Proceedings of the Fourteenth International AAAI Conference on Web and Social Media (ICWSM 2020)

When Does Trust in Online Social Groups Grow?

Shankar Iyer, Justin Cheng, Nick Brown, Xiuhua Wang
Facebook
{shankar94, jcheng, nabrown, xiuhuawangbeta}@fb.com

Abstract

The trust that people feel in their social groups is linked to important social outcomes such as member satisfaction and collective task performance. To understand the behaviors and conditions linked to trust, past studies of trust in groups have typically relied on cross-sectional surveys, but these are limited in their ability to identify causation. To better test the potential causal pathways between trust and behaviors or group properties, we paired a two-wave longitudinal survey of 2358 participants in Facebook Groups with logged activity on Facebook. Using latent change score modeling, we examined how trust may predict changes in behavior or group properties and how behaviors and group properties may predict changes in trust. On one hand, people who trust a group tend to contribute more written content to the group over time; and while groups that are more trusted tend to add more administrators and moderators over time, groups that have many administrators and moderators does not tend to be trusted more over time. On the other hand, people's trust in a group increases over time when the group is well-connected and active overall, while that trust decreases over time when that person is also actively involved in multiple other groups. These findings suggest a positive feedback loop related to active engagement and trust: seeing activity in a group drives trust, which in turn leads to increased individual activity and hence greater overall activity in the group. Overall, trust may be best promoted by encouraging both active engagement and friendship formation.

Introduction

Social exchange theory suggests that negotiated exchanges between two or more parties drive social change and stability (Cropanzano and Mitchell 2005). Feelings of trust mitigate the risks that the parties perceive when pooling their resources and skills and can enable greater levels of exchange, cooperation, and collective value generation (Ma 2019). Empirically, trust within social groups is associated with enhanced task performance, team satisfaction, and commitment to the group (Costa, Roe, and Taillieu 2001; Colquitt, Scott, and LePine 2007).

Online, social groups exist in places such as web forums, Reddit, and Facebook. As of May 2018, 1.4 billion people used the latter each month (Perez 2018). Though these services are popular, less is known about how trust in these online communities is built over time.

While substantial work on trust in (online) groups exists (Ma 2019), a majority is cross-sectional, making cause-and-effect relationships difficult to ascertain. Experimental and longitudinal studies exist (Cook et al. 2005; Glanville, Andersson, and Paxton 2013), but have typically focused on specific variables in offline settings. Applying a longitudinal approach to studying trust's relationship in online groups to a broader set of behaviors and properties may help disentangle the effects of the online instantiation of the group from confounding effects. For example, some social groups are formed entirely online, while others are built out of preexisting relationships in the offline world (e.g., local neighborhood groups). If (hypothetically) local groups are more trusted, some of the behavioral correlates of a cross-sectional trust measurement may simply be those that tend to occur in local groups, rather than those that are actually associated with growth in trust over time *as a result of the online platform*.

This work attempts to address this gap in the research literature through a longitudinal study of trust in Facebook Groups. Participants were surveyed about the same group twice, in two survey waves[1] separated by about two months. The survey results were joined to behavioral logs from the four weeks preceding each survey round, and latent change score models were used to address four distinct questions:

1. **Which behaviors or group properties are associated cross-sectionally with trust?** Active engagement (e.g., commenting and liking) and having a well-connected friendship-network structure within a group are positively correlated with trust. Meanwhile, engaging with many other groups is negatively correlated with trust.[2]
2. **Does trust predict changes in behaviors or group properties?** Having high trust in a group is associated with writing more text in a group over time and with in-

[1] Hereafter referred to as "Wave 1" and "Wave 2."

[2] We also observe several other cross-sectional associations of behaviors and group properties with trust that **do not** remain relevant in the longitudinal analyses of questions 2-4.

creases in the number of group administrators and moderators.

3. **Do behaviors or group properties predict changes in trust?** When a group is well-connected internally and active overall (e.g., a large proportion of members took an action in the group recently), its members tend to report trusting the group more over time. However, members who simultaneously engage with many other online groups are more likely to report decreases in trust over time.
4. **Which changes in behavior or group properties co-occur with changes in trust?** People who contribute more content (e.g., comments) over time also tend to report increases in trust over time.

Overall, these results have implications for the types of design interventions that are most likely to work and the situations in which they may be most effective. For instance, trust may be best fostered in online social groups by encouraging active engagement (particularly commenting), especially among less active members, and by increasing connection density (e.g., via Facebook friendships). By contrast, interventions that involve adding more administrators or moderators may be less effective at fostering trust.

Background

Defining and Measuring Trust

Social trust is "a belief in the honesty, integrity, and reliability of others" (Taylor, Funk, and Clark 2007) and can lower the perceived risks of pooling skills and resources (Gambetta 1988). Trust can further be divided into three "levels" (Ma 2019):

1. **Generalized trust**: the trust that people feel towards others generally.
2. **Dyadic trust**: the trust that one person feels towards a specific other person.
3. **Group trust**: the trust that one person feels towards a collective of other people (e.g., a religious community, an online social group, etc.).

Generalized trust is commonly viewed as an individual trait (Rotter 1971) and is traditionally measured through surveys (e.g., the World Values Survey). Prior work found that while generalized trust tends to increase with age and income and is higher in rural areas, several other individual-level attributes such as gender showed little association with generalized trust (Taylor, Funk, and Clark 2007).

Dyadic trust has often been measured through lab or web-based games where one individual has to place trust in another in a situation with some inherent risk. Recent research has demonstrated that reputation systems (e.g., reputation scores on Airbnb) can reinforce generalized trust and overcome the effects of homophily (Kuwabara 2015; Abrahao et al. 2017).

Group trust, our primary focus, has been studied in a variety of offline contexts. Prior work found that trust in social groups depends strongly on the size of the group (Brewer 1991), that people identify more with smaller groups (Simon and Brown 1987), and that more active groups tend to be trusted more (Cartwright and Zander 1953). In teams, trust is positively associated with successful completion of tasks, satisfaction of team members, and commitment to the team (Costa, Roe, and Taillieu 2001). Because generalized trust is likely associated with dyadic or group trust (Ma et al. 2019; Mayer, Davis, and Schoorman 1995), we control for it in our analyses, similar to previous work (Ma et al. 2019).

Trust in Online Communities

Questions about trust in offline settings have also been posed in online contexts, often focusing on virtual collaboration. Previous work has argued that trust is one of the most crucial social factors affecting the success of such collaborations (Handy 1995; Poole 1999). Case studies of geographically and culturally disperse collaborations found that while trust can emerge quickly, it is difficult to maintain (Jarvenpaa and Leidner 1999). Compared to offline groups, trust develops more slowly in computer-mediated groups as more time is needed for social information to be exchanged between participants (Wilson, Straus, and McEvily 2006), and long-term virtual collaboration without in-person interaction can lead to decreases in trust (Nandhakumar and Baskerville 2006).

Research has also examined the types of behaviors that can lead to trust and the types of behaviors that trust can induce. Imposition of rules upon a collaboration can help foster trust among the collaborators, nearly irrespective of what the rules actually are (Walther and Bunz 2005). Trust can also be promoted by crafting a clear identity for an online group, creating multiple opportunities for learning, having a credible and active moderator, and modeling appropriate behavior within the community (Booth 2012). Some work showed that in online health communities, trust measures were related to people's acceptance of advice (Fan and Lederman 2018). Trust is also positively associated with safety, both online (Zhang et al. 2010) and offline (Booth, Ayers, and Marsiglia 2012), and it is negatively associated with exposure to negative content online (Näsi et al. 2015). These studies point to the importance of understanding correlates of trust, mechanisms for promoting trust, and outcomes of the presence and growth of trust in online communities.

As noted previously, a large number of people worldwide now engage with online communities through Facebook Groups in particular, and a stream of recent research examines trust in these groups specifically. For instance, trust in mother-to-mother sale groups appeared to be fostered through exclusivity, a clear shared group identity, and active regulation of behavior by administrators and other group members (Moser, Resnick, and Schoenebeck 2017). Meanwhile, trust in buy-sell groups was positively associated with the density of the Facebook friendship connections within the group and the centrality of the seller in the friendship network (Holtz, MacLean, and Aral 2017). A general study of trust in groups found that generalized trust predicts trust in groups; that people trust groups more when those groups are smaller, more exclusive, and more homogeneous; that network structure and the truster's position within it predict trust; and that trust predicts downstream outcomes such as the formation of new friendship ties (Ma et al. 2019).

Longitudinal Survey Research

In cross-sectional studies, participants are surveyed on their feelings towards a group exactly once, and such studies can identify behaviors and group properties that remain associated with trust after controlling for various individual and group-level properties (Ma et al. 2019). However, it is difficult in such studies to disentangle the signatures of relationships that existed prior to the online instantiation of the group from the effects of that online instantiation. Longitudinal studies can help discriminate between these different contributors to trust and elucidate the causal drivers of trust.

To this end, previous studies have longitudinally examined specific predictors and outcomes of trust, mostly in offline settings. These studies have examined the relationship between measures of individual and dyadic trust and income (income increases trust but not vice versa) (Brandt, Wetherell, and Henry 2015; Bilson, Jetter, and Kristoffersen 2017), tie formation (social ties increase trust) (Glanville, Andersson, and Paxton 2013), justice (adhering to justice rules increases trustworthiness and vice versa) (Colquitt and Rodell 2011), and purchase intent (Kim, Ferrin, and Rao 2009). They have also examined the role of trust in collaborative settings (Webber 2008): trust in another team predicts future risk-taking with respect to that team (Serva, Fuller, and Mayer 2005), and trust also buffers against future conflict (Peterson and Behfar 2003).

Compared to prior research, which has tended to be smaller-scale, cross-sectional, or focused on offline behavior, the present work is a larger, longitudinal analysis of a broad diversity of online social groups. The analysis approach not only accounts for potential confounds such as generalized trust and demographics, but also identifies important differences between behaviors that are cross-sectionally correlated with trust and those that may precede or follow trust.

Method

To better understand the relationship between changes in group trust and changes in behavior, responses from a two-wave survey on trust were combined with demographic and activity logs in the four weeks prior to each wave of the survey. Latent change score models were used to identify significant effects. Survey participation was voluntary, and an internal research board reviewed the study design ethics and privacy practices prior to its start. No experiment was performed. Other than the survey, participants' experiences of Facebook were unchanged. All data was deidentified and analyzed in aggregate, and no researchers viewed any individual-level data.

Sampling Strategy

The sampling strategy for Wave 1 of our survey (which began on May 14, 2019) was as follows: we identified people belonging to Facebook Groups who were at least 18 years of age and who were from Brazil, Egypt, Indonesia, India, UK, Mexico, Thailand, US, the Philippines, or Vietnam. We additionally required that potential participants had spent at least five minutes in the group or took at least one action in the group during the two week period from April 29 to May 12, 2019. We then deduplicated these person-group pairs, first at the person level (choosing one group per person randomly) and then at the group level (again randomly, but excluding group administrators and moderators). Next, we filtered out pairs in which the person had been in the group for fewer than 29 days, which allowed us to consistently join to at least 28 days of logs. The end result was a set of pairs, deduplicated at both the person and group levels, which we then randomly limited to 750,000 per country. The sample for Wave 2 (which began on July 9, 2019) consisted of Wave 1 respondents who consented to be contacted again.

Survey Design

Our survey (Table 1) is modeled on one recently used in the research literature (Ma et al. 2019), which itself is based on prior surveys of generalized and interpersonal trust (World Values Survey 2018; Rotter 1971; Johnson-George and Swap 1982; Ashleigh, Higgs, and Dulewicz 2012). To understand trust's relationship to safety (Zhang et al. 2010), we incorporated questions on general safety and safety in groups.[3] All questions were randomized within blocks.

The "disposition to trust" question is based on the generalized trust question from the World Values Survey. Of the three other General Trust Attributes questions in Ma et al., the "general risk attitude" question was least correlated with the "disposition to trust" question, and so was retained here. A "trust in government" question was included as trust in public institutions may influence one's civic engagement (Keele 2007), and thus participation in some types of groups. The questions on Trust in Groups are a subset of those used in Ma et al.[4] The "safety in neighborhood" question is commonly used to measure perceived safety (Hinkle 2015), and the same question stem was used to measure safety on the internet and safety in groups. The "negative interactions" question is adapted from a survey used to measure online harm (Livingstone et al. 2010).

Major life events can have significant effects on subjective well-being (McCullough, Huebner, and Laughlin 2000), so we asked respondents if they experienced any major life events in the time period before each survey round.

Group Properties and Behavioral Metrics

Survey responses were joined with behavioral logs and properties of the person or group. Counts of behaviors were computed for a four-week period shortly preceding the launch of the survey round. For example, for Wave 1, the behaviors were measured for the four weeks leading up to and including May 12, 2019. Meanwhile, properties of the person or group (e.g., the group size) were measured two days before the launch of the survey wave.

Three classes of features were examined:

Person-Level Features

[3] As we later show, the survey results indicate that trust and safety are sufficiently different concepts that should be analyzed separately, and we focus on the former for space.

[4] One question was omitted to reduce the length of the survey.

General Trust Attitudes	
Disposition to trust	Most people can be trusted.
Trust in government	Most of the time we can trust people in government to do what is right.
General risk attitude	I'm willing to take risks.
General Safety Attitudes	
Safety in neighborhood	How safe do you feel in your neighborhood?
Safety on the internet	How safe do you feel when you are using the internet?
Trust in Groups	
Care	Other members of the group care about my well-being.
Reliability	Other members of this group can be relied upon to do what they say they will do.
Integrity	Other members of this group are honest.
Other Group Measures	
Safety in groups	How safe do you feel in this group?
Negative interactions	How often do interactions with other people in this group make you feel uncomfortable or upset?
Major Life Events	
Positive events	In the past month, have you experienced a positive life event such as starting a new job, going on vacation, or making a new friend?
Negative events	In the past month, have you experienced a negative life event, such as losing a job, or the death of a close friend or family member?

Table 1: Longitudinal Trust and Safety in Groups Survey.

- **Basics**: Basic demographic features (e.g., age, gender, etc.) and overall engagement metrics across Facebook.
- **Friendships**: Properties of the person's friendship graph, including its absolute size and diversity over demographic characteristics (e.g., the entropy of the distribution of countries that a person's friends are from).
- **Time Spent Across All Groups**: Measurements of how much time the person spent across all groups.[5]
- **Actions Across All Groups**: Measurements of how many actions (e.g., Likes given) the person performed in all groups.

[5]Metrics identified as "across all groups" include the group about which the respondent was surveyed.

Group-Level Features

- **Basics**: Basic properties of the group, such as size, privacy type[6], average site-wide engagement of members, etc.
- **Friendships**: Properties of the friendship network within the group, including various friendship counts but also metrics related to k-core[7] decomposition of the network.
- **Actions**: The total number of actions and action-takers in the group (e.g., the number of commenters).
- **Membership**: Properties of the group members, including diversity metrics and metrics related to when people joined the group (e.g., the entropy of the distribution of countries that members are from).
- **Growth**: Metrics about how many members the group has gained or lost.
- **Group Administration**: Metrics about admins and moderators and the actions they take in the group.
- **Time Spent Across All Members**: The amount of time all people spent in the group, split into various channels (e.g., via News Feed or the Group page itself).

Person-Group-Level Features

- **Membership**: Metrics about when the person joined the group (e.g., through which part of the Facebook product).
- **Friendships**: Metrics around how many friends the person has in the group.
- **Actions**: The number of actions the person took in the group (e.g., comments written).
- **Time Spent**: The amount of time the person spent in the group, divided over various channels.

Latent Change Score Models (LCSMs)

Latent change score models (McArdle and Hamagami 2001) are a class of structural equation model that can be used to simultaneously model multiple hypothesized relationships between two or more variables over time. They have been used to measure changes in brain and behavior (Kievit et al. 2018) and to quantify the relationship between loneliness and social engagement (McHugh Power et al. 2019).

This approach allows us to simultaneously test multiple causal relationships between two variables X and Y across two time periods while accounting for individual differences. A simplified path diagram is shown in Figure 1, where X is a latent variable (e.g., trust in a group) and Y is an observed variable (e.g., time spent in the group). X_{T_1} and X_{T_2} correspond to measurements of X in the first and second waves of the survey respectively. X_{1,T_1}, X_{2,T_1}, X_{3,T_1} represent the survey responses for each of the three questions on Trust in Groups. The rationale for treating these questions as

[6]Facebook Groups can be open (anyone can view the group's content but only members can post), closed (anyone can see that the group exists but only members can see the content), or secret (the group is not searchable and can can only be found by people who are invited by current members).

[7]The k-core of a network is what remains when nodes with degree less than k are repeatedly removed from the network.

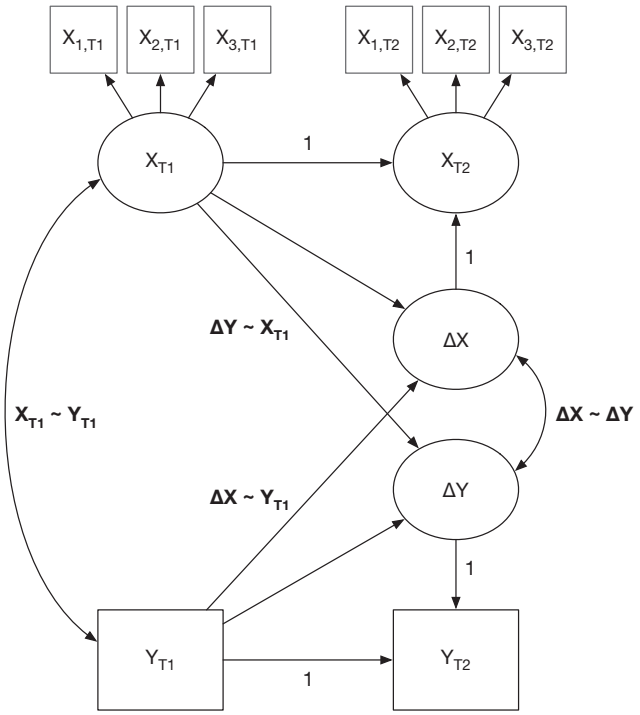

Figure 1: Simplified path diagram of the latent change score model used in our analysis. Lines indicate directional relationships between variables. Variables in ovals are measured indirectly, while variables in boxes are measured directly. X_{T_i} is the latent trust variable measured in Wave 1 (X_{T_1}) or Wave 2 (X_{T_2}), while Y_{T_i} is a measured behavioral metric or group property. Trust in Wave 1 (X_{T_1}) is dependent on each of the three survey measures of trust (X_{j,T_i}, $j \in \{1, 2, 3\}$). Trust in Wave 2 (X_{T_2}) is modeled exactly as $X_{T_1} + \Delta X$. Control variables (which are assumed to affect X_{T_1}, Y_{T_1}, ΔX, and ΔY) and self-loops have been omitted for clarity.

a single factor is detailed in the following section. Similarly, Y_{T_1} and Y_{T_2} correspond to measurements of Y at the time of the first and second waves of the survey respectively. ΔX and ΔY represent the changes in X and Y from T_1 to T_2.

The model allows us to measure four relationships that correspond to the four questions asked in the introduction:

- The cross-sectional relationship between X and Y ($X_{T1} \sim Y_{T1}$) (Q1)
- The change in Y as a function of the initial value of X ($\Delta Y \sim X_{T1}$) (Q2)
- The change in X as a function of the initial value of Y ($\Delta X \sim Y_{T1}$) (Q3)
- The relationship between the change in X and change in Y ($\Delta X \sim \Delta Y$) (Q4)

Survey Responses

Responses & Response Biases

10090 people completed Wave 1 of our survey, and 6736 consented to being contacted again. Of those, 2358 completed Wave 2. Figure 2 presents distributions of demographic, group, engagement and survey variables over the original sample and the sets of respondents. We observe the following response biases (association strength is reported using phi coefficients and Cohen's d):

- **Gender**: Wave 1 respondants were more likely to be female compared to the original sample ($\phi = 1.5 \times 10^{-3}$, $p = 2.2 \times 10^{-4}$). Wave 2 respondants were somewhat more likely to be female compared to Wave 1 respondants ($\phi = 0.016$, $p = 0.081$).
- **Age**: Wave 1 respondants were more likely to be 45 and older ($\phi = 2.9 \times 10^{-3}$, $p = 8.1 \times 10^{-13}$) compared to the original sample. Wave 2 respondants were also more likely to be 45 and older compared to Wave 1 respondants ($\phi = 0.027$, $p = 2.5 \times 10^{-3}$).
- **Country**: Wave 1 respondants were more likely to be from Great Britain, Mexico, or the United States compared to the original sample ($\phi = 5.8 \times 10^{-3}$, $p < 10^{-15}$). Similarly, Wave 2 respondants were more likely to be from those countries compared to Wave 1 respondants ($\phi = 0.063$, $p = 2.0 \times 10^{-12}$).
- **Time spent**: Wave 1 respondants spent more time in the group they were asked about, compared to the original sample ($d = 0.050$, $p = 1.7 \times 10^{-4}$). However, Wave 2 respondants spent a similar amount of time compared to Wave 1 respondants ($d = 0.034$, $p = 0.19$).
- **Group size**: Wave 1 respondants came from smaller groups compared to the sample ($d = 0.062$, $p < 10^{-15}$), but Wave 2 respondants came from groups of similar size as Wave 1 respondants ($d = 0.024$, $p = 0.31$).

To identify the metrics that are most robustly associated with changes in trust, we ran three sets of analyses:

1. LCSM analyses with X representing trust and Y representing each of the previously-listed features (without reweighting).
2. LCSM analyses where the Wave 2 respondent population was reweighted by age, gender, and time spent in the group to match the population in our initial sample to minimize response bias (Little 1993).
3. LCSM analyses without reweighting, but split on engagement level (i.e., the median time spent in a group). We performed a median split of the population by monthly time spent in the group before Wave 1 and constructed LCSMs separately for those with engagement less than and greater than the median value.

All three sets of analyses control for age, gender, country, major life events, group size, and generalized trust. In the "Results" section below, we report associations that are robust to the reweighting (i.e., they appear as statistically significant associations in analysis sets 1 *and* 2 above). Coefficients reported come from analysis set 2. Important differences between the engagement levels in analysis set 3 are noted in the text.

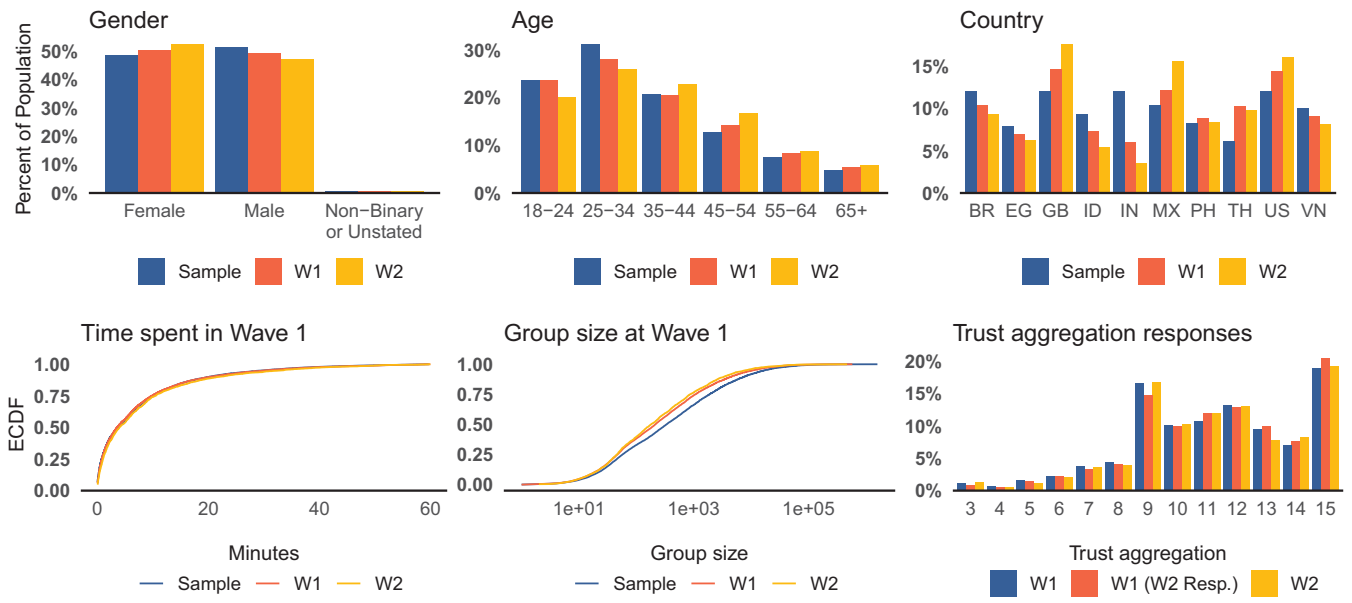

Figure 2: Plots summarizing response biases in our survey responses. The bottom-right plot shows the distribution of trust aggregation scores over the 10090 Wave 1 responses, the 2358 Wave 1 responses from people who also responded to Wave 2, and the Wave 2 responses. The remaining plots show either distributions or cumulative distributions of demographic, behavioral, or group properties over the initial sample, the 10090 Wave 1 responses, and the 2358 Wave 2 responses.

Relationships Between Survey Responses

Figure 3a shows the relationships between the 10090 Wave 1 responses to the Trust in Groups questions. The three responses are tightly correlated with one another (Cronbach's $\alpha = 0.82$). The strong agreement among these three questions provides support for treating them as measuring a single latent concept. For simplicity, this section sums these responses into a single "trust aggregation score". The "Trust aggregation responses" panel of Figure 2 shows the distributions of this aggregate score over Wave 1 responses, Wave 2 responses, and Wave 1 responses of Wave 2 respondents.

Figure 3b shows that the correlations between the Trust in Groups and General Trust Attitudes questions are substantially weaker than the correlations between the Trust in Groups questions. This is consistent with prior research: although Generalized Trust predicts some of the variance in trust in a particular group, it also leaves the majority of the variance unexplained (Ma et al. 2019).

Figure 3c shows correlations between the Wave 1 Trust and Safety in Groups responses. The correlations between the Trust in Groups questions and the "negative interactions" question are particularly weak, but there is a strong positive association between the responses to the "safety in groups" question and the Trust in Groups questions. We performed an exploratory factor analysis of the trust and safety questions (including the General Trust Attitudes and General Safety Attitudes questions from Table 1) to determine the factor structure. A parallel analysis and a scree plot both suggest a four-factor solution, roughly corresponding a "generalized safety" construct ("safety in neighborhood", "safety on internet"), a "generalized trust" construct ("disposition to trust", "trust in government"), a "trust and safety in groups" construct ("care", "reliability", "integrity", "safety in groups"), and a "negative interactions" construct. Despite this, we measure trust only using the three Trust in Groups questions for three reasons:

1. This approach retains more comparability to the rest of the literature on trust.

2. The "safety in groups" question loads more weakly ($r = 0.48$) onto the "trust and safety in groups" construct than the three Trust in Groups questions (≥ 0.72).

3. The relationship between trust and safety at the low end of both scales is ambiguous. For example, people who report low levels of trust in a group still report feeling safe in the group surprisingly often.

We exclude the "safety in groups" and "negative interactions" questions from further analysis in this paper for space, but future work may involve deeper analyses of the relationship between trust, safety, and negative experiences.

Figure 3d examines correlations between the Trust in Groups questions across the two waves in terms of the three individual questions. Meanwhile, Figure 4 shows conditional distributions of the Wave 2 trust aggregation score after conditioning on the Wave 1 trust aggregation score. While the two waves are highly correlated ($r = 0.60$), there is variance between them, especially at the lower end of the scale where there may be an opportunity for the trust aggregation score to rise by Wave 2.

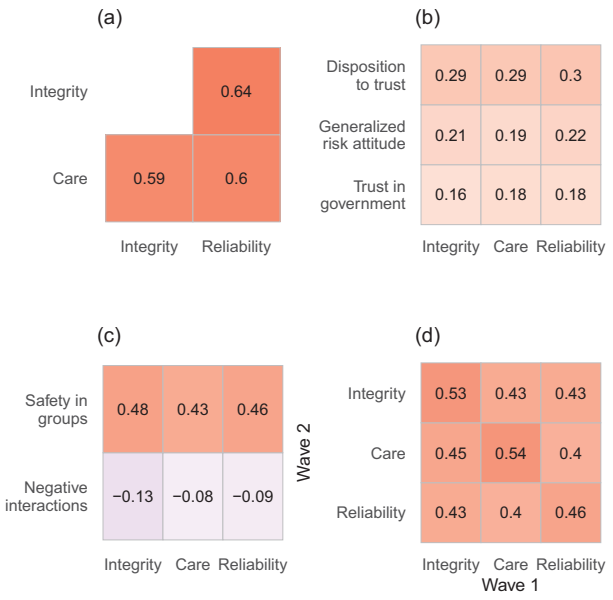

Figure 3: Relationships between responses to the survey questions, within Wave 1 and between Wave 1 and Wave 2. (a), (b), and (c) show Pearson correlations between responses to questions within Wave 1. These heatmaps are based on the 10090 Wave 1 responses. (d) shows Pearson correlations between the Wave 1 and Wave 2 responses to the Trust in Groups questions. This heatmap is based on the 2358 people who responded to both waves.

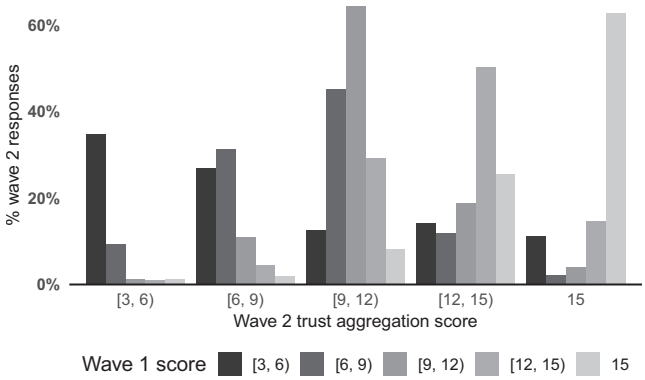

Figure 4: Distribution of the Wave 2 trust aggregation score, conditioned on the Wave 1 score.

Latent Change Score Model Analysis

We now address the four questions that we articulated in the Introduction. As the analysis requires complete data at both time points, subsequent findings are based on the subset of 2358 participants who completed both survey waves. Regression coefficients (β) and p-values are reported for each considered feature. All p-values reported in this section are false-discovery-rate adjusted to account for multiple comparisons. Figure 5 summarizes our findings.

Which behaviors or group properties are associated cross-sectionally with trust? (Q1)

On one hand, many behaviors and group properties have positive cross-sectional associations with trust:

- the respondent's activity in the group (e.g., time spent in the group with $\beta = 0.083$, $p = 3.5 \times 10^{-5}$, monthly likes with $\beta = 0.10$, $p = 8.0 \times 10^{-8}$, or monthly comments with $\beta = 0.086$, $p = 4.5 \times 10^{-6}$)
- the respondent's connections in the group (e.g., number of friends in the group with $\beta = 0.091$, $p = 3.1 \times 10^{-6}$)
- group-level activity (e.g., the number of monthly likes on the group's content with $\beta = 0.074$, $p = 3.0 \times 10^{-6}$ or the number of monthly comments in the group with $\beta = 0.064$, $p = 1.9 \times 10^{-4}$)
- global properties of the group's friendship network (e.g., graph density[8] with $\beta = 0.068$, $p = 4.7 \times 10^{-7}$ or the number of friendships where the friends are from the same country with $\beta = 0.036$, $p = 9.0 \times 10^{-4}$)

On the other hand, few behaviors and properties have negative cross-sectional associations with trust. The number of groups in which the respondent posted was negatively associated with trust ($\beta = -0.055$, $p = 0.010$), and there was some evidence of a negative association between country homophily (i.e., the Shannon entropy of the country distribution of members) and trust ($\beta = -0.038$, $p = 0.052$).

Though many of these observations (e.g., on time spent, likes and comments made, and friendship network density) corroborate past findings (Ma et al. 2019), two points are worth keeping in mind when proceeding from these cross-sectional results at Wave 1 to the longitudinal analyses that follow. First, far fewer behaviors and group properties show any significant associations in the longitudinal analyses. Second and more specifically, country homophily is not relevant in those longitudinal analyses. This suggests that, despite what cross-sectional correlations suggest, trust may not necessarily grow if groups are made more internationally homogeneous.

Does trust predict changes in behaviors or group properties? (Q2)

If trust is higher at Wave 1, the respondent is more likely to contribute more text in the group via posts and comments by Wave 2 ($\beta = 0.082$, $p = 5.3 \times 10^{-3}$), and the percentage of group members that are admins or moderators is more likely to increase ($\beta = 0.014$, $p = 5.5 \times 10^{-4}$)[9].

When we separate the analysis by Wave 1 engagement level, the finding for text contribution holds only among people who spent more than the median amount of time in a group. In other words, trust may only lead to greater downstream activity if the person is already fairly engaged with the group in the first place.

[8]The fraction of all possible friendships between group members that do exist before Wave 1.

[9]Conversely, the proportion of admins or moderators at Wave 1 is not associated with changes in trust, suggesting that the presence of admins and moderators may be a lagging indicator of trust.

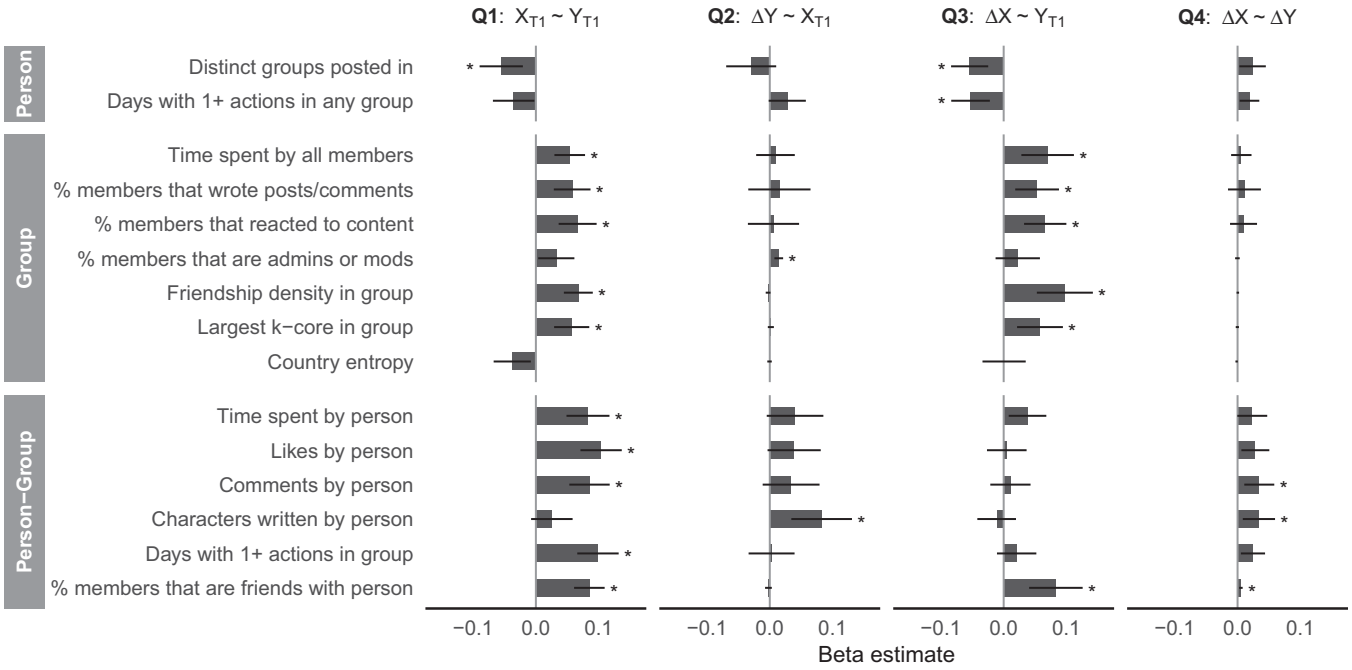

Figure 5: Associations between trust and behavior from the LCSM analysis. For example, high levels of trust in wave 1 are associated with positive changes in the number of characters written between wave 1 and wave 2. Stars ($\star$) indicate significance at $p < 0.05$ or better; error bars indicate 95% confidence intervals.

Do behaviors or group properties predict changes in trust? (Q3)

Several group-level features are associated with growth in trust over time. These include the total monthly time spent by all people in the group ($\beta = 0.070$, $p = 6.7 \times 10^{-3}$), the fraction of group members who reacted to content ($\beta = 0.066$, $p = 1.0 \times 10^{-3}$), and the fraction of group members who contributed content ($\beta = 0.053$, $p = 0.016$). In contrast, metrics relating to the respondent's activity across all groups, such as the number of groups that the respondent posted in ($\beta = -0.054$, $p = 2.5 \times 10^{-3}$) or the number of days on which the respondent person liked, commented, or posted in any group ($\beta = -0.052$, $p = 5.3 \times 10^{-3}$) in the past four weeks, are associated with declines in trust.

Properties of the friendship network structure, including the friendship density of the group ($\beta = 0.098$, $p = 2.1 \times 10^{-4}$), the size of the maximal k-core in the group's friendship network ($\beta = 0.058$, $p = 0.011$), and the percent of the group's members that are friends with the respondent ($\beta = 0.083$, $p = 1.1 \times 10^{-3}$), are also associated with growth in trust.

Splitting the analysis by Wave 1 engagement level, many of the group-level activity metrics are no longer statistically significantly associated with changes in trust, indicating that these may simply be correlates of person-group level engagement. Nevertheless, the overall friendship density in the group remains a positive correlate for both engagement cohorts, the number of friends that the respondent has in the group remains a positive correlate for the more-engaged cohort, and engagement with many groups remains a negative correlate for the less-engaged cohort.

Which changes in behavior or group properties co-occur with changes in trust? (Q4)

If the respondent writes more comments ($\beta = 0.035$, $p = 0.023$) or contributes more text via posts and comments ($\beta = 0.034$, $p = 0.043$) wave-over-wave, that is associated with growth in trust wave-over-wave. There was also some evidence that actively engaging with the group (via liking, commenting, or posting) on more days in the four weeks preceding Wave 2 than in the four weeks preceding Wave 1 is associated with growth in trust ($\beta = 0.025$, $p = 0.052$).[10] An increase in the percentage of group members who are friends with the respondent is also associated with an increase in trust, but this association is substantially weaker ($\beta = 5.3 \times 10^{-3}$, $p = 1.7 \times 10^{-3}$). The existence of these correlated changes suggests that other underlying variables may also be driving changes in both trust and these variables.

In the engagement-split analysis, wave-over-wave changes in monthly comments and the percentage of group members who are friends with the respondent remain positively associated with wave-over-wave changes in trust for the highly engaged cohort. Notably, for the lesser engaged cohort, growth in trust is positively associated with growth in the number of distinct groups in the respondent posts in.

[10]These findings may not be independent of one another. For example, an increase in characters added may reflect an increase in comments written.

Discussion

We now summarize our results and make recommendations for fostering trust in online communities.

Implications

Successful communities rely on their members to contribute resources (e.g., content) to ensure their continued existence (Kraut and Resnick 2012). Our findings suggest that one primary benefit of trust is in promoting more substantial contribution: higher levels of trust are associated more strongly with growth over time in writing (Q2).

But how is this trust created? We find two primary sets of properties are associated with increased trust over time: measures relating to connectedness (e.g., overall friendship density), and measures relating to overall group activity (e.g., the fraction of members that contributed or reacted to content) (Q3). In other words, in groups where members are friends with each other or where a large fraction of members are active, trust is likely to grow over time.

The above findings suggest a virtuous cycle in which active engagement drives trust and vice versa: seeing multiple others contribute can increase a person's trust in a group, which in turn leads to that person contributing more content that other group members may see, which then leads to them also trusting the group more.

At the same time, the number of groups that a person engages may influence trust negatively. In general, engaging with many groups is cross-sectionally associated with lower levels of trust in a particular group and is associated with decreases in trust in that particular group over time. Still, this does not mean that people should be discouraged from participating in multiple groups. In particular, this effect was stronger for people who spent less time in the group – they may have been exploring multiple groups to figure out the ones that they enjoy participating in more, and over time gravitated to one of their other groups.

Our findings also suggest types of interventions that may be less successful at promoting trust. For instance, trust is associated with an increase in the number of administrators or moderators in a group, but having a greater number of administrators or moderators in a group is not associated with any increases in trust. As such, interventions that increase the number of administrators or moderators in a group may not necessarily lead to group members trusting the group more. As another example, diversity in the international composition of groups appears negatively associated with trust cross-sectionally, but low homogeneity did not lead to changes in trust, suggesting that making groups more internationally homogeneous is unlikely to increase trust. Thus, the cross-sectional association may simply reflect relationships that preexist the online instantiation of groups (e.g., that local groups tend to have higher levels of trust).

In aggregate, the findings suggest that trust in online social groups can best be promoted by encouraging active engagement, especially among less active group members, as well as increasing within-group connectedness.

Limitations

LCSMs allow us to simultaneously test multiple hypotheses and control for many confounds. Though the analysis was longitudinal, as with any observational analysis, other confounds may exist. While we are able to show that changes in commenting remain associated with changes in trust after controlling for factors including demographics, we cannot make a completely conclusive causal statement. The relationship between trust and behavior in online groups may still vary with other unmeasured variables (e.g., changes in offline interactions) or demographics (e.g., non-binary gender identities). And as each variable of interest was studied in a separate model, understanding the interaction effects between them (e.g., between time spent and other variables, or between text contributed and comments made) remains future work.

The study included survey data from ten different countries, representing very different cultural contexts. The findings remained empirically similar after dropping the United States and United Kingdom from the analysis, suggesting that they are relatively robust. Still, future quantitative and qualitative analyses could better understand the nuances of how trust in groups varies across the world.

The analysis was conducted on Facebook Groups because of their relative popularity, but these findings may not generalize to other social media platforms, such as those where users are anonymous, as anonymity can affect both trust and participation (Friedman, Khan Jr, and Howe 2000).

Finally, trust is just one of the pro-social outcomes that developers may want to optimize for, and our recommendations should be balanced against other considerations where appropriate. For example, increasing the number of administrators and moderators in a group may be less effective at fostering trust, but could still promote perceptions of greater safety and reduce negative experiences. Relatedly, while we observed an association between trust and safety, understanding the differences in how behaviors drive trust and safety remains future work.

Conclusion

As predicted by social exchange theory, trust is a crucial contributor to the efficient functioning of social groups. A large number of people engage with social groups online through a variety of web products. Developers of these products may want to make design decisions that help foster trust over time. In previous work, trust and its behavioral correlates have been identified through cross-sectional surveys. However, correlates identified through such an approach may reflect the group type rather than the effects of the online instantiation of the group. The present work attempts to fill this gap and identify the behaviors and group properties that are associated with growth of trust over time.

The longitudinal study reported here allows us to identify both the potential causal drivers and outcomes of trust in online social groups and to identify substantial differences between the results of cross-sectional and longitudinal analyses. We find that many behaviors that show cross-sectional associations with trust are actually not relevant in the lon-

gitudinal analysis. In sum, we find that the most promising approach for helping participants in online communities to build trust in those communities is to both promoting certain forms of active engagement (especially commenting) and increase connectedness in these communities. In particular, the former may lead to a virtuous cycle, where active engagement grows trust and vice versa. We hope that this recommendation will help developers build better online communities, but also that the longitudinal aspect of this research will motivate further studies that explore the gap between cross-sectional, longitudinal, and ultimately causal approaches.

References

Abrahao, B.; Parigi, P.; Gupta, A.; and Cook, K. S. 2017. Reputation offsets trust judgments based on social biases among airbnb users. *Proceedings of the National Academy of Sciences* 114(37):9848–9853.

Ashleigh, M. J.; Higgs, M.; and Dulewicz, V. 2012. A new propensity to trust scale and its relationship with individual well-being: implications for hrm policies and practices. *Human Resource Management Journal* 22(4):360–376.

Bilson, J.; Jetter, M.; and Kristoffersen, I. 2017. Gender differences in the link between income and trust levels: Evidence from longitudinal data.

Booth, J.; Ayers, S. L.; and Marsiglia, F. F. 2012. Perceived neighborhood safety and psychological distress: Exploring protective factors. *J. Soc. & Soc. Welfare* 39:137.

Booth, S. E. 2012. Cultivating knowledge sharing and trust in online communities for educators. *Journal of Educational Computing Research* 47(1):1–31.

Brandt, M. J.; Wetherell, G.; and Henry, P. 2015. Changes in income predict change in social trust: A longitudinal analysis. *Political Psychology* 36(6):761–768.

Brewer, M. B. 1991. The social self: On being the same and different at the same time. *Personality and Social psychology Bulletin* 17(5):475–482.

Cartwright, D., and Zander, A. 1953. Group cohesiveness: introduction. *Group Dynamics: Research and Theory. Evanston, IL: Row Peterson*.

Colquitt, J. A., and Rodell, J. B. 2011. Justice, trust, and trustworthiness: A longitudinal analysis integrating three theoretical perspectives. *Academy of Management Journal* 54(6):1183–1206.

Colquitt, J. A.; Scott, B. A.; and LePine, J. A. 2007. Trust, trustworthiness, and trust propensity: a meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology* 92(4):909.

Cook, K. S.; Yamagishi, T.; Cheshire, C.; Cooper, R.; Matsuda, M.; and Mashima, R. 2005. Trust building via risk taking: A cross-societal experiment. *Social Psychology Quarterly* 68(2):121–142.

Costa, A. C.; Roe, R. A.; and Taillieu, T. 2001. Trust within teams: The relation with performance effectiveness. *European journal of work and organizational psychology* 10(3):225–244.

Cropanzano, R., and Mitchell, M. S. 2005. Social exchange theory: An interdisciplinary review. *Journal of management* 31(6):874–900.

Fan, H., and Lederman, R. 2018. Online health communities: how do community members build the trust required to adopt information and form close relationships? *European Journal of information systems* 27(1):62–89.

Friedman, B.; Khan Jr, P. H.; and Howe, D. C. 2000. Trust online. *Communications of the ACM* 43(12):34–40.

Gambetta, D. 1988. *Trust: Making and breaking cooperative relations*. Blackwell Publishing Limited.

Glanville, J. L.; Andersson, M. A.; and Paxton, P. 2013. Do social connections create trust? an examination using new longitudinal data. *Social Forces* 92(2):545–562.

Handy, C. 1995. Trust and the virtual organization. *Long Range Planning* 28(4):126–126.

Hinkle, J. C. 2015. Emotional fear of crime vs. perceived safety and risk: Implications for measuring "fear" and testing the broken windows thesis. *American Journal of Criminal Justice* 40(1):147–168.

Holtz, D.; MacLean, D. L.; and Aral, S. 2017. Social structure and trust in massive digital markets. In *Proceedings of the International Conference on Information Systems (ICIS)*.

Jarvenpaa, S. L., and Leidner, D. E. 1999. Communication and trust in global virtual teams. *Organization Science* 10(6):791–815.

Johnson-George, C., and Swap, W. C. 1982. Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology* 43(6):1306.

Keele, L. 2007. Social capital and the dynamics of trust in government. *American Journal of Political Science* 51(2):241–254.

Kievit, R. A.; Brandmaier, A. M.; Ziegler, G.; Van Harmelen, A.-L.; de Mooij, S. M.; Moutoussis, M.; Goodyer, I. M.; Bullmore, E.; Jones, P. B.; Fonagy, P.; et al. 2018. Developmental cognitive neuroscience using latent change score models: A tutorial and applications. *Developmental cognitive neuroscience* 33:99–117.

Kim, D. J.; Ferrin, D. L.; and Rao, H. R. 2009. Trust and satisfaction, two stepping stones for successful e-commerce relationships: A longitudinal exploration. *Information systems research* 20(2):237–257.

Kraut, R. E., and Resnick, P. 2012. *Building successful online communities: Evidence-based social design*. Mit Press.

Kuwabara, K. 2015. Do reputation systems undermine trust? divergent effects of enforcement type on generalized trust and trustworthiness. *American Journal of Sociology* 120(5):1390–1428.

Little, R. J. 1993. Post-stratification: a modeler's perspective. *Journal of the American Statistical Association* 88(423):1001–1012.

Livingstone, S.; Haddon, L.; Görzig, A.; and Ólafsson, K. 2010. Risks and safety for children on the internet: the uk report. *Politics* 6(1).

Ma, X.; Cheng, J.; Iyer, S.; and Naaman, M. 2019. When do people trust their social groups? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 67. ACM.

Ma, X. 2019. *Networked Trust: Computational Understanding of Interpersonal Trust Online*. Ph.D. Dissertation, Cornell University.

Mayer, R. C.; Davis, J. H.; and Schoorman, F. D. 1995. An integrative model of organizational trust. *Academy of Management Review* 20(3):709–734.

McArdle, J. J., and Hamagami, F. 2001. Latent difference score structural models for linear dynamic analyses with incomplete longitudinal data.

McCullough, G.; Huebner, E. S.; and Laughlin, J. E. 2000. Life events, self-concept, and adolescents' positive subjective well-being. *Psychology in the Schools* 37(3):281–290.

McHugh Power, J. E.; Steptoe, A.; Kee, F.; and Lawlor, B. A. 2019. Loneliness and social engagement in older adults: A bivariate dual change score analysis. *Psychology and aging* 34(1):152.

Moser, C.; Resnick, P.; and Schoenebeck, S. 2017. Community commerce: Facilitating trust in mom-to-mom sale groups on facebook. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 4344–4357. ACM.

Nandhakumar, J., and Baskerville, R. 2006. Durability of online teamworking: patterns of trust. *Information Technology & People* 19(4):371–389.

Näsi, M.; Räsänen, P.; Hawdon, J.; Holkeri, E.; and Oksanen, A. 2015. Exposure to online hate material and social trust among finnish youth. *Information Technology & People* 28(3):607–622.

Perez, S. 2018. Facebook is launching a new groups tab and plug-in. https://techcrunch.com/2018/05/01/facebook-is-launching-a-new-groups-tab-and-plugin. (Accessed Sep 2018).

Peterson, R. S., and Behfar, K. J. 2003. The dynamic relationship between performance feedback, trust, and conflict in groups: A longitudinal study. *Organizational behavior and human decision processes* 92(1-2):102–112.

Poole, M. S. 1999. Organizational challenges for the new forms. *Shaping organizational form: Communication, connection, and community* 453–471.

Rotter, J. B. 1971. Generalized expectancies for interpersonal trust. *American Psychologist* 26(5):443.

Serva, M. A.; Fuller, M. A.; and Mayer, R. C. 2005. The reciprocal nature of trust: A longitudinal study of interacting teams. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior* 26(6):625–648.

Simon, B., and Brown, R. 1987. Perceived intragroup homogeneity in minority-majority contexts. *Journal of Personality and Social Psychology*.

Taylor, P.; Funk, C.; and Clark, A. 2007. Americans and social trust: Who, where and why. *A Social Trends Report*.

Walther, J. B., and Bunz, U. 2005. The rules of virtual groups: Trust, liking, and performance in computer-mediated communication. *Journal of Communication* 55(4):828–846.

Webber, S. S. 2008. Development of cognitive and affective trust in teams: A longitudinal study. *Small group research* 39(6):746–769.

Wilson, J. M.; Straus, S. G.; and McEvily, B. 2006. All in due time: The development of trust in computer-mediated and face-to-face teams. *Organizational behavior and human decision processes* 99(1):16–33.

World Values Survey. 2018. World values survey database. *Social Psychology Quarterly*.

Zhang, Y.; Fang, Y.; Wei, K.-K.; and Chen, H. 2010. Exploring the role of psychological safety in promoting the intention to continue sharing knowledge in virtual communities. *International Journal of Information Management* 30(5):425–436.