysteries of morality. *Psychological Bulletin*, *139*(2), 477–496.

Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., ... & Seligman, M. E. (2015). Psychological language on twitter predicts county-level heart disease mortality. *Psychological science*, *26*(2), 159-169.

Graham, J. (2014). Morality beyond the lab. *Science*, *345*(6202), 1242.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029–1046.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S., & Ditto, P. H. (2013). Moral Foundations Theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, *47*, 55-130.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: how innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55–66.

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*(4), 613–628.

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, *97*(6), 963–976.

Hout, M., & Fischer, C. S. (2002). Why More Americans Have No Religious Preference: Politics and Generations. *American Sociological Review*, *67*(2), 165–190.

Hout, M., & Fischer, C. S. (2014). Explaining Why More Americans Have No Religious Preference: Political Backlash and Generational Succession, 1987--2012. *Sociological Science*, *1*, 423–447.

Koleva, S. P., Graham, J., Iyer, R., Ditto, P. H., & Haidt, J. (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. *Journal of Research in Personality*, *46*(2), 184–194.

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network1. *The*

American Journal of Sociology, *115*(2), 405–450.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2013). *Handbook of latent semantic analysis*. Psychology Press.

Lazarsfeld, P. F., Merton, R. K., & Others. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and Control in Modern Society*, *18*(1), 18–66.

Lee, S. W. S., & Schwarz, N. (2010). Dirty hands and dirty mouths: embodiment of the moral-purity metaphor is specific to the motor modality involved in moral transgression. *Psychological Science*, *21*(10), 1423–1425.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*, 415–444.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *American Society of Agricultural and Biological Engineers*, *50*(3), 885–900.

Motyl, M., Iyer, R., Oishi, S., Trawalter, S., & Nosek, B. A. (2014). How ideological migration geographically segregates groups. *Journal of Experimental Social Psychology*, *51*(0), 1–14.

Political Polarization in the American Public - Pew Research .. (n.d.). Retrieved from http://www.people-press.org/files/2014/06/6-12-2014-Political-Polarization-Release.pdf

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic inquiry and word count: LIWC [Computer software]. *Austin, TX: Liwc. Net*.

Preston, J. L., & Ritter, R. S. (2012). Cleanliness and godliness: Mutual association between two kinds of personal purity. *Journal of Experimental Social Psychology*, *48*(6), 1365–1368.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.

Rozin, P., Haidt, J., & Fincher, K. (2009). From oral to moral. *Science*. http://doi.org/10.1126/science.1170492

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, *88*(6), 895–917.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, *5*(2), 99–114.

Vaisey, S., & Lizardo, O. (2010). Can Cultural Worldviews Influence Network Composition? *Social Forces; a Scientific Medium of Social Study and Interpretation*, *88*(4), 1595–1618.

Vapnik, V. (2000). *The Nature of Statistical Learning Theory*. Springer.

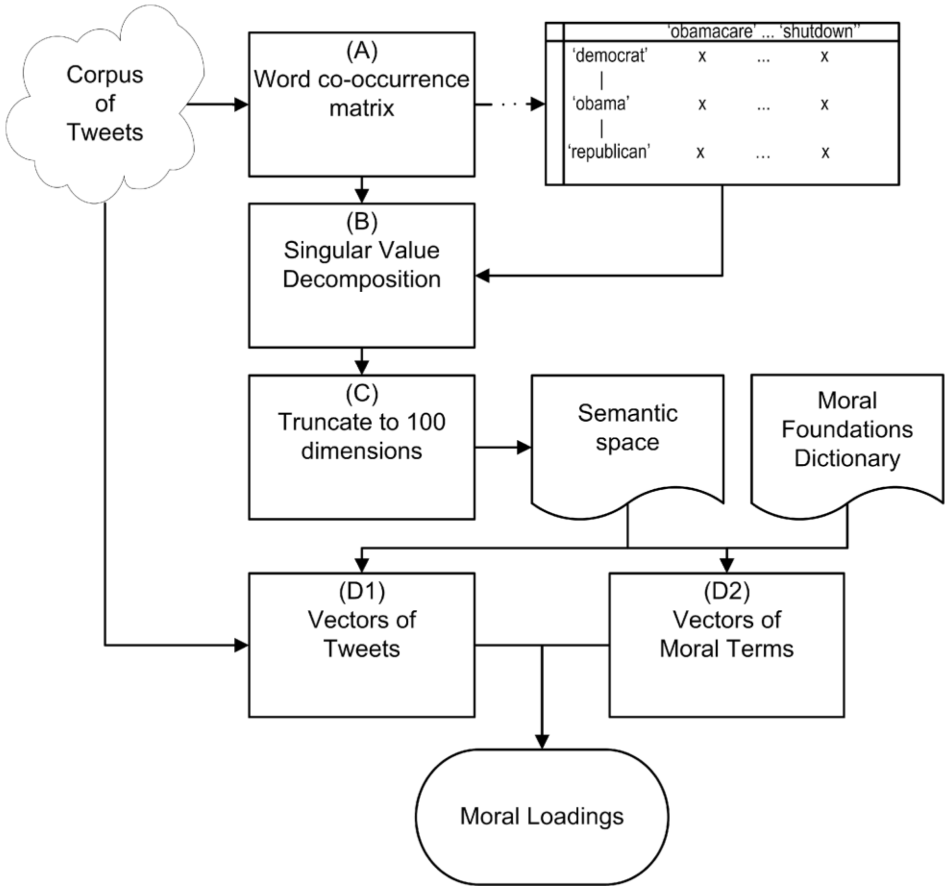

Figure 1. An algorithmic description of the method we used to measure moral rhetoric. We start by computing a matrix of word co-occurrence (A) and perform Singular Value Decomposition on it (B). The semantic space is created by truncating the matrix to 100 dimensions (C). Using vector addition we can then compute vectors for tweets (D1). We measure moral loadings of a tweet as the cosine of the mean angle between its vector with those of terms associated with each moral concern (D2).

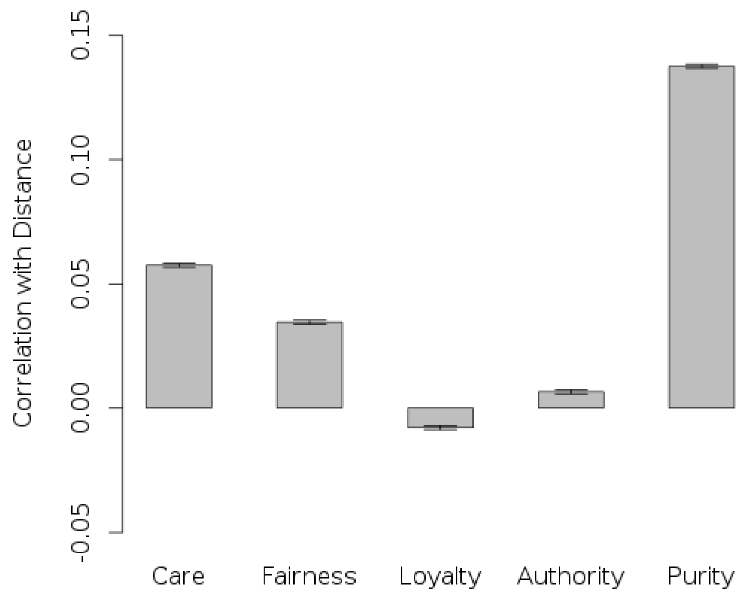

Figure 2. Correlation between node distance and differences in moral loadings. Error bars represent 99% CI.

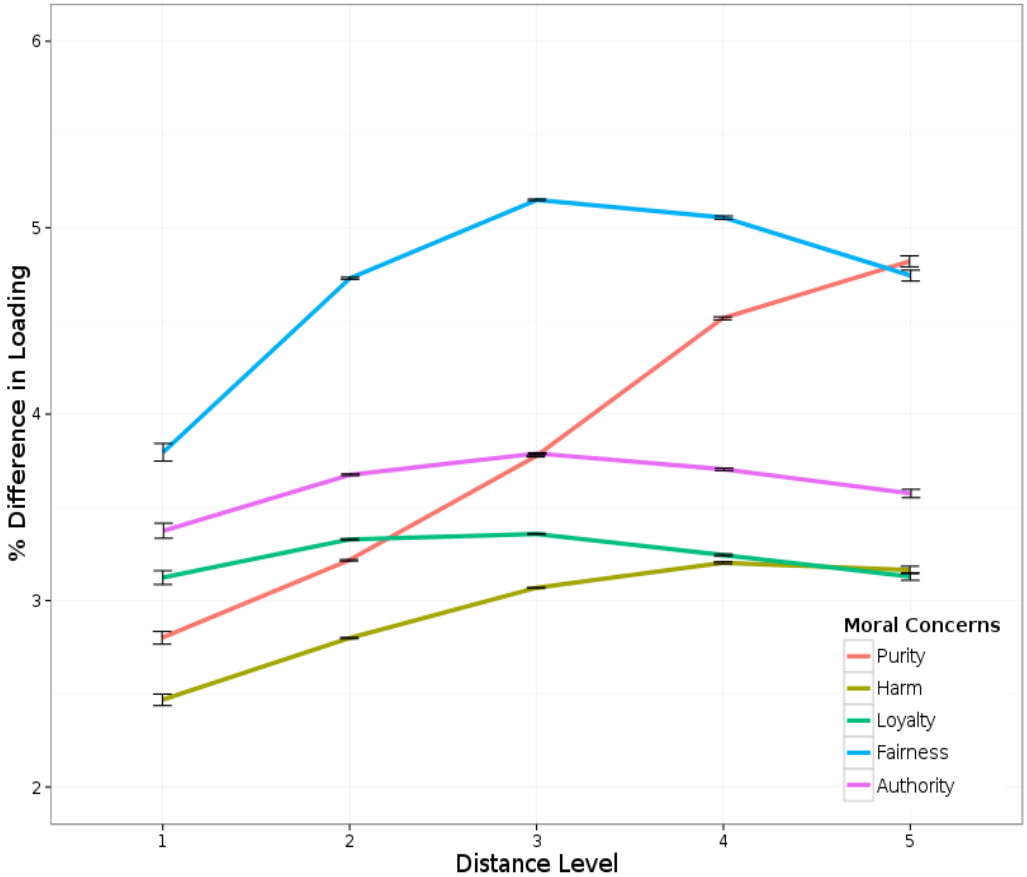

Figure 3. Change in Loading of Moral Concerns based on Distance in the Social Graph. A significant increase in difference in purity loading was observed with increase in distance. This difference increase at every distance level, and such linear increase is not observed with other moral concerns. Error bars represent 99% CI.

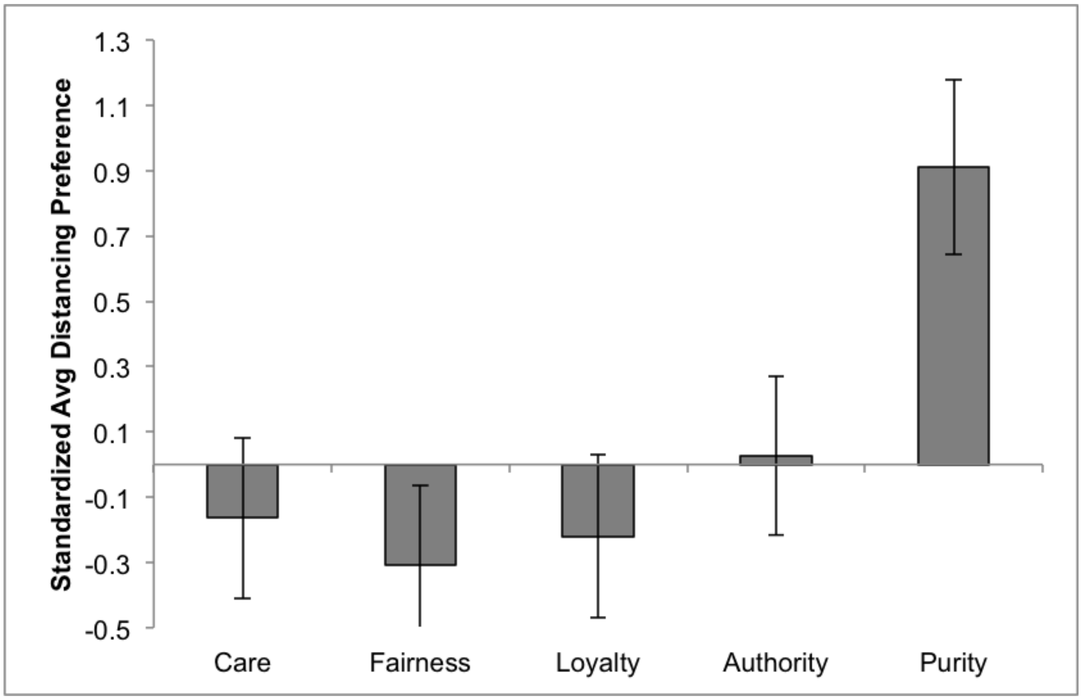

Figure 4. Study 2 results. Standardized average of social and physical distancing for the Moral Foundation feedback conditions. Error bars represent 95% CI.

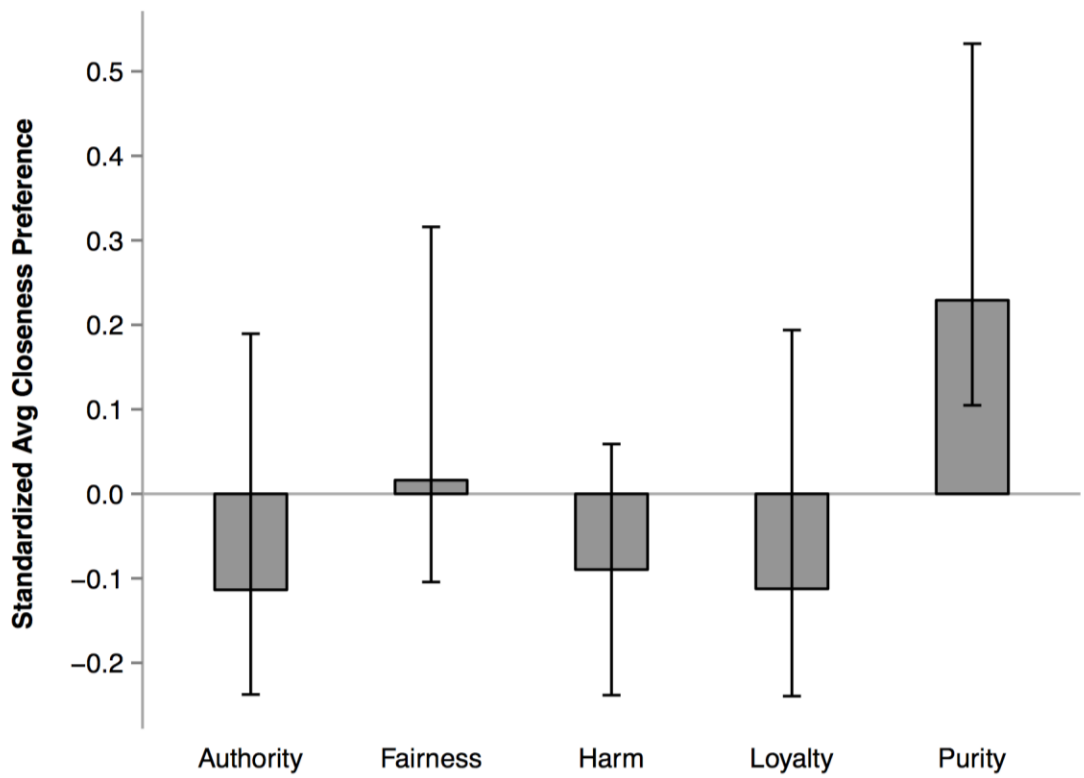

Figure 5. Study 3 results. Standardized average of social and physical closeness for the Moral Foundation feedback conditions. Error bars represent 95% CI.

Table 1. Tukey's HSD Comparisons of Purity Differences Between Distance Levels within Liberal and Conservative Clusters

User Set	Distance	$M_{Difference}$ (SE)	99% CI	*d*	*df*	*z*	Sig.
All Users	2 vs 1	0.0045 (0.0003)	[0.0036, 0.0053]	0.1774	9990232	16.95	< .001
	3 vs 2	0.0048 (0.00008)	[0.0046, 0.0051]	0.1702	9990232	57.02	< .001
	4 vs 3	0.0074 (.00009)	[0.0071, 0.0077]	0.2403	9990232	81.81	< .001
	5 vs 4	0.0023 (0.0002)	[0.0017, 0.0028]	0.0647	9990232	12.87	< .001
Liberals	2 vs 1	0.0030 (0.0003)	[0.0019, 0.0042]	0.1303	793533	8.355	< .001
	3 vs 2	0.0037 (0.0002)	[0.0030, 0.0044]	0.1534	793533	16.126	< .001
	4 vs 3	0.0031 (0.0006)	[0.0014, 0.0048]	0.1213	793533	5.429	< .001
Conservatives	2 vs 1	0.0053 (0.0004)	[0.0041, 0.0066]	0.2104	4687714	13.60	< .001
	3 vs 2	0.0053 (0.0001)	[0.0050, 0.0057]	0.1832	4687714	45.80	< .001
	4 vs 3	0.007 (0.0001)	[0.0069, .0.0077]	0.2306	4687714	59.83	< .001

Table 2. RMSE Comparisons for Bootstrapped SVM Regression Models for Moral Dimensions †

Model Test	$MRMSE_a$ †† (SD) [95% CI]	$MRMSE_{Purity}$ - $MRMSE_a$ ††† [95% CI]	*d*	*t*	Sig.
Harm	0.6677 (0.0053) [0.6663, 0.6691]	-0.0022 [-0.0008, -0.0038]	-0.2191	-3.0982	<.005
Fairness	0.6676 (0.0052) [0.6663, 0.6691]	-0.0023 [-0.0008, -0.0038]	-0.2193	-3.1014	<.005
Loyalty	0.6677 (0.0052) [0.6663, 0.6691]	-0.0023 [0.0008, -0.0038]	-0.2193	-3.069	< .005
Authority	0.6677 (0.0053) [0.6663, 0.6691]	-0.0023 [0.0008, -0.0038]	-0.2202	-3.1144	< .005
Purity	0.6654 (0.0052) [0.6640, 0.6668]	–	–	–	–

Note. †These estimates have been rounded from the 7th place to the 4th decimal place.
††Mean-Root-Mean-Squared-Errors (MRSME) were calculated by averaging across RMSEs from 100 k-fold-validated (*K*=10) models which were each estimated on randomly drawn *N*=10,000 samples.
†††Comparisons of model fit (Purity vs. all others) were conducted with independent *t* tests using the distribution of *N*=100 RMSEs obtained during the bootstrapping process.

Table 3. Full OLS Multiple Regression Model

Predictor	*b* (SE) [95% CI]	β	*t*	Sig.
Harm	0.95 (0.010) [0.93, 0.97]	0.03	94.33	<.001
Fairness	0.32 (0.006) [0.31, 0.34]	0.02	52.72	<.001
Loyalty	-1.28 (0.013) [-1.30, -1.25]	-0.05	-94.58	< .001
Authority	-0.06 (0.012) [-0.08, -0.03]	0.00	-4.63	< .001
Purity	3.00 (0.007) [2.98, 3.01]	0.13	405.57	< .001