

Predicting Trusting and Trustworthy Behavior

By

Lena Arai

Claremont Graduate University
2021

Approval of the Dissertation Committee

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Lena Arai as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy in Economics and Information Systems & Technology.

Paul Zak, Co-Chair

Claremont Graduate University

Professor of Economics

Yan Li, Co-Chair

Claremont Graduate University

Professor of Information Systems & Technology

Jorge Barraza

University of Southern California

Professor of (Consumer) Psychology - Ask Jorge

George D. Montañez

Harvey Mudd College

Professor of Computer Science

Abstract

Predicting Trusting and Trustworthy Behavior

By

Lena Arai

Claremont Graduate University: 2021

First sentence

Example:

"This work" develops the best linear model of residential real estate prices for 2003 through 2009 in Los Angeles County. It differs from other studies comparing models for predicting house prices by covering a larger geographic area than most, more houses than most, a longer time period than most, and the time period both before and after the real estate price boom in the United States. In addition, it open sources all of the software. We test designs for linear models to determine the best form for the model as well as the training period, features, and regularizer that produce the lowest errors.

I compare the best ...

and point to directions for further research.

and discuss areas of...

Dedication

tfhgjhjhjj

Acknowledgements

Small Roman numeral, page v or vi

Table of Contents

Broken

Predicting Trust using Measures [Statistics] of Physiological Synchrony	2
Introduction	2
Trust in Economics	2
Physiological Synchrony as a Measure of Behavioral Trust	3
Predicting Interpersonal Trust	3
Comparing Statistics of Physiological Synchrony	4
The Economics of Trust	5
A Positive Economic Science	5
Rational Choice Theory (Incomplete)	9
An Inductive Economic Science	10
The Neoclassical Theory of Trust in Economics	12
Interpersonal Trust	13
Experimental Economics and Experimental Game Theory	15
The Trust Game	17
Social Preferences	22
Physiological Synchrony as a Measure of Trust	25
Physiology and the Autonomic Nervous System (ANS)	25
Physiological Response to Decision Making (Incomplete)	26
Physiological Synchrony	29
Mechanism(s)? of Physiological Synchrony (Incomplete)	31
Statistics of Physiological Synchrony (Incomplete)	34
Pearson Correlation	36
Time Lagged Cross Correlation	37
Windowed Time Lagged Correlation	38
(1)	ISC
39	
Instantaneous Phase Synchrony (?)	40
Differential Equation Modeling	40
Dynamic Time Warping	40
MPdist Algorithm (?)	41
Predicting Dyad Payoffs in the Trust Game with Measures of Physiological Synchrony	42
Methodology (Incomplete)	
42	
Participants (Incomplete)	42
Procedure	43
	Data Collection (Revisit Surveys)
45	

	Predicting Interpersonal Trust (Incomplete)	
46		
Hypotheses (Incomplete)		48
	Data / Cleaning & Methods (Introduction)	
48		
	EDA Data Cleaning	
49		
	Game Results and Statistics	
51		
Results		51
	Inter-Subject Correlation (ISC) Analysis	
51		
EDA Inter-Subject Correlation (ISC)		52
ISC Results		52
	Dynamic Time Warping (DTW) Analysis	
52		
	K-Nearest Neighbors (KNN) Analysis	
52		
	Recurrent Neural Network (RNN) ?	
52		
(Conclusion) Comparing [Discussion of] Statistics of Physiological Synchrony		54
	Comparing Measurements [Statistics] of Synchrony	
54		
	Limitations	
54		
	Discussion / Conclusion	
54		
Figure 1: An example of the trust game, from Evans and Krueger (2010).		18
Figure 2: Cost, benefit, and temptation in the trust game, from Evans and Krueger (2010).		19
Figure 3: Example of a dictator game		21
Figure 4: Example of a trust game		22
Figure 5: Diagram of the proposed steps involved in decision making, from		28
Figure 6: Dictator game Decision Tree		44
Figure 7: Trust game Decision Tree		44
Figure 8: Experimental Design: Protocol B, Session 1		45

i.	Predicting Trusting and Trustworthy Behavior	1
ii.	Approval of the Dissertation Committee	3
iii.	Abstract	4
	Predicting Trust using Measures [Statistics] of Physiological Synchrony	13
	Introduction	13
	Trust in Economics	13
	Physiological Synchrony as a Measure of Behavioral Trust	14
	Predicting Interpersonal Trust	14
	Comparing Statistics of Physiological Synchrony	15
	The Economics of Trust	16
	A Positive Economic Science	
	Rational Choice Theory (Incomplete)	20
	An Inductive Economic Science	21
	The Neoclassical Theory of Trust in Economics	23
	Interpersonal Trust	24
	Experimental Economics and Experimental Game Theory	26
	The Trust Game	28
	Social Preferences	33
	Physiological Synchrony as a Measure of Trust	36
	Physiology and the Autonomic Nervous System (ANS)	36
	Physiological Response to Decision Making (Incomplete)	37
	Physiological Synchrony	40
	Mechanism(s)? of Physiological Synchrony (Incomplete)	42
	Statistics of Physiological Synchrony (Incomplete)	45
	Pearson Correlation	47
	Time Lagged Cross Correlation	48
	Windowed Time Lagged Correlation	49
	(1) ISC	50
	Instantaneous Phase Synchrony (?)	51
	Differential Equation Modeling	51
	Dynamic Time Warping	51
	MPdist Algorithm (?)	52
	Predicting Dyad Payoffs in the Trust Game with Measures of Physiological Synchrony	53
	Methodology (Incomplete)	
	53	
	Participants (Incomplete)	53
	Procedure	54

56	Data Collection (Revisit Surveys)
57	Predicting Interpersonal Trust (Incomplete)
Hypotheses (Incomplete)	59
59	Data / Cleaning & Methods (Introduction)
60	EDA Data Cleaning
62	Game Results and Statistics
Results	62
62	Inter-Subject Correlation (ISC) Analysis
EDA Inter-Subject Correlation (ISC)	63
ISC Results	63
63	Dynamic Time Warping (DTW) Analysis
63	K-Nearest Neighbors (KNN) Analysis
63	Recurrent Neural Network (RNN) ?
(Conclusion) Comparing [Discussion of] Statistics of Physiological Synchrony	65
65	Comparing Measurements [Statistics] of Synchrony
65	Limitations
65	Discussion / Conclusion

List of Figures

Figure 1: An example of the trust game, from Evans and Krueger (2010).	29
Figure 2: Cost, benefit, and temptation in the trust game, from Evans and Krueger (2010).	30
Figure 3: Example of a dictator game	32
Figure 4: Example of a trust game	33
Figure 5: Diagram of the proposed steps involved in decision making, from	39
Figure 6: Dictator game Decision Tree	55
Figure 7: Trust game Decision Tree	55
Figure 8: Experimental Design: Protocol B, Session 1	56

Introduction (Optional): First Arabic page number, start with page 1

I. Predicting Trust using Measures [Statistics] of Physiological Synchrony

I. 1. Introduction

- **1-A. Overview:** Briefly explain why the study is being undertaken and what main questions or foreshadowed problems will be addressed. Do this in a general manner, because it will be done more specifically in the following sections.

- Finally, compare the analyses used to measure physiological synchrony

I. 2. Trust in Economics

- **1-B. Statement of the Problem:** Discuss the problem to be addressed in the research— the gaps, perplexities, or inadequacies in existing theory, empirical knowledge, practice, or policy that prompted the study. The problem may be a theory that appears inadequate to explain known phenomena, the lack of empirical data on a potentially interesting relationship between X and Y, or a common practice that appears ineffective. First state the problem generally, and then state the specifics that your research will address. In quantitative research, the specifics will include the constructs studied.

That your favorite reading program is rarely used in schools does not constitute a problem; widespread impaired reading in inner-city elementary schools is a problem. That your favorite conjectures are not represented in prevailing theory does not constitute a problem; that the theory does not explain applicable phenomena is a problem. That a certain group has been omitted from prior studies can indeed constitute a problem, because theory, policy and practice have not been shaped by knowledge of that group.

Problems usually have underlying causes that may be well-known or the subject of speculation. They also have consequences that are often apparent. You should briefly discuss these causes and consequences.

1-C. Purpose

The purpose of research is to acquire knowledge to address the problem or certain aspects of it. Quantitative research tries to fulfill that purpose by answering questions and/or testing hypotheses. Qualitative research tries to fulfill that purpose by starting with foreshadowed problems, conjectures, or exploratory questions. Mixed-methods research may use both approaches.

I. 3. Physiological Synchrony as a Measure of Behavioral Trust

1-D.2. Significance of the Study: Discuss the potential significance of the research. Significance comes from the uses that might be made of your results—how they might be of benefit to theory, knowledge, practice, policy, and future research. The potential significance should be based upon your literature review in Chapter 2.

<i>Quantitative Research: Research Questions and/or Hypotheses</i>	<i>Qualitative Research: Foreshadowed Problems, Conjectures, or Exploratory Questions</i>
Present the research hypotheses stated fully—exactly as you state them in Chapter 3.	Present the foreshadowed problems, conjectures, or exploratory questions stated in 3-B below. State them fully—exactly as you state them in Chapter 3.

1-E. Conceptual Framework: Briefly summarize the theoretical foundation or conceptual framework(s) derived from the literature review that is reported in Chapter 2. Conceptual framework is the theoretical foundations helping us understand the problem situation and its dynamics. It includes your study variables and depicts the established or predicted relationship(s) among these variables. You may adopt an existing conceptual framework or develop your own modified version based on the literature review.

- Motivation for research
- Answer question: "So what?" --> Importance, significance of answering research question

The study of physiological synchrony began in the 1950s with the study of patient-therapist relationships and analyses of physiological responses of therapists and their clients during interviews (Levenson & Gottman, 1983).

I. 4. Predicting Interpersonal Trust

- **Summary of Methodology:** Briefly summarize the methodology of the research that is described fully in Chapter 3.

- Overview of experiment & analyses

I. 5. Comparing Statistics of Physiological Synchrony

- Overview of Ch. 5 -

II. The Economics of Trust

II. 1. A Positive Economic Science

Milton Friedman's seminal 1953 paper, *The Methodology of Positive Economics*, set precedent in neoclassical economic theory. Friedman (1953) advocated a framework for developing a positive science, with the ultimate goal of developing theories or hypotheses that yield valid and meaningful predictions. He sought to establish methodological criteria for constructing a positive economic science and deciding whether a hypothesis or theory should be tentatively accepted by the field. Further, he sought to clarify which aspects of theory development are relevant and which are immaterial/insignificant.

According to Friedman (1953), hypotheses are "designed to abstract essential feature of complex reality" (p. 6). If sound economic theories seek to abstract essential features of reality, then our theories should accurately represent reality. Designing theories with features that do not reflect reality has led to the current reputability crisis in economics (our field). It is too easy for a layperson to dismiss economic theories on the basis of their improbability. As economists we are doing the public and our field a disservice by continuing to develop theories with no basis in reality.

Friedman (1953) sought to clarify a common misunderstanding of the role of empirical evidence in testing economic hypotheses. His fundamental argument centers (centered b/c he is dead?) around which criteria should be deemed necessary to assess (for assessing?) the validity of a hypothesis. Friedman (1953) asserts that a theory's performance "is to be judged by the precision, scope, and conformity with experience of the predictions it yields." The evidence of a theory should conform to the implications of the hypothesis and accurately predict an outcome within "the class of phenomena it is intended to explain" (Friedman, 1953, pp. 7, 11) In other words, the principle measurement of the validity of an economic hypothesis or theory is its implications and predictive power. He refers to this as testing a hypothesis by its implications.

The methodological guidelines affirmed by Friedman (1953) sought to clarify the process by which economists construct hypotheses and test their validity. According to Friedman (1953), "...to suppose that hypotheses have not only 'implications' but also 'assumptions' and that the conformity

of the these 'assumptions' to 'reality' is a test of the validity of the hypothesis *different from* or *additional to* the test by implications...is fundamentally wrong and productive of much mischief" (p. 11).

Friedman (1953) asserts that the realism or plausibility of a theory's assumptions are irrelevant and to imply their relevance "promotes misunderstanding about the significance of empirical evidence for economic theory, produces a misdirection of much intellectual effort devoted to the development of positive economics, and impedes the attainment of consensus on tentative hypothesis in positive economics" (p. 11). In sum, assumptions and their conformity to reality are irrelevant in determining a theory's validity.

As the paper progresses, Friedman (1953) expands his line of reasoning against the importance of realistic assumptions. He questions whether theories have "assumptions" at all; if they do, there is actually an inverse relationship between the realism of an assumption and a theory's validity/predictive power. Unrealistic assumptions provide/lead to the most accurate/predictive theories. Since theories are meant to abstract from the complexity of a phenomenon and produce valid predictions, "a hypothesis must be descriptively false in its assumptions...[and] its very success shows them to be irrelevant for the phenomena to be explained" (p. 11). In fact, significant hypotheses will have "assumptions that are wildly inaccurate...representation of reality [and]...the more significant the theory, the more unrealistic the assumptions..." (p. 11). How can a theory with unrealistic assumptions result in valid predictions? How can hypotheses contain "wildly inaccurate description representations of reality," and provide "sufficiently good approximations for the purpose at hand?" Is there an example of this in economic theory? Friedman (1953) utilizes an unconvincing example from physics to justify/prove his point, rather than an example from economics, which only weakens his argument and leaves the reader with more questions than answers.

According to Friedman (1953), the realism of an assumption cannot and should not be judged separately from the theory's predictive power/ability to accurately predict outcomes. The relevant assessment of an assumption is whether the theory is a sufficiently good approximation for the purpose at hand; this can only be measured by whether or not the theory works. The latter

is a test of validity (he would argue a poor test of a theory's validity), the former a test of the implications or predictions. Additionally, he argues/contends that "the two supposedly independent tests" simply (in fact) reduce to one test (p. 12). A theory's predictive ability (i.e., correct implications) is paramount. If a theory yields sufficiently accurate predictions, then it's unrealistic assumptions should be ignored. Further, unrealistic assumptions often result in more accurate theories.

Friedman (1953) argues that the concept of "the assumptions of a theory" is an ambiguous concept, but every theory is based upon a set of assumptions or premises. A formal argument contains an internal structure comprised/consisting of a set of assumptions or premises, a method of reasoning or deduction, and a conclusion. A premise is a stated assumption that something is true. An assumption is unstated and when added to the premise(s) prove the conclusion. In other words, an assumption is contained in the premise. By definition, hypotheses and generalizations are deduced from assumptions (EconomicConcepts.com, 2015). Thus, an assumption must be true for the given evidence to logically lead to the stated conclusion.

Further, arguments are considered to be either valid or invalid. A valid argument consists of premises that are true, hence, by definition the conclusion must be true. A valid argument is either sound or unsound. A sound argument is logically correct and consists of true premises. The strongest formal argument is both sound and valid. The proof of a conclusion depends on the truth of both the premises and the validity of the argument. Evidence is required to determine if the conclusion is an accurate reflection of what-is.

Reasoning is the process of using existing knowledge to draw conclusions, make predictions, or construct explanations (Butte College, 2016). There are two predominant methods of reasoning for developing theories: deductive and inductive. Deductive reasoning starts with a set of axioms - universally accepted principles - about an economic phenomena and relies on these principles to explain said phenomena (an event) (Fehr & Krajbich, 2014). It employs/utilizes economic principles and theories that have been empirically verified through observation, research, and critical analysis (Fehr & Krajbich, 2014). Deductive reasoning/logic can be described as a top-down approach, since a conclusion is reached reductively by applying general rules, and narrowing

the domain under consideration until only the conclusion(s) is left (EconomicConcepts.com, 2015). In other words, it consists of deriving the conclusion(s) from general truths. Conversely, an inductive conclusion is reached by generalizing specific cases to general rules. It begins with a problem or question and proceeds to form a general principle based on the evidence observed in the real-world of economic activity (Guru, n.d.). Hence, inductive reasoning/logic is considered a bottom-up approach. One favorable aspect of deductive reasoning/logic is the use of mathematical techniques in deducing theories. The major drawback of deduction is the development of highly sophisticated theoretical models based on highly unrealistic assumptions, which do not have any practical significance (Guru, n.d.). Scientists use induction to develop hypotheses and theories (Bradford, 2017). The findings from inductive reasoning then form economic theories used in deductive analysis. Deductive reasoning allows economists to apply the theories to specific situations (Bradford, 2017).

Deductive arguments are evaluated on the basis of their *validity* and *soundness*. By definition, deductive arguments are *valid* if the premise(s) is true. An argument derived deductively is considered *valid* when the premise(s) is true because the conclusion is also true - the conclusion must be true if the premise(s) is true. It is impossible for the argument's premises to be true and the conclusion false because/since the truth of the conclusion is implicit in the premise(s). An argument is deemed *sound* if it is *valid* and the premise(s) is true. It is possible for a deductive argument to be logically *valid* but not *sound*. Therefore, deductive conclusions are accurate *as long as and only if* the underlying assumptions are valid (EconomicConcepts.com, 2015).

In deductive logic the premise(s) ensure the conclusion is true. Inductive reasoning does not have this property. Instead, an inductive argument may rest on a true premise(s) and be *valid*, but its conclusion can be false. The premise(s) of an inductive argument indicate support for the conclusion but do not entail it - the premise(s) suggest the truth but do not ensure the conclusion is true. The premise(s) from which inferences are drawn may not hold in all circumstances, time, and places - a conclusion derived deductively may not be applicable universally (EconomicConcepts.com, 2015). Often, assumptions(premises) are based on half truths or have no relation to reality, making conclusions drawn from such assumptions misleading

(EconomicConcepts.com, 2015).

Economists have relied predominantly on deductive reasoning to develop theories, which explains why the field maintains many erroneous theories. This is because the general truths that economists rely on to deduce their theories are not (actually) general truths; the majority of economists have preconceived notions and biases about economic issues. (This is due to the fact that the general truths economists rely on to deduce their theories are not (actually) general truths.) For example, economists from (belonging to) the neoclassical school (e.g., Ricardo, Senior, Cairnes, J.S. Mill, Malthus, Marshall, Pigou) applied deductive reasoning to derive their theories/hypotheses (EconomicConcepts.com, 2015). The validity of Friedman's (1953) arguments rest (solely) on his use of deductive logic. Most economists are well-educated, philosophical, and often mathematically inclined. This does not describe the majority of economic participants(agents). Neoclassical economists do not believe that individuals tend to act irrationally. It is a simple case of a myopic viewpoint.

Neoclassical economists adopted Friedman's stance on theory development and his framework for developing hypotheses has persisted among economists today (...hypotheses still remains popular today). The deductive method employed/utilized by neoclassical economists has led to many false conclusions due to their reliance on imperfect and incorrect assumptions (EconomicConcepts.com, 2015). The result has been a field rich with theories based predominantly on unrealistic (i.e., false) assumptions.

II. 2. Rational Choice Theory (Incomplete)

Rational choice theory states that the behavior of individuals (individual actors) is purely self-interested (Blume & Easley, 2018). In general, rational decision making consists of choosing the alternative that an individual prefers most. The premise of the theory is methodological individualism, the idea that social phenomena is explained in terms of individual actions, which in turn is explained by individuals' motivation (Blume & Easley, 2018). Individuals make choices based on their preferences and constraints. Further, rational choice theory assumes that individuals have full or perfect information about the choice set.

The most prominent version of rational choice theory is expected utility (EU) theory, which states that individuals choose not the highest expected value, but rather maximize their expected utility according to some utility function. It quickly became a dominant paradigm for choice under uncertainty (Blume & Easley, 2018). Expected utility theory stipulates four axioms that define a rational decision maker: completeness, transitivity, independence, and continuity (von Neumann & Morgenstern, 1953). Completeness assumes that an individual can compare all pairs of alternatives and decide between any two. The completeness axiom is the focus of this analysis because it assumes perfect information about the alternatives -- all parties have full information regarding the actions chosen by others.

- Rational Behaviour (Sen, 2018)

- Rationality, History of the Concept (Sen, 2018)

Extra:

- Too many economic theories have a narrow "class of phenomena the hypothesis is designed to explain" because they are formed (constructed) deductively.

Economic theories have been derived deductively

Economic theories derived deductively

A deductive argument can be logically *valid* but not *sound*.

II. 3. An Inductive Economic Science

(As discussed in the previous section,) inductive reasoning is the process of using particular facts to extrapolate to a general principle or theory. Conclusions are made based on current knowledge and predictions. An inductive argument's premise(s) is viewed as supplying evidence for the truth of the conclusion (Irving, Carl, Daniel, & Bradley, 2007).

Deductive reasoning seeks certainty while inductive reasoning seeks to formulate a probability (Harry, Supriya, & Shane, 2002). While the conclusion of a deductive argument is certain, provided the premise(s) is true, the truth of the conclusion of an inductive argument is probable, based upon the evidence given (Irving, Carl, Daniel, & Bradley, 2007). Inductive

reasoning/logic does not rely on universals holding over a closed domain of discourse to draw conclusions, so it can be applicable even in cases of uncertainty (Bart, 1990).

Uncertainty is an inherent property of inductive logic. In other words, inductive reasoning allows for the possibility that the conclusion is false and there is another explanation, even if all the premises are true (Henderson, n.d.). Inductive arguments are deemed(rated) *strong* or *weak* along a continuum, which describes how *probable* it is that the theory(conclusion) is true (Herms, n.d.). An inductive argument is strong in proportion to the probability that its conclusion is correct. Its reliability varies proportionally with the evidence. This is known as inductive probability. In addition, inductive arguments/conclusions are not defined as true or false. Instead, they are considered cogent - the evidence on the subject is/appears/seems complete, relevant, and generally convincing, therefore the conclusion is probably true - or not cogent (Butte College, 2016).

In non-axiomatic systems, such as reality, deductive certainty is impossible. This leaves inductive logic as the primary route to probabilistic knowledge of such a system (Garrath, 2016). Induction involves gathering evidence, seeking patterns, and forming a hypothesis or theory to explain what is seen (Butte College, 2016). Detailed data are collected regarding a specific economic phenomenon and an effort is made to arrive at generalizations that follow from the observations collected (Guru, n.d.). For example, a common approach to induction is inductive generalization, whereby a theory is developed empirically. This proceeds from a premise(s) about a representative sample to a conclusion about the population. This is accomplished by collecting data about the variables related to an (economic) phenomenon and thinking about the possible functional relationships between the relevant variables (Guru, n.d.). Hypotheses can also be tested by collecting data about a particular economic phenomenon by running an experiment, analyzing the data collected, interpreting the results, and drawing general conclusions from the (data) analysis. Thus, inductive theories are drawn from data (Bradford, 2017). How much the premises support the conclusion depends upon three factors: the number in the sample, the number in the population, and the degree to which the sample represents the population. There is no way to know that all the possible evidence has been gathered, and that there exists no further bit of unobserved evidence that might invalidate a hypothesis (Butte College, 2016).

While inductive reasoning cannot yield a certain conclusion, it can increase human knowledge because it is *ampliative* (Butte College, 2016). Deductive reasoning cannot increase human knowledge because it is *nonampliative* and the conclusions yielded by deduction are tautologies, statements that are contained within the premises and virtually self-evident (Butte College, 2016). Thus, deduction can be used to make observations and expand implications, but it cannot make predictions about the future or as-yet, unobserved phenomena (Butte College, 2016). Inductive reasoning can be used to make predictions about future events, making it a more powerful method of reasoning.

II. 4. The Neoclassical Theory of Trust in Economics

The standard, 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). Fortunately, behavioral economists have proposed theories of rationality that are more consistent with behavior (Evans & Krueger, 2009).

Neoclassical theory on trust in economics is just one example of theory that does not correspond (represent?) to observed behavior. "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. 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, behavior in 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).

Interpersonal trust, experimental game theory, and findings on social preferences are discussed below. In addition, physiology and the Autonomic Nervous System (ANS) are explained, as well as methods for measuring individuals' physiology in an experimental setting. Finally, the experiment, methodology, and data collection are explained.

-- The above paragraph was a transition so that's why it's worded this way. I think I need to look at this again / rework it. I could use it in the introduction or somewhere else where I need to summarize.

II. 5. Interpersonal Trust

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.

Interpersonal trust simply refers to trust that occurs between individuals. A formal and widely accepted definition from Rousseau et al. (1998) 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.

The determinants of trust are dependent upon the relevant characteristics of the individual, the situation, and their interaction (Evans & Krueger, 2009, p. 1003). Researchers have found consistent and significant individual differences when measuring trusting behavior (Evans & Krueger, 2009). The decision to trust is related to personality differences, social identity, and expectations, and these differences are in part predicted (and perhaps caused) by an underlying propensity to trust (Evans & Krueger, 2009; Rotter, 1967). "This propensity is most relevant in ambiguous situations; dispositions matters most when there is little first-hand knowledge or past experience" (Evans & Krueger, 2009, p. 1011). Provided that reliable and validated instruments are used, behavior is consistently found to be correlated with disposition (Evans & Krueger, 2009).

Dispositional trust is "the general expectation that others will behave fairly and responsibly" (Evans & Krueger, 2009). Dispositional theory posits that "trust is motivated by individual differences in a general expectation" (Evans & Krueger, 2009, p. 1014). For example, individuals develop differing amounts of *generalized expectancy* - the learned expectation that promises kept between other social agents will generalize from one social agent to another (Rotter, 1967). "The development of such a generalized attitude may be learned directly from the behavior of parents, teachers, peers, etc., and also from verbal statements regarding others made by significant people or trusted sources of communication such as newspapers and television" (Rotter, 1967, p. 653). Studies suggest that children that experience a higher proportion of promises kept by parents and authority figures in the past have a higher generalized expectancy for trust from other authority figures (Rotter, 1967).

The propensity to trust alone does not provide a complete explanation for trusting behavior. "The situation and context matter" (Evans & Krueger, 2009, p. 1012). Further, dispositional theories explain interactions between strangers, but trust in established relationships is more dependent upon context and experience (Evans & Krueger, 2009). Since "trust is associated with a

mental state where we give others the benefit of the doubt by default," established relationships change the process and determinants of trust significantly (Evans & Krueger, 2009, p. 1014). Economists focus on dispositional trust because trust in a one-time setting is typically derived by the estimation of gains and losses, weighed by perceived risks; thus, it is easier to study in a controlled setting (Rousseau et al., 1998).

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).

II. 6. Experimental Economics and Experimental Game Theory

Game theory examines the mechanism between individual decision making and group level outcomes and the manner in which incentives affect decisions in strategic economic environments (Houser & McCabe, 2014). An economic environment or interaction is considered *strategic* "when decisions made by one player influence the opportunities and payoffs available to another," and all players are aware of this (aware this is the case) (Houser & McCabe, 2014).

Game theory is the study of one or more strategic interactions and the way(s) in which incentives (can) affect the decisions made by the players of the game. Economic interactions are considered (to be) (defined as) strategic when decisions made by one player influence the opportunities and payoffs available to another player, and all players (are aware of this) / (know this is the case).

Experimental games allow us to measure "how much players are willing to sacrifice their own economic payoff to increase or decrease the payoffs of others" (Fehr & Krajbich, 2014, p. 194). Laboratory game theory experiments are designed to test game theory, and are particularly useful

for uncovering critical aspects of the human decision making process that might (may, can) be difficult to detect (determine) outside of a controlled environment (Houser & McCabe, 2014). An important aspect of laboratory game theory experiments is that participants' decisions can be impacted (influenced) by the experiment design & implementation. For example, the framing of instructions, randomization, anonymity, and salient rewards (real money incentives) are key (vital?) experimental economics procedures that can have profound impacts on the internal and external validity of an experiment (Houser & McCabe, 2014).

Games are represented visually as mathematical models by a game tree, which specifies: which player (or players) are able to move when, what moves can be made at a given stage, what information players possess when they make their move, and how moves (i.e. decisions) made by different players can impact further decisions and determine the final outcome(s) of other players.

In a Dictator game, one individual is randomly selected to be the 'dictator' and is allocated a sum of money. The dictator can divide the money between themselves and the other player, but they are not required to share any of their endowment. The 'dictator' has the final say. The Dictator game is designed to measure positive concern for the other player's material payoff that is independent of the other player's behavior (because the other player has no actions).

Neoclassical economics predicts that the 'dictator' will not share any of their endowment because it is their self-interest to keep the entire amount for themselves. Further, there is nothing compelling them to share. Instead, research shows that the dictator allocations tend to be a blend of 50% offers and 0% offers, with a few offers in between (Fehr & Krajbich, 2014, p. 194). Studies also show that allocations in the dictator game "are sensitive to details of how the game is described (Camerer, 2003), the dictator's knowledge of who the recipient is (Eckel and Grossman, 1996), and whether the recipient knows" that they are playing a Dictator game (Dana et al., 2006)" (Fehr & Krajbich, 2014, p. 194).

The Ultimatum game is a simple, one time take-it-or-leave-it bargain between two randomly and anonymously matched individuals. One person is randomly assigned to be the proposer, the other is the responder. The proposer is endowed with some predetermined amount of money and is instructed to suggest, or *propose*, a division of that money between themselves and

the responder. The responder then decides whether to accept or reject the proposed division of the money. If the responder accepts, then both players earn the amount decided upon by the proposer. If the responder rejects the offer, neither earn money. The Ultimatum game is identical to the Dictator game, except that the recipient can accept or reject the proposed allocation.

The key findings of ultimatum game laboratory experiments is that most proposers offer between 40-50% of the endowed amount, and this split is almost always accepted by the responder (Houser & McCabe, 2014). In addition, these results are highly robust to a variety of natural design manipulations, such as repetition, stake size, and degree of anonymity, as well as a variety of demographic variables (Camerer, 2003; Houser & McCabe, 2014, p. 28).

II. 7. The Trust Game

The trust game was introduced in 1995 by Joyce Berg, John Dickhaut, and Kevin McCabe in *Trust, Reciprocity, and Social History* to control for alternative explanations of trusting behavior. 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/making/taking home 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 (Evans & Krueger, 2010). 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. The social optimal outcome of the game is mutual cooperation" (Evans & Krueger, 2009, p. 1006)

Figure 1 provides an example of the trust game. With complete reciprocity, players receive outcome R , \$20; with betrayal, the trustor receives S , \$5, and the trustee receives T , \$55. The payoffs are ordered as follows: $T > R > P > S$. Ultimately, the outcome of trust depends upon the behavior of the trustee (Evans & Krueger, 2009). The key is the amount returned by the trustee, as this is a measure of trustworthiness.

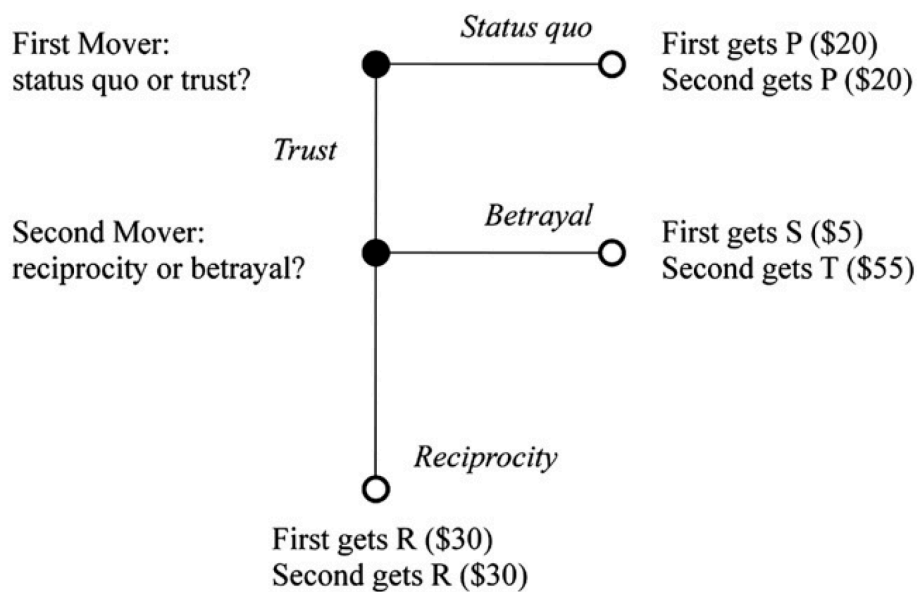


FIGURE 1: AN EXAMPLE OF THE TRUST GAME, FROM EVANS AND KRUEGER (2010).

The psychological situations and choices faced are notably different for the two players. The second mover faces a choice between fairness and selfishness: they can reciprocate and distribute

the money equally (each player receives R), or betray trust by taking more money for themselves (Evans & Krueger, 2010). "On the other hand, the trustor has a strategic dilemma that is not solvable without making assumptions about the trustee. The trustor knows that reciprocated trust yields a better outcome than the status quo ($R > P$), but there is no guarantee that reciprocity will occur" (Evans & Krueger, 2010, p. 171). Betrayal leads to an outcome that makes the trustor worse off than the status quo ($S < P$). The trustor must decide if the benefit outweighs the cost and if the risk is worth taking.

The payoff structure of the trust game is comprised of three distinct elements: cost, benefit, and temptation. Cost and benefit correspond to the concept of (personal) vulnerability or risk. "The trustor's potential cost is the difference between the status quo and betrayal ($P - S$), and the potential benefit is the difference between reciprocity and the status quo ($R - P$)" (Evans & Krueger, 2010, p. 171). Temptation captures the trustee's additional payoff if they defect, as it is measured as "the difference in the trustee's payoffs between betrayal and reciprocity ($T - R$)" (Evans & Krueger, 2010, p. 171). It is critical to the trustor because it implies the probability of reciprocity. "Empirically, temptation is the best predictor of reciprocity among strangers" (Evans & Krueger, 2010, p. 171). These three elements "capture the idea that trust presupposes both vulnerability and expectation. Without personal vulnerability, the expectation that others will act benevolently would be an ordinary social prediction (Luhmann, 2000). Conversely, trust without the expectation of reciprocity is self-defeating" (Evans & Krueger, 2010, p. 171). Figure 2 depicts the cost, benefit, and temptation of the trust game along a number line.

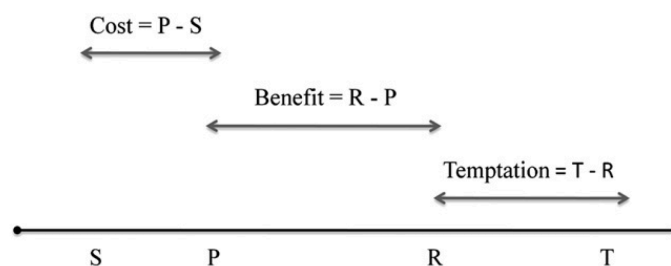


FIGURE 2: COST, BENEFIT, AND TEMPTATION IN THE TRUST GAME, FROM EVANS AND KRUEGER (2010).

Classic game theory fails to describe real-world behavior because it predicts that players will "reason by backward induction. In the trust game, this means individuals must view the payoffs in the final stages of the game and then reason backward to the game's beginning" (Evans & Krueger, 2010, p. 172). In practice, backward induction is counterintuitive (Evans & Krueger, 2010). Instead, players tend to focus on the payoffs in the first round and reason forward (Evans & Krueger, 2010). This is true even after individuals gain experience in the game (Evans & Krueger, 2010). Evans and Krueger (2010) proposed a revised rational model predicts that people calculate an estimate of the expected value of trust, and sought to provide direct evidence of their process-oriented description of trust. Unlike backward induction, their model allows for the subjective probability of reciprocity to be positive (Evans & Krueger, 2010).

Berg et al. (1995) found that trustors send about 50% of the endowment on average, and trustees generally return the same amount sent. The most variance was seen in the amounts returned, with 50% of trustees returning \$1 or less. Empirically, trustors send about half of their endowment and trustees send back about as much as the trustor (Camerer, 2003; Fehr & Krajbich, 2014). Research shows that players consistently choose to trust the stranger they have been randomly paired with, contradicting the neoclassical prediction and Nash equilibrium.

The trust game introduced by Berg et al. (1995) has become an influential experimental paradigm in both psychology and behavioral economics (Evans & Krueger, 2010). It has been used widely in neuroeconomics experiments and has played an important role in shaping economists' views of trust and reciprocity (Houser & McCabe, 2014). The trust game is an example of a non-cooperative game – players are unable to enforce cooperation nor punish non-cooperation, and any cooperation must be self-enforcing (Shor, 2005). The trust game also represents an economic situation with imperfect information – players do not know what the other player has chosen. As discussed in section I.5. Rational Choice Theory, expected utility theory assumes that all parties have perfect information. Neoclassical economists assume that at least some individuals are EU-rational and that an equilibrium, such as Nash, exists (Blume & Easley,

2018). Classic game theory assumes that players are strictly self-interested. "If so, a player reasoning by backward induction must conclude that trust is irrational" (Evans & Krueger, 2010, p. 171). These assumptions are false, which is why expected utility theory fails to predict behavior in the trust game. "Classic game theory does not fare well as a descriptive theory. Scores of experiments show that trust is a common choice, even under conditions of complete anonymity" (Evans & Krueger, 2010, pp. 171-172). Instead, "neoclassical rationality predicts an inefficient equilibrium: mutual defection" (Evans & Krueger, 2009, p. 1006). In other words, it predicts that the trustor will keep their initial endowment, ending the game after one round.

The figures below represent economic games as mathematical models known as game trees. In the figures time moves from left to right, arcs represent the possible actions that the player could take, the numbers represent the material payoffs for the players, and both are color coded to match the players (Fehr & Krajbich, 2014). The dotted lines represent one example choice. An example of the dictator game is depicted in Figure 1. The dictator, shown in red, has \$10 and can send any amount to the recipient, shown in blue. In this example, the dictator chooses to send \$3 and keep \$7 for themselves.

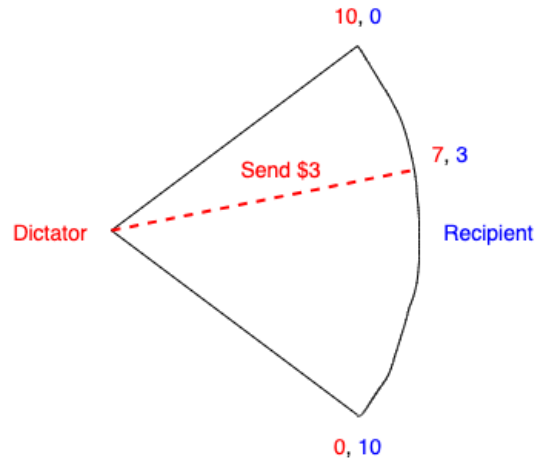


FIGURE 3: EXAMPLE OF A DICTATOR GAME

The Trust game in Figure 3 depicts the trustor sending \$3, which becomes \$9. The trustee returns \$5 and keeps \$4 for themselves.

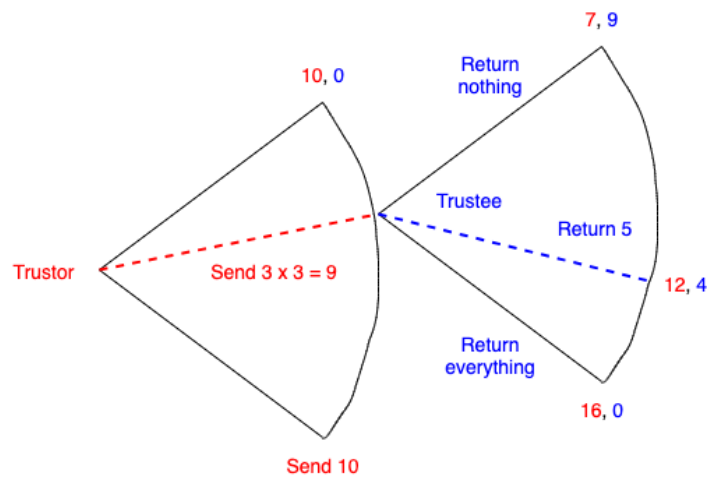


FIGURE 4: EXAMPLE OF A TRUST GAME

II. 8. Social Preferences

Social preferences are a characteristic of a individual's behavior or motivation that indicate whether the individual has positive or negative other-regarding motives (Fehr & Krajbich, 2014, p. 193). Social preference theories assume 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). In other words, they trade-off other-regarding behavior with selfish goals (Fehr & Krajbich, 2014). The majority of individuals display social preferences. They are considered one important aspect of individuals' utility functions (Fehr & Krajbich, 2014, p. 196). For example, most care about others' material payoff or 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). In addition, research shows that social preferences "may play a decisive role for aggregate social and economic outcomes in strategic settings" (Fehr & Krajbich, 2014, p. 193).

There is also evidence of considerable individual heterogeneity in social preferences: some

people display strong, positive social preferences, while others display little or not concern for the other person (Fehr & Krajbich, 2014, pp. 193, 195). This heterogeneity in social preferences "is a key reason why, in certain competitive environments, all individuals behave *as if* they were purely self-interested (Smith, 1962, 1982), while in strategic games the vast majority of individuals often deviate strongly from purely self-interested behavior" (Fehr & Krajbich, 2014, p. 195).

Social preferences are elicited predominantly using simple, one-shot games such as the Dictator game or Ultimatum game that involve real monetary payoffs and are played between two anonymous individuals (Fehr & Krajbich, 2014, p. 196). In a one-shot game two players play the game with each other only once; it does not consist of repeated play between the same two players. In one-shot games, an individual displays social preferences if they are "willing to forgo their own material payoff for the sake of increasing or decreasing another individual's material payoff" (Fehr & Krajbich, 2014, p. 196). In other words, an individual cares enough about helping or harming another to bear the associated cost.

A robust demonstration of social preferences requires that an individual's action be independent of their beliefs about the other player's action, since such beliefs affect behavior and thus represent a confound (Fehr & Krajbich, 2014, p. 196). Further, anonymity is an important factor because it provides a baseline measure of social preferences (Fehr & Krajbich, 2014, p. 196). A lack of anonymity also represents a confound. Both aspects, one-time play and anonymity, are crucial for accurately measuring social preferences in experiments. Finally, all theories of social preferences assume that individuals' utility functions depend not only on their own material payoff, but also on non-monetary payoff elements: concerns for fairness, reciprocity, equality, and efficiency (Fehr & Krajbich, 2014, p. 196).

The research stimulated by social preferences and norm-based theories has enabled behavioral economics to overcome the limitations of the neoclassical, self-interest hypothesis (Evans & Krueger, 2009). "Trust and reciprocity are not necessarily irrational when they are understood in these terms" (Evans & Krueger, 2009, p. 1009). Further, the experimental data demonstrate that other-regarding behavior often leads to socially desirable results (Evans & Krueger, 2009).

III. Physiological Synchrony as a Measure of Trust

III. 1. Physiology and the Autonomic Nervous System (ANS)

Physiological arousal is the subconscious bodily response to some stimuli. Emotions experienced are expressed through one's physiological arousal (Appelhans & Luecken, 2006). The Autonomic Nervous System (ANS) is responsible for the generation of physiological arousal by coordinating and managing a complex, highly differentiated network of nerves, organs, and biological sensors distributed throughout the body (Appelhans & Luecken, 2006; Levenson, 2014). "Functionally, the ANS plays a number of different roles, serving as regulator, activator, coordinator, and communicator" (Levenson, 2014, p. 101). It regulates vital bodily functions including heart rate, respiration, and digestion, and internal processes such as the cardiac, respiratory, and glandular systems (Palumbo et al., 2017). The ANS is also the primary mechanism of our fight-flight-or-freeze response. It is composed of two systems, the excitatory sympathetic nervous system (SNS) and the inhibitory parasympathetic nervous system (PNS). Together, they dynamically regulate physiological arousal. The SNS is activated during physical or psychological stress, and is characterized by an increase in pulse, sweat secretion, and alertness. During periods of SNS activation, the body chooses to divert energy away from secondary processes such as digestion. The immediate threat requires our full attention. The PNS is most active in non-stressful situations. The effects of the PNS include a slowed and steady heart rate, pupil constriction, and increased digestion (Marieb, 2002). Both systems work in tandem, dynamically changing as they regulate the body in preparation for and response to endogenous and exogenous environmental conditions" (Palumbo et al., 2017, p. 100).

There are a variety of techniques and measures used to study the complex interaction between the SNS and PNS. For example, the human skin displays many forms of bioelectric phenomena, particularly in the hands, fingers, and soles of the feet (Pflanzer & McMullen, 2012). The skin's ability to conduct electricity is known as electrodermal activity (EDA). This specifically refers to the changes in the electrical conductance of the skin, which depends on sweat secretion.

EDA "is an indirect measure of eccrine sweat glands, which are uniquely innervated by the SNS" (Palumbo et al., 2017). 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.

Since EDA has been closely linked to implicit emotional responses and cognitive processing that may occur without conscious awareness, it is widely used as an objective index of emotional processing and sympathetic activity (Braithwaite, Watson, Jones, & Rowe, 2015). For example, changes in one's skin conductance level (SCL) reflect general changes in autonomic arousal (Braithwaite, Watson, Jones, & Rowe, 2015). According to Braithwaite et al. (2015), it is "the most useful index of changes in sympathetic arousal that are tractable to emotional and cognitive states as it is the only autonomic psychophysiological variable that is not contaminated by parasympathetic activity" (p. 3). It is difficult to overstate the importance of ANS measures to the study of psychology and psychological research; "psychophysiological findings have contributed to almost every aspect of psychology, including cognition, emotion, and behavior" (Palumbo et al., 2017, p. 100).

III. 2. Physiological Response to Decision Making (Incomplete)

Bechara et al. (1997) were the first to examine physiological response during decision making and uncovering that individuals generate skin conductance responses (SCRs) in anticipation of a large loss. Their study compared healthy individuals and those with ventromedial prefrontal cortex (vmPFC) and decision making defects as they performed a gambling task in which participants are asked to draw cards from four decks, two "risky" decks in which repeated drawings lead to an overall loss, and two "safe" decks in which repeated drawings lead to an overall gain.

After encountering a few losses, participants without prefrontal cortex damage began to generate anticipatory skin conductance responses (SCRs) *before* selecting a card from the "risky" decks and also began avoiding these decks (Bechara et al., 1997). Those with vmPFC damage did neither. To determine whether "subjects choose correctly only after or before conceptualizing the nature of the game and reasoning over the pertinent knowledge" the researchers briefly interrupted the game to ask questions that would allow them to assess the participants' "account of how they conceptualized the game and of the strategy they were using" (Bechara et al., 1997, p. 1293). They found that although the individuals with vmPFC damage were able to accurately conceptualize the game and the correct strategy, they failed to generate autonomic responses and continued to select cards from the bad decks. In other words, they failed to make decisions based on their accurate conceptual knowledge.

To explain this novel finding, Bechara et al. (1997) proposed/presented/provided a "sensory representation of a situation that requires a decision" (Figure X) and showed how it "leads to two largely parallel but interacting chains of events" (1997, p. 1294). When an individual is presented with a situation, neural systems are activated that "hold nondeclarative [*sic*] dispositional knowledge related to the individual's previous emotional experience of similar situations" (Bechara et al., 1997, p. 1294). The ventromedial frontal cortices are among the brain structures that store dispositional knowledge, the activation of which activates the autonomic nervous system and neurotransmitter production (Bechara et al., 1997). These non-conscious signals act as "covert biases on the circuits that support processes of cognitive evaluation and reasoning" (Bechara et al., 1997, p. 1294). In the second chain of events, a situation triggers the overt recall of pertinent facts, e.g. possible responses and potential outcomes given a particular course of action, and the application of reasoning strategies to facts and options. Participants with vmPFC damage could be described as doing this chain of events while making decisions during the gambling task.

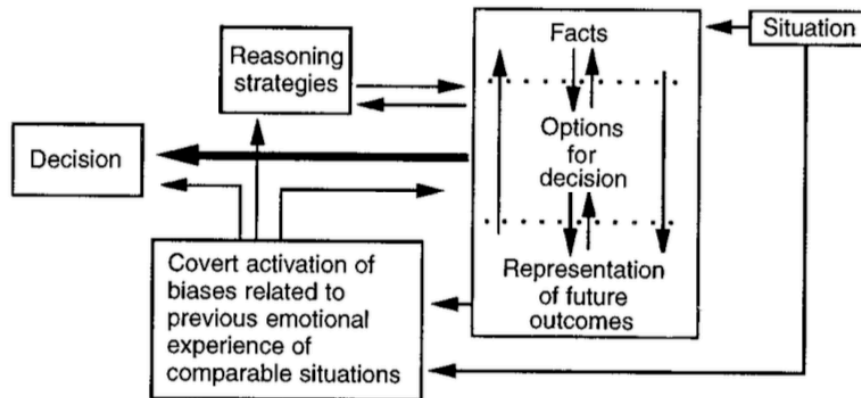


FIGURE 5: DIAGRAM OF THE PROPOSED STEPS INVOLVED IN DECISION MAKING, FROM (BECHARA ET AL., 1997)

Bechara et al.'s (1997) experiment demonstrates that brains of healthy individuals activate covert biases that precede overt reasoning on the available facts, and that these biases may assist the reasoning process by facilitating "the efficient processing of knowledge and logic necessary for conscious decisions" (Bechara et al., 1997, p. 1294). The researchers theorized that their results are evidence of "a complex process of nonconscious signaling, which reflects access to records of previous individual experience — specifically, of records shaped by reward, punishment, and the emotional state that attends them" (Bechara et al., 1997, p. 1294). Individuals with damaged vmPFCs are prevented from accessing these particular kinds of previous and related experiences; they are unable to produce nonconscious signals and covert biases that precede overt reasoning.

Bechara et al.'s (1997) theory/explanation matches (is aligned with) Damasio's (1994) Somatic Marker Hypothesis, which suggests that these bodily signals are an anticipatory arousal response corresponding to the value of (the?) choice (Glimcher & Ernst, 2013). Further, these bodily signals serve to deter participants away from "risky" choices (Glimcher & Ernst, 2013).

Damasio (1994) theorized

Damasio (1994) developed the Somatic Marker Hypothesis to explain

Researchers theorize

the Somatic Marker Hypothesis to explain results

(Damasio, 1994)

(Glimcher & Ernst, 2013)

Emotional arousal has also been identified with social decision tendencies in the the ultimatum game.

Emotional arousal has also been identified "with decision tendencies in social decision making in the ultimatum game" (Glimcher & Ernst, 2013)

Arousal responses

"Interpersonal physiological interactions are theorized to be ubiquitous social processes that co-occur with observable behavior" (Palumbo et al., 2017, p. 99).

III. 3. Physiological Synchrony

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). There is evidence that the ANS is responsive to, and in some instances, dependent on and/or shaped by, the nervous system of others between-subjects (Palumbo et al., 2017). 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). PS is also referred to as *interpersonal physiology*, *physiological linkage*, *physiological coherence*, and *physiological covariation* (Palumbo et al., 2017, p. 101). For example, Reed, Randall, Post, and Butler (2013) used the term *physiological linkage* "in reference to both concurrent and lagged interdependencies between participants' cardiac and electrodermal measures" (Palumbo et al., 2017, p. 102).

Following a review of the IAP literature, Palumbo et al. (2017) identified six parameters that have been used to define PS: *magnitude*, *sign*, *direction*, *lag*, *timing*, and *arousal*.

Magnitude is a measure of "the strength of synchrony, such as a regression or correlation coefficient. This typically represents the effect size of a given measure of synchrony" (Palumbo et al., 2017, p. 102).

Sign indicates the direction in which individuals' arousal levels move, either in the same or opposite directions. This is also referred to as concordant and discordant synchrony (Palumbo et al., 2017).

The *direction* parameter "refers to the predictability of one person's physiology from another's" (Palumbo et al., 2017, p. 102). Tests of predictability may reveal a unidirectional or multi-directional relationship (Palumbo et al., 2017).

Lag represents the shift in the temporal alignment of data. "The difference in time alignment from one data set to another is the specific lag, which may indicate millisecond differences, or much longer time offsets" (Palumbo et al., 2017, p. 102). According to Palumbo et al. (2017), "although lagged PS implied a unidirectional relationship, there is some evidence that the length of a lag (e.g., 1s vs. 10s time offset) can reflect psychosocial properties that are independent of direction and worthy of exploration" (102). Lagged PS will be explored in various/multiple analyses in Chapter 3.

Timing refers to the length of time that an interaction is assessed or analyzed. The time scale selected is a key element of the research question (Palumbo et al., 2017).

Arousal is commonly assessed as a covariate of synchrony. For example, mean arousal level can be tested as a covariate or moderator of PS, such as examining whether increasing PS correlates with decreasing arousal (Palumbo et al., 2017). Further, "specific physiological states in one person may contribute to reduced or amplified arousal levels in others, or the likelihood that a given type of PS will develop" (Palumbo et al., 2017, p. 103).

Asynchrony is the opposite construct of PS; it describes a lack of observable PS (Palumbo et al., 2017). *Asynchrony* is "difficult to substantiate without the use of multiple models to test for PS," but identification of interactions that lack PS can provide valuable insight (Palumbo et al., 2017, p. 103). "For example, periods of asynchrony may occur during an interaction if an individual is ignoring the state of a partner, or if one member does not correspond with others in a group"

(Palumbo et al., 2017, p. 103).

Palumbo et al.'s (2017) review of the IAP literature characterizes synchrony as a nonspecific construct, since its detection is dependent on the procedure used to test it. In general, analyses of synchrony identify both concurrent and/or lagged interdependencies between individuals' cardiac and/or electrodermal measures (Palumbo et al., 2017). Different analyses of synchrony address different components of data, therefore the results can differ substantially (Palumbo et al., 2017). Thus, it is necessary that an appropriate PS analyses correspond to the research question(s) at hand, since different approaches can alter study findings and implications (Palumbo et al., 2017). Additionally, the majority of studies on the ANS have been performed (primarily) at the intrapersonal level, or within-subjects (Palumbo et al., 2017). Instead, temporal changes are explored at the interpersonal level, or between-subjects.

Add:

studied married couples engaged in naturalistic interactions to determine the extent to which variation in marital satisfaction could be attributed to physiological and affective patterns between and within spouses (Levenson & Gottman, 1983). Using
They found that "60% of the variance in marital satisfaction was accounted for using measures of physiological linkage alone" (Levenson & Gottman, 1983),

III. 4. Mechanism(s)? of Physiological Synchrony (Incomplete)

Early research on interpersonal physiological synchrony began over half a century ago, "when a series of studies found significant positive and negative correlations in the EDA and HR of therapists and clients during therapy, interpreted as evidence of therapeutic rapport and empathy" (Palumbo et al., 2017, p. 110).

The study of physiological synchrony began in the 1950s with the study of patient-therapist

relationships and analyses of physiological responses of therapists and their clients during interviews.

Results suggested that there were periods of synchronization in the heart rates (HR) and skin conductance (SC) levels of the dyads. Contextual data indicated that sessions with higher levels of PL were experienced as more empathic, prompting the researchers to conclude that there was a physiological component of empathy." {Palumbo, 2015, PhD in Psychology, Interpersonal Physiology: Assessing Interpersonal Relationships Through Physiology} pg. 22

Contemporary research on physiological synchrony has provided evidence of a number of psychosocial constructs, including empathy, attachment, conflict, and emotional co-regulation, with empathy and attention considered the most common explanations (Palumbo et al., 2017).

For example, Marci et al. (2007) found that clients' reports of therapist empathy were positively correlated with the magnitude of PS. Clients therefore had a reportable experience of feeling more understood when PS was higher. {Palumbo et al., 2017, Personality and Social Psychology Review, 21, 99-141} pg.125

The presence of PS has been considered evidence of a number of psychosocial constructs, including empathy (Adler, 2002, 2007), attachment (Diamond, 2008), conflict (e.g., Levenson & Gottman, 1983), and emotional coregulation (Field, 2012; Sbarra & Hazan, 2008), although conflict- ing results suggest it is not specific to any of these. Empathy is the most commonly considered psychosocial explanation of PS.

From the original studies through recent reports, researchers and theorists have considered the possibility that experiential connections that define emotional empathy are mirrored in physiology.

These ideas suggest that the ANS reflects a component of shared experience, so PS may be an objective measure of internal processes accompanying an empathic interaction.

"have considered the possibility that experiential connections that define emotional empathy are mirrored in physiology"(Palumbo et al., 2017, p. 123).

Researchers have tried to determine whether "experiential connections that define emotional empathy are mirrored in physiology"(Palumbo et al., 2017, p. 123).

This suggests that the ANS

"These ideas suggest that the ANS reflects a component of shared experience, so PS may be an objective measure of internal processes accompanying an empathic interaction" (Palumbo et al., 2017).

It is unclear whether

"Although repeated findings show an association between PS and empathy, other results indicate these constructs are independent.

"Whereas PS may correlate with emotional empathy, physiological interactions in other contexts suggest it is not specific to that construct. Future research is needed to help disentangle this association, such as exploring whether a subtype of PS is specific to empathy" (Palumbo et al., 2017). pg. 23

"the development of PS between two or more people does not appear to be dependent on (a) shared conditions such as behavior or environment, (b) a specific sensory mode of communication, or (c) psychosocial conditions such as valence or relationship type. Multiple methods have been used to show that PS is not dependent on individuals being in the same conditions (e.g., Marci et al., 2007) or environments (e.g., McAssey et al., 2013), and that engaging in the same behaviors does not reliably lead to synchrony (e.g., Henning et al., 2001)" (Palumbo et al., 2017).

Second, the extant literature suggests that PS is a transient state. Studies showing differences in PS across contexts and conditions indicate that physiological relationships change over time. This is evident in studies such as Müller and Lindenberg's (2011) and Ghafar-Tabrizi et al.'s (2008), which show that during a given time period, parameters of PS are not static. This is an important consideration, as attempts to apply statistical models that assume a constant state may be problematic. For example, if a dyad shifts between periods of positive and negative PS during a trial, but the entire interaction is assessed using a single linear model, then results will be an aggregate of two heterogeneous processes and will misrepresent the patterns of both. Guastello et al. (2006) and Helm et al. (2014) addressed this issue well, highlighting the need for flexible statistical models capable of identifying multiple types of physiological relationships occurring during a single interaction.

PS can therefore be predictive of an outcome, although more research is needed to explore interactions between PS parameters and variables including context, valence, and arousal.

*Although there is evidence that these processes are not fully dependent on environment or behavior, **it is unclear how one individual's autonomic activity responds to or mimics another's**. One likely possibility is that these complex interactions can develop through a number of different mechanisms, including shared environment, coordinated behaviors, and matched responses to a third variable, as well as through interpersonal processes. Future research may be directed at experimentally determining the degree to which given variables are dri-*

ving interper- sonal physiological interactions during a given condition, as multiple modalities may be simultaneously involved.

Systematic, experimentally manipulated, research is needed to isolate the basic components that contrib- ute to PS for interpretations of group and conditional differences to be meaningful. Until controlled, systematic research addresses the basic components and outcomes of PS, interpretations of results will be limited (Sbarra & Hazan, 2008). (pg. 126)

Controlled experiments are required to explore mechanisms that generate PS, as well as to determine interactions with different variables

The application of an inductive strategy is recommended to identify and define a typology of PS, followed by systematic replication of studies across contexts and time, both within and across people. Although con- verging evidence suggests that PS is robust enough to be detected using correlational analyses and nomothetic methods, results from these strategies may be too broad to identify complex, time-varying components of PS. Combining idiographic designs with time-varying analyses offers the greatest potential to explore various processes involved in PS. (Palumbo, 2017, pg. 127)

III. 5. Statistics of Physiological Synchrony (Incomplete)

The critical component of studying / measuring interpersonal autonomic physiology is the type of statistical analysis used.

Statistical procedures used to capture/measure physiological synchrony have evolved and produce better accuracy (?)

The previous literature has focused predominantly on correlation calculations. Extensions

...

(Palumbo et al., 2017):

Another critical issue for IAP research is the type of statistical analysis used.

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).

Still, the number of viable statistical procedures applicable to IAP research is rapidly increasing, and many studies have developed strategies tailored to these data. Some example methods include dynamic systems models (Ferrer et al., 2013), cross-lagged panel models (Helm et al., 2014), multilevel models (Reed et al., 2013), wavelet analysis (Müller, & Lindenberger, 2011; Quer et al., in press), and functional data analysis (Liu et al., in press). Despite their promise, it is important that researchers adequately match these statistical approaches to appropriate research questions, as interpretation of results can differ substantially depending on the method used. (pg. 126)

co-occurring physiological interactions may exist at multiple time scales, so explorations of timing are critical to uncovering long- and short-term processes. Future work is needed to both assess the time range in which interpersonal physiological interactions occurs and to develop and adapt analyses capable of detecting when and how they change over time.

(pg. 127)

Finally, regardless of statistical procedures used, determining a valid null hypothesis is a necessary component for testing PS. Spurious synchrony can be detected in randomly paired data, and the probability of chance findings can vary due to contextually dependent data structures. To validate findings, a null hypothesis determining the potential for chance findings of PS in contextually matched, randomized

data is often necessary. Otherwise, it may be unclear whether results are valid, or due to chance. (pg. 127)

The complexities inherent to PL call for analyses with a temporal resolution that can only be achieved through idiographic methods. Newly developed idiographic analyses focus on patterns of relationships over time within a given unit (e.g., a dyad). {Palumbo, 2015, PhD in Psychology, Interpersonal Physiology: Assessing Interpersonal Relationships Through Physiology}

Various analytical methods have been developed to create an accurate 'statistic of physiological synchrony.' Some of the first analyses used to measure (identify) physiological synchrony ...

III. 5. 1. Pearson Correlation

- I think this is what many papers did in the 90s, and then they tried to adjust and make derivations of correlations with lags and window functions.
- McFarland (2001) – Cross-correlation
- Strang, Funke, Russell, Dukes, and Middendorf (2014) – teammates; cross-correlation, cross-recurrence quantification analysis, and cross-fuzzy entropy

Ran pearson correlations with zero lag on 15-second windows. Simplistic.

“Windowed correlation of slope” ~ window of percent changes?

See Palumbo 2017

3 papers used the same technique:

Marci and Orr 2006

Marci et al. 2007

Messina et al. 2013

- They used either window time lagged correlation, or a simple Pearson correlation

(Marci & Orr, 2006):

Pearson correlations with lag-zero were then calculated over successive, running 15- second windows (corresponding to 15 slope averages) between time-locked patient and interviewer SC slope values. Finally, a single session index was calculated from the ratio of the sum of the positive correlations across the entire interview divided by the sum of the absolute value of negative correlations across the interview. Because of the skew inherent in ratios, a final natural logarithmic transformation of the resulting index was calculated. Thus, an index value of zero reflects equal positive and negative correlations or balanced concordance for the interaction. Concomitantly, a value greater than zero reflects relatively more positive SC concordance over time while a value less than zero reflects relatively more negative SC concordance over time.

(Marci et al., 2007):

Physiologic Concordance (pg. 105)

Following data collection, continuous SC levels from the therapy sessions were used to derive a measure of SC concordance (Marci and Orr, 2006). First, a smoothing function using a 1.5-second Bartlett window was applied to the continuously recorded SC level values (Oppenheim and Schaffer, 1999). Next, the average slope of the SC level within a moving 5-second window was determined (i.e., after calculating the first 5-second slope value, the window moves forward 1 second and calculates the next 5-second slope value, producing successive 5-second SC slope values at 1-second increments). Slope was chosen as the index of SC activity because the rate of change in SC level (as an indicator of physiologic arousal) was of primary interest and was

expected to be a better determinant of psychophysiology concordance between interacting individuals than individual discrete peaks or absolute SC level. Individual peaks and absolute SC level tend to show idiosyncratic drift over time, and amplitude and frequency can vary considerably from person to person (Hugdahl, 1995).

Pearson correlations with lag-zero were then calculated over successive, running 15-second windows (corresponding to 15 slope averages) between time-locked patient and therapist SC slope values. Next, a single session index was calculated from the ratio of the sum of the positive correlations across the entire therapy session divided by the sum of the absolute value of negative correlations across the session. Because of the skew inherent in ratios, a natural logarithmic transformation of the resulting index was calculated. Thus, an index value of zero reflects equal positive and negative correlations or neutral concordance for the therapy session. Concomitantly, a value greater than zero reflects relatively more positive SC concordance over time, while a value less than zero reflects relatively more negative SC concordance

Pearson correlations with zero lag were calculated over successive, running 15-second windows on skin conductance slope values (percentage change in skin conductance).

"Next, a single session index was calculated from the ratio of the sum of the positive correlations across the entire therapy session divided by the sum of the absolute value of negative correlations across the session. Because of the skew inherent in ratios, a natural logarithmic transformation of the resulting index was calculated. Thus, an index value of zero reflects equal positive and negative correlations or neutral concordance for the therapy session. Concomitantly, a value greater than zero reflects relatively more positive SC concordance over time, while a value less than zero reflects relatively more negative SC concordance over time

III. 5. 2. Time Lagged Cross Correlation

- "Time lagged cross correlation (TLCC) can identify directionality between two signals such as a leader-follower relationship in which the leader initiates a response which is repeated by the follower."
- "...we can still extract a sense of which signal occurs first by looking at cross correlations."
- TLCC is measured by incrementally shifting one time series vector (red) and repeatedly calculating the correlation between two signals. If the peak correlation is at the center (offset=0), this indicates the two time series are most synchronized at that time. However, the peak correlation may be at a different offset if one signal leads another.

- (Cheong, 2019)

(Butler, 2011):

Time-lagged synchrony: Time-lagged covariation with level removed. The majority of the literature on synchrony has minimized the interpretive difficulties associated with concurrent assessment of partners' emotions by focusing instead on time-lagged covariation, most commonly by using cross-correlation functions (CCFs) obtained from time-series analysis (Feldman, 2003, 2006, 2007a; Field, 1985; Gottman, 1981). A CCF plot shows the strength of association between two time series for each possible time lag after removing any linear trends. Positive spikes show evidence of association with Partner-1 leading, negative spikes show evidence of association with Partner-2 leading, and spikes in both directions suggest association with both partners alternately leading (see Recommendations section).

Recommendations for assessing time-lagged synchrony. One approach for assessing time-lagged synchrony involves using time-series methods (Feldman, 2003). The strength of this approach is that covariation is simultaneously assessed across a wide range of time lags (e.g., Person1 predicted from Person2 1 minute earlier, 2 minutes earlier, 3 minutes earlier, etc.). Therefore, this approach is optimal if you do not have an a priori idea about where to look in time for the best evidence of synchrony. The disadvantages of this approach are that it requires a fairly large number of repeated observations (in the order of hundreds) and it requires separate preliminary analyses for each person in your sample, followed by preliminary analyses for each dyad. Specifically, autoregressive integrated moving average (ARIMA) models are applied one at a time to each person's data to partial out the autocorrelation. The residuals from these models are saved and then cross-correlation functions are used to assess the covariance of these residuals between dyad partners across a range of lag times. Various pieces of information from the CCF plots can be noted, such as the value of the largest cross-correlation for each dyad, and then used as synchrony indicators in subsequent analyses of variance (ANOVAs) or regression analyses comparing those indicators across dyads.

III. 5. 3. Windowed Time Lagged Correlation

- "This process repeats the time lagged correlation in multiple windows of the signal. Then we analyze each window or take the sum over the windows, [which] would provide a score comparing the difference between the leader follower interaction between two individuals."
- This sounds exactly like ISC
- List according to Palumbo (2017):

- Ham and Tronick (2009)
- Marci, Ham, Moran, and Orr (2007)
- Marci and Orr (2006)
- Messina et al. (2013)
- Stratford, Lal, and Meara (2009)
- Stratford, Lal, and Meara (2012)

A great challenge is to statistically capture the dynamics of social interactions with fluctuating levels of synchrony and varying delays between responses of individuals. Windowed Cross-Correlation analysis accounts for both characteristics by segmenting the time series into smaller windows and shifting the segments of two interacting individuals away from each other up to a maximum lag. Despite evidence showing that these parameters affect the estimated synchrony level, there is a lack of guidelines on which parameter configurations to use.

The current study aimed to close this knowledge gap by comparing the effect of different parameter configurations on two outcome criteria: (1) the ability to distinguish synchrony from pseudosynchrony by means of surrogate data analyses, and (2) the sensitivity to detect change in synchrony as measured by the difference between two within-subject conditions. Focusing on physiological synchrony, we performed these analyses on heartrate, skin conductance level, pupil size, and facial expressions data.

Results revealed that a range of parameters was able to discriminate synchrony from pseudosynchrony. Window size was more influential than the maximum lag with smaller window sizes showing better discrimination. No clear patterns emerged for the second criterion. Integrating the statistical findings and theoretical considerations regarding the physiological characteristics and biological boundaries of the signals, we provide recommendations for optimizing the parameter settings to the signal of interest.

(Behrens et al., 2020)

III. 5. 3. 1. ISC

Golland et al. (2014) applied an intersubject correlation (interSC) analysis to measure "the across-subject reliability of physiological responses by quantifying the commonalities of the response time courses among individuals..." (Golland et al., 2014, p. 1103). ISC ...

Based on...

III. 5. 4. Instantaneous Phase Synchrony (?)

Lastly, if you have a time series data that you believe may have oscillating properties (e.g. EEG, fMRI), you may also be able to measure instantaneous phase synchrony. This measure also measures moment-to-moment synchrony between two signals. It can be somewhat subjective because you need to filter the data to the wavelength of interest but you might have theoretical reasons for determining such bands. To calculate phase synchrony, we need to extract the phase of the signal which can be done by using the Hilbert transform which splits the signal into its phase and power (learn more about Hilbert transform [here](#)). This allows us to assess if two signals are in phase (moving up and down together) or out of phase.

The instantaneous phase synchrony measure is a great way to compute moment-to-moment synchrony between two signals without arbitrarily deciding the window size as done in rolling window correlations. If you'd like to know how instantaneous phase synchrony compares to windowed correlations, check out my earlier blog post [here](#).

(Cheong, 2019)

III. 5. 5. Differential Equation Modeling

(Ferrer & Helm, 2013)

III. 5. 6. Dynamic Time Warping

"Dynamic time warping (DTW) is a method that computes the path between two signals that minimize the distance between the two signals. The greatest advantage of this method is that it can also deal with signals of different length. Originally devised for speech analysis (learn more in this [video](#)), DTW computes the euclidean distance at each frame across every other frames to compute the minimum path that will match the two signals. One downside is that it cannot deal with missing values so you would need to interpolate

beforehand if you have missing data points."

(Cheong, 2019)

"

To compute DTW, we will use the `dtw` Python package which will speed up the calculation."

(Cheong, 2019)

III. 5. 7. MPdist Algorithm (?)

<https://link.springer.com/article/10.1007/s10618-020-00695-8>

IV. Predicting Dyad Payoffs in the Trust Gamewith Measures of Physiological Synchrony

Alt Chapter Titles:

- Applying Measures of Physiological Synchrony to Predict Dyad Payoff(s)? in the Trust Game
- Analyzing Measures of Physiological Synchrony to Predict Dyad Payoff(s)? in the Trust Game
- Analyzing Measures of Physiological Synchrony between Pairs in the Trust Game

IV. 1. Methodology (Incomplete)

Briefly re-introduce the problem and provide an overview of the methodological approach.

IV. 1. 1. Participants (Incomplete)

Am I supposed to be telling it as a story, like this:

Participants were recruited ... ?

When participants arrived, they were welcomed and given an explanation of the study.

Once they gave verbal and written consent, participants were asked to fill out a survey with demographic and personality trait questions.

Example from Hearts in Sync Paper:

METHODS

Participants.

We recruited 103 participants (57 female) from the Claremont colleges and the surrounding community through mass e-mails, posted fliers, and an existing online recruitment pool (ages 18-71, M= 30). Participants self-identified as Caucasian (53.4%), Asian (14.6%), Latino/Hispanic (13.6%), African American/Black (2.9%), and mixed ethnicity/other (15.5%). Participant payments varied depending on correctly answering narrative questions and voluntary donations made, with total possible of \$38. Study sessions were conducted at the Center for Neuroeconomics Studies (CNS) at Claremont Graduate University in Claremont, CA. Claremont Graduate University's Institutional Review Board and the U.S. Army Medical Research and Materiel Command's Office of Research Protections, Human Research Protection Office approved this study.

Experimental sessions were conducted between June 22 and July 11, 2013. Participation time averaged 55 minutes.

IV. 1. 2. Procedure

This exploratory study assessed interpersonal trust and the physiological responses associated with trust, using the Trust Game to measure trust behavior. 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. First, participants were fitted with the electrophysiological instrumentation and continuous EDA data were collected for the duration of the experiment, including several baseline quiet periods. Next, participants were awarded an endowment of \$120 to be used during the experiment. Participants played the Trust game and Dictator game with two other participants, both as the first mover and second mover, and as the dictator, respectively. 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. Figures 3 and 4 depict the game trees for the Dictator game and Trust game, respectively. Participants were also 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. Since there were two sessions, each with two dyads, there were a total of four games played and 28 decisions made: 8 Dictator game decisions (4 as the FM, 4 as the SM) and 20 Trust game decisions (4 as the FM, 4 as the SM with imperfect information, and 12 as the SM with perfect information). Trust was measured as the outcome across different payoff levels - the amount sent by the first mover in the trust game. The game order was counterbalanced between first and second mover order, and partner order across experimental sessions. Figure 5 shows the full experimental procedure for Protocol B, Session 1. "To prevent feedback effects, players did not receive information on the other player's decisions

until all ___ rounds of the game were completed" (Evans & Krueger, 2010)

At the end of the session, participants learned of their partners' decision for only one game that was randomly selected to determine payouts.

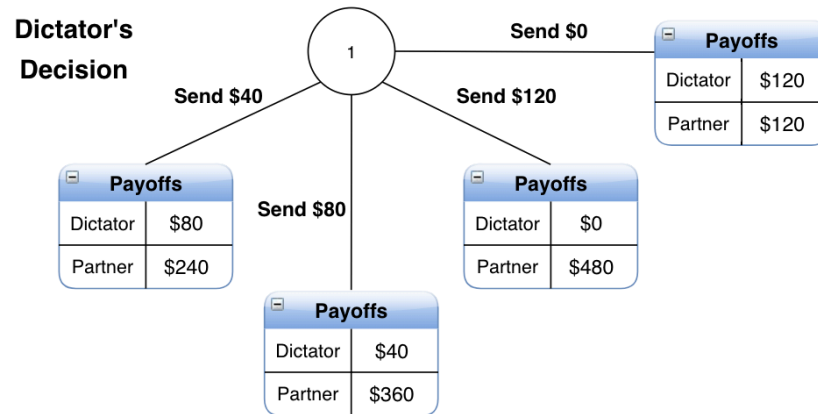


FIGURE 6: DICTATOR GAME DECISION TREE

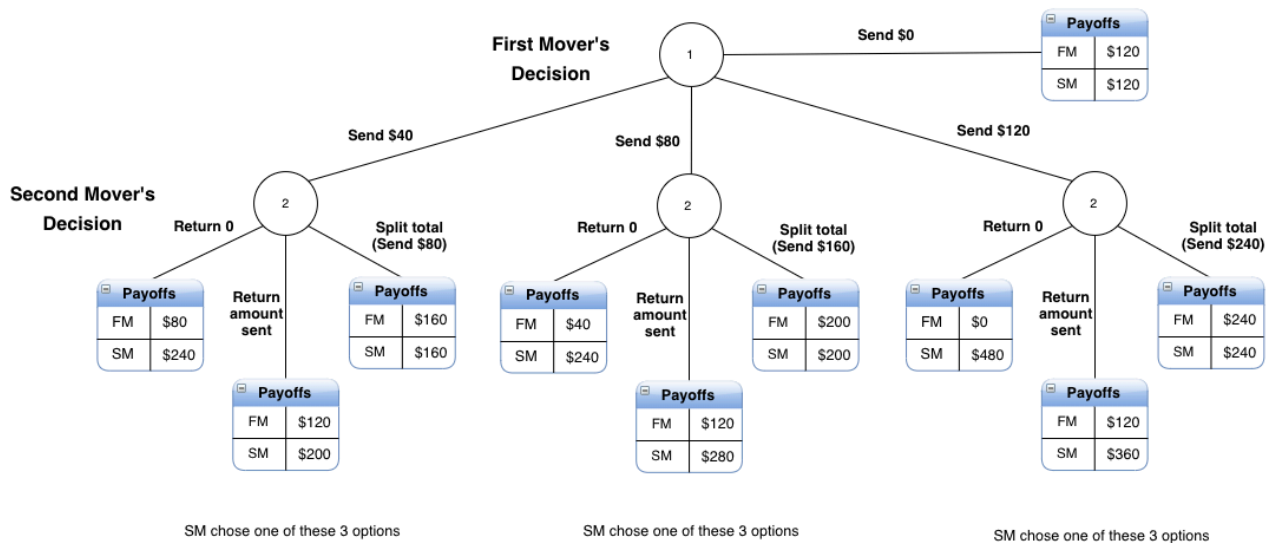


FIGURE 7: TRUST GAME DECISION TREE

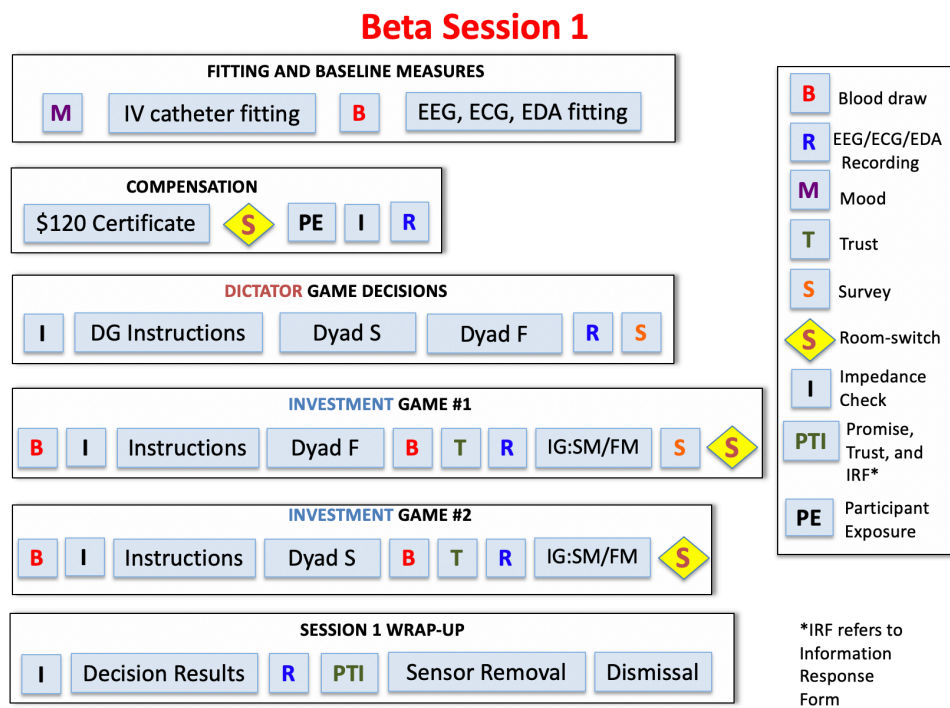


FIGURE 8: EXPERIMENTAL DESIGN: PROTOCOL B, SESSION 1

The study was conducted in 2012 and sponsored by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via a United States Air Force Research Laboratory grant. The study also gathered neuroendocrine responses (oxytocin, adrenocorticotrophic hormone, and cortisol) with blood draws at strategic points throughout the experiment. These measures were not analyzed for this paper.

IV. 2. Data Collection (Revisit Surveys)

Participants were greeted by a research assistant who explained the experiment, asked them to read and sign the study consent form. Participants were then instructed to wash their hands with non-detergent bar soap to remove residue and facilitate a good connection between the skin and electrodes. Next, two disposable Ag-AgCl electrodermal (EDA) electrodes were placed on the participants' index and middle fingertips of their non-dominant hand. Participants were asked to keep their hand steady for the duration of the study to reduce movement artifact.

EDA data were collected at a sampling rate of 1 kHz using an MP150 data acquisition system and Acknowledge software (BIOPAC; Goleta, CA). Electrodes were placed on the palmar surface of the middle phalanges of the index and middle fingers of the participant's non-dominant hand. 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.

Incomplete - Revisit after getting initial results

- Surveys collected: Not sure what I'll use; TBD after analyses

-Demographics

Data collection also consisted of demographic surveys...

From experiment paper:

Survey measures

Participants completed 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 demographic questionnaire, a Friendship Closeness survey, an Empathy survey, the Stephenson Multigroup Acculturation Scale, and an exit survey. Following their dyadic interactions, participants also indicated their subjective level of trust in their partner on a five- point scale.

Total: 8 surveys + demographics

IV. 3. Predicting Interpersonal Trust (Incomplete)

This analysis aims to predict the amount paired strangers earn in the trust game, based on the dyad's physiological synchrony. Can an individual's willingness and/or unwillingness to trust be classified based on their physiology? 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 these individuals display no correlation or a negative correlation with their partner? Can the pairs' physiological synchrony predict how much

money they will earn together? If so, which statistic of physiological synchrony most accurately predicts mutual trust, and thus the largest earnings?

Variables:

DV: Total amount earned (payoffs) by participant pair (total earned by dyad)

IVs:

- EDA time series (participants' physiology)
- Sex
- Partner pairings (same or different sex pairs)
- Control for altruism using the amount sent in the DG as a covariate (TBD)
- Additional demographic data TBD

Proposed Analysis 1 - Inter-Subject Correlation (ISC):

- Apply ISC analysis to classify EDA signals based on the amount each participant pairs made.
- Next, use bootstrapping as a robustness check to ensure distributions show statistical significance (comparing distributions of real ISC scores vs bootstrapped ISC scores).

Proposed Analysis 2 - Dynamic Time Warping (DTW):

- Apply dynamic time warping to classify EDA signals based on the amount each participant pair made.
- DTW analysis produces participant pair scores that can be classified using KNN, another clustering algorithm
 - 10 Clustering Algorithms (Python examples)
 - <https://machinelearningmastery.com/clustering-algorithms-with-python/>
#:~:text=Cluster%20analysis%2C%20or%20clustering%2C%20is,or%20clusters%20in%20feature%20space.

Proposed Analysis 3 - Neural Network (TBD)

"Theorists and researchers of trust may model the concept as an independent variable (cause), dependent variable (effect), or interaction variable (a moderating condition for a causal relationship). (Rousseau et al., 1998, p. 396).

When economic outcomes are of interest, researchers often conceptualize trust as a potential cause in choice scenarios framed around social dilemmas. High trust, perhaps based on previous experiences with a partner in a repeated game, tends to result in the decision to cooperate, which can lead to access to economic gains, as in the classic Prisoner's Dilemma (Axelrod, 1984; Miller, 1992). Trust, thus, is conceptualized as an independent variable" (Rousseau et al., 1998, p. 396).

IV. 3. 1. Hypotheses (Incomplete)

Introduction

Main hypotheses (for each statistic of physiological synchrony/analysis: ISC, DTW & KNN score):

H₁: Participant pairs that made the most money produced (displayed?) synchronous physiological signals, indicating trusting and trustworthy behavior.

H₂: Participant pairs that made the least money produced nonsynchronous physiological signals, indicating a lack of mutual trust. This may be no correlation or a negative correlation.

IV. 4. Data / Cleaning & Methods (Introduction)

Summary / Overview of cleaning necessary

- Necessary tests for time series ?

Summary / Overview of analysis to come in following sections: ISC, DTW + KNN, RNN ?

-Palumbo 2nd half of paper

Not really important....

Idiographic methods focus on an individual unit (e.g., one dyad or team), whereas nomothetic approaches combine data to assess group-level trends (e.g., multiple dyads or teams). Results from the two strategies only correspond when all conditions of the ergodic theorems are met (e.g., multivariate normal data with equal autocorrelation and trends across the data {Palumbo et al., 2017, Personality and Social Psychology Review, 21, 99-141})

IV. 5. EDA Data Cleaning

Using the Electrodermal Activity Analysis feature in AcqKnowledge (cite software)

transforms a given tonic EDA signal

What did you do to clean the data ~ within ACQ - Biopac lab manual

Incomplete

Copy Pasted from Jorge's Paper:

Following data collection, skin conductance waveforms were visually inspected for brief periods of signal loss, and data drop-offs shorter than 1 s in length were replaced with averages from adjacent parts of the waveform. Additionally, waveform noise due to experimenter-noted movement was smoothed using mean-value replacement from adjacent parts of the waveform. Next, a 10-Hz low-pass filter was applied to the waveform to remove high-frequency noise (Norris, Larsen, & Cacioppo, 2007), and a square root transformation was applied to adjust for skew inherent to skin conductance data (Dawson, Schell, & Filion, 1989; Figner & Murphy, 2011). After transformations, average skin conductance level (SCL) was extracted for the final 2 min of the baseline and for the time-span of the narrative. These values were used to calculate percent change in SCL from baseline to the narrative. For time series analyses, 1 s segments of SCL were taken from baseline and narrative stimulus. Non-specific skin conductance responses (NS-SCRs) were identified using a threshold of .01 μ S, and NS-SCR counts were taken for baseline, and narrative. Following extraction of NS-SCR counts, these values were used to calculate rate of NS-SCRs/min for baseline, narrative, and the three narrative segments.

(Golland et al., 2014):

The resulting two-Hz EDA and HR time series were manually examined for gross motion

artifacts and for detection of nonresponsive participants. This led to the exclusion of four participants that showed nonresponsive EDA signals. Two more participants were excluded due to gross artifacts, leaving 21 subjects for the subsequent analysis. For the remaining participants, the first 60 s were removed from the data to reduce nonspecific adaptation effects. EDA signals were smoothed, using a 10-s long moving average (Kettunen, Ravaja, & Keltikangas-Järvinen, 2000), and linear trends were removed. HR signals were low-pass filtered (below 0.04Hz) to remove the fast, idiosyncratic component of HR variability. Finally, all signals were z normalized.

Getting started (i) - pg. 16

Any analysis of a time-series signal begins with a visual inspection of the signal. Here the observer is looking at the overall characteristics and examining the integrity of the signal. This is best achieved by both examining the overall signal in the window (if the signal is not too long) and also zooming into the signal and moving through it in stages from start to finish. Crucially one is inspecting the signal for artefacts such as periods of poor contact or sharp square-wave spiking that may reflect contamination from an artificial source. These artefacts can be removed easily in a number of ways. Perhaps the simplest way of smoothing out and removing these artefacts is to down-sample the signal by a given constant (Transform > Smoothing > Samples). Care should be taken that the constant is sufficient to ensure a desired amount of smoothing without being so excessive as to distort the signal shape. On rare occasions noise / artefacts can be too excessive and so the data should be discarded. pg. 16

"An initial visual analysis"

(Braithwaite, Watson, Jones, & Rowe, 2015)

"Furthermore, sometimes it is necessary to smooth a signal due to factors such as noise, sudden deflections, etc. However, some signal smoothing algorithms require down-sampling." pg. 8

"With high sample rates one can remove a host of artefacts without altering the signal shape markedly and still have plenty of samples remaining for an appropriate analysis (the ideal scenario)." pg. 9

(Braithwaite, Watson, Jones, & Rowe, 2015)

"the FIR Filter (Low Pass) transformation is performed with Bartlett windowing, which reduces the signal amplitude" (BIOPAC Knowledge Base, 2020)

A low pass filter removes higher frequency variation from a lower frequency signal, improving the

signal quality. The selected frequency cutoff attenuates the components of the signal above a certain frequency. The frequency cutoff of the filter is set to 10Hz, which means the power levels of frequencies above 10Hz were diminished.

What did you do to clean the data ~ after you exported it cleaned from ACQ

- Baseline corrected: % change from baseline average
- Normalized, then standardized

IV. 6. Game Results and Statistics

IV. 6. 1. Results

The results of the study... preliminary?

Interaction pairs summary data... group/pairs payoff amounts

IV. 7. Inter-Subject Correlation (ISC) Analysis

Inter-subject correlation measures the strength of the EDA relationship between (each) individual participants and all others in the group. In this experiment ... pairs

Golland et al. (2014) applied an intersubject correlation (interSC) analysis to measure "the across-subject reliability of physiological responses by quantifying the commonalities of the response time courses among individuals..." (Golland et al., 2014, p. 1103). ISC ...

IV. 7. 1. EDA Inter-Subject Correlation (ISC)

IV. 7. 2. ISC Results

TABLE 1: ISC SUMMARY STATISTICS FOR PROTOCOL B, SESSIONS 1 & 2

IV. 8. Dynamic Time Warping (DTW) Analysis

(Berndet & Clifford, 1994)

(Keogh & Pazzani, 2002)

- compares each participant with every other participant
- takes the smaller of the two lengths

- squared

-absolute value

IV. 9. K-Nearest Neighbors (KNN) Analysis

The next step in

In order to categorize subjects by their ANS measures, we must group the DTW scores

IV. 10. Recurrent Neural Network (RNN) ?

V. (Conclusion) Comparing [Discussion of] Statistics of Physiological Synchrony

- Accuracy of the measures used; compare them; pros and cons of each

V. 1. Comparing Measurements [Statistics] of Synchrony

V. 2. Limitations

- Only one physiological measure was analyzed. A thorough study would also include an analysis of heart rate and EEG.

V. 3. Discussion / Conclusion

References

- Appelhans, B. M., & Luecken, L. J. (2006). Heart Rate Variability as an Index of Regulated Emotional Responding. *Review of General Psychology*, 10(3), 226-240.
- Bart, K. (1990). Fuzziness vs. Probability. *International Journal of General Systems*, 12(1), 211-240.
- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding Advantageously Before Knowing the Advantageous Strategy. *Science*, 275(5304), 1293-1295.
- Behrens, F., Moulder, R. G., Boker, S. M., & Kret, M. E. (2020). Quantifying Physiological Synchrony through Windowed Cross-Correlation Analysis: Statistical and Theoretical Considerations. *bioRxiv*.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10(1), 122-142.
- Berndt, D. J., & Clifford, J. (1994). Using Dynamic Time Warping to Find Patterns in Time Series. *Workshops on Knowledge Discovery in Databases*, 359-370.
- BIOPAC Knowledge Base. (2020). FIR Filter – Acqknowledge 3.9.2 (Mac). Retrieved Nov. 8, 2019, from <https://www.biopac.com/knowledge-base/fir-filter-acqknowledge-3-9-2-mac/>.
- Blume, L. E., & Easley, D. (2018). Rationality. In M. P. Ltd (Ed.), *The New Palgrave Dictionary of Economics* (Third ed.). London: Palgrave Macmillan.
- Bradford, A. (2017). Deductive Reasoning vs. Inductive Reasoning. Retrieved February 7, 2019, from <https://www.livescience.com/21569-deduction-vs-induction.html>.
- 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>.

- Butler, E. A. (2011). Temporal Interpersonal Emotion Systems. *Personality and Social Psychology Review*, 15(4), 367-393.
- Butte College. (2016). Tip Sheet: Deductive, Inductive, and Abductive Reasoning. Retrieved February 7, 2019, from <http://www.butte.edu/departments/cas/tipsheets/thinking/reasoning.html>.
- Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, N.J.: Princeton University Press.
- Cheong, J. H. (2019). Four Ways to Quantify Synchrony Between Time Series Data.
- Damasio, A. R. (1994). *Descartes' Error*.
- EconomicConcepts.com. (2015). Methods of Economic Analysis.
- 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.
- Ferrer, E., & Helm, J. L. (2013). Dynamical systems modeling of physiological coregulation in dyadic interactions. *International Journal of Psychophysiology*, 88(3), 296-308.
- Friedman, M. (1953). The Methodology of Positive Economics Essays in Positive Economics. In (pp. 3-43). University of Chicago Press.
- Garrath, W. (2016). Kant's Account of Reason" The Stanford Encyclopedia of Philosophy Tip Sheet: Deductive, Inductive, and Abductive Reasoning. In N. Z. Edward (Summer 2018 ed.).
- Glimcher, P. W., & Ernst, F. (2013). *Neuroeconomics: Decision Making and the Brain* (2nd ed.). Academic Press.
- Golland, Y., Keissar, K., & Levit-Binnun, N. (2014). Studying the dynamics of autonomic activity during emotional experience. *Psychophysiology*, 51(11), 1101-1111.

- Guru, S. (n.d.). Methods of Economic Analysis: Deductive Method and Inductive Method. Retrieved February 8, 2019, from <http://www.yourarticlelibrary.com/economics/methods-of-economic-analysis-deductive-method-and-inductive-method/36578>.
- Harry, J. G., Supriya, G., & Shane, H. (2002). *Introduction to Logic Methods of Economic Analysis: Deductive Method and Inductive Method Induction vs Deduction Economics*. Rutledge.
- Henderson, L. A. D. H. (n.d.). The Problem of Induction The Stanford Encyclopedia of Philosophy Logical Basis of Hypothesis Testing in Scientific Research. In N. Z. Edward(Spring 2019 ed.). Ohio State University.
- Herms, D. (n.d.). Logical Basis of Hypothesis Testing in Scientific Research. Retrieved February 7, 2019, from <http://www.dartmouth.edu/~bio125/logic.Giere.pdf>.
- Houser, D., & McCabe, K. (2014). Experimental Economics and Experimental Game Theory. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics: Decision Making and the Brain* (pp. 19-34). Elsevier Inc.
- Irving, C., Carl, C., Daniel, F., & Bradley, D. (2007). *Essentials of Logic Internet Encyclopedia of Philosophy: A Peer-Reviewed Academic Resource* (Second ed.). Taylor & Francis.
- Keogh, E., & Pazzani, M. (2002). *Derivative Dynamic Time Warping*. Proceedings from First SIAM International Conference on Data Mining.
- Levenson, R. W. (2014). The Autonomic Nervous System and Emotion. *Emotion Review*, 6(2), 100-112.
- Levenson, R. W., & Gottman, J. M. (1983). Marital Interaction: Physiological Linkage and Affective Exchange. *Journal of Personality and Social Psychology*, 45(3), 587-597.
- Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic Correlates of Perceived Therapist Empathy and Social-Emotional Process During Psychotherapy. *The Journal of Nervous and Mental Disease*, 195(2), 103-111.
- Marci, C. D., & Orr, S. P. (2006). The effect of emotional distance on psychophysiological

- concordance and perceived empathy between patient and interviewer. *Appl Psychophysiol Biofeedback*, 31(2), 115-128.
- Marieb, E. N. (2002). *Anatomy and Physiology*. San Francisco, CA: Pearson Education, Inc.
- 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.
- Pflanzer, R., & McMullen, W. (2012). Electrodermal Activity & Polygraph Introduction. In *BIOPAC Student Lab Manual* (pp. I-1). Goleta, CA: BIOPAC.
- 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.
- Sen, A. (2018). Rational Behaviour. In M. P. Ltd (Ed.), *The New Palgrave Dictionary of Economics* (Third ed.). London: Palgrave Macmillan.
- Sent, E.-M. (2018). Rationality, History of the Concept. In M. P. Ltd (Ed.), *The New Palgrave Dictionary of Economics* (Third ed.). London: Palgrave Macmillan.
- Shor, M. (2005). Dictionary of Game Theory Terms. *Non-Cooperative Game* 2020, from <http://www.gametheory.net/dictionary/Non-CooperativeGame.html>.
- Smith, V. L. (2003). Constructivist and Ecological Rationality in Economics. *American Economic Review*, 93, 465-508.
- von Neumann, J., & Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton, N.J.: Princeton University Press.