PARASYMPATHETIC RECOVERY FOLLOWING ANGER RECALL IN YOUNG ADULT FEMALES

By

CHARLES TODD SULLIVAN

A THESIS PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN EXERCISE AND SPORT SCIENCES

UNIVERSITY OF FLORIDA

2003
Copyright 2003

by

Charles Todd Sullivan
This document is dedicated to those individuals that have endured my numerous theories and ramblings regarding human biology. Some of my “second pitcher” thoughts, generally expressed to friends too inebriated to walk away, at The Swamp, are now on paper. If they are foolish enough to listen to me again, I may eventually write a dissertation.
ACKNOWLEDGMENTS

I wish to thank my friends and family for their support. I also wish to thank my committee, chaired by Milledge Murphey, Ph.D. This thesis could not have been accomplished without their help.
TABLE OF CONTENTS

ACKNOWLEDGMENTS ......................................................................................................................... iv
LIST OF TABLES ................................................................................................................................. ix
LIST OF FIGURES ..................................................................................................................................... x
ABSTRACT .............................................................................................................................................. xi

CHAPTER

1 INTRODUCTION .................................................................................................................................... 1

Overview of Topics ................................................................................................................................. 2
Cardiac Performance and Estrogen’s Influence .................................................................................... 2
Affective States, the Brain, and the Heart ................................................................................................. 3
Specific Methodological Concerns ......................................................................................................... 4
Relevance of a Gender-Specific Approach ............................................................................................. 5
Overview of the Menstrual Cycle ........................................................................................................... 7
Confusing Results Regarding Hormone Replacement Therapy ............................................................. 9
Statement of Purpose ............................................................................................................................... 12
Research Hypotheses ............................................................................................................................... 12
Hypothesis One ......................................................................................................................................... 12
Hypothesis Two ......................................................................................................................................... 13
Hypothesis Three ....................................................................................................................................... 14
Significance of the Study ......................................................................................................................... 14

2 REVIEW OF THE LITERATURE .................................................................................................. 15

Cardiac Performance and Estrogen’s Influence ..................................................................................... 15
Cardiovascular Disease ......................................................................................................................... 15
Baroreflex .................................................................................................................................................. 17
Integration of Executive Brain Centers ................................................................................................... 18
Heart Rate Variability ............................................................................................................................... 21
What Is Being Measured With Spectral Methodology ......................................................................... 22
How Else Is Variability Measured ........................................................................................................... 23
How Do We Differentiate Heart Rate Components at Rest and During Various States of Arousal ........ 24
Do Mental Stress and Affective State Show Unique Effects on Heart Rate Components ................. 25
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate Analysis</td>
<td>80</td>
</tr>
<tr>
<td>Overall Results of Hypothesis Testing</td>
<td>81</td>
</tr>
<tr>
<td>Hypotheses Testing</td>
<td>82</td>
</tr>
<tr>
<td>Power and Effect Size</td>
<td>85</td>
</tr>
<tr>
<td>Analysis Progression</td>
<td>85</td>
</tr>
<tr>
<td>Hypothesis One</td>
<td>85</td>
</tr>
<tr>
<td>Hypothesis Two</td>
<td>86</td>
</tr>
<tr>
<td>Best Subsets Analysis</td>
<td>89</td>
</tr>
<tr>
<td>Diagnostics</td>
<td>90</td>
</tr>
<tr>
<td>Assumption One: Linear Coefficients, Correct Model Specification</td>
<td>91</td>
</tr>
<tr>
<td>Additive Error Term</td>
<td></td>
</tr>
<tr>
<td>Assumption Two: Error Term Has a Zero Mean</td>
<td>92</td>
</tr>
<tr>
<td>Assumption Three: Error Term Has Constant Variance, no Heteroskedasticity</td>
<td>93</td>
</tr>
<tr>
<td>Assumption Four: No Serial Correlation Between the Observations</td>
<td>95</td>
</tr>
<tr>
<td>Assumption Five: Error Terms Distributed Normally</td>
<td>95</td>
</tr>
<tr>
<td>Assumption Six: Error Terms Not Correlated with Explanatory Variables</td>
<td>96</td>
</tr>
<tr>
<td>Assumption Seven: Predictor Variables not Linearly Related, no Multicollinearity</td>
<td>97</td>
</tr>
<tr>
<td>Other Diagnostic Issues</td>
<td>98</td>
</tr>
<tr>
<td>Hypothesis Three</td>
<td>102</td>
</tr>
<tr>
<td>Prediction Analysis</td>
<td>103</td>
</tr>
<tr>
<td>Supplementary Analysis</td>
<td>105</td>
</tr>
<tr>
<td>Overall Synopsis</td>
<td>107</td>
</tr>
<tr>
<td>5 CONCLUSION</td>
<td>109</td>
</tr>
<tr>
<td>Research Findings and Their Implications</td>
<td>109</td>
</tr>
<tr>
<td>Practical Implications</td>
<td>110</td>
</tr>
<tr>
<td>Secondary Implications</td>
<td>112</td>
</tr>
<tr>
<td>Agreements and Contrasts with Previous Findings</td>
<td>113</td>
</tr>
<tr>
<td>Limitations and Future Directions</td>
<td>115</td>
</tr>
<tr>
<td>Design Considerations</td>
<td>116</td>
</tr>
<tr>
<td>Final Conclusion</td>
<td>117</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>A STANDARD ESTROGEN VALUES</td>
<td>119</td>
</tr>
<tr>
<td>B QUESTIONNAIRES ONE AND TWO</td>
<td>121</td>
</tr>
<tr>
<td>Questionnaire One</td>
<td>121</td>
</tr>
<tr>
<td>Appropriate Answers for Inclusion</td>
<td>122</td>
</tr>
<tr>
<td>Questionnaire Two</td>
<td>122</td>
</tr>
</tbody>
</table>
LIST OF REFERENCES ................................................................................................ 123

BIOGRAPHICAL SKETCH .......................................................................................... 135
LIST OF TABLES

Table | Page
--- | ---
4-1 | Summary data for the anger/baseline ratio manipulation check 78
4-2 | Descriptive statistics for N=44 scores passing the manipulation check 82
4-3 | Summary statistics for hypothesis one. Baseline regressed on estrogen 83
4-4 | Summary statistics for hypothesis two. Full model 84
4-5 | Best subsets regression results 91
4-6 | Comparison with and without estrogen, with baseline a constant 105
4-7 | Regression of R/B ratio on Estrogen values 106
4-8 | Regression of recovery HF power on baseline HF power and menstrual phase.. 107
A-1 | Estrogen values for each day, shown as percentage of the maximum 120
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-1</td>
<td>Conceptual diagram of partial and semipartial correlation coefficients</td>
<td>85</td>
</tr>
<tr>
<td>4-2</td>
<td>Scatter plots for first order and second order predictors with recovery.</td>
<td>88</td>
</tr>
<tr>
<td>4-3</td>
<td>Scatter plots among interactions and recovery.</td>
<td>88</td>
</tr>
<tr>
<td>4-4</td>
<td>Scatter plot matrices of first order effects alone.</td>
<td>91</td>
</tr>
<tr>
<td>4-5</td>
<td>Residuals versus fits for model one, showing relatively equal dispersion of residuals about zero, indicating mean error of zero</td>
<td>93</td>
</tr>
<tr>
<td>4-6</td>
<td>Residual plot showing standardized residuals plotted against baseline HF, displaying possible curvature and heteroskedasticity</td>
<td>95</td>
</tr>
<tr>
<td>4-7</td>
<td>Plot of Cook’s D against each observation. Note highest amount associated with observation one</td>
<td>101</td>
</tr>
<tr>
<td>A-1</td>
<td>Estrogen values shown as percentage of maximum.</td>
<td>119</td>
</tr>
</tbody>
</table>
Abstract of Thesis Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Master of Science in Exercise and Sport Sciences

PARASYMPATHETIC RECOVERY FOLLOWING ANGER RECALL IN YOUNG ADULT FEMALES

By

Charles Todd Sullivan

December 2003

Chair: Milledge Murphey
Major Department: Exercise and Sport Sciences

Cardiovascular disease in females is often diagnosed with less precision than in males. A potential reason is the impact that fluctuating estrogen levels may have on indices of cardiovascular function. Prior work has shown that estrogen enhances parasympathetic effect, an important prognostic aid for predicting cardiovascular disease. Additionally, research has shown that cognitive and emotional stress can significantly alter autonomic function. This investigation tested the main hypothesis that parasympathetic recovery from an affective stressor, anger recall, could be modeled and predicted in young healthy females (N=43), using estimated values of estrogen levels, circadian status, and resting levels of parasympathetic function. Parasympathetic function was indexed using the power within the high frequency component of a heart rate spectral analysis. Participants followed a resting period with five minutes of anger recall,
followed by a recovery period of five minutes. Results showed that regression modeling explained 95% of the variance in high frequency power during the recovery period, with estimates of estrogen levels and baseline measures being the only significant predictors. Importantly, the addition of estrogen, to an equation consisting of baseline measures alone, reduced the prediction interval for all scores by 33%, making its estimation a valuable addition. Given the greater complexity found in diagnosing cardiovascular disease in women as compared to men, and the value of parasympathetic functioning in the prediction and prognosis of cardiovascular disease, accounting for estrogen levels may be of critical importance in the future. The results of this investigation indeed indicate that robust predictions can be made in a healthy cohort of females. This provides a solid foundation for further work in more varied populations, utilizing more rigorous approaches to the quantifications of estrogen levels. Future work should refine this methodology and make it suitable for clinical use.
CHAPTER 1
INTRODUCTION

Parasympathetic-effect is a broad term referring to physiological adjustments
subsequent to outflow of parasympathetic, or vagal, tone from autonomic centers in the
brain stem. Its abatement and return following various challenges is an important
prognostic tool in the cardiovascular sciences. In the past decade, biomedical research has
paid closer attention to gender specificity. One of the concepts to emerge from recent
research is that estrogen has the ability to enhance the effectiveness of parasympathetic
tone, generally regarded as protective for cardiovascular health. Utilizing an affective
challenge, this thesis seeks to exploit the fluctuation in hormonal levels that accompany
the menstrual cycle; and, model those effects upon parasympathetic return following an
anger recall task in a young, healthy female population.

The importance of this investigation stems from the lethality of cardiovascular
disease in males and females, as well as the inadequate knowledge regarding response
and recovery following affective challenges. In particular, the use of affective, or
emotional, challenges has shown considerable promise in the evaluation of
cardiovascular function. Because these modalities at least partially control for peripheral
metabolic effects on cardiac function, organic disease of the myocardium may be
detected with increased power, as compared to traditional exercise testing.

Cardiovascular disease remains the biggest mortal threat to both genders, and
specifically, cardiovascular disease is greatly underestimated in the female community.
Psychological tasking has the potential of providing greater diagnostic power, which is currently inadequate for the female population (Silber, 2003).

An added facet of the current investigation is that it attempts to model the effects of estrogen, utilizing a regression analysis. Despite the progress made in basic science knowledge, in order for emotional testing to become a useful clinical tool, an established set of criteria is needed to determine normal from abnormal. Regression analysis provides confidence intervals that can be used to evaluate the likelihood of a particular score. Though major theoretical breakthroughs are not anticipated as a result of this work, an extension of previous results and the evaluation of regression modeling will provide useful additions. Positive results of a regression analysis would provide the impetus for continued research using more precise measurement tools and the eventual inclusion of clinical populations.

**Overview of Topics**

**Cardiac Performance and Estrogen’s Influence**

Cardiovascular performance is one of the most widely studied areas in all of science. As such, any attempt at a comprehensive review would be foolish. In the sections to follow, a brief review of literature citing differences in cardiac parameters thought to be attributable to varying estrogen levels will be addressed.

At this time, there is debate regarding the true global influence that estrogen has upon human physiology. In addition, there are ongoing investigations and re-evaluations of previous studies in order to determine the effectiveness and safety of hormone replacement therapy following menopause. To say the overall analysis is ambiguous is an understatement. As such, caution must be utilized on the part of the reader in the following regard: while the most agreed upon consensus statements will be presented
here, along with the most reputable research reports, it is most likely the case that future knowledge will radically shift research paradigms and consensus statements.

A recent example relates to hormone replacement therapy (HRT), during and after menopause. The major authorities in women’s health and research have varied their stances in regard to the benefits and detriments HRT and estrogen supplementation in particular. A recent review highlights some of the conclusions and methodological shortcomings in the major clinical trials evaluating HRT (Genazzani et al. 2003). Many of the concerns expressed in that review are relevant to all studies looking at gender issues. Briefly, problems with sample selection, treatment allocation, and at times the statistical approach have all hindered development. A major reason for this is due to the fact that most large-scale studies targeted at gender issues have used populations that are not directly comparable. These differences have brought about contrasting results. However, the focus and protocol of this investigation have been devised in such a way so as to alleviate these potential pitfalls. Namely, the relatively simple statistical technique of multiple regression and the purposefully targeted, homogenous population of healthy females will provide results that are not the most powerful for clinical application, yet should not confound inferences made concerning the observed variables. This will become clear in the explanation of the methods.

**Affective States, the Brain, and the Heart**

The most recent series of clinical investigations evaluating affective challenges and the subsequent changes in cardiac performance were brought forward by the PIMI investigators (Goldberg et al., 1996). These results have clearly shown that psychological challenges have effects on the cardiovascular system that are measurable and at times severe. Additionally, these effects produce quick adjustments that can be measured in a
small amount of time, just as conventional methods currently in use, such as graded exercise testing. An overlooked asset of psychological testing is that abnormalities of cardiac performance cannot be attributed to metabolic demands at the periphery, or artifacts due to body movement. The advantage of this characteristic is due to the fact that cardiac performance is then dependent on only two factors, 1) the coding of autonomic signals, meaning the rate of impulses and the relative amount of parasympathetic and sympathetic tone, and 2) the ability of the myocardium and conduction systems to incorporate these signals accurately. Practically speaking, the second factor mentioned above is the only one clinically measurable. By removing influences from metabolic demands and body movement, abnormalities of cardiac performance can be most accurately attributed to organic disturbances within the myocardium, which is the location where cardiac disease most often begins, yet is arguably the costliest and most difficult place to be evaluated.

Currently, there is a lack of understanding in the mechanisms by which affective states may modulate cardiac activity. To begin understanding such relationships, some information of higher brain centers and findings linking their activation to emotional states and subsequent cardiovascular changes is required. Thus, a discussion will be presented on such findings, particularly focusing on the anterior cingulate cortex, and the orbitofrontal cortex, elements of executive centers.

**Specific Methodological Concerns**

Recovery measures are rarely utilized. Instead, reactivity measures are most often the targets of data analysis. Some aspects making recovery measures more powerful include the ability to use baseline and reactivity measures as covariates if necessary. Also, using recovery values and regression analysis is more powerful because it is
possible to account for and model data that may not be amenable to analysis of variance. It also allows prediction, not just the ability to say that a difference exists.

Additionally, the measures and methods of collection are perhaps the most agreed upon noninvasive techniques. A current problem in the literature is that variables and the parameters for their collection are varied and often authors do not adhere to published guidelines. Adding to this, much of the research relating emotions to cardiovascular function has been undertaken by those not well educated in cardiovascular function. Often, such research produces invalid results due to misappropriation of methods, measures, and invalid interpretations. It will be presented that the methods utilized in this thesis are standardized and will attain balance in attending to the clinical, psychological, and physiological rigor necessary for credibility and interpretability.

An important caveat is the following: Though many authors claim to utilize the same standards employed in this thesis, vastly differing results can be found in the literature. This is due to nuances in the data reduction and reporting processes. As such, the reader must be cautioned that these results will not be directly comparable to all others claiming the same standard analysis conditions. Unfortunately, the consensus guidelines used here are quite dated and are in need of revision. In addition, a concerted effort is required of all researchers to standardize the reduction and reporting of data so that efficient analysis of previous research is possible. While such issues do not invalidate the methods utilized here, it is only fair that the reader be informed of these shortcomings.

Relevance of a Gender-Specific Approach

Research targeted at women’s health issues has increased in amount and relevance over the past two decades. It is clear that male dominated protocols are not longer
generalizable to the female population in every instance. However, despite the increases in gender research, in general, mortality and morbidity statistics for women have not benefited to the same extent as those of males. For instance, since 1984, deaths due to cardiovascular disease have decreased in males, but increased in females. Many reasons can potentially be sought for this discrepancy. However, what appears to be most germane is the lack of research specifically addressing biological differences between the genders. This discrepancy remains in spite of the aforementioned increases in gender-specific research.

Despite the emergence of holistic medicine, and the perceived need to address psychosocial as well as purely physiological requirements, today’s medical system is almost completely based on traditional models of research. This is known as evidence based medicine. It serves as the foundation for today’s medical model. It’s shortcomings for the treatment of women stem primarily, in this author’s opinion, from a lack of scientific inquiry targeted at gender-specific hormonal differences. The argument, in its most reduced form, states that the contrasting hormonal milieu between males and females has far reaching effects not completely realized. The gonadal hormones are now known to produce nongenomic effects, a phenomenon not at all discussed in medical texts as recently as 20 years ago. Such discoveries have most certainly been quickened by the desire to pursue gender differences. While shortcomings do exist, as mentioned above, it is encouraging that current research emphasizes emerging gender specificity and the enhanced understanding of human biology, and such practices can only bolster the standard of care for both genders.
Most relevant to this thesis, differences in cardiovascular disease, cardiac performance, and links between affective state and cardiac performance will be addressed. Reviewing the literature will present an overall picture of current knowledge, yet in each major section, the impact of gender will be discussed.

The reader should be aware that although this thesis focuses solely on a female sample, the aim of this investigation is not to further expose gender differences in cardiovascular disease. (For quick and efficient access to such topics, the reader is referred to the American Heart Association’s web site, www.americanheart.org, for the latest statistical and practical information.) The aim is to model the effects of varying estrogen levels on the return of parasympathetic-effect following a task that commonly brings about a profound decrease in parasympathetic influence. However, to lay foundation for such a study, some necessary inclusion of epidemiological data parsing the cardiovascular differences between the sexes is mandatory.

Lastly, an overview of the menstrual cycle is given below. Though detailed working knowledge of the cycle is not necessary, rudimentary understanding of the cycle and the fluctuations of estrogen levels are valuable to the understanding of this investigation.

**Overview of the Menstrual Cycle**

The menstrual cycle in women is responsible for tremendous variation in blood chemistry. Menarche marks the initiation of puberty and the beginning of a cyclical, pulsatile rhythm of neuroendocrine operation designed to maintain reproductive function and viability. The termination of this sequence results in menopause. On average, girls reach menarche at age 12, and women enter a phase of progressive decline in fertility beginning approximately at age 35, commonly known as the perimenopausal period. The
eventual entrance into menopause generally occurs in the fifth decade of life. Once menopause is reached, the cyclic variation in reproductive hormones ceases.

The consistency of menstrual rhythms is less secure during the earliest and latest periods of a women’s reproductive life. This inconsistency can make research seeking to evaluate rhythmic influences problematic. In the main, most women attain consistent menstrual cycles, with a median intercycle length of 28 days. Importantly, such fluctuations in intercycle length generally disappear by late adolescence. This supports the need for protocols specifically addressing adolescent populations, and gives credibility to those protocols using females that are in the middle of their reproductive years. Although seemingly trivial, it is important to know that the standard reference point for determination of cycle initiation is the first day of menstrual bleeding. This event marks the end of the previous cycle and is readily obvious, and thus serves as the most accurate way to mark cycle timelines. Many researchers create artificial timelines for the menstrual cycle or fail to rigidly define the boundaries of the cycle as they pertain to their particular investigation. Attention to such detail should be paid by researchers and research consumers to avoid confusion.

Hormonal dynamics are conveniently divided into four phases. Each phase can be further subdivided, however, such subdivisions will not be explored here. The follicular phase shows an overall increase in levels of estrogen and constitutes approximately half of the total menstrual cycle. Graphically, it culminates with a sharp rise in estrogen levels. It commences with the first day of menstrual bleeding. Next is the ovulatory phase, lasting approximately two days. It immediately follows the sharp rise in estrogen levels in the last part of the follicular phase. The main effect associated with the
ovulatory phase is that it marks the beginning of a consistent rise in gonadotropins and progesterone levels. Also, plasma estrogen levels display a sharp decline during this phase. Following the ovulatory phase is the luteal phase, lasting approximately 12 days. Its most prominent feature is the sharp and sustained rise in progesterone. The last phase is the menstrual, specifically pertaining to the bleeding activity that takes place if pregnancy is not attained.

The primary focus here is upon the change in estrogen levels. Many authors utilize a dichotomy to characterize the participants. Typically authors classify participants as being in the follicular or luteal phase. However, most research pertinent to this thesis has been equivocal in defining specific roles for reproductive hormones. It is this author’s opinion that this fault has stemmed from the inability to adequately characterize estrogen fluctuations with such a dichotomous scheme (Appendix A provides a graphical display of estrogen fluctuations). Intuitively, dividing the cycle into two parts does not allow for even semi-precise accounting for estrogen levels and subsequent physiological effects. Recent research has shown that estrogen mediated effects are dose dependent, receptor dependent, and possible even without the presence of estrogen receptors (Tolbert & Oparil, 2001). Given this multiplicity, it is logical to attempt a more accurate accounting of estrogen levels.

Confusing Results Regarding Hormone Replacement Therapy

As a brief example regarding the debate of estrogen’s role in women’s health, some quick review of postmenopausal studies will be issued below.

In addition to defining the biological effects that vary through the normal menstrual cycle, much attention has been paid to health status during menopause. Menopause brings about a sharp decline in circulating estrogen levels. Prior to menopause, females have a
markedly lower rate of cardiovascular disease, as compared to males. Once menopause is reached, the risk of disease rises sharply (American Heart Association [AHA], 2001). The causal link for this effect is most often speculated to be the change in estrogen levels. Despite several large studies and many years of targeted research, no consensus statement is present.

What appears to be clear is that estrogen does indeed affect disease risk, whether directly or indirectly. However, what remains to be decided is the determination of appropriate dosages, time for therapy initiation, and other drug combinations that may be of benefit. These questions remain in lieu of thousands of study participants for several reasons. The primary pitfall has been the selection of study participants. For instance the recent discouraging results from the WHI (Women’s Health Initiative) and HERS (Heart and Estrogen/progestin Replacement Study) have been criticized for their inappropriate inclusion criteria (Genazzani et al., 2003). These studies have utilized inclusion criteria much too broad to allow for precise causal inferences.

Though basic science research supports a protective role for the use of hormone replacement therapy, results from the studies cited above seem to indicate otherwise. In the WHI study, researchers found that hormone replacement therapy provided an unexpected increased risk for heart disease when compared to the placebo. However, upon closer examination, the placebo group displayed an exaggerated and abnormal decrease in the incidence of heart disease, causing the treatment group to appear to be at a higher risk. Critics have argued that the study population included too many subjects with co-morbidities that were either unaccounted for or not possible to be account for statistically. Also, in the WHI, an accounting for time since the onset of menopause was
not performed. This may confound the findings because if health changes ensue as a result of estrogen decline, then these changes would exist in varying degrees at study initiation. Given the possibility of threshold effects for disease progression and remediation, such accounting may provide spurious results. In essence, the lack of experimental control due to participant selection has provided spurious results despite a large sample size.

In the HERS trial, similarly unexpected findings resulted. Treatment resulted again in an increased risk early in the trial, with a decreased risk towards the end of a four-year period. However, in subsequent analysis, treatment again provided an increase in risk in later years. It has been pointed out that the study population was extremely diverse, encompassing over thirty years difference in participant ages. Some reviewers have argued that the statistical analysis was to blame for the findings more than the treatment protocol (Genazzani et al., 2003). Also, these reviews point to high dropout rates, and when any adverse incident took place, appropriate statistical measures were not employed.

The main conclusion reached from the above examples is that while large amounts of raw data have been collected regarding menstrual function and life after reproductive functioning has ceased, so much of this data has been accumulated in heterogeneous samples that firm conclusions are hard to reach. Though an oversimplification would likely be inappropriate, it may be wise for many studies to greatly restrict participant variables. Some recommendations for restrictive criteria include age, body mass index, reproductive timeline status, prior disease incidence, parity, and possibly family history.
Most of these criteria will be utilized to restrict the sample for this thesis. Further information will be provided in the methods section.

**Statement of Purpose**

The purpose of this investigation is to establish the viability of modeling the varying levels of estrogen throughout the menstrual cycle in order to create precise regression equations to model recovery from an anger recall task.

**Research Hypotheses**

Briefly, this study utilized a young healthy cohort of females. They were asked to perform an anger recall task, bracketed by a baseline measure and a recovery measure. Estrogen values were estimated by their status within the menstrual cycle. Additionally, their circadian status was measured as the number of hours since awakening. A regression analysis was performed which included baseline, estrogen, and circadian values as independent variables, and recovery as the dependent variable. Prior research has shown that incidence of adverse cardiovascular events are occasionally linked to circadian rhythm. Also, as mentioned above, estrogen enhances parasympathetic tone in repeated measures. However, there are no studies known to this author showing that estrogen values predict baseline parasympathetic measures, across individuals. The hypotheses developed were primarily driven by the research focus and to validate statistical assumptions and nuances germane to regression analysis.

**Hypothesis One**

Not directly related to the primary objective of the study, this hypothesis is generated to add weight to the assumption that the predictor variables are not related by causal or chance mechanisms. Also, this regression is performed with the intent to add weight to the assertion that individual differences are most important for one-time
measures of spectral power. While menstrual status can affect spectral power in repeated measures, these measures are made within one hour of each other and thus there is no assumption of changing estrogen values. Simple bivariate regression will be performed with baseline HF power as the criterion and estrogen level serving as the predictor. The hypothesis is that a non-significant result will be obtained. That is, in a 90% confidence interval, the $R^2$ value of 0.00 will be contained therein.

Hypothesis one: baseline HF power cannot be regressed on estrogen level at a significance level of 0.10.

**Hypothesis Two**

In keeping with the assumed limited variation within the study sample, and a projected high degree of variance being explained by the regression equation, it is hypothesized that a 95% confidence interval for R2 will contain a value of 0.75, when regressing recovery scores of HF power onto estrogen levels and baseline HF power. A caveat to this assertion is that if the coefficient is substantially higher than 0.75, a confidence interval’s lower limit may indeed surpass this value, which would not serve to reject the hypothesis. A level of 0.75, preferably higher, is needed because the study sample should show a high degree of health and thus robust responses to tasking and recovery. With subsequent groups displaying increased age and potential for pathology, more variance is anticipate that would not be predicted by a regression equation. Therefore, assuming a fall in the predictive efficiency, a relatively high level should initially be attained to warrant further study. Also, because this represents a prediction method, it will only be accepted if proven to be as accurate and at least as cost effective as current measures. Thus, once again, a high level of predictive power should be present in initial regressions with healthy and apparently homogenous samples.
Hypothesis two: when regressing recovery scores on baseline, circadian, and estrogen levels, $R^2$ will meet or exceed 0.75, using a 95% confidence interval.

**Hypothesis Three**

Lastly, to concretely validate the inclusion of both variables, it is hypothesized that a 95% confidence interval for each partial coefficient, that is, for estrogen level and baseline HF power, will not include zero. This result validates the assertion that both predictors contribute significantly to the prediction of recovery.

Hypothesis three: partial coefficients will differ from zero, using a 95% confidence interval.

**Significance of the Study**

Significance is attained because of the current demand for more precise accounting of female biology, improvements in gender-specific knowledge, and improvements in medical treatment targeted at gender-specific issues.
CHAPTER 2
REVIEW OF THE LITERATURE

Cardiac Performance and Estrogen’s Influence

Cardiovascular Disease

Currently, cardiovascular disease (CVD) is the number one cause of mortality in the United States (American Heart Association [AHA], 2001). The widespread effects of CVD have been meticulously chronicled in the ongoing trials that make up the Framingham Study (Thom, Kannel, Silbershatz, & D'Agostino, 1998). The American Heart Association estimates that at least 50,000,000 individuals have hypertension. Hypertension is the most common comorbidity with cardiovascular disease. It is the number one predisposing factor for stroke, and it is especially prevalent within the African-American community. Of the deaths attributable to elevated blood pressure each year in the United States, 87.1% are found within the African-American community (American Heart Association AHA, 2001). Additionally, hypertension in women often goes undiagnosed until end-organ damage has reached an occult level. There are currently few studies that shed light on the genetic and environmental influences of hypertension, particularly primary hypertension, although candidate genes and respective hypotheses are beginning to emerge (Kotchen et al. 2000; Williams et al. 2000). There also is no clear consensus on the underlying physiological mechanisms explaining the etiology and pathophysiology of primary hypertension (Williams et al. 2000). However, what is clear is that certain elements within the autonomic cardiovascular control system are considered to behave pathologically. Although hypotheses have speculated that
hypertension may be linked or caused by certain affective, or temperamental dispositions, mechanistic evidence of such connections has yet to be defined.

Furthermore, cardiovascular disease is by far the most lethal process to women, surpassing all forms of cancer by 400% (Silber, 2003).

An important component of cardiovascular disease is the risk of fatal arrhythmia. Such electrical aberrations are common and signify a deranged myocardial electrical system, an inability of the structural elements to accommodate neural input, or deranged neural inputs to the heart. Since the seminal 1981 publication in *Science* (Akselrod, Gordon, Ubel, Shannon, & Cohen, 1981), spectral analysis of heart rate variability has been a valuable investigative tool. Spectral analysis gives a presentation of parasympathetic and sympathetic balance, and a potential window to the autonomic spectrum as it emerges from the brain stem. Additionally, the more generic analysis of heart rate variability has also shown prognostic value. Both techniques evaluate the amount of variation in the components of cardiac control. It has been established that the more varied the heart rate, the higher the index of cardiovascular health (Stein & Kleiger, 1999). Optimally, there is a constant and varied fluctuation of each autonomic component, which is thought to symbolize dynamic and precise control of cardiac function. The healthier the system, the more varied form of control it displays. Very often, cardiovascular pathology is accompanied by decreases in heart rate variability, or decreases in the power of the respective components (Tsuji et al., 1996).

The pathological characteristics of an individual with cardiovascular disease and/or hypertension are not unique to those two conditions alone. Many researchers have focused on autocrine, endocrine, and brain stem mechanisms. But, many of the so-called
higher brain structures typically associated with cognitive and emotional properties have recently been found to jointly influence autonomic and voluntary function (Paus, 2001; Verberne & Owens, 1998). At this time, there are a multitude of causes for CVD and prime candidates include a defective baroreflex system, overactive sympathetic activity, inappropriate endocrine function, failure to adequately regulate fluid volume, and many others (Alexander, Schlant, & Fuster, 1998). Until recently, little attention has been paid to the relationships among higher brain centers, negative emotions such as anger, and cardiac performance.

**Baroreflex**

The baroreflex, literally, is a "pressure-reflex." Areas within the carotid sinus, aortic arch, and elsewhere serve as pressure monitors that constantly inform the central nervous system about pulsatile pressure changes within the vascular system. This pulsatile pressure is a surrogate for cardiac output. Cardiac output is the main outcome variable for all factors affecting myocardial function; for, it is the maintenance of cardiac output that perpetuates homeostasis within the vascular system. Without cardiac output, a sufficient supply of oxygen and other vital species are not available to the rest of the organism.

There is no discrete algorithm for baroreflex operation. The system must constantly alter its sensitivity, or gain, depending on metabolic and flow requirements. For example, the system does not act to lower cardiac output when extreme exertion is required, such as during a sporting contest or, more primitively, a mortality threat. The receptors forming the forward elements of the baroreflex are not integration centers; and, they themselves do not modulate gains in the reflex. Instead, there is an interpretation and prioritization that occurs in the central circuitry. The specifics of this integration are
unknown at this time. However, what is known is that higher brain centers do influence cardiovascular function. As an example, Buchanon found that electrical stimulation of the anterior cingulate in rabbits and rats caused severe depression of cardiovascular function (Buchanon, & Powell, 1982). This depressor effect was abated or attenuated when the ACC was removed or chemically destroyed. In general, it appears there is a complex relationship between the inputs from baroreceptors and higher brain centers, most likely manifesting as differential coding of brain stem efferents to peripheral structures, such as the heart and blood vessels. Importantly, research to date is equivocal regarding the role of psychological and affective state upon baroreflex function.

Integration of Executive Brain Centers

Regarding the role of the brain centers above the brain stem, such as the anterior cingulate cortex (ACC) and orbitofrontal cortex (OFC), activation in both of these structures, located in the prefrontal cortex, is commonly seen in states of arousal (Rolls, 2002). Research utilizing patients with coronary artery disease has found that arousal states often lead to dysfunctional myocardial performance, and alterations of heart rate variability (Bernardi, Wdowczyk, Szulc, Valenti, Castoldi, & Passino, 2000; Delaney & Brodie 2000; McCraty, Atkinson, Tiller, Rein, & Watkins, 1995). However, to date there have been no studies suggesting that the pathology lies within the higher brain centers. The apparent meaning of these findings is that individuals with cardiac pathology respond to arousing stimuli much the same as unaffected individuals, the only difference being in the magnitude of response seen within brain centers, possibly due to atrophy. Due to organic disease within the myocardium, however, these individuals produce aberrant electrophysiologic profiles, marking the presence of disease. As such, the use of
psychologically relevant stimuli can be seen as a modality to survey the health of the cardiovascular system.

Overall, the literature suggests that negative emotions and negatively biased affect are detrimental to ventricular function, often indicated by a negative impact upon heart rate variability. This is in contrast to the effects of exercise, which generally enhances resting variability (Kramer, Plowey, Beatty, Little, & Waldrop 2000; Pagani, Somers, Furlan, Dell'Orto, Conway, Baselli, et al., 1988).

The question remains: how and why could the executive brain structures be involved in cardiac performance? No single investigation has, or likely will, answer this question.

This thesis will replicate specific parts of previous work, and combine new measures to help shed light on this connection.

Below, three basic areas of interest will be discussed: 1) anatomical connections between the higher brain centers and the brain stem autonomic sites; 2) direct neural connections between the heart and higher brain centers; and, 3) correlational evidence from various sources.

Connections between the ACC, OFC, and the brain stem autonomic centers have been established using neural tracing (Ter Horst, 1999; Ter Horst, Hautvast, De-Jongste, & Korf, 1996; Ter Horst, & Postema, 1997; Ter Horst, Van-den-Brink, Homminga, Hautvast, Rakhorst, Mettenleiter, et al., 1993). Unfortunately, there is limited human data to support this, although some evidence has been accumulated in non-human primates. Interestingly, some of our closest relatives on the genetic family tree display unique features within the ACC and OFC that appear to be shared only with humans (Devinsky
et al. 1995). This leads to the speculation that evolution has shaped these structures and perhaps research with other animals may not be as valuable as once thought to be. However, no clear evidence of such specificity has emerged. Thus, data from other animal species is still considered to be appropriate for this discussion.

Neural tracing techniques are either antegrade or retrograde. They essentially provide a picture of one-way traffic only. Therefore, both techniques have been used extensively, and connections between the brain stem and the ACC, for instance, have been found to be reciprocal. There are no known studies in which connections failed to be observed. For the most part, tracing studies illustrate consistent but sparse connections. However, given the multitude of functions the ACC and OFC are believed to participate in, it is not surprising that connections to any one area would be diffuse. A limitation of neural tracing techniques, and the literature that employs them, is that the spatial resolution is not equivalent to other measures, such as fMRI. This leads to the problem of findings that connections are present, but not being able to pinpoint highly specific locations, such as sub-nuclei known to be vital for specific functions such as respiratory gating or heart rate. In other words, are the connections found in one experiment the same, anatomically, as those in another? Is the data reaffirming, adding to, or subtracting from the previous work? Unfortunately, there is no answer for that question at this time.

The main evidence of direct connections between the heart and the ACC comes from the above-mentioned studies, and summarized by Ter Horst (1999). The main finding from these investigations has been tracing of parasympathetic nerves that connect the left ventricle and the ACC. Additionally, a substance known as tumor necrosis factor-alpha, commonly observed in the tissue of diseased or damaged myocardium, will
produce atrophy of the anterior cingulate if injected into the cerebral ventricular system (Ter Horst, 1999). This effect is identical to the atrophy witnessed in the ACC region after MI and reductions of heart rate variability are evident. The interpretation of such results is as follows: direct neural connections between the heart and higher brain centers known to become activated during affective states provides mechanistic evidence of feedback loops potentially bypassing traditional autonomic centers. Such connections could one day explain much of the variation currently present in the literature evaluating psychological challenge and myocardial function.

Finally, some correlational evidence stems not only from the above material, but also from the results of others. It has been detailed elsewhere that depression following MI or stroke is common and that the ACC and OFC are often involved (Aben, Verhey, Honig, Lodder, Lousberg, & Maes 2001; Strik, Honig, Maes, 2001). Thus, damage to peripheral or central structures may play a role in the onset of depression by strictly physiological relationships. There are also purely psychological causes for depression, such as the decrease in function and overall health, the threat of mortality, and others. However, the biological evidence is significant. Musselman (1998) has detailed several mechanisms in which depression and systemic illness are related. In that work, there are clear ties to the anterior cingulate and cardiovascular pathology. Again, though the causation is not established, the correlational evidence certainly provides a rationale for further investigation.

**Heart Rate Variability**

Since Einthoven began the inquiry into myoelectrical relationships, the attention paid to the electrical activity of the heart has remained a constant in applied, medical, and basic science literature. The nature of the electrical rhythm is of prime importance when
determining the adequacy of the heart's function. In the early 1980's, it was discovered that the methods of spectral analysis could be employed to survey the relative influence of the sympathetic and parasympathetic arms of the autonomic nervous system (Akselrod et al., 1981). Since that time, spectral analysis has been used extensively. Though there is still some debate regarding the exact significance of the spectral parameters, some facets have gained wide support. Overall, the assessment of vagal (parasympathetic) input is considered to be a reliable measure. However, documenting sympathetic fluctuations is more difficult. Because exaggerated activity of the sympathetic nervous system is often associated with pathological conditions, it is the absence of parasympathetic influence that is most often sought. Generally, though spectral analysis of the heart rate signal has been the topic of much debate by experts in the field of psychophysiology; its efficacy as a prognostic tool has been established and is without doubt (Maliani, Pagani, Lombardi, & Cerutti, 1992). At this time, using spectral power provides global information but does not give highly detailed results that can be stratified into minute categories of causal influence. Despite this limitation, spectral analysis is a powerful and legitimate research tool.

**What is Being Measured with Spectral Methodology**

Essentially, spectral methods are used to find two general components, high frequency and low frequency. The high frequency component is derived primarily from parasympathetic influence, while the low frequency component is derived from a combination of sympathetic and parasympathetic input. The high frequency range is between 0.15 Hz and 0.40 Hz. The low frequency range is from 0.03 Hz to 0.15 Hz. There is no current consensus on the meaning of the spectrum below 0.03 Hz.
One other measure derived from the frequency spectrum is overall variability. The total power found in the low and high ranges represents the global variability of heart rate. It has been known for many years that the autonomic nervous system controls cardiovascular function in a very dynamic way. Thus, a normal, healthy individual will display variability in his or her interbeat intervals. A person with cardiovascular pathology is likely to show a decrease in variability. These decreases in variability inherent to cardiovascular disease have been shown to be prognostic indicators of mortality (Foreman, 1999; Malliani et al., 1992).

In the end, spectral components are sought in order to gain insight into the differential nature of cardiac variability and autonomic functioning. It does not provide us with a tool that can supplant all others. As with most other research tools, spectral analysis is only one of many tools available to the researcher. Given its utility and potential in future research endeavors, it is a worthwhile measure.

**How Else is Variability Measured**

There are several measures of variability derived from the raw EKG record. The most commonly used calculation is the standard deviation of the inter-beat intervals. The norm for such a measurement is $139 \pm 40$ ms (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). Studies utilizing patients with cardiovascular disease have documented a standard deviation as low as 67 ms, and such a reduction in variability has been highly correlated with increases in mortality. One caveat to the use of spectral components and standard deviation calculation is the length of the record used. According to standardized guidelines, five minutes is the minimum reliable measurement time period necessary. Thus, all periods of measurement in this thesis will consist of five minute intervals.
How do we Differentiate Heart Rate Components at Rest and During Various States of Arousal

There is no patent answer for this question. Naturally, one would believe that there is an affinity for more sympathetic influence during arousal, which is the case. However, several studies have shown that total spectral power will decline during exercise, with a roughly linear decline with increasing workload. This is not always the case, as explained below. It would be expected that the body would decrease parasympathetic tone at the onset of exercise, thus lowering the total power within the spectrum; however, it would also be expected that total power would not continue to decline with the increase in workload, because, at some point there would be a call for increased sympathetic tone to step up the demand for performance. To date, there are no studies that outline theoretical, physiological, or mathematical explanations for this occurrence. Again, the emphasis goes back to the relative presence or disappearance of parasympathetic influence as indexed by the high frequency band. There is no reason to believe that spectral components are subject to the ceiling effects that may occur as the heart rate increases, though such characteristics are not impossible and have not been refuted by published data.

In contrast, a ceiling effect can be seen if simple interbeat standard deviations are used. As the intervals decrease, the standard deviation will decrease, yet this does not imply a lack of regulation by the autonomic nervous system. Because the parasympathetic system acts at a relatively fast rate, its influence can be present yet largely unnoticed by the interbeat interval standard deviation. There are no mathematical methods capable of weighting the intervals against absolute length, and making
comparisons to standard deviations at a particular interval length. Though this would appear possible, it is not feasible because of individual variability.

While heart rate variability studies have been extensively carried out with individuals at rest, exercise studies are less abundant and less clear. There have been minor differences documented in various types of exercise, such as static and dynamic (Gonzalez et al., 2000). Essentially, a more varied heart rate and a larger representation of high frequency components are seen with dynamic exercise, given the workload is not very high. The main inference from such findings is that a varied and relatively constant activity pattern produces the most variability and plasticity of autonomic function, generating more interbeat variability and increases in total spectral power. This clearly lends support to the suggestion that regular exercise is beneficial for one's health.

**Do Mental Stress and Affective State Show Unique Effects on Heart Rate Components**

This is an easy topic, as the overwhelming evidence points to a pattern of increased sympathetic influence with psychological stress. The main paradigms for this type of investigation are the use of interference tasks, such as the Stroop test, and emotion-producing situations such as those designed to bring about anger or anxiety. These studies have clearly shown that psychological stress increases sympathetic influence, and decreases overall variability and high frequency spectral power.

It is important at this point to acknowledge that the literature is quite varied. In addition to the normal explanations for high variability of findings, such as differences in methodology, study populations, statistical analyses, and interpretations, a constant remains. This constant is inter-individual variability. Obviously, individuals vary in their respective abilities and propensities to appraise and react in accord with a given pattern.
So prevalent is this variability, that, concrete patterns of reactivity to specific stressors have yet to be identified. Overall, as with most forms of arousal states, psychological challenge brings about an increase in the sympathetic component with a coincident decrease in the parasympathetic component. Placing quantitative, normalized boundaries on such changes have not been possible to this point.

A mentioned previously, in 1981, Akselrod published what appears to be the most often cited article regarding spectral analysis of heart rate variability. Briefly, a researcher records the basic heart rhythm using conventional electrocardiography techniques. Next, a graph is formed comparing the interval between each R wave against each respective beat, or absolute amount of time. This graph tends to have a chaotic appearance, not resembling conventional wavelet graphs. A subsequent Fast Fourier Transform is performed and the resulting transform graph is analyzed from 0.0 Hz to 0.50 Hz.

Recently, in 1995, a consensus report was issued regarding standards for the assessment of heart rate variability (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). Within this investigation, the recommendations by that task force are followed as closely as possible. Because the Akselrod article is so highly cited, a brief description of the article will be presented below, in order to acquaint the reader with concepts relevant to heart rate variability measurement.

Using conscious dogs as models, Akselrod pharmacologically blocked the parasympathetic and sympathetic neural discharges to the heart. It was found that the area under the lower frequencies was increased at any time the sympathetic or parasympathetic influence was increased. However, during parasympathetic blockade,
the diminution of higher frequencies was much greater than that of the lower frequencies. The main finding, and current consensus, is that the higher frequencies are sign of parasympathetic modulation, whereas the lower frequencies reflect both sympathetic and parasympathetic influences. Using the recommendations of the task force report, the lower frequency is defined as 0.03 Hz to 0.15 Hz; the higher frequency as 0.15 Hz to 0.40 Hz. When assessing heart rate variability, the overall power in each spectral component is a reflection of total variability. It is often most appropriate to express the area under each frequency component in normalized units, which demonstrates the total power in each band as compared to the total power contained within the spectrum from 0.03 Hz to 0.40 Hz. Another index commonly used is the low frequency to high frequency ratio (LF/HF). Most researchers use this a measure of sympathovagal balance. This is particularly useful when comparing acute changes over short time intervals.

In addition to spectral analysis, a simpler technique is the analysis of the standard deviation in R-R interval. Investigators may use this measurement as a global assessment of cardiac variability, and still attain prognostic power in clinical environments. Of course, this simple measure of variability does not provide mechanistic analysis, which is where the power of spectral analysis is seen.

Lastly, a few caveats of spectral analysis should be presented. First, as described above regarding the Akselrod article, high frequency influence appears to be derived entirely from parasympathetic discharge. But, the consensus is not so clear regarding the lower frequencies. The most acceptable position at this time is that the combined influences of sympathetic and parasympathetic discharges are concerned with the total area under the low frequency spectrum. Because there are no established algorithms to
divide the relative influence of each discharge pattern within the low frequency range, a
stable interpretation of the low frequency peak is problematic. Finally, the area of the
graph from 0.00 Hz to 0.03 Hz has eluded even the most rudimentary of consensus
statements. It is felt that this area may be simply related to sympathetic skin activity,
endocrine substances passing through the blood and thus into the heart tissue, or various
other minute changes within the internal milieu that may alter electrical conduction and
propagation. Overall, in the simplest analysis, spectral analysis serves as a powerful tool
for the well-tuned investigator capable of integrating many facets of physiological input,
in order to assess the relative importance of differing psychological sets, physical activity
patterns, and pharmacological manipulations.

Given the aforementioned evidence of parasympathetic ties between higher brain
centers and autonomic sites, as well as the known role of higher centers in cognition and
emotion, spectral analysis during affective challenge shows great promise. Coupling this
to empirical evidence showing decreases in high frequency power during psychological
arousal, and spectral analysis’ ability to index parasympathetic tone via the high
frequency component, there is a natural fit between the use of spectral methodology and
psychological tasking in the evaluation of autonomic cardiac regulation.

**Estrogen’s Impact on Cardiac Performance**

Overall, the main thrust of estrogen’s influence is to increase parasympathetic-
effect within the cardiovascular system. In the subsections below, these effects will be
discussed regarding cardiac electrophysiology and arrhythmia patterns, myocardial
structural characteristics, coronary health, cardiovascular response to challenge, and heart
rate variability. The summary section will detail the composite picture and why the
recovery of parasympathetic-effect is affected by estrogen’s influence.
Electrophysiology and Arrhythmia Patterns

The conduction of electrical impulses within the heart is a complex order of events dependent on temporal, structural, and electrochemical events, all of which must be synchronized to some degree.

Definitive research in human subjects is sparse, yet animal studies have found that estrogen can induce hyperpolarization in the canine heart (Harder & Coulson, 1979). Such an effect can potentially lead to a decrease in susceptibility to sympathetic influence, leading to the previously described parasympathetic-effect. Additionally, estrogen has been found to change the expression of proteins constituting specific ion channels, producing the same parasympathetic-effect (Tolbert & Oparil, 2001). Importantly, changes in electrophysiological measures and protein expression have been witnessed within three hours after administration of exogenous estrogen (Boyle, MacLusky, Naftolin, Kaczmarek, 1987). This provides evidences that estrogen generates genomic and non-genomic effects that alter electrophysiological properties.

The above-mentioned effects of estrogen lead to parasympathetic effects that overpower or stunt simultaneous sympathetic outflow. There is evidence that women display greater sympathetic tone to the peripheral vascular system at rest, as compared to men, which would seem to contradict the above-mentioned parasympathetic effect (Dart & Kingwell, 2001; Cooke, Creager, Osmundson & Shepherd, 1990). However, this has not been correlated with estrogen levels and direct cardiac effects have not been witnessed. This leads to speculation that such findings are spurious, or that differential allotment of autonomic tone to the periphery is possible. In other words, men and women may sustain chronic levels of tone in differing proportions to different regions of the body. For instance, men may generate higher sympathetic tone to the lower extremities,
and smaller amounts to the upper extremities, in comparison with females. Such findings have not been definitively shown within the same study, and certainly have not been correlated with varying levels of estrogen. Reasons for such disparate findings may be the differences in body mass, body fat distribution, and muscle mass differences between the genders and between upper and lower extremities. How these types of results relate to recovery measures following cardiovascular challenges is unknown. Given the propensity to bring about parasympathetic effects in studies at the myocardial level, the logical assumption is that rising estrogen levels mediate a stronger return of vagal power in the autonomic spectrum, when evaluated at the heart.

Evidence for these proposed recovery effects was found in large scale studies that found women more resistant to overdrive suppression of pacemaker activity within the heart (Kadish, 2001). In other words, the hearts of females showed less susceptibility to abandoning their normal rhythm for a faster, sympathetically dominated rhythm. Unfortunately, these finding were not correlated with estrogen level. However, they did persist after controlling for body mass and other parameters measured. This leads to speculation that estrogen had a role in the observed effect, but no mechanistic ties were established.

Data applying to arrhythmia susceptibility and subsequent mortality is equivocal as well. Little has been firmly established regarding estrogen’s effects. For instance, fewer women develop atrial fibrillation than men; but of those that do, a greater percentage subsequently die from complications (Steing & Kleiger, 1999). Unfortunately, no causal links between hormonal levels or disease patterns can explain such a discrepancy. The only known link between estrogen and arrhythmia production is found in paroxysmal
supraventricular tachycardia (PSVT), a fast-paced arrhythmia, most often benign. A study by Rosano and his colleagues found that increasing levels of estrogen were found to correlate with decreasing episodes of PSVT (Rosano, Leonardo, Sarrel, Beale, De Luca & Collins, 1996). Given that PSVT is a sympathetically dominant rhythm, and that increasing estrogen levels produced a decrease in its incidence, the most logical causal explanation is that estrogen serves to blunt sympathetic activity.

Lastly, further evidence of estrogen’s parasympathetic-effect comes from early studies describing the effect on neurotransmitter release and disposition. As reported by Larsen and Kadish (1998), work from Ball and his colleagues found that estrogen enhances the degradation of sympathetically associated neurotransmitters. This is important because if not degraded, these chemicals will continue to stimulate the conduction system and enhance sympathetic influence. Additionally, rising estrogen levels have been found to attenuate release of these neurotransmitters (Hamlet, Rorie & Tyce, 1980). Combined, these effects may explain the ability of the sinus node (the heart’s pacemaker) to resist overdrive from sympathetic tone, and assist in the recovery of parasympathetic influence following cardiovascular challenge.

One last research finding connects myocardial structure, coronary health, and electrophysiological function. Recently, much attention has been paid to vasoactive substances that can also modulate neural activity, such as angiotensin-II (Staniloae et al., 2003) This substance causes profound vasoconstriction and increases the potentiality of sympathetic neural transduction events. This and other actions make angiotensin-II deleterious for coronary health, myocardial structure, and rhythm homeostasis. It has been found that estrogen can suppress the activity of the key enzyme responsible for
production of angiotensin-II. Thus, within electrophysiological analyses, estrogen would again enhance parasympathetic-effect. Other aspects concerning structural and vascular health will be addressed below.

In sum, if individual variability proves to be the dominant factor in determining the level of parasympathetic withdrawal and sympathetic activation following psychological challenge, then estrogen levels may prove to be the deciding factor in the extent of return for parasympathetic influence. Key to the argument is the role that structural integrity of the myocardium plays in conveying the autonomic signals, which will be discussed below. If predictable influences of estrogen levels are indeed present, then deviation from normal behavior could be a sensitive marker for the presence of cardiac disease.

Though a discussion of heart rate variability would easily fit within a section on electrophysiology, it will be withheld for later discussion because the parasympathetic component of heart rate variability will constitute the criterion variable in this thesis.

**Myocardial Structure**

In recent years, several authors have brought to light the importance of myocardial structure for cardiac performance and health (Lab & Dean, 1991; Horner et al., 1996). These authors highlighted several pertinent facts regarding estrogen’s effects on the myocardium. In general, estrogen has tonic effects on the heart and its structure, which make the tissue more resistant to fibrous deposits, the development of hypertrophy, and deleterious effects of wall stress. This is evidenced by reduced cardiac mass and patterns of hypertrophy that are different between males and females (Luchner et al., 2002; Xu et al., 2003).

Epidemiological data from the Framingham study have shown that females are more resistant to the development of cardiac hypertrophy, when compared to their age-
matched controls (Levy et al., 1988). Furthermore, these differences stand when accounting for body surface area, a measure related to the overall amount of work the heart must perform. Some research has shown that estrogen can attenuate the hypertrophic response to either volume overload or pressure overload, the two most common causes for cardiac hypertrophy (Weinberg, Thienelt, & Katz, 1999). The exact mechanism for such a relationship has not been discovered. Possible explanations range from the inhibition of fibroblast and myocyte proliferation to alterations in pulse waves reflected back to the heart from the periphery which may set up turbulent flow patterns and establish a mechanical stimulus for such growth and proliferation (Dart & Kingwell, 2001).

To date, a rigorous accounting for estrogen levels and hypertrophic status has not been accomplished. In other words, researchers have not charted estrogen levels and cardiac mass longitudinally. Whether this represents a pitfall in the analysis of estrogen’s structural effects is up to interpretation. It seems highly unlikely the general notion of a protective effect concerning hypertrophy is unwarranted. This is due to the fact that epidemiological evidence from sources such as the Framingham Study, have documented reduced numbers of hypertrophy in females during the middle of their reproductive years. This is a time when the menstrual patterns and estrogen levels should be at their most consistent during a women’s lifetime.

Of course, the future is likely to add more information and refine the relationship; yet, it is suffice at the moment to say that estrogen is protective against unnecessary hypertrophy of the myocardium.
An important point is that myocardial structure and mass is an important prognostic indicator of future cardiovascular events (Levy et al., 1988). This is part of the reason that ultrasound is such an important asset for the cardiologist, as it is able to provide indices of cardiac mass and size. However, ultrasound studies are often subjectively graded and always expensive. In addition, cardiac structural properties may affect heart rate variability measures (MacFadyen, Barr & Struthers, 1997). These structural properties are not visible with ultrasound. Therefore, the presence of structural abnormalities may be sensitively detected by HRV analysis.

**Coronary Health**

Coronary artery disease can be roughly characterized as a process by which the lining of the coronary blood vessels is infiltrated by oxidized lipids with subsequent reactions that lead to plaque formation and eventual blockage of the vessels. This in turn can lead to starving the downstream myocardium of blood flow, which produces an infarction. An important event in this process is the oxidative damage to lipids. Estrogen has been shown in numerous studies to have antioxidant properties (Shwaery, Vita & Keaney, 1997). This effect is thought to be largely responsible for the differences in CAD between males and their age-matched pre-menopausal female counterparts (Tolber & Oparil, 2001).

More recent studies have shown important effects of estrogen’s presence on the activity of angiotensin converting enzyme. The blunting of this enzyme’s activity is observed in the presence of estrogen (Staniloae et al., 2003). This enzyme catalyzes a reaction that produces angiotensin-II (ANGII), a potent vasoconstrictor and proliferative agent in the coronary circulation. ANGII has repeatedly been found to increase vascular tone, reduce the health of the heart’s vessels, cause fibrotic depositions, and instigated the
growth of smooth muscle cells within the coronary vasculature, greatly reducing the luminal cross sectional area that blood that can traverse, subsequently starving the downstream tissue.

Though less publicized, coronary tone and rigidity is important for the structural rigidity of the heart. If vascular tone is increased, there is in effect a tense meshwork of vessels that make the heart a more rigid vessel. This increase in rigidity forces the heart to work harder to eject the same amount of blood. Myocardial oxygen consumption is directly proportional to the workload. Thus, an increase in workload brings about a need for increased oxygen supply. However, if the vessels are diseased, an increase in oxygen supply via the blood is not possible. This sets up a spiraling cascade whereby the heart works harder to accomplish its mission, demands greater blood flow, cannot achieve this demand, vessels become more diseased and more constrictive, and eventual infarction ensues. These same relationships between heart rigidity, workload, oxygen demand, and an ensuing cascade of failure also apply to defects that originate in the myocardium, outside of the coronary vasculature.

Another important aspect of estrogen’s function in the coronary vasculature is its ability to induce relaxation in the smooth muscle cells that brings about a reduction in tone, primarily due to its liberation of nitric oxide, a well known vasodilator (Mendelsohn, 2002; Muller-Delp et al., 2003). This makes vessels more compliant and accommodating to flow. This reduction in tone also produces a more yielding heart structure, and a reduction in workload, as outlined above. An added bonus to reductions in vessel tone is that blood flow is increased to the more distal segments in the vascular tree.
Lastly, it has recently been uncovered that the family of steroid hormones, to which estrogen belongs, may create fast-acting alterations in physiology, not mediated by genomic activation (Hodgin & Maeda, 2002). This is news because steroid hormones generally attach to receptors at a cell’s nucleus, which in turn creates a sequence of events within the nucleus aimed at changes in gene expression. Such activities take many hours to begin showing results. However, if gene regulation is not the end response, then effects may be witnessed in minutes. Some of these non-genomic effects, as they are called, are linked to calcium channel activity. Most of the reported effects are thought to produce a decrease in vascular tone, further increasing estrogen’s potential as a vaso-regulative substance.

In sum, estrogen produces a more compliant vascular system, which when combined with estrogen’s effects at the myocardial level, creates an environment that is more compliant, suitable for blood flow and movement, with a subsequent reduction in workload. These effects are cardio-protective. Given that heart rate variability is influenced by structural properties, attending to estrogen’s structurally mediated effects is worthwhile. This is due to the possibility that non-genomic actions may one day prove influential in day-to-day structural properties that may in turn affect such measures as heart rate variability.

**Cardiovascular Response to Challenge**

The literature contains data regarding the exercise response of females, largely ascribing any female-specific characteristics to the presence of estrogen. Women respond to exercise training in the same manner as men. Women are able to build muscle, run marathons, and the same principles of training that apply to men also apply to women. The primary reason women do not add the same muscle mass is due to testosterone.
There is no data to indicate that estrogen is catabolic to muscle tissue or exercise benefits. However, there are nuances to the acute exercise response that should be addressed.

Treadmill testing is a less sensitive marker for disease presence in the female population, compared to the male population (Weiner et al., 1995). Several reasons have been postulated, including the disparity in the numbers of women that regularly participate in strenuous exercise, as compared to men. This is thought to lead to more women being unable to truly perform at a maximal effort level, which lowers the sensitivity of a maximal test, for obvious reasons. Also, because of the considerations regarding activity levels, it is thought that exercise testing may lead to greater anxiety in women than in men, which can alter baseline and testing EKG readings (Hlatky et al., 1984). These alterations are due to catecholamine responses. Catecholamines are released from the adrenal glands, and from there freely circulate throughout the blood where they affect virtually every tissue in the body. Included in this distribution is the heart. The catecholamine response may affect peripheral vascular tone, coronary tone, and heart rate dynamics. The end result is an increase in the variability of critical EKG components necessary for diagnosis. High variation in some EKG parameters often leads to false-positive results (Myers, 1996). Fortunately, the differences in response to challenge appear to be limited to exercise testing more so than psychological testing. However, because exercise testing has been the dominant paradigm for measuring response to challenge within the CV system, there has been less credibility in disease prognostication for the female population as a whole. Additionally, it has not been determined whether varying estrogen levels need to be accounted for during exercise testing, which may also lower the power of these tests.
Heart Rate Variability

Serial measures in females have provided some debate regarding the effect of estrogen upon heart rate variability measures. In reviewing the literature, the difference appears to be due to inconsistencies in measurement methodology and a failure to fully account for day-to-day fluctuations during the normal menstrual cycle. Also, many studies have not taken circadian variation into account. While circadian variation does not appear to affect response to, or recovery from psychological challenge, it has been shown to affect resting measures (Evans et al., 2001).

Regarding the division of the menstrual cycle into phases, as was described in the introduction, the menstrual cycle produces fluctuations in estrogen levels that cannot be described by any biphasic criteria. For instance, some authors have decided to divide the menstrual cycle into luteal and follicular phases. In addition, differing criteria for the delineation of each phase have been utilized. Lastly, in most studies, small samples have been utilized.

Summary

This investigation seeks to model the return of parasympathetic-effect following an anger recall task. As such, a case must be made for the use of a recovery measure as a criterion variable. In physiological interventions, recovery of baseline levels is often used as a measure of fitness. Intuitively, the use of recovery measures seems to be an efficient variable, for if there is a high degree of health and fitness, it follows that a person would recover from a challenge much faster. However, the use of such measures has been complicated and the effectiveness reduced in the female population. Reasons for this are similar to those expressed above for differences in the genders in their respective exercise response. Discrepant effort levels, anxiety associated with testing protocols, and
differences between true fitness and activity-specific fitness all affect the usefulness of physiological testing, in particular, recovery measures. Residual catecholamine levels may affect cardiac rate, and spectral parameters. As detailed above, estrogen has been linked to the suppression of catecholamine release and faster degradation rates after release. This effect can dramatically alter the electrophysiologic profile of cardiac activity, as catecholamines may spur on early depolarization and a sympathetic effect on the EKG. This blunting effect by estrogen may prevent the adequate characterization of responses, and bias the measurement of recovery in a positive or negative direction, pending the level of estrogen present.

The above array of facts still does not provide the clearest rationale for the use of recovery measures. This detail lies mainly in the temporal constraints of exercise testing and psychological testing. Typically, an exercise testing session lasts more than an hour, and given the differing anxiety levels that may ensue over such a period, it is logical that a confounded picture will result. However, psychological testing can be as short as 15 minutes, and should be no longer than 30 minutes. More importantly, it involves tasks that are ecologically valid, meaning that the interventions are common to all humans. Not all testing participants are familiar with exercise, yet all would be expected to be familiar with anger, and other psychological constructs. As such, it follows that psychological testing may be more valid than physiological testing.

Lastly, the specific use of a recovery measure stems from the fact that there is not strong consensus in how to model normal responses during psychological testing. At this time, conflicting data exist on the response characteristics for differing tasks (Lovallo, 1997).
Affective States, the Brain, and the Heart

Studies of emotions and cardiovascular function have remained relatively separate endeavors. While many volumes have been written concerning the constructs of various emotions, the most relevant emotions for cardiovascular function appear to negatively valenced ones, such as anger, anxiety, and depression. Debating the exact nature of emotion, if and how anger, anxiety, and depression are separate emotions, and other questions are beyond the scope of this thesis. Fortunately, because the above three affective states are generally understandable in layman’s terms, intricately studying and defining each for study populations has not been required.

In the paragraphs below, empirical evidence of affective modulation of two executive centers will be discussed. Also, a brief description of executive brain involvement with heart health and performance will be provided.

Descriptions of the executive brain generally relate to the frontal and prefrontal areas. There are two areas particularly important in the executive brain that have recently come to light for their global involvement in human biology. These areas may function during cognitive, or emotional, or autonomic functions: separately or simultaneously. In addition, one of these two structures, the anterior cingulate cortex, has been shown to have direct myocardial connections. Another structure, the orbitofrontal cortex, is closely allied with the anterior cingulate cortex. Due to the limited scope of many studies, it is commonly believed that the orbitofrontal cortex is most often associated with reward. However, it is involved in emotion on a global level. There is less evidence linking the OFC to cardiovascular function, but where appropriate, pertinent facts will be discussed. Consequently, the majority of the review on brain structures will center on the ACC.
**Orbitofrontal Cortex**

The OFC is part of the primate prefrontal cortex, receiving inputs from all senses, and is part of the global response to emotionally relevant stimuli (Rolls, 2001). In humans, damage to the OFC has been seen to remove the ability to express emotion (Damasio, 1994). Also, demonstrated connections to autonomic sites in the brain stem have been shown to exist. Unfortunately, these studies did not evaluate the OFC’s influence on autonomic function, merely demonstrating anatomical connections (Nauta, 1972).

The OFC is most commonly associated with reward because in primate studies, its activation is accompanied with stimuli prompting the primates to pursue various rewards such as food. However, more generally, it is the salience of consequence that the OFC is involved in. Therefore, the perception of consequence to action, if it generates emotion, will produce activation in the OFC (Rolls p.114, 2001). Relating to higher brain functions, the OFC is particularly prominent in reinforcement behavior. Again, this has been mostly tied to reward, and obsessive disorders, such as drug addiction. Though this role is pertinent, again recalling that the attachment of emotion to consequence is responsible for activation of the OFC, along with its connections to autonomic centers, it is wise to speculate that it may also prove to have a role in cardiovascular dynamics in response to or recovery from psychological challenge.

Material on the OFC has been included largely based on speculation of its future relevance to cardiac psychology. However, given the evidence that emotions profoundly affect cardiac behavior, and the OFC’s connection to emotion and autonomic sites, referencing is warranted. Additionally, research into psychological manipulation of cardiovascular function is bound to increase dramatically in the near future; and, as such,
alerting and calling attention to potentially relevant structures should prove most useful to those interested readers.

**Anterior Cingulate Cortex**

The anterior cingulate cortex is commonly grouped into the limbic system, although the concept of the limbic system has undergone quite a bit of revision since the early work of Papez in the 1930's. Often times, the entire cingulate gyrus is referred to in the literature. The dorsal, or posterior, segment is considered to be important in visuospatial and memory tasks, but relatively uninvolved in matters that concern affect. The anterior segment is implicated in affective characteristics as well as motor activity. Conceptually speaking, the combination of affective and motor capabilities suits the perspective of an evolutionary-directed stratification of functions. If the theories of Panksepp and others (Panksepp, 1998), regarding a hard-wired emotional system, shaped by evolution and the needs for survival and procreation are correct, then combining affective and motor actions is logical. Part of the affective judgment system corresponds with fear, which is in response to threatening stimuli. If there are clear anatomical and functional associations between the perception, processing, and acting upon fearful stimuli, then survival chances would be improved due to enhancing processing efficiency.

The human cortex is, in part, functionally described in terms of its layers. These layers often show homogenous anatomical and physiological relationships between separate areas and nuclei. Within the ACC, layer five is especially well developed and has extensive connections to motor areas within the brain. According to Devinsky (1995), the anterior cingulate separates itself from other structures by virtue of the numerous afferents it receives from the thalamus. The thalamus receives input from virtually every
sensory modality, making it a hub for the reception and distribution of sensory information to other sites in the central nervous system. The efferents sent from the thalamus to the ACC provide a vast array of sensory information. While there seems to be little debate about that fact, there is much debate regarding what actually happens in the ACC, how it happens, how it may be modulated by other structures, and what specific efferent impulses are then sent from the ACC to other sites. This has made the anterior cingulate a popular target of recent investigations. However, an important point to remember is that the anterior cingulate is only one of many structures in the central nervous system that affects human biology. The similarities among findings regarding ACC activation merely imply correlational ties, and not causal influences. Unfortunately, at this point, the status of anterior cingulate research is more precisely categorized as descriptive rather than mechanistic.

In the following, an attempt has been made to reference the most current literature, given the rapid advances in imaging, cell biology, and other methodological issues. Where necessary, more dated research will be presented.

**Connections to Viscero-and Skeletomotor Areas**

Prior to addressing direct motor connections, it is important to stress that the ACC shares neural pathways with other areas involved in motor selection and control, such as the thalamus (Devinsky et al. 1995). These connections lend support to the integrative duties of the cingulate, given prior research implicating the various thalamic nuclei in subsequent motor response selections (LeDoux 2001).

It is believed that Vogt was the first to label the anterior portion of the brain housing the ACC as the anterior executive region (Devinsky et al., 1995). This name was given to account for the important role of the ACC in planning and executing. Upon
further review, the ACC can be divided into two other divisions, theoretically rather than
anatomically or physiologically based, with one region labeled as affective and the other
being cognitive (Devinsky et al., 1995). Alerting, defending, and fleeing are all aspects of
instantaneous emotional reactions (LeDoux, 2001) that lead to motor action. Most of
these actions are mediated by the amygdala, its subnuclei, and the thalamus (LeDoux
2001). Importantly, the ACC receives input from the amygdala (Paus, 2001) and the
thalamus (Paus, 2001). Also, these emotional reactions take on autonomic-effector status
as well, again substantiating support for the involvement of the ACC.

The cognitive region of the ACC does not appear to be connected to limbic
structures. However, there is mild support for its connection to motor areas, as suggested
by Yeterian & Van Hoesen (1978) and Baleydier & Mauguiere (1980). Again, the
distinction between a cognitive and affective region is theoretical only, and not yet
substantiated in the literature. The truly elusive aspect lies in the potential ability to
control each region, and prevent unnecessary overlap and intrusion of one region into that
of another.

Relating to autonomic functioning within the motor and muscle spectrums, animal
studies have shown that the ACC sends projections to the solitary tract within the brain
stem's parasympathetic region (Terreberry & Neafsey 1983), and dorsal motor nucleus of
the vagus (Hurley, 1991). There are also suggestions of connections to brain stem regions
associated with sympathetic activation (Hurley 1991), though these findings appear to be
less common than those connections to the parasympathetic regions. In addition,
myocardial connections have been found by Ter Horst several times, illuminating direct
parasympathetic connections between the heart and the ACC (Ter Horst, 1999).
Functional evidence of such connections is provided by Kaada, showing depressor responses to electrical stimulation in the anterior cingulate region (Kaada 1949; Kaada 1951). Though Kaada's work is quite dated, its simplicity makes it a valuable resource. Essentially, Kaada showed that blood pressure and heart rate were greatly depressed when the anterior cingulate region was electrically stimulated.

Though the above findings illustrate a predominant association with parasympathetic and vagus centers, some studies suggest that emotional stress in animals produced activity in the anterior cingulate (Neafsey, 1993). However, some of these studies, such as the one cited above, used footshock or other painful conditioning stimuli; and, nociceptive stimuli are transmitted to the ACC (Foreman, 1999). For instance it is known that the ACC is an integral structure in the processing of angina (Foreman, 1999). Interestingly, these animal models of aversive conditioning show depressed cardiorespiratory stimulation during presentation of the conditioned stimulus once the ACC has been surgically lesioned (Frysztak & Neafsey 1991). Given the conditioning model, and the fact that presentation of the conditioned stimulus provides a means for anticipatory anxiety, evidence has accumulated supporting the role of the ACC in emotion and anticipation (Chua, Krams, Toni, Passingham, & Dolan, 1999). Overall, the literature outlines an influence upon voluntary and involuntary motor actions by the ACC.

When evaluating the ACC's influence upon the skeletomotor system, some interesting findings have been presented using the nonhuman primate. In monkeys, Biber et al. (1978) found corticospinal projections emanating from the anterior cingulate region and terminating in the cervical and lumbar areas of the spinal cord. This would suggest
that direct stimulation of the anterior cingulate would produce movement. Unfortunately, studies conducted during the past ten years were not located to substantiate this finding. However, work dating back over 30 years provides credible evidence for such an effect. Relying on the interpretations of experts in the field, findings from both monkeys and man have shown this to be the case. In monkeys, Showers (1959) and Hughes & Mazurowski (1962) found movements to be elicited when the anterior cingulate region was electrically stimulated. This occurred in conscious and anesthetized animals. Unfortunately, most findings from studies of that period cannot be reliably applied to autonomic functioning because of the anesthetic protocols used, since they often altered autonomic parameters. However, the finding of skeletal muscle activation upon stimulation serves to further verify the significance of neural connections found by previous investigators (Morecraft, & Van Hoesen, 1993).

In humans, stimulation of the dorsal region of the anterior cingulate produced movement. This region of the anterior cingulate has also been termed the supplementary motor area (Paus, 2001).

One shortcoming of the above studies is that they do not determine if the ACC is responsible for the processing of information related to movement or if it is simply a point through which neurological signals may pass. Thus, the studies do not determine if consciousness and intention is necessary for the anterior cingulate activation. This question has been at least partially answered by several studies indicating that different regions within the ACC are activated depending on the task evaluated (Pardo, Pardo, Janer, & Raichle, 1990). Furthermore, the requirement for conscious decision processes and intention come from Paus, Petrides, Evans, & Meyer (1993), where it was shown that
a conscious decision making process that was stratified among different task conditions, including motor and cognitive, produce activation in the ACC. Also, Pardo et al. (1990) found activation when performing a Stroop task. The Stroop results are grouped in with motor activity because of the motor control necessary for speech. As a side note, the finding of spontaneous vocalizations resulting from electrical stimulation of the cingulate region has been repeatedly demonstrated in monkeys (Jurgens, & Plog, 1970; Muller-Preuss, Newman, & Jurgens, 1980).

**Connections to Limbic Structures**

Given the high density of dopamine receptors within the anterior cingulate, it is not surprising that it shares circuits with many structures, including those germane to the limbic system (Paus, 2001). A more extensive view of these connections will be presented below. To provide a cursory view here, the ACC receives afferents from the thalamus, ventral striatum, brain stem, amygdala, and all parts of the frontal areas.

The significance of its dopamine receptors lies in the prevalence of dopamine signaling within the limbic system. Dopamine release and reception is a strategic maneuver designed to increase the impact of stimulus perception. Dopamine will decrease the excitatory threshold for many limbic structures (Schultz, Tremblay, & Hollerman 2000; Waelti, Dickinson, & Schultz, 2001), thus enabling a potentiation of effect. Another consequence of potentiation is the subsequent cellular changes that result. The changes often include intracellular signaling that alters gene expression and organelle production, among other facets of cell biology. These changes affect a cell's ability to respond, and shape subsequent responses, thus altering the nature of input-output relationships. Dopamine signaling and reception provides notice that a significant
event has or has not occurred and provides the foundation for changes at the cellular
level, which may translate into changes in gross biological functioning.

The main point for this subsection is that affective relationships are seen at cellular
and gross anatomical levels.

**Affective Relationships**

The ACC receives afferents from many structures; those with the most significance
for affective salience are the orbitofrontal cortex, thalamus, ventral striatum, and
amygdala (Devinsky, et al., 1995). Also, the prefrontal region with the highest density of
dopamine receptors is the ACC (Gaspar, Berger, Febvret, Vigny, & Henry, 1989).
Dopamine and dopamine receptors are important due to their role in the reward and
limbic circuitries (Schultz, Tremblay, & Hollerman, 2000). The reward circuit serves as a
route for alerting to the presence and receipt of a reward. Contrary to popular belief, the
reward system is not identical to the addiction system. Newer interpretations of the
literature suggest that addiction may be more complicated than any obsessive-compulsive
drug-craving hypothesis may be able to support. In accord with the work of Schultz,
Tremblay, and Hollerman (2001), the reward and dopamine systems act to alert the brain
about a prediction error. In other words, there are built-in expectations regarding the
nature of certain expected events. When those events are more pleasant than expected,
the dopamine and reward systems are activated. This activation assists the organism in
learning about the reward and it also shapes contingencies based upon behavior. In
addition, it allows the prediction of a pleasurable event based upon prior experience and
leads to behavioral activation, which prompts action and exploration.

Specific studies addressing affective manipulations and the anterior cingulate
have recently been published. Most closely related to this thesis are those concerning
anger and anxiety. A 1999 study from *Biological Psychiatry* used imagery to invoke feelings of anger in eight healthy males (Dougherty, Shin, Alpert, Pitman, Orr, Lasko, et al., 1999). The investigators documented blood flow and activation changes subsequent to imagery and during a control condition. The main findings of significance were the activation of the orbitofrontal cortex and the anterior cingulate cortex.

One aspect of imaging studies that should be mentioned concerns the spatial resolution. Davidson has recently expressed concern about the spatial resolution and statistical techniques utilized with the more advanced imaging modalities now available (Davidson, Abercrombie, Nitschke, & Putnam, 2001). His caution comes from the relative newness of these tools and the lack of tried-and-true statistical analyses. Despite the fact that such simple calculations as subtraction are often used, the processing and set-up of the data is called into question. Davidson also cautions against investing too much faith in the ability of such imaging equipment to resolve structures at the level reported. In other words, readings may be inaccurate. In addition, compatibility of data obtained from different techniques is not firmly established. Thus, spatial data from fMRI may not be perfectly correlated with that taken from PET or MEG.

With the above paragraph as a disclaimer, it will be noted that the Dougherty et al. (1999) study above cited the right anterior cingulate as the primary source of increase in activity. While this may lead to speculation regarding asymmetry among structures, it must be noted that the asymmetrical activation findings have not been consistent enough to make such an inference. To be more transparent, studies using the same or similar protocols have found differences in left versus right activation, despite activation still occurring in the same structural body.
Providing one other example of anterior cingulate activation for affective stimuli, a paper from Kimbrell, George, Parekh, Ketter, Podell, Danielson, et al. (1999), used a recall paradigm to elicit anger and anxiety in adults. They acquired images using positron emission tomography with H$_{2}^{15}$O, in order to assess changes in blood flow relating to affective states. Increased activation in the anterior cingulate was found during both emotive tasks. Furthermore, increased activation appeared to be localized to the left anterior cingulate in the anger condition. Although no significance was attached to this finding by the authors, it provides another example of regionalized activation. As mentioned above, no solid inferences should, as yet, be made from regional findings. Future research should attempt to clarify these region-specific activations.

**Cognitive Relationships**

As mentioned previously, the anterior cingulate shows an increase in activation pending the initiation of any focused state. Obviously, the various acts that constitute cognition represent a focused state. Within the literature concerning ACC activation and cognition, the same issues regarding regionalization with advanced imaging modalities still persist. Again, there are no definitive conclusions to be made from such findings. In this section, some specific results from studies utilizing arithmetic tasks will be presented.

Because most of the ACC literature concerns cognition, a special note should be made regarding potential problems when interpreting these findings. As pointed out by McIntosh, Fitzpatrick, and Friston (2001), most psychology programs do not emphasize complex statistical techniques in their respective curriculums. This has lead to the use of t-tests and ANOVA's for most analyses. According to McIntosh et al. (2001), these tests are insensitive to, and incapable of tackling the numerous measures presented with
advanced imaging techniques. Therefore, contextual relationships among structures anatomically or physiologically linked are often unnoticed. This may not lead to erroneous results; however, it may lead to an inability to dissect complex relationships and subsequent failure to detect key occurrences. The answer, according to the authors mentioned above, is a more stringent statistical analysis conducted by current researchers. Also, there is a need to increase the attention paid to statistical techniques within psychological and psychophysiological training programs.

Dated and recent work has detailed many on the anatomical connections between emotional centers in the brain. More recent work has elucidated connections between these emotional centers and autonomic sites in the brain stem. Also, some work has shown that direct neural connections are also present. These emotional centers show remarkably consistent activation patterns through a myriad of affective states, and are coincident with alterations in cardiac function. Taken together, clear neurophysiological and neuropsychological evidence is present to link alterations in cardiovascular performance to the generation of varying affective states.

What these findings do not provide is a parsing or quantitative explanation of how such relationships can be exploited for further gain. Behavioral manipulation and tasking may prove useful in this quest. Below, empirical findings linking heart rate variability and affective state will be discussed. Where available, findings pertinent to gender differences, and more specifically, estrogen level, will be added. It will be seen that separately, neurophysiological and neuropsychological findings, as well as behavioral manipulations in humans provide much unexplained variation. However, combining the two lines of evidence may prove useful.
Cardiac Measures, Affective State, and Mental Challenge

Although traditional methods inquiry have benefited those with cardiovascular disease a great deal, it is estimated that risk factors derived from such studies explain a mere 40% of the incidence rate (Goldstein & Niaura, 1992). Recently, longitudinal results from large scale studies have demonstrated that emotional states are indeed important risk factors for health problems in general, and cardiovascular disease specifically (Booth-Kewley & Friedman, 1987). At this time, there are two primary theories explaining this relationship. The specifics of these theories will not covered in great detail, because evidence of the credibility of each is equivocal at best. Essentially, the two theories are, 1) top down and, 2) feedback loop (Lovallo, 1997).

Top down theories suggest that personality traits and emotional biases cause autonomic functioning that progressively leads to disease states in the periphery. Feedback loop theories assert that derangements in the periphery are fed back to central sites and the subsequent interpretation and response to this information generates a chronic cycle in which detrimental autonomic function sets the stage for insidious disease progression.

Credibility for these theories stem primarily from studies in which strong correlational evidence is presented. However, because strict timelines have not been established in human cohorts, detailing central, peripheral, and affective functioning, establishing credible causal links has yet to be determined.

One aspect of affective-cardiac research has been regarded with no refute; that is, that negative emotions produce the most exaggerated responses and are the most closely linked with cardiovascular pathology (Kubzansky & Kawachi, 2000). The majority of research linking negative emotions deals with anxiety. Unfortunately, such studies
generally administer psychological inventories probing for trait level anxiety, then perform long term follow up to determine correlations with subsequent cardiac health, or the authors choose to evaluate those with clinically diagnosed anxiety disorders. Obviously, such results are not transferable to normal disease-free populations.

In analyzing negative emotions in apparently normal, disease-free populations, two excellent examples are discussed below.

In the Northwick Park Heart Study, 1457 men were followed for 10 years, after determining their relative levels of anxiety (Haines, Imeson & Meade, 1987). Results indicated a relative risk of 3.77 for men with higher than average anxiety levels. Importantly, this risk elevated to 8.64 in a dose-dependent manner, suggesting that anxiety acts on a continuum to affect heart health. Also, such results bolster theories asserting top down influences.

In a segment of the Framingham Study, 749 women were found to have a relative risk of 7.8 for all forms or cardiovascular disease, depending on their respective levels of anxiety (Eaker, Pinsky & Castelli, 1992). What makes this particular result important is that serial measurements were taken, and other factors such as socioeconomic status were controlled for. This study provided a 20-year follow up period. It was speculated by the authors of this study that degree of anxiety expressed during serial measures represented acute responses to life events, and that this could lend support for theories adhering to feedback mechanisms. Although this is pure speculation, this could provide evidence to support feedback driven mechanisms.

In the above two studies, adjustments were made for known risk factors, making that presence of anxiety statistically independent. Further work aimed at determining just
how much unexplained variance in disease rates can be attributed to these types of results is warranted and will likely be addressed by large cohort studies in the future.

The forthcoming sections will highlight some of the more pertinent facts from studies evaluating the impact of emotions and psychological stress on cardiac performance and health. The reader should note that many studies confine measures to those of blood pressure or heart rate changes. Changes in blood pressure do not provide adequate insight into autonomic regulation, particularly at the level of the heart. While it can be argued that such changes may affect baroreflex function and subsequently affect autonomic discharge, it is important to remember that the baroreflex operates in response to pulsatile pressure and that oscillations in blood pressure are more valid for interpretations of reflex function and any subsequent autonomic fluctuations. Second, changes in heart rate are in no way precise measures of autonomic balance. Also, germane to this study, such changes may be greatly biased by varying levels of estrogen, which as discussed previously, can alter autonomic spectrums. With this in mind, discussion will be limited to measures of utmost clinical relevance and spectral measures of variability.

Empirical Results

The Psychophysiological Investigations of Myocardial Infarction Study group was created to provide insight into cardiovascular function and disease incidence following exercise and psychological testing protocols with a study sample free of CVD upon entry to the study. Both men and women were included in the analysis. The PIMI studies have utilized three basic paradigms for psychological assessment, in addition to standardized exercise stress testing. Mental arithmetic, Stroop testing, and anger provoking situations have all been used and all have shown significant effects on cardiovascular function.
Mental arithmetic generally consists of serial addition or subtraction. Stroop testing is the classic Color-Word test commonly seen in psychological literature. Anger provocation has been either a simple anger recall task or a role-playing scenario.

Clinically useful variables commonly assessed have been wall motion utilizing echocardiography, heart rate, blood pressure, electrocardiography, and outcome measures of disease incidence and mortality rates.

Though the PIMI studies include males and females, to date, none of its results have been targeted at gender-specific issues. Below, results from PIMI-related studies and others are presented. Because much of the literature has either pooled data across genders or failed to explicitly mention gender effects, the data presented below will be an aggregate for both genders.

During mental stress, such as arithmetic challenges, ejection fraction is seen to decline to a greater extent than during exercise, yet ischemia is provoked more often during exercise (Goldberg et al., 1996). With mental stress, 58% of the participants displayed ischemic patterns via EKG analysis, versus 92% of the participants during exercise testing. The reasons for such findings are most likely related to the metabolic demand placed on the body and myocardium during exercise. Central command mechanisms are mostly likely responsible for ejection fractions not dropping during exercise, as compared to mental stress. Metabolic demand, and oxygen requirements increase greatly during exercise, and as such, the ischemic episodes would be expected to be higher during exercise. However, the significant decline in ejection fraction, despite a low metabolic demand of mental stress signifies that aberrant heart motion was generated. This aberration in ventricular dynamics has been attributed to defective, or
uncoordinated, neural signaling to the heart (Stein & Kleiger, 2001). It is highly unlikely that any individual person would be incapable of generating such a chaotic autonomic spectrum from higher brain centers. Thus, the most likely cause for uncoordinated myocardial activity is the failure of the heart to adequately respond to the input it receives. Such data, and the opinions of other authors, suggest that heart rate variability measures would also show abnormal patterns (Becker et al., 1996).

Another arm of the PIMI study evaluated the differences in mortality rates between those individuals with positive exercise stress tests who subsequently underwent psychological stress testing (Sheps et al., 2002). All 196 participants had positive signs of ischemia during exercise stress testing. No variables related to the exercise test proved significant in predicting mortality rates in the five-year follow up period. Of the 196 participants, 42 displayed ischemia during mental stress testing. In total, 17 deaths occurred within five years. Thus, less than 10% of individuals with positive stress tests died within five years, while 40% of those with positive mental stress testing did. What these results point to is the potential of psychological testing to uncover organic disease within the myocardium that may not be evident with exercise testing alone. Again, measures of heart rate variability would be expected to show abnormal behavior during such maneuvers.

A third example of PIMI results specifically addresses anger recall. The most prominent finding was that anger was able to generate heart rate variability changes in all subjects. Using participants with established cardiovascular disease, heart rate variability measures were analyzed for sixty minutes following the onset of anger. A precipitous fall in the HF component was immediately brought about by anger and all subjects during the
study period that displayed signs of ischemia showed these decreases in HF power prior to ischemic presentation and for up to 20 minutes after electrocardiographic indices of ischemia had vanished (Kop et al., 2001). Once again, the impact of psychological stress displays its ability to manifest in noticeable changes before, during, and after the onset of ischemia. Also, not all individuals developing significant drops in HF power developed ischemia, while all with ischemia did display drops in HF power. This would imply that HF power indices provide a highly sensitive marker for the onset of ischemia, which is necessary for infarction.

The vast majority of studies not conducted by the PIMI investigators have not been as credible. However, two exceptions will be provided here. The first comes from the journal *Psychosomatic Medicine*, in 1996; the purpose being to evaluate relationships between circulating catecholamines and low frequency power during several different psychological challenges. The primary conclusion was that low frequency power was not a strong indicator of stress response. The authors reported the raw data and significance levels for changes in several other parameters, high frequency power being one of them. Attaining a significance level of less than 0.001, they found profound decreases in HF power during the anger recall task (Sloan et al., 1996). Like the other studies above, males and females were included.

In the journal *Circulation*, researchers asked patients with implantable cardioverter-defibrillators were asked to record emotional state for the 15 minutes preceding shocks (Lampert et al., 2002). These devices sense a patient’s electrical rhythm and, if necessary, issue a shock to convert the rhythm to normal. Patients are generally aware of such occurrences and thus can make diary entries immediately following shocks. Of all events
preceding cardioversion, anger was the strongest non-exercise predictor. There was no reported interaction between anger and physical exertion, and both factors emerged as independent predictors.

A last example specifically relates to recovery of vagal influence (Mezzacappa et al., 2001). Before describing this study, it must be noted that it is typical of most recovery measures available in the literature, meaning that the methods do not closely adhere to published guidelines and that the rationale generally ignores basic facets of cardiovascular physiology. The preponderance of such inadequacies has greatly diminished the respectability of journals that commonly publish this type of research. Essentially, though many psychological and psychophysiological researchers are correct to pursue integrated protocols, a relative ignorance of physiological data and an extraordinary degree of methodological variation has led to a failure to reach consensus on many issues.

Researchers in the Mezzacappa study used mental stress protocols and then evaluated the recovery of participants for five minutes following the cessation of testing. One arm of the study used a Stroop test and an arithmetic test. Because these tests are consistent with most of the available literature, those results will be discussed.

The marker chosen for vagal balance was the natural log of the root mean square of the standard deviation in R-R intervals (lnRMSDNN). The authors calculated this measure for each of the five minutes of recovery time. Though this measure is acceptable, it is not recommended by published guidelines (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). Guidelines recommend durations greater than five minutes for time domain measures.
This may hinder strict interpretation; however, obvious trends are present, notably a sharp rise in vagal influence during the five minute recovery period. Unfortunately, the authors did not use an affective manipulation in this study.

There are two basic, but important points to come from the Mezzacappa article. First, measuring the return of vagal influence is easily accomplished. Such an assertion is made possible by noting the trend in lnRMSDNN, rising sharply from 3.7 in the first minute to 3.9 by minute five, significant at the 0.05 level. Given that anger recall tasks generally produce greater reactivity, it follows that a greater amount of vagal return would ensue during the recovery phase, making trends more prominent. Also, measuring spectral components over a five-minute period should produce more accurate and efficient results according to accepted guidelines. The last item of note is the deletion of scores reported by the authors. There were a total of 15 males and 15 females tested. The authors reported that three of the scores for females were dropped (20%) because recovery scores were found to be too high. Because the scores were computed across gender, it is unknown whether or not the three outliers would have been considered as such if compared solely against other female scores. At the very least, noting that the female scores were higher lends support for the parasympathetic enhancement by estrogen, consistent with the material presented thus far.

In balance, the literature contains many examples of cardiovascular reactivity to psychological stress. As mentioned above, most do not contain measures with strong clinical significance. Those that do measure heart rate variability or similar measures designed to parse autonomic input, generally include other interventions that preclude accurate interpretation of anger- or cognitively-mediated effects. For example, a study by
McCraty studied anger recall in an elderly population, yet there was not form of control employed and the participants were coached in methods to negate the effects of affective stimuli (McCraty, Atkinson, Tiller, Rein & Watkins, 1995). Unfortunately, many studies such as this are present. In essence, many researchers have attempted to study emotional modulation of heart rate variability without establishing concrete foundations beforehand. Thus, the results of such work present no basic information that can be transferred to other research.

The examples provide above, with the exception of the McCraty study, do provide sound evidence linking heart rate variability or other useful measures to mental stress and anger. Importantly, the above studies used both males and females and participants without cardiovascular disease were entered into the studies to serve as their own controls where necessary.

Thus far, material presented has illustrated that higher brain centers, mental stress, and affective state can powerfully alter cardiovascular function. However, some questions remain to be answered. This information may not be evident from first inspection of the literature review. Therefore, in the next section, methodological issues will be addressed that should help the reader understand how such apparently diverse material comes together to provide rationale for this study. Additionally, some new material will be presented to further support the rationale.

**Methodological Issues**

There are several questions that need to answered when establishing the basis for a study. Though it is impossible to address the concerns of every reader, and certainly every protocol has its weaknesses, it is hoped that the discussion below will help to summarize the available literature and provide additional informative material. This
section will address the following questions: 1) why use an anger recall task, 2) why use females only, 3) why use a recovery measure rather than a reactivity measure, 4) why use HF power, and 5) what is the potential impact of fitness level.

**Why use an Anger Recall Task**

The weakness when choosing anger recall is that a relatively small number of studies specifically use such a task. However, the potential strengths appear to outweigh this drawback. Strengths of using anger recall is that it is simple, appears to be repeatable, does not bias against any type of learning disability, generally displays stronger responses, and it is more often implicated in cardiovascular disease than any other affective challenge suitable for laboratory use. The primary reason for its use relates to the repeatability issue. For instance, in repeated measurements of arithmetic, Stroop testing, and anger recall, it was found that anger recall scores were virtually identical when assessments separated by two months were taken. In contrast, a great deal of variability was found in the other two tests over the same duration (Jain et al., 2001).

In many instances, serial measures of arithmetic testing or Stroop testing breeds complacency and desensitization on the part of the participants (Booth-Kewley & Friedman, 1987). Therefore, the reliability of those tests and their abilities to provide insight into cardiovascular performance and disease would seem rather limited. It must be remembered that most testing protocols are repeated over the lifetime of a person in order to assess changes over time, such as blood pressure recordings during annual visits to a physician’s office. Because anger is an emotion that everyone is likely to experience, and because explaining how to be angry is apparently unnecessary, using anger recall should prove to be the most ecologically valid measure for serial assessments.
Why use Females Only

Overall, the literature is riddled with contrasting findings in the female population. The amount of variance in female results has prevented clear consensus opinions regarding mechanistic links between affective state and cardiovascular function. More importantly, gender-specific knowledge is desperately needed to improve the care of females with cardiovascular disease. Gender-specific studies should improve the diagnostic and treatment abilities, as well as advance biological knowledge.

Why use a Recovery Measure Rather than a Reactivity Measure

This is perhaps the most difficult aspect of the study to justify, primarily because its rationale is based upon shortcomings in the literature and potential methods to alleviate these items, as opposed to empirical evidence. First, many issues have arisen concerning the proper way to assess reactivity. At this time, there are no infallible guidelines for this measure. It is of course, in essence, the analysis of change from baseline conditions. Currently, analyzing change-scores and changes in scores is subject to much interpretation (Collins & Sayer, 2001). By utilizing recovery measures and regression analysis, this potential pitfall is avoided. Also, of equal importance is the fact that patterns of reactivity are still not agreed upon (Lovallo, 1997). Much less statistical and methodological debate surrounds the use of recovery measures and regression analysis to model and predict their subsequent values. Hence, recovery measures would appear to be the most powerful index, producing the least amount of problems for analysis and interpretation.

Why use HF Power

This is perhaps the easiest question to answer. In addition to the study presented in the previous section, which found low frequency power to be an unreliable marker of
reactivity, numerous studies have found strong ties between vagal tone and cardiovascular mortality (Goldberger et al., 2001; Tsuji et al., 1996). Thus, vagal indices are the clear choice for strict scientific inquiry and clinical utility.

**What is the Potential Impact of Fitness Level**

There is clear evidence that increasing fitness levels increase resting parasympathetic power and enhance recovery (Myers, 1996). However, the evidence that fitness level mediates recovery from psychological or affective stress testing is less well known. Several studies provide equivocal results. Though studies generally do utilize both males and females, separate measures for each gender have not been calculated. Given the equal abundance of positive and negative results, it does not appear that fitness level will bias recovery from psychological testing protocols.

As stated above, higher fitness levels produce higher resting levels of parasympathetic power. Given this relationship, it would stand to reason that during psychological testing, a high-fit individual would withdrawal a greater amount of HF power to respond, and would likewise produce higher absolute level of recovery in terms of HF power. Therefore, using resting levels of parasympathetic power as a covariate would tend to account for this. At this time, there are no empirical results to either support or refute such assertions. However, it is commonly accepted by statistical experts, to employ such covariates if these types of relationships are thought to exist (Cohen & Cohen, 2003). Additionally, including covariates in a regression analysis allows the researcher to use polynomial adjustments if required, while strict application of ANCOVA techniques does not provide this flexibility.
Summary of Literary Evidence

This area of knowledge is truly an integrative enterprise, combining tenets of psychological theory with the concrete, quantitative results from physiological data. The preponderance of physiological data suggests that estrogen has powerful biologic effects that sustain and enhance parasympathetic-effect. This data alone is enough to justify the need for such an investigation. This is further supported by the numerous calls to improve the state of gender-specific knowledge and medical treatment. At this time, psychological theory, and its respective literature, provides more of a confounding influence than would be preferred. However, given that research accounting for varying estrogen levels proves fruitful, psychological theory provides a great potential in its ability to generate powerful inventories that could easily be employed in clinical as well as laboratory settings. As stated previously, current medical knowledge of cardiovascular disease and its risk factors accounts for a mere 40% of the variance in disease incidence. Assuming that modeling estrogen’s effects is worthwhile, then it follows that combining precisely targeted psychological evaluations may explain the majority of unexplained variance currently present. Of course, the steady advances in genetic research should prove to complete the analysis. However, the debate between nature and nurture is not likely to subside, despite gains in genetic research; thus, psychological theory will need to continue progressing within this research area.

With the above considerations and caveats in place, accounting for varying estrogen levels is certainly warranted at this present time.
CHAPTER 3
METHODS: EXPERIMENTAL DESIGN AND ANALYSIS

Design and Rationale

This was an observational study, generating a regression equation for the prediction of HF power following an anger recall task.

In brief, the study consisted of a random sample from a female subset of students attending the University of Florida. For each participant, a five-minute, resting baseline measure of heart rate was acquired, and spectral power determined for the HF component. Following a brief set of instructions, each female performed a five-minute anger recall task, during which, the same HF power measurement was taken. After the anger recall task, a one-minute transition period was allowed and another five-minute resting measurement acquired, representing the recovery phase. The same spectral power measure was calculated during the final phase.

The primary criterion in this investigation was the HF power during recovery. The measures of baseline HF power, estrogen levels, and circadian status were utilized as predictor variables. The serum estrogen levels were estimated from standard values (Yen, 2001).

Historically, the primary objective of observational designs has been the formulation of predictive regression equations; such that, in the clinical setting, individuals can be tested and evaluated further if actual scores are not bound by prediction intervals. An analogous goal was pursued in this investigation.
In addition, gender specific research has gained much popularity in recent years, for good cause. The proportion of female specific research has traditionally been suppressed by logistic concerns such as controlling or allowing for fluctuating hormonal levels. It is now known that female hormonal profiles lead to differing biological profiles. Regression allows the researcher to maximize the potential use of observations by modeling and predicting future criterion scores, based on continuous predictor variables that represent the varying biological effects of fluctuating hormone levels.

**Participants**

All participants (N=50) were selected from volunteers solicited from classes taught within the Department of Exercise and Sports Sciences at the University of Florida. Females alone were asked to participate due to the gender-specific focus of this investigation. There was selective screening for age, caffeine intake, smoking status, and history of cardiovascular risk factors, in addition to current use of medication. Also, a subjective history of menstrual status was obtained.

Age was limited to those females ranging from 18 to 25 years. All participants were asked to refrain from caffeine intake for 24 hours prior to arrival at the laboratory. Inclusion necessitated a self-reported abstinence from smoking of any kind within the previous six months. Participants with a history of any type of cardiovascular pathology, including the current use of cardiovascular medication were to be excluded; however, no such circumstances presented during the solicitation of participants. Questions regarding current medications being used were asked of each participant. If medications were reported, the investigator then determined the potential for side effects that could bias the responses. Inference of any such side effects precluded participation. Further, each volunteer was assessed for resting, seated blood pressure by the researcher, using
standard sphygmomanometry techniques. Normal blood pressure was a requirement for participation. Those volunteers with systolic blood pressure above 139 mmHg and/or diastolic pressures above 99 mmHg were to be excluded from participation. No such measures were obtained. Those with systolic pressure lower than 100 mm Hg or diastolic pressure lower than 60 mm Hg were also to be excluded (Alexander, Schlant & Fuster, 1998). Again, no such measures were obtained.

This study required an accounting of menstrual status. Participants were asked to provide the number of days since the start of the last menstrual cycle. Further, it was necessary for inclusion that each participant to have had normal menstrual cycles for at least six months prior to the investigation and that no pharmacological contraceptives were used in that time frame. Failure to provide a certain accounting of menstrual status precluded participation.

Due to the observational design, there was no assignment protocol for the participants.

**Measurement**

Predictor variables observed were the high frequency component of the heart rate variability spectrum, status within the menstrual cycle (serving as a surrogate for serum estrogen values), and circadian status represented by the time since awakening.

Heart rate variability is a spectral measurement of the variations in the intervals between successive R-waves. A graph comparing interval time to absolute time was constructed. After plotting the above-mentioned graph, a Fast Fourier Transform was performed over an interval of five minutes, as per standardized guidelines (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). The resulting transform was evaluated from 0.00 Hz to 0.50
Hz. The very low frequency (VLF) is characterized as 0.00 Hz to 0.04 Hz; low frequency as 0.04 Hz to 0.15 Hz; high frequency as 0.15 Hz to 0.40 Hz. It has been recommended that the VLF be omitted from recordings of five minutes or less due to uncertainty of the factors responsible for its occurrence. The LF is thought to represent an interplay of parasympathetic and sympathetic activity, though no clear way to divide the relative influence of each exists; and the HF is considered to be almost solely due to parasympathetic input (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). This investigation focused on the high frequency component alone.

The integral of the Fourier Transform within a specified frequency range represented the total power within that frequency range.

Serum estrogen levels were represented by standard values collected, pooled, and published in the literature (Yen, 2001). Because different authors have chosen to present data in different units, for example, some authors report estrogen in picograms while others report micrograms; the values were transformed to percentages of the maximum.

Many investigations have described a consistent pattern for the menstrual cycle. Though the menstrual cycle can be interrupted or perturbed by many processes, absent of pathology, it does proceed in an orderly manner, providing predictable values based upon the number of days since the last menses. As reported in the pooled results mentioned above (Yen, 2001), the menstrual cycle typically varies by only one day, from cycle to cycle, within the age group used in this investigation.

**Rationale for use of HF, E, and Circadian Status**

The high frequency component is thought to most closely represent the parasympathetic arm of the autonomic nervous system. There is evidence for estrogen’s
effect in the periphery and at central sites, as outlined in chapter two. However, to date, there is no clear conclusion regarding estrogen’s effect upon the measures of parasympathetic tone following an acute stressor. Typically, during an anger recall task, a drop of at least 15% in the HF component is observed, more commonly the deviations drop below 20% (Jain et al., 2001). This produces a relative dominance in sympathetic tone. The parasympathetic system is capable of blunting this sympathetic effect, serving as a potentially protective mechanism against the adverse effects of sympathetic dominance. During periods of arousal, parasympathetic influence is withdrawn as an organism prepares to mobilize resources in response to challenge. However, parasympathetic tone does not completely abate, and it is believed that parasympathetic input serves to increase the efficiency of an organism’s response. The effectiveness of estrogen to enhance parasympathetic effect is potentially mediated by structural characteristics within the myocardium. In addition, pathology may create a myocardium that is resistant to vagal input, allowing sympathetic influence to act in an unrestrained manner. Therefore, throughout the lifespan, utilizing measures of parasympathetic effect, mediated by estrogen, may provide a tracking mechanism for changes within the myocardium.

Finally, it has been realized for many years that circadian rhythms alter hormonal and electrocardiographic profiles. For instance, it is typical for sympathetic tone to peak during the first couple of hours after awakening. To date, there is no literature known by this author to lend support to the notion that circadian status affects reactivity or recovery from psychological challenge. Because the omission of predictor variables can be
devastating to a regression analysis, circadian status, taken as the number of hours since awakening, was recorded for each participant and utilized as a predictor variable.

**Signal Processing**

The electrocardiographic signal was amplified by a factor of 5,000 using Biopac (3.1.1, Santa Barbara, CA) bioamplifiers (ECG100A) and bandpass filtered from 1 to 35Hz. Equipment unique to the heart rate analysis included the following: 12 mm Biopac, Ag-AgCl electrodes (ADD212), with the ECG100A amplifier set to filter from 1 to 35 Hz, with a gain of two thousand.

Artifacts were visually identified to allow for more accuracy and judgment by the author. This method, vice algorithmic methods, is chosen because standardized guidelines have recommended visual inspection, due to high probability of wandering baselines, muscle artifacts, and lead repositioning during movement (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). Any records found to have artifact that precludes accurate discernment of R waves, or loss of signal integrity, were excluded from analysis.

To minimize the potential for artifact, participants were seated in a chair and asked to refrain from speaking, moving, or opening the eyes during the experimental session. They were instructed that if any discomfort, illness, anxiety, or any other need arose, to notify the experimenter immediately.

**Procedure**

Participants were solicited from classes taught at the University of Florida. A brief oral description of the investigation was provided, and then potential participants were asked to complete a simple questionnaire, shown in Appendix B. The investigator quickly surveyed the answers and if in accord with inclusion criteria, the investigator
asked the volunteer to consent to participation. Those volunteering were provided with an
informed consent form and asked to schedule an appointment with the investigator for the
experimental period. After, the volunteer was given a second questionnaire and asked to
complete it just prior to arrival at the Human Performance Psychology Laboratory on the
University of Florida campus. Both questionnaires provided to participants are displayed
in Appendix B, along with information pertinent to each questionnaire.

Once at the laboratory for the experimental session, and after receipt of the signed
informed consent and second questionnaire, each participant was evaluated for resting
blood pressure. After verification of normal blood pressure, participants were escorted to
the testing area for introduction to the equipment and the set-up procedure, followed by
set-up.

The participants were given a brief outline of the procedure without notification
of the exact nature of the tasks to be employed, to avoid biasing the baseline measures.
After collection of the baseline, an oral description of the anger recall task was given to
each participant. The experimenter asked if the participant was ready and willing to begin
the task and notified that once begun, the session would continue until the end with little
interruption. If an affirmative answer was received, the experimenter instructed the
participant to begin the anger recall task, and then left the laboratory. All participants
were told that at any time they could terminate the session by calling out to the
experimenter.

After five minutes of the anger recall task, the investigator returned to the testing
area and asked the participants to cease the anger recall task. The investigator used a
stop-watch to monitor the time while out of the testing area. Exactly six minutes after the
beginning of the anger recall task, the investigator reset the data acquisition system to begin collection of the recovery period electrocardiograph.

After completing the second resting measure, the participant was notified that the session was over and that she may ask questions or view the records of her physiological measures.

**Manipulation Check**

Each participant was asked to perform the anger recall at her own volition, with full intent. This provided a lack of experimental control. However, in previous research, it has been established that quantitative markers of effort do not show significant variation despite marked variation in outcome measures. However, in surveying the literature, anger recall is most often associated with a minimum 15% drop in HF power, regardless of age and gender. Therefore, each record was assessed for a minimum 10% drop in HF power during the anger recall task. In other words, the ratio of the anger recall HF power to baseline HF power could not meet or exceed 0.90. If this requirement was not met, the data was not entered into the full regression analysis.

**Statistical Analyses**

**Significance, Power and Sample Size Determination**

Using Cohen’s recommended formula (Cohen, Cohen, Aiken & West, 2003), and a medium effect size with a significance level of 0.05, the necessary sample size was 78 participants. Due to the small variation assumed within the study sample, there is an assumption of a much larger $R^2$; and, it was expected that the effect size would be much closer to 1.0, and most likely higher. The midpoint between sample sizes for the traditional medium effect (0.13) and the minimally anticipated 1.0 is 45, thus a sample size of 45 participants was sought. To account for attrition, 50 participants were sought.
All statistical tests utilized 95% confidence intervals, equating to a significance level of 0.05, with the exception of Hypothesis one. Hypothesis one utilized a significance level of 0.10.

**Analysis**

Bivariate and multiple linear regression using ordinary least squares was employed. The predictor variables for the multiple regression analysis were estrogen levels, circadian status, and baseline HF power. The bivariate regression analysis regressed baseline HF power on a predictor of serum estrogen level. For both analyses, the assumption of linearity was made because baseline measures of HF power have been shown to be amenable to transformation when needed in past research.

Other assumptions and practical issues regarding linear regression include detection of outliers, absence of multicollinearity, normality of residuals, homoscedasticity of residuals, independence of errors, and the correct model specification.

Detection of outliers was accomplished via graphical and numerical methods. The graphical measures were employed via simple inspection of scatterplot matrices. The main use of graphics was to guide and supplement numerical diagnostics and to aid decisions regarding the true status of a data point.

Numerically, estimates of each value’s leverage, discrepancy, and influence were calculated.

Multicollinearity is a term describing independent variables that are highly correlated. In short, multicollinearity creates a situation in which variance of the criterion is shared by independent variables too highly correlated with one another. Thus, the confidence intervals and predictive power of an equation is greatly reduced. Strictly,
multicollinearity exists when there are perfect correlations between predictor variables. However, there exists a continuum such that high correlation reduces the effectiveness of the regression equation as the correlations increase. In this investigation, it is assumed that low correlation will exist between baseline HF and estrogen level, as there is no literary evidence to the contrary. The same applies for circadian status. Therefore, multicollinearity is not an anticipated problem. Variance inflation factors were surveyed to detect multicollinearity, using guidelines published in Cohen, Cohen, Aiken, & West (2003). The optimal value for variance inflation factors is 1.0 for each variable. The general recommendation is that a value exceeding 10 requires close examination and remedial action. Remedial action for multicollinearity was not required in this investigation.

Normality of residuals establishes that there are no trends in the regression equation that biases the predicted values. A simple histogram of the residuals was evaluated visually for normality, and the Kolmogorov-Smirnov statistic was used to formally assess normality. With the exception of heavy tailed distributions, regression equations are not affected by slight departures from normality.

Homoscedasticity refers to an equal variance among the residuals. Again, visual inspection was utilized. Formally, the Breusch-Pagan test was used in that evaluation.

Bias in regression analysis is a serious concern. A common tool for evaluating a model’s potential bias is Mallow’s Cp. The Cp statistic will ideally be equal to the to the total number of parameters in the regression equation. Mallow’s Cp was used to identify the best model in a subsets analysis. The best subset identified was based upon Mallow’s Cp, the $R^2$ value, and the mean squared error.
Lastly, and most importantly, it is assumed that the correct model has been proposed. After a prudent diagnostic analysis, an overall test for lack of fit was carried out. Because it is possible for assumptions to be met, while not providing a properly fitted model, this test was the last item calculated.

**Regression Model**

As stated previously, the criterion variables were HF power during the recovery phase of the laboratory session for the multiple regression, and HF power during baseline for the bivariate regression. The predictor variables were circadian status, baseline HF power and estrogen levels for the multiple regression, and estrogen values for the bivariate regression.

**Hypotheses**

Regression analysis does not yield to efficient hypothesis testing. However, three germane hypotheses were developed and tested.

**Hypothesis One**

Not directly related to the primary objective of the study, this hypothesis was generated to add weight to the assumption that the predictor variables were not related by causal or chance mechanisms. Simple bivariate regression was performed with baseline HF power as the criterion and estrogen level serving as the predictor. The research hypothesis was that in a 90% confidence interval, the $R^2$ value of 0.00 would be contained therein.

Hypothesis one: baseline HF power cannot be regressed on estrogen level at a significance level of 0.10.
Hypothesis Two

This was the primary hypothesis, consisting of testing the strength of the full regression model. In keeping with the assumed limited variation within the study sample, and a projected high amount of variance being explained by the regression equation, it was hypothesized that a 95% confidence interval for $R^2$ would contain a value of 0.75, which was deemed a necessary requirement for the practical success of this study. A caveat to this assertion was that if the coefficient was substantially higher than 0.75, a confidence interval’s lower limit may indeed surpass this value, which would not serve to reject the hypothesis.

The criterion in this multiple regression analysis was recovery HF power. The predictors were estrogen levels, baseline HF power, and circadian status. A best subsets analysis was carried out to make an initial model selection. The best model from this analysis was subjected to diagnostic and remedial procedures to satisfy the assumptions of multiple regression analysis.

Hypothesis two: in the full model, $R^2$ would meet or exceed 0.75, using a 95% confidence interval.

Hypothesis Three

Lastly, to concretely validate the inclusion of all variables selected for the final model, it was hypothesized that a 95% confidence interval for each partial coefficient will not include zero.

Hypothesis three: partial coefficients would differ from zero, using a 95% confidence interval.
CHAPTER 4
RESULTS AND DISCUSSION

The results of this investigation will be presented in a concise manner at the beginning of this chapter, including the formal hypothesis testing results. Following will be a more detailed discussion of how the final regression model was attained.

Overall, the results suggest that a valid regression function was established. As Myers (1990) has said, “One must understand that an adroit job of model selection does not require that one actually locate the correct model. Indeed the correct model may never be found.” Thus, the full model outlined below may require extensive modification for further use and potential clinical application. However, it constituted the best available choice given the \textit{a priori} constraints outlined prior to data collection.

Though it cannot be assumed that the best regression function was produced, the power, effect size, \( R^2 \), and assumption checks were all positive results.

\textbf{Units of Reporting}

All results are presented with a maximum of four significant figures. Most of the units are self-explanatory, with age reported in years, blood pressure in mmHg, and circadian status as hours since awakening. All measures for HF power are calculated as msec\(^2\), however the units will not be attached to each reported value, for convenience. Only ratios with the same units are calculated and as such, the units cancel in the calculation, leaving a simple numerical score.
Manipulation Check and Data Screening

Fifty participants were recruited for this investigation, with 43 remaining in the final model. Two individual recordings showed evidence of artifact and were not analyzed. Four individuals did not pass the manipulation check. One score was dropped from the final model because of its influence on the regression equation.

As stated in chapter three, previous studies have found consistent drops in HF power during anger recall. Though these studies have primarily utilized diseased participants, profound and consistent drops in excess of 30% are not uncommon. Due to the assumed health and increased resistance to such precipitous declines in HF power, a minimum 10% decline was set as the cutoff for this investigation. The manipulation check consisted of dividing the anger recall HF power by the baseline HF power. Values less than 0.90 were considered acceptable. Four individuals failed to pass the manipulation check. Two voluntarily stated the inability to generate a substantially angering recall. Two others made no such statements. Table 4-1 presents the results of the manipulation check. The table presents the aggregate of all 48 available scores, and the segregated scores of those passing. Of the two individuals stating an inability to generate an angering event, their anger/baseline ratios were 0.97 and 0.95; while, those not making such statements had ratios of 0.92 and 0.90.

Table 4-1 Summary data for the anger/baseline ratio manipulation check.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Manipulation Check Before</th>
<th>Manipulation Check After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.81</td>
<td>0.80</td>
</tr>
<tr>
<td>Median</td>
<td>0.82</td>
<td>0.82</td>
</tr>
<tr>
<td>$SD$</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>N</td>
<td>48</td>
<td>44</td>
</tr>
</tbody>
</table>
To ascertain whether any systematic bias was involved in the scores that did not pass the manipulation check, inspection of each failing score was conducted. Scores on all measures were well within normal ranges, as compared to other participants. The only discernable trend was that HF power continued to fall during the recovery period, as all four participants’ scores during recovery were lower than during the anger recall task. This may signal that the individuals were aware that the anger recall task was not effective. Subsequent rumination on the task during the recovery period would be expected to produce HF values that did not show a trend toward returning to baseline values.

**Data Screening**

Regression analysis requires that the residuals from the regression equation be normally distributed when utilizing least squares calculations. However, it does not require the individual explanatory variables to be normally distributed. Therefore, the most important aspect of initial data management is the identification of points that preclude inclusion for further analysis.

Participants were screened for inclusion criteria as described in chapter three, via questionnaires shown in Appendix B. Other screening criteria were targeted at the identification of artifacts, error-free data collection and error-free data entry.

Two experimental sessions were discarded from analysis due to excessive artifact. The artifact was believed to be a consequence of the sensing leads losing contact with the skin. This artifact was witnessed after the termination of the baseline period in one participant and during the anger recall period in the other. In order to alleviate the potential for biasing, the experimental sessions for these two participants were terminated.
and none of their data was entered into the final regression analysis. These two participants were not asked to repeat the protocol.

All scores were analyzed immediately after collection, and repeated two additional times to ensure that scores were identical and accurate. Data entry was made using two separate files, which were compared to assure each entry was identical. No errors in the collection, calculation, or data entry were detected.

**Univariate Analysis**

Visual and numerical inspections of each measured or reported quantitative variable were conducted after initial data screening. There were 44 sets of scores that were allowed into this analysis. Univariate analysis was conducted on the following variables: age, systolic blood pressure, diastolic blood pressure, circadian status, estrogen level, baseline HF power, anger recall HF power, recovery HF power, recovery/baseline ratio, and anger/baseline ratio. Of these, circadian status, anger HF power, and anger/baseline ratio did not appear to represent normal distributions. Additionally, \( p \) values for K-S tests attained or approached significance, with respective \( p \) values of 0.099, 0.106, and 0.010.

A trimodal shape was apparent with circadian status. The most popular scores were four, nine, and eleven hours since awakening. There is no apparent explanation for this since many of the participants were evaluated at varying times during the weekends, in effort to alleviate trends that may arise from students conforming to class and work schedules.

HF power during the anger recall task displayed a bimodal appearance, with peaks centered approximately at the 600 and 800 intervals. Other scores throughout the range were represented with no gaps. The visual appearance of anger recall scores is suggestive
of a bimodal distribution; however, formal K-S testing provided a $p$ value of 0.106, which is not rigidly significant. There is no readily available rationale for this trend toward bimodal distribution. Given that such a distribution does not violate assumptions of least squares analysis, the scores were deemed appropriate and acceptable.

Finally, the anger/baseline ratios were obviously skewed in the negative direction. This is expected because the anger recall task has been shown to be quite provocative and consistent. Previous findings have also shown that some scores do drop extremely during the task. In this analysis, two scores appeared to be potential outliers, with ratios of 0.60 and 0.63. Using two different numerical analyses, neither score was considered to be an outlier. One method of analysis is to use the interquartile range, and the other is simply using the mean and standard deviation, assuming a normal distribution. When utilizing the interquartile range, and the recommendations of Ott & Longnecker (2001), the cutoff for outlying values of the anger/baseline ratio would have been 0.46. Those authors recommend tripling the interquartile range and subtracting this from the first quartile to set a limit for extreme outliers on the lower end. Using conventional standard deviation testing, and the fact that within three standard deviations, approximately 99.7% of all measurements will be contained, the most extreme score attained was 2.99 standard deviations from the mean. Therefore, though the anger/baseline ratios were negatively skewed, such a distribution is logically expected, and numerical means of assessment do not lend support for the presence of outliers.

**Overall Results of Hypothesis Testing**

Brief descriptions of the formal hypothesis tests are presented in this section. Table 4-2 provides descriptive statistics for the collected measurements, after the manipulation check and data screening processes were completed.
Table 4-2 Descriptive statistics for N=44 scores passing the manipulation check.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>44</td>
<td>21.2</td>
<td>21.0</td>
<td>2.17</td>
<td>18-25</td>
</tr>
<tr>
<td>Circadian status</td>
<td>44</td>
<td>7.50</td>
<td>8.50</td>
<td>4.00</td>
<td>1-15</td>
</tr>
<tr>
<td>Diastolic pressure</td>
<td>44</td>
<td>74.5</td>
<td>74.0</td>
<td>7.44</td>
<td>62-92</td>
</tr>
<tr>
<td>Systolic pressure</td>
<td>44</td>
<td>114</td>
<td>113</td>
<td>8.83</td>
<td>102-134</td>
</tr>
<tr>
<td>Estrogen level</td>
<td>44</td>
<td>59.1</td>
<td>58.0</td>
<td>19.5</td>
<td>25-100</td>
</tr>
<tr>
<td>Baseline HF</td>
<td>44</td>
<td>920</td>
<td>922</td>
<td>158</td>
<td>623-1209</td>
</tr>
<tr>
<td>Anger recall HF</td>
<td>44</td>
<td>738</td>
<td>736</td>
<td>150</td>
<td>488-1001</td>
</tr>
<tr>
<td>Recovery HF</td>
<td>44</td>
<td>768</td>
<td>762</td>
<td>143</td>
<td>453-1046</td>
</tr>
<tr>
<td>Anger/Baseline</td>
<td>44</td>
<td>0.80</td>
<td>0.82</td>
<td>0.06</td>
<td>0.60-0.89</td>
</tr>
<tr>
<td>Recovery/Baseline</td>
<td>44</td>
<td>0.83</td>
<td>0.83</td>
<td>0.06</td>
<td>0.70-0.98</td>
</tr>
</tbody>
</table>

**Hypotheses Testing**

For all hypotheses, the coefficient of multiple correlation is given as the adjusted coefficient, which is more conservative and adjusts for the sample size and number of parameters in the model that was tested. For most of this section, a “%” will be added to the coefficient of multiple correlation to emphasize the value explains a percentage of the variance in the criterion.

Hypothesis one stated that baseline HF power could not be regressed on estrogen levels, using a significance level of 0.10. The $R^2\%$ value obtained from the regression was 3.4%, the adjusted value falling to 0.90%. The subsequent $F$ test produced a test statistic of 1.43 with a $p$ value of 0.238, thus failing to reach significance, supporting the research hypothesis. The 90% confidence interval for $R^2$ was (0.0586 – 0.1226), which clearly includes zero.

Though the validation of the research hypothesis does not eliminate the possibility that estrogen and baseline HF power may interact, it does allay potential multicollinearity worries and further supports the notion that both variables theoretically and independently contribute to the recovery value of HF power.
Table 4-3 Summary statistics for hypothesis one. Baseline regressed on estrogen.

<table>
<thead>
<tr>
<th></th>
<th>$R^2_{adj}$</th>
<th>$F_{(0.05, 1.42)}$</th>
<th>$p$</th>
<th>MSE</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>0.90%</td>
<td>1.39</td>
<td>0.245</td>
<td>24726</td>
<td>1.23</td>
</tr>
</tbody>
</table>

Baseline HF = 835 + 1.45 Estrogen

Hypothesis two set forth that the overall regression function would produce an $R^2_{adj}$ that reached 75%, as a minimum. The 95% confidence interval for $R^2_{adj}$ was 95.5% ± 0.0002%, which obviously surpasses the 75% minimum required to support the research hypothesis. The final model retained only two of the three originally proposed predictors, the two being estrogen level and baseline HF power.

Partial correlation coefficients ($pr^2$) and semipartial correlation coefficients ($sr^2$) were calculated for the coefficients in the final model. These coefficients are listed in table 4-4. Partial correlation coefficients provide an index of the amount of variance in the criterion that can be explained by that specific regressor, once the other regressors have been accounted for. In other words, how much of the variance in the criterion, not accounted for by all other regressors, is uniquely accounted for by the specific regressor of interest.

Semipartial correlation coefficients provide an index of the unique amount of variance in the criterion associated with a specific regressor, once the effects of other regressors have been partialed from the regressor of interest. In a simpler definition, once the full amount of variance explained by a set of predictors has been calculated, that part unique to a specific regressor is expressed by the square of its semipartial correlation coefficient.
Interpretation of the partial and semipartial correlation coefficients is occasionally difficult, conceptually and mathematically. Figure 4-1 presents a diagram, similar to one presented in Cohen, Cohen, Aitken and West (2003), which helps to explain the concepts. What the resulting calculations show is that the addition of estrogen as a predictor helps to explain an additional 49.6% of the variance not accounted for by a model that already includes baseline HF power. Also, estrogen contributes approximately 10% to the understanding of the variance in recovery values, independent of baseline. In prediction analyses, it is the partial correlation coefficient that has the most utility, because it clearly shows how much additional variance can be accounted for, and as a result, provides a rough index of how much more refined a prediction can be. More detailed analysis of the prediction analysis is provided below in the supplementary analysis.

Table 4-4 Summary statistics for hypothesis two. Full model.

<table>
<thead>
<tr>
<th>N</th>
<th>$R^2_{(adj)}$</th>
<th>$F_{(0.05, 2, 40)}$</th>
<th>$p$</th>
<th>$C_p$</th>
<th>MSE</th>
<th>Pred.$R^2%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>95.50%</td>
<td>446.02</td>
<td>&lt;0.001</td>
<td>4.3</td>
<td>939</td>
<td>94.92%</td>
</tr>
</tbody>
</table>

Predictors

<table>
<thead>
<tr>
<th>S.E.</th>
<th>V.I.F.</th>
<th>$sr^2$ %</th>
<th>$pr^2$ %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrogen</td>
<td>0.249</td>
<td>1.0</td>
<td>10.6%</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.030</td>
<td>1.0</td>
<td>71.8%</td>
</tr>
</tbody>
</table>

Hypothesis three asserted that the partial regression coefficients in the final model would each be significant, lending support to their inclusion in the final regression function. The joint 95% confidence intervals for each partial coefficient included in the final model were $2.47 \pm 0.676$ for estrogen and $0.779 \pm 0.081$ for baseline HF power. Both estrogen and baseline HF power were significant predictors.
Figure 4-1 Conceptual diagram of partial and semipartial correlation coefficients.

**Power and Effect Size**

The overall power of the investigation was calculated by evaluating the $R^2_{adj}$ for hypothesis two, given that it resulted in the final model that would be used in the future for any further research. The $R^2_{adj} = 0.955$, and the power analysis exceeded 0.99, with an effect size of 21.22. The formula for the power calculation can be found in Cohen, Cohen, West, and Aiken (2003).

**Analysis Progression**

In the following sections, a more detailed analysis is provided regarding the criteria used when evaluating each hypothesis. These sections were constructed in order to give the reader a clearer interpretation of the data analysis, guidelines, and decision-making processes.

**Hypothesis One**

Hypothesis one stated that baseline HF power could not be regressed on estrogen level, using a 0.10 significance level. This hypothesis was made to support prior
empirical results showing that estrogen mediates HF power within individuals; yet, is not a strong predictor for interindividual variation (Evans et al., 2001).

The results of this analysis support the above claim. Though a regression format was used in the calculation, it is helpful to remember that a bivariate regression is similar to correlational analysis. The benefit of the Pearson correlation coefficient is that it is scaleless. Both of the variables, estrogen level and baseline HF power, were normally distributed, continuous, with a bivariate normal distribution. Thus, the $R^2$ obtained from regression analysis is the same as the squared value of the Pearson-Product moment correlation. A value of 0.032 was obtained for $R^2$. The corresponding $p$-value was 0.245, and no remedial measures were capable of increasing the utility of the analysis. Lastly, the calculated slope for regressing baseline HF power on estrogen was 1.45, and the 90% confidence interval was $1.45 \pm 2.07$, yielding a confidence interval for the slope that included zero (-0.62 to 3.52). This is in agreement with the Pearson product moment calculations and the failure to attain significance, pointing to the suggestion that no linear relationship existed between estrogen level and baseline HF power.

**Hypothesis Two**

The main hypothesis of this investigation proposed a strong predictive ability for recovery values of HF power, based on the baseline HF power, circadian status, and estrogen levels. To date, there is no empirical evidence to suggest that any of the other variables measured in this investigation are potential candidates.

As Wilkinson (1999) points out, researchers should not indiscriminately analyze all possible comparisons, or models. In plain language, he suggests that exploratory analysis is acceptable, but fishing expeditions are not. Also, he advises that researchers state, *a priori*, the comparisons, models, and rationale used in the analysis, based upon strong
trends in the literature, not small nuggets plucked from many different sources. The methods and analysis utilized here have sought to adhere to his guidelines, with one exception. Given the equivocal results of previous studies, this investigator found it prudent to survey many different areas of the literature, in order to develop a coherent paradigm, which may explain the lack of consensus. This was especially prudent given the integrative nature of psychophysiological investigations.

The literature suggested that estrogen be rigidly accounted for, in a quantitative sense, as opposed to dichotomization. With the natural assumption that one’s baseline would serve as a strong covariate for the subsequent recovery period, a regression analysis became the obvious choice. Combined with evidence that circadian status has been shown to affect cardiac electrophysiological measures, that variable was also added to the equation. Finally, in keeping with Wilkinson’s recommendation (Fisher’s as well), a simple analysis, with few predictors was employed.

Most guidelines stipulate a minimum of ten observations for each predictor in a given regression equation. This analysis meets those minimums. However, from a practical point of view, it is unwise to place much confidence in an equation targeted for prediction without a substantially larger sample, given the importance of prognostic prediction in medical care. That does not invalidate the conclusions of this investigation; rather, it serves to temper a rush to judgment. Also, larger samples would allow for the adequate characterization of polynomial and interaction effects, which, if present would seriously lower the power of the current investigation given its sample size. This does not mean that polynomial and interaction effects were not included in the exploratory analysis of matrix plots. These matrix plots are presented in figures 4-3 and 4-4.
Evidenced by the scatter plots, and later numerical diagnostics, the baseline scores did serve as a strong predictor. Estrogen status was not as strong independently, but
greatly added to the predictive power. Circadian status did not emerge as significant predictor. In the sections below, a rigorous accounting of the regression analysis will be presented. In addition, scatter plots and statistical tests will be presented to support the notion that the model selection was optimal.

**Best Subsets Analysis**

Given that only three variables and their respective powers and interactions were selected for the main regression analysis, the best subsets were derived using MINITAB software (Release 13). Most authors have agreed that with rare exception, finding best subsets or stepwise procedures that evaluate first order coefficients only are acceptable methods, given that scatter plot examinations do not strongly suggest nonlinear relationships. Further examination of first order predictors may include interactions, but for initial variable selection, first order effects are sufficient (Harrell, 2001). Previously shown scatter plots showed linear relations for most predictors except those including circadian status (C). Though subtle nonlinear relations may exist, they are more easily identified after subset selection, via appropriate diagnostic procedures. Table 4-5 shows the results of the subset analysis. It is clear, as stated previously, that baseline HF power is a powerful predictor of recovery HF power, which was expected.

The subsets analysis shows several important pieces of information. First, the adjusted value of $R^2\%$ provides the best measure of explained variance, accounting for the number of predictors and the sample size. Thus, for equal amounts of variance explained, adjusted values help to discern simpler models from more complex models. Another statistic listed is Mallow’s Cp, which is used to evaluate the degree of bias in a regression equation. Classically, bias results from poorly fitted models when either the predictors themselves introduce bias based upon their relationship with the criterion, or
when the model is not fully fitted. Because it is not possible to know when a model is in fact fully fitted, the Cp statistic is better used to compare models in their respective introductions of bias, or overall inability to adequately characterize the data. Desirable values of Cp should be close to the number of predictors included in the model. Therefore, a model with three predictors should have a value as close to or below three for its Cp statistic. As presented in table 4-5, subsets two and four have nearly identical values for the $R^2_{adj}$ and Cp statistics. This signifies bias that cannot be removed by simply adding circadian status to baseline as predictors. Based upon the results in the table, it is clear that subset one combines an acceptable Cp value with a high degree of explained variance. While it is comparable to subset three, subset one has fewer variables, and thus is the preferred model. In addition, as mentioned above, the addition of circadian status does not remove any bias in the model and thus its omission is justified. Consequently, subset one was chosen for further analysis. However, to be precise, hypothesis two stated that the regression model would produce an $R^2$ of 0.75 or better. As stated in the previous chapter, estrogen status, baseline HF power, and circadian status were the proposed predictors. This is clearly met by subset three. Yet, as stated above, subset one was chosen for rigorous diagnostic evaluation and as the final litmus test to ascertain whether this regression procedure does indeed provide high predictive power. The reason for the delay in fully evaluating the hypothesis is that it should be tested using the best available model, in order to give credibility. In further discussion, subset one will be referred to as model one, for convenience.

**Diagnostics**

The classical assumptions of regression analysis, tabled by Studenmund (2001), will be addressed individually, using model one.
Assumption One: Linear Coefficients, Correct Model Specification, Additive Error Term

The specification that the coefficients be linear translates to the regression function being displayed as a linear transformation of variables, to produce a dependent variable score. The equation for model one from the subset analysis is certainly linear, with the coefficients being –88.3, 2.27, 0.784, for the intercept, estrogen, and baseline respectively. This assumption is rarely violated, except with special regression procedures evaluating Poisson distributions, growth curves, and other circumstances encountered in the physical and social sciences.

Table 4-5 Best subsets regression results
Criterion variable is recovery. Baseline is used in all models.

<table>
<thead>
<tr>
<th>Subset</th>
<th>$R^2%$</th>
<th>$R^2_{(adj)}$</th>
<th>Cp</th>
<th>s</th>
<th>Included Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>94.2</td>
<td>93.9</td>
<td>3.50</td>
<td>35.2</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>85.1</td>
<td>84.4</td>
<td>68.8</td>
<td>56.4</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>94.4</td>
<td>94.0</td>
<td>4.00</td>
<td>35.0</td>
<td>X</td>
</tr>
<tr>
<td>4</td>
<td>84.9</td>
<td>84.5</td>
<td>68.4</td>
<td>56.2</td>
<td>X</td>
</tr>
</tbody>
</table>

Figure 4-4 Scatter plot matrices of first order effects alone.
Meeting the assumption that the model is correctly specified is never fulfilled, for
two reasons. First, due to the continuing advancements inherent to research, it is always
possible that new variables and measures may be discovered that provide better
modalities for regression uses. Second, there is no absolute way to eliminate the
possibility that chance, or random factors have combined to make any one model appear
correct. For both of these reasons, being certain of model specification is never
guaranteed.

Meeting the assumption of an additive error term is somewhat vague, and not
directly testable. With appropriate diagnostic evaluations, and ensuring the model is the
best available, assuming an additive error term exists is justified. The additivity of the
error term is generally affirmed by homogeneity in the variances of the errors, which is
discussed below under the appropriate assumption subheading. This assumption rarely
invalidates the use of a model and its violation generally leads to prediction intervals
varying in width and precision. If this assumption is not met, it is up to the researcher to
determine if the model adequately performs its prescribed role.

Assumption Two: Error Term has a Zero Mean

A zero mean in the error term is not strictly observable. However, the error term is
a function of residuals generated at each prediction run. Systematic bias in the model will
show as departures from this assumption. However, the differentiation techniques used to
generate least squares estimates will minimize the errors, and given rough linear
relationships, it will produce errors that balance throughout and generate a mean error of
zero. The primary way researchers evaluate and test this assumption is graphically, which
is done by plotting the standardized residuals against the fitted values obtained.
Therefore, a visual examination was made of such a plot. Figure 4-3 shows this
relationship. It is a reasonable statement that the mean value of the standardized residuals is zero. Indeed, the mean standardized residual for the model is –0.00282, which is close to zero.

**Assumption Three: Error Term has Constant Variance, no Heteroskedasticity**

Because graphing the standardized residuals against the fitted values depends on bias present in the whole model, it may not be as efficient as plotting the standardized residuals against each predictor variable. Highlighting how the residuals vary according to the level of the predictor will demonstrate the existence of patterns in the residuals according to the levels of each predictor. If any patterns exist, appropriate transformations can then be suggested.

Figure 4-5 Residuals versus fits for model one, showing relatively equal dispersion of residuals about zero, indicating mean error of zero.

After plotting the residuals from model one against each of the predictors, an apparent curvature was noted for the baseline variable. Also, the residual plot suggested a
heteroskedastic relationship with a megaphone shape. Generally, such distributions require log transformation or perhaps stabilizing maneuvers like the Box-Cox and Box-Tidwell transformations, or possibly weighted least squares analysis. However, after closer examination, the presence of a large residual was detected, shown as a larger point, at the top of the figure 4-4.

Conceptually, a residual, such as that shown in figure 4-4, can force the least squares estimates to minimize the other errors by clustering and pulling them toward the influential point. In effect, slight curvature may result, but this does not always produce heteroskedasticity. When testing the error variance, using the Breusch-Pagan test (identical to the Cook Weisberg test, which was independently developed), a nonsignificant result was attained ($p > 0.15$). Often, the threat of heteroskedasticity poses serious resistance to remediation, and it is very common for it to be present, regardless of transformations or remedial measures. Therefore, it’s testing once all diagnostics were completed, was mandated, and the results of that test will be discussed.

Importantly, heteroskedasticity is not as significant in this investigation because it will serve to increase the standard error of the coefficients, making the tests of the coefficients more conservative. Again, it will not bias the predictive ability of the regression function, so long as the residuals are normally distributed. This is because the regression coefficients are still centered on the true values; therefore, the predicted values do not change. The change is in the width of confidence intervals for each regression coefficient.
Assumption Four: no Serial Correlation Between the Observations

This is generally assessed by the Durbin-Watson statistic. However, this test is often criticized and can be difficult to interpret. In this investigation, there are no reasons for serial correlations to exist between measurements. However, as outlined by Neter (2001), the statistic was generated and compared according to the upper limit of 1.54, using 45 degrees of freedom and a 0.05 significance level. The value obtained from this sample was 1.66, thus the assumption of no serial correlation was not violated.

Assumption Five: Error Terms Distributed Normally

This assumption does not affect the OLS mechanics, however it is needed for most hypothesis testing procedures, and is thus a prudent task. It is also vital to prediction analysis. Prediction of future scores requires knowledge of the standard deviation of the errors around the forecasted variables. Therefore, exceedingly wide confidence intervals
may be present if the error terms are nonnormally distributed. Such effects make
prediction inefficient and invalid from a practical standpoint.

Standardized residuals are generally assessed using normal probability plots and
tests such as the Kolmogorov-Smirnoff test. The normal probability plot shows a heavy
value in the top right hand corner, which corresponded to a standardized residual score of
3.20. While some authors recommend that any value over 2.0 standard errors from the
mean be evaluated, other authors have stipulated that concern should not arise unless
there are values in excess of four standard errors (Ott & Longnecker, 2001). This high
residual is also the same point discussed earlier regarding assumption two,
heteroskedasticity. It was decided to use the Kolmogorov-Smirnoff test statistic to
facilitate this analysis. The test did not reach significance with a \( p \) value in excess of
0.15, and as such there was no remedial action taken at this point in the analysis.

**Assumption Six: Error Terms not Correlated with Explanatory Variables**

This assumption states that clear patterns should not exist between the distribution
of the explanatory variables and the subsequent residuals. Violations of this assumption
indicate that a possible interaction exists, important variables may need to be introduced,
or simply that the model is poorly fitted. When testing this assumption, the residual plot
comparing residuals to baseline HF power was again the suspicious plot. Typically, it
would be expected that no clear relationship exists. However, clearly a curvature appears
to be present. When testing a quadratic model using the residuals from model one
regressed on a second order equation with baseline as the predictor, the quadratic term
did reach significance with a \( p=0.013 \), giving validity to the notion that curvature exists
in the relationship between the residuals from model one and the baseline variable. As
mentioned previously, the errant data point could possibly pull the residuals toward it,
creating this curvature. Up to this point in the analysis, it was not clear that the data point in question should be dropped from the analysis, and actually, there seems to be logical support for its inclusion. Thus, at this point in the analysis, this assumption was the only one not met. As discussed below, this was remedied after the lack of fit test.

**Assumption Seven: Predictor Variables not Linearly Related, no Multicollinearity**

High correlations among the explanatory variables often leads to serious problems, not totally amenable to conventional remedial measures, and quite possibly can remove the ability to carry on with a specific model. This design was conceived with multicollinearity in mind, with the empirical evidence suggesting that estrogen levels and baseline HF power would not be correlated among individuals. Although potential correlations with circadian status were unknown, thus far, circadian status has shown no predictive quality. Referring back to figure 4-2, the scatter plot matrices for estrogen and baseline display no clear relationship between the two.

Also, from a quantitative testing standpoint, one of the most useful diagnostic tests for multicollinearity is the variance inflation factor (VIF). These inflation factors show how much confidence intervals for a given variable are expanded, compared to an optimum situation of no collinearity. This amount is found by taking the square root of the obtained VIF. The most important aspect of multicollinearity is that if variables are correlated with one another, it is often impossible to ascertain if this correlation is consistent from sample to sample. What this means is that bias will exist in the sample at hand, and if this bias is not of equal magnitude in other samples, then the predictive ability of a regression function may be greatly diminished.

For a given variable its VIF is directly proportional to the correlation between that variable and all other variables, wholly or singly. Ideally, VIF values for each variable
are equal to one. In this investigation, both estrogen and baseline HF power have variance inflation factors equal to one. Thus, the model as chosen shows optimal values for collinearity and the assumption is met.

**Other Diagnostic Issues**

There are other issues to consider in the credibility of a regression analysis. The two primary ones considered in this analysis are measurement error and the overall fit test. The assumptions discussed thus far are put in place to give validity to least squares techniques. Even when assumptions are met, and least squares principles are applied, the creation of an acceptable model is not assured. Though measurement error is an important issue, it is commonly ignored in many investigations. Its role has been taken into account in this investigation and will be discussed in practical terms during the next chapter. The most important issue in this section is the lack of fit test, for if a model does not fit, it is imprudent to apply it for future analysis.

Measurement error in regression practice is a tough issue. Ideally, the regression equations used for clinical purposes would be based on blood levels of estrogen present during the day of the exam. However, it is not practical to do such analysis here. Therefore, surrogate measures of estrogen have been utilized; pooled from numerous studies and presented by Yen (2001). There are two issues with measurement error. One is whether these errors should be accounted for and the other is how to account for them if so chosen. The values provided by Yen show a 95% confidence interval that is symmetrically distributed around the predicted values. As put forth by other investigators, if errors are randomly and normally distributed, the investigator may choose not to model these errors, with the belief that the random distribution will serve to alleviate any potential bias. However, there is still the potential to increase the standard
errors of the regression coefficients. This does not hinder the predictive ability of the equation, but it does make it more difficult to determine specific values for each coefficient. Within the scope of this analysis, that does not pose a problem and actually makes the hypothesis testing more conservative. Therefore, it was decided that no modeling of the potential measurement errors would be done.

The second important issue in regression diagnostics is testing the fit of the model. This is perhaps the culminating and deciding test of all assumptions and theoretical work done thus far. Models shown to reach significance in lack of fit tests are essentially void because further testing of hypotheses and development of inferences are not reliable. Therefore, attaining model fit is of utmost importance.

Unfortunately, the model lack of fit test was significant, in practical terms, with a \( p \) value of 0.054. There are several remedial measures that may be taken for such circumstances, such as transformations, refitting a different model, adding polynomial terms, or testing interactions. A model was tested with all possible quadratic terms and all possible interactions. The lack of fit became nonsignificant and the curvature in the residual plots disappeared; however, the VIF’s for each variable reached a minimum of 250 and some were above 1500, making such a choice a poor option. Add to this the constraints on interpretation of polynomial and interaction effects, plus the simplicity of the chosen model, and it is an easy decision to pursue remedial measures for the current set of predictors.

After selecting only one interaction or quadratic term to enter the model at one time, the minimum VIF inflation attained was 38.6, while the curvature remained for all except those that included the quadratic expression of baseline HF power. When the
curvature was removed, the megaphone appearance of the residuals was still apparent in all models. The lone extreme residual discussed previously remained so in all models ranging from a standardized value of 3.27 to 3.38, giving credibility to the thought that it was possibly pulling the regression function in it’s direction. Therefore, leverage values and influence statistics were evaluated.

Leverage and influence values determine how influential a particular observation is, related to the others obtained. Cook’s $D$ was evaluated because it represents how much an observation affects the regression function. The point of interest did not attain significance, according to recommended guidelines (Neter, 2001); however, it was approximately double that of any other value, as shown in figure 4-6. As pointed out by Neter (2001), values clearly above all others, yet not reaching significance according to published guidelines, must be subjectively judged by the analyst. To aid in this decision, values for the Studentized residuals were evaluated. Referring to recommendations from Cohen, Cohen, Aiken & West (2003), in a data set of this size, three or fewer Studentized residuals should exceed a value of 2.0, with only the rarest of circumstances producing values in excess of 3.0. The point in question produced a Studentized residual of 3.78, clearly making it a strong candidate as an outlier. It was determined that this observation would be discarded if no remedial procedures were able to improve the fit.

Several transformations were attempted to remedy the problem. The Box-Tidwell and Box-Cox procedures were utilized and unfortunately, neither increased the fit of the model. Both procedures actually increased the standardized residual of the data point in question, and adversely affected other model parameters such as $R^2_{adj}$, PRESS, and Cp. Additionally, the normality of the residuals was adversely affected as well, which as
mentioned above, is a serious threat to prediction equations. Finally, a model was constructed with a quadratic term for the baseline predictor. Though the fit test became nonsignificant, the improvement in the $R^2$ value was not significant, increasing by only 0.0043.

After discarding the observation the lack of fit test was no longer significant. Additionally, as noted before, there was apparent curvature between the residuals and the baseline variable, indicating a quadratic relationship. This curvature was greatly removed after removal of the suspected case. However, the basic megaphone shape of the residuals did not completely disappear.

![Figure 4-7 Plot of Cook’s D against each observation. Note highest amount associated with observation one.](image)

To formally test whether the heteroskedasticity was significant, the Breusch-Pagan was again employed. The Breusch-Pagan test was designed specifically for residuals with cone shaped distributions. The test statistic obtained conforms to a chi-square
distribution, with degrees of freedom equal to the number of parameters being evaluated. After discarding the potential outlier, this test showed a trend to heteroskedasticity, obtaining a test statistic of 5.841, just slightly less than the critical value of 5.991. This test was carried out with a significance level of 0.05. Though not technically significant, clearly there is practical significance. Yet, as discussed above, heteroskedasticity is not a vital threat to regression analysis used for prediction, and its presence actually made the hypothesis testing more conservative. As such, no additional remedial measures were taken.

Going back to assumption six, that of the residuals not being overtly related to any specific predictor, after removing the outlying data point, the curvilinear trend between residual scores and the baseline variable was greatly removed. A fitted line plot was constructed with the quadratic element tested yet again. The results of this test proved to be nonsignificant. Lastly, another fit test was performed with a $p$ value $> 0.015$.

At this point the regression function was determined fully developed. Formal hypothesis testing results could now begin. The $F$ test for a significant $R^2_{adj.}$ value was carried out with a test statistic of 446.02, attaining a $p$-value $< 0.0001$. The 95% confidence interval for $R^2_{adj.}$ was $95.5\% \pm 0.0002\%$.

**Hypothesis Three**

Hypothesis three stated that the regression coefficients in the final model would each reach significance. In other words, a joint confidence interval for both partial regression coefficients would not include zero for either one. This hypothesis seems trivial, but it is possible to have a significant and properly fitted regression function that includes predictors that may not bear a significant linear effect. Therefore, the calculations were made to produce a joint confidence interval for both estrogen and
baseline, simultaneously, which is more conservative and precise than individual confidence intervals.

The partial regression coefficient for estrogen, with a 95% confidence interval, was determined to be $2.47 \pm 0.676$; that for baseline HF power was $0.779 \pm 0.081$. Obviously, both partial regression coefficients would not include zero in their respective intervals, thus verifying a significant linear relationship, supporting the research hypothesis.

**Prediction Analysis**

Because this investigation was concerned with the prediction of recovery values of HF power, an estimation of the predictive utility of the final regression equation was made. There are no concrete guidelines for the evaluation of predictive ability. However, two procedures are commonly employed. One is the use of the test statistic, PRESS, and the other is shrinkage analysis.

The PRESS statistic is calculated by omitting the $i^{th}$ observation, and using the remainder of the data set to create a regression equation; repeated for all $n$ observations. It is also known as the “leave one out” method, and comes under the umbrella of Jacknife procedures. These equations are then used to predict a score at every $i^{th}$ position. By calculating the residual between the true value and the predicted value at the omitted positions, a sum of squares for the prediction errors is generated and constitutes the PRESS statistic. The PRESS statistic for model one was 44278, which was second best to the model which added circadian status to the equation, which generated a PRESS statistic of 44020, meaning that model one generated a PRESS value 0.59% higher than that by including circadian status. This further supports the idea that adding circadian status does not yield greater results.
When employing Jacknife methodology, the PRESS statistic is again utilized. It is divided by the total sum of squares for the data set, and this ratio is subtracted from the original $R^2$. In this investigation the Jacknife produced a loss of 0.058% yielding a predicted $R^2\%$ of 94.92, compared to the $R^2\%_{adj}$ of 95.50%. Such a small loss was not deemed to be a significant threat to the utility of the final model.

The idea of shrinkage is based on the notion that the $R^2\%_{adj}$ will most likely shrink from the sample it was derived from to the population it is claimed to represent. Obviously, these corrections are theoretically, not empirically based. However, it is widely accepted and routine to calculate measures of shrinkage when evaluating the usefulness of any prediction equation. Stevens has suggested that the best way to estimate this is by use of the Stein formula (Stevens, 2001). In general, if the Cp value is ideal, the predicted $R^2\%_{adj}$ should not drop by an appreciable degree. In using the Stein formula, the estimated $R^2$ value is obtained, given the hypothetical situation that the derived equation could be used on the population it represents. Recalling that the $R^2$ value obtained from model one was 0.955, after use of the Stein formula, the shrinkage was found to be 0.0058, resulting in an $R^2$ value of 0.949, a drop of only 0.6%. It is a matter of subjective judgment regarding the significance of the newly computed $R^2$, and such a small decrease did not appear to lessen confidence in model one’s predictive power. It is interesting to note that the Jacknife adjustment and the Stein formula produced identical results.

An additional means for evaluating the prediction of a model and the influence of a specific predictor is to compare residual values and prediction intervals before and after adding a specific predictor. Because baseline HF power was by far the most strongly
correlated predictor with recovery HF power, a model with baseline HF power only (reduced model) was compared to one containing both baseline HF power and estrogen levels (full model). Table 4-6 shows the comparison.

Though the true significance of estrogen’s presence cannot be determined because there is no clinical population to validate with, the results demonstrate that the addition of estrogen provides a refined prediction equation. By adding estrogen values, the mean prediction interval was reduced by 33%, which obviously leads to greater precision. Also, the mean residual value fell from 54.96 to 29.56, a 46% reduction. The clinical significance of this remains to be seen, and as with all research, replication will be the ultimate test of estrogen’s utility.

Table 4-6. Comparison with and without estrogen, with baseline a constant.

<table>
<thead>
<tr>
<th></th>
<th>Mean confidence interval width</th>
<th>Mean residual value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced Model</td>
<td>47.66</td>
<td>54.96</td>
</tr>
<tr>
<td>Full Model</td>
<td>31.91</td>
<td>29.56</td>
</tr>
<tr>
<td>Percent difference</td>
<td>-33%</td>
<td>-46%</td>
</tr>
</tbody>
</table>

**Supplementary Analysis**

Three additional analyses were carried out in order to gain more information about the extent of recovery attained, the degree of reactivity, and the dichotomization of the menstrual cycle.

The extent of recovery is defined by the ratio of recovery HF power to baseline HF power (R/B ratio). Because it was posited that estrogen enhances recovery, the simple correlation between estrogen levels and the R/B ratio was calculated. Additionally, a scatter plot of the two variables was evaluated to screen for the presence of curvilinear effects. The scatter plot revealed a linear shape. Subsequent regression of the R/B ratio on estrogen values produced highly significant values, with no evidence of lack of fit. Table
4-7 provides the details of this analysis. Clearly, in this sample, increasing levels of estrogen correlated with increases in the degree of recovery.

Table 4-7. Regression of R/B ratio on Estrogen values.

<table>
<thead>
<tr>
<th>N</th>
<th>$R^2_{\text{adj}}%$</th>
<th>$F_{(0.05,1.42)}$</th>
<th>$p$</th>
<th>$R^2_{\text{pred.}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>70.0%</td>
<td>99.18</td>
<td>0.000</td>
<td>0.6749</td>
</tr>
</tbody>
</table>

The degree of reactivity is expressed as the anger recall HF power divided by the baseline HF power, with lower ratios signifying a greater reactivity. When regressing the reactivity values on estrogen levels, a weakly linear function was developed, with the $p$ value for the regression function equating to 0.06, and an $R^2_{\text{adj}}\%$ value of 6.1%, which is hardly significant from a practical perspective. There was a positive correlation between estrogen and the reactivity score, meaning that increasing estrogen levels lead to decreases in reactivity. However, given that only 6.7% of the variance in reactivity could be linked to estrogen values, the author found no practical worth in that assessment.

Lastly, the data set was split into two groups representing the follicular and luteal phases of the menstrual cycle. This reduced the total number of evaluations from 43 to 29. A regression was performed using recovery HF power as the criterion. The predictors were baseline HF power and the dichotomized menstrual phases. The results are presented in table 4-8. It is clear that the explained variance is nearly identical to that obtained by regressing the full data set on baseline HF power alone. Additionally, the test for significance regarding the phase variable was highly nonsignificant, generating a $p$ value of 0.771. The obvious conclusion is that the dichotomization of the menstrual cycle provided no explanatory power for this sample, and the inference taken is that the dichotomization does not provide a useful tool to researchers.
Table 4-8. Regression of recovery HF power on baseline HF power and menstrual phase.

<table>
<thead>
<tr>
<th>N</th>
<th>Recovery HF = -12.3 – 6.1 Phase + 0.837 Baseline HF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$%$_{(adj)}$</td>
</tr>
<tr>
<td>29</td>
<td>85.80%</td>
</tr>
</tbody>
</table>

Predictors

<table>
<thead>
<tr>
<th>S.E.</th>
<th>V.I.F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase</td>
<td>20.63</td>
<td>1.0</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.8366</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Overall Synopsis

A special note must be applied to the high coefficient of multiple correlation, $R^2$. Though it was hypothesized that the coefficient would be high, reaching a value in excess of 0.95 was not expected, especially given the uncertainty in estrogen estimations. However, once realizing that the population chosen for this investigation was healthy, the high correlation is not such a surprise. If it is assumed that the $R^2$ value is too high, there are three possible reasons: 1) chance factors combined in such a way to form the results. This is always a possibility and is the primary reason replication is the statistician’s greatest tool; 2) a population such as the one chosen in this investigation is free of disease and as such, the recovery of HF power is robust, regardless of reactivity; 3) the manipulation used was not vibrant enough in a healthy population to tax the recovery abilities in such a manner so as to produce more variability. More discussion is presented in the following chapter.

Model one was developed based upon *a priori* theoretical considerations and subsequent statistical screening. All assumptions were met and research hypotheses affirmed. Despite the encouraging results from the diagnostics, the most important test of
this equation will be through further analysis against a similar sample. While efforts are
made to construct a design and analysis that is transferable to the general population, it is
impossible to do so, and the final conclusion, as always, rests on replication.
CHAPTER 5
CONCLUSION

The primary question addressed during this investigation was whether the recovery of parasympathetic effect, via the measurement of HF power, could be modeled and predicted by baseline HF power, circadian status, and estrogen level. The answer is yes, but as with most investigations, there are important caveats that require attention and discussion.

The research hypotheses set forth in chapter three were clearly validated in the statistical analysis. It is clear that baseline HF power can vary greatly between individuals and its inclusion in a regression equation is vital, due to its powerful covariate effect. Furthermore, accounting for estrogen levels was shown to be a significant contributor to regression analysis and variance accountability. Given the complexities inherent to behavior and physiology, the simplicity of the model developed in chapter four makes it a valuable starting point for further investigation.

Research Findings and Their Implications

The primary finding was that recovery of vagal power after anger recall in healthy females can be modeled and predicted using baseline vagal power and estimates of estrogen values. Secondary findings included a lack of correlation between the estrogen values and baseline values, and the statistical significance attained by both estrogen level and baseline power in the final model.

As indicated by the results from hypothesis one, baseline power was not correlated with estrogen values. Though estrogen increases HF power throughout a female’s cycle,
in a random sample, the two variables did not strongly correlate between individuals. This is important because it would lead to multicollinearity and greatly hinder statistical modeling and prediction. Given the gravitation toward simpler models, which is reasonably justified, having a model with a minimum of complications is to be preferred. This is particularly so given the number of diagnostic and prognostic tools currently utilized by clinicians in cardiovascular medicine.

Analysis of hypothesis three demonstrated that both estrogen and baseline were significant contributors to the coefficient of multiple determination. This is an important distinction because it is common for the $R^2_{adj}$ value to increase with the addition of predictors; however, such an increase is not always significant. Showing that both predictors were significant contributors to the regression equation solidifies their respective roles and inclusion.

**Practical Implications**

Though the investigation is considered a success because of the ability to generate a regression equation that included baseline and estrogen values, some practical implications are worthy of further discussion. Chief among these is whether or not the estimated estrogen values are superior to the conventional practice of dichotomizing the menstrual cycle into follicular and luteal phases. Also, though estrogen values showed statistical significance in the final model, does their addition lead to any practical significance when added to the covariate of baseline? These issues and others are discussed below.

As shown in the previous chapter, when comparing the use of estrogen values against the dichotomization of the menstrual cycle, estimates of estrogen values were clearly superior. The dichotomization produced an $R^2_{adj}$ identical to that attained by using
baseline power alone, and the menstrual phase factor did not show statistical significance. However, using estrogen estimates provided an $R^2_{adj}$ nearly identical to that obtained in the full model, and clearly superior to the dichotomized model. Therefore, the continued estimation of estrogen values in subsequent investigations is justified.

The next major question is whether or not the addition of estimated estrogen levels contribute, in a practical way, to the regression equation once baseline power has already been accounted for. To answer this question, partial and semipartial coefficients were calculated, as shown in chapter four. Most importantly, almost 50% of the variance not accounted for by baseline alone could be explained by the addition of estrogen estimates. From a practical position, this would be highly valuable, because explanation of variance is key to modeling.

Given that the main future utility of this analysis would rest on the prediction results, further discussion of prediction results is warranted as well.

To determine the effect that estrogen had on the prediction results, separate regressions and 95% prediction intervals were calculated for the 43 observations retained in the final model. As shown in chapter four, adding estrogen to the equation reduced the mean confidence interval for predictions from 47.66 to 31.91, a reduction of 33%, meaning that more precise predictions and decisions can be made from the addition of estrogen. However, because this analysis utilized healthy females with no indication of cardiac disease, it is impossible to determine how much latitude should be given to score fluctuations before clinical significance is attained. Overall, it would be prudent to consider a 33% reduction in the 95% prediction intervals to be clinically useful, as confounds are likely to be present in older populations, most likely having a continuum
of disease and comorbid conditions. Unfortunately, there exists no literary evidence on which to gauge the effectiveness of the prediction limits generated by this regression equation. This is a shortcoming in the practical significance, but does not take away from the scientific merit of this study.

**Secondary Implications**

Most of the literature exploring psychological manipulations and physiological measures focuses on reactivity. There are no consensus statements regarding the variables that influence reactivity. In fact, the pursuit of reactivity explanations has apparently hindered the brain-body-behavior paradigm. As such, no specific *a priori* attempt was made to identify explanations for the degree of reactivity or recovery. However, an exploratory approach was undertaken to evaluate the reactivity and recovery magnitudes.

The degree to which one recovers from stress is conventionally linked to parasympathetic function, most often cited in articles related to physiological stressors such as exercise. In line with this notion, and the previously described research findings showing that estrogen enhances clearance of catecholamines (Kadish, 1998), the degree of recovery attained was a linear function of the estrogen levels. Also, when the recovery/baseline ratio was evaluated using menstrual phase rather than estrogen level, it was found that only a quadratic fit was significant, showing that recovery for those in phase two, with significantly higher estrogen values, was enhanced. It must be remembered that phase two is the middle group of days not categorized as either follicular or luteal. The relative scatter of recovery/baseline ratios in phases one and three (follicular and luteal respectively) highlights the imprecision of phase-based analyses. Thus, just as estrogen levels were found to augment the prediction and modeling of raw
HF power during the recovery phase, they were positively correlated with the degree of recovery attained.

When assessing the degree or reactivity, the anger/baseline ratio, results were less robust. The clearest relation was between reactivity and day within the cycle, quadratic in nature, peaking during phase two, and suggesting that estrogen level may have indeed affected the reactivity. However, visual inspection of the scatter plot between the two variables clearly shows a high degree of variability, and combined with an $R^2_{adj}$ value of only 0.236, it is unwise to attach a large amount of confidence in these findings. Though it is possible that estrogen levels could blunt the decline in HF power during anger recall, without a larger degree of explained variance, concrete conclusions do not seem appropriate.

In the results section, an adjunct ratio was developed to look at the ratio of the recovery to anger ratio. Because participants were expected to relax after the anger recall period, naturally one would assume that HF power increases and such a ratio exceeds one. However, 15 individuals did not increase HF power during the recovery period. A possible explanation for this could be that the individuals continued to ruminate about the event and could not relax. Indeed rumination is speculated to be one of the links between affective state and negative health consequences. Future work may address the time required to reach a specified HF value, or integrate inventories to potentially parse out individual traits in the hopes of explaining a greater degree of variance in the recovery score.

**Agreements and Contrasts with Previous Findings**

Though the results of this work can be seen as extensions of previous work, there are ways in which the findings can be compared to those previously reported.
In general, the emergence of estrogen levels augmenting the extent of recovery, and possibly lessening the decline in HF power during anger recall, are in agreement with physiological studies showing how estrogen can enhance the parasympathetic effects such as decreases in catecholamine release and enhanced catecholamine uptake and breakdown (Hamlet, Rorie & Tyce, 1980), as well as augmentation of parasympathetic autonomic functioning via central locations (Saleh & Connell, 2000). These effects would result in less sympathetic dominance at the neuromuscular level and thus increase the relative degree of parasympathetic effect and HF power. Such increases in parasympathetic effect is commonly seen when postmenopausal women are given estrogen supplementation. Burleson et al. (1998) found that chronic estrogen treatment in postmenopausal women significantly enhanced parasympathetic measures during psychological stress, and blunted the sympathetic measures. In this regard, the present study is in agreement with prior research.

Contrasting, slightly, with previous research is the finding that reactivity could not be significantly associated with either the follicular or luteal phase. Sita and Miller (1996) found that during the follicular phase, females showed decreased reactivity and attributed this effect to alleged chronic increases in estrogen levels during that phase. In this investigation, no such effect for phase was found. Given that the reactivity literature is still rather contentious, it is difficult to speculate on the worth of reactivity findings. Most, if not all, of the reactivity literature dichotomizes the menstrual cycle without any type of biochemical measurement. It can be argued that such broad dichotomization actually leads to a greater degree of measurement error than the estimates used in this investigation. This claim being made based on the fact that the established fluctuations
for estrogen do not prove to be statistically different when comparing the daily values of the follicular and luteal phases.

**Limitations and Future Directions**

The main limitation of this investigation is the use of estrogen estimates rather than biochemically determined values. However, given the consistency of young females to follow a prescribed pattern of hormonal levels, the lack of consensus regarding whether or not absolute values of estrogen or relative values of estrogen are most important, and finally the precision and significance of estrogen values in the regression equation, it seems likely this error is not of sufficient magnitude to invalidate future research that utilizes estrogen values and estimates. However, blood or saliva measurements of estrogen, combined with indices marking status within the menstrual cycle would certainly be an improvement.

Less significant to this specific investigation, a limitation is the use of a healthy cohort of females, with no longitudinal tracking of health status. Longitudinal tracking would permit a logistic regression analysis to evaluate what factors and values led to disease and death. Without such tracking and analysis, it is unknown if the addition of estrogen values to psychological and physiological measures will affect epidemiological data and subsequent treatment protocols.

Future research should pick up where this investigation has ended, namely, the continued analysis of recovery utilizing measured or estimated values of estrogen. In addition, researchers should begin longitudinal analysis so that disease trends and the factors influencing those trends can be tracked and correlated. A broader scope of demographics could be investigated to determine if other predictors need to be added in order to generate a clinically useful tool.
Lastly, as briefly mentioned above, future work should involve well targeted psychological inventories, matched to task conditions, and explore the role of traits and rumination in the variance of recovery scores. Such work could provide useful predictors and eventually mechanistic evidence linking behavior and health.

**Design Considerations**

No research endeavor is perfect, and every design has its pitfalls. Though many psychologists refer to the Hawthorne effect to describe the fact that individuals may alter their behavior while under observation, it is actually the phenomenon of individuals amending their behavior to meet expectations, whether these expectations are real or imagined. Indeed, the Hawthorne effect has been criticized and questions raised regarding its mere existence (Adair 1984 and Parsons 1974). It would be expected for many in the psychology community to cite the Hawthorne effect as a means to stump an investigator and raise speculation concerning the validity of results, as is popularly done, but it would be incorrect to do so in the context of this investigation. More correctly, it would be wise to speculate that reductions in HF power could be due to anxiety, tension, and the general feeling of unease possible for someone asked to sit still for approximately 20 minutes. This is certainly different from the expectation and behavior linkage proposed by the Hawthorne effect.

Concerns over anxiety levels within the laboratory environment is especially prudent considering the participants were aware that their respective heart rates were being analyzed. In clinical environments, it is often the case that an individual’s blood pressure will be artificially elevated when it is measured by a health care professional, due to the awareness the general public now has concerning the importance of blood pressure in cardiovascular health. This phenomenon is referred to as the “white coat
effect”. Similarly, in this investigation, the participants were aware that cardiovascular parameters were being measured, and combined with the simple task of remaining still in a strange environment, it is logical to assume an increase in anxiety levels from start to finish. Increases in anxiety often produce decreases in HF power. Therefore, the trend of seeing HF power decline from baseline, to anger recall, to recovery would be expected. However, this trend was not witnessed, as HF power tended to increase when going from anger recall to the recovery phase. Such does not alleviate potential confounds inherent to the laboratory environment, and there still exists the possibility that anxiety levels may blunt the recovery. The addition of a control group in future studies may provide insight into this potential problem. Unfortunately, a control group would only account for potential anxiety levels minus the manipulation of a recall task. Therefore, truly accounting for the influence of the laboratory environment may not be possible.

While methodological and design considerations must be addressed in future investigations, this author does not believe such potential pitfalls have weakened the results and conclusions of the current investigation.

**Final Conclusion**

The primary finding during this work was that accounting for estrogen values is a productive maneuver, generating a more powerful modeling and prediction analysis. The simplicity of the developed model should lend itself to further study and allow for additional predictor variables and more varied populations. While the clinical utility of the protocol is unknown, additional studies with diseased and healthy participants, along with creative methods to account for potential confounds such as the “white coat effect”, should shore up this issue. Eventually, the arena of psychological testing will be brought into clinical relevance for disease states outside of the psychiatric field. The simplicity of
this study, combined with the strength of its results, should only bolster the effort to integrate psychological manipulations into conventional medical diagnostic and treatment models.
APPENDIX A
STANDARD ESTROGEN VALUES

Standard values from Yen (1999) were restructured to show each day’s estrogen value as a percentage of the maximum estrogen value during the menstrual cycle. Figure A-1 provides a graphical display of the estrogen fluctuations.

Figure A-1: Estrogen values shown as percentage of maximum.
Table A-1: Estrogen values for each day, shown as percentage of the maximum.

<table>
<thead>
<tr>
<th>Day</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estrogen</td>
<td>38</td>
<td>33</td>
<td>25</td>
<td>30</td>
<td>30</td>
<td>45</td>
<td>35</td>
<td>38</td>
<td>38</td>
<td>45</td>
<td>55</td>
<td>75</td>
<td>100</td>
<td>88</td>
<td>78</td>
</tr>
<tr>
<td>Day</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>Estrogen</td>
<td>68</td>
<td>55</td>
<td>58</td>
<td>68</td>
<td>53</td>
<td>53</td>
<td>60</td>
<td>58</td>
<td>55</td>
<td>53</td>
<td>45</td>
<td>35</td>
<td>33</td>
<td>30</td>
<td>35</td>
</tr>
</tbody>
</table>
APPENDIX B
QUESTIONNAIRES ONE AND TWO

Questionnaire One

Four digit number: ______________

Please make up a four-digit number of your choosing and write it in on the top right hand corner of this page, where the blank line is. This number will be used on a subsequent questionnaire if you choose to participate in this study. So, please remember this number and remember to annotate it on the second questionnaire, exactly as on this one. This number system is intended to protect your identity and not allow any individual to have an opportunity to connect your personal identity with any of the information you provide. Please use a number that does not identify you in any way. For instance, please do not use the last four digits of your social security number.

Please answer yes or no to the questions below. However, you are not obligated to answer any of the questions.

1. Have you refrained from smoking of any kind for the last six months? __________
2. Are you currently using any prescribed medications other than prescribed birth control? __________ If yes, please let the investigator know what these medications are. You are not obligated to inform the investigator what medications you are on. Some medications may affect your behavior or your heart and that is why the investigator would like to know what medications you are taking, if any. Again, you are not obligated to provide any of this information.
3. Are you currently using any prescribed means of birth control? __________
4. Does anyone in your immediate family have cardiovascular / heart disease? For instance, stroke, heart attack, hypertension, coronary artery disease, or others. If you are unsure, please feel free to ask the investigator and he will clarify this for you. __________
5. Have you had regular menstrual cycles for the previous six months? In this investigation, regular menstrual cycles are those that take place within 35 days of the previous cycle. __________
6. Have you been diagnosed with, or informed by a physician that you have hypertension? ________
Appropriate answers for inclusion

Questions one and five required a “yes”; all other questions required a “no”.

**Questionnaire Two**

**Questionnaire 2**  Four-digit number: _______________________

This questionnaire requires the same four-digit number that you created for the first questionnaire. Please put that number in the space provided at the top right hand corner of this form.

When you arrive at the laboratory, you will be asked to have your blood pressure taken. In order for the investigator to not connect your identity with your data, you will be asked to write down the results of your blood pressure measurement in the blanks below. The investigator will notify you about which number goes in right spot.

For this investigation, it is asked that you provide two other pieces of information. The first relates to the number of hours you have been awake prior to arriving at the laboratory. You will see a line labeled “Time of awakening”. Please write down the time you awoke on the day you report to the laboratory, making sure to include “a.m.” or “p.m.” as appropriate. Below that you will see the title, “Time of reporting”. Please write down the time you arrive at the laboratory in this blank, using “a.m.” or “p.m.” as appropriate. The second piece of information relates to your status within the menstrual cycle. Estrogen values fluctuate from day-to-day during the menstrual cycle. Because the modeling of these values is important to this investigation, an accurate accounting is required. You may fill in this blank at any time prior to arriving at the laboratory. You will see a heading titled, “Cycle day”. This refers to the day that you began your last menstrual cycle. The beginning of the menstrual cycle is marked by the first day of menstrual bleeding. If you are currently menstruating, the date you began is the date you would mark on the line. Please do your best to make this as accurate as possible.

When you are completely done with the study, you will be asked place your questionnaire face down. The investigator will check to make sure the numbers on the questionnaires one and two match and the data will then be transferred to a spreadsheet. The data will be double-checked for accuracy, and then the questionnaires will be shredded, removing any connection between the information you provide, your heart rate measurements, and your identity.

**Blood Pressure:**

SBP __________

DBP __________

**Time of awakening:**

________________________

**Time of reporting:**

________________________

**Cycle date:**

________________________

**Date of arrival at the lab:** __________
LIST OF REFERENCES


Decety, J., (1996). Do imagined and executed actions share the same neural substrate? *Cognitive Brain Research, 3*, 87-93


Morecraft, R.J., Van Hoesen, G.W. (1993). Frontal granular cortex input to the cingulate (M3), supplementary (M2) and primary (M1) motor neurons in the rhesus monkey. *Journal of Comparative Neurology*, 337, 669-689


Schultz, W., Tremblay, L., Hollerman, J.R. (2000). Reward processing in the primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex*, 10, 272-284


Trejo, J.L., Carro, E., Torres-Aleman, I. (2001). Circulating insulin like growth factor I mediates exercise induced increases in the number of new neurons in the adult hippocampus. *Journal of Neuroscience*, 21, 1628-1634


BIOGRAPHICAL SKETCH

Charles Todd Sullivan was born on July 30th, 1971. He is a Navy veteran. In 1997, he received an associate’s degree from Santa Fe Community College, in Gainesville, Florida. He received a Bachelor of Science in Exercise in Sport Sciences (BSESS) degree from the University of Florida in 2001, and continued on at the University of Florida to pursue a Master of Science in Exercise and Sport Sciences, beginning in 2001. His BSESS has a specialization in exercise physiology, while the MSESS has a concentration in exercise and sport psychology.