Predicting Dyad Payoffs in the Trust Game with Measures of Physiological Synchrony

A. Trust in Economics

Psychologists and economists study trust because it is an integral part of society and our daily lives. Trust is essential to personal well-being and economic success: customers trust that purchased goods will work as promised; businesses trust their employees to do honest work; and investors trust corporations to report accurate quarterly numbers (Evans & Krueger, 2010, p. 171). Nobel laureate Kenneth Arrow (1974) labeled trust "a lubricant for social systems" (Evans & Krueger, 2010, p. 171). Without trust, we could not survive.

"Trust is an important example of a generalized human tendency towards unselfish behavior" (Evans & Krueger, 2009, p. 1004). However, "neoclassical economic theory considers trust in strangers to be irrational" (Evans & Krueger, 2009, p. 1003). Neoclassical economic theory defines rational individuals as purely self-interested, and trusting and trustworthy behavior would be counter to one's own self-interest. The neoclassical economic model of decision making depicts a rational, purely self-interested man, "homo economicus." Behavior that deviates from self-interest is viewed as irrational, but this view of rationality fails to predict how people act (Berg et al., 1995). In a world with only rational actors "there is no reason to take the risk of trusting someone else. In a world of only self-interested individuals, every transaction would need to be enforced by legal contract or even brute force" (Evans & Krueger, 2009, p. 1004).

Fortunately, experiments conducted by behavioral economists has shown that "trust occurs even when it opposes self-interest" (Evans & Krueger, 2009, p. 1004; Smith, 2003). Research on trust demonstrates that the neoclassical model of rationality does not sufficiently describe human behavior (Evans & Krueger, 2009). Instead, the majority of individuals display social preferences, which indicates that at least some people regard the payoffs of others as positive utilities. "Individuals with benevolent preferences trust, cooperate, and reciprocate because it makes them feel good" (Evans & Krueger, 2009, p. 1007). For example, most care about others' well-being, and take the welfare of other individuals into account (Fehr & Krajbich, 2014, p. 193). "There is consistent empirical evidence that people often engage in prosocial behavior and prefer equality"(Evans & Krueger, 2009, p. 1008). Further, research shows that social preferences "may play a decisive role for aggregate social and economic outcomes in strategic settings" (Fehr & Krajbich, 2014, p. 193).

B. Physiological Synchrony as a Measure of Interpersonal Trust

This research aims to determine whether physiological synchrony can be found between two individuals who display interpersonal trust by choosing to mutually cooperate in a strategic setting. Interpersonal trust is simply trust that occurs between individuals. Rousseau et al. (1998) formally defines trust as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (p. 395). "Trust is not a behavior (e.g., cooperation), or a choice (e.g., taking a risk), but an underlying psychological condition that can cause or result from such actions" (Rousseau et al., 1998, p. 395). An earlier definition from Rotter (1967) defined interpersonal trust as an expectancy held by an individual or a group that the word, promise, verbal, or written statement of another individual or group can be relied upon (p. 651). Any definition of trust must contain two critical elements: vulnerability and expectation (Evans & Krueger, 2009, p. 1004). Trust is inherently risky because of the possibility of loss and regret. On the other hand, we trust because we expect a positive outcome. "The path-dependent connection between trust and risk taking arises from a reciprocal relationship: risk creates an opportunity for trust, which leads to risk taking" (Rousseau et al., 1998, p. 395). Trust cannot occur without taking a risk and accepting the possibility of betrayal.

Interpersonal autonomic physiology (IAP) is defined as the interconnection between individuals' physiological dynamics, as indexed by continuously measuring the autonomic nervous system (ANS) (Palumbo et al., 2017). One commonly used measure of physiology is electrodermal activity (EDA), which refers to the changes in the electrical conductance of the skin. "Changes in electrodermal activity (EDA) and skin conductance are related to changes in eccrine sweating which are, in turn, related to activity in the sympathetic branch (SNS) of the autonomic nervous system (ANS)" (Figner, Bernd & Murphy, 2011). During arousal the excitatory SNS is active, causing subtle changes in sweat production, which can be measured and analyzed in a laboratory setting. The most commonly studied property of EDA is skin conductance, which is quantified "by applying an electrical potential between two points of skin contact and measuring the resulting current flow between them" (Braithwaite, Watson, Jones, & Rowe, 2015). Increased sweat gland activity in response to SNS activation results in an increase in skin conductance.

Research on IAP and PS began over half a century ago, "when a series of studies found significant positive and negative correlations in the electrodermal activity (EDA) and heart rate (HR) of therapists and clients during therapy, interpreted as evidence of therapeutic rapport and empathy" (Palumbo et al., 2017, p. 110). Findings from IAP research indicate that physiological activity between two or more people can become associated or interdependent: physiological synchrony (PS) refers to "any interdependent or associated activity identified in the physiological processes of two or more individuals" (Palumbo et al., 2017, p. 100). Studies have expanded to include PS analyses of: mother-child relationships, couples, teammates, and groups (Palumbo et al., 2017).

C. Methodology

Economists study trust in laboratory experiments with games because they are "designed to represent certain critical features of real-life situations" (Evans & Krueger, 2009, p. 1004). Social scientists investigate the conditions relating to interpersonal trust using game theory (Rotter, 1967). "Experimental games are useful for studying trust because they provide an external, quantifiable measure of the underlying psychological state of trust" (Evans & Krueger, 2009, p. 1004). This exploratory study assesses interpersonal trust and the physiological responses associated with it, using the Trust Game to measure trust behavior.

The Trust Game was introduced in 1995 in order to control for alternative explanations of trusting behavior. The social optimal outcome of the game is mutual cooperation" (Evans & Krueger, 2009, p. 1006). It is played as follows: two participants are randomly and anonymously matched to play a one-shot game -- one as the investor or trustor, and one as the trustee. Both participants receive an initial endowment, *P*, for example, $10. The investor first decides to send some, all, or none of their $10 to the trustee. Every dollar sent by the investor is tripled. The first mover, or trustor, chooses between 'trust' - taking a risk with the expectation of earning more than $10, and the 'status quo,' $10 (Evans & Krueger, 2010). With the status quo, the game ends and both players receive their initial endowments. If the trustor chooses to take the risk, the game enters a second stage in which the second mover, or trustee, chooses between reciprocity and betrayal (Evans & Krueger, 2010).

The game is designed to capture the two critical features of trust: the trustor must assume the risk in anticipation of the trustee reciprocating. The trustee observes the amount sent and decides whether to keep the amount they received, or send some of the tripled amount back to the investor. Due to the tripling of the trustor's transfer, both players are better off collectively if more money is transferred. In other words, players can profit from friendly behavior.

C. 1. Participants

Researchers recruited 116 participants from the Claremont Colleges and the surrounding community (ages 25-50, mean age = 32, 52 females). After providing verbal and written consent, participants were asked to fill out a survey with demographic and personality trait questions. Participant payments varied depending on correctly answering narrative questions and participants' voluntary donations, with a maximum payoff of $38.

C. 2. Procedure

Two variations of the experimental protocol were developed: one session, Protocol A, with friends playing friends and one, Protocol B, with only strangers. Protocol A consisted of friends playing the trust game against each other, which resulted in a lack of within-subject variability. Thus, to eliminate this confound, this analysis is restricted to Protocol B, where participants had no relationship history.

Groups of four participants were asked to complete two four-hour experimental sessions, one week apart. Participants were first fitted with the electrophysiological instrumentation and continuous EDA data were collected for the duration of the experiment. Next, participants were awarded an endowment of $120 to be used during the experiment. Participants played the Trust game with two other participants, both as the first mover and second mover. To limit decision variability, there were only three second mover options: keep all the money, return the same amount the first mover sent, or split it 50/50. Participants were asked to decide how much to send back as the second mover for each possible first mover decision. In other words, participants chose how much to return if the first mover sent $40, $80, and $120. Figure 1 depicts the decision tree. The game order was counterbalanced between first and second mover order, and partner order across experimental sessions. Participant pairs played two rounds of the the Trust Game. Figure 2 shows the full experimental procedure for Protocol B, Session 1. At the end of the session, participants learned of their partners' decision for only one game that was randomly selected to determine payouts.

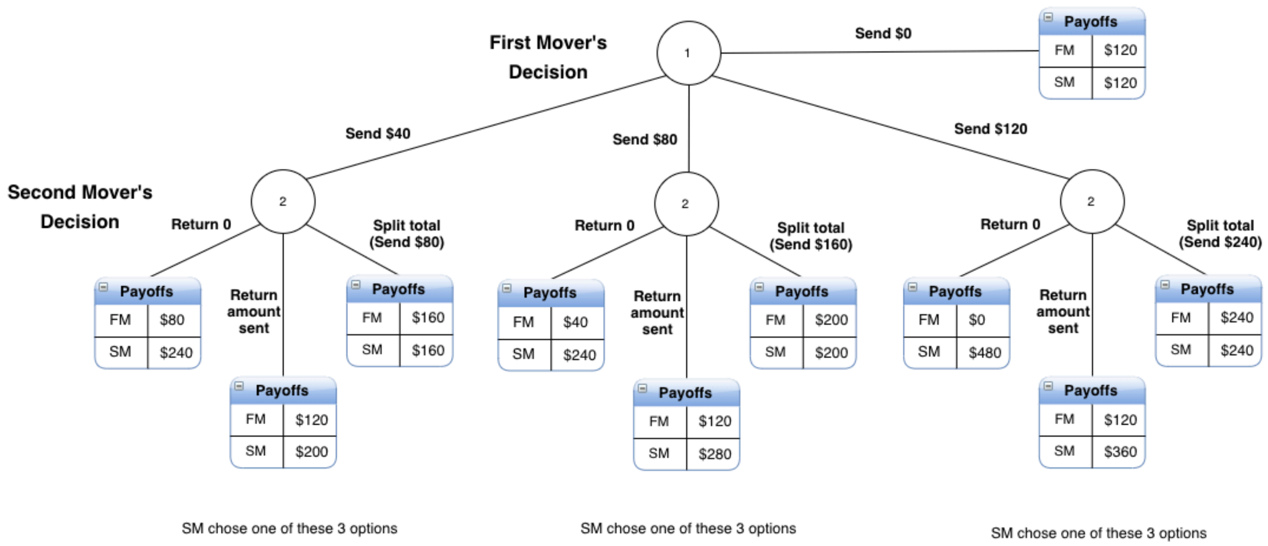

FIGURE 1: TRUST GAME DECISION TREE

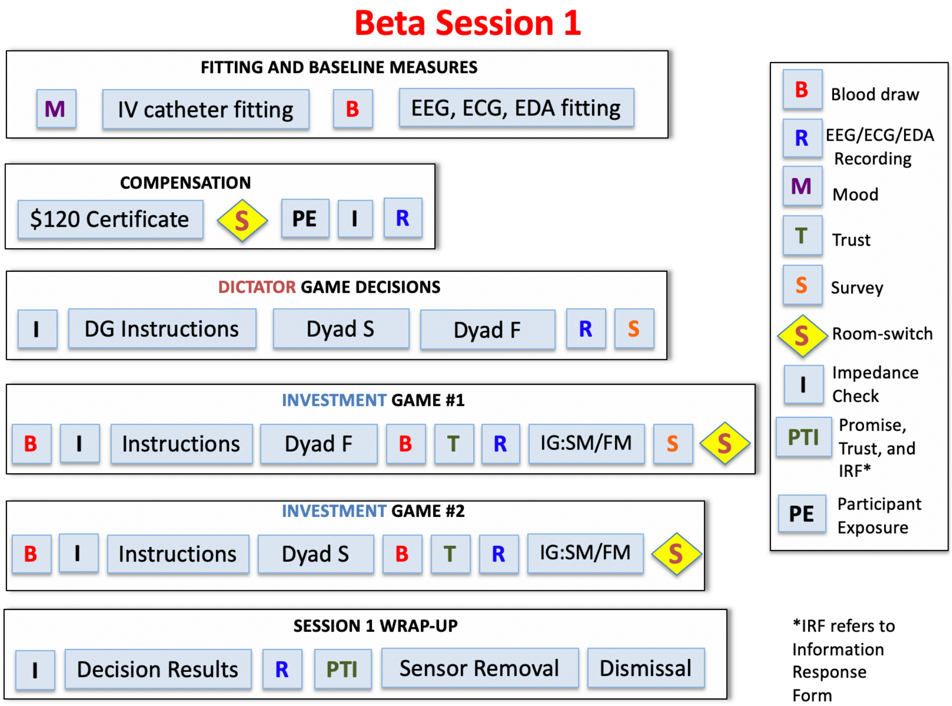

FIGURE 2: EXPERIMENTAL DESIGN: PROTOCOL B, SESSION 1

C. 3. EDA Data Collection

EDA data were collected at a sampling rate of 1 kHz using an MP150 data acquisition system and Acknowledge software (BIOPAC; Goltea, CA). A two minute baseline measurement period was collected while the participant was looking straight ahead at a blank screen and not engaged in any task. The EDA data analyzed for this paper were collected during the dyadic interactions, when participants were given a few minutes to talk, and potentially strategize.

D. Hypotheses

In this experiment, trust as a behavioral measure is defined as the participant pairs' total session payoff for the game that was randomly selected to determine payouts. Participant pairs who chose to strategize with each other in order to maximize their total payoffs exhibit the social optimal outcome of the game, mutual cooperation. These participants agreed to cooperate, and took the risk with no guarantee that their partner would honor their word.

This analysis aims to predict the amount paired strangers earn in the trust game, based on the dyad's physiological synchrony. Since participants were given time to talk to each other, is it possible that EDA analysis of participants who were more agreeable and cooperative with each other display evidence of psychological synchrony? Could pairs with higher psychological synchrony have also benefitted from earning the highest payoffs? Do individuals whose physiology synchronizes with their partner earn more? Are individuals who display less trust during the game less likely to show physiological concordance with their partner? Do the participant pairs with the lowest payoffs display no correlation or a negative correlation with their partner? Do these individuals display "opposite" or "negative" psychological synchrony? Can the pairs' physiological synchrony predict how much money they will earn together? Can an individuals' willingness and/or unwillingness to trust be classified based on their physiology? These hypotheses are summarized below:

H_1: Participant pairs with payoffs greater than the mean amount earned exhibited synchronous physiological signals, indicating trusting and trustworthy behavior.

H_2: Participant pairs with payoffs less than the mean amount earned had nonsynchronous physiological signals, indicating a lack of mutual trust. This may be no correlation or a negative correlation.

E. Preliminary Game Results & Statistics

Fifty-eight participant pairs played two rounds of the Trust Game, for a total of 116 game outcomes. Trust is measured as the sum of the pairs' payoff, the total session payout between partners, for a given game round. Participant pairs earned between \$160-\$680, with an average earning of \$343, ± \$118, and a median earning of \$340. Descriptive statistics of the 116 participant pairs' payouts are summarized in Table 1. Figure 3 shows the frequency distribution of pair payouts. Figure 4 shows a box plot of the pair payoffs.

	Pair Payoff
Mean	$343
Median	$340
SD	$118
Min	$160
25%	$280
50%	$340
75%	$480
Max	$680

TABLE 1: PAYOFFS OF 116 PARTICIPANT PAIRS'

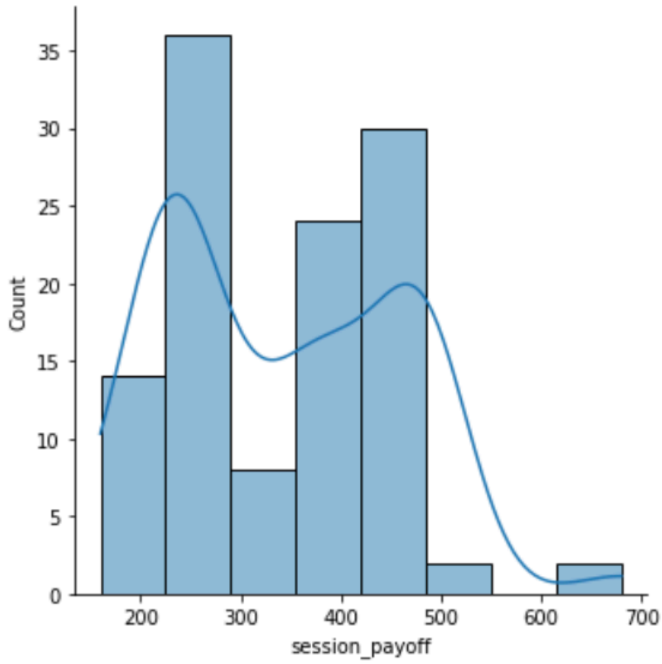

FIGURE 3: DISTRIBUTION OF 116 PARTICIPANT PAIRS' PAYOFFS

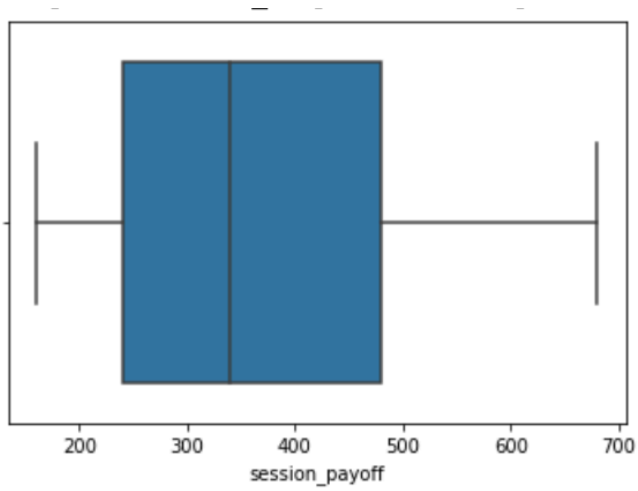

FIGURE 4: BOX PLOT OF 116 PARTICIPANT PAIRS' PAYOFFS

E. 1. Same vs Opposite Sex Pairs

Tables 2 and 3 show summary statistics for session payoffs by gender pairings: same vs opposite sex, and gender pairings (FF, FM, MF, MM).

	Same Sex	**Opposite Sex**
Count	56	60
Mean	$377	$312
Median	$400	$320
SD	$123	$105
Min	$160	$160
25%	$240	$240
50%	$400	$240
75%	$480	$400
Max	$680	$480

TABLE 2: PAYOFFS BY SAME VS OPPOSITE GENDER PAIRS

	FF	**FM**	**MF**	**MM**
Count	22	30	30	34
Mean	$378	$312	$312	$376
Median	$400	$320	$320	$400
SD	$123	$105	$105	$140
Min	$240	$160	$160	$160
25%	$240	$240	$240	$240
50%	$400	$240	$240	$360
75%	$480	$400	$400	$480
Max	$680	$480	$480	$680

TABLE 3: PAYOFFS BY GENDER PAIRS

The bar charts in Figures 5 and 6 show the distribution of payoffs based on gender pairs (same vs different sex) and gender pairings (FF, FM, MF, MM), respectively. The box plots in Figures 9 and 10 provide further insight on the distribution of payoffs based on gender pairs.

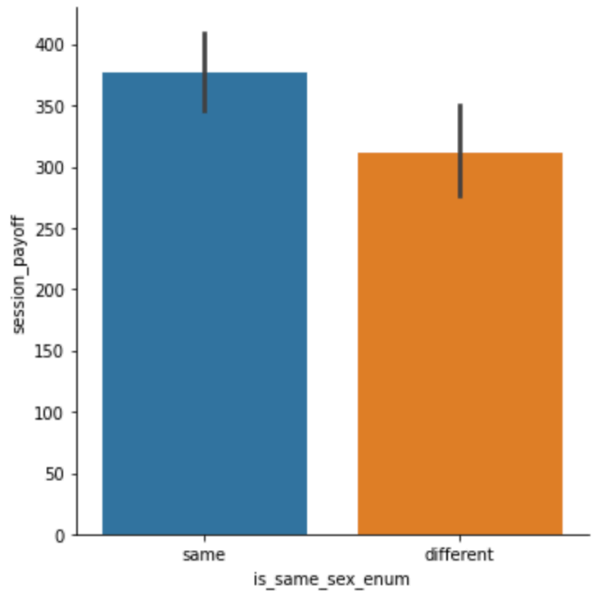

FIGURE 5: SAME VS OPPOSITE GENDER PAIRS' PAYOFFS

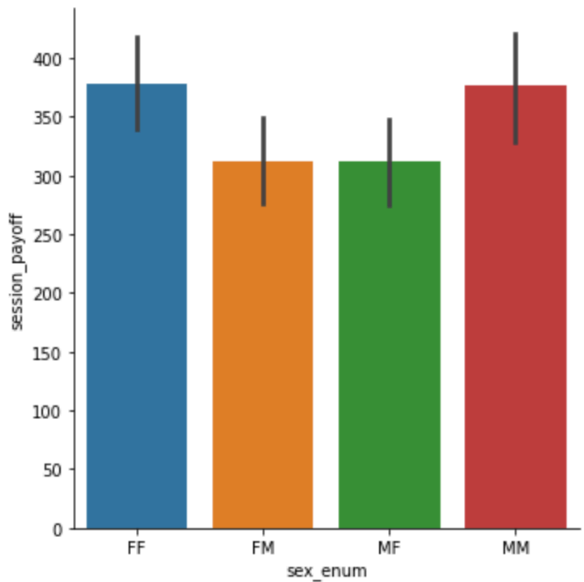

FIGURE 6: DIFFERENT GENDER PAIRS' PAYOFFS

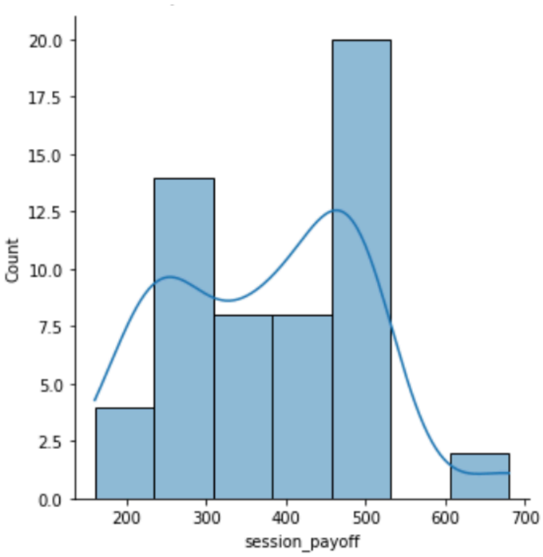

FIGURE 7: PAYOFFS OF SAME SEX PAIRS

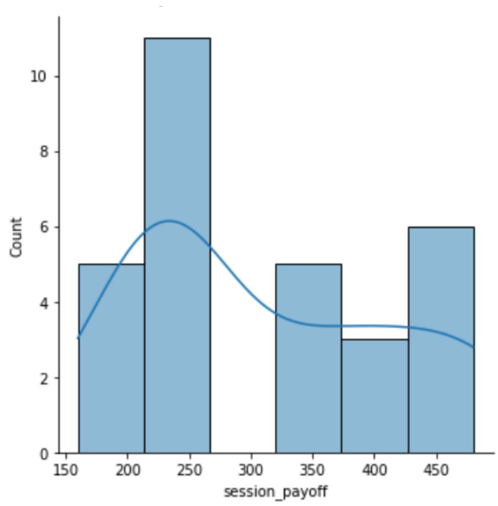

FIGURE 8: PAYOFFS OF OPPOSITE SEX PAIRS

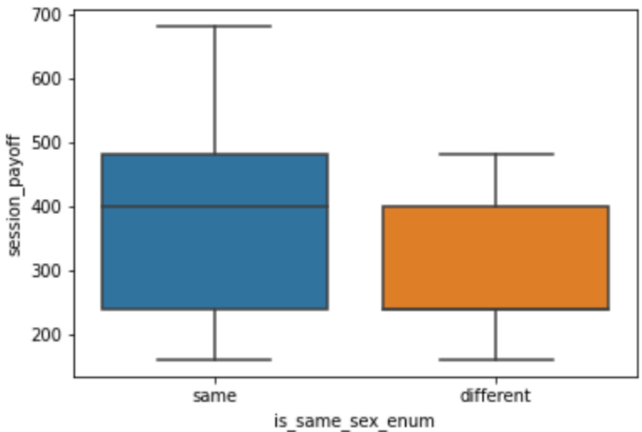

FIGURE 9: BOX PLOT OF PAYOFFS OF SAME SEX PAIRS

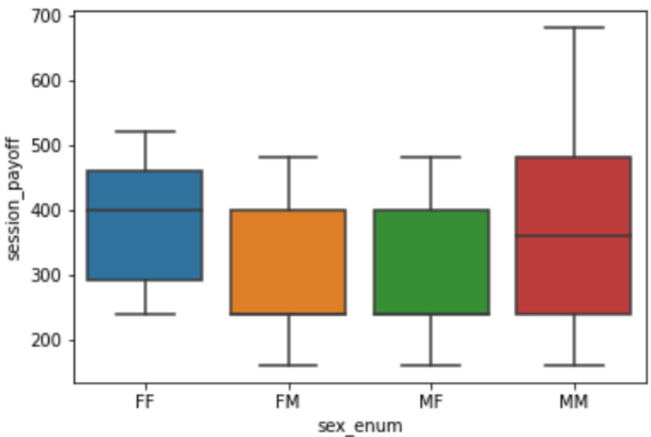

FIGURE 10: BOX PLOT OF OPPOSITE SEX PAIRS

A closer look at the payoffs by gender pairings reveals differences in earnings. Pairs of the same gender earned $65 more, on average. Payoffs differed significantly based on same versus opposite gender pairings (Welch's t-statistic = 2.58, p-value = 0.012). Figures 5 and 10 show that the payoffs differed significantly based on gender pairs (same vs different sex).

E. 2. Surveys and Demographics

Participants completed a demographics questionnaire and the following 8 surveys: the Social Values Orientation survey, the Emotion Regulation questionnaire, the Big Five personality questionnaire, the Positive and Negative Affect Schedule (PANAS), a Risk-Taking questionnaire, a Friendship Closeness survey, an Empathy survey, and the Stephenson Multigroup Acculturation Scale. Following their dyadic interactions, participants also indicated their subjective level of trust in their partner on a five- point scale. I plan to review this demographic and survey data after obtaining initial results analyzing only EDA data.

F. Proposed Analysis: Predicting Trust with Statistics of Physiological Synchrony

This research aims to predict the level of subject pairs' trust displayed in an initial round of a trust game, by measuring the strength of physiological synchrony with their EDA. To do this, dynamic systems analyses should be used, since they provide the ideal mathematical models for analyzing dyadic interactions (Gates & Liu, 2016). In this analysis, the participant pairs' physiology (EDA) is the independent variable, while the payoff amount earned by participant pairs, i.e. the total earned by the dyad, is the dependent variable.

The critical component of predicting interpersonal autonomic physiology is the type of statistical analysis used. The ideal methods are those that best match the research question(s) regarding the dyadic interaction(s) (Gates & Liu, 2016). Statistical procedures used to measure PS have evolved since the first studies on relied solely on Pearson correlations and cross-correlations. With advances in computing computer power, researchers have been applying new statistical techniques, with the goal of finding a measure that most accurately predicts PS, without over-fitting.

Determining an accurate measure of physiological synchrony is an ongoing process within IAP research. Various analytical methods have been developed to create an accurate 'statistic of physiological synchrony.' Overviews of statistical techniques for quantifying physiological synchrony were conducted by (Gates & Liu, 2016) and (Schoenherr et al., 2019), reviewing and dynamic systems approaches and linear time series analysis methods (TSAMs), respectively. (Gates & Liu, 2016) review the methods for dyad-level modeling, including: cross-correlation, vector autoregression (VAR), state-space models (SSMs), unified structural equation models (uSEMs), hidden Markov models (HMMs), state-space grids (SSGs) (GridWare), recurrence quantification analysis (RQA), differential equations, and spectral analysis. The linear time series analysis methods covered by (Schoenherr et al., 2019) include: cross-lagged correlation (CLC), cross-lagged regression (CLR), windowed cross-correlation (WCC), various versions of windowed cross-lagged correlations (WCLCs), and windowed cross-lagged regression (WCLR). It should be noted, "the analysis of multivariate, non-stationary, intensive time series of physiology is wrought with complexities, as these data violate a number of assumptions of parametric statistics (e.g., stationarity, independence of measures)" (Palumbo et al., 2017, p. 126).

F. 1. Inter-Subject Correlation (ISC) Analysis

The first proposed analysis for this research is Inter-Subject Correlation (ISC), which "provides a measure of the across-subject reliability of physiological responses by quantifying the commonalities of the response time courses among individuals" (Golland et al., 2014, p. 3). ISC analysis is an increasingly popular method to measure synchrony between individuals or groups. It is commonly used in physiological and neurological research, including EKG, EEG, and fMRI. For this research, it will be used to determine whether continuous ANS signals, measured by EDA, exhibit time-locked response patterns that are consistent across individuals (Golland et al., 2014).

ISC analysis calculates linear correlations between participant pairs and derives an ISC statistic from these correlations to measure the level of physiological synchrony (Lahnakoski & Chang, 2021). ISC is a version of a windowed cross-lagged correlation (WCLC) and also requires four parameters to be specified: window size, window increment, maximum lag, and lag increment. A participant pairs' ISC score is defined as $r_{i,j}= r(x_i, \bar{x}_j)$, where x_i is the maximal correlation within

±10 lags, between r_i, the response time course of individual i, and $\bar{x}_j$, the average response time course of individual j (Golland et al., 2014). As a robustness check, bootstrapping will be used to ensure distributions show statistical significance (distributions of real ISC scores vs bootstrapped ISC scores).

In order to run the ISC analysis, an ISC score for each participant pair needs to be calculated. The base ISC code for one participant pair is complete, and a program will be written to automatically run the EDA files through the ISC code, matching the interaction pairs & outputting the ISC scores. The next step is to test the null of the ISC scores, and export the bootstrapped data for each participant pair.

F. 2. Dynamic Time Warping (DTW) & K-Nearest Neighbors (KNN) Analysis

The second proposed analysis is two-fold. First, dynamic time warping (DTW) will be used to measure the similarity between the pairs' EDA time series. The DTW algorithm computes the minimum path that will match the two series, by minimizing a measure of distance of the set of functions to their warped average (Cheong, 2019). One advantage of DTW is that it can also deal with series of different lengths (Cheong, 2019). The Python package, dtw, package will be used to conduct the DTW analysis. DTW analysis produces participant pair scores that can be classified using K-Nearest Neighbors (KNN), which will cluster the scores based on the minimum distance of the scores.

F. 3. EDA Data Cleaning

EDA waveforms were visually inspected for brief periods of signal loss, and data drop-offs shorter than 1s in length were replaced with averages from adjacent parts of the waveform. Next, a 10-Hz low-pass filter was applied to the waveform to remove high-frequency noise (Norris et al., 2007), and a square root transformation was applied to adjust for skew inherent to skin conductance data (Dawson, M. E., et al., 1989; Figner, Bernd & Murphy, 2011). After extracting all participant data, the series of interest - the interaction session - was baseline corrected, normalized, and standardized.

G. Conclusion: Comparing Measures of Physiological Synchrony

The final question this research will aim to answer is, which statistic of physiological synchrony most accurately predicts mutual trust, and thus the largest earnings? A comparison of the two methods will comprise the conclusion chapter.

References

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, *10*(1), 122-142.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2015). A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments. *Technical Report, 2nd Version* https://www.biopac.com/wp-content/uploads/EDA-SCR-Analysis.pdf.

Cheong, J. H. (2019). Four Ways to Quantify Synchrony Between Time Series Data.

Dawson, M. E., Filion, D. L., & Schell, A. M. (1989). Is elicitation of the autonomic orienting response associated with allocation of processing resources? *Psychophysiology*, *26*(5), 560-572.

Evans, A. M., & Krueger, J. I. (2010). Elements of trust: Risk and perspective-taking. *Journal of Experimental Social Psychology*, *47*, 171-177.

Evans, A. M., & Krueger, J. I. (2009). The Psychology (and Economics) of Trust. *Social and Personality Psychology Compass*, *3*(6), 1003-1017.

Fehr, E., & Krajbich, I. (2014). Social Preferences and the Brain. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics: Decision Making and the Brain* (pp. 193-218). Elsevier Inc.

Figner, Bernd, & Murphy, R. O. (2011). Using skin conductance in judgment and decision making research. In *A handbook of process tracing methods for decision research*. New York, NY: Psychology Press.

Gates, K. M., & Liu, S. (2016). Methods for Quantifying Patterns of Dynamic Interactions in Dyads. *Assessment*, *23*(4), 459-471.

Golland, Y., Keissar, K., & Levit-Binnun, N. (2014). Studying the dynamics of autonomic activity during emotional experience. *Psychophysiology*, *51*(11), 1101-1111.

Lahnakoski, J., & Chang, L. (2021). Intersubject Correlation. 2020, from https://naturalistic-data.org/content/Intersubject_Correlation.html.

Norris, C. J., Larsen, J. T., & Cacioppo, J. T. (2007). Neuroticism is associated with larger and more prolonged electrodermal responses to emotionally evocative pictures. *Psychophysiology*, *44*(5), 823-826.

Palumbo, R., Marraccini, M., Weyandt, L., Wilder-Smith, Oliver, A McGee, H. et al. (2017). Interpersonal Autonomic Physiology: A Systematic Review of the Literature. *Personality and Social Psychology Review*, *21*(2), 99-141.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality*, *35*(4), 651-665.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. F. (1998). Not So Different After All: A Cross-Discipline View of Trust. *The Academy of Management Review*, *23*(3), 393-404.

Schoenherr, D., Paulick, J., Worrack, S., Strauss, B. M., Rubel, J. A., Schwartz, B. et al. (2019). Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony. *Behav Res Methods*, *51*(1), 361-383.

Smith, V. L. (2003). Constructivist and Ecological Rationality in Economics. *American Economic Review*, *93*, 465-508.