MODELING PSYCHOPHYSIOLOGICAL PROCESSES IN DYADIC INTERACTIONS

A Dissertation in
Human Development and Family Studies

by
Siwei Liu

© 2012 Siwei Liu

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2012
The dissertation of Siwei Liu was reviewed and approved* by the following:

Michael J. Rovine  
Professor of Human Development & Family Studies  
Dissertation Advisor  
Co-Chair of Committee

Peter C. M. Molenaar  
Professor of Human Development & Family Studies  
Co-Chair of Committee

Susan M. McHale  
Director, Social Science Research Institute  
Director, Children, Youth and Families Consortium  
Professor of Human Development & Family Studies

Runze Li  
Professor of Statistics and Public Health Sciences

Matthew S. Goodwin  
Assistant Professor of Health Sciences and Computer & Information Sciences

J. Douglas Coatsworth  
Professor-in-Charge of Graduate Program, Human Development & Family Studies

*Signatures are on file in the Graduate School.
ABSTRACT

Psychophysiological interactions constitute an important aspect of interpersonal relationships. Although behavioral scientists often conceptualize interpersonal units (e.g., dyad, family, or group) as dynamic systems, in which social partners are interwoven in terms of their physiology, psychology, and behavior, relatively little research has examined these dynamic processes, especially at the physiological level. This is partly due to the underdevelopment of statistical methods to analyze repeated measures data from interdependent individuals. The general purpose of this dissertation is to apply state-of-the-art methodology in longitudinal dyadic analysis to address empirical questions related to interpersonal physiological processes, and to develop new statistical approaches that researchers can use to model dyadic interactions.

This dissertation consists of three studies. Study 1 investigated coregulation of diurnal cortisol pattern in married couples using data from the Penn State Hotel Work and Well-Being project. Saliva samples containing cortisol were obtained from 28 heterosexual couples four times a day for four consecutive days. Multilevel modeling was conducted to examine whether husband and wife coregulate in their cortisol awaking responses (CAR) and diurnal cortisol slopes (DCS). Results showed synchrony (concurrent covariation) in DCS between spouses. For CAR, the strength of synchrony was stronger in couples characterized by higher levels of spousal strain and disagreement. In addition, cross-lagged models revealed that an individual’s diurnal cortisol pattern on a particular day depended on his/her own, but not his/her spouse’s, cortisol pattern on a previous day. Stability in CAR was stronger in couples reporting higher levels of spousal support. The discussion highlights the family systems perspective and the methodological advantages of the analytic technique.
Study 2 and 3 were based on a project examining dynamic interactions in physiological arousal between children with Sensory Processing Disorder (SPD) and their therapists, as indicated by electrodermal activity (EDA). Study 2 compared three approaches for handling missing data in multivariate time series, which is common in studies of physiological interactions. A recursive prediction routine developed by the author was found to outperform listwise deletion and data imputation using sample means and variances in recovering the parameters of the time series models. Study 3 introduced an advanced signal processing technique - time-frequency analysis (TFA) - to study interpersonal dynamic interactions. A program called SAM-MOW (Spectral Analysis for Multivariate data using a MOving Window) was developed by the author to carry out TFA automatically. The method was used to analyze EDA data from one child-therapist dyad during therapy. The analysis revealed a unidirectional influence from the therapist to the child (i.e., the child’s EDA was predicted by the therapist’s EDA). In addition, the magnitude of influence appeared to be contingent on guided activities during the therapy. The discussion focuses on the strengths and weaknesses of this novel methodological approach.

In sum, this dissertation focuses on the application and development of statistical methods to study interpersonal physiological interactions from a dynamic systems perspective. It contributes both to our knowledge of physiological processes underlying dyadic interactions, and to methodological development in this area.
# TABLE OF CONTENTS

LIST OF TABLES.................................................................................................................. viii

LIST OF FIGURES .................................................................................................................... ix

AKNOWLEDGEMENT ............................................................................................................. xi

CHAPTER 1 ............................................................................................................................. 1

Introduction ............................................................................................................................. 1

CHAPTER 2 ............................................................................................................................. 5

Coregulation of Diurnal Cortisol Pattern in Married Couples ........................................... 5

2.1 Introduction ....................................................................................................................... 5

2.1.1 Physiological Coregulation and Adult Attachment ................................................. 5

2.1.2 Using Cortisol to Study Physiological Coregulation between Spouses ............... 6

2.1.3 Modeling Diurnal Cortisol ....................................................................................... 8

2.1.4 Research Hypothesis ............................................................................................... 9

2.2 Methods ........................................................................................................................ 10

2.2.1 Participants .............................................................................................................. 10

2.2.2 Measures ............................................................................................................... 11

2.2.3 Analysis .................................................................................................................. 12

2.3 Results .......................................................................................................................... 15

2.3.1 Synchrony in Cortisol Awakening Response (CAR) ............................................. 15

2.3.2 Synchrony in Diurnal Cortisol Slope (DCS) .......................................................... 16

2.3.3 Cross-Lagged Relations in Cortisol Awakening Response and Diurnal Cortisol Slope ......................................................................................................................... 17

2.4 Discussion ....................................................................................................................... 18

CHAPTER 3 ............................................................................................................................ 25
Missing Data Procedures in Vector Autoregressive Modeling of Multivariate Time Series Data
........................................................................................................................................25

3.1 Background ...................................................................................................................... 25
  3.1.1 Vector Autoregressive Model .................................................................................. 27
  3.1.2 Methods for Handling Missing Data in Time Series ............................................. 28

3.2 Methods .......................................................................................................................... 30
  3.2.1 Simulation ................................................................................................................ 30
  3.2.2 Analysis ................................................................................................................... 32

3.3 Results ............................................................................................................................ 34
  3.3.1 Estimation of Complete Data .................................................................................. 34
  3.3.2 Directed Transfer Function from \(Y_1\) to \(Y_2\) (DTF21) .......................................... 35
  3.3.3 Directed Transfer Function from \(Y_2\) to \(Y_1\) (DTF12) ............................................. 37

3.4 Discussion ....................................................................................................................... 38

CHAPTER 4 ............................................................................................................................ 42

Dynamical Modeling of Child-Therapist Interactions in Electrodermal Activity Using
Time-Frequency Analysis ........................................................................................................42

4.1 Introduction ...................................................................................................................... 42
  4.1.1 Physiological Interactions in Psychotherapy ......................................................... 42
  4.1.2 Time Series Analysis in Time and Frequency Domains .................................... 43
  4.1.3 Transforming Time Domain Model into the Frequency Domain ....................... 45

4.2 Empirical Application .................................................................................................... 47
  4.2.1 The Study ............................................................................................................... 47
  4.2.2 Time-Frequency Analysis (TFA) ........................................................................... 49
  4.2.3 Results from One Therapy Session ................................................................. 51
  4.2.4 Interpretation and Confounding Factors ............................................................ 54

4.3 Discussion ....................................................................................................................... 57

CHAPTER 5 ............................................................................................................................ 60

Discussion ............................................................................................................................. 60
References ........................................................................................................................................... 65

APPENDIX A: Table 1 .................................................................................................................. 78
APPENDIX B: Table 2 .................................................................................................................. 79
APPENDIX C: Table 3 .................................................................................................................. 80
APPENDIX D: Table 4 .................................................................................................................. 81
APPENDIX E: Table 5 .................................................................................................................. 82
APPENDIX F: Table 6 .................................................................................................................. 83
APPENDIX G: Table 7 .................................................................................................................. 84
APPENDIX H: Table 8 .................................................................................................................. 85
APPENDIX I: Figure 1 .................................................................................................................. 86
APPENDIX J: Figure 2 .................................................................................................................. 87
APPENDIX K: Figure 3 .................................................................................................................. 88
APPENDIX L: Figure 4 .................................................................................................................. 89
APPENDIX M: Figure 5 .................................................................................................................. 90
APPENDIX N: Figure 6 .................................................................................................................. 91
APPENDIX O: Figure 7 .................................................................................................................. 93
APPENDIX P: Figure 8 .................................................................................................................. 94
APPENDIX Q: Figure 9 .................................................................................................................. 95
APPENDIX R: Figure 10 ............................................................................................................... 96
APPENDIX S: Figure 11 ............................................................................................................... 97
APPENDIX T: Figure 12 ............................................................................................................... 98
APPENDIX U: Figure 13 ............................................................................................................... 99
APPENDIX V: Figure 14 ............................................................................................................. 100
APPENDIX W: Figure 15 ............................................................................................................. 101
APPENDIX X: Figure 16 ............................................................................................................. 102
LIST OF TABLES

Table 1. Descriptive Statistics of Cortisol and Time of Collection. ........................................ 78

Table 2. Descriptive Statistics of Cortisol Awakening Response (CAR) and Diurnal Cortisol Slope (DCS), Broken Down by Gender. ......................................................................................... 79

Table 3. Descriptive Statistics and Correlations for Spouse Support, Strain, and Disagreement. 80

Table 4. Synchrony in Cortisol Awakening Response (CAR) between Spouses. ..................... 81

Table 5. Synchrony in Diurnal Cortisol Slope between Spouses.................................................. 82

Table 6. Cross-Lagged Models of Cortisol Awakening Response (CAR) and Diurnal Cortisol Slope (DCS) between Spouses .................................................................................................... 83

Table 7. Estimated VAR Models on Complete Data, Broken Down by Conditions. .................. 84

Table 8. Effects of Technique, AR_LARGE, and AR_POSITIVE on DTF21 ............................... 85
LIST OF FIGURES

Figure 1. Marital Relationship Moderates Synchrony in Cortisol Awakening Response (CAR). ......................................................... 86

Figure 2. Estimated Marginal Means of DTF21 by Technique, AR_LARGE, and AR_POSITIVE. Complete = Complete Data. Listwise = Listwise Deletion. Random = Imputation Using Sample Means and Variances. Recursive = Recursive Prediction. ............... 87

Figure 3. Estimated Marginal Means of DTF21 by Technique and AR_LARGE. Complete = Complete Data. Listwise = Listwise Deletion. Random = Imputation Using Sample Means and Variances. Recursive = Recursive Prediction ........................................ 88

Figure 4. Estimated Marginal Means of DTF21 by Technique and AR_POSITIVE. Complete = Complete Data. Listwise = Listwise Deletion. Random = Imputation Using Sample Means and Variances. Recursive = Recursive Prediction ........................................ 89

Figure 5. Estimated Marginal Means of DTF21 by Proportion of Missing Data .......................................................... 90

Figure 6. Estimated Marginal Means of DTF21, Broken Down by Technique, Proportion of Missing, AR_LARGE, and AR_POSITIVE......................................................................................... 92


Figure 8. Estimated Marginal Means of DTF12 by Proportion of Missing Data .......................................................... 94

Figure 9. Electrodermal Activity (EDA) Levels (μ mhos) from Child and Therapist during One Therapy Session (N=5338). ........................................................................................................... 95

Figure 10. Autospectra of the Child. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes. .................................................................................................................. 96

Figure 11. Autospectra of the Therapist. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes. .................................................................................................................. 97

Figure 12. Power Estimates of the First Frequency Component across Time ......................................................... 98

Figure 13. (a) Coherence between Child and Therapist across Time and Frequency. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes. (b) Average Coherence over Frequencies across Time ........................................................................................................... 99
**Figure 14.** Directed Transfer Function from the Therapist to the Child. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes. ................................................................. 100

**Figure 15.** Directed Transfer Function from the Child to the Therapist. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes. ................................................................. 101

**Figure 16.** Guided Activity and Directed Transfer Function between Child and Therapist Averaging over Frequencies. .................................................................................................................. 102
AKNOWLEDGEMENT

I am deeply indebted to many people who have made my graduate study a fulfilling and memorable experience. First and foremost, I would like to thank my advisor, Dr. Michael Rovine. I would not have achieved what I have done today without your guidance and support along the way. Thank you for patiently teaching me the basics and answering my naïve questions from the very beginning. Thank you for the long hours you spent meeting with me to discuss ideas, and guiding me through research and professional matters. Thank you for always believing in my abilities and encouraging me to pursue my goals. Thank you for offering help whenever it is needed. I am very fortunate to have such a supportive and nurturing advisor as you, and I will strive to follow your example as I mentor my own students in the future. I am also very grateful to the co-chair of my committee, Dr. Peter Molenaar. You have showed me what a great scholar is like. I have learned from you to enjoy and take pride in my work and to strive for excellence. These will keep me moving forward in the many years to come. I would like to thank other members on my committee, Drs Susan McHale, Runze Li, and Matthew Goodwin, and to Dr. David Almeida, who have made my dissertation work possible and provided valuable advice to make it better.

My labmates and friends have been a great source of support through my graduate study. Juan Peng, thank you for being there and sharing my laughter and tears for the past few years. Thanks to Ian Lam, for listening to my complaints and complaining with me when I was stressed out. Big thanks to Mary Lai Rose, for your kind heart and generous help to me and other international students. To all my other friends, my Penn State years would have been less joyful without your company. I will always remember our time together.
Finally, I would like to thank my parents for their unconditional love and care all these years. I can only pay you back by continuing to make you proud.

The first study of this dissertation was supported by the Alfred P. Sloan Foundation and conducted as part of the Work, Family and Health Network (WorkFamilyHealthNetwork.org), which is funded by a cooperative agreement through the National Institutes of Health and the Centers for Disease Control and Prevention: National Institute of Child Health and Human Development (Grant # U01HD051217, U01HD051218, U01HD051256, U01HD051276), National Institute on Aging (Grant # U01AG027669), Office of Behavioral and Science Sciences Research, and National Institute for Occupational Safety and Health (Grant # U010H008788). Special acknowledgement goes to Extramural Staff Science Collaborator, Rosalind Berkowitz King, Ph.D. and Lynne Casper, Ph.D. for design of the original Workplace, Family, Health and Well-Being Network Initiative. The second and third studies were supported by the 2012 dissertation award from the Society of Multivariate Experimental Psychology (SMEP). I sincerely thank Elliot Hedman, Richard Palumbo, and Dr. Wayne Velicer for their contributions to the third study.
CHAPTER 1

Introduction

Physiological dynamics constitute an important aspect of interpersonal interactions. Physiological responses such as cardiovascular reactivity (e.g., heart rate, blood pressure) and electrodermal activity (EDA) are closely associated with how we feel, think, and behave, which are contingent on social stimuli and crucial for our understanding of interpersonal relationships (Cacioppo, 1982; Grove, 2007). Empirical studies have found evidence that social partners’ physiologies move up or down together, a phenomenon known as physiological *synchrony* (Feldman, 2007) or *concordance* (Marci, Ham, Moran, & Orr, 2007). From a dynamic systems perspective of interpersonal relationships, the term *coregulation* has been used to describe the process by which autonomic physiological responses of individuals change in reaction to those of a social partner to maintain system (i.e., group) homeostasis. This process is thought to be the underlying mechanism for the formation of normative attachment relationships, such as mother-infant dyads and long-term romantic partners (Sbarra & Hazan, 2008). Hence, the study of psychophysiology provides a promising way to reveal important information about social interactions.

Our knowledge on interpersonal physiology, however, is very limited. Until recently, few studies have directly looked at the relations between individuals’ physiological activities during social interactions (e.g., Guastello, Pincus, & Gunderson, 2006; Marci et al, 2007). Moreover, these studies tend to be descriptive and focus on concurrent covariation. Many important questions in interpersonal physiology remain largely unanswered. For example, are the influences of the two parties on each other symmetric or unidirectional? Besides concurrent
relations, are there lagged relations such that an individual’s physiological state depends on that of the social partner at a previous time point? What are some possible correlates of interpersonal physiological interactions? From a developmental perspective, how does synchrony or coregulation in interpersonal physiology emerge and develop over time? Furthermore, due to limited research in this area, it is unknown whether and how results from previous studies generalize to a broader range of physiological measures, across individuals, and between social contexts. Addressing all these questions is beyond the scope of this dissertation. Instead, it is the hope of the author that the studies presented herein will provide new examples of how psychophysiological processes during dyadic interactions may be quantified, and that the methodology developed will be valuable for future research in this area.

Data analysis in the study of interpersonal physiology has been a challenge for researchers. Specifically, data required to address relevant research questions on this topic often involve repeated measures on at least two dimensions: repeated measures over time and repeated measures on the interpersonal unit (e.g., dyad, triad, family, or group). Repeated measures are interdependent, which violates the basic assumption of most traditional statistical methods and requires special treatments. The double dimensions of repeated measures further increase the complexity of modeling. Currently, analyzing longitudinal dyadic (or family/group) data in the behavioral sciences remains an advanced methodological topic and its application is rarely seen in empirical studies (Kenny, Kashy, & Cook, 2006).

This dissertation aims to apply state-of-the-art methodology in longitudinal dyadic analysis to address empirical questions related to interpersonal physiological processes, and develop new statistical approaches that researchers can use to model dyadic interactions. It consists of three studies. Chapter 2 used a series of multilevel models to examine coregulation of
cortisol in married couples and determine whether and how this process is moderated by characteristics of marital relationships. Cortisol is a steroid hormone controlled by the hypothalamic-pituitary-adrenocortical (HPA) axis, which regulates the body’s reaction to stress. The diurnal pattern of cortisol has been established as an index of neuroendocrine function and has important implications on health and aging (Dickerson & Kemeny, 2004; Piazza, Almeida, Dmitrieva, & Klein, 2010). Using repeated measures on cortisol obtained from a daily diary study, the current study examined whether husband and wife synchronize their physiological stress. Synchrony was operationalized as a within-family covariation: on days when an individual experiences more or less pronounced reactivity to stress, does his/her spouse also show more or less pronounced reactivity? Moreover, this study investigated whether and how perceived spousal support, strain, and disagreement moderate the degree of synchrony. In addition, cross-lagged regression models were used to investigate whether an individual’s stress could be predicted by his/her own stress and his/her spouse’s stress on a previous day, and whether and how these lagged influences depended on characteristics of the marital relationships. Results of this study provide evidence for the coregulation of neuroendocrine activities in married couples. The study also demonstrates the use of advanced multilevel modeling techniques to study dyadic interactions.

Chapter 3 presents a simulation study to compare procedures for handling missing data in multivariate time series. Time series analysis is widely applied in dynamic modeling (Valsiner, Molenaar, Lyra, & Chaudhary, 2009) and hence has great potential for the study of interpersonal interactions, which are often conceptualized as dynamic systems (Butler, 2011). Missing data are a common problem in time series physiological signals due to interruptions during recordings and/or removal of problematic values. The methods considered in this study included listwise
deletion, data imputation using sample means and variances, and data imputation using a recursive prediction routine developed in a study of child-therapist interactions, presented in Chapter 4. The three methods were evaluated in terms of their ability to recover parameter estimates in the frequency domain representing directional influences between time series.

Chapter 4 introduces an advanced signal processing methodology - time-frequency analysis (TFA) - to study dynamic interactions in EDA between individual children with Sensory Processing Disorder (SPD) and their therapists. Frequency analysis aims at revealing underlying cycles in time series data. Applying frequency analysis to the study of interpersonal physiology enables us to examine the degree of covariation, possible lead-lag relations, and directional influences between social partners. Furthermore, the adoption of time-frequency techniques allows us to examine change in these underlying dynamics over the course of a therapy session. In this study, a program called SAM-MOW (Spectral Analysis for Multivariate data using a MOving Window) was developed and used to conduct TFA on multivariate time series data automatically. The methods and programs presented in this study can be used by others to analyze data with similar structures, and thus contribute to the methodological repertoire available to study interpersonal interactions.
CHAPTER 2

Coregulation of Diurnal Cortisol Pattern in Married Couples

2.1 Introduction

Being in close relationships is a central experience for human beings. For many adults, living with a romantic partner or spouse provides essential social interactions and defines the basic context of everyday life. Despite many studies indicating the effects of a romantic partner or spouse on individual well-being (Koball, Moiduddin, Henderson, Goesling, & Besculides, 2010; Uchino, 2004), the underlying processes of how couples influence each other remain largely unknown. This study examines coregulation of stress hormones in married couples, and how these daily dynamics are associated with characteristics of the marital relationships. Using daily measurements of cortisol, a biomarker of the body’s response to stress, marital processes at the physiological level are investigated. The results provide implications for understanding the marital dyad system and the well-being of individuals in close relationships.

2.1.1 Physiological Coregulation and Adult Attachment

Physiological activities play an important role in the formation of attachment bonds. In human, attachment is thought to emerge when a caregiver consistently soothes an infant in stressful situations (Bowlby, 1958, 1982). Over time, the soothing from a caregiver facilitates the infant’s regulation of emotional and physiological arousal. Research has shown that skin-to-skin contact accelerates the autonomic functioning and neurobehavioral maturation in preterm babies (Feldman & Eidelman, 2003). In addition, synchrony in physiological systems and behaviors appears to be a crucial feature of attachment relationships. Contingency between maternal behavior and infant alertness is observed in the first hours after birth. Furthermore, parent-child
synchrony in infancy is found to be associated with a variety of developmental outcomes, including children’s self regulation ability, IQ scores, and attachment behavior (Feldman, 2007).

Researchers of close relationships have also argued that physiological coregulation and synchrony is a defining feature of adult attachment (Hazan, Gur-Yaish, & Campa, 2004). Coregulation refers to the process whereby one person up- or down-regulates the partner’s psychophysiological arousal (Sbarra & Hazan, 2008). Specifically, long-term romantic relationships are conceptualized as involving an interwoven physiological system which helps individuals maintain their psychological and physiological homeostasis. This mechanism is believed to underlie biobehavioral dysregulation such as sleep disruption that often occurs when one experiences separation from or loss of a long-term romantic partner (Sbarra & Hazan, 2008).

Despite speculation about physiological coupling between long-term romantic partners, research that directly addresses this issue is limited. Interrelatedness in physiology was found in married couples during lab managed conflictual interactions (Levenson & Gottman, 1983). Helm, Sbarra, and Ferrer (2011) also found that couples synchronized their respiration and heart rate during a series of lab tasks. In more natural settings, covariation and transmission of positive and negative affect have often been observed within marital relationships (e.g., Berg, Wiebe, & Butner, 2011; Butler, 2011; Ferrer & Nesselroade, 2003; Goodman & Shippy, 2002; Larson & Almeida, 1999).

2.1.2 Using Cortisol to Study Physiological Coregulation between Spouses

Diamond (2001) identified two biological systems that are particularly important for adult attachment relationships - the autonomic nervous system (ANS) and the hypothalamic-pituitary-adrenocortical (HPA) axis. The HPA axis controls our reactions to stress. Its end product, cortisol, has been established as a biomarker of stress reactivity and an indicator
of general neuroendocrine health (Hellhammer et al., 2007; Miller, Chen, & Zhou, 2007).
Cortisol increases in response to stressful stimuli and helps to increase blood sugar, providing energy needed to handle the stressor (Sapolsky, Romero, & Munck, 2000). This mechanism is adaptive in the face of acute stress. Chronic exposure to stress and frequent activation of the system, however, are associated with negative health outcomes such as depression (Belmaker & Agam, 2008; Bhagwagar, Hafizi, & Cowen, 2005) and cognitive decline (Seeman, McEwen, Singer, Albert, & Rowe, 1997). Cortisol is also known to be highly sensitive to social stimuli. Animal research has found abundant evidence that the presence of an attachment figure reduces cortisol secretion (e.g., Mendoza, Lyons, & Saltzman, 1991). In humans, social support from romantic partners has been shown to reduce cortisol reactivity induced by a public speaking task for men (Kirschbaum, Klauer, Filipp, & Hellhammer, 1995). Research in this line indicates that cortisol is a promising biomarker to look at when studying physiological coregulation between spouses.

Cortisol has a strong diurnal pattern. It typically peaks in the morning soon after waking, and gradually declines through the rest of the day. Previous research examining the association of cortisol between spouses has focused on the covariation in cortisol levels at a particular time of the day. For example, Schreiber and colleagues (2006) found a positive correlation between spouses in afternoon cortisol. Saxbe and Repetti (2010) obtained cortisol samples four times per day from married couples and found a positive association between husband and wife when sampling time was controlled for. These studies reveal important information about the interwoven neuroendocrine activities between spouses. However, both studies focused on the synchrony of cortisol levels at a particular time point, which is a static measure. It is possible that, besides synchrony in “state”, romantic partners also covary in their
movements in cortisol over time, or synchrony in “change”, which better represents regulation of the HPA axis. Moreover, both of the previous studies on this topic only considered concurrent relations between spouses’ cortisol and did not examine lagged relations. An individual’s biological stress may not only covary with his/her spouse’s stress at the same moment, but also depend on the spouse’s stress at previous time points. These are the research questions addressed in the current study.

2.1.3 Modeling Diurnal Cortisol

Researchers of cortisol have developed a variety of measures to quantify different aspects of its diurnal pattern. The most commonly used measures include cortisol awakening response (CAR) and diurnal cortisol slope (DCS). CAR refers to the surge in cortisol 30-45 minutes after waking. DCS measures the decline in cortisol from early morning to late evening. In some studies, DCS is modeled from the peak (e.g., Kudielka, Broderick, & Kirschbaum, 2003; Stawski et al., 2011), whereas in other studies it is modeled from waking (e.g., Adam, Hawkley, Kudielka, & Cacioppo, 2006). Researchers in favor of DCS from waking argue that decline from the peak may depend on the magnitude of CAR. In other words, individuals experiencing a more pronounced increase soon after waking are likely to experience a more dramatic decrease from the peak to the evening. Other researchers, however, speculate that decline from the peak may more accurately capture regulatory functioning of the HPA axis, because it reflects the body’s ability to “recover” from a high cortisol level. This study assesses coregulation between spouses in CAR and DCS from the peak values controlling for CAR on the same day. These measures are considered to be indices of general neuroendocrine health, with blunted CARs and flattened DCSs representing less healthy “profiles” (Adam & Kumari, 2009; Piazza et al., 2010).
Previous research also suggests that physiological coregulation between spouses is likely to be moderated by quality of the marital relationships. Levenson and Gottman (1983) found that married couples with lower marital satisfaction tend to show greater physiological concordance when discussing about problem areas in the marriage. They argued that “physiological linkage reflects the ebb and flow of negative affect, the escalation and de-escalation of conflict, and the sense of being ‘locked into’ the interaction and unable to ‘step back’ that can occur when spouses in dissatisfied marriages attempt to solve problems.”

Similarly, Saxbe and Repetti (2010) found that lower marital satisfaction is associated with a higher degree of coupling in cortisol between spouses. Therefore, in this study, measures of several marital relationship aspects are considered as moderators of coregulation. These include perceived support and strain from the spouse, and the degree of disagreement between spouses.

**2.1.4 Research Hypothesis**

This study investigates two possible ways that married couples coregulate in their patterns of diurnal cortisol - concurrent relations and lagged relations. Two sets of hypotheses are tested. First, it is hypothesized that husbands and wives synchronize their diurnal cortisol patterns over days. Synchrony is operationalized as a within-couple association; that is, on days when an individual experiences a steeper CAR or DCS than his/her average level, the spouse also experiences a steeper CAR or DCS than his/her average level, and vice versa. The moderating effects of three measures of marital relationships (spousal support, strain, and disagreement) on the strength of synchrony are also examined. Second, it is hypothesized that lagged relations exist between spouses, such that an individual’s diurnal cortisol pattern (CAR and DCS) on a particular day is predicted by the spouse’s cortisol pattern on the previous day.
Again, moderating effects of three different aspects of marital relationships on the lagged coefficients are investigated.

2.2. Methods

2.2.1 Participants

Data were drawn from the Penn State Hotel Work and Well-Being Study (Almeida, Davis, & Crouter, 2012), a project investigating connections between work stress, health, and family relations of employees from the hotel industry. Hotel managers completed a baseline interview and provided information about their work arrangement, family life, and so on. Respondents who were married or in a cohabiting relationship were invited to participate in a daily diary study together with their spouses/partners. The daily diary study aimed to obtain in-depth information about daily work experiences and personal well-being. Specifically, after a baseline interview, hotel managers and their spouses/partners were telephoned on eight consecutive days and asked about their daily experiences, including time use, physical symptoms, mood, stressful events, and so on. On Day 2 to Day 5, a subset of couples participated in a biomarker study, in which they provided saliva samples containing cortisol.

This study used data from 28 heterosexual couples that participated in the biomarker study. Men in this sample had an average age of 38.32 years, with a standard deviation (SD) of 9.58, and 57.1% had a college degree or more. Women had an average age of 36.11 years (SD = 9.20), and 60.7% had a college degree or more. Seventy-three percent of the participants were Caucasian, 16% were Hispanic, 5.5% were Black, and 5.5% were Asian. The couples had been married or living together for an average of 9.47 years (SD = 10.23) and had an average household income of $87,393 (SD = 30,279).
2.2.2 Measures

_Cortisol._ Cortisol levels were measured from saliva samples that participants provided four times a day (upon wakening, 30 minutes after wakening, early afternoon, and before bed at the end of the day) for four consecutive days (Day 2, 3, 4, and 5 of the daily interview). Saliva samples were excluded from the analysis if participants woke up after 12 pm on that day, if they were taking medicines known to affect cortisol secretion (e.g., Estrogen, Depo-Provera), if they were pregnant/breastfeeding, or if they had been pregnant/breastfeeding in the past year. They accounted for 17.9% of the total person-days and were all taken from women. Samples with cortisol > 60 nmol/L were considered invalid. Therefore cortisol values were top coded to 60 nmol/L. They accounted for 1.4% of the total cortisol values. Descriptive statistics of cortisol and the collection time of saliva samples are shown in Table 1.

CAR was computed by subtracting sample 2 from sample 1 on the same day and dividing the difference by the time lag between the two samples (Almeida, Piazza, & Stawski, 2009). Because CAR is known to be sensitive to participant compliance (Kudielka, Broderick, & Kirschbaum, 2003), it was excluded if the time lag between the first and second samples was smaller than 15 minutes or larger than 60 minutes. These excluded values made up 3.6% of all CAR values. DCS was computed by subtracting the bedtime value (sample 4) from the peak value (sample 2) and dividing the difference by the time lag between them. A flag was created if there was an increase between sample 2 and 3 that was larger than 10 nmol/L, as this would

1 Each individual contributed one person-day each day when he or she provided at least one saliva sample. There were 223 person-days in total.
indicate that cortisol assessed in sample 2 did not reflect peak value of the day. These problematic values constituted 1% of all DCS values and were excluded from the analysis. The means of valid CAR and DCS, broken down by gender, are presented in Table 2. The time unit is in hours, therefore these values represent changes in cortisol level per hour.

*Spousal support, strain, and disagreement.* Both the hotel managers and their spouses/partners answered questions about their perception of spousal support, strain, and disagreement in the baseline interview prior to the daily diary study. The three scales used were adapted from the Midlife in the United States (MIDUS) study (Grzywacz & Marks, 1999; Schuster, Kessler, & Aseltine, 1990). The spousal support scale contained seven items, such as “*How much does your spouse or partner really care about you?*” The spousal strain scale had six items, such as “*How much do you feel your spouse or partner makes too many demands on you?*” The spousal disagreement scale had five items, including “*How much do you and your spouse or partner disagree on money matters such as how much to spend, save, or invest?*” They were answered on a 4-point scale from “*not at all*” to “*a lot*”. Item scores in each scale were summed for each individual and then averaged over spouses within each family to create overall indices of the marital relationship for each couple. Descriptive statistics of these variables are presented in Table 3.

### 2.2.3 Analysis

Because the data were nested at multiple levels (across days and within dyads), multilevel modeling (MLM) was the appropriate statistical framework. To test the first hypothesis, cortisol slopes (i.e., CAR or DCS) for each person on a particular day was modeled as an intercept (a person-mean over the four days) plus deviation from the intercept. The
deviation was then predicted by the spouse’s deviation from his/her person-mean. The coefficient thus represents strength of synchrony in diurnal cortisol patterns. The basic model can be written as:

\[
\begin{align*}
\text{Level 1:} & \quad Y_{ti} = \beta_{0i} \text{HUSBAND}_i + \beta_{1i} \text{WIFE}_i + \beta_{2i} \text{SPOUSE}_{ti} + \varepsilon_{ti} \\
\text{Level 2:} & \quad \beta_{0i} = \pi_{00} + \upsilon_{0i} \\
& \quad \beta_{1i} = \pi_{10} + \upsilon_{1i} \\
& \quad \beta_{2i} = \pi_{20} + \upsilon_{2i}
\end{align*}
\]

(1)

where \(Y_{ti}\) represents the cortisol slope (CAR or DCS) of couple \(i\) on day \(t\), \(\text{HUSBAND}_i\) and \(\text{WIFE}_i\) are dummy coded indicators for husband and wife, and \(\text{SPOUSE}_{ti}\) is the spouse’s person-centered cortisol slope of couple \(i\) on day \(t\). The use of two dummy codes at level 1 to identify the individuals allows the estimation of separate intercepts for husband and wife, a technique commonly adopted in dyadic analysis (Kenny, Kashy, & Cook, 2006; Laurenceau & Bolger, 2005; Raudenbush, Brennan, & Barnett, 1995). Therefore, \(\beta_{0i}\) and \(\beta_{1i}\) represent the average cortisol slope for the husband and wife in couple \(i\). Because \(\text{SPOUSE}_{ti}\) is the difference between the spouse’s cortisol slope on day \(t\) and his/her person-mean, and it is used to predict the residual of an individual’s cortisol slope on day \(t\) after partialling out the intercept, \(\beta_{2i}\) represents strength of synchrony between the husband and wife in couple \(i\). At level 2, the level 1 coefficients are decomposed into a sample mean and a random component representing the difference between the couple and the whole sample. Therefore, \(\pi_{00}\) and \(\pi_{10}\) are the average cortisol slopes for the men and women in the sample. \(\pi_{20}\) is the average strength of synchrony across couples. If \(\pi_{20}\) is significant, we can say that when an individual’s spouse had a steeper slope on day \(t\) compared to other days, the individual also tended to have a steeper slope on day \(t\) compared to other days, and vice versa. The random components for the intercepts are allowed to
be correlated, and the correlation represents a between-couple association between husbands and wives. The variance for the synchrony coefficient is also estimated, but if the term is not significant, it is fixed to zero. For CAR, waking time was controlled for because it was likely to affect CAR (William, Magid, & Steptoe, 2005). CAR was controlled for in the model of DCS. These were done by adding person-means of the control variables to level 2 and within-person deviations to level 1. For example, person-means of waking time were constructed by taking the average of waking time for each individual over the four days, and used to predict the intercept of the individual (level 2). Within-person deviations of waking time were the differences in waking time between day \( t \) and person-means, and were used to predict CAR on day \( t \) (level 1).

After the basic model was fit, measures of marital relationships were added into level 2 to test their moderating effects on strength of synchrony. For example, the equation where spousal support moderated the synchrony coefficient can be written as:

\[
\beta_{2i} = \pi_{20} + \pi_{21}\text{Support}_i + \nu_{2i} \quad (2)
\]

To test the second hypothesis of lagged relations, cross-lagged regression analysis was conducted in the multilevel modeling framework. The cross-lagged models tested whether an individual’s cortisol slope was influenced by the spouse’s cortisol slope on the previous day, after controlling for the individual’s cortisol slope on the previous day. The basic model can be written as:

**Level 1:**

\[
Y_{ti} = \text{HUSBAND}_i [\beta_{0i} + \beta_{2i} Y_{h(t-1)i} + \beta_{4i} Y_{w(t-1)i}] + \text{WIFE}_i [\beta_{1i} + \beta_{3i} Y_{w(t-1)i} + \beta_{5i} Y_{h(t-1)i}] + \epsilon_{ti}
\]

**Level 2:**

\[
\begin{align*}
\beta_{0i} & = \pi_{00} + \nu_{0i} \\
\beta_{1i} & = \pi_{10} + \nu_{1i} \\
\beta_{2i} & = \pi_{20} + \nu_{2i}
\end{align*}
\]
\[ \beta_{3i} = \pi_{30} + \nu_{3i} \]
\[ \beta_{4i} = \pi_{40} + \nu_{4i} \]
\[ \beta_{5i} = \pi_{50} + \nu_{5i} \]  

(3)

In this model, \( Y_{h(t-1)i} \) and \( Y_{w(t-1)i} \) represent cortisol slopes for husband and wife in couple \( i \) on day \( t-1 \), where \( t \) is an integer from 3 to 5. Hence, \( \beta_{0i} \) and \( \beta_{1i} \) are intercepts for the husband and wife in couple \( i \). \( \beta_{2i} \) and \( \beta_{3i} \) are lag-1 autoregressive coefficients representing stability of cortisol slopes for the husband and wife in couple \( i \). \( \beta_{4i} \) and \( \beta_{5i} \) are cross-lagged coefficients representing the influence of the spouse’s cortisol slopes on a previous day on one’s cortisol slopes, after controlling for one’s own cortisol slopes on a previous day in couple \( i \). At level 2, the variances and covariances of these coefficients are tested, and non-significant random components are fixed to zero. Gender differences are tested to examine whether men and women differ in their stability and in their susceptibility to their spouses. If gender difference is not significant, one single coefficient is estimated for both men and women. Finally, spousal support, strain, and disagreement are added to level 2 to examine whether they moderate the stability or cross-lagged coefficients. An example of the multilevel lagged regression model was given by Rovine and Walls (2006), wherein a univariate case was considered.

2.3 Results

2.3.1 Synchrony in Cortisol Awakening Response (CAR)

Results from the models examining synchrony in CAR between spouses and the moderating effects of marital relationships are summarized in Table 4. Looking at the baseline model (with no moderator), the estimated intercepts for men and women in the sample were 4.19 and 15.68, respectively. A post-hoc comparison indicated that gender difference in the intercept
was significant \( (p < .05) \). Hence, women in the sample tended to have more pronounced CARs than men. This is consistent with findings of Almeida, Piazza, and Stawski (2009) using a national sample of 1,143 adults. Wake time did not predict CAR either at the between-person or within-person level. The coefficient associated with spouse’s deviation from person-mean was also not significant, indicating couples did not show synchrony in CAR when averaged over the whole sample. However, tests of moderating effects of the three measures of marital relationships revealed that spousal strain and disagreement significantly predicted degree of synchrony \( (p < .01) \). Figure 1 shows the plots of the moderations. Couples with greater spousal strain and disagreement (one standard deviation above the average) showed a positive association in CAR, whereas couples with lower than average spousal strain and disagreement showed a negative association. In other words, couples characterized by high levels of marital tension travelled up and down together in their CAR, whereas couples with low levels of marital tension showed a complementary pattern, such that when one had a steeper rise in cortisol after waking than usual, his/her spouse showed a less pronounced response than usual.

### 2.3.2 Synchrony in Diurnal Cortisol Slope (DCS)

Table 5 summarizes the results from the models looking at synchrony in DCS between spouses. Both men and women had negative intercepts \( (p < .001) \), indicating that cortisol did show a declining pattern throughout the day. Post-hoc analysis indicated no significant gender difference in average DCS. CAR was negatively related to DCS, both at the between- and within-person levels. In other words, individuals who showed more pronounced CARs, on average, tended to have steeper declines in their cortisol for the rest of the day than individuals with less pronounced CARs. In addition, on days when an individual experienced a more
pronounced CAR than his/her usual level, he or she also tended to experience a steeper decline in cortisol. The coefficient associated with spouse’s deviation from person-mean was positive and significant ($p < .05$). This suggests that spouses synchronized their cortisol decline slopes. On days when an individual experienced faster or slower cortisol decline compared to his/her usual level, the spouse also experienced faster or slower decline compared to his/her usual level. None of the measures of marital relationships significantly moderated strength of synchrony. Thus, the significant coupling in DCS between spouses did not vary in a systematic way depending on spousal support, strain, or disagreement.

2.3.3 Cross-Lagged Relations in Cortisol Awakening Response and Diurnal Cortisol Slope

Results from the cross-lagged regression models for CAR and DCS are shown in Table 6. A series of models were tested for each dependent variable as outlined in the methods section, but only the final models are presented. For CAR, there was no gender difference in the lag-1 autoregressive coefficient or the cross-lagged coefficient. Spousal support significantly moderated the lag-1 autoregressive coefficient ($p < .05$). For couples with spousal support one standard deviation above the average, their expected lag-1 autoregressive coefficient was 0.64, whereas for couples whose perceived spousal support was one standard deviation below the average, the expected coefficient was 0.20. Thus, higher levels of spousal support were associated with higher stability in CAR. The cross-lagged coefficient, however, was not significant and not moderated by any measure of marital relationships. This indicates that spouse’s CAR on a previous day did not net any significant influence on one’s CAR, after controlling for one’s own CAR on the previous day.
Looking at DCS, we see a marginally significant lag-1 autoregressive coefficient ($p < .1$), and the cross-lagged coefficient was not significant. There were no gender differences in these coefficients. The moderating effects of the measures on marital relationships were also not significant.

2.4 Discussion

This study examined the daily processes in diurnal cortisol patterns between spouses in a naturalistic setting. Diurnal cortisol pattern is conceptualized as consisting of an awakening response (CAR), characterized by an increase in cortisol level 30-45 minutes after waking, and a decline from the peak value to the bedtime value (DCS). It was found that spouses synchronize their DCS, such that on days when one experiences faster or slower decline in his/her cortisol level than usual, the spouse also experiences faster or slower decline than usual. For CAR, the positive association is only observed in couples reporting high levels of strain and disagreement with their partners. Couples with low levels of spousal strain and disagreement show a negative association, reflecting a complementary pattern in their CARs.

These results are consistent with previous work by Saxbe and Repetti (2010), which found coupling in cortisol values between spouses after controlling for sampling time\textsuperscript{2}. Together, these findings provide evidence that long-term romantic partners are interwoven in their endocrine responses to stress. Although conjectures of this phenomenon have long existed and

\textsuperscript{2} In Saxbe and Repetti’s (2010) study, cortisol was measured early morning upon wakening, late morning just before lunch, afternoon just before leaving work, and evening before going to bed. These measures do not capture CAR.
research on the transmission of stress within family has received considerable attention in recent years, this is the first study after Saxbe and Repetti’s pioneering work to look at coupling between spouses in stress at the physiological level. Moreover, this study extends the previous study and shows that husbands and wives not only covary in their cortisol levels at a particular time point, but also in their diurnal cortisol pattern. This synchrony in movement provides strong support for physiological coregulation in the marital dyad system. Given that diurnal cortisol pattern is associated with a large variety of psychosocial and health outcomes, including immunologic function, coronary calcification (Matthews, Schwartz, Cohen, & Seeman, 2006), burnout (Pruessner, Hellhammer, & Kirschbaum, 1999), depressive symptoms (Bhagwagar, Hafizi, & Cowen, 2005), and cognitive functioning (Stawski et al., 2011), this study also provides insights to the processes linking chronic stress to the physical and mental well-being of individuals in long-term romantic relationships.

The finding that couples characterized by higher levels of spousal strain and disagreement tend to show stronger synchrony in their CAR is also consistent with previous research. Saxbe and Repetti (2010) found that individuals who were less satisfied with their marriage were more reactive to fluctuations in their spouses’ cortisol. When discussing problem areas in their marriage, more distressed couples tended to show higher degree of “interrelatedness” in their physiology and emotions, reflecting escalated negative interactions (Levenson & Gottman, 1983). The stronger positive physiological linkage among distressed couples can be explained in two ways - perhaps poor marital quality exacerbates the impact of spouse’s stress on an individual, or perhaps individuals who are easily affected by their spouses’ stress tend to report poor marital quality. In this study, it is also found that couples reporting low levels of spousal strain and disagreement actually showed negative synchrony in their CARs.
This indicates that when an individual experienced higher than average reactivity to stress upon wakening, his/her spouse tends to be less reactive. Somehow this complementarity in CAR seems to be helpful in keeping couples away from marital tension.

Interestingly, the differential pattern in synchrony is only observed in CAR, but not in DCS. Although the physiological significance of CAR and DCS is not completely understood, some researchers have argued that they may reflect distinct biological and neuroendocrine processes (Clow, Thorn, Evans, Hucklebridge, 2004). Some evidence suggests that CAR is sensitive to the anticipation of a potentially stressful day (Kunz-Ebrecht, Kirschbaum, Marmot, & Steptoe, 2004; McHale et al., in press). It is possible that distressed spouses travel together in their anticipated stress on a day-to-day basis because of shared anticipation of family stressors (e.g., argument with each other). In low marital tension couples, in contrast, perhaps when an individual is under stress, the spouse is able to keep calm and soothe the individual. Due to limited sample size, there is not enough power to test whether the negative coefficient in low marital tension couples is significant. Future research with a larger sample size is needed to investigate this interesting phenomenon.

Cross-lagged regression analysis indicates that an individual’s diurnal cortisol pattern on a particular day is not associated with the spouse’s diurnal cortisol pattern on the previous day, if the individual’s own diurnal pattern on the previous day is taken into account. There is, however, considerable stability in diurnal cortisol pattern. Averaging across individuals, both CAR and DCS have a positive lag-1 autoregressive coefficient, indicating that a steeper slope on day \( t \) is predictive to a steeper slope on day \( t+1 \), and vice versa. In addition, stability in CAR is found to be stronger in couples characterized by higher levels of spousal support. Given that CAR is linked to anticipated stress of the day, this relation indicates that individuals who receive
more support from their spouses tend to have higher regularity in their anticipated stress from
day to day. On the other hand, it may also indicate that high levels of spousal support buffer the reaction of the HPA axis to stressful events.

This study contributes to the literature in a number of ways. Despite theories and empirical evidence linking physiological coupling and adult attachment (Sbarra & Hazan, 2008), few studies have explicitly examined the interrelatedness between spouses at the physiological level, and fewer have focused on the HPA axis which plays a significant role in health and aging. This study fills a gap in the literature and suggests that research of this type is promising. Similar studies may help us better understand the effects of marriage or a close partner on individuals’ psychological and physical well-being. Researchers may also extend the current framework to explore relations between health disparity across families and their exposure to differential levels of stress by conducting longitudinal studies with a greater time span. In addition, this study highlights the utility of family system theory, which conceptualizes marital dyads as self-organizing dynamic systems in which spouses mutually influence each other. According to this perspective, family processes evolve in a reciprocal way such that short-term interactions both depend on and reinforce relatively stable characteristics of relationships. Results from this study are in line with this notion. Specifically, the strength of coupling between spouses in their stress was associated with characteristics of the marital relationships, including spousal strain and disagreement, and the degree of stability in CAR was linked to the general level of spousal support in the relationships.

Methodologically, this study demonstrates a way to capture the coupling of a construct within the unit of analysis (e.g., family, group) using relatively few data points. The methods used represent an extension to the standard multilevel dyadic model (Kenny, Kashy, & Cook,
2006; Laurenceau & Bolger, 2005; Raudenbush, Brennan, & Barnett, 1995), where covariation between dyad members is modeled as covariance. In this study, synchrony between spouses is modeled as a fixed effect, instead of a random effect. This is achieved by computing a spouse deviation score from his/her person-mean, and using it as a predictor for the dependent variable after controlling for the person-mean (i.e., the intercept). This extension greatly increases the flexibility of the model in studying within-family processes, because it allows the degree of synchrony to be predicted by between-family covariates, such as measures of marital relationships used in this study. In other words, whereas the standard multilevel dyadic model provides a way to examine *whether* coupling exists between dyad members, the extended approach presented in this study enables us to further investigate *what* factors explain the variation in the strength of coupling. The same technique has been applied in Butner, Diamond, and Hicks (2007) to look at coupling in positive and negative affect between spouses and moderating effects of attachment style.

It should be noted that the models presented in this study also differ from the analysis in Saxbe and Repetti’s (2010) earlier work, although the data in the two studies have similar structures. In the previous study, cortisol data measured over days were collapsed and treated as measures obtained on one single day. The assumption behind the analysis is that the day-to-day variation is negligible when looking at synchrony in cortisol at a particular time point. This strategy is common in cortisol research (Adam, 2006; Adam & Gunnar, 2001; Saxbe, Repetti, & Nishina, 2008). In contrast, day-to-day variation plays a key role in the current study. In fact, synchrony was conceptualized as husbands and wives moving up or down together in their day-to-day variation. The differences in the model set-up have allowed us to address the same research question in different ways, which is illuminating.
Finally, there are several caveats in the current study. First, the sample size is rather small and participants were recruited from a single industry. Statistically, the small sample size has limited the power of between-family analysis. With a larger sample, we may be able to reveal significant effects of other between-family covariates. The relatively high homogeneity of the sample has also prevented us from generalizing the findings to a larger population. Given a more diverse sample, we may be able to investigate some other interesting questions, such as whether the strength of physiological synchrony depends on the stage of the relationships (e.g., people who just started dating vs. people who have been together for a long time) or sexual orientation (heterosexual vs. homosexual couples). Second, this study is cross-sectional and does not allow us to disentangle the direction of influences between strength of synchrony and characteristics of the marital relationships. Although measures of marital relationships were obtained prior to the cortisol data, they are likely to be relatively stable over time and may result from the history of how spouses influence each other in their physiology. Or, there may be bidirectional influences between characteristics of the marital relationships and physiological synchrony. Longitudinal data would help us determine the direction of the relation and gain deeper insight to spousal regulation. Finally, although the measurement of cortisol is intensive compared to most behavioral research and studies in the neuroendocrine health area, the number of repeated measures from each marital dyad does not allow us to model synchrony as a time-varying phenomenon. It is possible, for example, that degree of synchrony changes with the occurrence of some important life events (e.g., the birth of a child). More intensive data collection will allow us to investigate this interesting question.

In sum, this study provides empirical support for coregulation in neuroendocrine reactivity to stress in married couples. It contributes both to our knowledge of marital processes
and biobehavioral health, and to the advancement of methodology for longitudinal dyadic analysis. Future research with a larger and more diverse sample and/or more intensive measurements appears to be very promising.
CHAPTER 3

Missing Data Procedures in Vector Autoregressive Modeling of Multivariate Time Series Data

3.1 Background

A time series is a sequence of data points measured repeatedly from the same entity over time. Whereas in traditional longitudinal studies, time is fixed and subjects are random, in time series studies subjects are fixed and time is random. Accordingly, typical longitudinal analysis pools information across individuals, whereas time series analysis pools information across time. Because time series data are of greater length, they contain more detailed information about how an entity changes or develops. Therefore, time series design and analysis are uniquely suited for research aiming at investigating behavioral sequences or the dynamics of a physiological or psychological process (Velicer & Fava, 2003). Time series analysis has been applied in the study of smoking behavior (Velicer, Redding, Richmond, Greeley, & Swift, 1992), treatment adherence patterns (Aloia et al., 2008), affective processes (Molenaar, Sinclair, Rovine, Ram, & Corneal, 2009; Steele & Ferrer, 2011), social interactions (Gottman & Ringland, 1981), and many other aspects in the study of human behavior and development. Despite its fruitful results and increasing popularity, relatively little attention has been paid in the behavioral sciences to methodological issues related to the analysis of time series data (for exceptions, see e.g., Chow, Ferrer, & Nesselroade, 2007; Du Toit & Browne, 2007; Molenaar & Nesselroade, 2001; Zhang, Chow, & Ong, 2010).

This study examines procedures for handling missing data in time series analysis. Because time series data are obtained from the same individual(s) repeatedly, missing data are a
common problem. In behavioral studies where participants are usually measured on a daily basis over a number of weeks, missing data often occur when participants fail to comply with the research protocol. In studies involving physiological measures such as electrodermal and cardiovascular activity, missing values may result from technical problems during recording (e.g., disconnection of the sensors) or removal of problematic values that exceed a realistic range. Methods for handling missing data in time series have received considerable attention in the engineering and econometric literatures. Typical approaches include estimation of model parameters using maximum likelihood or least square methods (e.g., Dunsmuir & Robinson, 1981; Harvey & Pierse, 1984; Jones, 1980; Ljung, 1989) and missing value interpolation (e.g., Lu & Hui, 2003; Sorjamaa, 2010). In the behavioral sciences, however, little research has been conducted on this topic and researchers have relied on relatively naïve but convenient methods such as listwise deletion and mean imputation to handle missing data.

This study uses simulation to compare three methods for handling missing data in multivariate time series analysis. Simulation is ideal for missing data research because the mechanism of missingness and the distribution of the data are known. Moreover, complete data are available to be compared with results from the analysis of incomplete data. In the following, I first present the vector autoregressive (VAR) model, which provides the framework for this study. Next, I review popular methods in the behavioral sciences for handling missing data in time series and introduce the design of the current study. Finally, results are presented and the implications for the analysis of empirical data are discussed.
3.1.1 Vector Autoregressive Model

In recent years, multivariate time series analysis has gained increasing popularity in psychology because it allows researchers to examine dynamic relations between multiple constructs or people over time. One of the basic approaches for analyzing stationary multivariate time series is the vector autoregressive (VAR) model, which is an extension to the univariate autoregressive (AR) model. The general form of a VAR model of order \( p \) is:

\[
\mathbf{Y}_t = \mathbf{V} + A_1 \times \mathbf{Y}_{t-1} + A_2 \times \mathbf{Y}_{t-2} + \ldots + A_p \times \mathbf{Y}_{t-p} + \mathbf{U}_t
\]

(1)

where \( \mathbf{Y}_t \) is an \( m \times 1 \) vector of observed data at time \( t \), \( A_k \) are \( m \times m \) matrices of autoregressive coefficients at lag \( k \), \( \mathbf{V} \) is an \( m \times 1 \) vector of intercepts, and \( \mathbf{U}_t \) represents an \( m \)-dimensional white noise process with \( E(u_t) = 0, \ E(u_t,u_{t'}) = \Sigma_u \), and \( E(u_t,u_{t'}) = 0 \) for \( s \neq t \) (Lütkepohl, 2006). The diagonal coefficients in \( A_k \) capture the serial dependency of a variable on itself, and the off-diagonal coefficients capture the serial dependency of a variable on other variables.

It can be shown that any VAR(\( p \)) process can be written as a first order VAR process by extending the dimension of \( \mathbf{Y}_t \) (Lütkepohl, 2006). Hence, only the VAR(1) process is considered in this study. Further, the means of the data can be removed prior to the analysis by centering the data. Thus, a bivariate VAR(1) process can be rewritten as:

\[
\begin{bmatrix}
    y_{1t} \\
    y_{2t}
\end{bmatrix} =
\begin{bmatrix}
    a_{11} & a_{12} \\
    a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
    y_{1(t-1)} \\
    y_{2(t-1)}
\end{bmatrix} +
\begin{bmatrix}
    u_{1t} \\
    u_{2t}
\end{bmatrix}
\]

(2)

Here, \( a_{11} \) and \( a_{22} \) are the lag-1 autoregressive coefficients of \( y_1 \) and \( y_2 \) predicting themselves, \( a_{12} \) is the lag-1 coefficient of \( y_2 \) predicting \( y_1 \), and vice versa for \( a_{21} \). \( a_{12} \) and \( a_{21} \) are measures of Granger causality, which refers to the idea that lagged values of a time series can predict another time series after all other relevant pieces of information (e.g., history of the dependent time series).
are taken into account (Granger, 1969; Lütkepohl, 2006). In empirical studies, assessing Granger causality allows us to determine the direction of influences between multiple time series, such as electroencephalography (EEG) signals recorded from different channels or measures obtained from different people simultaneously.

3.1.2 Methods for Handling Missing Data in Time Series

The current study considers three methods for handling missing data. The first two are ad hoc procedures that are relatively easy to implement. The third procedure is a recursive routine developed to complement SAM-MOW (described in Chapter 4), an R program designed to conduct time-frequency analysis based on VAR estimates.

The first ad hoc procedure is listwise deletion, i.e., the whole vector of observations at a particular time point is deleted if one of the elements in the vector is missing. This approach shortens the length of the data and results in a decrease in sample size\(^3\). In time series analysis where the lag between data points is important, listwise deletion can distort autoregressive relations in the data. Despite these drawbacks, this method is extensively used by researchers. In R, listwise deletion is the only option for handling missing data when estimating autoregressive models using the basic stats package (R Development Core Team, 2011). In SAS, users can only choose between listwise deletion, replacing missing values by the first contiguous observations, listwise deletion may also refer to the strategy that, if one of the elements in a vector is missing, the whole vector is replaced by a vector of zero. This alternative strategy does not shorten the length of the data. In R, listwise deletion is implemented in the way described in the text.
and mean imputation in the estimation of VAR models (SAS/IML® 9.22 User’s Guide, 2010). Two studies have examined the consequences of listwise deletion in time series. Rankin and Marsh (1985) found that the negative effects of this approach are small as long as missing data do not exceed 20%. Velicer and Colby (2005) conducted a simulation study based on an ARIMA (Autoregressive Integrated Moving Average) model and found that listwise deletion generally yields accurate estimates of intercepts and error variance, but tends to overestimate linear trends in the data. It also underestimates the degree of dependency when the autocorrelation is negative. Both studies, however, only consider univariate time series. It is unclear how listwise deletion affects the estimation of VAR, which involves multiple time series.

The second ad hoc procedure considered in this study is data imputation using sample means and variances. Specifically, for each dimension of the time series, missing values are substituted by replacing the missing value with the sum of the sample mean and a random component which is normally distributed with a mean zero and a variance equivalent to the sample variance. Compared to listwise deletion, this approach retains the sample size and thus may be more appealing when the proportion of missing data is large. Compared to the more common practice of simply substituting by sample mean, the addition of a random component also retains the variance of the time series. The autoregressive relations, however, may be distorted because the random component is assumed to have no serial dependency.

The third method for handing missing data is a recursive procedure that imputes missing values based on information obtained from a moving window containing previous observations. In this approach, a window width \( n \) is specified. Missing values at time \( t \) are replaced by one-step ahead predictions using a VAR model based on observations in the window.
of \([t-n, t-1]\) plus a random component with the prediction error variance. For a bivariate time series, the one-step ahead predictions of a VAR(1) process can be written as:

\[
\begin{bmatrix}
\hat{y}_{1t} \\
\hat{y}_{2t}
\end{bmatrix} =
\begin{bmatrix}
\hat{a}_{11} & \hat{a}_{12} \\
\hat{a}_{21} & \hat{a}_{22}
\end{bmatrix}
\begin{bmatrix}
y_{1(t-1)} \\
y_{2(t-1)}
\end{bmatrix}
\]

This recursive routine runs from the beginning of the time series towards the end of the time series, replacing each missing value. Whereas this approach is computationally more intensive, it retains the autoregressive relations in the data. Hence, it is hypothesized that this method will outperform the other two methods in recovering the parameters of a VAR(1) process.

More in-depth treatments of missing data in time series appear in the econometric and engineering literatures (Harvey & Pierse, 1984; Jones, 1980; Kohn & Ansley, 1986). A particularly interesting engineering-based approach describes how to rewrite a VAR model into state-space form. Missing values can then be incorporated into a Kalman recursive routine (e.g., Brockwell & Davis, 1996; Phong & Singh, 2008). This approach, however, has not been fully implemented in popular statistical programs, such as SAS and R. Hence it is not included in the current study.

3.2 Methods

3.2.1 Simulation

Data were simulated in R 2.11.1 (R Development Core Team, 2011) based on the bivariate VAR(1) model represented by equation (2):

\[
\begin{bmatrix}
y_{1t} \\
y_{2t}
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
y_{1(t-1)} \\
y_{2(t-1)}
\end{bmatrix}
+ 
\begin{bmatrix}
u_{1t} \\
u_{2t}
\end{bmatrix}
\]

The \(A_1\) matrix determines the stability of the process. A VAR(1) process is stable (i.e., does not explode when time goes to infinity) when
\[
\det (I_m - A_1 z) \neq 0 \quad \text{for } |z| < 1 \quad \text{(4)}
\]

where \(I_m\) is an identity matrix of dimension \(m\) (Lütkepohl, 2006). To satisfy this condition, \(a_{11}\) and \(a_{22}\) have to be in the range of (-1, 1). Four values were chosen within this range, including -0.8, -0.4, 0.4, and 0.8. These values represent both the positive and negative domains and the scenarios of large and small autoregressive relations. In this simulation, \(a_{12}\) was set to zero to represent the scenario of no influence. \(a_{21}\) was set to be either 0.3 or -0.3. These values were chosen based on the consideration that, in behavioral and physiological data, the magnitude of the lag-1 influence from \(y_1\) to \(y_2\) is usually smaller than the magnitude of the lag-1 influence from \(y_2\) to itself, i.e., \(a_{22}\). The simulation of \(A_1\) thus followed a fully crossed design with three factors: (1) Magnitude of \(a_{11}\) and \(a_{22}\); these two parameters were set to be identical and had absolute values of 0.4 or 0.8; (2) Sign of \(a_{11}\) and \(a_{22}\); they could be either positive or negative; and (3) Sign of \(a_{21}\); it could be either 0.3 or -0.3. In total, there were eight \((2 \times 2 \times 2)\) different patterns for the \(A_1\) matrix. The error process was simulated to be a bivariate white noise process with unit variances and zero cross-dimensional covariances \((\Sigma_u = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix})\).

The length of the time series was set to 240, which was used in an empirical study described in Chapter 4 that involves the recursive prediction procedure. Because the recursive prediction procedure required a starting model, the length of the data that were actually simulated was 480. Missing data were created by randomly eliminating values from the second half of the time series (i.e., time point 241 to 480), with the proportion of missingness being 10%, 20%, or 30%. The first half of the time series remained complete and was only used in the recursive prediction procedure. This procedure started from estimating a VAR model based on the first 240 observations, and then progressed towards the end of the time series.
The purpose of this study was to compare the ability of missing data procedures to recover the parameter values, rather than to obtain distributions of the parameter estimates. Therefore, a small number of replications were needed for each condition. Previous studies on univariate time series suggested that 10 replications were enough for each condition (Harrop & Velicer, 1990; Velicer & Colby, 2005). In this study, 100 datasets were generated for each condition to ensure sufficient power to detect differences among the methods of interest.

3.2.2 Analysis

Estimation of the models was conducted in R 2.11.1 using the Yule-Walker estimation method (Lütkepohl, 2006; R Development Core Team, 2011). The order of the models was determined by the Akaike Information Criterion (AIC; Akaike, 1974). Because the estimated models were not necessarily of order 1, direct comparisons of VAR estimates were not appropriate. Therefore, the estimated VAR models were transformed into the frequency domain using:

\[ A(f) = I - \sum_{k=1}^{p} A_k \cdot \exp(2\pi kf) \]  \hspace{1cm} (5)

where \( A_k \) is the autoregressive coefficient matrix at lag-\( k \) in the estimated VAR of order \( p \), \( f \) is the square root of -1, and

\[ f = \frac{n}{N}, \hspace{1cm} n = 0, 1, ..., N-1 \]  \hspace{1cm} (6)

where \( N \) is the number of time points (Schlögl & Supp, 2006). The statistic of interest for comparing the three missing data techniques was the directed transfer function (DTF):
\[ DTF_{ij}(f) = \frac{|H_{ij}(f)|}{\sqrt{\sum_{k=1}^{m}|H_{ik}^{2}(f)|}} \]

\[ = \frac{|H_{ij}(f)|}{\sqrt{H_{i}(f)H_{i}^{H}(f)}} \] (7)

where \( H(f) = A^{-1}(f) \) and \( H_{i}(f) \) is the \( i \)-th row of \( H(f) \) (Kamiński & Blinowska, 1991). It can be shown that \( H(f) \) is the transfer function matrix of the model:

\[ Y(f) = H(f)U(f) \] (8)

where \( Y(f) \) is the observed data and \( U(f) \) is a white noise error process. Hence, DTF is the scaled version of the transfer function. Linking it to the time domain VAR model, DTF can also be viewed as a composite measure of the directional influences between time series combining all lags; hence, it is an index of Granger causality in the frequency domain. It is normalized between 0 and 1 with higher values indicating larger influences. For this study, the average DTF over frequencies was computed for both directions between \( y_1 \) and \( y_2 \). DTF12 denotes the influence of \( y_2 \) on \( y_1 \), and vice versa for DTF21.

The statistical approach used to compare the three methods was repeated measures analysis of variance (ANOVA). There were three between-subject factors: magnitude of \( a_{11} \) and \( a_{22} \) (denoted as AR_LARGE, with 0 representing small and 1 representing large), sign of \( a_{11} \) and \( a_{22} \) (denoted as AR_POSITIVE, with 0 representing negative and 1 representing positive), and sign of \( a_{21} \) (denoted as CR_POSITIVE, with 0 representing negative and 1 representing positive). There were two within-subject factors: missing data technique and proportion of missing data. The complete data were also included in the comparison, making it an unbalanced design. Hence, for each outcome variable, three sets of repeated measures ANOVA were conducted. The first
analysis excluded the complete data scenario and looked at whether any interaction terms containing both within-subject factors were significant. The second analysis examined main effects of missing data technique (four levels, including complete data) and the three between-subject factors, and their interactions. The final analysis examined main effects of proportion of missing data (four levels, including complete data) and the three between-subject factors, and their interactions. Sphericity was tested in all analyses and F- tests were adjusted using the Huynh-Feldt epsilon if sphericity was violated.

Because p-values of the effects are influenced by the number of replications, which is arbitrary in a simulation study, effect sizes were estimated by $\eta^2$. According to Cohen (1988), an effect size of $\eta^2 = .01$ can be considered “small”, an effect size of $\eta^2 = .06$ can be considered “medium”, and an effect size of $\eta^2 = .14$ can be considered “large”. To focus on the most important effects, only results that were both significant and had an effect size of $\eta^2 \geq .06$ are reported and interpreted.

3.3 Results

3.3.1 Estimation of Complete Data

Table 7 shows results from the complete data analysis, broken down by conditions. The first row represents the percentage of first order VAR in all estimated models. We can see that even with complete data, not all estimated VAR models were of order 1. The percentage of VAR(1) models ranged from 82% to 93% across conditions. Looking at the average DTFs over frequencies, we see that DTF12 (influences from $y_1$ to $y_2$) was close to zero, whereas DTF21 (influences from $y_2$ to $y_1$) ranged from .25 to .31.
3.3.2 Directed Transfer Function from \( Y_1 \) to \( Y_2 \) (DTF21)

A \( 3 \times 3 \times 2 \times 2 \times 2 \) (technique \( \times \) proportion of missing \( \times \) AR_LARGE \( \times \) AR_POSITIVE \( \times \) CR_POSITIVE) repeated measures ANOVA was first conducted to look at the effects of the five factors on the estimation of DTF21. The complete case scenario was not included in this model. Results show that none of the interaction effects involving both technique and proportion of missing data had an effect size of \( \eta^2 \geq .06 \). In other words, all effects involving technique are applied similarly across proportions of missing data, and all effects involving proportion of missing data are applied similarly across techniques. This justifies later analysis including only one within-subject factor.

In this model, there was a three-way interaction of technique, AR_LARGE, and AR_POSITIVE, \( F(1.95, 1550.47) = 58.14, p < .001, \eta^2 = .07 \). A plot of the interaction is shown in Figure 2. For comparison purposes, the complete case scenario was also included in the figure. We can see that the estimates appeared to be most stable for the recursive prediction procedure across conditions. Regardless of magnitude and sign of the autoregressive parameters, this approach led to underestimation of the influence from \( y_1 \) to \( y_2 \). Averaging over proportions of missing data, the size of the bias was 15\% to 22\% of the true values. Imputation using sample means and variances also consistently led to underestimation of DTF21, with greater bias – 30\% to 48\% of the true values when averaged over proportions of missing data. In contrast to the two data imputation techniques, the performance of listwise deletion was highly dependent on the simulation conditions. When \( a_{11} \) and \( a_{22} \) were negative, it severely underestimated DTF21, with about 45\% bias. When \( a_{11} = a_{22} = 0.4 \), it also underestimated DTF21, but with only 11\% bias. When \( a_{11} = a_{22} = 0.8 \), listwise deletion overestimated DTF21, with 19\% bias.
To statistically compare the three missing data techniques to the complete data scenario, a \(4 \times 2 \times 2 \times 2\) (technique \(\times\) AR_LARGE \(\times\) AR_POSITIVE \(\times\) CR_POSITIVE) repeated measures ANOVA was conducted. No interaction effect involving CR_POSITIVE had an effect size of \(\eta^2 \geq .06\), indicating that the sign of \(a_{22}\) did not affect the magnitude of DTF21. In this model, the three-way interaction mentioned above had a small effect size \((\eta^2 = .04)\), and hence was not included. Table 8 shows the results of the final model. There were two significant two-way interactions with \(\eta^2 \geq .06\): technique by AR_LARGE, \(\eta^2 = .11\), and technique by AR_POSITIVE, \(\eta^2 = .27\). The interactions are plotted in Figure 3 and Figure 4. Post-hoc contrasts indicated that AR_LARGE had different effects on listwise deletion \((\eta^2 = .07)\) and imputation using sample means and variances \((\eta^2 = .07)\) than on complete data. Looking at Figure 3, we can see that for both complete data and listwise deletion, the estimates of DTF21 were smaller when \(a_{11}\) and \(a_{22}\) were \(\pm 0.4\) as compared to \(\pm 0.8\), but the difference was greater for listwise deletion. For imputation with sample means and variances, the estimates of DTF21 were larger when \(a_{11}\) and \(a_{22}\) were \(\pm 0.4\) as compared to \(\pm 0.8\). For the other two-way interaction, post-hoc contrasts indicated that AR_POSITIVE had different effects on listwise deletion \((\eta^2 = .38)\) than on complete data. Looking at Figure 4, we can see that when \(a_{11}\) and \(a_{22}\) were positive, listwise deletion tended to overestimate DTF21, whereas when \(a_{11}\) and \(a_{22}\) were negative, listwise deletion severely underestimated DTF21. Combining Figure 2, we can see that these two-way interactions were primarily driven by the unusual pattern of overestimation by listwise deletion when \(a_{11} = a_{22} = 0.8\).

Another \(4 \times 2 \times 2 \times 2\) repeated measures ANOVA was conducted to examine the effects of proportion of missing data, AR_LARGE, AR_POSITIVE, and CR_POSITIVE on the
estimation of DTF21. No interaction effect had an effect size of $\eta^2 \geq .06$. The main effect of proportion of missingness had a large effect size, $F(2.70, 2155.28) = 350.26, p < .001, \eta^2 = .31$. Post-hoc polynomial contrasts indicated that both the linear ($\eta^2 = .45$) and quadratic ($\eta^2 = .12$) trends were apparent. From Figure 5, we can see that bias in DTF21 increased as the proportion of missing data increased. The first 10% of missing data appeared to have a larger impact on the estimates than any additional 10% of missing data.

To get a detailed picture of the severity of bias, the estimates of DTF21 are plotted in Figure 6, broken down by technique, proportion of missing, AR_LARGE, and AR_POSITIVE. The solid grey lines represent the estimates from the complete data, and the dotted grey lines represent 20% bias. In most scenarios, estimates yielded by the recursive prediction approach fell between the two lines when missing data did not exceed 20%. The listwise deletion estimates only fell between the two lines when $a_{11}$ and $a_{22}$ were positive. Estimates yielded by the other imputation technique were almost always out of the range.

3.3.3 Directed Transfer Function from $Y_2$ to $Y_1$ (DTF12)

Figure 7 and Figure 8 show the estimated marginal means of DTF12 by technique and by proportion of missing data. Compared to DTF21, the estimates of DTF12 were substantially smaller and close to zero. All three techniques overestimated DTF12, and the bias increased as the proportion of missing data increased. In the repeated measures ANOVA analyses, however, no main effect or interaction effect had an effect size of $\eta^2 \geq .06$. Hence, the estimate of the influence from $y_2$ to $y_1$ was not severely affected by missing data technique, proportion of missing data, or the pattern of the $A_1$ matrix.
3.4 Discussion

This study investigated the impact of three missing data techniques on the estimation of bivariate VAR(1) models. The three techniques considered were listwise deletion, data imputation with sample means and variances, and data imputation using a recursive prediction procedure described in section 3.1.2. Because time domain estimates (e.g., autoregressive coefficients) could not be directly compared when the orders of the VAR models were not equal, comparison was done in the frequency domain. Data were simulated such that a unidirectional influence existed between time series $y_1$ and $y_2$. Specifically, $y_2$ was predicted by $y_1$ at lag-1 with a coefficient of -0.3 or 0.3, and $y_1$ was not predicted by previous values of $y_2$. The corresponding frequency domain statistic for assessing directional influences was the directed transfer function (DTF). DTF21 reflected influences from $y_1$ to $y_2$, combining all lags, and vice versa for DTF12. Hence, an accurately estimated VAR model should have a relatively large DTF21 and a DTF12 close to zero.

Results indicated that all three techniques yielded acceptable estimates of DTF12, but led to bias in the estimation of DTF21. The recursive prediction procedure was the most stable over simulation conditions and consistently yielded relatively small downward bias. Data imputation using sample means and variances consistently led to large downward bias. Listwise deletion showed substantial variation across simulation conditions. Its performance was acceptable when $a_{11}$ and $a_{22}$ were positive, but very poor when $a_{11}$ and $a_{22}$ were negative. In the positive case, the direction of bias also depended on the magnitude of the autoregressive coefficients.

The behavior of data imputation using sample means and variances can be expected because the imputed values do not have serial dependency with previous or later values, which
naturally leads to underestimation of the autoregressive coefficients. The large bias across simulation conditions suggests that this technique is not appropriate in the time series context. The behavior of listwise deletion is consistent with findings of Velicer and Colby (2005) and is not surprising. In this approach, missing data are replaced by the first full vector (i.e., no missing) of values at a later time point. If the time series is spiky (i.e., highly variable), such that observations at time $t$ substantially differ from observations at time $t-1$ and time $t+1$, the estimates of VAR will be severely affected. This is the case when $a_{11}$ and $a_{22}$ are negative. In contrast, if the time series is smooth, such that observations at time $t$ are similar to observations at time $t-1$ and time $t+1$, listwise deletion would have a less detrimental impact. This is the case when $a_{11}$ and $a_{22}$ are positive. The direction of the bias for DTF21 can be explained by mathematical deduction. With $a_{12} = 0$, equation (2) can be written as:

$$
\begin{align*}
  y_{1t} &= a_{11} y_{h(t-1)} + u_{1t} \\
  y_{2t} &= a_{21} y_{1(t-1)} + a_{22} y_{2(t-1)} + u_{2t}
\end{align*}
$$

(9)

Hence,

$$
\begin{align*}
  y_{2t} &= a_{21} (a_{11} y_{l(t-2)} + u_{l(t-1)}) + a_{22} (a_{21} y_{l(t-2)} + a_{22} y_{2(t-2)} + u_{2(t-1)}) + u_{2t} \\
  &= (a_{21} a_{11} + a_{22} a_{21}) y_{l(t-2)} + a_{22} y_{2(t-2)} + a_{21} u_{l(t-1)} + a_{22} u_{2(t-1)} + u_{2t}
\end{align*}
$$

(10)

and so on for larger lags. Therefore, when $a_{11} = a_{22} = 0.8$, the regression coefficient for using $y_1$ at a larger lag to predict $y_2$ would be greater than $a_{21}$, leading to an overestimation of DTF21. When $a_{11} = a_{22} = 0.4$, DTF21 would be underestimated. Since the sign of the autoregressive coefficients is often not known prior to model fitting, and the underlying VAR process for empirical data can have both positive and negative autoregressive coefficients, listwise deletion is also not appropriate for handling missing data in time series analysis.
The recursive prediction procedure outperformed the other two approaches. When proportion of missing data was less than or equal to 20%, the estimates of DTF21 were mostly within ±20% of the true values. When proportion of missing data exceeded 20%, however, this approach led to large bias and should not be used. In this approach, missing values are generated by an estimated VAR model based on the previous time window of length 240. Results from the complete data analysis show that only 82% to 93% of the estimated VAR models had the correct order. In other words, the estimated models are not always accurate even when complete data are available. The one-step ahead prediction based on the estimated VAR models, in this case, will deviate from data generated using the true model, leading to bias in DTF21. Moreover, because missing data at later time points are predicted with an estimated model which may itself depend on imputed data, the prediction error accumulates as the recursive procedure runs from the beginning to the end of the time series. Thus, it appears that the accuracy of the estimated VAR model in the first window is crucial for the performance of the recursive prediction procedure. Since the accuracy can be improved by increasing the length of the data, this finding suggests that a larger window width for the first window may help reduce the bias in DTF21.

It should be noted that the validity of the recursive prediction procedure relies on the important assumption that the underlying process generating the data is homogeneous over time. In this study, the recursive prediction started from a model estimated based on data that were simulated according to the true model. Hence, the performance described in this study represents the best possible performance of the recursive prediction procedure given the length of the time series. In most empirical data, the homogeneity assumption may not hold. Violation of this assumption is likely to impair the performance of the procedure, and the magnitude of the impact needs to be investigated in future research.
The current study provides important implications for handling missing data in multivariate time series analysis, which has gained increasing popularity in behavioral and physiological research. As the first study in the behavioral sciences to compare missing data techniques in the vector autoregressive modeling framework, this study points to some future directions that are worth exploring. First, due to difficulty in implementation, several alternative techniques (e.g., pairwise deletion, Kalman filtering, multiple imputation) for handling missing data have not been included in the current comparison. A follow-up study focusing on these techniques will help us better understand different ways to handle missing data and make sounder decisions in the analysis. Second, the current study focuses on bivariate processes and unidirectional influences. Future research can expand the dimension of the models and consider mutual influences between time series (i.e., reciprocal Granger causality). In the brain imaging area, researchers have also argued that including both contemporaneous and lagged relations may better capture the dynamics of a process (Gates, Molenaar, Hillary, Ram, & Rovine, 2010). Yet, it is not known whether and how this model may be influenced by missing data and what technique(s) work(s) best in this context. Finally, this study only considers normally distributed data and data that are missing completely at random (MCAR). Research that examines other distributions and missing data mechanisms would provide deeper insight to this topic.
CHAPTER 4

Dynamical Modeling of Child-Therapist Interactions in Electrodermal Activity Using Time-Frequency Analysis

4.1 Introduction

Physiological arousal is an important aspect of psychotherapy for children with Sensory Processing Disorder (SPD). Coregulation in physiological arousal between individual children and their therapists reflects social-emotional processes during psychotherapy and is likely to affect therapy outcomes. In this study, a novel approach is presented for the dynamic modeling of physiological interactions between children with SPD and their therapists using time series measures of electrodermal activity (EDA). This approach, called *time-frequency analysis* (TFA), allows us to investigate the underlying cycles in each individual’s EDA signal, as well as the covariation, lead-lag relations, and directional influences at the dyadic level (i.e., child and therapist jointly). Moreover, a program called SAM-MOW (Spectral Analysis for Multivariate data using a MOving Window) is developed to conduct TFA on multivariate time series data automatically. The program provides a new tool for researchers interested in assessing the interrelations between multiple time series data.

4.1.1 Physiological Interactions in Psychotherapy

Researchers and clinical practitioners have long been interested in understanding interpersonal processes between patient and therapist during psychotherapy, which seem to affect the effectiveness of the therapy (Henry, Schacht, & Strupp, 1990; Teyber & McClure, 2010). Physiological dynamics is an important yet understudied aspect of interpersonal interactions. The earliest attempt to capture patient-therapist physiological interactions is by Di Mascio and
colleagues (1955, 1957), who recorded electrocardiogram measures (EKG) of a therapist and patient dyad over 38 psychiatric interviews and reported observed “concordance” and “discordance” in their pulse rates. Based on the therapist’s post-session notes, they found that the therapist was more likely to be distracted from the session when the physiological concordance was low. Recently, Marci and colleagues (2006, 2007) recorded skin conductance levels from therapist-patient dyads and found that their physiological coherence was associated with patient’s post-session rating of therapist empathy.

Despite research suggesting that physiological interactions between patient and therapist may reflect therapeutic rapport and have implications for success of the therapy, research in this area is scarce. This may in part be due to a lack of appropriate methodological tools and statistical programs to study interpersonal physiology. In this study, a unique approach is presented that combines multiple techniques in time series analysis and allows us to quantify these underlying dynamics.

4.1.2 Time Series Analysis in Time and Frequency Domains

Time series analysis can be conducted in the time domain and the frequency domain. Time domain analysis examines serial dependency in the data. For stationary multivariate data, a basic time-domain model is the vector autoregressive model (VAR):

\[
\hat{Y}_t = \tilde{V} + A_1 \times \hat{Y}_{t-1} + A_2 \times \hat{Y}_{t-2} + \ldots + A_p \times \hat{Y}_{t-p} + \tilde{U}_t
\]  

(1)

As noted in the previous chapter, the autoregressive coefficients in \(A_k\) represent the degree of serial dependency of one dimension of a time series on itself and on other dimensions at lag \(k\). The covariance matrix of the errors \(\Sigma_u\) contains the variances and covariances of the data that are not explained by the autoregressive relations.
Frequency domain analysis, or spectral analysis, is used to study cyclical behaviors in time series data. Cyclical patterns are inherent in many physiological signals, such as measures of brain activities (e.g. EEG) and cardiovascular responses (e.g. ECG). In frequency analysis, time series data are decomposed into asymptotically statistically independent sine and cosine waves of different frequencies. The distribution of variance accounted for by the different frequency components is described by the *spectral density function*. A visual tool to represent the estimated spectral density function is the *spectrogram*, which is a graph of estimated power (i.e. contributed variance) against frequency, i.e., the *spectrum*. For example, a time series composed with a sine wave of 0.5 Hz will have one major peak in its spectrogram at 0.5 Hz.

Frequency domain analysis is convenient for deriving estimates that directly tap the concept of coregulation. For bivariate data, the *coherence spectrum* describes the degree of covariance for two dimensions of the time series at each frequency. The *phase spectrum* describes the lead-lag relation at each frequency. Gottman (1981) illustrated the use of bivariate frequency techniques to examine synchrony in mother-infant behaviors during face-to-face interactions using data from a study by Tronick, Als, and Brazelton (1977). Lester, Hoffman, and Brazelton (1985) conducted a similar study and found that coherence in mother-infant behaviors increased from three to five months of age. They also found higher mother-infant coherence in dyads involving full term babies than in dyads involving preterm babies. After these initial attempts, however, applications of multivariate frequency analysis have been rare in psychological research. This may be partly due to difficulties in obtaining time series data in the behavioral sciences and in implementing the method.
4.1.3 Transforming Time Domain Model into the Frequency Domain

Estimates in the frequency domain can be computed using raw data. Raw estimates, however, have poor bias and variance properties and require advanced handling techniques (Chatfield, 2004; Gottman, 1981). The time series literature has provided a large variety of options for reducing bias and obtaining better variance properties of these estimates (e.g., smoothing and tapering, see Percival & Walden, 1993). Yet, choosing an optimal method for empirical data requires substantial time series modeling experience, and the decision making process is often a daunting task. An alternative way to obtain frequency domain estimates involves the combination of time and frequency domain analysis. Specifically, estimates from a time domain model can be easily transformed into the frequency domain by Fourier transform methods. This constitutes the basis of the maximum entropy spectrum estimation method proposed by Burg (1972, 1975).

Schlögl and Supp (2006) introduced methods for transforming a VAR model into the frequency domain. For a particular frequency $f$, we can compute

\[
A(f) = I - \sum_{k=1}^{p} A_k \cdot \exp(2\pi k fi)
\]  

where $A_k$ is the autoregressive coefficient matrix at lag $k$ of a VAR model, $p$ is the order of the VAR, $i$ is the square root of -1, and

\[
f = \frac{n}{N}, \quad n = 0, 1, \ldots, N-1
\]

where $N$ is the number of observations and assumed to be even. Let

\[
H(f) = A^{-1}(f)
\]

and the model in equation (1) can then be written as
\[ Y(f) = H(f)U(f) \]  

where \( Y(f) \) and \( U(f) \) represent the observed data and innovation (errors) in the frequency domain. Hence, \( H(f) \) is the transfer function matrix of the model. We can compute

\[ S_Y(f) = H(f) \Sigma_u H^H(f) \]

where the superscript indicates the Hermitian operator (transposed complex conjugate of matrix \( H \)). The diagonal elements of \( S_Y(f) \) give the estimates of the power for each time series at frequency \( f \), which combine to form the *autospectrum*.

The complex-valued coherency matrix is defined as

\[ C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f) \cdot S_{jj}(f)}} \]

where \( i \) and \( j \) are integers from 1 to \( m \) for a \( m \)-dimensional VAR model (Nolte et al., 2004). The real-valued coherence between series \( i \) and series \( j \) at frequency \( f \) is the absolute value of \( C_{ij}(f) \):

\[ |C_{ij}(f)| = \sqrt{\text{Re}(C_{ij}(f))^2 + \text{Im}(C_{ij}(f))^2} \]

where \( \text{Re} \) and \( \text{Im} \) represent the real and imaginary part of a complex number.

The phase difference between series \( i \) and \( j \) is:

\[ \phi_{ij}(f) = \arctan(\text{Im}(S_{ij}(f))/\text{Re}(S_{ij}(f))) \]

A positive phase difference indicates that series \( i \) leads series \( j \) in time.

Schlögl and Supp (2006) reviewed several options to derive measures that uncover directional influences between multiple time series. One of them is the *directed transfer function* (DTF), defined by Kamiński and Blinowska (1991) as:
\[
DTF_{ij}(f) = \frac{|H_{ij}(f)|}{\sqrt{\sum_{k=1}^{m}|H_{ik}^2(f)|}}
\]

\[
= \frac{|H_{ij}(f)|}{\sqrt{H_{ij}(f)H_{ji}^H(f)}}
\]  

(10)

where \(H_i(f)\) is the \(i\)-th row of \(H(f)\). We can see that DTF is the scaled version of the transfer function in the model represented by equation (5). Unlike coherence, DTF is not symmetric, and it is related to the concept of Granger causality. In the time domain, Granger causality is established if the prediction error of a time series is reduced by including past measurements from another time series (Granger, 1969). In the frequency domain, it can be defined as the off-diagonal elements of the transfer function matrix (Caines & Chan, 1975). Kamiński, Ding, Truccolo, and Bressler (2001) showed that DTF is an index of Granger causality for multivariate time series data.

In the following, an application of TFA is demonstrated with empirical data. In this application, the physiological dynamics in EDA between a child and a therapist during psychotherapy is investigated by obtaining estimates of their autospectra, coherence, phase difference, and DTFs.

4.2 Empirical Application

4.2.1 The Study

EDA is a measure of eccrine sweat gland activity and is controlled primarily by the sympathetic nervous system (SNS; Dawson, Schell, & Filion, 2007). When eccrine sweat glands open, the ability of the skin to conduct electricity increases, leading to higher recordings of EDA.
In general, EDA is considered an indicator of psychophysiological arousal, and has been found to be associated with a large variety of psychological processes, including emotional regulation (Boucsein, 1992). Thus, EDA provides a possible way to monitor the arousal of children with SPD, a psychological disorder characterized by over-response or under-response to a sensation (Bundy, Lane, & Murray, 2002). Moreover, dyadic EDA data obtained from both a child and a therapist allow us to examine the dynamic interactions in their arousal during psychotherapy.

In this study, EDA data were obtained from 22 pairs of children with SPD and their therapists at the STAR (Sensory Therapists and Research) Center in Greenwood Village, Colorado. Fifteen children provided demographic information. Their age varied from 3 to 10 years, with an average of 6.5 years. Among them, seven were female. Fourteen were Caucasian, and one was Hispanic. Each child was paired with a specific therapist. All therapists had a master’s degree and had received advanced training and mentorship before participating in the study. In addition, all child-therapist pairs had gone through at least three therapy sessions before being enrolled in the study (Hedman, 2010).

During the study, each child-therapist dyad was observed in at least two therapy sessions, which typically lasted for one hour and involved three to eight guided activities. Both the child’s and the therapist’s EDA were recorded by iCalm, a wireless sensor developed by Fletcher and colleagues (Fletcher et al., 2010). Each individual wore iCalm on both their right and left ankles, and their EDA levels were recorded every 0.5 second simultaneously. In total, 77 therapy sessions were recorded.
4.2.2 Time-Frequency Analysis (TFA)

A person-specific approach (Molenaar & Campbell, 2009) was adopted to investigate the dynamic interactions in EDA for each individual dyad within each therapy session separately. Because the physiological patterns were likely to change during the therapy session, a TFA approach was adopted. TFA is a set of methods used in signal processing to characterize signals whose statistical properties vary in time. In this study, TFA was conducted using a moving-window technique. Specifically, the entire time series was divided into multiple time windows, and frequency analysis described in the previous section was applied within each window. In other words, within each window, estimates of the autospectra, coherence, phase difference, and DTFs between the child’s and the therapist’s EDA were obtained based on an estimated VAR model. As the window moved forward in time, this technique allowed us to explore how physiological interactions changed throughout the therapy session.

The estimation of VAR was conducted using the ordinary least square (OLS) method via the \textit{vars} package in R 2.11.1 (R Development Core Team, 2011), developed by Pfaff (2008). Because the procedure was computationally intensive, a program named SAM-MOW (Spectral Analysis for Multivariate data using a MOving Window) was developed to automatically carry out the analysis. Additional routines were created to accommodate several problems relating the child-therapist data. First, the data contained missing values. Currently, the \textit{vars} package does not allow missing data. Thus, a program named MMISS was developed to replace missing values by recursive one-step ahead predictions based on a VAR model estimated using previous observations. This procedure has been shown to outperform listwise deletion and data imputation using sample means and variances in a simulation study (cf. Chapter 3), although it tends to lead to small downward bias in non-zero DTFs. Second, the EDA data contained deterministic trends,
even when divided into windows. Frequency analysis is based on the assumption of stationarity, meaning that the time series needs to have a constant mean and autocovariance function over time. To accommodate this problem, a sub-routine was added to SAM-MOW to remove polynomial trends within each window. Furthermore, tests of stationarity including the Dickey-Fuller test (Dickey & Fuller, 1979) and the KPSS test (Kwiatkowski, Phillips, Schmidt, & Shin, 1992) were incorporated into the program to help users determine the order of polynomial trends that need to be removed.

In the following, the analysis is described using data from an exemplary child-therapist dyad in one therapy session. This therapy session was selected for demonstration because the EDA recordings were relatively problem-free. Specifically, there were no apparent technical problems (e.g., loss of signal) or artifacts (e.g., sudden jumps or drops in the values; see Hedman 2010 for details). The analytic procedure for this exemplary dataset consists of four steps:

1. Recordings of EDA from one ankle of each person (i.e., child; therapist) were selected based on the number of missing values and problematic values⁴. Data from the two dyad members formed a bivariate time series.

2. The time series was divided into non-overlapping windows of equal length⁵. Missing values were imputed using MMISS.

---

⁴ Technical problems and/or movements of participant(s) during recording may result in problematic EDA values, most commonly in the forms of sudden jumps or drops (cf. Hedman, 2010). In this exemplary dataset, no such problem was detected. Hence for each dyad member, the time series with fewer missing values was selected.

---

50
3. Within each window, polynomial trends were removed in increasing order (i.e., linear, quadratic, etc.) until the residuals were stationary. Stationarity was determined by the Dickey-Fuller test and KPSS test.

4. TFA was conducted using SAM-MOW. Specifically, a VAR model was fit within each window. Estimates of the model were transformed into the frequency domain to obtain the autospectra, coherence, phase difference, and DTFs between the child’s and therapist’s EDA.

4.2.3 Results from One Therapy Session

Figure 9 shows the plot of the EDA values from the child and therapist in the exemplary dataset. The total length of the time series was 5338 time points, which was approximately 44.5 minutes. Both the child and therapist showed an overall increase in EDA throughout the therapy session. A window width of 240 data points was used based on joint considerations of local stationarity and analytic power\(^6\). Thus, the time series was divided into 22 non-overlapping windows, each representing two minutes of child-therapist interactions. Among

\(^5\) SAM-MOW allows users to specify whether the time windows overlap with each other and by how many data points. In this example, the windows were specified to be non-overlapping. In addition, the current version of SAM-MOW only allows a fixed window width as part of its automatic recursive routine. However, the core of the program can be extracted to conduct TFA on any part of the time series (of any length) if so desired.

\(^6\) The larger the window width, the more power there is to estimate the VAR model. Nevertheless, stationarity is also more likely to be violated with a larger window width.
the 22 windows, 5 had complete data, and 14 had less than 5% missing values. Among the three remaining windows, window 2 and window 9 had 5.2% missing, and window 3 had 7.1% missing. Missing values were imputed using MMISS. In this analysis, the linear and quadratic trends were removed within each window. Residual time series in all but two windows (window 5 and 6) were stationary according to the Dickey-Fuller test and KPSS test. TFA was carried out using SAM-MOW. Within each window, a VAR model was fit and the adjusted portmanteau test was used to test whether the residuals conform to a white noise process with no serial dependency (Lütkepohl, 2006). All residuals were white noise except for window 17, indicating that in all but one window the autoregressive coefficients in the VAR successfully captured all serial dependency in the data.

Figure 10 and 11 show the 3-dimensional graphs as well as the vertical views of the autospectra for the child and therapist. In the 3-dimensional graphs, the x-axis represents time (window), the y-axis represents frequency, and the z-axis represents power. Markers on the y-axis represent the number of cycles within the 2-minute window, hence, the lowest frequency ($f = 1$) represents 1 cycle per 2 minutes, or approximately 0.0083 Hz, whereas the highest frequency ($f = 120$) represents 120 cycle per 2 minutes, or 1 Hz. Because the estimated power depends on the variance of the data (Brillinger, 1981), which was not constant across windows, the autospectrum within each window was standardized by the first frequency component ($f = 1$) to facilitate comparison across time. It should be noted that the scale of the z-axis varies across different figures. The power estimates corresponding to $f = 1$ are plotted in Figure 12. From Figure 12, we can see that the power of the first frequency component peaked at window 6 for both the child and the therapist. This is likely to be an artifact resulting from the non-stationarity of the data in this window. In other words, the large peak at the lowest frequency band is likely
to reflect a deterministic trend in the data. We also see that, in general, the child’s EDA had larger power than the therapist’s EDA, indicating that the variance was larger in the child’s data. There was an overall increase in power towards the end of the session. These patterns are not surprising given that the child’s EDA level was higher than the therapist’s EDA level, and that both dyad members’ EDA increased during the therapy session. Higher levels of the data are typically associated with larger variances.

Looking at Figure 10 and 11, we see that for both the child and the therapist, the power concentrated in the frequency range of 40 to 80 in about half of the windows. This frequency range corresponds to 0.33 Hz to 0.67 Hz, or one cycle per 1.5 to 3 seconds. In most of the remaining windows, the autospectra appeared to be flat.

Figure 13a shows the 3-dimentional graph of the coherence between the child’s and therapist’s EDA. Overall, the coherence was low, except for a small peak at the lowest frequency component in window 6. As mentioned before, the data were not stationary in this window, thus the peak is likely to be an artifact instead of an indicator of strong covariation between the child and therapist. An average coherence over frequencies was computed for each window and plotted in Figure 13b. We see that none of the windows had an average coherence above 0.2. Because the phase difference between time series is only meaningful when coherence is high (Warner, 1998), the phase spectrum is not presented here.

DTFs across time and frequency are plotted in Figure 14 and 15. It should be noted that the z-axis in the two figures has dramatically different scales. To facilitate comparison, average DTFs over frequencies were computed and plotted in Figure 16. It is clear that DTFs from the therapist to the child were in general much higher than DTFs in the opposite direction, indicating a unidirectional relationship. Specifically, in most windows the child’s EDA was influenced by
the therapist’s EDA, but not vice versa. Comparing across time, this unidirectional influence was stronger in the first half of the therapy session than in the second half. Time-synced video records of the therapy session showed that the child was playing with tactile toys during the entire window 2 and the first 70% of window 3. From Figure 16, it appears that the influence of the therapist on the child’s EDA peaked as this activity started, but declined when it ended. In addition, during approximately 30% of the time in window 7, the child was playing an interactive game called Makoto. This activity corresponds to a peak in the influence of the therapist on the child in the same window, as well as in the next window (window 8).

4.2.4 Interpretation and Confounding Factors

This analysis provides valuable information about the physiological dynamics between child and therapist during the therapy session under study. The autospectra suggested that in some time windows EDA values did not show cyclical patterns and the autospectra were relatively flat, whereas in other windows EDA appeared to be characterized by cycles with periods of 1.5 to 3 seconds. In the psychophysiological literature, researchers have developed a large variety of measures for characterizing EDA activity. For example, a skin conductance response (SCR) can be characterized by its latency (temporal interval between stimulus onset and SCR initiation), rise time (temporal interval between SCR initiation and SCR peak), amplitude (increase in level), and half recovery time (temporal interval between SCR peak and point of 50% recovery of SCR amplitude). Within a period of time, EDA can also be quantified by the skin conductance level (SCL), change in SCL, and/or the number of SCRs. These components are associated, such that frequent SCRs usually correspond to large SCR amplitude, short latency, and short rise time, etc. (Dawson, Schell, & Filion, 2007). Thus, the concentration
of power in a higher frequency band may reflect more pronounced EDA activities. However, the correlations between different measures on EDA are usually not very high, and the unique information each measure contains is not fully understood by scientists. Therefore, the substantive significance of the power distribution, which is a composite measure that combines different components of EDA, needs to be investigated and validated in future research.

Although the autospectra showed that EDA of the child and the therapist appear to have similar power distributions in some windows, the coherence statistics suggested their covariation was not strong. It should be noted that the conceptualization of coherence in this study is different from the “concordance” assessed in Marci et al.’s (2006, 2007) studies on patient-therapist EDA interactions. In their studies, concordance was calculated as the correlation between dyad members’ SCL slopes, whereas in the current study coherence assesses whether the child’s and the therapist’s EDA are characterized by the same frequency components, after removing trends in the SCL. In the psychophysiological literature, researchers have developed different methods to quantify “synchrony” or “concordance” (e.g., McAssey, Helm, Hsieh, Sbarra, & Ferrer, in press). The use of different measures would likely lead to different findings. Readers should pay careful attention to the construction of these measures when interpreting results.

DTFs suggested there was a unidirectional influence between the child and the therapist. In most of the time windows, the child was influenced by the therapist, but the therapist was not influenced by the child. The direction of the relation is in line with the goal of the therapy, i.e., the therapist was modeling responses for the child and helping him or her achieve a desirable state. Moreover, the magnitude of the influence appeared to be contingent on guided activities during the therapy session. In particular, the initiations of the two activities corresponded to
peaks in the therapist-to-child influence. The interpretation of DTFs, however, needs to take into account at least two confounding factors. The first confound factor is accuracy of measurement. In EDA recordings, extreme low values (< 0.5 μ mhos) may be problematic because the sensors may not be able to pick up nuances in skin conductance and hence the signal-to-noise ratio may be impaired (Hedman, 2010). From Figure 9, we see that the first half of the therapist’s EDA consists of extreme low values. This may have led to elevated DTFs from the therapist to the child if the two EDA recordings were characterized by similar noise processes. Another confounding factor is variance of the data. Individuals sharing the same experience may differ from each other in their skin conductance levels, and for a specific individual, EDA level can also differ across contexts (i.e., EDA level is a factor of hydration and thermo regulation and does not necessarily reflect reactivity to psychological events). In this study, the child’s EDA ranged from 0.95 to 13.04 μ mhos, whereas the therapist’s EDA ranged from 0.04 to 3.68 μ mhos. Accordingly, the child’s EDA had a larger variance (SD = 3.71) than the therapist’s EDA (SD = 0.98). Variance of the data is influential in the estimation of the VAR model because the OLS estimator tries to minimize total variance of the errors (Lütkepohl, 2006). Specifically, there may be an upward bias in the autoregressive coefficients from the therapist’s EDA to the child’s EDA simply because this would lead to a more pronounced reduction in total error variance.

A way to enhance proper interpretation of physiological dynamics between the child and therapist is to compute standard errors of the estimates, especially for DTFs. Kamiński et al. (2001) demonstrated the use of surrogate data methods to generate empirical distributions of estimates, thus allowing assessment of the significance of DTFs. Schlögl and Supp (2006) introduced another approach, which uses a “jackknife” resampling method to estimate standard errors. Besides these computationally intensive supplements, an alternative to gain more insight
to the substantive meaning of the estimates is to conduct simulation studies. Specifically, a simulation study looking at how variance of the data influences estimation of the VAR may enable us to disentangle “true” physiological influences from statistical artifacts.

4.3 Discussion

This study applies time-frequency analysis to EDA data obtained simultaneously from a child with SPD and a therapist during psychotherapy. The approach presented has several strengths. First, by dividing the time series data into windows and conducting separate analyses within each window, this approach highlights and accommodates substantial heterogeneity (non-stationarity) in interpersonal processes during therapy. For instance, it is obvious that influences from the therapist to the child vary from window to window. If the windows are collapsed in the analysis, as is often the case in psychophysiological studies, there would be a loss of information, and researchers may reach incorrect conclusions.

In addition, this study uses a person-specific approach, which enables us to address and examine heterogeneity across therapy sessions and dyads. Most psychological research is conducted using sample statistics. Whereas findings obtained from sample statistics are informative about trends at the group or population level (i.e., nomothetic), they only apply to the individual level when strict homogeneity assumptions hold (i.e., ergodic theorems, Molenaar, 2004). Clinical research findings are often used to facilitate diagnosis and treatment, which is an individual level endeavor. In this context, conducting person-specific analysis is necessary. In this study, the dynamic interactions in EDA for a specific child-therapist dyad was modeled, which partials out possible confounding effects of individual characteristics such as age, gender, symptoms, and severity of the disorder. Based on these results, we would expect greater
confidence in drawing conclusions and making recommendations for this particular child-therapist dyad than if the analysis was done by pooling across a group of individuals.

The current study also contributes to the literature by providing an alternative way to look at interpersonal physiological dynamics. In recent years, there has been increasing interest in this topic and several dynamic modeling techniques have been developed. For example, Ferrer and colleagues (e.g., Ferrer & Helm, under review; Helm, Sbarra, & Ferrer, 2011) have used differential equation models to examine dynamic interactions in heart rate and respiration between romantic partners, and obtained promising results. This approach, however, requires researchers to make explicit hypotheses regarding ways dyad members relate to each other. McAssey et al. (in press) presented two non-parametric techniques to identify synchrony in non-stationary respiration and thoracic impedance signals, which are less dependent on solid theoretical foundations. The current study provides another exploratory approach. Specifically, the method presented utilizes a data-driven VAR fitting procedure which can be standardized. Hence, the current approach is especially suited for studying phenomena which are not well understood by researchers.

As a novel method, the approach presented here also has several weaknesses. Some of them have already been mentioned in section 4.2.4. Specifically, the exact interpretation of the autospectra of EDA data needs to be further studied, and there are confounding factors that may affect interpretations of the results. Moreover, although DTFs allow us to examine the direction of the influences between dyad members, they do not contain information about whether the influences are positive or negative. Hence, it remains unknown whether an increase in the therapist’s EDA is associated with an increase or decrease in the child’s EDA. Addressing this
question is important for further understanding the physiological processes during therapy, and it can be achieved by conducting impulse response analysis in the time domain (Lütkepohl, 2006).

A shortcoming of conducting frequency analysis by Fourier transform is its reliance on stationary data. Although the moving window technique provides a way to overcome this problem, selection of an appropriate window width is not always obvious (Boashash, 2003). The techniques introduced by McAssey et al. (in press) represent alternative ways to model non-stationary signals. Another possibility is wavelet analysis (e.g., Walnut, 2001). These methods may complement the current approach.

Finally, the current study presents results from only one child-therapist dyad in one therapy session. Analyses of data from additional therapy sessions and/or dyads will be a next step. Conducting person-specific analyses, however, brings about the challenge of summarizing results from individual models and making inferences to a larger population. Recent development in brain imaging provides a way to accommodate this interpersonal heterogeneity problem by first identifying a common model for the group and then allowing individuals to vary based upon the group model (Gates & Molenaar, under review). This work suggests a promising direction for bridging the gap between statistics at individual and aggregate levels.
CHAPTER 5

Discussion

The primary goal of this dissertation is to apply state-of-the-art methodology and develop new statistical methods to study dyadic interactions from a dynamic systems perspective. The studies presented were guided by the view that social interactions constitute the central context of human development, and that studying the ways social partners influence each other over time is crucial for the understanding of individual and family development. In the following, the results of each study are summarized and differences in analytic approaches are discussed.

The first study focused on the marital dyad system and examined whether and how husbands and wives coregulate in their patterns of diurnal cortisol, a biomarker of stress and an index of overall neuroendocrine health. Twenty-eight heterosexual couples provided saliva samples containing cortisol four times a day for four consecutive days. The daily diary design allowed investigation of couple coregulation in cortisol on a day-to-day basis. Multilevel models showed that husbands and wives move up or down together in their diurnal cortisol slope (DCS) at the within-couple level (i.e., across days). Couples characterized by high levels of spousal strain and disagreement also showed positive covariation in their cortisol awakening response (CAR), whereas couples with low levels of spousal strain and disagreement showed negative covariation in CAR. Spouse’s diurnal cortisol pattern on day t-1 was not influential to individual’s diurnal cortisol pattern on day t when individual’s own cortisol pattern on day t-1 was controlled for. Individuals in families characterized by higher spousal support, however, showed higher stability in CAR across days. These results highlight the family system theory, and provide evidence that husbands and wives influence each other in biological stress from day
to day and that these daily dynamics are related to characteristics of the marital relationships. The study demonstrates the advantage of using daily diary data and multilevel modeling in family research. Specifically, the study design and statistical approach have made it possible to examine family phenomenon at both the between- and within-family levels. Within-family analysis allows us to investigate the processes whereby family members interact over time, whereas between-family analysis addresses the heterogeneity across families in these interactions. Both are important questions and complement each other to provide a comprehensive picture of family interactions.

The second study addresses a methodological issue related to time series analysis, which is commonly used in psychology to model dynamic systems. Specifically, this study compared three procedures for handling missing data in multivariate time series. A simulation study was conducted in the vector autoregressive (VAR) modeling framework, a popular technique to model dynamic interactions between several constructs or people over time. It is found that a recursive prediction routine that replaces missing values based on VAR modeling outperformed listwise deletion and data imputation using sample means and variances in recovering model parameters. The results have implications for selecting a proper way to handle missing data in VAR estimation. In addition, a program (MMISS) was developed to carry out the recursive prediction routine. It can be used to impute missing data in multivariate time series when more advanced techniques such as Kalman filtering cannot be conveniently implemented.

The third study presented an example of using time-frequency analysis (TFA) to model psychophysiological processes between a child with Sensory Processing Disorder (SPD) and a therapist during therapy. Electrodermal activity (EDA), an index of psychophysiological arousal, was measured every 0.5 second from the child and the therapist simultaneously. TFA allows us
to examine characteristics of the child’s and therapist’s EDA and their dynamic interactions. Most of the time, the therapist’s EDA predicted the child’s EDA, but there was no influence in the opposite direction. This suggests the therapist was modeling the child’s arousal to the desired state, a goal that the therapist was trying to achieve. In addition, peaks in therapist-to-child influences occurred with the initiation of guided activities during therapy, such as when the child started to play with toys. To the author’s knowledge, this is the first study to conduct frequency analysis on EDA data; hence the study also raises some unanswered questions. For example, the interpretation of power distribution of EDA (i.e., the autospectrum) is not straightforward. The impact of several confounding factors, such as variance of the data, on TFA results also needs to be explored in future research. Nevertheless, this study demonstrates great potentials of TFA in monitoring physiological processes and identifying effective strategies during psychotherapy. The program developed (SAM-MOW) to carry out TFA on multivariate time series data automatically can also be used to study dynamic interactions between multiple constructs or people in other settings.

The studies that comprise this dissertation reveal a rich and complicated world of psychophysiological exchanges that underlie social interactions. Beyond the scope of the studies presented here, there are certainly other aspects of psychophysiology that needs to be explored. These processes can occur on multiple time scales, from milliseconds (e.g., EEG) to days (e.g., diurnal cortisol), and even to years. Understanding these underlying dynamics will enhance our knowledge of interpersonal relationships and how they affect individual well-being. Future research in this area is warranted.

This dissertation also highlights the importance of methodology development in the study of social interactions. Despite the common theme, the studies presented here have adopted
different analytic strategies, which allow us answer different research questions. The first study used multilevel modeling, which pools across dyads under the assumption that all dyads follow a common pattern (i.e., the basic model is the same for all couples). Heterogeneity across dyads is addressed by estimating random effects, which are post hoc predictions based on group parameters under the assumption of normality. In other words, the random coefficient for a specific dyad depends on all other dyads in the sample. Hence, multilevel models are in essence group-based models. In studies that aim at making inferences at the group or population levels, such as the first study, multilevel modeling represents a proper and powerful modeling technique. Results from multilevel models, however, only sufficiently describe each individual dyad when the assumption of ergodicity holds (Molenaar, 2004). Hence, in studies that aim to draw inferences to individual dyads, multilevel modeling is not optimal, and other modeling techniques need to be considered.

The second and third studies illustrate the use of time series analysis. This approach pools across time rather than units of analysis. Because time series data consist of intensive measurements from each individual dyad, models can be estimated for each dyad and compared across dyads. This allows testing of ergodicity and ensures validity of making inferences at the individual level. In the third study, for example, time series analysis has enabled us to identify patterns of interactions for a specific child-therapist dyad and potentially develop interventions or treatments that are tailored to this dyad. This cannot be accomplished in the multilevel modeling framework.

It should be noted that despite their apparent differences, time series analysis and multilevel modeling can share the same basic model. For example, the VAR model in the second and third studies have the same form with the cross-lagged regression model in the first study. In
both cases, the models involve repeated measures from two social partners, and an individual’s score on a particular occasion is regressed on his/her own and the partner’s score on the previous occasion (for VAR models with order 1). The only difference between the two is that the cross-lagged model pools information from multiple dyads, thus requiring fewer measurements from each dyad but a larger number of dyads, whereas the VAR model focuses on one dyad at a time and pools information from multiple occasions, thus requiring more measurements from each dyad over time but a smaller number of dyads. The selection of analytic approach and the corresponding study design has to be made based on the research question and purpose of study. Given the difficulty in obtaining time series data in the behavioral sciences, multilevel models may be more cost-efficient when inferences are to be made at the group or population levels. In the case where individual level inferences and predictions are desired, time series design and analysis are preferred.
References


and network architecture for wirelessly communicating and logging autonomic activity.


in afternoon cortisol levels: Twin and parent-offspring designs.  

*Psychoneuroendocrinology, 31*(9), 1131-1137.


*Journal of Clinical Endocrinology & Metabolism, 82*(8), 2458-2465.


*Journal of Communication, 27*(2), 74-79.


**APPENDIX A**

**Table 1.** Descriptive Statistics of Cortisol and Time of Collection.

<table>
<thead>
<tr>
<th>Occasion</th>
<th>N</th>
<th>Mean (nmol/L)</th>
<th>SD</th>
<th>Mean Collection Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (upon wake)</td>
<td>171</td>
<td>26.86</td>
<td>11.73</td>
<td>6:48 a.m.</td>
</tr>
<tr>
<td>2 (30 min after wake)</td>
<td>176</td>
<td>32.38</td>
<td>11.90</td>
<td>7:25 a.m.</td>
</tr>
<tr>
<td>3 (afternoon)</td>
<td>165</td>
<td>14.70</td>
<td>9.07</td>
<td>1:00 p.m.</td>
</tr>
<tr>
<td>4 (before bed)</td>
<td>170</td>
<td>6.39</td>
<td>5.88</td>
<td>11:35 p.m.</td>
</tr>
</tbody>
</table>
### APPENDIX B

**Table 2.** Descriptive Statistics of Cortisol Awakening Response (CAR) and Diurnal Cortisol Slope (DCS), Broken Down by Gender.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>CAR</td>
<td>103</td>
<td>7.25</td>
</tr>
<tr>
<td>DCS</td>
<td>101</td>
<td>-1.69</td>
</tr>
</tbody>
</table>

*Note: Values represent changes in cortisol (nmol/L) per hour.*
APPENDIX C

Table 3. Descriptive Statistics and Correlations for Spouse Support, Strain, and Disagreement.

<table>
<thead>
<tr>
<th></th>
<th>Support</th>
<th>Strain</th>
<th>Disagreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strain</td>
<td>-0.26</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Disagreement</td>
<td>-0.31</td>
<td>0.69***</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>26.43</td>
<td>10.91</td>
<td>10.07</td>
</tr>
<tr>
<td>SD</td>
<td>1.34</td>
<td>1.72</td>
<td>2.10</td>
</tr>
<tr>
<td>N</td>
<td>28</td>
<td>28</td>
<td>23</td>
</tr>
</tbody>
</table>

***p<.001.
## APPENDIX D

### Table 4. Synchrony in Cortisol Awakening Response (CAR) between Spouses.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>No Moderator</th>
<th>Moderator: Spouse Strain</th>
<th>Moderator: Spouse Disagreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept for Husband</td>
<td>4.19</td>
<td>4.04</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>(3.94)</td>
<td>(4.03)</td>
<td>(4.01)</td>
</tr>
<tr>
<td>Intercept for Wife</td>
<td>15.68***</td>
<td>15.14**</td>
<td>12.01*</td>
</tr>
<tr>
<td></td>
<td>(3.85)</td>
<td>(4.21)</td>
<td>(4.41)</td>
</tr>
<tr>
<td>Wake Time (Between-Person)</td>
<td>3.38</td>
<td>3.64</td>
<td>3.07</td>
</tr>
<tr>
<td></td>
<td>(2.81)</td>
<td>(2.92)</td>
<td>(3.10)</td>
</tr>
<tr>
<td>Wake Time (Within-Person)</td>
<td>2.54</td>
<td>2.26</td>
<td>-0.52</td>
</tr>
<tr>
<td></td>
<td>(2.09)</td>
<td>(2.08)</td>
<td>(2.30)</td>
</tr>
<tr>
<td>Spouse’s Deviation from Person-Mean</td>
<td>0.19</td>
<td>-0.03</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.11)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Moderator</td>
<td>-</td>
<td>0.51</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.73)</td>
<td>(1.37)</td>
</tr>
<tr>
<td>Moderator × Spouse’s Deviation from Person-Mean</td>
<td>-</td>
<td>0.16**</td>
<td>0.16**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>N</td>
<td>124</td>
<td>124</td>
<td>96</td>
</tr>
</tbody>
</table>

**Note:** Wake time (between-person), spouse strain, and spouse disagreement were centered by sample mean. *p < .05; **p < .01; ***p < .001.
### APPENDIX E

**Table 5.** Synchrony in Diurnal Cortisol Slope between Spouses.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept for Husband</td>
<td>-1.81***</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept for Wife</td>
<td>-1.77***</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAR (Between-Person)</td>
<td>-0.02**</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAR (Within-Person)</td>
<td>-0.02***</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spouse’s Deviation from Person Mean</td>
<td>0.18*</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>112</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* CAR (between-person) was centered by sample mean. *p < .05; **p < .01; ***p < .001.
### APPENDIX F

**Table 6.** Cross-Lagged Models of Cortisol Awakening Response (CAR) and Diurnal Cortisol Slope (DCS) between Spouses.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>CAR</th>
<th>DCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept for Husband</td>
<td>0.88</td>
<td>-1.09**</td>
</tr>
<tr>
<td></td>
<td>(3.54)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>Intercept for Wife</td>
<td>6.84†</td>
<td>-1.33***</td>
</tr>
<tr>
<td></td>
<td>(3.58)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>Lag-1 Autoregressive</td>
<td>0.42***</td>
<td>0.25†</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Spouse Support</td>
<td>-0.42</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(1.68)</td>
<td></td>
</tr>
<tr>
<td>Spouse Support × Lag-1 Autoregressive</td>
<td>0.16*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>Cross-Lagged</td>
<td>0.16</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>N</td>
<td>90</td>
<td>84</td>
</tr>
</tbody>
</table>

*Note:* Spouse support was centered by sample mean. †p < .1; *p < .05; **p < .01; ***p < .001.
## APPENDIX G

### Table 7. Estimated VAR Models on Complete Data, Broken Down by Conditions.

<table>
<thead>
<tr>
<th>$A_l$</th>
<th>[ \begin{array}{cc} -0.4 &amp; 0 \ -0.3 &amp; -0.4 \end{array} ]</th>
<th>[ \begin{array}{cc} -0.4 &amp; 0 \ 0.3 &amp; -0.4 \end{array} ]</th>
<th>[ \begin{array}{cc} -0.8 &amp; 0 \ -0.3 &amp; -0.8 \end{array} ]</th>
<th>[ \begin{array}{cc} -0.8 &amp; 0 \ 0.3 &amp; -0.8 \end{array} ]</th>
<th>[ \begin{array}{cc} 0.4 &amp; 0 \ -0.3 &amp; 0.4 \end{array} ]</th>
<th>[ \begin{array}{cc} 0.4 &amp; 0 \ 0.3 &amp; 0.4 \end{array} ]</th>
<th>[ \begin{array}{cc} 0.8 &amp; 0 \ -0.3 &amp; 0.8 \end{array} ]</th>
<th>[ \begin{array}{cc} 0.8 &amp; 0 \ 0.3 &amp; 0.8 \end{array} ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR(1)$^a$</td>
<td>88%</td>
<td>90%</td>
<td>89%</td>
<td>82%</td>
<td>89%</td>
<td>85%</td>
<td>85%</td>
<td>93%</td>
</tr>
<tr>
<td>DTF12$^b$</td>
<td>.0529</td>
<td>.0459</td>
<td>.0423</td>
<td>.0417</td>
<td>.0465</td>
<td>.0509</td>
<td>.0435</td>
<td>.0367</td>
</tr>
<tr>
<td></td>
<td>(.0395)</td>
<td>(.0375)</td>
<td>(.0360)</td>
<td>(.0319)</td>
<td>(.0377)</td>
<td>(.0382)</td>
<td>(.0446)</td>
<td>(.0338)</td>
</tr>
<tr>
<td>DTF21$^b$</td>
<td>.2827</td>
<td>.2867</td>
<td>.2930</td>
<td>.2790</td>
<td>.2673</td>
<td>.2552</td>
<td>.2759</td>
<td>.3008</td>
</tr>
<tr>
<td></td>
<td>(.0953)</td>
<td>(.0707)</td>
<td>(.0840)</td>
<td>(.0937)</td>
<td>(.0795)</td>
<td>(.0853)</td>
<td>(.0924)</td>
<td>(.0713)</td>
</tr>
</tbody>
</table>

*Note*: N = 100.

$^a$Percentage of first order VAR models in all estimated models.

$^b$Means and standard deviations (in parentheses) of the estimated directed transfer functions (DTF).
# APPENDIX H

## Table 8. Effects of Technique, AR_LARGE, and AR_POSITIVE on DTF21.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
<th>p-value</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Within-Subject Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technique</td>
<td>2.79</td>
<td>492.68</td>
<td>.000</td>
<td>.38</td>
</tr>
<tr>
<td>Technique by AR_POSITIVE</td>
<td>2.79</td>
<td>296.67</td>
<td>.000</td>
<td>.27</td>
</tr>
<tr>
<td>Technique by AR_LARGE</td>
<td>2.79</td>
<td>97.62</td>
<td>.000</td>
<td>.11</td>
</tr>
<tr>
<td>Error (Technique)</td>
<td></td>
<td>2224.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Between-Subject Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR_POSITIVE</td>
<td>1</td>
<td>52.59</td>
<td>.000</td>
<td>.06</td>
</tr>
<tr>
<td>AR_LARGE</td>
<td>1</td>
<td>7.01</td>
<td>.000</td>
<td>.01</td>
</tr>
<tr>
<td>Error</td>
<td></td>
<td>797</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: N = 100.*
APPENDIX I

(A) Spouse Strain Moderates Synchrony in CAR

(B) Spouse Disagreement Moderates Synchrony in CAR

Figure 1. Marital Relationship Moderates Synchrony in Cortisol Awakening Response (CAR).
Figure 2. Estimated Marginal Means of DTF21 by Technique, AR_LARGE, and AR_POSITIVE. Complete = Complete Data. Listwise = Listwise Deletion. Random = Imputation Using Sample Means and Variances. Recursive = Recursive Prediction.
Figure 3. Estimated Marginal Means of DTF21 by Technique and AR_LARGE. Complete = Complete Data. Listwise = Listwise Deletion. Random = Imputation Using Sample Means and Variances. Recursive = Recursive Prediction.
APPENDIX L

Figure 5. Estimated Marginal Means of DTF21 by Proportion of Missing Data.
APPENDIX N

AR_{\text{large}} = 0

AR_{\text{large}} = 1
Figure 6. Estimated Marginal Means of DTF21, Broken Down by Technique, Proportion of Missing, AR_LARGE, and AR_POSITIVE.
APPENDIX O

Figure 8. Estimated Marginal Means of DTF12 by Proportion of Missing Data.
Figure 9. Electrodermal Activity (EDA) Levels (μ mhos) from Child and Therapist during One Therapy Session (N=5338).
**APPENDIX R**

*Figure 10.* Autospectra of the Child. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes.
Figure 11. Autospectra of the Therapist. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes.
**Figure 12.** Power Estimates of the First Frequency Component across Time.
Figure 13. (a) Coherence between Child and Therapist across Time and Frequency. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes. (b) Average Coherence over Frequencies across Time.
**Figure 14.** Directed Transfer Function from the Therapist to the Child. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes.
Figure 15. Directed Transfer Function from the Child to the Therapist. Markers on the Frequency Axis Represent Number of Cycles Per 2 Minutes.
Figure 16. Guided Activity and Directed Transfer Function between Child and Therapist Averaging over Frequencies.
CURRICULUM VITAE

Siwei Liu

EDUCATION

Ph.D. Human Development and Family Studies, August 2012
The Pennsylvania State University, University Park, PA

M.A.S. Applied Statistics, August 2012
The Pennsylvania State University, University Park, PA

M.S. Human Development and Family Studies, December 2008
The Pennsylvania State University, University Park, PA

B.S. Psychology, July 2006
Fudan University, Shanghai, China

PUBLICATIONS

Liu, S., Rovine, M., & Molenaar, P. C. M. (in press), Using fit indices to select a proper covariance structure for longitudinal data. Structural Equation Modeling.


Liu, S., & Hynes, K., (2012), Are difficulties balancing work and family associated with subsequent fertility? Family Relations. 61, 16-30.


HONORS AND AWARDS

2012 Dissertation Award, Society of Multivariate Experimental Psychology