Designing for Trust: A Behavioral Framework for Sharing Economy Platforms

Natã M. Barbosa
School of Information Sciences
University of Illinois at Urbana-Champaign
natamb2@illinois.edu

Emily Sun
Airbnb Inc.
San Francisco, CA
emily.sun@airbnb.com

Judd Antin
Airbnb Inc.
San Francisco, CA
judd.antin@airbnb.com

Paolo Parigi
Airbnb Inc.
San Francisco, CA
paolo.parigi@airbnb.com

ABSTRACT

Trust is a fundamental prerequisite in the growth and sustainability of sharing economy platforms. Many of such platforms rely on actions that require trust to take place, such as entering a stranger's car or sleeping at a stranger's place. For this reason, understanding, measuring, and tracking trust can be of great benefit to such platforms, enabling them to identify trust behaviors, both online and offline, and identify groups which may benefit from trust-building interventions. In this work, we present the design and evaluation of a behavioral framework to measure a user's propensity to trust others on Airbnb. We conducted an online experiment with 4,499 Airbnb users in the form of an investment game in order to capture users' propensity to trust other users on Airbnb. Then, we used the experimental data to generate both explanatory and predictive models of trust propensity. Our contribution is a framework that can be used to measure trust propensity in sharing economy platforms via online and offline signals. We discuss which affordances need to be in place so that sharing economy platforms can get signals of trust, in addition to how such a framework can be used to inform design around trust in the short and long term.

KEYWORDS

trust, sharing economy, modeling, Airbnb

ACM Reference Format:
Natã M. Barbosa, Emily Sun, Judd Antin, and Paolo Parigi. 2020. Designing for Trust: A Behavioral Framework for Sharing Economy Platforms. In *Proceedings of The Web Conference 2020 (WWW '20), April 20–24, 2020, Taipei, Taiwan.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3366423.3380279

1 INTRODUCTION

Successful sharing economy platforms have transformed society. Today, hundreds of millions of people depend on these platforms, whether it is simply to go from point A to point B (e.g., Uber, Lyft), travel to and discover new places (e.g., Airbnb, Vrbo), or even rent one's own car to someone else (e.g., Turo). In one way or another, such platforms rely largely upon strangers trusting strangers. Therefore, one could argue that while such platforms provide reliable infrastructures (e.g., guarantees, insurance), ultimately, trust between people is their "fuel" and something of utmost importance in their growth and sustainability. Consequently, studying how trust operates within these platforms can yield powerful insights to inform design in ways that promote the further development of these platforms with user trust in mind.

We define trust as the belief that the other person will not cause harm, even though they may be in the position to do so [18]. That is, a Uber or Lyft driver must trust their passenger, an Airbnb guest must trust their host, and a Turo car owner must trust the stranger who will drive their car temporarily to places unknown. According to Alarcon *et al.* [3, 4], trust actions are informed and motivated by mainly four constructs: (1) one's propensity to trust; (2) perceived trustworthiness of the other (e.g., perceived ability, benevolence, and integrity); (3) dyadic influences and (4) previous trust actions. According to several theories and empirical studies (e.g., [8, 10, 24, 25, 32, 36, 37]), general trust propensity is a fundamental construct of trust which operates early on in one's decision to trust others. Trust propensity may be demonstrated by people's propensity to think others are generally trustworthy [8], or may indicate skepticism and cynicism towards others [10, 37]. The former, will indicate a high propensity to trust and will lead individuals to believe the trustee will not do harm. The latter will indicate a low propensity to trust and will lead individuals to behaving more carefully and conservatively towards others. Propensity to trust is particularly useful early in new scenarios, where information about the trustee which can signal trustworthiness is lacking [3].

In this work, we propose a methodological framework grounded in experimental data and user behavior that can capture and predict an individual's propensity to trust others when making decisions in a sharing economy platform. Through reproducing the steps outlined in this framework, one can more closely understand how the use of certain system features and user behavior may be related to trust propensity, which can help identify design changes that may better serve users on the topic of trust. We believe that collecting data about trust via an experiment captures the behavior of users in the heavily optimized environment of an app better than asking about beliefs toward trust. Understanding what button users click and look at requires data that are similarly centered on behavior.

WWW '20, April 20–24, 2020, Taipei, Taiwan

ACM ISBN 978-1-4503-7023-3/20/04.
https://doi.org/10.1145/3366423.3380279

We conducted an online experiment in the form of an investment game with 4,499 Airbnb users from the US and Canada in order to capture their propensity to trust other Airbnb users. Using data collected through this experiment, we used an Airbnb data set containing signals from online and offline behaviors (e.g., clicks on stay listings: online, number of stays in shared spaces: offline) to model their relationship with one's propensity to trust other users. Finally, we combined the experiment data set with each user's respective behavior to create explanatory models of trust along the user journey of the Airbnb user, as well as predictive models of trust that can be used on aggregate levels to compare groups of users and identify groups who may benefit from design artifacts that can help inform their trust decisions.[1]

We show how our framework can explain and predict propensity to trust on Airbnb, and most importantly, how one can use this framework within the context of sharing economy platforms to (1) identify affordances that signal trust; (2) generate explanatory models of trust; and (3) predict propensity to trust among different groups of users. Therefore, we see the methodological framework as the main contribution of our work. By having a more accurate measure for trust that relates to behavior rather than to beliefs, sharing economy platforms will be better able to serve their users and protect their preferences.

2 RELATED WORK

2.1 Trust on Airbnb

Trust is essential to the functioning of sharing economy services like Airbnb. As a complex system, Airbnb relies on trust in a multi-faceted rather than singular dimension. Ma *et al.* [21] showed the impact that self-disclosure in the profile description has on the trustworthiness of Airbnb hosts, and similarly, Ert *et al.* [16] explored the effect of trustworthiness of host photographs on listing price and probability of being booked. This research highlights how hosts' presentation on Airbnb influences their trust and success on the platform. A different signal of trust on Airbnb comes from the reputation system through which guests leave public reviews and star ratings of listings that they have stayed. Previous research by Qiu *et al.* [31] shows the impact that increased review count and higher ratings have on the trustworthiness of Airbnb listings. The reputation system has been shown to have such an importance on Airbnb that common social biases like homophily are relatively weak drivers of trust [1]. Taken together, this body of work highlights the difficulty of understanding how a given affordance may impact trust in such a complex system as Airbnb.

These prior works have approached trust from the perspective of the social interactions of users on the platform, often considering a dyad of users as the unit of analysis. Differently, the scope of our work is in understanding how each individual user's propensity to trust others operates with regards to their online and offline behaviors on the platform.

[1]The paper does not present the detailed findings of our model, such as exact variable names. Instead, we present more generic terms such "engagement" and "communication," to show the family of potentially interesting metrics correlated with trust propensity.

2.2 Complexity of Trust

Further, while there is a known relationship between trust, risk, and familiarity, the underpinning mechanics of how trust works is still largely not understood. Uncertainty reduction theory applied to the assessment of online profiles purports that the more information shared, the lower the risk and therefore the higher the trust [7, 19]. However, the type of information can have a differential impact on trust [21]. As Luhmann explained, "the relation between confidence and trust is not a simple zero-sum game in which the more confidence is given the less trust is required and vice versa" [20]. It is unclear how trust is developed through familiarity, whether it is a gradual or step-changing function. Personal experience also affects trust, not only the current milieu, making it difficult to study trust in isolation. Past research shows how important different contextual factors such as rarity and familiarity are for interpersonal trust [23]. On the other hand, behavior alone is also not always an accurate proxy for trust; Cheshire *et al.* [13] showed a *negative* association between online activity and website trust. One would infer incorrectly from log data alone that greater usage on a website indicates a higher propensity of trust. From these studies, it becomes apparent that obtaining a snapshot indicator of trust or relying on high-level behavioral data alone do not account for the complexities of trust.

Consequently, we aimed to take a more nuanced approach to obtaining measurements of trust from behavioral data, where we consider a large number of fine-grained behavioral signals obtained both from online and offline interactions surrounding the Airbnb platform, combined with experimental data obtained from users. This approach enables us to unpack how trust propensity operates. In addition, the scope of our work is not around measuring trust toward the Airbnb platform, but instead it is attempting to capture users' propensity to trust other users, given each user's behavior signals from using the affordances provided by the Airbnb platform which facilitate users' decisions about staying and hosting.

2.3 Data Triangulation

Given the intricacy of trust, especially in the context of the sharing economy, triangulation of data sources provides a potential opportunity to generate more accurate models of trust. Data triangulation is largely accepted as a methodological way of reducing bias in research [9, 17, 34]. An increasingly common form of data triangulation combines survey data with qualitative interviews [30]. However, behavioral traces online are a meaningful source of data which have been used for modeling future behavior [11, 33]. Combining behavioral data with an online experiment is a way of accounting for complex historical contexts as well as the effects of individual affordance manipulations at a large scale. Experiments are commonly conducted at technology companies through simple A/B testing, but this method is often limited to guiding specific feature development [15]. An online experiment, and specifically the investment game, is a way to directly measure trust that does not rely on self-report [6]. The game has been used in the past for a variety of applications, including measuring mood disorder on cooperation [28], the effect of trust on arousal [26], and the impact on trust of consumption priming [38].

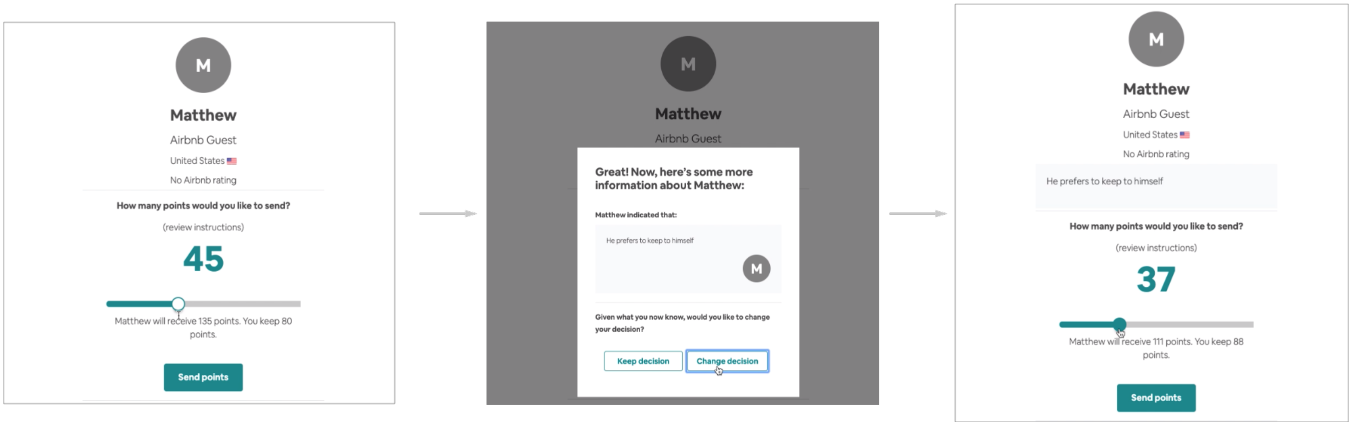

Figure 1: Experiment workflow. Participants saw a profile and were asked how many credits they wish to invest. Then, participants were given an additional piece of information about the profile and were asked whether to keep or change their original decision. We used the first investment as an individual's (i.e., the player) propensity to trust others.

To our knowledge, this is the first time the investment game is being combined with behavioral data to arrive at a model of trust.

3 METHOD

The end goal of our work was to establish a number of steps to capture trust propensity of users on the Airbnb platform, and more broadly, on sharing economy platforms. More specifically, we wanted to create both explanatory and predictive models of trust propensity. The former is a way to explain which behaviors signal low or high propensity to trust other users, captured from the different stages of the user journey (e.g., when searching, before the trip, after the trip, etc.). The latter is a tool for making predictions about trust propensity for each user, which can be useful at an aggregate level, enabling platforms to identify which groups may be better served with features that can inform trust decisions, such as users located in specific countries or cities where trust propensity may be relatively low. Such interventions can surface and prioritize information that helps inform trust decisions. While trust is an important aspect for both hosts and guests, it operates differently for them. For this reason, we built distinct models for hosts and guests. If this framework were to be used within the context of a ride-hailing sharing economy platform, for example, it would need to have distinct models for drivers and passengers. In general, our approach distinguished between service providers, e.g. the host, the driver, etc, and consumers, e.g. the guest, the rider, etc.

Since the behavior we were interested in modeling consisted of actions in the app (clicks, scrolls, etc.), we opted to not use a survey for collecting beliefs about trust. Surveys have been shown to have poor correlation with choices users made in apps [5]. For example, users often report about the importance of privacy, yet they click "I agree" buttons to grant access to large amounts of their personal data. In light of all of this, the first step of our approach consisted of conducting an online experiment with Airbnb users to capture propensity to trust via an investment game. After that, we collected online and offline behaviors that we thought might have a relationship with trust propensity. We created a data set for modeling trust propensity made of the trust measure from the experiment and dozens of features taken from log behavior on the platform. We then selected the most important features using machine learning, and finally, we created and evaluated models of trust propensity within the context of Airbnb using a data triangulation approach. While the scope of this work is the Airbnb platform, we believe this methodology can be replicated in other sharing economy platforms in a similar fashion.

3.1 Online Experiment

In order to capture propensity to trust, we conducted an online experiment with Airbnb hosts and guests, recruited from two countries, the United States and Canada, to play an investment game.

The investment game is a behavioral game built on game theory and widely used for measuring trust [6]. In the experiment, participants played the role of investor and started with a number of credits (125 points). We showed investors a single profile that had a rating (either 0 or 5 stars), a role (either a guest or a host), a country (picked from a small set), a name (either male or female) and an avatar (either colored or grey, with no picture). We chose these variables because they are part of the Airbnb profile page users will have to fill out when joining the platform. We presented the profiles as if they belonged to randomly selected Airbnb users that were assigned the role of recipients. In reality the recipients were synthetic profiles we created in order to see which combination of characteristics gathered the most investment. The user interface of the experiment was purposefully made to replicate Airbnb's visual language, including button colors and fonts, as well as micro-impressions such as animations and transitions. While great effort was placed to mimic real Airbnb profiles for ecological validity, participants knew they were playing a game outside of the Airbnb platform. The countries of the profiles were selected as the two most common countries guests from the US and guests from Canada travelled to in 2018 and as the two most common countries guests from abroad travelled in the US and in Canada in 2018. These countries were, for Canadian guests: United States and Italy; for American guests: Canada and Great Britain; for Canadian hosts: United States and France; for American hosts: Canada and Great

Britain. If the participant was a guest, he would see the profile of a host; if the participant was a host, he would see the profile of a guest. So, to illustrate, a Canadian guest participating in the game would see either the profile of an American host or of an Italian host because these were the two most common destination countries in 2018 among Canadian guests.

Participants were asked to invest any of their credits in a given profile. In order to create the simulation of a game we told participants that any point they invested would be tripled and given to the recipient. Figure 1 shows the workflow of the experiment. Further, we told participants that the recipient had up to 2 weeks to decide how many points to return to the investor. This setting created the illusion that the profile was a real player chosen at random for the role. The deception component of our study was deemed minimal, as reviewed and approved by Stanford University's IRB.

In the spirit of game theory, the receiver could cooperate with the participant and return a good portion or all of the credits received or defect and keep most or all of these credits. This setting placed participants in a vulnerable position and turned the amount invested in each receiver into a proxy for how much trust they placed that the given profile would return their credits.

In contrast to lab experiments, the external environment challenges the attention of online experiment participants. Thus, we gamified our experiment in order to win participants' attention. In particular, we presented the incentive structure not as a lottery but as a game. Our gamification approach mirrors the approach described in Abrahao *et al.* [1]. Particularly, to motivate participants to pay attention to how they allocated points, the top-15 players in total points accumulated would win a $50 gift card (and participants knew this). To test our game, we ran a pilot with 100 users, in which some participants played the role of receivers and manually returned the points invested. From the pilot results, we modeled the return mechanics using the cumulative distribution of investment returns, improved on the game instructions, and added a tutorial section before the actual investment. We recognize that gamification could potentially encourage investing and thus, in our written instructions about the game to participants, we emphasized that participants did not need to invest any points. Consequently, 114 participants invested no points, 44 invested 1 point, 28 invested 2 points. The first quartile of the investment distribution was 25 points out of 125 points. That is, the gamification did not appear to significantly bias participants toward investment.

Therefore, the more the participants trust that the given receiver would cooperate, the more they invested because that would potentially increase the original amount of points invested and thus, increase their chances to win the gift card. The presence of a monetary incentive tied to performance in the game forced participants to pay attention to how they decide to invest points. The experiment user interface also contained instructions and rules accessible at all times, and every participant was guided through a tutorial run explaining the user interface and the rules before they were asked to make the real investments.

We collected data in two waves. We launched the experiment in Canada in February 2019 and in the United States in April 2019. For each country we sent out 50,000 emails to Airbnb users split between hosts (15,000) and guests (35,000). We had a total of 4,762 completed games. 52% of the completed games were from Canada with 56% of the completed games from guests. About 60% of participants were female. After filtering out incomplete responses, our data set contained 2,628 responses from guests and 1,871 responses from hosts. All users participating in this study signed a consent form approved by Stanford University's IRB.

We took each participant's original investment given in the game as a proxy for their propensity to trust other users, effectively being the dependent variable in our modeling work. We discounted the effect of 5-star ratings on profiles into the investment by adding the estimate of the main effect of the 5-star rating from a linear model to the responses where the profile included a profile with no star ratings. Since the average difference between trust investment between the 5-star condition and the 0-star condition was about 35 points, we added these points to the 0-star conditions to make the results comparable. This was needed in order to discount the variance observed in trust propensity due to having seen a profile with a 5-star rating in the game.

3.2 Data Set Creation

After collecting the measure for propensity to trust from the experiment, we joined the experiment data set with a wide data set containing behavioral signals of the same Airbnb users who participated in the online experiment. This data set of behavioral signals was extracted from Airbnb logs and contained signals for 85 different behaviors, with a combination of online and offline signals. For example, for guests, one online signal is the historical review-reading behavior on an Airbnb listing, and one offline signal is the number of nights that they stayed in a shared space before (e.g., a private or shared room on Airbnb). For hosts, one online signal is their historical guest-profile visiting behavior, and one offline signal is how many new guests they have hosted before, meaning guests with no prior reviews/experience on the Airbnb platform.

3.2.1 Choice of Behavioral Signals. The individual decisions about which behavioral signals should be included in this data set were guided by theories of trust, more specifically, trust constructs as synthesized in Alarcon *et al.*'s review [3, 4] of prior works on trust behavior. For example, as a host, the number of guests hosted with no prior reviews by other hosts signals a prior trust action – given the associated risk involved – therefore it is a relevant signal to include if we were to model trust propensity for hosts. Likewise, prior works [4] suggest that one's general propensity to trust others can be target-agnostic – that one may have "baseline" trust intentions regardless of who the trustee is – and follow a rational path that optimizes for the best personal outcome. For this reason, historical review-reading behavior of guests when choosing a place to stay is a reasonable signal to include, since this behavior can arguably be taken as a rational warranting action undertaken to mitigate risk.

Once these behavioral signals were mapped and included in the data set, we conducted feature selection and importance analyses by running a Lasso regression analysis. This was so that we would find the optimal, smallest number of signals to use as a predictor of trust propensity. Lasso penalizes a model for having too many predictors, and also determines feature importance, which gives us the explanatory model with relative importance of each behavior signal into trust propensity. The Lasso regression analysis searched for the best outcome using different alpha levels in the range [0.0001..1] to select the variables that would best serve as independent variables to explain and/or predict trust propensity (as captured by the online experiment). We used the *LassoCV* implementation of Scikit-learn [29] for this analysis.

These procedures resulted in 18 behavioral signals selected for guests, and 11 signals selected for hosts. These served as independent variables used in our models.

3.2.2 Triangulation with Trust Attitudes. An additional step we followed to ensure we were capturing the proper signal was to triangulate the trust propensity signal captured via the online experiment with trust attitudes captured in previous surveys of Airbnb hosts and guests. These data sets were generated by the Survey Science team at Airbnb and include Likert-scale (1-5) responses to questions such as *"How trustworthy are Airbnb guests?"* for hosts and *"How safe do you feel when staying at Airbnb?"* for guests. While self-reported perceived trustworthiness and perceived safety can be different from trust propensity, these data sets were useful in a triangulation analysis, where the direction of the relationships should hold if we made predictions about trust propensity for users who took these surveys. In other words, we would expect users with low trust propensity to also show low perceived trustworthiness of other users. Likewise, guests with low trust propensity might also show low self-reported perceived safety when staying at Airbnb. These survey data sets had responses for over 200,000 Airbnb users.

This triangulation was done in two steps. The first step consisted of creating a data set with the 85 behavioral signals from Airbnb logs for the users who took the attitudinal surveys and running the Lasso regression to determine if similar features would be selected. The second step consisted of calculating the average prediction of trust propensity from a trained model on the online experiment data for individuals within the Low (1-2 responses) and High (4-5 responses) buckets, and verifying that the average propensity to trust is higher for users reporting higher general perceived trustworthiness of other users, for example. We note here that we did not use these attitudinal, self-reported signals to train our models, and only used them to triangulate data and obtain further evidence that we were capturing the intended signal of trust from the behavioral variables selected by the Lasso regression.

3.3 Model Generation

Once we established confidence in the features selected by Lasso, we wanted to use them to generate both explanatory and predictive models of propensity to trust others. This section outlines the steps we took to create them.

3.3.1 Explanatory Model. From the features selected and their respective relative importance, we created models explaining the behavior of hosts and guests with low and high propensity to trust others, mapped to the user journey of an Airbnb user. For example, when searching and comparing Airbnb listings, a guest with low propensity to trust would engage with user profiles differently than a user with higher propensity to trust. This conclusion was made from the results of the Lasso regression, where the coefficient for each behavior was estimated. Likewise, a guest with high propensity to trust others will engage in messaging behavior in a different way compared with guests that have lower trust propensity.

For guests, the user journey we mapped consisted of *Search & Compare*, *Request for engagement*, *Pre-engagement*, and *Post-engagement*. For hosts, the user journey consisted of *Listing services*, *Request for engagement*, *Engagement*, and *Post-engagement*. We did not map any steps of the journey for which we thought we did not have behavioral variables for them. In other words, the user journey steps we mapped for the explanatory model are not comprehensive, and may not include all the points along the journey a user goes through when using the Airbnb platform or other sharing economy platforms. We also purposefully attempted to make them generic so that they are also suitable to other platforms.

3.3.2 Predictive Model. In addition to the explanatory models, we also wanted to generate a model that could "calculate" a user's propensity to trust other users on the platform. More specifically, we wanted to be able to categorize users as having *Low*, *Medium*, and *High* trust propensity. This consists of an ordinal classification problem where *Low* < *Medium* < *High*.

We started by bucketing the dependent variable given by users in the online experiment into the appropriate categories (namely *Low*, *Medium*, and *High*). This resulted in a nearly balanced data set, each group belonging in the proper range. We then split the data set containing 4,499 rows into 60% for training, 30% for validation, and 10% for test. We had one data set for hosts and one for guests. The data set for guests contained 18 numerical features and the data set for hosts contained 11 numerical features. These are the same features selected by the Lasso analyses.

The next step consisted of evaluating a potential predictive model. For this, we used a neural-network based approach introduced by Cheng *et al.* used for ordinal classification [12]. We implemented this algorithm with Keras [14], using the MXNet back-end. Our performance metric was the multiclass F-1 score. This approach was suitable for our classification problem, which consists of an ordinal multilabel classification where *Low* < *Medium* < *High*.

Our final model was trained to predict the three categories of propensity to trust with a deep neural network architecture using a Sigmoid activation function with binary cross-entropy loss. We used a batch size of 32 over 1,000 epochs with a learning rate of 0.01, with Early Stopping to prevent overfitting with a tolerance of 3 epochs. Our network had one hidden layer with the same size as the input layer, and three nodes in the output layer (representing the three categories). Our dependent variable was encoded as a multilabel problem where *Low* propensity to trust was encoded as [1,0,0], *Medium* as [1,1,0], and *High* as [1,1,1]. This encoding enables the Sigmoid function to learn the proper relationship in the ordinal multilabel classification problem. Adapting the binary cross entropy loss function for multilabel classification (instead of using categorical cross entropy) requires an additional step before a class can be assigned, which is to exhaustively find the optimal probabilistic threshold for each label individually. Thresholding (or threshold tuning) is often necessary in multi-class classification (see Al-Otaibi *et al.* [2] for a review of methods). Score-based, label-wise thresholding on a hold-out or cross-validation set is one of the approaches used, so we applied this. This is needed mainly for two reasons (1) situations of class imbalance, which in our ordinal classification case applies when labels are translated into [1,0,0], [1,1,0], and [1,1,1] and (2) the use of the Sigmoid activation function

Guest Model			
[Negative or Positive] Behavioral Signal	**% of most important**	**[On/Off]Line**	**Stage in User Journey**
[-] Reading reviews in full	1	Online	1. Search & Compare
[+] Engaging with full descriptions	0.8	Online	1. Search & Compare
[+] Length of communication with other user	0.299	Online	2. Request for engagement
[+] Engaging with photos	0.296	Online	1. Search & Compare
[-] Engaging with other user's user-generated content	0.258	Online	1. Search & Compare
[-] Engaging with profiles of authors of past reviews	0.248	Online	1. Search & Compare
[-] Engaging with other user's brief intro	0.185	Online	1. Search & Compare
[+] Engaging with past reviews beyond the first page	0.165	Online	1. Search & Compare
[-] Engaging with rules specified by other user	0.129	Online	1. Search & Compare
[-] Requests approved in previous engagements	0.129	Online	2. Request for engagement
[-] Cancellations	0.121	Offline	3. Pre-engagement
[-] Requests for engagement denied	0.064	Online	2. Request for engagement
[+] Level of exposure to other user in previous engagements	0.035	Offline	4. Post-engagement
[-] Engaging with other user's profile	0.031	Online	1. Search & Compare
[-] Intended duration of previous engagements	0.027	Offline	3. Pre-engagement
[+] Contacting the other user	0.025	Online	2. Request for engagement
[-] Engaging with "contact" affordances	0.015	Online	1. Search & Compare
[+] Engaging with photos	0.015	Online	1. Search & Compare
Host Model			
[Negative or Positive] Behavioral Signal	**% of most important**	**[On/Off]line**	**Stage in User Journey**
[+] Total engagements in the past	1	Offline	4. Post-engagement
[-] Communications exchanged	0.662	Online	2. Request for engagement
[-] Rejected requests for engagements in the past	0.608	Online	2. Request for engagement
[+] Prior requests for engagement received in the past	0.484	Online	2. Request for engagement
[+] Relaxing approval requirements for engagements	0.434	Online	1. Listing services
[+] Highest possible ratings received in the past	0.223	Online	4. Post-engagement
[+] Previous engagements with new users	0.214	Offline	4. Post-engagement
[-] Engaging with other user's profile	0.160	Online	2. Request for engagement
[+] Party size of previous engagements	0.120	Offline	3. Engagement
[+] Intended duration of previous engagements	0.059	Offline	3. Engagement
[-] Number of listings managed	0.019	Offline	1. Listing services

Table 1: Selected behavioral features ordered by feature importance. These can be used to explain behavior of users with low and high propensity to trust others.

instead of Softmax gives independent class probabilities and so output prediction values must be used to find optimal thresholds for each class. See Maxell *et al.* [22] for a reference. The classifier output then returns an array of class probabilities corresponding to each instance, and a class can be assigned to the output by giving a binary outcome to each class defined by $[p(low) >= T_l, p(medium) >= T_m, p(high) >= T_h]$, where the thresholds T_l, T_m, and T_h were found via exhaustive search of the best F-1 score within the range of probability thresholds [0-1] at increases of .0001. This threshold finding procedure was implemented during model selection, using the validation data set, and later saved for making predictions on the test split. The best thresholds found for the guest model were [0, .6622, .4022] and [0, .6388, .4159] for the host model. We used the multiclass weighted F-1 metric as calculated by implementations provided in Scikit-learn.

Once the best model was trained, we saved it, along with its best probability thresholds, for making predictions for new data points.

4 RESULTS

4.1 Descriptive Statistics

We provide next a description of the distribution of points invested by users in the investment game. The overall mean investment was 56.65 points, with Canadian participants being more trusting than American participants: 59.29 points versus 53.74 points respectively. This difference was statistically significant (p<.001). About 700 participants invested all their points, i.e. 125, while a bit more than 100 participants kept all their points, i.e. invested 0 points. The 1st quartile of the investment distribution was 25 points exactly. The average number of points invested was 56.65 (Min=0, 1Q=25, Median=50, 3Q=90, Max=125).

4.2 Feature Selection

The feature selection process resulted in 18 features being selected for guests and 11 for hosts. For the host model, five out of the 11

Attitudinal Survey Question	Avg. Prediction
How trustworthy are Airbnb guests? (N=4,449)	
High (4-5)	2.14
Low (1-2)	2.07
How safe do you feel when hosting guests in your listing(s) with Airbnb? (N=52,141)	
High (4-5)	2.16
Low (1-2)	2.06
Trust in Airbnb if things go wrong (N=65,317)	
High (4-5)	2.14
Low (1-2)	2.07

Table 2: Triangulation of attitudes and experimental data on trust for hosts. The average trust propensity prediction for users in the High groups is greater than those in the Low groups. All differences are statistically significant.

features represented offline behavior, and for guests, three out of the 18 did so. Table 1 shows the relative importance of each feature along the user journey for online or offline behaviors.

From here on, we use the noun "engagement" to refer to the object of interest of a sharing economy platform. For instance, on Airbnb, the engagement can be a stay booked through the platform, whereas on a ride-hailing platform, the engagement is the ride. Differently, when using the verb "engaging," we mean behavior that can be associated with clicking, visualizing, or reading online content. Table 1 shows the selected features for hosts and guests. For hosts, the five most important behaviors associated with increased propensity to trust are the number of prior engagements (offline), the number of prior requests for engagement received in the past (online), the relaxing of approval requirements for engagements (online), highest possible ratings received in the past (online), and previous engagements with new users (offline). On the other hand, the behaviors negatively impacting trust propensity are communications exchanged (online), rejected requests for engagement in the past (online), engaging with other user's profile (online), and the number of listings managed (offline).

For guests, the most important behaviors positively associated with trust propensity are: engaging with full descriptions (online), length of communication with the other user (online), engaging with photos (online), engaging with reviews past the first page (online), and level of exposure to other user in previous engagements (e.g., in the case of Airbnb, staying in a shared space may imply more exposure than an entire home) (offline). The five most important behaviors associated with decreased propensity to trust for guests are reading reviews in full (online), engaging with other user's user-generated content (online), engaging with profiles of authors of past reviews (online), engaging with other user's brief intro (online), and engaging with rules specified by other user (online).

4.3 Triangulating Attitude and Behavior

We triangulated the experimental data with attitudinal data by training our model using the experimental data and making predictions for hosts who took the attitudinal survey data, using their behavioral features according to the ones included in our model. The trained model made predictions, assigning individuals to the Low (1), Medium (2) and High (3) trust propensity classes. We then bucketed attitudinal survey respondents into High and Low groups (excluding 3-Medium), and calculated the average predicted trust propensity within each group.

Table 2 shows the results of the triangulation for hosts, with the averages within each group across different survey questions. A Mann–Whitney U test showed a statistically significant difference of the average trust propensity prediction between groups (perceived guest trustworthiness $p<.05$, perceived safety $p<.001$, trust in Airbnb $p<.05$). In other words, the predictions of trust propensity from behavioral data corresponded to attitudes of users toward perceived trustworthiness of other users, trustworthiness of Airbnb, and safety when hosting at Airbnb. This was an important result for us which provided further evidence of the nature of the signal we captured via the online experiment, along with the behavioral data used in the modeling.

While there are obvious distinctions between perceived safety, trustworthiness, and trust in the Airbnb entity, we believe these constructs are correlated and the fact that the differences between groups held the expected differences is an additional indicator that the variables we used were capturing the intended signal of trust.

4.4 Explanatory Models of Trust Propensity

In order to create explanatory models of trust propensity, we mapped each selected feature to possible steps in the user journey of a sharing economy platform. We purposefully attempted to make the steps in such user journey – along with their respective features – generic in ways that can be applicable to other sharing economy platforms. By considering the signal of the coefficients found via our Lasso analyses, we were able to design, for example, vignettes about how a user with low or high propensity to trust others would behave along the journey at Airbnb, and potentially other platforms.

As an example, if we take Turo [2], a car sharing platform, as the sharing economy platform, and use Table 1 to create vignettes aimed at explaining actions of users with low and high propensity to trust along the user journey, when dealing with requests to book (2. Request for engagement), car owners with low trust propensity will exchange more messages with guests, reject more requests to book, and visit the guest driver's profile more before booking. Conversely, a car owner with high propensity to trust others will automatically pre-approve requests to book when first listing their car (1. Listing services), will be one that has had many prior engagements, and has received many requests for engagement. With regards to driver reservations (3. Engagement), a car owner with high propensity to trust others will also book more new guest drivers (i.e., drivers with no prior reviews by other car owners) and book longer reservations. Finally, signals taken from the post-booking stage (4. Post-engagement), a high-trust-propensity car owner will continue booking, and will receive more 5-star reviews.

If, for example, we take a ride-hailing sharing economy platform and use Table 1 to explain trust propensity behavior, then one could say a driver with low propensity to trust others, when dealing with requests to ride (2. Request for engagement), will exchange

[2]https://turo.com/

Performance Metric	Guest Model	Host Model
Training Accuracy	0.792	0.781
Validation accuracy during training (20%)	0.767	0.763
Validation Precision	0.773	0.784
Validation Recall	0.756	0.774
Validation F-1	0.762	0.770
Validation Hamming Loss	0.280	0.249
Test Precision	0.791	0.711
Test Recall	0.761	0.718
Test F-1	0.773	0.715
Test Hamming Loss	0.268	0.293
Best Thresholds	[0, 0.6622, 0.4022]	[0, 0.6388, 0.4159]

Table 3: Performance of predictive models.

messages more with passengers, reject more requests to ride, and read information about the passenger more. Conversely, a driver with high propensity to trust others will automatically pre-approve requests to ride, and is likely one that received many ride requests before. With regards to rides (3. Engagement), a driver with high propensity to trust others will accept rides from new passengers (i.e., passengers with no prior reviews by other drivers), accept rides for larger party sizes, and accept longer rides. Finally, signals taken from the post-ride stage (4. Post-engagement), a high-trust-propensity driver will continue driving passengers, and will receive more 5-star reviews. Vignettes for guests, in the case of Airbnb, or passengers, in the case of a ride-hailing app, could be generated in similar fashion, using the explanatory framework to translate behavior within each context.

An important point on the generalization of our framework is that the features we identified are largely dependent upon the context and features implemented on Airbnb at the time of our analyses. This is important because our framework in practice enables one to capture relationships of trust propensity from the current affordances of each platform. To illustrate this difference, it may not make sense, for example, that a driver of a ride-hailing platform will message users back and forth due to the cost and risk of doing so within the context of driving. Also, one could argue that the consequences of accepting an undesirable passenger may be relatively less worse in terms of property damage than the risk of accepting an undesirable guest to a half million dollar property. In summary, there are many factors underlying the nature of the engagement that will determine the level of warranting needed by users with low propensity to trust. Nonetheless, we believe the methodology we propose will unearth features that are relevant to each platform and therefore our work can be generalized.

One final note about our explanatory models is that they can also unearth affordances (e.g., viewing profiles, reading reviews, viewing photos, viewing user-generated content) which when implemented into a sharing economy platform, will enable one to capture how trust propensity operates via the use of such affordances.

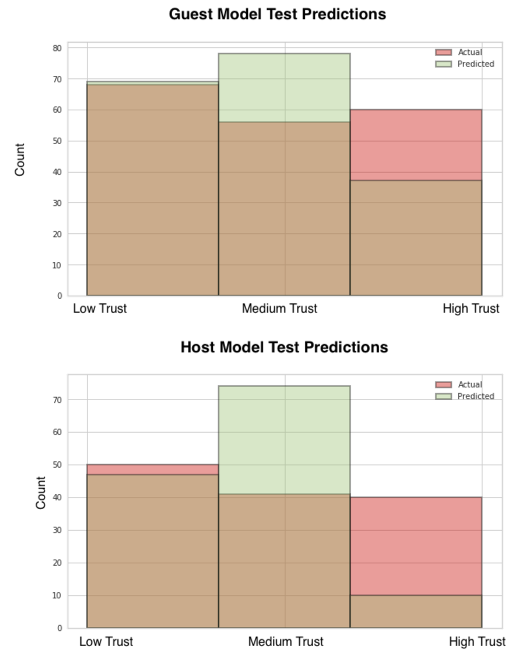

Figure 2: Distribution of predictions on the test data set for each model.

4.5 Predictive Models of Trust Propensity

We trained a neural network model for each user type, namely guests and hosts. The guest model had 18 features, and the host model had 11 features. All the features were numeric, and they were properly scaled as a preprocessing step.

The guest model was trained over 65 epochs with early stopping, with a training accuracy of .792 and validation accuracy during training of .767. The validation F1-score of the guest model on the 30% validation split was .762, and .773 for the 10% test split.

The host model trained over 41 epochs, and had a training accuracy of .781, and a validation accuracy during training of .763. The validation F-1 score (evaluated via the hold-out 30% validation set) of the host model was .770, and .715 for the test split (evaluated via the hold-out 10% test set). Table 3 shows all the performance metrics used to evaluate the performance of the predictive models, along with their results. Figure 2 shows the counts of predicted and actual classes on the test split. While the data sets are nearly balanced, the models "over predict" for the *Medium* class and "under predict" for the *High* class. This was expected given the nature of the data, but also can be an acceptable limitation if the priority is to identify individuals with low trust propensity with confidence. That means one may prioritize identifying low trust propensity users over those with high propensity to trust because doing so may lead to actionable steps in helping users gather pieces of information needed to build trust where they might otherwise not have them. For example, knowing that hosts within a certain region have a lower general propensity to trust guests, a message could remind these hosts that they can manually approve their bookings, or that they can message guests before approving a request, as informed by the explanatory models. The ultimate goal would be to equip users with the information needed to make informed trust decisions, according to the features available in the platform.

We note that while the resulting performance is reasonable, this performance may not be enough to tailor the experience of individual users based on predictions made at the individual level, given a relatively low F-1 score. However, at the aggregate level, for example, grouping users based on countries, cities, or other grouping factors, the ratio High:Low users classified can be used to compare trust propensity between groups. At the aggregate level, using trust as a reference, one can also compare other variables not originally included in the initial set of behaviors to identify "with regards to what behaviors" users with low and high propensity behave differently. In other words, in what ways do the behavior of users with low and high propensity to trust is different *beyond* the behavioral signals included originally in the model? Answering this question can inform further development of the behavioral set.

5 DISCUSSION

Our proposed methodology enables one to measure users' propensity to trust other users on a sharing economy platform. In our work, this platform was Airbnb, but we hope researchers and developers can reproduce our methodology to gain knowledge about and capture trust signals on other platforms as well. This proxy measure for trust can inform design insights and evaluate whether design interventions are successful in the long-term. For example, knowing that the use of a certain affordance on the user interface signals low propensity to trust other users, designers could implement the affordance in ways that better support these users in making their decisions, such as by providing additional information or by guiding them on how to use a feature more effectively. Likewise, when launching a feature aimed at promoting trust, teams could verify its effectiveness by looking at whether the feature is used mostly by individuals with low or high propensity to trust. Perhaps such new feature is targeted in ways that would reduce the reliance of users on features that are used because of low propensity to trust, and therefore their usage would be reduced by launching said new feature. These and many other uses are possible once the knowledge of users' propensity to trust others is obtained, effectively enabling teams to design for trust. Nonetheless, in using the framework in this manner, one must be careful not to suppress a healthy suspicion on the users' part, since there could be situations in which real risks are present, such as in the case of scams or heightened safety concerns. For this reason, a better approach might be to consider users' trust propensity in combination with the determination of dangers in any given scenario. For example, building trust may not be appropriate when the predicted risk of fraud is high.

The scope of this work was one's propensity to trust others, a proxy for trust associated with each individual. However, this framework can be expanded to also assign a measure for trust to user dyads. In the case of Airbnb, for example, this dyad would be a host and a guest. In the case of a ride-hailing platform, such dyads would be composed by a driver and a passenger. In considering the dyad as the unit of analysis, one could identify low-trust dyads which may benefit from design interventions. Features of such models could include features that affect perceived trustworthiness of other users, such as perceived benevolence, ability, and integrity (e.g., [21]). In fact, our experimental data can capture variance in the ratio of the initial to the final investment due to different interventions showing additional information about the profile in ways that can be operationalized for this purpose. For example, typos in the body of message or profile description may negatively affect perceived trustworthiness, which is information that can be used by models in real time to assign a trust score to dyadic interactions, where one's general propensity to trust can become a predictor, along with features of the interactions within the dyad.

One important aspect of the framework is to diversify the baseline trust captured via the experiment. Our experiment was conducted with users in the US and Canada, and trust propensity predictions made from models using these data may not generalize well in other countries where trust is in decline or historically low. This means that a model trained on a country where most people have highest propensity to trust others, would predict people from most other countries as having low propensity to trust. For this reason, it may be beneficial to gather experimental data from users in countries at different points in the low-high trust spectrum.

While making changes to the user interface based on findings from the framework may sound reasonable, by modifying the user interface, signals can start being conflated and no longer useful for measuring trust, since they have now been potentially biased. For this reason, careful consideration must be given to interventions and perhaps decisions on which features are better "left untouched" so that trust signals are reliable and consistent over time.

In our work, we purposefully stayed away from including demographic characteristics in our models (e.g., country, gender). This was made so that models would not learn from such features, which may lead to biases, unfair experiences, and other ethical issues. However, we suspect that characteristics such as country of origin, gender, or age, may be highly informative of trust propensity, while user interface behavior such as the ones we leveraged in our work are not as informative. One way to leverage such likely relevant features in an ethical fashion would be to operationalize them in indirect ways. For instance, it is well known that homophily can influence trust decisions. A feature could perhaps signal a gender-match or country-match within the dyad, which would inform the model without learning from each individual level of a demographic. A concrete example of indirectly leveraging demographics in trust dyads could be a continuous "homophily score," where homophily information is embedded. For example, this value is increased in a female-to-female user match in shared room or private room listings, where a guest who identifies as a female may be more willing to trust host who identifies as a female for safety reasons. This score could also change based on other homophily effects such as sharing the *alma mater* or the country of origin, all without actually adding the raw demographic information into the model.

When triangulating the experimental data with attitudinal data collected via surveys, we noticed a greater level of noise for guests. We suspect this had to do with the fact that, from the guest's perspective, there are many confounds factoring into a decision to stay at a place on Airbnb, and trust in the host may not be a primary factor in the decision. On the other hand, hosts are indeed choosing to trust a person or a group of people coming to stay in their listing, therefore, their behavior on the platform with regards to how they make hosting decisions may be more likely to associate with trust.

5.1 Implications for Design

The main advantage of implementing this framework is the unearthing of behaviors of users with low and high propensity to trust others on a sharing economy platform. This new knowledge can inform designers in ways that enable them to design with users' trust in mind. For example, one could argue that user interface affordances used by users with low propensity to trust others are in fact helping them build trust. In identifying such features, designers can prioritize information and features in ways that empower users to make more informed trust decisions. One way to do this is by first using the predictive models to identify groups who may benefit from such interventions, and then making the affordances more accessible/prominent to users in those groups. Still, interventions should be carefully considered *vis-a-vis* risk assessments.

In being able to identify groups of users who have low and high propensity to trust others, one could measure the ratio of high to low propensity users over time and understand if certain changes are in fact helping users grow trust. These measurements can be a powerful tool in identifying when trust may be decaying or increasing as sharing economy platforms grow and reshape to accommodate their external demands and environments.

Another use for this framework is in being able to measure users' propensity to trust at times of organizational crises, which in turn could give insights about trust in the brand. Brand value is an important concept which is highly correlated with trust [35].

In using both the explanatory and predictive models to inform design, sharing economy platforms could structure their affordances and user interfaces in ways that prioritize trust. For example, knowing that hosts with high propensity to trust others tend to host new users without prior reviews more often than other hosts, search results for a new user could be ordered by each listing's host propensity to trust. This will in turn increase the likelihood that a booking will be successful for the right reasons, namely, that a host with high propensity to trust was comfortable taking in a guest with no prior reviews by others. Design decisions around promoting trust in scenarios involving new users can translate directly into growth for sharing economy platforms.

Trust propensity captures only one level of how trust operates, and is a measurement assigned to each individual. From our experimental data, we could expand the framework to also assign a trust measurement to dyads. This would enable real-time shaping of experiences of users that may involve filtering, matching, and ordering artifacts along the user journey in ways that build and promote trust. We encourage more research in this area.

We conducted this work under experimental settings in order to carry out our research project. One open and important challenge that must be addressed before considering using this framework in real settings is how to be transparent about design that incorporates and measures trust propensity, recognizing that providing too much detail may actually result in less understanding of online data use [27]. A potential open question is how to design consent and control mechanisms for this purpose.

5.2 Limitations

The training data used in our models were collected from individuals in the US and Canada. These users' propensity to trust are likely affected by western views and perspectives, and are not representative of trust-related behavior in other countries (nor in their own countries). For this reason, we recommend experimental data be collected from countries at different places on the "trust intentions" spectrum. We also recruited hosts and guests matching certain conditions which may not be representative of all users on the platform. Nonetheless, our framework establishes a foundation for sharing economy platforms to measure their users' trust propensity.

What we observed in our experiments is correlation between certain user behaviors and their propensity to trust other users. We do not claim causation in such relationships nor do we know in reality the direction of these relationships. For example, do users engage with other users' profiles because they have a low propensity to trust others, or do they have a lower propensity to trust others because they engage frequently with other users' profiles?

Finally, we note that there are likely confounds between users' propensity to trust other users and their propensity to trust the sharing economy platform. Our framework does not make this distinction, and it would be useful for future works to investigate this in more detail, effectively untangling trust within the provider–platform–consumer triad.

6 CONCLUSION

Trust is fundamental to the existence of sharing economy platforms. While these platforms provide the infrastructure for making ride-hailing and home-sharing possible, which include guarantees and insurance, they ultimately depend upon people's willingness to trust strangers. Trust is a complex topic which operates along many different dimensions. One such dimension is a person's general propensity to trust others. Unpacking this dimension can yield meaningful knowledge to inform design for trust on sharing economy platforms. Using experimental, attitudinal, and behavioral data from Airbnb users, we introduce a behavioral framework through which sharing economy platform users' general propensity to trust others can be captured and measured. More specifically, we conducted an online experiment with 4,499 Airbnb users in the form of an investment game in order to capture users' propensity to trust other users on Airbnb. Then, we used the experimental data to generate both explanatory and predictive models of trust propensity. We showed how our methodological framework can be used to measure trust propensity in sharing economy platforms, and discussed how our framework can be implemented to inform design tailored to trust.

ACKNOWLEDGMENTS

We thank Nate Kugland, Robert Morris, Kieu Vuong, Real Canty, Fan Guo, Juca Lopes, Karen Cook, and many others from the Trust team at Airbnb for their individual contributions to this project. We also thank anonymous reviewers for their insightful comments and suggestions.

REFERENCES

[1] Bruno Abrahao, Paolo Parigi, Alok Gupta, and Karen S. Cook. 2017. Reputation offsets trust judgments based on social biases among Airbnb users. *Proceedings of the National Academy of Sciences* 114, 37 (2017), 9848–9853. https://doi.org/10.1073/pnas.1604234114 arXiv:https://www.pnas.org/content/114/37/9848.full.pdf

[2] Reem Al-Otaibi, Peter Flach, and Meelis Kull. 2014. Multi-label classification: A comparative study on threshold selection methods. In *First International Workshop on Learning over Multiple Contexts (LMCE) at ECML-PKDD 2014*.
[3] Gene M Alarcon, Joseph B Lyons, and James C Christensen. 2016. The effect of propensity to trust and familiarity on perceptions of trustworthiness over time. *Personality and Individual Differences* 94 (2016), 309–315.
[4] Lyons J. B. Christensen J. C. Klosterman S. L. Bowers M. A. Ryan T. J. Wynne K. T. Alarcon, G. M. 2018. The effect of propensity to trust and perceptions of trustworthiness on trust behaviors in dyads. *Behavior research methods* 50 (2018), 1906–1920. Issue 5.
[5] Susan Athey, Christian Catalini, and Catherine Tucker. 2017. *The Digital Privacy Paradox: Small Money, Small Costs, Small Talk*. Working Paper 23488. National Bureau of Economic Research. https://doi.org/10.3386/w23488
[6] Joyce Berg, John Dickhaut, and Kevin McCabe. 1995. Trust, reciprocity, and social history. *Games and economic behavior* 10, 1 (1995), 122–142.
[7] Charles R Berger and Richard J Calabrese. 1974. Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human communication research* 1, 2 (1974), 99–112.
[8] John K Butler Jr. 1999. Trust expectations, information sharing, climate of trust, and negotiation effectiveness and efficiency. *Group & Organization Management* 24, 2 (1999), 217–238.
[9] Donald T Campbell and Donald W Fiske. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin* 56, 2 (1959), 81.
[10] Jennifer A Chatman. 1989. Matching people and organizations: Selection and socialization in public accounting firms.. In *Academy of Management proceedings*, Vol. 1989. Academy of Management Briarcliff Manor, NY 10510, 199–203.
[11] Ye Chen, Dmitry Pavlov, and John F. Canny. 2009. Large-scale Behavioral Targeting. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '09)*. ACM, New York, NY, USA, 209–218. https://doi.org/10.1145/1557019.1557048
[12] Jianlin Cheng, Zheng Wang, and Gianluca Pollastri. 2008. A neural network approach to ordinal regression. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. IEEE, 1279–1284.
[13] Coye Cheshire, Judd Antin, Karen S Cook, and Elizabeth Churchill. 2010. General and familiar trust in websites. *Knowledge, Technology & Policy* 23, 3-4 (2010), 311–331.
[14] François Chollet et al. 2015. Keras. https://keras.io.
[15] Brian Christian. 2012. The A/B Test: Inside the Technology That's Changing the Rules of Business. https://www.wired.com/2012/04/ff-abtesting/ [Online].
[16] Eyal Ert, Aliza Fleischer, and Nathan Magen. 2016. Trust and reputation in the sharing economy: The role of personal photos in Airbnb. *Tourism Management* 55 (2016), 62 – 73. https://doi.org/10.1016/j.tourman.2016.01.013
[17] U. Flick. 2013. *The SAGE Handbook of Qualitative Data Analysis*. SAGE Publications. https://books.google.com/books?id=R-6GAwAAQBAJ
[18] Diego Gambetta and Heather Hamill. 2005. *Streetwise: How Taxi Drivers Establish Customer's Trustworthiness*. Russell Sage Foundation. http://www.jstor.org/stable/10.7758/9781610442350
[19] Jennifer L Gibbs, Nicole B Ellison, and Chih-Hui Lai. 2011. First comes love, then comes Google: An investigation of uncertainty reduction strategies and self-disclosure in online dating. *Communication Research* 38, 1 (2011), 70–100.
[20] Niklas Luhmann. 2000. Familiarity, confidence, trust: Problems and alternatives. *Trust: Making and breaking cooperative relations* 6 (2000), 94–107.
[21] Xiao Ma, Jeffery T Hancock, Kenneth Lim Mingjie, and Mor Naaman. 2017. Self-disclosure and perceived trustworthiness of Airbnb host profiles. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 2397–2409.
[22] Andrew Maxwell, Runzhi Li, Bei Yang, Heng Weng, Aihua Ou, Huixiao Hong, Zhaoxian Zhou, Ping Gong, and Chaoyang Zhang. 2017. Deep learning architectures for multi-label classification of intelligent health risk prediction. *BMC bioinformatics* 18, 14 (2017), 523.
[23] Julia M. Mayer, Quentin Jones, and Starr Roxanne Hiltz. 2015. Identifying Opportunities for Valuable Encounters: Toward Context-Aware Social Matching Systems. *ACM Trans. Inf. Syst.* 34, 1, Article 1 (July 2015), 32 pages. https://doi.org/10.1145/2751557
[24] Roger C Mayer, James H Davis, and F David Schoorman. 1995. An integrative model of organizational trust. *Academy of management review* 20, 3 (1995), 709–734.
[25] D Harrison McKnight, Larry L Cummings, and Norman L Chervany. 1998. Initial trust formation in new organizational relationships. *Academy of Management review* 23, 3 (1998), 473–490.
[26] Panagiotis Mitkidis, John J McGraw, Andreas Roepstorff, and Sebastian Wallot. 2015. Building trust: Heart rate synchrony and arousal during joint action increased by public goods game. *Physiology & behavior* 149 (2015), 101–106.
[27] Helen Nissenbaum. 2011. A contextual approach to privacy online. *Daedalus* 140, 4 (2011), 32–48.
[28] Desmond C. Ong, Jamil Zaki, and June Gruber. 2017. Increased cooperative behavior across remitted bipolar I disorder and major depression: Insights utilizing a behavioral economic trust game. *Journal of abnormal psychology* 126 1 (2017), 1–7.
[29] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
[30] Ingrid Pettersson, Florian Lachner, Anna-Katharina Frison, Andreas Riener, and Andreas Butz. 2018. A Bermuda Triangle?: A Review of Method Application and Triangulation in User Experience Evaluation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 461, 16 pages. https://doi.org/10.1145/3173574.3174035
[31] Will Qiu, Palo Parigi, and Bruno Abrahao. 2018. More stars or more reviews? Differential effects of reputation on trust in the sharing economy. In *Proceedings of the CHI*.
[32] Julian B Rotter. 1967. A new scale for the measurement of interpersonal trust 1. *Journal of personality* 35, 4 (1967), 651–665.
[33] Jeffrey M. Rzeszotarski and Aniket Kittur. 2011. Instrumenting the Crowd: Using Implicit Behavioral Measures to Predict Task Performance. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 13–22. https://doi.org/10.1145/2047196.2047199
[34] Terri A Scandura and Ethlyn A Williams. 2000. Research methodology in management: Current practices, trends, and implications for future research. *Academy of Management journal* 43, 6 (2000), 1248–1264.
[35] Harald Van Heerde, Kristiaan Helsen, and Marnik G Dekimpe. 2007. The impact of a product-harm crisis on marketing effectiveness. *Marketing Science* 26, 2 (2007), 230–245.
[36] Maria Yakovleva, Richard R Reilly, and Robert Werko. 2010. Why do we trust? Moving beyond individual to dyadic perceptions. *Journal of Applied Psychology* 95, 1 (2010), 79.
[37] Dale E Zand. 1972. Trust and managerial problem solving. *Administrative science quarterly* (1972), 229–239.
[38] Michael Zürn and Fritz Strack. 2017. When More Is Better- Consumption Priming Decreases Responders' Rejections in the Ultimatum Game. *Frontiers in Psychology* 8 (2017), 2226. https://doi.org/10.3389/fpsyg.2017.02226