Psychophysiology, 39 (2002), 292–302. Cambridge University Press. Printed in the USA.

DOI: 10.1017.S0048577201393046

Cardiovascular patterns associated with threat and challenge appraisals: A within-subjects analysis

KAREN S. QUIGLEY,[a] LISA FELDMAN BARRETT,[b] AND SUZANNE WEINSTEIN[c]

[a]Department of Psychiatry, University of Medicine and Dentistry of New Jersey and VA Medical Center, East Orange, New Jersey, USA
[b]Department of Psychology, Boston College, Boston, Massachusetts, USA
[c]Center for Excellence in Learning and Teaching, Pennsylvania State University, University Park, Pennsylvania, USA

Abstract

Previous studies demonstrated distinct cardiovascular patterns associated with threat and challenge appraisals for groups of participants. We extend these results by assessing whether appraisals continue to be associated with these cardiovascular response patterns within an individual as appraisals change. Participants completed four verbal mental arithmetic tasks for which they made appraisals before and after each task. Cardiac reactivity and total peripheral resistance (TPR) were calculated for the first and last minutes of each task, and the number of responses and percent correct were measured for each task. In line with our prediction, pretask appraisals were related to some task-related cardiac responses across the four tasks. In addition, task-related cardiovascular reactivity and behaviors both influenced appraisals following the task. Our findings suggest that an idiographic analysis of appraisals, cardiovascular physiology, and task-related behaviors provides a richer understanding of the appraisal process and reveals sex differences deserving further assessment.

Descriptors: Mental arithmetic, Hierarchical linear modeling, Heart period, Cardiac output, Preejection period, Total peripheral resistance

Adaptation involves constant assessment of environmental demands and selection of effective strategies for dealing with those demands. Lazarus and Folkman (1984) presented a theory of stress and coping based on the central tenet that both stressors and coping ability must be understood from the perspective of each individual's appraisal of the environment. In response to an environmental event, primary appraisal occurs when individuals assess whether the event will be impactful or stressful, and secondary appraisal occurs when individuals assess whether or not they have the resources to cope with the event. Recently, several studies by Blascovich, Tomaka, and colleagues have reformulated the appraisal concept using the ratio of primary to secondary appraisal to reflect threat and challenge appraisals (Blascovich, Mendes, Salomon, & Hunter, 1999; Tomaka & Blascovich, 1994; Tomaka, Blascovich, Kelsey, & Leitten, 1993; Tomaka, Blascovich, Kibler, & Ernst, 1997). In this modification, a threat appraisal is made when the individual perceives environmental demands as excessive relative to his or her resources or ability to cope (i.e., primary appraisal is large relative to secondary appraisal). In contrast, a challenge appraisal is made when the individual perceives that coping abilities are high relative to environmental demands. Blascovich and Mendes (2000) have further modified the threat and challenge conceptualization. They now refer to demand and resource appraisals rather than primary and secondary appraisals, respectively. Demand appraisals are postulated to derive from cognitive assessments of potential harm and required future effort, but also affective and perceptual cues that inform this assessment. Similarly, resource appraisals are cognitive assessments of putative available resources that are influenced by affective and perceptual factors.

In a series of studies, Blascovich, Tomaka, and colleagues demonstrated that threat and challenge appraisals led to distinct patterns of cardiovascular response during a goal-relevant, motivated performance task (Blascovich et al., 1999; Tomaka & Blascovich, 1994; Tomaka et al., 1993, 1997). Changes in cardiovascular

The authors would like to thank Robert Kelsey and Bill Guethlein for the use of their program for acquiring and processing the impedance cardiographic data, Gary Berntson for the use of his program for processing heart period data and John Ernst for advice on the arithmetic task. In addition, the authors thank the following research assistants who helped with data collection and reduction: Tabitha Charles, Yvonne Chia, Cory Clouser, Sangeeta Day, Heather Deaner, Mike Dudis, Tim Gallagher, Sean Gibbs, Shira Goldstein, Stephanie Orner, Jessica Stephens, Amy Sutter, Audrey Wang, Dave Weikert, and Connie Witmer. Preparation of this article was facilitated by a grant from the National Science Foundation to Lisa Feldman Barrett (SBR 9727896) as well as a grant from the Research and Graduate Studies Office of the College of Liberal Arts at Pennsylvania State University to Karen Quigley and Lisa Feldman Barrett.

Note: The ordering of the first two authors for this manuscript was determined randomly. This work represents equivalent contributions by these two authors.

Address reprint requests to: K.S. Quigley, VA Medical Center, 385 Tremont Ave. Stop #127B, East Orange, NJ 07018, USA. E-mail: quigl@umdnj.edu. L. Feldman Barrett, Department of Psychology, McGuinn Hall, Boston College, Chestnut Hill, MA 02167, USA. E-mail: barretli@bc.edu.

by larger decreases in heart period (HP; i.e., larger increases in heart rate) and preejection period (PEP), larger increases in cardiac output (CO), and a shift from increased to decreased (or no change in) total peripheral resistance (TPR). In addition, we explored the relationship between pretask appraisals and respiratory sinus arrhythmia (RSA) reactivity.

The Pretask Appraisal Hypothesis: Task Demand and Performance Relationships

Following from previous group-level findings (Tomaka & Blascovich, 1994; Tomaka et al., 1993), we predicted that pretask appraisals would be related to task demand and performance, such that greater appraisals of threat would be associated with lower demand (fewer problems attempted) and poorer performance (lower percent correct).

The Reappraisal Hypothesis: Posttask Appraisals

Following previous results showing that eliciting physiological patterns like those produced by threat and challenge was not sufficient to produce threat and challenge appraisals (Tomaka et al., 1997), we predicted that physiological changes in the final minute of the arithmetic tasks would not relate to posttask appraisals. However, consistent with the suggestion that experience informs reappraisals (Blascovich & Tomaka, 1996; Tomaka et al., 1993), we predicted that having just attempted more problems and having more correct responses would be associated with greater challenge appraisals. This hypothesis is based on our expectation that task performance was likely the most salient information available for making a posttask appraisal.

The Reappraisal Hypothesis: Pretask Appraisals of the Upcoming Task

To further explore the reappraisal process, we explored how physiological changes, task behaviors, and posttask appraisals from a previous arithmetic task influenced pretask appraisals of an upcoming task.

Methods

Participants

Seventy-four psychology undergraduates at the Pennsylvania State University (37 women; ages 17–26, mode = 18 years) participated in this study for course extra credit. Potential participants were excluded if they reported cardiovascular or respiratory illness, used recreational drugs, were more than 35% over or under ideal weight for height, or if their parents had a history of hypertension. In addition, participants were rescheduled for later testing if they had ingested alcohol in the preceding 12 hours, or were ill when they arrived at the lab.

Measures

Physiological. We recorded the ECG, the impedance cardiogram (ZCG), blood pressure, and respiration. The ECG and ZCG were recorded using a tetrapolar band electrode configuration with bands placed around the neck and torso following the recommendations of Sherwood et al. (1990). The inner two recording electrodes were placed around the base of the neck and around the thorax at the level of the xiphisternal junction. The outer two current electrodes were placed at least 3 cm above and below the inner recording electrodes. These current electrodes passed a 4-mA, 100-kHz alternating current across the thorax. Basal thoracic impedance (Z_0), the first derivative of the change in thoracic impedance (dZ/dt), and the ECG were measured by a Minnesota Impedance Cardiograph (Model 304B, IFM). Blood pressure was recorded using a Dinamap (Model 1846 SX; Critikon) automated oscillometric blood pressure monitor. The ECG and ZCG signals were passed to a microcomputer with an A/D converter (12 bit) with ECG and dZ/dt sampled at 500 Hz, and Z_0 sampled at 250 Hz. Digitized data were stored for off-line reduction and analysis.

Impedance-derived physiological measures were reduced using software that permitted visual inspection of impedance cardiographic waveforms, and provided computer-aided event detection and ensemble averaging (Kelsey & Guethlein, 1990). ECG and ZCG data were inspected for movement artifact, and affected beats were not used in the ensemble averages. Movement artifact affected less than 1% of the data. One-minute means were derived from the ensemble-averaged waveforms where fewer than 1% of ensemble averages were comprised of 45 or fewer heart beats. In contrast to the analyses of heart rate change conducted by Tomaka et al. (1993), we instead chose to assess heart period, or the time between successive heart beats. Heart period demonstrates a more linear relationship with the underlying autonomic changes that likely mediate short-term cardiac changes than does heart rate (Berntson, Cacioppo & Quigley, 1995). Thus, across different initial basal heart periods, equivalent changes in autonomic input to the heart will result in nearly equivalent changes in heart period, a relationship which does not hold for heart rate.

To derive cardiac output, stroke volume was calculated from the ensemble averaged waveform using Kubicek's equation (Sherwood et al., 1990). Cardiac output then was calculated as the product of stroke volume and heart rate (the inverse of heart period), and total peripheral resistance (in dyne-seconds$\cdot$cm^{-5}) was calculated as: (mean arterial pressure/cardiac output) $\times$ 80. Preejection period was taken as the time between the Q wave of the ECG and the B point of the dZ/dt waveform. It should be noted that PEP is not a measure of changes in sympathetic effects on chronotropic (rate) function, but rather inotropic (contractility) function. However, a previous study using very similar methods showed that task-induced changes in PEP were strongly related to sympathetically mediated changes in HP (Berntson et al., 1994). Respiratory sinus arrhythmia (RSA) was derived using the method of Porges and Bohrer (1990; MXEdit, ver. 2.21, Delta Biometrics) for each minute of heart period data.

Self-report. Pretask appraisals were assessed immediately following instructions for each task using the method outlined by Tomaka et al. (1993). Primary appraisal was assessed by asking the participant to "Please rate how stressful you think the upcoming task will be on a scale of 1 to 5 where 1 is not at all stressful and 5 is very stressful." A 5-point Likert scale with anchors at 1 = *not at all*, 3 = *moderately*, and 5 = *very much* was placed on the wall next to the participant to aid in making ratings. Secondary appraisal was assessed by asking the participant to "Please rate how well you think you can cope with the upcoming task on a scale from 1 to 5 where 1 is I cannot cope at all with the task, and 5 is I can cope very well with the task." Posttask appraisals were assessed immediately following each task by asking participants to rate the stressfulness of the preceding task and how well they coped with it. The same 5-point Likert scale was used as for pretask appraisals.

Task demand and performance. Task demand was operationalized as the number of subtraction attempts made for each arithmetic task. Task performance was measured as the percent of correct responses for each task. Kelsey and colleagues (1999, 2000) demonstrated that error rates do not vary across repetitions of a self-paced serial subtraction task; thus the percent correct response measure here is most likely affected by response rate, not error rate.

Procedures

Participants were initially informed only that they would be performing a counting task in order to prevent early appraisals of the task as particularly stressful. Following the initial briefing, participants gave informed consent, and filled out a health questionnaire for screening purposes. Participants then completed a battery of personality questionnaires for approximately 10 min. Following completion of the measures, the participant's weight and height were measured, and the electrodes and blood pressure cuff were placed.

For testing, the participant sat quietly in a sound-attenuated testing room in an upholstered chair. The experimenters sat in another room, and communicated with the participant via an audio system. The participant sat quietly for a 10-min electrode stabilization period during which a single blood pressure measurement was taken to check that resting blood pressure did not exceed 150/90 mmHg. Following electrode stabilization, a 4-min resting baseline measurement of the ECG and ZCG was recorded. In addition, a blood pressure recording was made once per minute during the 4-min baseline. After baseline recordings, participants were instructed that they would be performing a mental arithmetic task that would require completing serial subtractions and reporting the answers aloud. Participants were told that they would be given the three-digit number 725 to subtract from, and the number 7 to subtract by (i.e., the subtrahend) and that they were to continue subtracting for 4 min. Immediately following these instructions, participants were asked to give primary and secondary appraisal ratings, and were told to work as quickly and accurately as possible because their answers were being recorded. During the task, the answers given by the participant were recorded by an experimenter. The experimenter did not speak to the participant during this task unless the participant stopped for more than about 5 s, at which point the participant was prompted to please continue with the last number reported. Physiological measures were recorded for the entire 4-min task.

Following the initial 4-min task, participants gave posttask appraisal ratings, and the procedure was repeated (beginning with a second 4-min baseline recording). Task 2 was identical to Task 1 except that the participant was given a different number to subtract from. As before, physiological, appraisal, and performance measures were recorded for the 4-min task epoch, and following the task, posttask appraisals were made.

A second set of two baselines and two tasks (Tasks 3 and 4) followed this first set of tasks. To ensure that the task remained impactful and enhance the likelihood that individuals would make different appraisals across tasks, Tasks 3 and 4 were calibrated to the participant's performance. Instructions for Tasks 3 and 4 indicated that the participant would be subtracting from a three-digit number by a smaller number, and that both numbers would be changed several times during the task. The number being subtracted from, as well as the subtrahend, were changed at the beginning of each minute, at which time the difficulty of the task was adjusted so that difficulty remained moderately high for each participant (e.g., a subtrahend of 7 in the first minute changed in the second minute to 3, 4, 8, 11, or 13 based upon performance). Furthermore, participants received immediate feedback such that when they made a mistake, the experimenter indicated this and provided the participant with the last correct answer given, and asked her or him to continue. As with the previous tasks, participants were asked to work as quickly and accurately as possible. Following Task 4, a final 4-min baseline was recorded, and, using instructions like those given for Task 4, a final pretask appraisal for a presumed fifth task was made. Following this pretask appraisal, the participant was informed that there was not another task forthcoming, and that the experiment was complete. Participants were then disconnected from the recording equipment, debriefed, thanked and given credit for participating.

Data Reduction and Analysis of Physiological Measures

Mean CO, PEP, HP, TPR, and RSA were calculated for each minute of the rest and task periods. A measure of baseline physiological levels was computed by averaging the recordings for the 4 min of the rest period immediately preceding each task. A 4-min baseline was computed to provide a maximally reliable assessment of baseline function. Tests using baseline physiological measures revealed no relationship between pretask appraisal ratings and baseline physiological measures across all four tasks. Following from analyses reported by Tomaka, Blascovich, and colleagues (Tomaka & Blascovich, 1994; Tomaka et al., 1993, 1997), psychophysiological reactivity scores were computed as differences from baseline. Two sets of reactivity scores were computed. First-minute reactivity scores were computed by subtracting the average score for the 4-min baseline from the score for the first minute of each task epoch. Last-minute reactivity scores were computed by subtracting the average score for the 4-min baseline from the score for the last minute of each task epoch. We decided that the advantages of the proximity of the first- and last-minute reactivity scores to the pretask and posttask appraisals, and the fact that previous investigators of this phenomenon also used the first task minute, outweighed the disadvantage of poorer reliability for the TPR reactivity scores based on 1-min epochs.

General data analysis strategy. This study conformed to a multilevel data structure because multiple lower-level observations (physiological measures, appraisals, task demand, and performance scores) were measured on a task by task basis and nested within upper-level units or participants (Kenny et al., 1997). Hierarchical Linear Modeling (HLM; Bryk & Raudenbush, 1992; Bryk, Raudenbush, & Congdon, 1996) was used to analyze the current data set because it allowed us to analyze within-subject (lower-level) and between-subject (upper-level) variation simultaneously, thus enabling us to model each source of variation while taking the statistical characteristics of the other level into account. HLM allowed us to estimate the average relationship between lower-level variables (e.g., TPR reactivity for the first task minute and pretask appraisal ratios) as well as the amount of variation that individual participants display around this average. Upper-level variables that describe participants (in this case, sex of participant) were used to model this observed variation (e.g., allowing us to determine whether the pretask appraisal–TPR first minute reactivity relationship is stronger for some individuals than for others).

HLM does not treat the two levels of data (within a person over tasks and across persons) as separate for purposes of analyses, but instead statistically links them together, simultaneously estimating effects at both levels. Lower level (Level-1) variance is modeled taking into account the variance at the upper level (Level-2) through the use of empirical Bayesian estimation techniques. With empirical Bayesian estimation, the Level-1 regression coefficients for each person are based on the maximum-likelihood solution for a particular participant and the overall solution for all participants (Bryk et al., 1996; Kreft & de Leeuw, 1998). Similarly, Level-2 variation is estimated while accounting for Level-1 variation. This is accomplished through (1) treatment of Level-1 regression coefficients as random (rather than fixed), along with (2) use of weighted least squares procedures to estimate the Level-2 model (Bryk et al., 1996; Kreft & de Leeuw, 1998). Level-1 regression coefficients from each participant may or may not be equally efficient for purposes of estimating an average relationship among the predictor and outcome variables across all participants. For example, some participants may have missing data, or some may have outliers, causing person differences in the standard errors of the estimate corresponding to their Level-1 coefficients. By treating the Level-1 coefficients as random and calculating the Level-2 model using weighted least squares, the Level-1 coefficients with lower standard errors of the estimate are given more weight (and those with larger standard errors less weight) when computing the average regression weight. Moreover, because the two levels of variance are being modeled simultaneously to estimate regression parameters in HLM, and because parameters are being tested against the null hypothesis in a multivariate fashion, familiar relationships between regression weights and t tests seen in OLS estimation procedures do not hold for HLM (for more information, see Bryk & Raudenbush, 1992, Chapter 3). Degrees of freedom for all t tests from HLM analyses were 71.

Results

We first briefly present baseline data, followed by the analyses for the pretask appraisal hypotheses, and then those for the reappraisal hypotheses. Analyses associated with our exploratory questions about sex follow the main analyses within each section where appropriate. For the sake of comparison, we also assessed the data using the between-subjects approach of Tomaka, Blascovich, and colleagues using their analytic strategy of multivariate analysis of variance (MANOVA) to test for appraisal group differences.

Baselines

Baseline scores were estimated using an HLM analysis with the following Level-1 model:

$$Baseline_{ij} = b_{0j} + r_{ij}, \quad (1)$$

where $Baseline_{ij}$ is participant j's cardiovascular baseline score for the first minute of the ith task, b_{0j} is the average baseline score for participant j, and r_{ij} is a within-subjects residual component. There was significant variability in all baseline estimates across individuals, so we examined whether sex of participant accounted for any of this variance. Sex of participant was dummy coded (male = 0) and added to the between-subjects level (or Level-2 aspect) of the model, as follows:

$$b_{0j} = b_{00} + b_{01}\,(\text{sex}) + u_{0j} \quad (2)$$

where the b_{00} term is the average baseline score for men, the b_{01} term indicates the difference between the magnitude of the b_0 coefficients for male and female participants, u_{0j} is the between-subject error term and represents the degree to which the Level-1 (i.e., lower level) parameters for participants continued to vary even after sex was taken into account. A parallel set of analyses was run with sex dummy coded where female = 0, so the significance of the average baselines for female participants could be computed.

Baseline values are as follows (men vs. women): HP = 929.6 ± 21.4 ms (or 64.5 bpm) versus 845.8 ± 22.1 (or 70.9 bpm), PEP = 98.7 ± 2.2 ms versus 94.1 ± 1.9, RSA = 7.3 ± 0.18 l_n band variance versus 7.3 ± 0.14, CO = 7.37 ± 0.31 l/min versus 7.00 ± 0.24, and TPR = 930.1 ± 48.9 dyne-s·cm^{-5} versus 937.1 ± 32.8). Mean baseline values across the four tasks never fell outside the 95% confidence interval bounds on any variable. The only significant effect of sex on any baseline measure was that men had longer baseline HP (lower heart rate) than women, $t = 2.69$, $p < .01$.

The Pretask Appraisal Hypothesis: Cardiovascular Reactivity Relationships

The within-subjects level of the data analytic model estimated the relationships between pretask appraisal ratios and first task minute reactivity scores using the formula

$$FTM_{ij} = b_{0j} + b_{1j}PreA_{ij} + r_{ij} \quad (3)$$

where FTM_{ij} is participant j's autonomic reactivity score for the first task minute of the ith task, b_{0j} is the average first task minute reactivity score for participant j (when his or her pretask appraisal is at its mean), b_{1j} is the magnitude of the relationship between first task minute reactivity scores and pretask appraisal ratios for participant j, $PreA_{ij}$ is participant j's (centered) pretask appraisal ratio for the ith task, and r_{ij} is a within-subjects residual component. The between-subjects level (or Level-2 aspect) of the model allowed us to assess the average reactivity for the first task minute (the average value of b_{0j}) as well as the average relationship between reactivity for the first task minute and pretask appraisal ratios (the average value of b_{1j}), as follows:

$$b_{0j} = b_{00} + u_{0j} \quad (4)$$

$$b_{1j} = b_{10} + u_{1j} \quad (5)$$

Table 1. *Average Associations Between Pretask Appraisals and First Task Minute Reactivity*

	Average Relationship					Variation in Relationship		
	b_{10}	*SEM*	B_{10}	t	p	*SD*	χ^2	p
HP	32.02	9.68	0.19	3.31	.001	12.80	81.19	.17
PEP	2.39	0.88	0.15	2.69	.008	2.77	112.09	.001
RSA	0.20	0.12	0.10	1.61	.108	0.25	85.57	.10
CO	−0.17	0.10	−0.09	1.68	.092	0.17	137.70	.001
TPR	−5.06	20.28	−0.02	0.25	n.s.	60.08	102.86	.007

Note: bs indicate unstandardized regression coefficients and Bs indicate standardized coefficients. SEM is given for the bs. The degrees of freedom for the χ^2 = 70. HP = heart period. PEP = preejection period. RSA = respiratory sinus arrhythmia. CO = cardiac output. TPR = total peripheral resistance.

where the upper-level estimate, b_{00}, represents the average reactivity score for the first task minute (across all four tasks) for all participants, and the upper-level estimate, b_{10}, represents the average relationship between first task minute reactivity and pretask appraisal ratios for all participants, u_{0j} and u_{1j} are between-subject error terms and represent the degree to which the Level-1 (i.e., lower level) parameters for participants varied around the average for each coefficient. Although we modeled the average of the intercept (b_{0j}) for all analyses, we do not report them because they were not relevant to our hypotheses. Similar HLM analyses were used to estimate the relationship between pretask appraisals and all reactivity scores, and the results are presented in Table 1. In these analyses where there is no Level-2 moderator, the Level-1 predictors were grand-mean centered. Where there was a Level-2 moderator, the Level-1 predictors were group-mean centered (Hofmann & Gavin, 1998).[1]

HLM analyses revealed that pretask appraisals were related to the cardiac reactivity measures (HP, PEP, and RSA) and to cardiac output (CO). The parameters listed in Table 1 represent the average within-subject relationships among the lower level variables. As pretask appraisals moved from threat towards challenge (i.e., appraisal ratios became smaller), both HP and PEP shortened more, RSA showed a marginally larger decrease, and CO showed a marginally larger increase. Taken together, these findings are as predicted and indicate that as individuals became more challenged, their hearts beat faster, most likely due to more sympathetic activation (suggested by the greater shortening of PEP) and more parasympathetic withdrawal (greater decreases in RSA). Although not all of the effects reached conventional levels of significance, they were all in the predicted direction.

The finding that activity in both branches may have contributed to the relationship between pretask appraisals and cardiac reactivity led us to assess the multivariate relationship between pretask appraisals and both PEP and RSA. Here we used a multivariate hierarchical linear model where pretask appraisal served as the predictor, and RSA and PEP were used simultaneously as criterion variables (MacCallum, Kim, Malarkey, & Kiecolt-Glaser, 1997; Raudenbush, Brennan, & Barnett, 1995). We did not use multivariate HLM procedures for all reactivity analyses because HLM is unable to deal with highly collinear variables such as characterize many cardiovascular measures in the current study (except PEP and RSA). In that multivariate analysis, only the pretask appraisal–PEP relationship remained significant, $B_{10} = 2.74$, $SEM = 0.69$, $t = 3.97$, $p < .001$.

There was significant variation in the magnitude of the relationship between pretask appraisals and PEP, CO, and TPR change (i.e., significant variation in the size of the b_1 coefficients; Table 1). This variation indicated that the relationship between these cardiovascular measures and pretask appraisals varied in size, and it was likely the relationships were statistically significant for some individuals, but not for others. We investigated the degree to which sex of participant accounted for this variability. One HLM model was run for each reactivity score. Sex of participant was dummy coded (male = 0) and added to the Level-2 HLM model for each analysis in the following manner:

$$b_{1j} = b_{10} + b_{11}\,(\text{sex}) + u_{1j} \tag{6}$$

$$b_{0j} = b_{00} + b_{01}\,(\text{sex}) + u_{0j} \tag{7}$$

where the b_{10} term represents the average relationship between first task minute physiological reactivity and pretask appraisal ratings for men, and the b_{11} term indicates the difference between the magnitude of the b_1 coefficients for male and female participants. The b_{00} term is the mean first task minute reactivity score for men and the b_{01} term indicates the difference between the magnitude of the b_0 coefficients for male and female participants. A parallel set of analyses was run with sex dummy coded where female = 0, so the significance of the average simple slopes for female participants could be computed. In analyses where there are both Level-1 and Level-2 predictors, Level-1 predictors were group-centered to control for the possibility that a Level-2 predictor (e.g., sex) might be correlated with a Level-1 predictor (e.g., pretask appraisals; Hofmann & Gavin, 1998).

Sex accounted for individual variation in the pretask appraisal–PEP, CO, and RSA reactivity relationships, and was marginally significant for the pretask appraisal–TPR reactivity relationship. Men and women differed in the extent to which their pretask appraisals predicted PEP reactivity, $t = 2.00$, $p < .05$. Women had a strong positive association between their pretask appraisals and their PEP reactivity during the first task minute. Thus, PEP shortened more as women's pretask appraisals moved towards challenge, b_1 for women = 4.98, $B_1 = 0.30$, $t = 3.25$, $p < .002$. Conversely, men showed no relationship between pretask appraisals and initial PEP change in the task, b_1 for men = 0.91, $B_1 = 0.06$, $t = 0.55$, n.s. Men and women also differed in the extent to which their pretask appraisals predicted CO reactivity, $t = 2.20$, $p < .03$. Women had a marginal negative association between their pretask appraisal scores and their CO reactivity during the first task minute. Thus, for women, CO tended to increase more as pretask appraisals moved towards challenge, b_1 for women = −0.29, $B_1 = -0.16$, $t = 1.82$, $p < .068$. As with PEP, there was no relationship between pretask appraisals and CO reactivity for men, b_1 for men = 0.15, $B_1 = 0.08$, $t = 1.20$, n.s. Men and women differed in the extent to which pretask appraisals predicted RSA reactivity, $t = 2.09$, $p < .05$. Women had a marginally significant pretask appraisal–RSA reactivity relationship such that as women's pretask appraisals moved toward challenge, RSA tended to decrease more, b_1 for women = 0.31, $B_1 = 0.15$, $t = 1.80$, $p < .07$. This is consistent with previous studies demonstrating greater cardiac reactivity as pretask appraisals move toward challenge. For men, there was no pretask appraisal–RSA reactivity relationship, b_1 for men = −0.20, $B_1 = -0.09$, $t = 1.16$, n.s. Finally, men and women marginally differed in the extent to which their pretask appraisals predicted TPR reactivity, $t = 1.38$, $p < .17$. There was a marginally significant pretask appraisal–TPR reactivity relationship, but only for men, b_1 for men = −56.76, $B_1 = -0.19$, $t = 1.65$, $p < .10$; b_1 for women = 11.06, $B_1 = 0.04$, $t = .31$, n.s., and it was opposite the predicted direction.

The findings thus far demonstrated sex differences across several of the pretask appraisal–reactivity relationships. It is possible that men and women differed in their cardiovascular reactivity because they also differed in their pretask appraisals. Indeed, a set of exploratory analyses indicated that women were significantly

[1]Multilevel models such as those computed by HLM inherently contain interactions between Level-1 and Level-2 predictors. When asking the question: "Does a Level-2 predictor account for the variance in Level-1 regression coefficients?", we essentially are testing an interaction between the Level-1 and Level-2 predictors in their effect on the Level-1 criterion variable. Thus, in regression equations with interaction terms in which the predictor variables have been centered, the Level-1 regression coefficients are not main effects, but instead represent the effect of the predictor on the criterion at the mean of the Level-2 predictor variable.

Table 4. *Replication Check Using Multivariate Tests for Task 1 First Task Minute Reactivity Scores*

	Challenge (ratio < 1; $N = 58$)			Threat (ratio ≥ 1; $N = 15$)		
	Men ($N = 30$)	Women ($N = 28$)	Mean	Men ($N = 7$)	Women ($N = 8$)	Mean
HP	−195.4 (18.9)	−181.8 (19.6)	−188.8	−151.7 (39.1)	−222.8 (36.6)	−188.9
PEP	−7.0 (2.1)	−10.3 (2.2)	−8.6	−8.0 (4.4)	−8.9 (4.1)	−8.5
CO	0.57 (0.23)	0.81 (0.24)	0.69	0.38 (0.48)	0.58 (0.45)	0.49
TPR	135.3 (34.1)	65.4 (35.3)	101.6	157.3 (70.7)	122.7 (66.1)	138.8

Note: Standard errors are in parentheses. HP = heart period. PEP = preejection period. CO = cardiac output. TPR = total peripheral resistance.

is important to note that this between-subjects approach is quite different from that of the main analyses using HLM. Because HLM provides the ability to simultaneously model both between-subjects and within-subjects error, it provides a more powerful test that in this case, was sufficient to observe effects even with a more modest manipulation (Bryk & Raudenbush, 1992; Kenny et al., 1997).

We also assessed the percent correct and number of attempts as a function of challenge and threat groups using analyses parallel to those of Tomaka, Blascovich, and colleagues. There was no significant main effect of appraisal group (MANOVA group effect: $F(2,69) = 0.62$, n.s.), but there was a significant effect for sex of participant, $F(2,69) = 5.80$, $p < .005$. Specifically, men made more attempts than did women (64.7 vs. 44.5), $F(1,70) = 11.45$, $p < .001$. The interaction was not significant. There were no significant effects for percent correct responses. The fact that differences in performance were not significantly different across the challenge and threat groups further suggests that the first task was relatively easy for many participants.

Discussion

The Pretask Appraisal Hypothesis: An Idiographic View

The results of this study suggest that several of the predicted relationships between appraisals, cardiovascular reactivity, and task-related behaviors previously observed at a group level also can be observed within individuals. By assessing changes in appraisal and cardiovascular reactivity to task demands over multiple tasks, we demonstrated that as individuals become more challenged, they displayed more cardiac (HP, PEP, and marginally RSA and CO) reactivity. Moreover, the data suggest that this increased cardiac reactivity probably resulted primarily from greater sympathetic reactivity, although the task was also associated with marginally greater parasympathetic withdrawal. Although both autonomic branches may contribute to the cardiac reactivity associated with the serial subtraction task, only the relationship between pretask appraisals and PEP reactivity was significant when both PEP and RSA were in the same analysis, suggesting that pretask appraisals are predominantly associated with the initial sympathetic response to the task. The within-subjects findings for TPR were not consistent with previous between-subjects findings, and, indeed, were opposite the predicted effect for men. Note that because the change scores for the initial TPR response to the task were based on only one minute of data during the task, the potential unreliability of these scores could have contributed to our failure to find significant effects.

We also observed that women were more likely to report being threatened than were men. The lack of an objective criterion for threat introduces ambiguity into interpreting the results, especially for men. On the one hand, if we assume their self-reported appraisals are the best measure of subjective threat, then we might conclude that our male participants were simply not threatened sufficiently, which in turn limited our ability to find a relationship between pretask appraisals and cardiovascular changes. Consistent with this view is the fact that participants seemed to find the arithmetic task relatively easy. On the other hand, if we assume that men may not admit to feeling threatened by the arithmetic task (either to themselves or to an experimenter) even if they appraise the task as threatening at a less conscious level, then we might conclude that our male participants failed to display the expected relationships between pretask appraisals and cardiovascular changes because of a reporting bias. Indeed, appraisals in this study were operationalized as consciously available self-reports of stressfulness and coping ability; however, appraisals need not, and often may not be conscious (Blascovich & Mendes, 2000; Lazarus & Folkman, 1984). Because only some aspects of the appraisal process are likely to be available to conscious awareness, and only some subset of those perceptions may be "acceptable" for the participant to report, we might expect some incongruence between implicit appraisals and explicit reports of those appraisals (Greenwald & Banaji, 1995). The interpretation that reporting bias of some sort precluded an observable pretask appraisal–cardiovascular reactivity relationship in men is consistent with the finding that men and women did not differ systematically in their cardiovascular reactivity during the first minute of each task.

It is important to note, then, that we might not always expect congruence between consciously reportable appraisals and physiological changes, particularly from a within-subjects or idiographic perspective. We believe it will be important to further explore the interaction between person variables that may influence the willingness to report threat and assess task contexts that could alter the relationship between self-reported appraisals and cardiovascular reactivity. For example, in the current study, all experimenters to whom stress and coping ratings were reported were women, and the role of this particular contextual factor may require attention in future studies.

In contrast to the relatively strong support for our hypothesis that pretask appraisals would be related to cardiovascular reactivity, we did not find support for the proposed relationship between pretask appraisals and task behaviors. In the current study, men did not differ either in number of attempts or percent correct responses as a function of their pretask appraisals, and women showed results opposite those of some previous studies (Tomaka & Blascovich, 1994; Tomaka et al., 1993). Taken together with our findings, it seems that pretask appraisals do not invariably predict performance from either a between-subjects (e.g., Tomaka et al., 1997) or within-subjects perspective.

The Reappraisal Hypothesis

Our findings on the reappraisal hypotheses suggest that individuals may use information from multiple sources to make reappraisals. For example, greater cardiac reactivity during the task, combined with poorer performance, led participants to reappraise a task as more threatening after it was over. This finding suggests that the appraisal–reactivity relationship is not necessarily reciprocal. Thus, although more threatening pretask appraisals predict less cardiac reactivity early in the task, less reactivity in the task led to a more challenging posttask appraisal. Similarly, the appraisal–behavior relationship also can be nonreciprocal. So, although here pretask appraisals predicted task behaviors only for women, behaviors in the task predicted posttask appraisals across all individuals. As a further example of the dynamic perspective of the appraisal process possible from these results, we observed that when women make threat appraisals prior to a task, they perform better in the task and then make more challenged posttask appraisals. Although the current study cannot be used to draw conclusions about the temporal unfolding of reappraisals across time, it does suggest that there are important changes in appraisal that take place in motivated performance situations. More importantly, this study provides some initial data concerning the information that individuals may use to make those reappraisals, both before and after a task, as well as from one task to the next. For example, the lagged analyses showed that, on average, appraisals of an upcoming task were based both on the previous number of attempts made and on posttask appraisals of the preceding task, although posttask appraisals appeared to be the stronger predictor. These findings suggest that across iterations of the task, some information remembered from previous tasks may be used to make judgments about upcoming tasks.

The finding that both greater cardiac reactivity and poorer performance during a task led to more threatening appraisals after the task is interesting, but should be viewed cautiously. Although increased cardiac reactivity (more shortening of HP and PEP) during the last minute of the task was related to more threatening posttask appraisals even when the effects of behavior were controlled, the effect size for HP was small, suggesting that the primary effect was that for PEP. This is interesting given previous results in a heartbeat detection task suggesting that those with better detection abilities showed stronger inotropic reactivity, but not stronger heart rate reactivity during a mental arithmetic task (Eichler & Katkin, 1994). These data suggest that those with stronger contractility responses to stressors may have a better ability to detect some internal states, although the ability to detect heartbeats even for good detectors is far from perfect. Moreover, the lagged analyses in the current study indicated that cardiovascular reactivity during the last minute did not influence pretask appraisals of a succeeding task. Thus, since perceptions of cardiovascular function in many individuals are typically available in relatively impoverished, or nonconscious form (e.g., Adam, 1998; Fahrenberg, Franck, Baas, & Jost, 1995), and because reactivity in the current study influenced only the appraisal made immediately following the task, it seems likely that cardiovascular reactivity alone will have a modest and fleeting influence on appraisals in the stressful situations that individuals typically experience in everyday life or in a laboratory setting. However, if individuals do use this information, it appears that they interpret the experience of increased cardiac reactivity as a signal of threat. These considerations suggest there is not a simple answer to the question of whether physiological changes influence subjective experience. Indeed, our data suggest a very transactional view such that expectations about an upcoming task led to particular cardiovascular changes during a task, and later components of the cardiovascular response predicted subjective experience following the task (at least briefly).

The fact that task demand and performance influenced posttask appraisals more strongly for women than men suggests there are individual differences in the information that people rely on during the reappraisal process. Furthermore, it suggests that a productive avenue for future research would focus on how individuals calibrate new appraisal judgments based on previous experience. There are two issues that likely warrant further scrutiny, namely, what information is available to a given individual for making a reappraisal, and when does a given individual use available information? We speculate that individuals vary in their propensities to seek out and use environmental information in the course of making reappraisals, and as a result, individuals will differ in how successfully they calibrate their judgments to the environmental context (Quigley & Feldman Barrett, 1999).

Conclusion

This study represents more than a simple within-subjects conceptual replication of some previously observed between-subjects effects. Instead, it begins to examine the dynamic and recursive appraisal process. Our findings strongly support the theoretical proposition that cardiovascular and behavioral stress responses are an important determinant of reappraisals. Thus, appraisals made before a series of motivated performance tasks predict cardiovascular changes in the task, and in turn, performance and reactivity in the tasks relate to an individual's posttask appraisals. Each succeeding experience with the potential stressor provides the individual with new information and may help in making future judgments and adapting to changing environmental conditions. Our data suggest that a more complete picture of the within-individual relationships between appraisals, physiology, and behavior will emerge by taking a more dynamic view of the appraisal process, and using analytical tools to model both within-subject and between-subjects variance. Moreover, these findings demonstrate that sex is likely an important factor moderating these relationships, which deserves further exploration.

REFERENCES

Adam, G. (1998). *Visceral perception: Understanding internal cognition.* New York: Plenum Press.

Berntson, G. G., Cacioppo, J. T., Binkley, P. F., Uchino, B. N., Quigley, K. S., & Fieldstone, A. (1994). Autonomic cardiac control: III. Psychological stress and cardiac response in autonomic space as revealed by pharmacological blockades. *Psychophysiology, 31,* 599–608.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1991). Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological Review*, *98*, 459–487.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1995). The metrics of cardiac chronotropism: Biometric perspectives. *Psychophysiology*, *32*, 162–171.

Blascovich, J., & Mendes, W. B. (2000). Challenge and threat appraisals: The role of affective cues. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 59–82). New York: Cambridge University Press.

Blascovich, J., Mendes, W. B., Salomon, K., & Hunter, S. B. (1999). Social facilitation as challenge and threat. *Journal of Personality and Social Psychology*, *77*, 68–77.

Blascovich, J., & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 1–51). New York: Academic Press.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Bryk, A. S., Raudenbush, S. W., & Congdon, R. T. (1996). *Hierarchical linear and nonlinear modeling with the HLM/2L and HLM/3L programs*. Chicago: Scientific Software.

Dienstbier, R. (1989). Arousal and physiological toughness: Implications for physical and mental health. *Psychological Review*, *96*, 84–100.

Eichler, S., & Katkin, E. S. (1994). The relationship between cardiovascular reactivity and heartbeat detection. *Psychophysiology*, *31*, 229–234.

Fahrenberg, J., Franck, M., Baas, U., & Jost, E. (1995). Awareness of blood pressure: Interoception or contextual judgment? *Journal of Psychosomatic Research*, *39*, 11–18.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Journal of Personality and Social Psychology*, *102*, 4–27.

Grossman, P., Stemmler, G., & Meinhardt, E. (1990). Paced respiratory sinus arrhythmia as an index of cardiac parasympathetic tone during varying behavioral tasks. *Psychophysiology*, *27*, 404–416.

Grossman, P., & Svebak, S. (1987). Respiratory sinus arrhythmia as an index of parasympathetic cardiac control during active coping. *Psychophysiology*, *24*, 228–235.

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, *24*, 623–641.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly*, *14*, 299–324.

Kelsey, R. M., Blascovich, J., Leitten, C. L., Schneider, T. R., Tomaka, J., & Wiens, S. (2000). Cardiovascular reactivity and adaptation to recurrent psychological stress: The moderating effects of evaluative observation. *Psychophysiology*, *37*, 748–756.

Kelsey, R. M., Blascovich, J., Tomaka, J., Leitten, C. L., Schneider, T. R., & Wiens, S. (1999). Cardiovascular reactivity and adaptation to recurrent psychological stress: Effects of prior task exposure. *Psychophysiology*, *36*, 818–831.

Kelsey, R. M., & Guethlein, W. (1990). An evaluation of the ensemble averaged impedance cardiogram. *Psychophysiology*, *27*, 24–33.

Kenny, D. A., Kashy, D. A., & Bolger, N. (1997). *Quantitative methods in social psychology*. In D. Gilbert, S. Fiske, & G. Lindsey (Eds.), Handbook of social psychology (4th ed.). New York: McGraw-Hill.

Kreft, I. G. G., & De Leeuw, J. (1998). Introducing multilevel modeling. Newbury Park, CA: Sage.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer Publishing.

MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, *32*, 215–253.

Porges, S. W., & Bohrer, R. (1990). The analysis of periodic processes in psychophysiological research. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of psychophysiology, Physical, social and inferential elements* (pp. 708–753). New York: Cambridge University Press.

Quigley, K. S., & Feldman Barrett, L. (1999). Emotional learning and mechanisms of intentional psychological change. In J. Brandtstädter, & R. M. Lerner, (Eds.). *Action and self-development: Theory and research through the life span* (pp. 435–464). Thousand Oaks, CA: Sage Publications, Inc.

Ratliff-Crain, J., & Baum, A. (1990). Individual differences and health: Gender, coping, and stress. In H. Friedman (Ed.). *Personality and disease* (pp. 226–253). New York: John Wiley, & Sons.

Raudenbush, S. W., Brennan, R. T., & Barnett, R. C. (1995). A multivariate hierarchical model for studying psychological change within married couples. *Journal of Family Psychology*, *9*, 161–174.

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R., & van Doornen, L. J. P. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology*, *27*, 1–23.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A. R., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, *107*, 411–429.

Tomaka, J., & Blascovich, J. (1994). Effects of justice beliefs on appraisal of and subjective, physiological, and behavioral responses to potential stress. *Journal of Personality and Social Psychology*, *67*, 732–740.

Tomaka, J., Blascovich, J., Kelsey, R. M., & Leitten, C. L. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *65*, 248–260.

Tomaka, J., Blascovich, J., Kibler, J., & Ernst, J. M. (1997). Cognitive and physiological antecedents of threat and challenge appraisal. *Journal of Personality and Social Psychology*, *73*, 63–72.

(RECEIVED September 18, 2000; ACCEPTED September 4, 2001)