

Dispositional Empathy and Autonomic Reactivity During a Comfort Interaction

Alaina Tiani, B.S.

Thesis submitted to the
Eberly College of Arts and Sciences
at West Virginia University

in partial fulfillment of the requirements for the degree of

Master of Science
in the Department of Psychology

Kevin Larkin, Ph.D. – Chair
Julie Hicks Patrick, Ph.D.
Amy Gentzler, Ph.D.

Department of Psychology

Morgantown, West Virginia
2019

Keywords: empathy; cardiovascular response to stress; cardiovascular reactivity; interpersonal stress; heart rate; blood pressure

Copyright 2019 Alaina Tiani

ABSTRACT

Dispositional Empathy and Autonomic Reactivity During a Comfort Interaction

Alaina Tiani, B.S.

The purpose of the current study was to examine the association between dispositional (trait) empathy and autonomic reactivity during dyadic interactions among women enrolled in undergraduate psychology courses. The primary research question was: do individuals who are deemed higher in dispositional empathy according to self-report exhibit differential autonomic reactivity patterns during a task in which they are asked to comfort another student experiencing emotional distress than those who are deemed lower in dispositional empathy? The literature on this relation has been mixed and has utilized photos or videos to evoke emotion; thus this study sought to examine the relation between empathy and autonomic reactivity using an in vivo interaction between a participant and a confederate. Measures of heart rate (HR), heart rate variability (HRV), and blood pressure (BP) were recorded continuously during two separate interactions (the comfort task and a neutral interaction) and displays of empathy and prosocial behavior were coded. Self-reported measures of state affect, interpersonal goals, and task appraisals were also measured. Results revealed very few differences between women higher and lower in dispositional empathy. Women in the higher empathy group displayed more social network support behaviors like offering to spend time with the confederate or sharing that they can personally relate to the situation and reported encountering the types of interactions used in this study more frequently than women in the lower empathy group. Several differences were observed between the comfort and neutral tasks. Specifically, although HR responses were higher in response to the neutral than comfort tasks, the neutral task was associated with lower self-reported ratings of negative affect, stressfulness, and difficulty, than the comfort task. Although findings failed to confirm study hypotheses, several methodological issues were uncovered that should be considered in future empirical work examining the positive and negative impact of empathy.

TABLE OF CONTENTS

Introduction	i
Aims of Proposed Study	v
Method	vii
Participants	vii
Measures	ix
Experimental Design	xiii
Procedure	xiii
Results	xvi
Cardiovascular Measures at Rest	xvi
Primary Analyses: Cardiovascular Reactivity to the Task	xvii
Measures of Affect	xx
Post-Experimental Questionnaire	xxi
Behavioral Measures	xxii
Discussion	xxvi
Strengths and Limitations	xxxiv
Future Directions	xxxvii
Summary and Implications	xl
References	xlii
Tables	liii
Figures	lxi
Appendices	lxiii
Appendix A: Literature Review	lxiii

Appendix B: Measures, Forms, and Scripts	lxxvii
Appendix C: Preliminary Analyses	lxxxix
Appendix D: Confederate Effects	civ
Appendix E: Task Order Effects	cxvi

Dispositional Empathy and Autonomic Reactivity During
a Comfort Interaction

The relation between health care provider and patient is a critical variable in the provision of evidence-based care. Patients who trust their providers are more apt to seek their assistance when advice is needed, follow recommended treatment plans, and return for follow-up visits than those who do not (Piette, Heisler, Krein, & Kerr, 2005; Thom, Kravitz, Bell, Krupat, & Azari, 2002). Although there are several factors known to promote the quality of these professional relationships, provider empathy has consistently been shown to be an important parameter in establishing positive doctor-patient relationships (Dixon, Sweeney, & Gray, 1999; Kim, Kaplowitz, & Johnson, 2004; Mercer & Reynolds, 2002). Empathy is commonly defined as both a cognitive (e.g., perspective-taking or the ability to accurately understand another's emotions; Duan & Hill, 1996; Teding van Berkhout & Malouff, 2016) and affective phenomenon (e.g., personal experience or vicarious arousal of another's emotions; Mehrabian & Epstein, 1972) that facilitates the development of positive human relationships. Empathy is often accompanied by the portrayal of prosocial behaviors (e.g., supportive touch or body language, facial expressions, or verbalizations; Batson et al., 1991; Eisenberg & Miller, 1987; Preston & DeWaal, 2002) also known to enhance the quality of interpersonal relationships of all types. According to Allgood (1992), empathy is a trait that reflects an involuntary, unlearned, innate human attribute to feel for others, and that some individuals are more empathic by nature than others.

Although essential for establishing positive relationships between health care providers and patients, exhibiting empathy with all patients seen on a daily basis could also be quite exhausting for health care providers. In this regard, empathy may have both positive and negative consequences for health care professionals of all types, including those with interests in

providing behavioral health care. Students in health care professional training programs, eventually pursue careers that involve developing therapeutic relationships constructed on empathic listening and the validation of patient's life experiences and often are drawn to or required to take introductory courses in psychology. It is important then to examine factors that both foster and potentially deter the development of skills in empathy that might influence health profession students' future success in establishing effective provider-patient relationships. This study was designed to examine the effects of being empathic towards another by exploring both the positive and negative aspects of empathy. Thus, we observed young adults with higher and lower levels of trait empathy as they engaged in an interaction in which they were asked to comfort a person in distress.

Empathy in Health Care Providers

Most of the empirical work conducted on the relation between empathy and overall health has focused on studying empathy and health benefits for the *receiver*. Several studies examining effects of provider empathy in medical settings have shown that provider empathy and perception of empathy by patients is associated with improved health outcomes related to diseases such as diabetes (Hojat et al., 2011). Several systematic reviews of effectiveness of provider empathy has shown it to be associated with better clinical outcomes, including patient satisfaction, adherence, psychological adjustment, and feelings of distress (Derksen, Bensing, & Lagro-Janssen, 2013; Lelorain, Brédart, Dolbeault, & Sultan, 2012).

Although empathy of health care providers has consistently been shown to be beneficial for patient outcomes, numerous studies have found that, over time, empathy may actually decrease in health care providers and other human service workers. In a systematic review of the progression in empathy among medical students and residents, Neumann et al. (2011) discovered

a paradoxical reduction in empathy in this population as a result of training phase, training progression, and specialty selection. Several studies in the review found that empathy significantly decreased as student's entered into the clinical training phase of their educations. They proposed that because this transition exposes students to suffering and mortality, perhaps they begin to "over-identify" with patient emotions, make less effective health care decisions, and perform less effectively. Consequently, these health professionals-in-training often learn to distance themselves from their patients emotionally as they launch their professional careers.

Though the systematic review mentioned above was conducted on medical students and residents, the literature contains studies examining empathy in a wide range of human service-related positions. Grevin (1996) observed lower scores on an empathy scale among the experienced paramedics and the paramedic students compared to a control sample, and Miller, Stiff, and Ellis (1988) reported that empathy was a significant precursor to burnout among human service workers in a hospital setting. Similarly, in a study of hospital staff, Lief and Fox (1963), identified "detached concern" as a strategy used to portray concern while remaining emotionally unattached allowing staff to perform their duties without becoming emotionally exhausted.

If empathy is beneficial for patient well-being in medical settings, why does it tend to decline among providers of care for those in need? Overall, the results of these studies suggest that a reduction in empathy may be adaptive for health care providers, protecting them from the emotional strain of their everyday work. Perhaps expressing empathy frequently with numerous patients each day is emotionally-taxing, leading to more frequent or intense autonomic nervous system responses to these emotional stimuli. To the extent that experiencing frequent, exaggerated, and/or prolonged autonomic responses to life events places an individual at risk for

one's own health problems (e.g., Krantz & Manuck, 1984; Treiber et al., 2003), adopting a less empathic style of interaction style may be associated with a more adaptive physiological stress profile than those who consistently display higher empathy. In fact, there is some support for this hypothesis from studies that have measured empathy in relation to exposure to emotional provocation.

Empathy and Autonomic Reactivity

Numerous studies have examined the relation between empathy and autonomic response to emotions using various means for eliciting emotion, including imagination, photos, videos, and live dyadic interactions (See Appendix A). However, no consistent pattern of findings has emerged. Most studies found that increased empathy was associated with increased sympathetic or decreased parasympathetic activity in response to stressful or emotionally-provocative stimuli (Ardizzi et al., 2016; Balconi & Bortolotti, 2012; de Wied, et al., 2012; Geringer, 2015; Liew et al., 2011; Miller, et al., 2016; Oliveira-Silva & Gonclaves, 2011; Tracy & Giummarra, 2017; Truzzi et al., 2016; Wiesenfeld et al., 1984; Zahn-Waxler et al., 1995). In contrast, Stellar and colleagues (2015) found that engagement in compassion was associated with decreased sympathetic responses to stress. Vanderpool and Barratt (1970) reported mixed findings such that higher empathy was associated with increased skin conductance responses (SCR) and respiration rate, but decreased HR. Likewise, Miu and Baltes (2012) found increased empathy to be related to both increased respiration rates and decreased SCR responses when manipulated, but no differences associated with participant's trait empathy. Finally, three studies found no significant associations between autonomic reactivity measures and empathy (Anastassiou-Hadjicharalambous & Warden, 2007; Ono et al., 2012; Perrone-McGovern, 2014). As noted in Appendix A, the pattern of findings did not differ based on stimulus type (standard mental stress

stimuli, photos of emotional expression, auditory emotionally-provocative stimuli, video stimuli, and in vivo interpersonal interactions) or physiological measure selected as a measure of sympathetic or parasympathetic responding (HR, HRV, SCR, and respiration rate). For a more in-depth review of the literature, see Appendix A.

Aims of the Study

The purpose of the current study was to examine the association between dispositional (trait) empathy and autonomic reactivity during dyadic interactions among women enrolled in undergraduate psychology courses. Findings from the current literature regarding autonomic reactivity to various emotion inducing stimuli have been mixed, but most of these studies have supported the finding that at least one measure of physiological responding to emotion-evoking stimuli was associated with dispositional empathy. Those who possessed more empathy exhibited larger autonomic reactions than those with less empathy. The bulk of this literature, however, has employed static emotion-laden photos or dynamic video portrayals of emotionally evocative situations. Although these types of studies possess good internal validity by exposing all study participants to identical test stimuli, their attention to generalizing results to how people perform in real-life situations is limited. The current study focused on extending our knowledge of the relation between empathy and physiology into more ecologically-valid situations. In this regard, this study aimed to explore the external validity of findings that persons higher in empathy exhibit larger autonomic responses to emotionally-evocative stimuli than persons lower in empathy. This study utilized an in vivo interaction task between a participant and a confederate in order to examine the nature of the empathy and autonomic reactivity relation in a more naturalistic dyadic interaction. Because sex differences have been reported on measures of empathy (Baron-Cohen & Wheelwright, 2004; Eisenberg & Lennon, 1983; Hoffman, 1977) and

in the relation between empathy and physiological response to emotion (Tracy & Giummarra, 2017; Vanderpool & Barratt, 1970; Zahn-Waxler et al., 1995), the proposed study only used female participants and female confederates. This had the added benefit of removing any confounds associated with potential influences of comforting or interacting with same-sex or other-sex persons. Measures of HR, HRV, and BP were recorded continuously during two interactions to obtain resting and task period cardiovascular reactivity data. Interactions were recorded and coded to permit collateral assessment of displays of empathy and prosocial behavior and self-reported measures of state affect, interpersonal goals, and task appraisals were obtained.

The primary research question which served as the basis for the current study was: do individuals who are deemed higher in dispositional empathy according to self-report exhibit differential autonomic reactivity patterns during a task in which they are asked to comfort another student experiencing emotional distress than those who are deemed lower in dispositional empathy? A related research question of whether the differential pattern of reactivity between students reporting higher and lower dispositional empathy, if indeed it was observed, was unique to situations requiring empathic listening and provision of support or whether it generalized to other types of interactions is also addressed in this study. For comparison purposes, a neutral task was used that involved no expressions of comfort. Finally, both observed behavioral responses and covert emotional responses were measured to determine whether any differences in physiological responses to these two interaction tasks extended to other response parameters.

Method

Participants

Thirty-two female participants were ultimately recruited from undergraduate classes at West Virginia University for participation in the laboratory study. Participants were excluded from the initial pool of participants if they were male, younger than 18 years old, older than 25 years old, had any chronic major health concerns (e.g., heart disease, cancer, diabetes), currently used tobacco products, or were taking medications that influenced heart rate or blood pressure. Regarding age, participants who fell outside of the 18-25 year age range were excluded as the focus of this study was on relationships and interactions between peers of a similar age group. Participants were asked to abstain from caffeine, alcohol, and vigorous exercise for 2 hours prior to the experiment. Only females were recruited for this study to eliminate confounding of sex differences in emotional responding to same-sex versus other-sex laboratory confederates.

Sample size for the current study was based on prior studies similar in nature (Balconi & Baltes, 2012; Miu & Baltes, 2012; Truzzi et al., 2016). Balconi and Bartolotti (2012) was the only study to compare two groups (one higher and one lower empathy) on physiological measures, which is the most similar to the current study design. Based on the three studies mentioned, the total sample size chosen for the current study was 32 (with 16 in the higher empathy and 16 in the lower empathy groups) in order to conduct a 2x2 Factorial ANOVA on the data and examine the interaction between task type and empathy level.

Participants were selected from an initial screening sample of undergraduate students enrolled in psychology courses at West Virginia University. Students completed an initial empathy questionnaire, the Basic Empathy Scale (BES), through SONA, an online survey

system, and participants were invited to the laboratory portion of the study based on their BES score. The total number of students who completed the questionnaire through SONA was 362.

Participants were excluded from the initial pool of participants if they were male ($N = 57$), or did not indicate gender ($N = 2$), were younger than 18 years old ($N = 3$), older than 25 years old ($N = 2$), had any chronic major health concerns (e.g., heart disease, cancer, diabetes) ($N = 18$), or were taking medications that influenced heart rate or blood pressure ($N = 4$). The remaining number of participants eligible for laboratory participation was $N = 276$ (See Table B.1 for complete demographic information).

Scores on the BES were split into tertiles of the SONA screening sample and participants whose score on the BES was 73 or lower fell in the bottom tertile or “lower” empathy group, while those with BES scores of 80 or higher fell into the “higher” empathy group. Participants were invited via email to participate in the laboratory portion of the study if their scores fell into either of these two categories. Mean BES score for the higher empathy group was 84.75 (± 5.05) and mean BES score for the lower empathy group was 66.00 (± 6.13), which were similar to mean empathy scores for females on the BES ($M = 75.3$, $SD = 8.3$) reported by Joliffe and Farrington (2006).

For the 32 female participants who completed the laboratory portion of the study, the mean age was 18.72 years ($SD = 0.92$). Of these participants, 28 identified as White/Caucasian, 1 identified as African-American/Black, 1 identified as Hispanic/Latino(a), and 2 identified as multiple races. For comparisons of higher and lower empathy groups on demographic characteristics, see Table B.2. Across demographic variables, the sample of women who completed the laboratory session represented the screening sample of women well.

Physiological Measures

Blood pressure. An *Industrial and Biomedical Sensors, Inc.* Model SD-700A (Waltham, MA) automated sphygmomanometer was used to measure SBP and DBP. This device utilizes an automated occluding cuff positioned on the brachial artery of the participants' non-dominant arm to detect Kortokoff sounds (via a microphone), ensuring accurate BP measurement. Maximum cuff inflation was set at 165 millimeters of Mercury (mm Hg) and rate of deflation was set at 3 mm Hg per sec. SBPs and DBPs were displayed digitally and recorded by the experimenter.

Heart rate (HR). HR was measured using a Model 810i *Polar* heart rate monitor (Lake Success, New York). This device measured HR continuously throughout data collection by sending ECG signals from a sensor strapped below the participants' chest to a receiver attached to a computer. Three measures of heart rate variability (HRV) were also determined from the continuous HR signals: the root mean square of successive R-R intervals (RMSSD) (Esco & Flatt, 2014; McNames & Aboy, 2006; Nussinovitch et al., 2011; Thong, Li, McNames, Aboy, & Goldstein, 2003), low frequency (LF) HRV, and high frequency (HF) HRV. Kubios Premium HRV v3.1 software was used to examine HR signals for clarity, reduce HR data, and calculate measures of HRV (Niskanen, Tarvainen, Ranta-Aho, & Karjalainen, 2004).

Self-report Measures

Demographic form. A short demographic form used in previous studies in the Behavioral Physiology laboratory (e.g., Stephenson, 2015) was used in this study. This questionnaire included items pertaining to age, sex, height, weight, race/ethnicity, family socioeconomic status, and year in school. The form also included general questions about participants' health status and behaviors. (See Appendix B).

Basic Empathy Scale (BES). The BES is a 20-item self-report questionnaire which utilizes a 5-point Likert-type scale ranging from “Strongly Disagree” to “Strongly Agree.” The BES includes items which measure two empathy-related factors, cognitive empathy (9 items) and affective empathy (11 items) (Jolliffe & Farrington, 2006). The BES adopts the definition of empathy published by Cohen and Strayer (1996), which describes empathy “as the understanding and sharing in another’s emotional state or context” (p. 523). Examination of the scale’s psychometric properties show good internal reliability for both the cognitive and affective factors separately, ($\alpha = .79$ to $.85$, respectively) and no differences in factor structure between males and females. Significant sex differences were found between males and females on the overall scale, with women scoring significantly higher than men. Literature on test-retest reliability of the BES in adults suggests that scores on the BES do not significantly differ over retest durations of 7 weeks/49 days (Carré, Stefaniak, D’ambrosio, Bensalah, & Besche-Richard, 2013).

Positive and Negative Affect Schedule (PANAS). The PANAS is a 20-item self-report questionnaire that asks the participant to indicate how they feel in the present moment, utilizing a 5-point Likert-type scale ranging from 1 (very slightly or not at all) to 5 (extremely). Ten items measure negative affect (NA), and ten items measure positive affect (PA). Watson, Clark, and Tellegan (1988) define PA as “the extent to which a person feels enthusiastic, active, and alert,” with high levels of PA representing a state of “high energy, full concentration, and pleasurable engagement,” and low PA as “characterized by sadness and lethargy” (p. 1063). The authors define NA as “a general dimension of subjective distress and unpleasurable engagement that subsumes a variety of aversive mood states, including anger, contempt, disgust, guilt, fear, and nervousness, with low NA being a state of calmness and serenity” (p. 1063). Psychometric evaluation of the PANAS revealed moderately good reliability and validity. Cronbach’s alpha

coefficients for the PA scale have ranged from 0.86 to 0.90, and 0.84 to 0.87 for the NA scale. Test-retest reliability correlations over an 8-week retest interval were 0.47-0.68 for the PA scale and 0.39-0.71 for the NA scale. The scale has shown good external validity as it is highly correlated with measures of distress and general psychopathology (.74 with NA, -.19 with PA), state anxiety (.51 with NA, -.35 with PA), and depression (.58 with NA, -.36 with PA), as measured by the Hopkins Symptom Checklist (HSCL), Beck Depression Inventory (BDI), and STAI State Anxiety Scale (A-State) over the past few weeks, respectively (Watson, Clark & Tellegan, 1988).

Post-Task Questionnaire. The post-experimental questionnaire consisted of items that assessed the degree of emotion experienced during the two interactions, perceived difficulty of the two interactions, semantic attributions of the two interactions, and other factors related to the participants' experiences during the stress tasks. (See Appendix B).

Behavioral Measures

Behaviors during participant-confederate interactions were videotaped. Verbal and nonverbal behaviors were coded using a combination of behavioral codes taken from the Marital Interaction Coding System (MICS) (Hops, Wills, Weiss, & Patterson, 1972) and the Social Support Behavior Code (SSBC) by Suhr, Cutrona, Krebs, and Jensen (1992). The MICS was developed as a tool to objectively measure both verbal and non-verbal units of interaction between married couples in a dyadic interaction discussing their marital problems. The SSBC was developed to assess social support behaviors in the context of help-intended dyadic interactions in which one member of the couple discloses a personal problem to the other, and to understand the ways in which people communicate social support to one another. All behavior codes from the SSBC were selected, and several non-verbal codes from the MICS were selected.

Coded behaviors included the following: Informational Support, Emotional Support, Esteem Support, Tangible Aid, Social Network Support, Tension Reduction, Attentiveness, Negative Behavior, “Other,” and MICS non-verbal codes. Interrater agreement of each behavioral code was calculated and ranged from 83.3% to 100% agreement (See Appendix B for the behavioral coding form. See Tables C.9 and C.10 for interrater agreements for each code.)

Experimental Tasks

Each participant engaged in two interaction tasks with a confederate. Two female confederates were used throughout the study to remove any potential influence of different versus same sex interaction effects. The same confederate participated in both the comfort and neutral tasks for each participant.

Comfort Task. In this task, the confederate expressed sadness that her significant other had just broken up with her and seemed as though she would like to talk about it. The participant was asked to do her best to comfort the confederate, “her friend.” The confederate only responded to the participant’s verbalizations using a set of standardized prompts that express sadness and hopelessness at the loss of the relationship (e.g., “It just won’t be the same without him”).

Neutral Task. In this task, the participant was asked to imagine that her friend was trying to decide whether or not she should attend WVU for an undergraduate degree. The participant was asked to convince the confederate that WVU is the best school for her. The confederate only responded to the participant’s verbalizations using a set of standardized prompts to express non-commitment or hesitation about attending WVU (e.g., “None of my friends are going to WVU”).

Experimental design

The study was quasi-experimental in design. Participants were assigned to one of two levels of trait empathy (higher empathy, lower empathy) based on her BES total score; participants scoring in the top and bottom tertiles were invited to participate in the laboratory portion of the study. All participants completed both the comfort and the neutral interaction. The key interaction was the comfort interaction, while the neutral interaction served as a non-comfort comparison interaction. The two interactions were counterbalanced across both higher and lower empathy participants to control for order effects.

See Appendix B for a complete description of participant demographics, measures (physiological, self-report, and behavioral), and confederate scripts for the experimental tasks.

Procedure

Participants completed the demographic and other questionnaires online using the SONA system. Those who scored either in the top or bottom tertile on the BES and met all other inclusion criteria (no medications that would affect cardiovascular activity, etc.) were invited to schedule a laboratory session. Upon entering the laboratory, the participant met the experimenter. The experimenter described the components of the study and reviewed the potential risks and benefits with the participant and obtained her informed consent for study participation, as well as video recording of the session, using an approved consent agreement. The experimenter explained that the study involved two role play interactions. Then, the experimenter measured the height and weight of the participant, and confirmed that she had abstained from caffeine, alcohol, and exercise for the previous two hours. The experimenter then left the room to allow the participant to attach the *Polar* heart rate monitor around her chest privately. The experimenter re-entered the room and attached the blood pressure cuff to the

participants' non-dominant upper arm. HR and BP measurements were subsequently examined to assure signal clarity. Participants were then instructed to sit quietly with both feet on the floor for a 5-minute adaptation period. Blood pressure was measured continuously throughout the entirety of the protocol, beginning at the start of the adaptation period. After this period, the first 5-minute rest period began. Blood pressure measurements began immediately at the start of the rest period and were taken every two minutes (at minute one, three, and five) for the remainder of the rest period.

At the conclusion of the rest period, the participant was asked to complete a baseline measure of the PANAS. Then, the experimenter introduced the confederate as another student who was helping with the study, and the experimenter attached a blood pressure cuff to the confederate's non-dominant upper arm. The experimenter distributed note cards to both the participant and the confederate with a description of the first interaction scene (either the Comfort Task or the Neutral Task) and the participant goal, and explained the protocol and relevant task to both the participant and the confederate. The participant was instructed to respond as she normally would if the interaction occurred in real life. The experimenter began the video recording using a Dell laptop and Microsoft LifeCam software. The experimenter then returned to the equipment room and announced over the intercom that it was time to begin the interaction, cuing the participant to start the conversation to meet her goal. BP measures were recorded during minutes one and three of the three-minute interaction. After the first interaction had concluded, the experimenter entered the testing room and detached the blood pressure cuff from the confederate and asked her to wait in the adjoining room. The participant was given the first post-task questionnaire and the PANAS. This period of questionnaire completion also served as a recovery period from the first task.

After completion of the questionnaire, the second 5-minute rest period began. BP measures were recorded at one, three, and five minutes. At the end of the rest period, the confederate reentered the testing room and a blood pressure cuff was reattached to her arm. The experimenter distributed note cards to both the participant and the confederate with a description of the second interaction scene and the second participant goal. The participant was given the same instructions for responding as in the first interaction. The experimenter then began the second video recording and returned to the equipment room to announce the start of the second interaction. BP measures were recorded at minutes one and three of the interaction. After the second interaction, the confederate left the room, and the second recovery period began. The experimenter distributed the second post task questionnaire and PANAS to the participant. After completing these two questionnaires, the experimenter distributed the self-report questionnaire of empathy (previously completed during pre-screen) to examine whether empathy score changed during the time period between the SONA pre-screen and the laboratory portion of the study.

After completing the questionnaires, the experimenter removed the blood pressure cuff, and the participant was asked to remove the heart rate monitor. Upon leaving the testing room, the participant met with the experimenter for a debriefing of the study and payment of \$10.

Participants in the current study completed the BES online between late August and early September 2018, and laboratory sessions occurred between early October to early December of 2018. The average number of days between BES completion times for the current sample was 58.2 ± 20.3 days. Overall test-retest reliability for the BES scores of the study sample was $r = .703, p < .001$. For those with a test-retest duration of 49 days or less ($n = 10$), test-retest reliability was $r = .756, p < .05$. For those with a test-retest duration of 50 days or greater ($n = 22$), test-retest reliability was $r = .707, p < .001$. Preliminary analyses revealed that there was a

significant change in empathy score between laboratory and SONA BES administration times (see Appendix C: SONA versus Laboratory Empathy Score Comparison), and findings are interpreted in light of this information.

Results

Preliminary analyses were conducted to examine reactivity during rest and task periods, Confederate effects, as well as Task Order Effects. (See Appendix C, D, and E for a complete description of these analyses). Overall, few differences were observed between participants when interacting with Confederate A versus Confederate B; however, differences were seen for SBP, feelings of being upset, and esteem support behaviors. In regard to Task Order, differences were observed between groups for DBP and HF-HRV during rest periods, positive/negative affect, ratings of realism, stress, feeling upset, and perceived task difficulty; however, most of these order effects were qualified by task type. Because assignment to confederate and task order were counterbalanced and half of the participants in each empathy group were assigned to interact with each confederate and complete the tasks in one of the designated orders, confederate and task order were not controlled in the study's primary analyses.

Primary Analyses: Resting Measures

The primary data analytic strategy for evaluating the effects of trait empathy of study participants as they interacted during the comfort and neutral tasks was a 2 x 2 mixed factor's analysis of variance (ANOVA), with the between subject's factor being Empathy Group (Higher, Lower) and the within subject's factor being Task Type (Comfort Task, Neutral Task). When dependent measures were available from pre-task rest periods (e.g., all cardiovascular parameters and measures of self-reported affect), respective pre-task values were used as covariates to control for differences in resting parameters.

Cardiovascular Measures at Rest. An Empathy Group by Task Type ANOVA was conducted on each resting cardiovascular parameter prior to each task (pre-neutral and pre-comfort) to determine whether there were significant differences in resting measures between the higher and lower empathy groups. No significant differences between groups were detected on any resting measure of cardiovascular functioning. (See Table 1).

Self-Report Measures of Affect at Rest. A comparable ANOVA was conducted on resting self-report measures of positive and negative affect (pre-neutral and pre-comfort) to determine whether there were significant differences in resting measures between the higher and lower empathy groups, and no significant differences between groups were detected. (See Table 2).

These results indicated that both the higher and lower empathy groups were equivalent in cardiovascular and self-report parameters prior to engaging in both interaction tasks.

Primary Analyses: Cardiovascular Reactivity to Tasks by Empathy Group

Results for the main effects of both Empathy Group and Task Type, as well as Empathy Group by Task Type interactions, are reported here.

Systolic Blood Pressure. Analyses revealed no significant main effect of Empathy Group on SBP, $F(1, 29) = 1.64, p = .21, \eta_p^2 = .05$, and no significant main effect of Task Type, $F(1, 29) = 3.16, p = .09, \eta_p^2 = .10$. Results did reveal a significant interaction between Empathy Group and Task Type, $F(1, 29) = 4.54, p = .04, \eta_p^2 = .14$. Post-hoc simple main effect analyses were conducted to examine which mean differences were statistically significant; however, no individual mean comparisons yielded significant results. Specifically, no significant difference was found between higher and lower empathy groups for SBP during the comfort task, $F(1, 29) = .63, p = .44, \eta_p^2 = .02$, or the neutral task, $F(1, 29) = 1.44, p = .24, \eta_p^2 = .05$. For the comfort

task, covariance adjusted mean SBP for the lower empathy group was 125.8 mm Hg, $SE = 2.66$, and for the higher empathy group was 128.8 mm Hg, $SE = 2.66$. For the neutral task, covariance adjusted mean SBP for the lower empathy group was 127.2 mm Hg, $SE = 3.12$, and for the higher empathy group was 132.5 mm Hg, $SE = 3.12$. (See Figure 1). The comparison between the neutral and comfort task SBP reactivity for those in the lower empathy group revealed no significant difference, $F(1, 14) = .13, p = .72, \eta_p^2 = .01$. Likewise, the comparison between the neutral and comfort task SBP reactivity for those in the higher empathy group revealed no significant difference, $F(1, 14) = 3.55, p = .08, \eta_p^2 = .20$.

Diastolic Blood Pressure. Analyses revealed no significant main effect of Empathy Group on DBP, $F(1, 29) = .10, p = .76, \eta_p^2 = .00$, no significant main effect of Task Type, $F(1, 29) = .01, p = .91, \eta_p^2 = .00$, nor a significant interaction between Empathy Group and Task Type, $F(1, 29) = .17, p = .69, \eta_p^2 = .01$.

Heart Rate. Analyses revealed no significant main effect of Empathy Group on HR, $F(1, 29) = .89, p = .35, \eta_p^2 = .03$, nor a significant interaction between Empathy Group and Task Type, $F(1, 29) = .09, p = .76, \eta_p^2 = .00$. There was a significant main effect of Task Type, $F(1, 29) = 7.73, p < .01, \eta_p^2 = .21$. Specifically, mean covariance adjusted HR for the comfort task was $M = 85.7$ bpm, $SE = 1.17$, and mean covariance adjusted HR for the neutral task was $M = 88.4$ bpm, $SE = 1.10$.

Low Frequency Heart Rate Variability. Analyses revealed neither a significant main effect of Empathy Group on LF-HRV, $F(1, 29) = .57, p = .46, \eta_p^2 = .02$, nor a significant main effect of Task Type on LF-HRV, $F(1, 29) = 1.11, p = .30, \eta_p^2 = .04$. There was also no significant interaction between Empathy Group and Task Type, $F(1, 29) = .60, p = .44, \eta_p^2 = .02$.

High Frequency Heart Rate Variability. Analyses revealed no significant main effect of Empathy Group on HF-HRV, $F(1, 29) = 1.35, p = .26, \eta_p^2 = .04$, or a significant main effect of Task Type, $F(1, 29) = .01, p = .94, \eta_p^2 = .00$, but revealed a significant interaction between Empathy Group and Task Type, $F(1, 29) = 4.92, p = .04, \eta_p^2 = .15$. Post-hoc simple main effect analysis was conducted and yielded no significant findings. Specifically, no significant difference was found between means of higher and lower empathy groups for HF-HRV measures during the comfort task, $F(1, 29) = 0.00, p = .98, \eta_p^2 = .00$, and during the neutral task, $F(1, 29) = 3.08, p = .09, \eta_p^2 = .10$. For the comfort task, covariance adjusted mean HF-HRV for the lower empathy group was 6.29 log ms², SE = .157, and for the higher empathy group, mean HF-HRV was 6.28 log ms², SE = .157. For the neutral task, covariance adjusted mean HF-HRV for the lower empathy group was 6.10 log ms², SE = .159, and for the higher empathy group, mean HF-HRV was 6.50 log ms², SE = .159 (See Figure 2). The comparison between the neutral and comfort task reactivity for those in the lower empathy group revealed no significant difference, $F(1, 14) = 2.02, p = .18, \eta_p^2 = .13$. Likewise, the comparison between the neutral and comfort task reactivity for those in the higher empathy group revealed no significant difference, $F(1, 14) = 1.75, p = .21, \eta_p^2 = .11$.

Root Mean Square of Successive Differences (between normal heartbeats). Analyses revealed neither a significant main effect of Empathy Group on RMSSD, $F(1, 29) = 3.97, p = .06, \eta_p^2 = .12$, nor a significant main effect of Task Type on RMSSD, $F(1, 29) = 1.40, p = .25, \eta_p^2 = .05$. In addition, there was no significant interaction between Empathy Group and Task Type, $F(1, 29) = .29, p = .59, \eta_p^2 = .01$.

Primary Analyses: Affective Reactivity to Tasks by Empathy Group

For the positive PANAS scores, the main effect of Empathy Group was not significant, $F(1, 29) = .52, p = .48, \eta_p^2 = .02$, and there was no significant interaction between Empathy Group and Task Type, $F(1, 29) = .22, p = .65, \eta_p^2 = .01$. There was, however, a significant main effect of Task Type, $F(1, 29) = 10.23, p < .01, \eta_p^2 = .26$. The covariance adjusted mean positive PANAS score during the comfort task was $M = 22.4, SE = .91$ and the covariance adjusted mean positive PANAS score during the neutral task was $M = 26.1, SE = 1.14$.

For the negative PANAS scores, the main effect of Empathy Group was not significant, $F(1, 29) = .08, p = .78, \eta_p^2 = .00$, and there was no significant interaction between Empathy Group and Task Type, $F(1, 29) = .64, p = .43, \eta_p^2 = .02$. There also was a significant main effect of Task Type, $F(1, 29) = 4.83, p < .05, \eta_p^2 = .14$. The covariance adjusted mean negative PANAS score during the comfort task was $M = 14.9, SE = .65$, and the covariance adjusted mean negative PANAS score during the neutral task was $M = 12.7, SE = .63$.

Primary Analyses: Interpersonal Goals during Tasks by Empathy Group

Results indicated that for both agentic and communal striving, there was a significant main effect for Task Type, $F(1, 30) = 18.99, p < .001, \eta_p^2 = .39$, and $F(1, 30) = 26.76, p < .001, \eta_p^2 = .47$, respectively. For agentic striving, mean scores were greater for the neutral task, $M = 9.6, SE = .48$, as compared to the comfort task, $M = 7.1, SE = .39$. For communal striving, mean scores were greater for the comfort task, $M = 13.0, SE = .37$, as compared to the neutral task, $M = 10.7, SE = .49$. (See Table 3). Neither main effects for Empathy Group nor the Empathy Group by Task Type interactions were significant for measures of agentic or communal striving.

Primary Analyses: Task Appraisals by Empathy Group

Analyses were conducted on self-reported appraisals of the task, and a series of mixed design 2 (Empathy Group: Higher, Lower) by 2 (Task Type: Neutral, Comfort) ANOVAs were conducted to determine whether there were differences in appraisals across Task Types and between higher and lower Empathy Groups. The main effect of Empathy Group was only observed for the item measuring likelihood of encountering the situation in daily life, and main effects for Task Type were observed for the same item, as well as three other items regarding the stress caused by the interaction. All other main effects and interactions were not significant. (See Table 4).

“How likely is it that you would encounter this situation in your daily life?” The main effect of Task Type and the main effect of Empathy Group were significant. Specifically, for Task Type, mean rating of likelihood for the comfort task was $M = 4.3$, $SE = .15$, and mean rating of likelihood for the neutral task was $M = 3.7$, $SE = .20$. For Empathy Group, the mean rating of likelihood for the lower empathy group was $M = 3.7$, $SE = .18$, and the mean rating of likelihood for the higher empathy group was $M = 4.3$, $SE = .18$.

“How stressful would it be to handle an interaction like this?” The main effect of Task Type was significant. For the comfort task, the mean rating of stressfulness was $M = 2.7$, $SE = .20$, and for the neutral task, the mean rating of stressfulness was $M = 1.8$, $SE = .16$.

“How stressful was it for you to engage in this interaction?” The main effect of Task Type was significant. For the comfort task, the mean rating of stressfulness was $M = 2.8$, $SE = .19$, and for the neutral task, the mean rating of stressfulness was $M = 2.3$, $SE = .23$.

“How difficult was the interaction?” The main effect of Task Type was significant. Specifically, the mean rating of difficulty for the comfort task was $M = 2.7$, $SE = .23$, and the mean rating of difficulty for the neutral task was $M = 2.2$, $SE = .15$.

Behavioral Measures

A series of 2 (Empathy Group) x 2 (Task Type) mixed factors ANOVAs were conducted on 6 of the 8 major behavioral parameters (not including Tangible Aid or Tension Reduction). (See Table 5). Mixed factors ANOVAs were utilized when possible to compare behavioral differences within task types between empathy groups; however, some parameters and codes required individual analysis due to issues with skewness of scores for one or both task types. In instances where the skewness of parameters or codes was not able to be corrected with a transformation, Mann-Whitney U non-parametric tests were utilized to compare groups.

Tension Reduction scores for the comfort task were heavily skewed and could not be corrected using a transformation, requiring a Mann-Whitney U test. For the neutral task, the scores were log transformed and then subjected to a one-way ANOVA. For Tangible Aid, both comfort and neutral task scores were heavily skewed and could not be corrected using a transformation; thus, this parameter was subjected to a Mann-Whitney U test.

Several *individual* behavioral codes were subjected to repeated measures one-way ANOVAs (when possible) or Mann-Whitney U tests, including: Verbal Agreement, Disagree/Disapprove, Not Tracking, Normative Verbal, and Normative Non-Verbal codes. For Verbal Agreement specifically, scores for the comfort task were heavily skewed and could not be corrected by a transformation, requiring a Mann-Whitney U test. For the neutral task, the scores were successfully log transformed and subjected to a one-way ANOVA. Additionally, Disagree/Disapprove was only analyzed for the comfort task using a one-way ANOVA, as this

code did not occur for any participant during the neutral task. Finally, scores for the Not Tracking code were log transformed for both the comfort and the neutral tasks. Normative Verbal, Normative Non-Verbal, and Not Tracking were all subjected to a mixed design ANOVA analyses.

Information Support. Neither the main effect of Empathy Group nor the interaction between Empathy Group and Task Type were significant; however, the main effect of Task Type was significant, $F(1, 30) = 4.33, p < .05, \eta^2 = .126$. Specifically, the mean score on information support for the comfort task was $M = 7.1, SE = .55$, which was less than the mean score for the neutral task ($M = 8.4, SE = .61$).

Emotional Support. Neither the main effect of Empathy Group nor the interaction between Empathy Group and Task Type were significant; however, the main effect of Task Type was significant, $F(1, 30) = 11.10, p < .01, \eta^2 = .270$. Specifically, the mean score on emotional support for the comfort task was $M = 5.4, SE = .48$, less than the mean score for the neutral task ($M = 7.1, SE = .48$).

Esteem Support. Neither the main effect of Empathy Group nor the interaction between Empathy Group and Task Type were significant; however, the main effect of Task Type was significant, $F(1, 30) = 69.47, p < .001, \eta^2 = .698$. Specifically, the mean score on esteem support for the comfort task was $M = 3.1, SE = 1.10$, which was greater than the mean score for the neutral task ($M = 1.3, SE = 1.07$).

Social Network Support. Neither the main effect of Task Type nor the interaction between Empathy Group and Task Type were significant; however, the main effect of Empathy Group was significant, $F(1, 30) = 5.09, p < .05, \eta^2 = .145$. Specifically, the mean score on social

network support for those in the higher empathy group was $M = 1.7$, $SE = 1.10$, which was higher than the mean score for those in the lower empathy group ($M = 1.3$, $SE = 1.10$).

Attentiveness. Neither the main effect of Task Type, nor the main effect of Empathy Group, and the interaction between Empathy Group and Task Type were significant.

Positive Non-Verbal. Neither the main effect of Empathy Group nor the interaction between Empathy Group and Task Type were significant. However, the main effect of Task Type was significant, $F(1, 30) = 49.56$, $p < .001$, $\eta^2 = .623$. Specifically, the mean score on positive nonverbal codes for the comfort task was $M = 11.0$, $SE = .83$, lower than the mean score for the neutral task at $M = 16.6$, $SE = .69$.

Tangible Aid. A Mann-Whitney non-parametric analysis was used to determine whether there were differences in Tangible Aid scores between those deemed higher and lower in empathy. For both the comfort task ($U = 120.00$, $p = .780$) and the neutral task ($U = 120.00$, $p = .780$), no significant differences between scores were detected between empathy groups. Specifically, mean score for the higher empathy group was $M = 0.1$, $SE = .04$ for the comfort task and $M = 0.0$, $SE = .09$ for the neutral task. For the lower empathy group, mean score was $M = 0.0$, $SE = .04$ for the comfort task and $M = 0.1$, $SE = .09$ for the neutral task.

A Wilcoxon Signed Ranks Test was then conducted to determine whether there were differences in Tangible Aid scores across task types. Results indicated that there was no statistically significant difference between scores across task types ($Z = -.447$, $p = .655$).

Tension Reduction. A Mann-Whitney non-parametric analysis was used to determine whether there were differences in Tension Reduction scores on the comfort task (the neutral task was subjected to a one-way ANOVA) between those deemed higher and lower in empathy. No

significant difference between scores was detected between empathy groups, $U = 101.50$, $p = .323$.

A one-way ANOVA was utilized to determine whether there were differences in Tension Reduction scores on the neutral task between Empathy Groups. Results indicated no significant differences between empathy groups, $F(1, 30) = .976$, $p = .331$.

Verbal Agreement. The Verbal Agreement code was analyzed separately for the comfort and neutral task due to issues with the skewness of the data from the comfort Task. For the comfort task, no significant results were found for Empathy ($U = 120.00$, $p = .780$).

The neutral task was analyzed utilizing a one-way ANOVA. No significant results were found for Empathy Group, $F(1, 30) = .102$, $p = .752$.

Disagree/Disapprove. The Disagree/Disapprove code was only analyzed for the comfort task, as this code did not occur during the neutral task. A one-way ANOVA was conducted but no significant main effect was found for Empathy Group, $F(1, 30) = .047$, $p = .830$,

As previously stated, several of the individual codes were either normally distributed or fixed via a log transformation for both tasks and thus were able to be analyzed utilizing a mixed design repeated measures ANOVA. These three codes included Not Tracking (Negative Non-Verbal), Normative Verbal, and Normative Non-Verbal. (See Table 6).

Not Tracking. Scores for the Not Tracking code were log transformed for both the neutral and the comfort task prior to analyses. An Empathy Group by Task Type mixed design ANOVA revealed no significant difference in scores between Empathy Groups, across Task Types, nor a significant interaction between the two.

Normative Verbal. A mixed design ANOVA revealed no significant difference in scores between Empathy Group, but did reveal a significant difference between Task Types, $F(1, 30) =$

4.90, $p < .05$, $\eta^2 = .140$. Specifically, the mean score for the neutral task was $M = 1.91$, $SE = .346$, lower than the mean score for the comfort task ($M = 2.75$, $SE = .319$).

Normative Non-Verbal. A mixed design ANOVA revealed no significant difference in scores between Empathy Group or across Task Type, nor a significant interaction between the two.

Number of Words Spoken. In order to determine whether there was a difference in the number of words spoken by each participant during the 3-minute interactions as a function of empathy level, a 2 (Empathy Group by 2 (Task Type) mixed factors ANOVA was conducted. Results indicated that the main effect of Task Type was significant, $F(1, 30) = 82.17$, $p < .001$, $\eta^2 = .733$, but no significant main effect was detected for empathy level, and no significant interaction was found. For the Task Type main effect, mean number of words spoken for the neutral task was $M = 396.2$, $SE = 13.42$, greater than the mean number of words spoken for the comfort task ($M = 291.4$, $SE = 14.77$).

Overall Empathy Rating. Finally, empathy ratings of each participants' perceived level of empathy displayed (as rated by the 2 independent coders after viewing each interaction) were subjected to a Mann-Whitney non-parametric analysis to determine whether there were differences in empathy ratings between those deemed higher and lower in empathy on a self-report measure. For both the comfort task ($U = 110.50$, $p = .463$) and the neutral task ($U = 108.50$, $p = .418$), no significant differences between behavioral empathy ratings were detected between empathy groups.

Discussion

The literature examining the relation between empathy and cardiovascular reactivity is broad and employs a wide range of participant demographics, measures of cardiovascular

reactivity, and emotionally-evocative stimuli. The purpose of this study was to examine the association between dispositional empathy and autonomic reactivity and to address the shortcomings of previous studies by utilizing an in vivo, dyadic interaction between a participant and a confederate as the stimuli. It was hypothesized that during the comfort task specifically, those in the higher empathy group in contrast to lower empathy counterparts would expend a greater amount of “emotional effort” and thus would display increased autonomic reactivity as measured by measures of HR, HRV, and BP. No evidence to support this primary hypothesis was observed, as none of the six measures of cardiovascular reactivity significantly differed between higher and lower empathy groups. There were, however, two instances in which significant interactions between Empathy Group and Task Type were detected, but simple main effects were not detectable. Specifically, although no significant differences were detected between higher and lower empathy groups for the comfort task or the neutral task, a trend indicated that those in the higher empathy group had slightly higher SBPs during the neutral versus comfort tasks, a trend not observed among those in the lower empathy group. For HF-HRV, again, simple main effects were not significant, but a trend revealed such that during the neutral task, HF-HRV was higher for the higher empathy compared to the lower empathy group. This pattern of findings is perplexing because it suggests that those higher in dispositional empathy exhibit slightly both higher SBP and HF-HRV responses to the neutral but not comfort tasks. Although these observed findings suggest that those in the higher empathy group *may* have been slightly more physiologically reactive to the neutral interaction, the response profile indicates that both branches of the autonomic nervous system were involved. Although increased SBP is known to be associated with both increased sympathetic nervous system activity and/or reduced parasympathetic nervous system activity, increased HF-HRV is known to

be associated with increased parasympathetic activation (Berntson et al., 1997). The observed response profile indicates that both branches of the autonomic nervous system were activated for higher empathy participants while engaged in the neutral interaction, and the sympathetic activation was more robust than the parasympathetic response. The fact that this response profile was only detected during the neutral interaction and not the comfort task is also noteworthy. Because the neutral task did not involve instructions to empathize with the confederate or include themes of empathy, it is possible that differences in responding are only observed in contexts that do not involve being empathic. Again, simple main effect analyses were not significant so this finding should be interpreted with caution.

Previous studies examining the relation between empathy and autonomic reactivity have found mixed results. The majority have found that higher trait empathy was associated with increased autonomic nervous system activity in response to emotional stimuli (Ardizzi et al., 2016; Balconi & Bortolotti, 2012; de Wied, et al., 2012; Geringer, 2015; Liew et al., 2011; Miller, et al., 2016; Oliveira-Silva & Gonclaves, 2011; Tracy & Giummarra, 2017; Truzzi et al., 2016; Wiesenfeld et al., 1984; Zahn-Waxler et al., 1995). Others like Stellar and colleagues (2015) found that increased expression of compassion for another was associated with decreased autonomic reactivity. Still others found mixed results such that increases in empathy were associated with increases in some measures of reactivity but decreases in others (Miu & Baltes, 2012; Vanderpool & Barrat, 1970). Still others failed to detect an association between autonomic activity measures and empathy (Anastassiou-Hadjicharalambous & Warden, 2007; Ono et al., 2012; Perrone-McGovern, 2014). Of the three prior studies that employed in vivo interactions, one found a significant positive association between empathy and autonomic reactivity (Zahn-Waxler et al., 1995), but the other two studies failed to find a significant association between the

two variables (Ono et al., 2012; Perrone-McGovern, 2014). In this regard, findings from the current study are consistent with the results observed by Ono et al. and Perrone-McGovern.

Because self-reported measures of affect, interpersonal goals, and task appraisals were measured along with coding of interpersonal behaviors in the current study, it is important to examine any differences observed between higher and lower empathy groups on any of these parameters when interpreting findings on the physiological responses. In brief, very few differences between higher and lower empathy groups were observed. No group differences were observed on either positive or negative affect responses or agentic or communal interpersonal strivings to the two tasks. The only difference between higher and lower empathy groups on task appraisals was for the item assessing likelihood of encountering these situations in daily life. Participants in the higher empathy group rated a significantly higher likelihood of encountering these sorts of situations in their daily life than participants in the lower empathy group, perhaps reflecting a greater propensity of others to seek out support or advice from friends displaying higher empathy traits. The only difference between groups on behavioral measures was a higher score on provision of social network support for participants in the higher empathy group in comparison to those in the lower empathy group. Specifically, behaviors which fell under the social network support parameter included Presence (offer to “be there” or spend time with the person), Access (offer to introduce the individual to new companions), and Companions (discussing the experience of others that have been through the same situation). In all other regards, women in the higher empathy and lower empathy groups responded comparably to the two interpersonal tasks.

The current study failed to detect significant differences between higher and lower empathy groups on all measures of cardiovascular reactivity as well as affective reactivity and

interpersonal strivings during each task. Ratings of likelihood of encountering tasks of this type and behavioral codes involving the provision of social support suggest that women higher in empathy engage in specific types of supportive behaviors and are more likely to encounter situations involving support and advice than women lower in trait empathy.

Task Differences

In addition to understanding whether cardiovascular reactivity differed between persons exhibiting higher and lower dispositional empathy, study analyses also explored whether participants displayed differential reactivity between empathy and neutral tasks. The only cardiovascular parameter for which a significant main effect of Task Type was observed was the measure of HR; mean HR was higher during the neutral task overall as compared to the comfort task. This may be due to the fact that the neutral task required and elicited more spoken words from the participant than the comfort task, as the neutral task involved persuasion and discussion of pros and cons of attending a university, while the comfort task tended to require more listening.

In addition, there were significant differences in self-reported affect between tasks. Positive PANAS scores were higher overall for the neutral task compared to the comfort task, and Negative PANAS scores were higher for the comfort task as compared to the neutral task. This intuitively makes sense based on the differing themes of those interactions, as the neutral task involved a discussion of college choice while the comfort task involved discussing heartbreak with a sad confederate. Also, analyses of interpersonal goals revealed that agentic strivings were higher for the neutral task than the comfort task, and communal strivings were greater during the comfort than the neutral tasks. These results were also expected, as the neutral task involved achieving a goal of persuading the confederate to agree to attend a college, while the comfort task required listening and understanding on behalf of the participant.

For the post-task questionnaire task appraisal items, several items specifically differed between tasks. All participants reported that they were more likely to encounter the comfort task than the neutral task but also rated the comfort task as being more stressful to handle, more stressful to engage in, and more difficult than the neutral task.

Several behavioral codes significantly differed between tasks. Specifically, participants offered more Information and Emotional Support, and displayed more Positive Nonverbal behaviors during the neutral task as compared to the comfort task. In contrast, participants offered more Esteem Support and displayed more Normative Verbal behaviors during the comfort task as compared to the neutral task.

Because participants viewed the comfort task as being more stressful and more difficult than the neutral task, as well as rating it as eliciting more negative and less positive affect, it is unlikely that these factors contributed to the increased HR reactivity that was observed during the neutral task. It is possible that simply speaking more during the neutral task than during the comfort task led to an increase in HR. It is also possible that tasks associated with more agentic goals evoke greater HR reactions than tasks associated with communal goals. Finally, the different types of support provided during the two tasks (information and emotion support for the neutral task and esteem support for the comfort task) may be associated differentially with HR responding to stress.

Confederate Effects

Two confederates were used in this study. Although they were both young, White/Caucasian undergraduate women comparable in age to study participants and were extensively trained in the study protocol, it was important to examine whether they elicited comparable responses from study participants. Few differences were observed between

participants when interacting with Confederate A versus Confederate B. Overall results indicated that during the comfort task, participants interacting with Confederate A showed less Esteem Support, reported getting less upset, and exhibited lower SBP responses than participants interacting with Confederate B. Because no other Confederate main effects were observed, the age and other demographic characteristics of the two confederates appeared to be well matched. However, because a few Confederate by Task Type interactions revealed differences in participant responding to the comfort task, it appears that interacting with Confederate A potentially evoked less emotion than interacting with Confederate B. Although the current study was not powered adequately to examine whether this influenced the primary study hypotheses, future empirical work could address this interesting possibility.

Task Order Effects

Like the examination of confederate effects, it was also important to examine any potential order effects, given that half the participants completed the neutral task prior to the comfort task (Order 1) and the other half completed the comfort task prior to the neutral task (Order 2). Most of the order effects observed were qualified by Task Type. During the rest periods, those in Task Order 2 exhibited higher DBP during the pre-neutral as compared to the pre-comfort rest period, and higher HF-HRV during the pre-comfort as compared to the pre-neutral rest period. These findings suggested that participants were likely experiencing residual arousal from the comfort task during the pre-neutral rest period. Participants who completed the neutral task first did not experience this sustained arousal that influenced the rest period following the first task.

There were mixed observations regarding order effects. On the one hand, there was some indication that completing the neutral task first (Order 1) elicited greater reactions than

completing the comfort task first (Order 2). For example, participants completing the neutral task first exhibited higher HR during the neutral task than those completing the comfort task first. Furthermore, self-reported positive affect of participants completing the study in Order 1 reported lower positive affect across both tasks than participants completing the study in Order 2.

On the other hand, there were several Order by Task Type interactions that only revealed significant task effects for participants completing the study in Order 2. For example, participants assigned to Order 2, reported higher ratings of negative affect during the comfort task than the neutral task, a difference not observed for participants assigned to Order 1. Likewise, participants completing the study in Order 2 rated the neutral task as being more positive, less negative, more realistic, less stressful, less upsetting, and less difficult compared to the comfort task. This was expected as the comfort task involved a vague goal to “Comfort this person” and was designed to elicit emotion. In contrast, the neutral task resembled a less emotional discussion focusing on the sharing of facts and information. It may have also felt less stressful to engage in the neutral task because participant knowledge of whether they achieved their goal was evident (the confederate either agrees or disagrees to attend WVU for school), whereas understanding whether you comforted someone is not so clear. It is noteworthy that significant differences in task appraisal only occurred among those completing the study in Order 2. It is unclear why those participants who completed the neutral task first (Order 1) did not exhibit the same findings.

The significant interaction between Task Order and Task Type on number of words spoken revealed that participants in both orders spoke significantly more words during the neutral task as compared to the comfort task. This difference, however, was more pronounced for those completing the study in Order 2. Perhaps being asked to comfort someone you have

just met is more taxing than being asked to comfort someone after you have discussed a different “neutral” topic and developed some understanding of the other individual’s personality and background. Being asked to comfort someone before getting to know them even at a basic level may lead to feeling more uncomfortable and saying less as a result of not knowing what to say to the individual.

Strengths and Limitations

The current study exhibited several strengths. This study employed a series of in vivo dyadic interactions that had been largely lacking in the previous literature. Use of this type of emotionally evocative stimuli approaches greater ecological validity as the tasks sought to mimic the types of interactions that individuals might have in their daily lives. An additional strength of the current study was the use of not only physiological measures of cardiovascular reactivity but also the addition of both self-reported affect, interpersonal goals, and task appraisals as well as behavioral measures. Combining the results of these three methods of measurement provides a more comprehensive understanding of not only the physiological mechanisms at play but also the affective, cognitive, and behavioral components as well.

Though the current study exhibited a number of strengths, it also contained several limitations or weaknesses. The most obvious limitation is a result of the generalizability of the results to other samples or populations. This study was conducted on a sample of college-aged, predominantly White/Caucasian females (with two White/Caucasian female confederates). As a result, we are unable to predict how results may differ if the same study had been conducted using an all-male sample (with male confederates), or with a mix of both male and female participants along with both male and female confederates. Perhaps empathic behavior looks different between two females and two males, or between a female and male dyad. Likewise, it is

unknown whether comparable findings would be observed among samples with different racial/ethnic compositions or ages. Another implication of the use of only females is that we may have artificially limited the range of BP and HR measurements available for analysis, as female measures of these physiological parameters generally differ from males. Another limitation of the study is that the main effect of empathy measurement time revealed that self-reported empathy scores were not entirely consistent across SONA and laboratory session measurement times, with some individuals even exhibiting changes in scores that moved them from the lower to the higher empathy group. Specifically, the SONA (pre-laboratory session) scores were lower than mean empathy scores on the same empathy questionnaire measured in the laboratory. For those in the Lower Empathy Group only, SONA scores were significantly lower than laboratory scores (no difference was observed for those in the Higher Empathy Group). As the laboratory measure of empathy was given at the end of the laboratory session, perhaps being asked or instructed to comfort the confederate in the lab (meant to evoke empathic behaviors) influenced participants to indicate that they are more empathic as these themes were made more salient from their recent interactions and artificially inflated self-report ratings of empathic tendencies. Of course, another plausible explanation could be due to demand characteristics or social desirability in that participants did not wish to label themselves as being “un-empathic.” Another limitation is that the test-retest period between initial SONA and final laboratory BES completion was on average 9 days longer than the 7 week/49 day test-retest period described in the literature. Self-report ratings of trait empathy may not be reliable after a period of time greater than 7 weeks; however, the test-retest correlation coefficients observed on our study sample appeared comparable to those reported in the prior study. A related limitation of the study is a potential weakness of the empathy scale that was used to categorize participants. As scores

were not consistent between SONA and laboratory completion, it is unclear whether the questionnaire adequately measures empathy as a “trait” or disposition in this sample. As only one measure of trait empathy was utilized in this study, there were no additional empathy questionnaire results with which to compare. Though participants were classified into Empathy Groups based on the expectation that higher and lower empathy scale scores represented *trait* empathy, perhaps participants answered the items according to their empathic *state* at the time of questionnaire completion. It is important to consider the difference between state and trait empathy, as perhaps one type of empathy dominates in familiar versus unfamiliar contexts and interactions. An additional limitation is related to the fact that the researchers who rated participants on behavioral displays of empathy were not blinded to participant empathy scores, a factor which may have biased researcher behavioral ratings; this bias if present, however, did not appear to influence study findings as very few behavior codes showed significant group differences between higher and lower empathy participants. Another limitation of this study is that participants were asked to interact with and comfort an individual whom they had never met. More often than not, individuals are placed in situations in which they are more familiar and have a history of a personal relationship with the individual that comes to them for comfort. Perhaps no reactivity differences were observed between Empathy Groups because both groups were asked to interact with and comfort a stranger, a task that may be difficult regardless of the participants’ level of empathy. An additional study limitation is related to the sample size of the study. Given that the sample size was small ($N = 32$), there was adequate statistical power to test main hypotheses but limited power with which to test interactions between Empathy Group, Task Order, and Task Type or Confederate, for example. Thus results based on these interactions are unknown. A final limitation is that participants were asked to engage in an interaction with

an unknown confederate, in a well-lit room while being video-recorded. Feelings of being “on the spot” may have caused participants to feel self-conscious, and thus influenced participants to behave more or less empathic than usual.

Future Directions

The results and limitations of this study indicate that there are a number of variables both interpersonal and situational that may influence the way that individuals do or do not express empathy for another. One area to target in future work in this area may be the type of interaction that the participant is asked to complete. Our study implemented a comfort task as the main task of interest, but this task only tapped one human emotion – sadness. Perhaps a true empathy task requires the participant to empathize with not only sadness, as in the comfort task, but with a variety of other emotions as well (e.g., a confederate expressing joy, anger, or fear). Findings may have been different if higher and lower empathy groups had been asked to empathize with some other category of emotion. Studies in the literature that have found significant associations between empathy and autonomic reactivity have included scenes with either sad/distressing stimuli (e.g. Zahn-Waxler et al., 1995) or a mixture of joyous/sad stimuli (e.g. Ardizzi et al., 2016). Although it is important to consider the nature of the emotion elicited in the current study, the emotional valence of the task stimuli has not been shown to influence the pattern of findings observed in previous studies. In addition, another way to conceptualize differences between the task types is describing the comfort task as communal in nature and the neutral task as agentic in nature. In this way perhaps it was not the specific content or theme of the interactions that influenced results but the overall type of goal that the interaction instructions evoke (being there for someone versus persuading them). Relatedly, future studies could consider the way in which the tasks or interactions are introduced to the participant. Perhaps

offering instructions to comfort the confederate influenced findings, and simply introducing the confederate and letting conversation flow naturally would feel more realistic. Use of deception should be considered in future work in this area as this could influence empathic behaviors and increase ecological validity. For example, stating that the confederate had truly gone through a break-up may make the interaction more real than just being a “role play.”

In addition to modifying the primary task of interest, changes may also be warranted for the “control” task, in this case, the neutral task. Though the task was meant to serve as a control interaction to compare to the comfort task, it is possible that this task did not function as a truly “neutral” task. Participants were asked to *persuade* the confederate into making an important academic decision, and maybe giving a more ambiguous instruction for the participant to “get to know” the confederate would have made for a better “non-emotional” neutral comparison task.

Another variable that future studies might consider is examining different demographic study samples. As previously mentioned, the exclusive use of women in the current study obviously decreases the generalizability of the findings, and future research could consider not only male-male interactions but female-male interactions as well. In addition, demographic variables of age and race may influence the nature of these types of interactions. As both confederates as well as the majority of the sample were primarily Caucasian, future studies should examine interactions in a more diverse sample and investigate whether race affects self-report and behavioral measures of empathy, as well as physiological measures. In addition, studies could employ same versus different race interactions in order to understand how results might differ in relation to different racial confederate/participant pairings. An additional variable of interest is age of the participant and confederates. Future studies may investigate whether

small versus large age differences between confederates and participants influence results, and how expression of empathic behavior may differ as a result of age.

The nature of the relationship between the participant and the confederate is an additional variable to consider in future work in this area. The current study required participants to interact with an individual with whom they were not familiar; however, perhaps empathy and empathic behaviors differ as a result of familiarity with the target individual. Future studies might recruit participants who are higher or lower in self-reported empathy who come to the laboratory with a friend in order to observe empathic behaviors and reactivity in these types of interactions. Future studies could alternatively consider the benefit of implementing a pre-study practice interaction in which unfamiliar participants and confederates have a chance to get to know one another before engaging in the interaction of interest. Relatedly, future studies could consider ratings of participant levels of anxiety and how it relates to engagement in empathic behaviors or reactions to emotional stimuli. Perhaps participants deemed higher or lower in empathy feel more or less social or general anxiety when interacting with an unknown confederate and consideration of moderation by this potentially important construct is worthy of future study.

Another avenue for future studies would be to examine behavioral measures of actual observed empathy in order to more accurately and objectively classify individuals as higher or lower in empathy for others, rather than relying on self-reported measures of empathy. The current study observed an inconsistency in self-reports of empathy between screener and laboratory measurement times; thus, relying on self-report measures alone may not represent the most objective measure of empathy.

Summary and Implications

The current study was designed to incorporate an in-vivo dyadic interaction between participant and confederate to allow for a more ecologically valid testing of emotional reactivity to interpersonal stress and empathy. In addition to measuring empathy and cardiovascular reactivity, this study also incorporated measures of self-reported affect, task appraisals, and behaviors to understand how participants responded to each task emotionally, cognitively, behaviorally, and physiologically. Though no significant effects of Empathy Group were observed on measures of cardiovascular reactivity, this study provides a framework for understanding the ways in which interpersonal interactions with another may influence our autonomic reactivity and ultimately our health.

As students in health care professional training programs eventually pursue careers that require them to interact daily with numerous patients, it is important to consider the factors that may increase or decrease their tendency to offer empathic listening and understanding. Though the current study did not find any support for negative influences of empathy, discovering that those higher in empathy are more susceptible to increased autonomic reactivity, for example, would be useful in informing training health professionals about the physiological consequences of their future careers. Should future research confirm this hypothesis, those identified as being higher in empathy might undergo skills training designed to help them *internally* process patient outcomes objectively while *externally* expressing warmth and concern. It is also important to consider situational factors, such as level of familiarity with the patient and the patient's family. Familiarity and quality of previous interactions with others may lead one to exhibit increased or decreased empathic responding. Just as well, it is relevant to consider thematic factors related to the nature of patient-provider interactions that involve different emotionally charged situations

(e.g., a provider offering news of a feared diagnosis versus a provider offering news of disease remission). Finally, it is important to consider these and more factors to understand the ways in which empathic behaviors affect not only patient outcomes, but provider stress and cardiovascular health parameters as well.

References

- Alligood M. R. (1992) Empathy: the importance of recognizing two types. *Journal of Psychosocial Nursing* 30, 14-17. doi.org/10.3928/0279-3695-19920301-06
- Allport, G. W., Vernon, P. E., & Lindzey, G. (1951). *Study of values manual of directions* (Rev. ed.). Boston: Houghton Mifflin, 1951
- Anastassiou-Hadjicharalambous, X., & Warden, D. (2007). Convergence between physiological, facial and verbal self-report measures of affective empathy in children. *Infant and Child Development*, 16(3), 237-254. doi: 10.1002/icd.464
- Ardizzi, M., Umiltà, M. A., Evangelista, V., Di Liscia, A., Ravera, R., & Gallese, V. (2016). Less empathic and more reactive: the different impact of childhood maltreatment on facial mimicry and vagal regulation. *PLoS One*, 11(9), e0163853. <https://doi.org/10.1371/journal.pone.0163853>
- Balconi, M., & Bortolotti, A. (2012). Resonance mechanism in empathic behavior: BEES, BIS/BAS and psychophysiological contribution. *Physiology & Behavior*, 105(2), 298-304. <https://doi.org/10.1016/j.physbeh.2011.08.002>
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34(2), 163-175. <https://doi.org/10.1023/B:JADD.0000022607.19833.00>
- Batson, C. D., Batson, J. G., Slingsby, J. K., Harrell, K. L., Peekna, H. M., & Todd, R. M. (1991). Empathic joy and the empathy-altruism hypothesis. *Journal of Personality and Social Psychology*, 61(3), 413. <http://dx.doi.org/10.1037/0022-3514.61.3.413>

- Berntson, G. G., Thomas Bigger Jr, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., ... & Van Der Molen, M. W. (1997). Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, *34*(6), 623-648.
- Boehm, J. K., & Kubzansky, L. D. (2012). The heart's content: the association between positive psychological well-being and cardiovascular health. *Psychological Bulletin*, *138*(4), 655. <http://dx.doi.org/10.1037/a0027448>
- Brems, C. (1989). Dimensionality of empathy and its correlates. *Journal of Psychology*, *123*, 329-337. <https://doi.org/10.1080/00223980.1989.10542989>
- Carré, A., Stefaniak, N., D'ambrosio, F., Bensalah, L., & Besche-Richard, C. (2013). The Basic Empathy Scale in Adults (BES-A): Factor structure of a revised form. *Psychological Assessment*, *25*(3), 679.
- Cohen, D., & Strayer, J. (1996). Empathy in conduct-disordered and comparison youth. *Developmental Psychology*, *32*, 988-998
- Cole, P. M., Jordan, P. R., & Zahn-Waxler, C. (1990). Mood induction stimulus for children. Bethesda, MD: National Institute of Mental Health.
- Derksen, F., Bensing, J., & Lagro-Janssen, A. (2013). Effectiveness of empathy in general practice: a systematic review. *British Journal of General Practice*, *63*(606), e76-e84. <https://doi.org/10.3399/bjgp13X660814>
- de Wied, M., van Boxtel, A., Matthys, W., & Meeus, W. (2012). Verbal, facial and autonomic responses to empathy-eliciting film clips by disruptive male adolescents with high versus low callous-unemotional traits. *Journal of Abnormal Child Psychology*, *40*(2), 211-223. [doi:10.1007/s10802-011-9557-8](https://doi.org/10.1007/s10802-011-9557-8)

- Dillard, J. P., & Hunter, J. E. (1986). Questions about the construct validity of three scales: Emotional empathy, self-consciousness scales, and self-monitoring. Unpublished manuscript, Department of Communication Arts, University of Wisconsin-Madison.
- Dixon, D. M., Sweeney, K. G., & Gray, D. J. P. P. (1999). The physician healer: Ancient magic or modern science? *British Journal of General Practice*, *49*,309-312.
- Duan, C., & Hill, C. E. (1996). The current state of empathy research. *Journal of Counseling Psychology*, *43*(3), 261. <http://dx.doi.org/10.1037/0022-0167.43.3.261>
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, *94*(1), 100. <http://dx.doi.org/10.1037/0033-2909.94.1.100>
- Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, *101*(1), 91.
- Esco, M. R., & Flatt, A. A. (2014). Ultra-short-term heart rate variability indexes at rest and post-exercise in athletes: Evaluating the agreement with accepted recommendations. *Journal of Sports Science & Medicine*, *13*(3), 535.
- Friedberg, J. P., Suchday, S., & Shelov, D. V. (2007). The impact of forgiveness on cardiovascular reactivity and recovery. *International Journal of Psychophysiology*, *65*(2), 87-94. <https://doi.org/10.1016/j.ijpsycho.2007.03.006>
- Geringer, E. A. (2015). Empathy: Emotional or cognitive? Patterns in physiological response, empathy scales and personality traits (Doctoral dissertation, California Lutheran University).
- Gladstein, G. A. (1983). Understanding empathy: Integrating counseling, developmental, and social psychology perspectives. *Journal of Counseling Psychology*, *30*, 467–482. <http://dx.doi.org/10.1037/0022-0167.30.4.467>

- Greenwood, D. C., Muir, K. R., Packham, C. J., & Madeley, R. J. (1996). Coronary heart disease: a review of the role of psychosocial stress and social support. *Journal of Public Health, 18*(2), 221-231. <https://doi.org/10.1093/oxfordjournals.pubmed.a024483>
- Grevin, F. (1996). Posttraumatic stress disorder, ego defense mechanisms, and empathy among urban paramedics. *Psychological Reports, 79*(2), 483-495. <https://doi.org/10.2466/pr0.1996.79.2.483>
- Hemingway, H., & Marmot, M. (1999). Clinical evidence: Psychosocial factors in the etiology and prognosis of coronary heart disease: systematic review of prospective cohort studies. *Western Journal of Medicine, 171*(5-6), 342.
- Hendrick, S. S., Dicke, A., & Hendrick, C. (1998). The relationship assessment scale. *Journal of Social and Personal Relationships, 15*(1), 137-142. doi:10.1177/0265407598151009.
- Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33*, 307-316. <http://dx.doi.org/10.1037/h0027580>
- Hojat, M., Louis, D. Z., Markham, F. W., Wender, R., Rabinowitz, C., & Gonnella, J. S. (2011). Physicians' empathy and clinical outcomes for diabetic patients. *Academic Medicine, 86*(3), 359-364. doi: 10.1097/ACM.0b013e3182086fe1
- Hoffman, M. L. (1977). Sex differences in empathy and related behaviors. *Psychological Bulletin, 84*(4), 712. <http://dx.doi.org/10.1037/0033-2909.84.4.712>
- Hops, H., Wills, T., Weis, R., & Patterson, G. Marital interaction coding system. Unpublished manuscript, 1971. Copies may be obtained by writing Dr. R. Weiss, Psychology Clinic, Department of Psychology, University of Oregon, Eugene, Ore. 97403.
- Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence, 29*(4), 589-611.

- Kim, S. S., Kaplowitz, S., & Johnston, M. V. (2004). The effects of physician empathy on patient satisfaction and compliance. *Evaluation & the Health Professions, 27*(3), 237-251. <https://doi.org/10.1177/0163278704267037>
- Krantz, D. S., & Manuck, S. B. (1984). Acute psychophysiologic reactivity and risk of cardiovascular disease: a review and methodologic critique. *Psychological Bulletin, 96*(3), 435. <http://dx.doi.org/10.1037/0033-2909.96.3.435>
- Kubzansky, L. D., Sparrow, D., Vokonas, P., & Kawachi, I. (2001). Is the glass half empty or half full? A prospective study of optimism and coronary heart disease in the normative aging study. *Psychosomatic Medicine, 63*(6), 910-916.
- Kurdek, L. A. (1994). Areas of conflict for gay, lesbian, and heterosexual couples: What couples argue about influences relationship satisfaction. *Journal of Marriage and the Family, 56*(4), 923–934. DOI: 10.2307/353603
- Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press on Demand.
- Lelorain, S., Brédart, A., Dolbeault, S., & Sultan, S. (2012). A systematic review of the associations between empathy measures and patient outcomes in cancer care. *Psycho-Oncology, 21*(12), 1255-1264. <https://doi.org/10.1002/pon.2115>
- Leone L, Pierro A, Mannetti L, Validità della versione italiana delle scale BIS/BAS di Carver e White. (2002). Generalizzabilità della struttura e relazioni con costrutti affini. *Giorn Ita Psicolog 29*, 413–34. doi: 10.1421/1245
- Lief, H. I., & Fox, R. C. (1963). Training for "detached concern" in medical students. In Lief, H. I., Lief, V. F., and Lief, N. (Eds.), *The psychological basis of medical practice* (pp. 12-35). New York: Harper and Row.

- Liew, J., Eisenberg, N., Spinrad, T. L., Eggum, N. D., Haugen, R. G., Kupfer, A., ... & Baham, M. E. (2011). Physiological regulation and fearfulness as predictors of young children's empathy-related reactions. *Social Development, 20*(1), 111-134.
<https://doi.org/10.1111/j.1467-9507.2010.00575.x>
- Lin, C., Li, L., Wan, D., Wu, Z., & Yan, Z. (2012). Empathy and avoidance in treating patients living with HIV/AIDS (PLWHA) among service providers in China. *AIDS Care, 24*(11), 1341-1348. <https://doi.org/10.1080/09540121.2011.648602>
- Marler, M. R., Jacob, R. G., Lehoczky, J. P., & Shapiro, A. P. (1988). The statistical analysis of treatment effects in 24-hour ambulatory blood pressure recordings. *Statistics in Medicine, 7*(6), 697-716. <https://doi.org/10.1002/sim.4780070608>
- McNames, J., & Aboy, M. (2006). Reliability and accuracy of heart rate variability metrics versus ECG segment duration. *Medical and Biological Engineering and Computing, 44*(9), 747-756. doi: 10.1007/s11517-006-0097-2
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*, 525-543. <https://doi.org/10.1111/j.1467-6494.1972.tb00078.x>
- Mercer, S. W., & Reynolds, W. J. (2002). Empathy and quality of care. *British Journal of General Practice, 52*(Suppl), S9-12.
- Miller, J. G., Nuselovici, J. N., & Hastings, P. D. (2016). Nonrandom acts of kindness: Parasympathetic and subjective empathic responses to sadness predict children's prosociality. *Child Development, 87*(6), 1679-1690. <https://doi.org/10.1111/cdev.12629>
- Miller, K. I., Stiff, J. B., & Ellis, B. H. (1988). Communication and empathy as precursors to burnout among human service workers. *Communications Monographs, 55*(3), 250-265.
<https://doi.org/10.1080/03637758809376171>

- Miu, A. C., & Balteş, F. R. (2012). Empathy manipulation impacts music-induced emotions: A psychophysiological study on opera. *PloS One*, 7(1), e30618.
<https://doi.org/10.1371/journal.pone.0030618>
- Neumann, M., Edelhäuser, F., Tauschel, D., Fischer, M. R., Wirtz, M., Woopen, C., ... & Scheffer, C. (2011). Empathy decline and its reasons: A systematic review of studies with medical students and residents. *Academic Medicine*, 86(8), 996-1009. doi: 10.1097/ACM.0b013e318221e615
- Niskanen, J. P., Tarvainen, M. P., Ranta-Aho, P. O., & Karjalainen, P. A. (2004). Software for advanced HRV analysis. *Computer Methods and Programs in Biomedicine*, 76(1), 73-81.
<https://doi.org/10.1016/j.cmpb.2004.03.004>
- Nussinovitch, U., Elishkevitz, K. P., Katz, K., Nussinovitch, M., Segev, S., Volovitz, B., & Nussinovitch, N. (2011). Reliability of ultra-short ECG indices for heart rate variability. *Annals of Noninvasive Electrocardiology*, 16(2), 117-122. <https://doi.org/10.1111/j.1542-474X.2011.00417.x>
- Oliveira-Silva, P., & Gonçalves, Ó. F. (2011). Responding empathically: A question of heart, not a question of skin. *Applied Psychophysiology and Biofeedback*, 36(3), 201. doi: 10.1007/s10484-011-9161-2
- Ono, M., Fujita, M., & Yamada, S. (2012). Physiological and psychological responses induced by expressing empathy with others. *Japan Journal of Nursing Science*, 9(1), 56-62.
<https://doi.org/10.1111/j.1742-7924.2011.00186.x>

- Perrone-McGovern, K. M., Oliveira-Silva, P., Simon-Dack, S., Lefdahl-Davis, E., Adams, D., McConnell, J., ... & Gonçalves, Ó. F. (2014). Effects of empathy and conflict resolution strategies on psychophysiological arousal and satisfaction in romantic relationships. *Applied Psychophysiology and Biofeedback, 39*(1), 19-25. doi: 10.1007/s10484-013-9237-2
- Piette, J. D., Heisler, M., Krein, S., & Kerr, E. A. (2005). The role of patient-physician trust in moderating medication nonadherence due to cost pressures. *Archives of Internal Medicine, 165*(15), 1749-1755. doi:10.1001/archinte.165.15.1749
- Plutchik, R. (1962). *The emotions: Facts, theories and a new model*. New York, NY, US: Crown Publishing Group/Random House.
- Preston, S. D., & De Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(1), 1-20. <https://doi.org/10.1017/S0140525X02000018>
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21*(2), 95. <http://dx.doi.org/10.1037/h0045357>
- Rogers, C. R. (1986). A client-centered/person-centered approach to therapy. In H. Kirschenbaum & V. L. Henderson (Eds.), *The Carl Rogers reader* (pp. 135-152). London: Constable.
- Rosenman, R. H., Brand, R. J., Sholtz, R. I., & Friedman, M. (1976). Multivariate prediction of coronary heart disease during 8.5 year follow-up in the Western Collaborative Group Study. *American Journal of Cardiology, 37*(6), 903-910. [https://doi.org/10.1016/0002-9149\(76\)90117-X](https://doi.org/10.1016/0002-9149(76)90117-X)

- Shen, L. (2010). Mitigating psychological reactance: The role of message-induced empathy in persuasion. *Human Communication Research, 36*(3), 397-422.
<https://doi.org/10.1111/j.1468-2958.2010.01381.x>
- Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology, 24*(18), R875-R878. <https://doi.org/10.1016/j.cub.2014.06.054>
- Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personality Assessment, 91*, 62–71.
<https://doi.org/10.1080/00223890802484381>
- Stellar, J. E., Cohen, A., Oveis, C., & Keltner, D. (2015). Affective and physiological responses to the suffering of others: Compassion and vagal activity. *Journal of Personality and Social Psychology, 108*(4), 572. <http://dx.doi.org/10.1037/pspi0000010>
- Stiff, J. B., Dillard, J. P., Somera, L., Kim, H., & Sleight, C. (1988). Empathy, communication, and prosocial behavior. *Communications Monographs, 55*(2), 198-213.
<https://doi.org/10.1080/03637758809376166>
- Stotland, E. (1969). Exploratory studies of empathy. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 271-314). New York: Academic Press.
- Strayer, J. (1987). Affective and cognitive perspectives on empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 218-244). New York: Cambridge University Press.
- Suhr, J. A., Cutrona, C. E., Krebs, K. K., & Jensen, S. L. (2004). The social support behavior code. In P.K. Kerig & D.H. Baucom (Eds.), *Couple observational coding systems* (pp. 311-318). New York: Routledge.

- Teding van Berkhout, E., & Malouff, J. M. (2016). The efficacy of empathy training: A meta analysis of randomized controlled trials. *Journal of Counseling Psychology, 63*(1), 32. <http://dx.doi.org/10.1037/cou0000093>
- Thom, D. H., Kravitz, R. L., Bell, R. A., Krupat, E., & Azari, R. (2002). Patient trust in the physician: relationship to patient requests. *Family Practice, 19*(5), 476-483. <https://doi.org/10.1093/fampra/19.5.476>
- Thong, T., Li, K., McNames, J., Aboy, M., & Goldstein, B. (2003, September). Accuracy of ultra-short heart rate variability measures. In *Engineering in Medicine and Biology Society, Proceedings of the 25th Annual International Conference of the IEEE* (Vol. 3, pp. 2424-2427). IEEE. doi: 10.1109/IEMBS.2003.1280405
- Tracy, L. M., & Giummarra, M. J. (2017). Sex differences in empathy for pain: What is the role of autonomic regulation? *Psychophysiology, 54*(10), 1549-1558. doi:10.1111/psyp.12895
- Truzzi, A., Setoh, P., Shinohara, K., & Esposito, G. (2016). Physiological responses to dyadic interactions are influenced by neurotypical adults' levels of autistic and empathy traits. *Physiology & Behavior, 165*, 7-14. <https://doi.org/10.1016/j.physbeh.2016.06.034>
- Treiber, F. A., Kamarck, T., Schneiderman, N., Sheffield, D., Kapuku, G., & Taylor, T. (2003). Cardiovascular reactivity and development of preclinical and clinical disease states. *Psychosomatic Medicine, 65*(1), 46-62.
- Vanderpool, J. P., & Barratt, E. S. (1970). Empathy: Towards a psychophysiological definition. *Diseases of the Nervous System, 31*(7), 464-467.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063.

- Wiesenfeld, A. R., Whitman, P. B., & Malatesta, C. Z. (1984). Individual differences among adult women in sensitivity to infants: evidence in support of an empathy concept. *Journal of Personality and Social Psychology*, *46*(1), 118. <http://dx.doi.org/10.1037/0022-3514.46.1.118>
- Woodall, W., & Kogler-Hill, S. (1982). Predictive and perceived empathy as predictors of leadership style. *Perceptual and Motor Skills*, *54*, 800-802.
<https://doi.org/10.2466/pms.1982.54.3.800>
- Yasuma, F., & Hayano, J. I. (2004). Respiratory sinus arrhythmia: Why does the heartbeat synchronize with respiratory rhythm? *Chest*, *125*(2), 683-690.
<https://doi.org/10.1378/chest.125.2.683>
- Zahn-Waxler, C., Cole, P. M., Welsh, J. D., & Fox, N. A. (1995). Psychophysiological correlates of empathy and prosocial behaviors in preschool children with behavior problems. *Development and Psychopathology*, *7*(1), 27-48. doi:10.1017/S0954579400006325
- Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion* *8*: 494–521. doi: 10.1037/1528-3542.8.4.494

Table 1. Summary ANOVA Table for Cardiovascular Task Baseline Comparisons between Higher and Lower Empathy Groups: Task Type (Pre-Neutral, Pre-Comfort) x Empathy Group (Lower, Higher) Repeated Measures ANOVAs

Source	SS	df	MS	F	p	η^2
SBP						
Empathy Group	45.56	1	45.56	.210	.650	.007
Error (Empathy Group)	6518.38	30	217.28			
Task Type	52.56	1	52.56	1.94	.174	.061
Task * Group	76.56	1	76.56	2.82	.104	.086
Error (Task Type)	814.88	30	27.16			
DBP						
Empathy Group	66.69	1	66.69	.426	.519	.014
Error (Empathy Group)	4692.50	30	156.42			
Task Type	40.11	1	40.11	1.50	.231	.048
Task * Group	16.00	1	16.00	.597	.446	.020
Error (Task Type)	803.78	30	26.79			
HR						
Empathy Group	3.65	1	3.65	.015	.903	.001
Error (Empathy Group)	7245.24	30	241.51			
Task Type	1.78	1	1.78	.390	.537	.013
Task * Group	2.70	1	2.70	.591	.448	.019
Error (Task Type)	137.21	30	137.21			
LF-HRV						
Empathy Group	2.04	1	2.04	1.38	.249	.044
Error (Empathy Group)	44.23	30	1.47			
Task Type	.019	1	.019	.174	.680	.006
Task * Group	.093	1	.093	.830	.370	.027
Error (Task Type)	3.35	30	.112			
HF-HRV						
Empathy Group	.435	1	.435	.170	.683	.006
Error (Empathy Group)	76.85	30	2.56			
Task Type	.110	1	.110	1.02	.320	.033
Task * Group	.041	1	.041	.378	.543	.012
Error (Task Type)	3.24	30	.108			

RMSSD

Empathy Group	.016	1	.016	.110	.743	.004
Error (Empathy Group)	4.41	30	.147			
Task Type	.002	1	.002	.472	.497	.015
Task * Group	.001	1	.001	.367	.549	.012
Error (Task Type)	.098	30	.003			

Table 2. Summary ANOVA Table for PANAS Rest Period Comparisons between Higher and Lower Empathy Groups: Task Type (Neutral, Comfort) x Empathy Group (Higher, Lower) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
PANAS Positive Total						
Empathy Group	6.25	1	6.25	.072	.790	.002
Error (Empathy Group)	2598.50	30	86.62			
Task Type	36.00	1	36.00	2.12	.156	.066
Task * Group	.250	1	.250	.015	.904	.000
Error (Task Type)	508.75	30	16.96			
PANAS Negative Total						
Empathy Group	34.52	1	34.52	1.31	.262	.042
Error (Empathy Group)	792.09	30	26.40			
Task Type	28.89	1	28.89	3.73	.063	.110
Task * Group	.016	1	.016	.002	.964	.000
Error (Task Type)	232.59	30	7.75			

Table 3. Summary ANOVA Table for Post-task Agentic and Communal Score Comparisons between Higher and Lower Empathy Groups: Task Type (Neutral, Comfort) x Empathy Group (Higher, Lower) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
PTQ Agentic						
Empathy Group	4.00	1	4.00	.569	.456	.019
Error (Empathy Group)	210.75	30	7.03			
Task Type	100.00	1	100.00	18.99	.000	.388
Task * Group	9.00	1	9.00	1.71	.201	.054
Error (Task Type)	158.00	30	5.27			
PTQ Communal						
Empathy Group	.391	1	.391	.044	.835	.001
Error (Empathy Group)	264.22	30	8.81			
Task Type	83.27	1	83.27	26.76	.000	.471
Task * Group	1.89	1	1.89	.608	.442	.020
Error (Task Type)	93.34	30	3.11			

Note. p-values for significant effects are bolded

Table 4. Summary ANOVA Table for Post-task Individual Item Comparisons between Higher and Lower Empathy Groups: Task Type (Neutral, Comfort) x Empathy Group (Higher, Lower) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
How realistic?						
Empathy Group	2.25	1	2.25	1.52	.228	.048
Error (Empathy Group)	44.50	30	1.48			
Task Type	3.06	1	3.06	3.93	.057	.116
Task * Group	.563	1	.563	.722	.402	.023
Error (Task Type)	23.38	30	.779			
How likely to encounter?						
Empathy Group	5.64	1	5.64	5.67	.024	.159
Error (Empathy Group)	29.84	30	.995			
Task Type	5.64	1	5.64	5.69	.024	.160
Task * Group	.141	1	.141	.142	.709	.005
Error (Task Type)	29.72	30	.991			
How stressful to handle?						
Empathy Group	.563	1	.563	.371	.547	.012
Error (Empathy Group)	45.44	30	1.52			
Task Type	10.56	1	10.56	18.44	.000	.381
Task * Group	.250	1	.250	.436	.514	.014
Error (Task Type)	17.19	30	.573			
How stressful to engage?						
Empathy Group	1.56	1	1.56	.717	.404	.023
Error (Empathy Group)	65.38	30	2.18			
Task Type	5.06	1	5.06	7.45	.010	.199
Task * Group	1.56	1	1.56	2.30	.140	.071
Error (Task Type)	20.38	30	.679			
How meaningful?						
Empathy Group	.391	1	.391	.293	.592	.010
Error (Empathy Group)	39.97	30	1.33			
Task Type	.016	1	.016	.029	.867	.001
Task * Group	.141	1	.141	.258	.615	.009
Error (Task Type)	16.34	30	.545			

How positive?

Empathy Group	.250	1	.250	.170	.683	.006
Error (Empathy Group)	44.19	30	1.47			
Task Type	2.25	1	2.25	3.82	.060	.113
Task * Group	.063	1	.063	.106	.747	.004
Error (Task Type)	17.69	30	.590			

How negative?

Empathy Group	1.56	1	1.56	1.33	.259	.042
Error (Empathy Group)	35.38	30	1.18			
Task Type	1.56	1	1.56	3.79	.061	.112
Task * Group	.063	1	.063	.152	.700	.005
Error (Task Type)	12.38	30	.413			

How upset?

Empathy Group	2.64	1	2.64	1.77	.193	.056
Error (Empathy Group)	44.72	30	1.49			
Task Type	3.52	1	3.52	3.38	.076	.101
Task * Group	.766	1	.766	.736	.398	.024
Error (Task Type)	31.22	30	1.04			

How difficult?

Empathy Group	5.06	1	5.06	2.28	.142	.071
Error (Empathy Group)	66.69	30	2.22			
Task Type	4.00	1	4.00	4.22	.049	.123
Task * Group	1.56	1	1.56	1.65	.209	.052
Error (Task Type)	28.44	30	.948			

Note. p-values for significant effects are bolded

Table 5. Summary ANOVA Table for Behavioral Parameters: Task Type (Neutral, Comfort) x Empathy Group (Higher, Lower) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
Information Support						
Empathy	18.06	1	18.06	1.25	.272	.040
Error (Empathy)	432.94	30	14.43			
Task Type	30.25	1	30.25	4.33	.046	.126
Task * Empathy	3.06	1	3.06	.438	.513	.014
Error (Task Type)	209.69	30	6.99			
Emotional Support						
Empathy	.563	1	.563	.055	.816	.002
Error (Empathy)	307.44	30	10.25			
Task Type	49.00	1	49.00	11.10	.002	.270
Task * Empathy	.563	1	.563	.127	.724	.004
Error (Task Type)	132.44	30	4.42			
Esteem Support						
Empathy	.057	1	.057	1.18	.287	.038
Error (Empathy)	1.44	30	.048			
Task Type	2.42	1	2.42	69.47	.000	.698
Task * Empathy	.014	1	.014	.412	.526	.014
Error (Task Type)	1.05	30	.035			
Social Network Support						
Empathy	.207	1	.207	5.09	.032	.145
Error (Empathy)	1.22	30	.041			
Task Type	.020	1	.020	.494	.488	.016
Task * Empathy	.017	1	.017	.417	.523	.014
Error (Task Type)	1.20	30	.040			
Attentiveness						
Empathy	31.64	1	31.64	1.39	.247	.044
Error (Empathy)	680.97	30	22.70			
Task Type	3.516	1	3.516	.577	.453	.019
Task * Empathy	1.27	1	1.27	.208	.652	.007
Error (Task Type)	182.72	30	6.09			
Positive Non-Verbal						
Empathy	12.25	1	12.25	.455	.505	.015
Error (Empathy)	807.69	30	26.92			
Task Type	495.06	1	495.06	49.56	.000	.623
Task * Empathy	2.25	1	2.25	.225	.639	.007
Error (Task Type)	299.69	30	9.99			

Table 6. Summary ANOVA Table for Individual Behavior Codes: Task Type (Neutral, Comfort) x Empathy Group (Higher, Lower) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
Not Tracking						
Empathy	.085	1	.085	.313	.580	.010
Error (Empathy)	8.19	30	.273			
Task Type	.006	1	.006	.070	.793	.002
Task * Empathy	.032	1	.032	.369	.548	.012
Error (Task Type)	2.61	30	.087			
Normative Verbal						
Empathy	.141	1	.141	.030	.865	.001
Error (Empathy)	142.47	30	4.75			
Task Type	11.39	1	11.39	4.90	.035	.140
Task * Empathy	.391	1	.391	.168	.685	.006
Error (Task Type)	69.72	30	2.32			
Normative Non-Verbal						
Empathy	39.06	1	39.06	2.61	.117	.080
Error (Empathy)	448.88	30	14.96			
Task Type	6.25	1	6.25	2.35	.136	.073
Task * Empathy	4.00	1	4.00	1.51	.229	.048
Error (Task Type)	79.75	30	2.66			

Note. p-values for significant effects are bolded

Figure 1. Systolic Blood Pressure (SBP) Reactivity by Empathy Level and Task Type

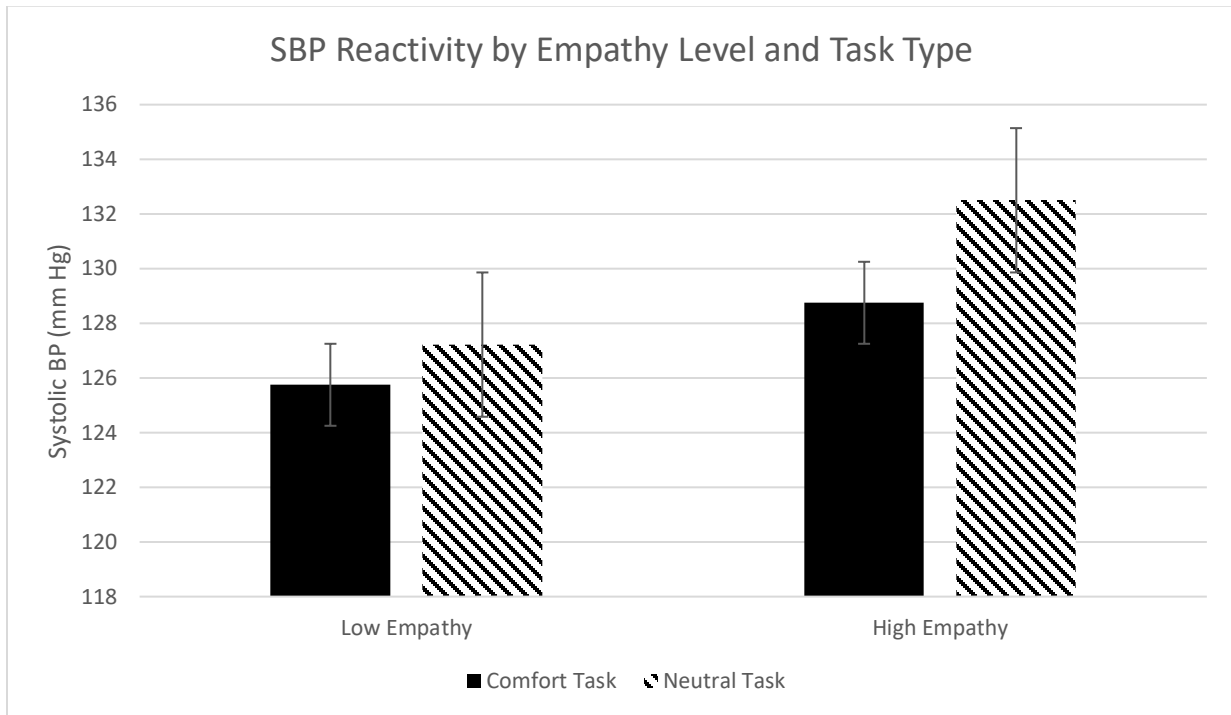
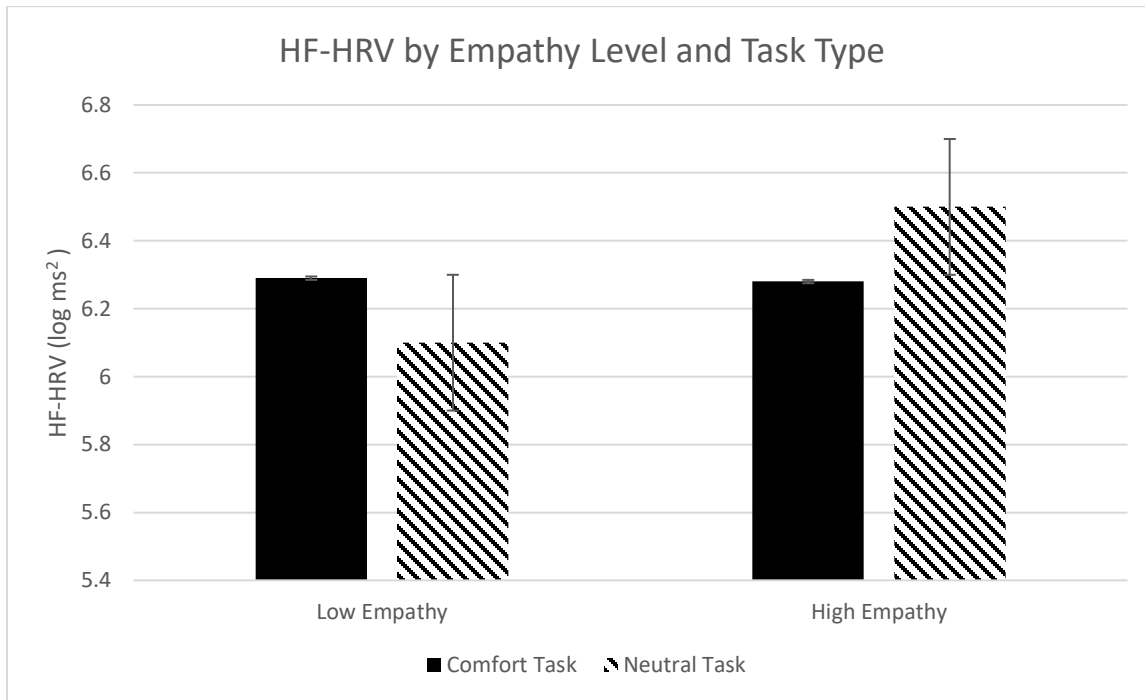


Figure 2. High-Frequency Heart Rate Variability (HF-HRV) Reactivity by Empathy Level and Task Type



APPENDIX A: Literature Review of Empathy and Autonomic Reactivity

Studies examining the relation between empathy and autonomic response to emotions have used a variety of stimuli including imagination, photos, videos, and live dyadic interactions. The literature includes studies that have employed a range of measures of autonomic nervous system response to emotionally provocative stress, including traditional measures of heart rate (HR) and skin conductance response (SCR) as well as measures of HR variability (HRV), primary indicators of parasympathetic activity.

Standard Mental Stress Stimuli. Tracy and Giummarra (2017) examined the role of autonomic regulation on sex differences in empathy for pain. Male and female participants completed a scale measuring empathic reactions to imagining someone in pain to assess for sex differences in empathic concern for pain. Measures of HRV were obtained using the square root of the mean squared differences of successive R-R intervals (RMSSD) as a measure of vagally mediated HRV, while participants engaged in two laboratory challenges, a paced breathing task and serial subtraction. Results indicated that women self-reported significantly higher empathic concern on the EPS than men. Although there were no differences in resting parasympathetic activity (PA) measured via RMSSD, both sexes exhibited a decrease in PA during the stress task. The HRV response was moderated by participant sex. Among women, lower levels of resting RMSSD were associated with higher empathic concern, suggesting that empathic concern was associated with lower parasympathetic tone. There was no association between resting HRV and empathic concern among men.

Photos of Emotional Expression. Ardizzi et al. (2016) utilized photo stimuli to examine the impact of childhood maltreatment on facial mimicry, considered to be a measure of empathic understanding of another's emotions, and vagal regulation in a sample of maltreated children and

age-matched controls. Spontaneous facial electromyography activation corrugator and zygomaticus muscle and respiratory sinus arrhythmia (RSA) responses were recorded during exposure to visual facial expressions of anger, fear, joy, and sadness. Ardizzi et al. found that maltreated children showed lower facial mimicry than controls to negative facial expressions and increased RSA suppression (less PNS activation) in response to angry facial expressions, but controls did not. The authors concluded that the pattern of responding exhibited by the maltreated children was the result of their adaptation to a hostile environment, in which vagal regulation becomes linked with threatening social cues, leading to increased sympathetic responding. The authors proposed that the greater RSA suppression was indicative of the ability of maltreated children to engage selectively with their environment, perhaps by adapting to the threatening and adverse experiences that they experienced earlier in life than the control group.

Auditory Emotionally-Provocative Stimuli. Vanderpool and Barratt (1970) utilized auditory stimuli to define a pattern of psychophysiological response that could predict empathy. Participants included psychiatry and psychology residents who listened to patient interview tapes depicting a hysterical individual and an individual with a psychotic break. Residents were asked to match the emotions of the individual in each of the tapes to responses on the Plutchik Emotion Profile Index (Plutchik, 1962) and Allport-Vernon-Lindzey Scale of Values (Allport, Vernon, & Lindzey, 1951), from the perspectives of the patients from the tapes. The extent to which they did this accurately was considered a measure of empathy for this study. Physiological measures of HR, SCR, and respiration were recorded, and results indicated the HR measures recorded during the word list were significantly negatively correlated with empathy scores in men only. The authors also observed that residents higher in empathy showed higher autonomic responses measured by SCR and respiration at the beginning of the tape, but habituated faster than those

with lower in empathy over the course of the tape. They noted that those with increased empathy scores actually showed slower acceleration in HR over the course of the tape than those with lower empathy scores. The authors suggested that higher empathizers had a “thicker skin” that allowed them to orient themselves to external stimuli and thus respond differently than lower empathizers.

Video Stimuli. Wiesenfeld, Whitman, and Malatesta (1984) utilized video stimuli to examine differences among adult women in sensitivity and empathy to infants. Higher and lower empathy grouped females watched videos of infants smiling, crying, and during a quiescent state, while their physiological and facial reactions were recorded, followed by a description of the perceived emotion of the infant and verbalization of personal feelings during the film. Females in the higher empathy group showed a greater cardiovascular response to the emotional stimuli, measured by HR change from baseline to stimuli, but this finding only approached statistical significance. Females in the higher empathy group also exhibited larger electrodermal SCR as well as similar facial expressions in response to the videos. The authors concluded that higher empathic females are more emotionally responsive to infant emotions than lower empathic females.

Anastassiou-Hadjicharalambous and Warden (2007) conducted a study in children aged 8-10 years assessing convergence between physiological, facial, and verbal self-report measures of affective empathy. Children watched a stimulus film that depicted several clips of differing emotional valence (i.e., sadness, fear, happiness). Psychophysiological measures were recorded via ECG and facial expressions were recorded. Following the video, children verbally described how they felt while watching the video. Findings showed no significant convergence between physiological and verbal measures; however, significant but low positive convergence was found

between physiological and facial measures and facial and verbal measures. The authors concluded that attention must be given to the specific measurement method used before generalizing results across measurement methods.

Liew et al. (2011) examined physiological regulation and fearfulness as predictors of empathy-related reactions to an unfamiliar person's distress in young children at 18 months and again at 30 months of age. The children's HR and respiration were recorded while viewing films depicting either neutral babies or distressed, crying babies. They were then exposed to an unfamiliar female who acted distressed, and their empathy-related reactions were recorded. The study found that resting RSA was negatively related to empathic concern, so those with more PNS activity exhibited less empathy. Resting RSA was associated with less distress at 30 months, but not 18 months, and RSA suppression (less PNS activation) at 18 months predicted helping, prosocial behavior at the 30-month time point.

Oliveira-Silva and Gonçalves (2011) examined the psychophysiological correlates of empathy via measures of skin conductance level (SCL) and inter-beat interval (IBI) as a function of the empathic response. Undergraduate students viewed emotionally charged positive or negative vignettes of interactions between two actors. Participants were then asked to think about the response they would give to the situation viewed and asked to categorize that response among a list of provided response options. Responses represented three empathic levels: subtractive (a response unrelated to the expressed emotion), interchangeable (a response mirroring the expressed emotion), and additive (a response helping with the clarification of the expressed emotion). No significant differences occurred in SCL across the three levels of empathic response, though significant differences were observed in IBI between all three empathic response levels. Additive empathy responses were significantly related to increased

cardiac activity (decreased interbeat intervals) as compared to the interchangeable empathic response. During the response phase only, additive empathic responses were significantly related to decreased IBIs more than subtractive empathic responses. As such, the authors concluded that these findings suggested that cardiovascular reactivity changes with differing types or degrees of empathic responses.

de Wied, van Boxtel, Matthys, and Meeus (2012) conducted a study to examine autonomic responses to empathy-eliciting videos in adolescent males (and age-matched controls) with disruptive behavior disorder and higher versus lower callous-unemotional (CU) traits. Parents and teachers completed a questionnaire measuring psychopathic and CU traits in adolescents, which were then used to classify participants as higher versus lower in CU. Adolescents first viewed a relaxation video for purposes of measuring resting RSA and HR. Next, emotional film clips were presented to the adolescents and included sadness, anger, and happiness clips. At the end of the clips, the main character expressed intense vocal and facial emotion, and this was the “target episode” during which HR was analyzed. Participants rated feelings of empathy towards the clip subject by reporting on the intensity and quality of their emotions. Responses were deemed empathic if the adolescent reported that they experienced the same emotion as observed in the main character of the clip. Results revealed a significant difference between the higher CU and lower CU groups in HR responses to sadness, but not to happiness or anger. Higher CU disruptive adolescents exhibited less empathic sadness than controls. Higher CU adolescents not only reported feeling less empathy, but showed less of a change in HR between baseline to the target period of sadness than controls and those with lower CU.

Using a musical context, Miu and Baltes (2012) examined psychophysiological measures of autonomic reactivity and empathy in response to videos of emotional opera compositions. Participants observed two opera performances, one describing the pain of a mother about to lose her sons (negative), and one describing a happy march of a boy-drummer (positive). The Positive and Negative Affect Schedule (PANAS) was completed to assess for participant mood pre-experiment. The Toronto Empathy Questionnaire (Spreng, McKinnon, Mar, & Levine, 2009) was used to measure trait empathy. The Geneva Emotional Music Scales (Zentner, Grandjean, & Scherer, 2008) was used after video viewing to measure music-induced emotions related to wonder, transcendence, sadness, peacefulness, etc. Physiological measures of SCL, respiration rate (RR), HR, and HRV were measured during the video session. Individuals viewed the videos and were instructed using two different empathy (higher or lower) scripts. In the higher empathy condition, participants were asked to imagine how the performer felt about what was described in the music and to attempt to feel those same emotions. In the lower empathy condition, participants were asked to view the video objectively and not to empathize with the feelings of the performer. SCL was lower among participants in the empathy condition than those in the lower empathy condition during the negative emotion opera, but RR was higher among those in the empathy condition during the positive emotion opera. Analyses showed no significant differences in mood or trait empathy between individuals in the higher or lower empathy conditions. Trait empathy was not found to be significantly associated with physiological activity during each aria.

Balconi and Bortolotti (2012) examined resonance mechanisms underlying empathic behavior. Undergraduate students viewed a video of an interaction between two actors and were instructed to empathize. Four scene types were viewed and included cooperative, non-

cooperative (implicitly opposite), conflictual (explicitly opposite), and neutral (indifference). Participants completed the Balanced Emotional Empathy Scale (Mehrabian & Epstein, 1972) and were categorized into higher and lower trait empathy groups. After viewing the interaction, behavioral inhibition system (BIS) and behavioral activation system (BAS) scores were calculated for each participant, measures of approach-withdrawal attitudes. Theoretically, individuals with higher BAS levels respond more to positively-valenced situations and tend to approach them, while individuals with higher BIS levels respond more to negatively-valenced or threatening situations and tend to avoid them. A higher and lower group of participants was created for each measure. Measures of autonomic activity included SCR and HR. Results demonstrated increased SCR and HR during the non-cooperative and conflictual scenes in comparison to the cooperative and neutral scenes. Higher and lower empathy participants showed corresponding levels of responding, both on self-report measures of empathy and on the autonomic measures of SCR and HR. BIS and BAS measures were shown to be related to empathy scores and autonomic responses, with those higher in BAS showing increased autonomic response and empathy in response to the cooperative (positive) situation, and those higher in BIS showing more empathy in response to the conflictual (negative) situation.

Geringer (2015) examined physiological reactions to videos displaying happiness, sadness, or a neutral emotion. Participants categorized as higher or lower empathy viewed three films while HR, respiratory rate, HRV, and electrodermal activity were recorded. The only physiological response that significantly differed between higher and lower empathy groups was HR, with those higher in emotional empathy experiencing a significant increase in HR while watching video clips in contrast to those lower in emotional empathy.

Stellar, Cohen, Oveis, and Keltner (2015) conducted four studies in which undergraduate participants viewed videos of the suffering of others while physiological indices of HR, respiration, SCR, and RSA were measured. To induce compassion, participants viewed a video of a grieving student discussing the death of her grandfather. Although compassion represents a somewhat different construct than empathy, to express compassion for another person requires empathy (Singer & Klimecki, 2014). The first study found that RSA was greater during the induction of compassion (increased PNS activation) as compared with a control or “non-emotional” scene about nature. The second study also found RSA to be greater during compassion induction as compared to another positive emotion, and the third study found RSA to be greater during compassion induction than another “prosocial emotion lacking appraisals of another’s suffering” (p. 572). Compared to a neutral or control emotion, greater RSA during compassion induction also was associated with lower HR and respiration, but no difference in SCR. In the fourth study, greater RSA during compassion induction predicted prosocial behaviors and compassionate non-verbal and verbal behaviors. In all, these four studies showed that compassion was linked with increased parasympathetic activity.

Miller, Nuselovici, and Hastings, (2016) examined children aged 4 to 6 years old and found that, when exposed to empathy-inducing videos, RSA response patterns showed suppression, (less PNS activation) followed by RSA recovery and another smaller suppression following resolution of the empathic event. The children’s pattern of RSA activity during the empathic event was associated with self-reported feelings of empathic concern. Dynamic RSA changes in response to the videos additionally were found to predict middle childhood prosocial behaviors longitudinally.

Truzzi, Setoh, Shinohara, and Esposito (2016) examined how physiological responses to dyadic interactions were influenced by levels of both autistic and empathy traits. Neurotypical undergraduates completed a questionnaire to assess for autism trait levels, as well as an assessment of empathic abilities. Participants then watched clips of dyadic interactions between actors. HR (an index of parasympathetic and sympathetic nervous system activity) and peripheral surface body temperature (an index of sympathetic activation) were recorded during the clips. Results showed that there was a significant main effect of empathy scores on left cheek temperature, such that higher trait empathy was associated with increased skin temperature. The authors mentioned that facial temperature increases and decreases can both be associated with sympathetic activation. Results also showed that participants with lower levels of autistic traits and higher empathic traits showed an increase in HR, while participants with higher levels of autistic traits and lower empathic traits showed decreased HR. The authors concluded that levels of autistic as well as empathic traits influence autonomic nervous system responses to observing interpersonal interactions, and that perhaps physiological responses could act as biomarkers of social ability and autistic traits.

In Vivo Interpersonal Interactions. Ono, Fujita, and Yamada (2012) examined physiological responses (via second finger skin temperature, an indicator of sympathetic nerve activation) of participants while expressing empathy. Female undergraduates were partnered into pairs, and one member of the pair completed a stress task (requiring them to complete the Stroop color-word test simultaneous with a listening test). After task completion, the other partner inquired about the task experience, and the task completer discussed her negative emotions during the task. In one condition, the partner was instructed to display empathy with the task completer by “listening with empathy” but making no verbalizations. In a second condition, the

partner made no empathic response. Results showed no significant difference in sympathetic nerve tone between the empathy and control conditions, as both conditions resulted in decreased fingertip skin temperature. Subjective participant stress was significantly higher after the completion of empathy-listening, but not during it, when compared with the control condition. The authors concluded that perhaps subjective stress was not immediately recognized while expressing empathy, even if physiological changes were occurring, and proposed that individuals may need to disregard their personal emotions while engaging in empathy in order to fully attend to the emotions of another.

Perrone-McGovern et al. (2014) examined the effects of empathy and conflict resolution strategies on psychophysiological arousal and satisfaction in romantic relationships. Participants included male and female undergraduates, and each participant was accompanied by his or her romantic partner. SCL and IBI were measured continuously throughout the interaction task. A baseline task where participants were shown neutral valence and low arousal photos was used to establish resting levels of cardiovascular activity. The second task involved an interaction where couples were asked to discuss the strengths and weaknesses of their relationship. Partners took turns speaking about and then listening to one another while discussing their relationship. After listening, the listener was asked to explain what he or she heard the partner say about the relationship. Results indicated no significant effects of empathy on IBI or SCL.

Zahn-Waxler, Cole, Welsh, and Fox (1995) examined psychophysiological (HR and SCL) correlates of empathy and prosocial behaviors in preschoolers with behavior problems. Children ranged from high, moderate, to low in risk for developing future disruptive behavior disorders. To assess for empathy, experimenters simulated scripted pain and emotional distress scenarios naturally into the ongoing activities of the session (i.e., dropping a box on their foot

and expressing pain). Children were scored on prosocial behavior (offering help or comfort), hypothesis testing (attempts to comprehend the others' distress), empathic concern (facial, vocal, or gestural-postural expressions), arousal (level of activation and body tension), and avoidance (aloofness, withdraw from the distress). Children were also asked to engage in hypothetical dilemmas (distress and conflict themes) in which the experimenter used a script to describe an event and asked the child to complete it, verbally and while using props to represent characters. Prosocial (hugging, kissing) and aggressive (harming, pushing) behaviors were coded. Children then participated in an astronaut game and wore astronaut suits with electrodes. Baseline measures of HR were recorded while children read books about outer space. Children then participated in a computer-based vigilance task, and a mood induction task. Children viewed the Mood Induction Stimulus for Children (Cole, Jordan, & Zahn-Waxler, 1990), which portrayed a child from space who expressed emotions to induce joy, anger, fear, or sadness in the children. Results showed that girls expressed more prosocial behavior than boys, and boys showed more anger during the hypothetical dilemmas than girls. Higher HR and higher HR deceleration predicted empathic concern and prosocial behavior. Lower HR was associated with avoidance and aggression. During sadness mood inductions, girls showed higher SCL than boys, and high-risk girls showed the highest levels of SCL.

Summary. The literature examining the relation between empathy and autonomic nervous system reactivity reveals a range of findings. The majority of these studies, however, found that increased empathy was associated with increased sympathetic or decreased parasympathetic activity in response to stressful or emotionally-provocative stimuli (Ardizzi et al., 2016; Balconi & Bortolotti, 2012; de Wied, et al., 2012; Geringer, 2015; Liew et al., 2011; Miller, et al., 2016; Oliveira-Silva & Gonclaves, 2011; Tracy & Giummarra, 2017; Truzzi et al.,

2016; Wiesenfeld et al., 1984; Zahn-Waxler et al., 1995). The series of studies conducted by Stellar and colleagues (2015), in contrast, found that engagement in compassion was associated with decreased sympathetic response to stress. It should be noted that this study compared those instructed to experience compassion with those instructed to remain distant rather than comparing effects of trait levels of empathy on physiological response measures. Vanderpool and Barratt (1970) reported mixed findings such that higher empathy was associated with increased SCR and respiration, but decreased HR. Likewise, Miu and Baltes (2012) found increased empathy to be related to both increased respiration and decreased SCR response when manipulated, but no differences in empathy and trait empathy. Finally, several studies found no significant association between autonomic reactivity measures and empathy (Anastassiou-Hadjicharalambous & Warden, 2007; Ono et al., 2012; Perrone-McGovern, 2014).

The pattern of findings did not appear to be influenced by stimulus type (standard mental stress stimuli, photos of emotional expression, auditory emotionally-provocative stimuli, video stimuli, and in vivo interpersonal interactions). The only study to use a standard mental stress stimuli found a positive association between empathy and autonomic response (Tracy & Giummarra, 2017), but the only study to use an auditory stimulus reported mixed findings (higher empathy was associated with increased SCR and respiration and decreased HR) (Vanderpool & Barratt, 1970). The only study to utilize photos of emotional expression found that higher empathy was associated with increased sympathetic response (Ardizzi et al., 2016). Of the eleven studies that used video stimuli, eight studies found that increases in empathy were associated with increases in sympathetic activity (Balconi & Bartolotti, 2012; de Wied, et al., 2012; Geringer, 2015; Liew et al., 2011; Miller et al., 2016; Oliveira-Silva & Gonclaves, 2011; Truzzi et al., 2016; Wiesenfeld et al., 1984), one study found that higher empathy was associated

with a decrease in sympathetic activity (Stellar et al., 2015), and two studies found no significant association between trait empathy and autonomic response (Anastassiou-Hadjicharalambous & Warden, 2007; Miu & Baltes, 2012). Of the three in vivo interpersonal interaction studies, one found a positive association between empathy and autonomic reactivity (Zahn-Waxler et al., 1995), and two found no significant associations between empathy and autonomic response (Ono et al., 2012; Perrone-McGovern, 2014).

The pattern of findings also did not appear to vary based on the physiological measure that was selected as a measure of sympathetic or parasympathetic responding (HR, HRV, SCL, and respiration). Of the ten studies that measured HR, six found that higher empathy was associated with increased HR (Balconi & Bortolotti, 2012; de Wied et al., 2012; Geringer, 2015; Truzzi et al., 2016; Wiesenfeld et al., 1984; Zahn-Waxler et al., 1995), two found a negative association between empathy and HR (Stellar et al., 2015; Vanderpool & Barratt, 1970), and two found no significant association between empathy and HR (Anastassiou-Hadjicharalambous & Warden, 2007; Miu & Baltes, 2012). Of the nine studies that measured HR variability (via RMSSD, RSA, or IBI), five studies found a significant positive association between empathy and HRV (Ardizzi et al., 2016; Liew et al., 2011; Miller et al., 2016; Oliveira-Silva & Gonclaves, 2011; Tracy & Giummarra, 2017), one found a negative association between empathy and HRV (Stellar et al., 2015), and three found no significant association between empathy and HRV (Geringer, 2015; Miu & Baltes; Perrone-McGovern et al., 2014). Of the eleven studies that examined SCL (or temperature), six studies found a significant positive association between empathy and SCL (Balconi & Bartolotti, 2012; Oliveira-Silvas & Gonclaves, 2011; Truzzi et al., 2016; Vanderpool & Barratt, 1970; Wiesenfeld et al., 1984; Zahn-Waxler, et al., 1995), one found a negative association between empathy and SCL (Miu & Baltes, 2012) and four studies

found no significant association between empathy level and SCL (Geringer, 2015; Ono et al., 2012; Perrone-McGovern et al., 2014; Stellar et al., 2015). Finally, of the four studies that measured respiration, two studies found a significant positive association between empathy and respiration (Miu & Baltes, 2012; Vanderpool & Barratt, 1970), one study found a significant negative association between empathy and respiration (Stellar et al., 2015), and one study found no significant association between empathy and respiration (Geringer, 2015). In sum, it does not seem that the type of stimulus examined or the physiological parameter measured is responsible for the mixed pattern of results observed in this body of literature.

APPENDIX B: Measures, Forms, and Scripts

- 1. Demographic Questionnaire**
- 2. Post Task Questionnaire**
- 3. Behavioral Coding Form**
- 4. Comfort Interaction Script**
- 5. Neutral Interaction Script**

Demographic Questionnaire

Participant #: _____

Date: _____

Height(in.): _____

Weight(lbs): _____

Please provide your email address so that we can contact you for part 2 of the study: _____

Your Information:

Age _____ yrs

Biological sex

- Male
- Female

Do you identify as:

- Male
- Female
- Prefer not to say

Are you Hispanic, Latino, or of Spanish origin?

- No, not of Hispanic, Latino, or Spanish origin
- Yes, Mexican, Mexican Am., Chicano
- Yes, Puerto Rican
- Yes, Cuban
- Yes, another Hispanic, Latino, or Spanish origin (please indicate) _____

Race- check all that apply

- White
- Black or African American

- American Indian or Alaska Native
- Asian
- Native Hawaiian or Other Pacific Islander
- Other (please indicate) _____

What is the highest level of education you have completed to date?

- High school
- 1 year of college
- 2 years of college
- 3 years of college
- 4 or more years of college

What is your intended major (s) at WVU? _____

Please describe any cardiovascular related illness that you may have, including high blood pressure (if none, please write "N/A"):

Please list any other medical or psychiatric problems that you have: _____

Please list any major surgeries and medical, or psychiatric illnesses you have had *in the past year*: _____

Females: Are you currently pregnant?

- Yes
- No

Females: Are you currently on birth control (contraceptives).

- Yes
- No

What type of birth control are you taking? _____

Please list any drugs (legal or otherwise) that you are currently taking including; birth control (contraceptives), heart medications, cold or allergy medications, over the counter medications, asthma medications, Beta-Blockers (i.e. Inderal, Tenormin), psychoactive drugs (i.e. Adderall, Xanax, Haldol, Lithium, Prozac), or diet pills: _____

Do you currently smoke cigarettes?

- Yes
- No

Have you smoked cigarettes within the last month?

- Yes
- No

Do you currently use smokeless tobacco?

- Yes
- No

Have you used smokeless tobacco within the last month?

- Yes
- No

Do you currently vape (with nicotine) ?

- Yes
- No

Have you vaped (with nicotine) within the last month?

- Yes
- No

Do you currently smoke electronic/e-cigarettes?

- Yes
- No

Have you smoked electronic/e-cigarettes within the last month?

- Yes
- No

How often do you drink alcohol?

- Never
- Infrequently (a few drinks per year)
- Occasionally (1-2 drinks per month)
- Weekly (1-3 drinks per week)
- Weekly (3-6 drinks per week)
- Daily (1-2 drinks per day)
- Daily (more than 2 drinks per day)

How many cups of **caffeinated** coffee, tea, soda, or energy drinks (e.g. Red Bull, 5-hr Energy) do you have on a typical day?

- None
- 1-2 cups per day
- 3-4 cups per day
- 5-6 cups per day
- 7-8 cups per day
- Greater than eight cups per day

How many times per week do you engage in aerobic physical activity?

- Never
- 1-2 times
- 3-6 times
- 7 or more times

For how long do you typically exercise on each occasion?

- 5-10 minutes
- 11-15 minutes
- 16-30 minutes
- 31-60 minutes
- More than 60 minutes

Family Information:

Imagine a ladder that represents where people stand in the United States.

At the **top** of the ladder are the people who are the best off – those who have the most money, the most education, and the most respected jobs. At the **bottom** are the people who are the worst off – who have the least money, least education, and the least respected jobs or no job. The higher up you are on this ladder, the closer you are to the people at the very top; the lower you are, the closer you are to the people at the very bottom.

On which rung of the ladder (1 being the lowest rung and 10 being the highest rung) would you place your family?

1.....2.....3.....4.....5.....6.....7.....8.....9.....10

Is your father currently living?

- yes
- no

Approximately how old is your father? _____

Did/does your father have high blood pressure (hypertension)?

- yes
- no

How certain are you that he did, or did not, have high blood pressure (hypertension)?

- Absolutely (100%) certain
- Almost (75%) certain
- Not sure at all (25%)
- No information by which to judge (0%)

Did/does your father have any heart problems such as angina (chest pains), a heart attack, or coronary heart disease?

- yes
- no

If yes, please specify if you are able: _____.

How certain are you that he did, or did not, have a heart problem as indicated above?

- Absolutely (100%) certain
- Almost (75%) certain
- Not sure at all (25%)
- No information by which to judge (0%)

Is your mother currently living?

- Yes
- No

Approximately how old is your mother? _____

Did/does your mother have high blood pressure (hypertension)?

- Yes
- No

How certain are you that she did, or did not, have high blood pressure (hypertension)?

- Absolutely (100%) certain
- Almost (75%) certain
- Not sure at all (25%)
- No information by which to judge (0%)

Did/does your mother have any heart problems such as angina (chest pains), a heart attack, or coronary heart disease?

- Yes
- No

If yes, please specify which problem(s) (if unsure, write "Unsure"):

_____.

How certain are you that she did, or did not, have a heart problem as indicated above?

- Absolutely (100%) certain
- Almost (75%) certain
- Not sure at all (25%)
- No information by which to judge (0%)

Behavioral Coding Form

Participant # _____
 Coded Date _____
 Scene _____

Coder _____
 Taped Date _____
 Interaction # _____

Informational Support

- SA Suggestion/Advice
- SI Situation Appraisal
- TE Teaching
- INF Information

Emotional Support

- RL Relationship
- SY Sympathy
- UE Understanding/Empath
- PE Personal Experience
- EC Expresses concern
- R Reassurance
- PA Physical Affection
- CF Confidentiality
- PY Prayer

Esteem Support

- CM Compliment
- VA Validation
- RB Relief of Blame

Tangible Aid

- DT Direct Task
- IT Indirect Task
- AP Active Participation
- WI Willingness
- LO Loan
- CR Complies with request

Social Network Support

- PR Presence
- AC Access
- CP Companions

Tension Reduction

- H Humor
- DE Distraction/Escape

Other:

- AG Verbal Agreement
- NOV Normative Verbal

Attentiveness

- LI Responsiveness
- IN Inquiries
- SP Statement of Problem

Negative Behavior

- IP Interrupt
- CN Complain
- DD Disagree/Disapprove
- CT Criticism
- IS Isolation

MICS Nonverbal

- AT Attention
- LA Laugh
- NO Normative
- NR No Response
- NT Not Tracking
- TO Turn Off

Minute 15 second segment

#1 1 _____
 2 _____
 3 _____
 4 _____

#2 1 _____
 2 _____
 3 _____
 4 _____

#3 1 _____
 2 _____
 3 _____
 4 _____

* How much did the participant display **empathy** during this interaction? (circle one)

Little/None (0)

Moderate (1)

High (2)

Comfort Interaction Script

Your friend tells you that their significant other has just suddenly broken up with her. Your friend seems to want to talk about it.

INSTRUCTIONS/GOAL: This is (confederate name), and she has just gone through a break-up with her significant other. Your goal for this interaction is to comfort her.

CONFEDERATE PROMPTS:

-But I miss/love him/her so much

-I don't want to move on

-I just want him/her to come talk to me

-I just wish I knew what he/she was doing right now

-It just won't be the same without him/her

-I just wish I would have been a better girlfriend, maybe he/she would have stayed

-There must be something wrong with me for him/her to leave like that

-I'll just have to wait for him/her to come back to me

-I don't understand why this keeps happening to me

-I didn't even see it coming

-This just hurts so bad

-I'll never be happy again

Neutral Interaction Script

Your friend is considering attending WVU, but they are unsure if it will be the right fit for them.

INSTRUCTIONS/GOAL: This is (confederate name), and she is trying to decide what college she should attend as an undergraduate. Your goal for this interaction is convince her that WVU is the best option.

CONFEDERATE PROMPTS:

- I don't know if I would fit in there.
- None of my other friends are going to WVU.
- It seems like a good school, but I hear Pitt is good too.
- A lot of people from my high school went to Pitt or Penn State.
- I don't know anyone in Morgantown.
- My parents think I should weigh all of my options.
- It's a really hard decision to make.
- I just wish I knew what to do.
- I don't know what I'll do for fun on the weekends.
- I've never been to West Virginia before.
- I don't even know what I want my major to be yet.
- There's so many schools to choose from, it's hard to decide where to go.

APPENDIX C: PRELIMINARY ANALYSES

Data Cleaning and Reduction

Initial data cleaning of the physiological data involved inspection of the data set for BP measurements that did not meet the criteria established by Marler, Jacob, Lehoszky, and Shapiro (1988). SBPs below 70 mmHg or above 250 mm Hg or DBPs below 45 mm Hg or above 150 mm Hg were considered for deletion or replaced by the most recent valid BP measurement during that specific period of the lab session. In addition, when pulse pressure (PP) (the difference between SBP and DBP) was less than 30 mm Hg, inconsistent BPs used to calculate this PP were replaced with the mean of the most proximate, valid BP(s). Of the 832 total BP measurements, 53 BP measurements (6.37% of the total) were adjusted according to these criteria.

Heart rate data and interbeat intervals (IBIs) were also inspected for abnormal data points and artefacts, or extreme deviations from the pattern of IBIs using the automatic artefact correction algorithm in the Kubios Premium HRV v3.1 software. After correction, mean HRs were calculated for each minute of the baseline, task, and rest periods, respectively.

Demographic Variables: SONA and Laboratory Higher/Lower Empathy Groups

Due to an inadequate sample size with which to conduct statistical analyses, SONA and laboratory samples as well as laboratory higher and lower empathy groups were visually compared on a number of demographic variables (See Tables C.1 and C.2). In sum, the laboratory sample resembled the SONA screening sample across all demographic variables; both samples largely consisted of young, single, Caucasian women with a heterosexual orientation. Furthermore, the demographic characteristics of women in the higher empathy group who participated in the laboratory session resembled those of women in the lower empathy group.

Preliminary Analyses: Cardiovascular Parameters

Baseline. Data for the baseline period was used to test whether there were significant differences in cardiovascular measures of heart rate and blood pressure across each minute of baseline using one-way repeated measures ANOVAs. Measures of baseline DBP, $F(2, 62) = .54$, $p = 0.58$, revealed no significant differences across each minute of the baseline period; thus DBP values were averaged to create a mean baseline DBP value. Analysis of baseline SBP, $F(2, 62) = 14.96$, $p < .001$, $\eta_p^2 = .33$, revealed a significant main effect, and mean comparisons revealed that SBP at minute 0 (116.1 mm Hg) was significantly larger than SBP at both minute 2 (111.7 mm Hg) and minute 4 (110.0 mm Hg). This result suggested a gradual laboratory habituation effect on SBP, so the first minute of the baseline SBP measure was removed and a mean baseline SBP was calculated by averaging minutes 2 and 4. Additionally, the repeated measures ANOVA of baseline HR with a Greenhouse-Geisser correction, $F(2.69, 83.48) = 13.67$, $p < .001$, $\eta_p^2 = .31$, revealed a significant main effect. Mean comparisons revealed that HR at minute 1 (75.0 bpm) was significantly lower than HR at minute 2 (79.0 bpm), minute 4 (80.5 bpm), and minute 5 (79.0 bpm). Though the mean comparisons revealed significant differences, this pattern did not suggest habituation occurred across the baseline period, so all HR baseline measures were averaged to obtain a mean baseline HR.

Rest 1 Period. Cardiovascular data for each minute of the rest period following the first task (Rest 1 Period) were analyzed in the same manner as the baseline data, using one-way repeated measures ANOVAs. Analyses of Rest 1 HR, $F(2.76, 85.56) = 1.58$, $p = 0.20$, and Rest 1 DBP, $F(1.60, 49.67) = .10$, $p = .86$, both with a Greenhouse-Geisser correction, revealed no significant main effects; thus, all Rest 1 HR measures and DBP measures were averaged, respectively. Analyses of SBP, $F(2, 62) = 3.28$, $p < .05$, $\eta_p^2 = .10$, revealed a significant main

effect. Mean comparisons revealed that SBP at minute 0 (113.4 mm Hg) was significantly higher than SBP at minute 4 (110.9 mm Hg). This seemed to indicate that participants were still recovering from Task 1 during the first minute of Rest 1, thus only minute 2 and minute 4 measures of SBP in Rest 1 were utilized to calculate average SBP for Rest 1.

Heart rate variability (HRV) measures for each protocol period were not subjected to one-way repeated measures ANOVAs because HRV measures are unreliable across one-minute time periods. For this reason, interbeat intervals from all valid HR measures underwent HRV spectral analysis using the Kubios software to obtain one baseline measure of HF-HRV, LF-HRV, and RMSSD, respectively. HRV measures (LF, HF, and RMSSD) for baseline, rest, and task periods were log transformed prior to analysis in order to ensure a normal distribution of these variables.

In order to measure cardiovascular reactivity to each of the interaction tasks, HR and BP values across each minute of the task were averaged to create a mean HR and BP measure for the Comfort Task, and a mean HR and BP for the Neutral Task.

Task Period. To test and confirm that participants exhibited increased autonomic activity to the interaction tasks, a 2 (Task Type: Neutral, Comfort) x 2 (Period: Baseline, Task) repeated measures ANOVA was conducted for each physiological measure (See ANOVA Summary Tables C.3-C.8). For SBP, the main effect of Task Type was significant and SBP was larger during the neutral task ($M = 121.1$ mm Hg, $SE = 2.1$) as compared to the comfort task ($M = 118.9$ mm Hg, $SE = 2.0$). The main effect for Period was also significant, as SBP was lower during the rest period for each task ($M = 111.5$ mm Hg, $SE = 1.8$) as compared to SBP during the task periods overall ($M = 128.5$ mm Hg, $SE = 2.5$). There was no significant interaction between Task Type and Period for SBP measures. For DBP, the main effect of Task Type and interaction

of Task Type and Period were not significant. The DBP main effect of Period was significant, and DBP was lower during the rest period for each task ($M = 62.8$ mm Hg, $SE = 1.5$) as compared to the task periods overall ($M = 72.5$ mm Hg, $SE = 1.8$).

The same analyses were conducted on measures of HR. For HR, the main effect of Task Type was significant, as HR was greater during the neutral task ($M = 83.0$ bpm, $SE = 1.9$) as compared to the comfort task ($M = 81.5$ bpm, $SE = 2.0$). In addition, the main effect for Period was also significant, and HR was lower during the rest period for each task ($M = 77.5$ bpm, $SE = 1.9$) as compared to the task period overall ($M = 87.0$ bpm, $SE = 2.1$). There was no significant interaction between Task Type x Period for HR.

For measures of LF-HRV, the main effect of Task Type and the interaction of Task Type and Period were not significant. The main effect for Period was significant, as LF-HRV was lower during the rest period for each task ($M = 6.9$ log ms², $SE = .15$) as compared to the task period overall ($M = 7.3$ log ms², $SE = .14$).

For measures of HF-HRV, the main effect of Task Type and the interaction of Task Type and Period were not significant. The main effect for Period was significant, and HF-HRV was higher during the rest period for each task ($M = 6.8$ log ms², $SE = .20$) as compared to the task period overall ($M = 6.3$ log ms², $SE = .20$).

For measures of RMSSD, the main effect of Task Type was not significant. The main effect for Period was significant, and RMSSD was higher during the rest period for each task ($M = 1.7$ log ms, $SE = .05$) as compared to the task period overall ($M = 1.5$ log ms, $SE = .04$). In addition, there was no significant interaction between Task Type x Period for RMSSD measures.

In sum, all measures of cardiovascular functioning revealed that participants reacted to the two interpersonal tasks in predicted ways, showing increases in BP, HR, and LF-HRV,

parameters that revealed increased sympathetic nervous system arousal, and decreases in HF-HRV and RMSSD, parameters that revealed reduced parasympathetic activity. For both measures of HR and SBP, reactions were greater for the neutral task than the comfort task, indicating less sympathetic nervous system activation during the comfort task.

Preliminary Analysis: Behavioral Coding

Task behaviors were coded by two independent raters to 80% agreement or better. Interrater agreement percentages were calculated by dividing the frequency with which the observers agreed that a category should be coded by the number of agreements plus disagreements, and codes that achieved an 80% agreement rate or better were deemed acceptable for analysis. Several categories of behavioral responding were not observed during either task or were only used infrequently in one of the tasks. Several of the categories never occurred (e.g. Teaching, Relationship, Loan) while others occurred more frequently (e.g. Suggestion/Advice, Information Giving, Personal Experience). As depicted in Tables C.9 and C.10, the coding of behavioral categories was determined to be reliable.

Based on The Social Support Behavior Code (SSBC) by Suhr, Cutrona, Krebs, and Jensen (2004), individual behavioral codes were summed to create 8 major behavioral parameters (Informational Support, Emotional Support, Esteem Support, Tangible Aid, Social Network Support, Tension Reduction, Attentiveness, and Positive Non-Verbal). Individual codes that did not fall under one of the 8 behavioral parameters were analyzed individually and included Verbal Agreement, Normative Verbal Response, Normative Nonverbal Response, Disagree/Disapprove, and Not Tracking (Negative Nonverbal).

All behavioral parameters were checked for normality, skewness, and kurtosis. Two main behavioral parameters were not normally distributed and required a log transformation prior to

analysis: Esteem Support (neutral task), and Social Network Support (neutral task). In order to run a repeated measures analysis, comfort tasks were also log transformed for these two parameters. However, for purposes of presenting descriptive statistics, non-transformed means and standard errors were reported.

Preliminary Analysis: SONA versus Laboratory Empathy Score Comparison

A 2 (Measurement Time: SONA Empathy, Lab Empathy) x 2 (Empathy Group: Higher vs. Lower) repeated measures ANOVA was conducted to determine whether there were significant differences in empathy scores between higher and lower empathy groups during initial SONA survey completion and lab survey completion. Results indicated that there was a significant main effect of Empathy Group, $F(1, 30) = 91.02, p < .001, \eta_p^2 = .75$. Mean empathy scores for the lower empathy group were $M = 69.9, SE = 1.15$, and mean empathy scores for the higher empathy group were $M = 85.4, SE = 1.15$, confirming that the categorization scheme employed resulted in groups that differed in dispositional empathy. There was also a significant main effect of Measurement Time, $F(1, 30) = 12.26, p < .01, \eta_p^2 = .29$. The mean empathy score for all participants on the SONA empathy questionnaire was $M = 75.4, SE = .99$, and the mean empathy score on the laboratory empathy questionnaire was $M = 79.9, SE = 1.1$, indicating that empathy scores were higher when gathered in person versus empathy scores obtained via SONA administration. Finally, there was a significant interaction between Measurement Time and Empathy Group, $F(1, 30) = 6.40, p < .05, \eta_p^2 = .18$. For those in the lower empathy group, there was a significant main effect of Measurement Time, $F(1, 15) = 13.36, p < .01, \eta_p^2 = .47$. Specifically, the mean SONA empathy score was $M = 66.0, SE = 1.53$, and the mean laboratory empathy score was $M = 73.8, SE = 1.73$. For the higher empathy group, there was no significant main effect of Measurement Time, $F(1, 15) = .740, p = .40, \eta_p^2 = .05$. Among the 16 lower empathy participants, 6 still obtained scores that met study selection criteria for being

categorized as lower dispositional empathy; of the remaining 10 lower empathy participants, 6 exhibited scores that no longer met the selection criteria but did not meet criteria for higher empathy participants, and 4 met criteria for higher empathy participants. For purposes of analysis, the initial categorization of dispositional empathy based on SONA scores was retained.

Table C.1. Demographic comparison of women in the SONA sample and lab sample.

Demographic Variable	SONA Sample (N = 276)	Lab Sample (N = 32)
Age	M = 19.02, SD = 1.61	M = 18.72, SD = 0.92
Marital Status		
Single	268	32
Married	3	0
Widowed	0	0
Separated	1	0
No Response	4	0
Ethnicity		
White/Caucasian	234	28
African-American/Black	7	1
Asian	6	0
Hispanic/Latino(a)	6	1
Other	1	0
Multiple races	19	2
No response	3	0
Sexual orientation		
Heterosexual	239	31
Lesbian/gay	7	0
Bisexual	21	1
Other	4	0
No response	5	0
Political orientation		
Democrat	72	11
Republican	79	4
Independent	61	11
Libertarian	7	1
Other	16	0
No response	41	5
Religion		
Christian	180	17
Not religious	53	7
Atheist	11	0
Agnostic	5	4
Jewish	2	0
Muslim	4	0
Other	8	1
Hindu	0	0
Buddhist	5	0
No response	8	3
Hometown		
Rural	35	2
Small town	94	15
Small city	49	3
Medium-sized city	29	2

Suburban	49	9
Large city	16	1
No response	4	0
Political characterization		
Strongly liberal	16	3
Liberal	44	3
Slightly liberal	30	5
Moderate	101	12
Slightly conservative	30	3
Conservative	26	2
Strongly conservative	3	0
No response	26	4
Year in college		
First	120	20
Second	90	8
Third	35	3
Fourth	25	1
Fifth or above	3	0
No response	3	0
Major *		
Healthcare	103	17
Behavioral & Social Sci.	99	8
Education	7	1
Basic Science	25	2
Other	1	1
Business	16	3
Arts	5	0
Agriculture & Natural Sci.	8	0
No response	3	0
Language	1	0
Engineering	3	0
Undecided	5	1

*Note: Double majors were counted in more than one category

Table C.2. Demographic comparison of higher empathy and lower empathy groups.

Demographic Variable	Higher Empathy (N = 16)	Lower Empathy (N = 16)
Age	M = 18.75, SD = 1.06	M = 18.69, SD = 0.79
Marital Status		
Single	16	16
Married	0	0
Widowed	0	0
Separated	0	0
Ethnicity		
White/Caucasian	16	12
African-American/Black	0	1
Asian	0	0
Hispanic/Latino(a)	0	1
Other	0	0
Multiple races	0	2
Sexual orientation		
Heterosexual	15	15
Lesbian/gay	0	0
Bisexual	1	0
Other	0	0
No response	0	1
Political orientation		
Democrat	7	4
Republican	3	1
Independent	3	8
Libertarian	1	0
Other	0	0
No response	2	3
Religion		
Christian	9	8
Not religious	4	3
Atheist	0	0
Agnostic	1	3
Jewish	0	0
Muslim	0	0
Other	1	0
Hindu	0	0
Buddhist	0	0
No response	1	2
Hometown		
Rural	1	1
Small town	7	8
Small city	1	2
Medium-sized city	2	0
Suburban	5	4
Large city	0	1

Political characterization		
Strongly liberal	1	2
Liberal	2	1
Slightly liberal	4	1
Moderate	5	7
Slightly conservative	0	3
Conservative	2	0
Strongly conservative	0	0
No response	2	2
Year in college		
First	9	11
Second	5	3
Third	1	1
Fourth	1	1
Fifth or above	0	0
Major		
Healthcare	8	9
Behavioral & Social Sci.	5	3
Education	1	1
Basic Science	0	2
Other	0	0
Business	2	1
Arts	0	0
Agriculture & Natural Sci.	0	0
Language	0	0
Engineering	0	0
Undecided	0	1

Tables C.3-C.8. Summary Tables for Task Period Reactivity: Task Type (Neutral, Comfort) x Period (Task, Baseline) Repeated Measures ANOVAs**3. SBP.**

Source	SS	df	MS	F	p	η^2_p
Task Type	156.42	1	156.42	4.39	.045	.124
Error (Task Type)	1105.64	31	35.67			
Task Period	9273.52	1	9273.52	104.05	.000	.770
Error (Task Period)	2762.80	31	89.12			
Type * Period	5.08	1	5.08	.12	.729	.004
Error (Type * Period)	1290.98	31	41.65			

4. DBP.

Source	SS	df	MS	F	p	η^2_p
Task Type	8.60	1	8.60	.186	.669	.006
Error (Task Type)	1433.62	31	46.25			
Task Period	2972.53	1	2972.53	29.35	.000	.486
Error (Task Period)	3139.36	31	101.27			
Type * Period	36.30	1	36.30	.636	.431	.020
Error (Type * Period)	1768.25	31	57.04			

5. HR.

Source	SS	df	MS	F	p	η^2_p
Task Type	71.99	1	71.99	8.43	.007	.214
Error (Task Type)	264.76	31	8.54			
Task Period	2910.36	1	2910.36	92.63	.000	.749
Error (Task Period)	973.97	31	31.42			
Type * Period	43.51	1	43.51	4.13	.051	.118
Error (Type * Period)	326.38	31	10.53			

6. LF-HRV.

Source	SS	df	MS	F	p	η^2_p
Task Type	.047	1	.047	.228	.636	.007
Error (Task Type)	6.44	31	.208			
Task Period	4.16	1	4.16	8.55	.006	.216
Error (Task Period)	15.06	31	.486			
Type * Period	.172	1	.172	1.36	.253	.042
Error (Type * Period)	3.94	31	.127			

7. HF-HRV.

Source	SS	df	MS	F	p	η_p^2
Task Type	.036	1	.036	.231	.634	.007
Error (Task Type)	4.87	31	.157			
Task Period	8.40	1	8.40	31.15	.000	.501
Error (Task Period)	8.36	31	.270			
Type * Period	.078	1	.078	.373	.546	.012
Error (Type * Period)	6.46	31	.208			

8. RMSSD.

Source	SS	df	MS	F	p	η_p^2
Task Type	.008	1	.008	1.94	.173	.059
Error (Task Type)	.124	31	.004			
Task Period	.493	1	.493	43.63	.000	.585
Error (Task Period)	.350	31	.011			
Type * Period	.001	1	.001	.244	.625	.008
Error (Type * Period)	.137	31	.004			

Note. p-values for significant effects are bolded

Table C.9. Interrater Agreement of Behavioral Data: Neutral Task

Informational Support		Tension Reduction	
Suggestion/Advice	94.8%	Humor	95.8%
Situation Appraisal	97.9%	Distraction/Escape	100%
Information Giving	88.5%		
Emotional Support		Attentiveness	
Sympathy	100%	Responsiveness	92.7%
Understanding/Empathy	96.9%	Inquiries	97.9%
Personal Experience	93.8%	Statement of Problem	100%
Express Concern	100%		
		Positive Nonverbal	
Reassurance	95.8%	Attention	100%
		Laugh	83.3%
Esteem Support		Negative Nonverbal	
Compliment	100%	Not Tracking	90.6%
Validation	100%		
Relief of Blame	100%		
Tangible Aid		Other	
Direct Task	100%	Verbal Agreement	96.9%
Willingness	100%	Normative Verbal response	87.5%
		Normative Non-Verbal response	91.7%
Social Network Support			
Presence	98.9%		
Access	100%		
Companions	98.9%		

Table C.10. Interrater Agreement of Behavioral Data: Comfort Task

Informational Support		Tension Reduction	
Suggestion/Advice	96.9%	Humor	95.8%
Situation Appraisal	92.7%	Distraction/Escape	96.9%
Information Giving	98.9%		
Emotional Support		Attentiveness	
Sympathy	96.9%	Responsiveness	92.7%
Understanding/Empathy	95.8%	Inquiries	94.8%
Personal Experience	96.9%	Statement of Problem	100%
Express Concern	100%	Negative Verbal	
Reassurance	94.8%	Disagree/Disapprove	100%
Esteem Support		Positive Nonverbal	
Compliment	95.8%	Attention	100%
Validation	97.9%	Laugh	86.5%
Relief of Blame	94.8%	Negative Nonverbal	
Tangible Aid		Not Tracking	92.7%
Direct Task	100%	Other	
Willingness	100%	Verbal Agreement	96.9%
Social Network Support		Normative verbal response	87.5%
Presence	96.9%	Normative Non-Verbal response	90.6%
Access	96.9%		
Companions	100%		

*Note: Codes not utilized during both scenes: Teaching, Relationship, Physical Affection, Confidentiality, Prayer, Indirect Task, Active Participation, Loan, Complies with Request, Interrupt, Complain, Criticism, Isolation, Turn Off, No Response. Codes not utilized during neutral task only: Disagree/Disapprove, Express Concern

APPENDIX D: ANALYSES OF CONFEDERATE EFFECTS

Confederate Effects on Cardiovascular Reactivity. To assess for potential confederate effects on physiological measures during task periods, a 2 (Confederate: A, B) x 2 (Task Type: Neutral, Comfort) mixed factors ANOVA was conducted for each physiological measure, covarying for resting physiological measures (see Summary Tables D.1-D.6). Only results pertaining to the main effect for Confederate and the Confederate x Task Type interaction are reported below because results for the main effect for Task Type are reported in the primary study analyses.

For SBP, the main effect of Confederate was not significant, but there was a significant interaction between Confederate and Task Type, $F(1, 29) = 6.25, p < .05, \eta_p^2 = .18$. Post hoc analyses of simple main effects revealed a significant difference between SBP reactivity during the neutral and comfort task for those who interacted with Confederate A, $F(1, 14) = 7.00, p < .05, \eta_p^2 = .33$. Specifically, mean SBP for the comfort task was $M = 124.3$ mm Hg, $SE = 2.99$, and mean SBP for the neutral task was $M = 130.8$ mm Hg, $SE = 3.13$, when interacting with Confederate A. Comparisons between the neutral and comfort task reactivity for those who interacted with Confederate B revealed no significant difference, $F(1, 14) = .27, p = .61, \eta_p^2 = .02$. Specifically, mean SBP for the comfort task was $M = 130.2$ mm Hg, $SE = 2.23$, and mean SBP for the neutral task was $M = 129.0$ mm Hg, $SE = 3.26$, when interacting with Confederate B.

For all other cardiovascular parameters (DBP, HR, LF-HRV, HF-HRV, RMSSD) the main effects of Confederate and the interactions between Confederate and Task Type were not significant.

Confederate Effects on Self-reported Affect. A 2 (Confederate: A, B) x 2 (Task Type: Comfort, Neutral) mixed factors ANCOVA was conducted, covarying pre-task self-report

scores to determine whether there were significant differences between those assigned to Confederate A and Confederate B on negative and positive PANAS measures. Like cardiovascular measures, only results pertaining to the main effects for Confederate and the Confederate by Task Type interaction effects are reported.

For the positive PANAS scores, the main effect of Confederate was not significant, $F(1, 29) = .00, p = .97, \eta_p^2 = .00$, and there was no significant interaction between Confederate and Task Type, $F(1, 29) = 1.23, p = .28, \eta_p^2 = .04$.

Similarly, for the negative PANAS scores, the main effect of Confederate was not significant, $F(1, 29) = 1.12, p = .30, \eta_p^2 = .04$, and there was no significant interaction between Confederate and Task Type, $F(1, 29) = .05, p = .83, \eta_p^2 = .00$. (See Table D.7).

Confederate Effects on Self-reported Interpersonal Goals. A similar 2 (Confederate) x 2 (Task Type) mixed factors ANOVA was conducted on measures of agentic and communal striving (interpersonal goals). Like prior analyses, only results pertaining to the main effects for Confederate and the Confederate by Task Type interaction effects are reported. No main effects for Confederate or Confederate by Task Type interactions were significant for either agentic or communal striving (see Table D.8).

Confederate Effects on Self-reported Task Appraisals. The rest of the items on the Post-Task Questionnaire were analyzed individually using a comparable series of Confederate by Task Type mixed factors ANOVAs. Only results regarding the main effects for Confederate and the Confederate by Task Type interaction effects are reported. (See Table D.9). The only significant confederate effect across all of the self-reported task appraisals was a significant Confederate by Task Type interaction for the item: “How upset are you by your performance in the interaction?” $F(1, 30) = 4.93, p = .04, \eta_p^2 = .14$. Post hoc simple main effects analyses were

conducted to determine the nature of the interaction. For only those assigned to Confederate A, there was no significant difference between Task Types, $F(1, 15) = .03, p = .87, \eta_p^2 = .00$. Mean ratings of being upset for the comfort task was $M = 1.9, SE = .24$, and for the neutral task was $M = 1.9, SE = .30$. For those assigned to Confederate B, however, there was a significant difference between Task Types, $F(1, 15) = 10.91, p < .01, \eta_p^2 = .42$, such that ratings of being upset were higher for the comfort task, $M = 2.5, SE = .35$, as compared to the neutral task, $M = 1.5, SE = .20$.

Confederate Effects on Behavioral Measures. Analysis of only one behavioral parameter yielded a significant Confederate by Task Type interaction, Esteem Support, $F(1, 30) = 5.27, p < .05, \eta^2 = .149$. For those assigned to Confederate A, there was a significant difference between esteem support scores of the two tasks, $F(1, 15) = 21.14, p < .001, \eta^2 = .585$. Mean score for the comfort task was $M = 2.5, SE = 1.12$. and mean score for the neutral task was $M = 1.3, SE = 1.10$. For those assigned to Confederate B, there was also a significant difference between esteem support scores across Task Types, $F(1, 15) = 67.25, p < .001, \eta^2 = .818$. Mean score for the comfort task was $M = 3.9, SE = 1.12$, and mean score for the neutral task was $M = 1.3, SE = 1.10$. Mean comparison for the comfort task across Confederates revealed a significant simple main effect of Confederate for the Esteem Support parameter, $F(1, 30) = 6.41, p < .05, \eta^2 = .176$; however, for the neutral task, no significant simple main effect was detected, $F(1, 30) = .016, p = .899, \eta^2 = .001$ (See Tables D.10 and D.11).

Summary of Confederate Effects. Across all cardiovascular, self-report, and behavioral variables, very few differences were observed between participants when interacting with Confederate A versus Confederate B. SBPs were lower during the comfort task, but only for participants interacting with Confederate A. Self-reported ratings of being upset were higher in

response to the comfort task than the neutral task, but this was only observed for participants interacting with Confederate B. Finally, participants interacting with Confederate B displayed more esteem support than those interacting with Confederate A, but only during the comfort task. Based on this pattern of findings, it appears that during the comfort task, participants interacting with Confederate A showed less esteem support, reported getting less upset, and exhibited lower SBP responses than participants interacting with Confederate B. However, because no other Confederate effects were observed and half of the Higher and Lower Empathy participants interacted with each confederate, Confederate was not included as a covariate in the primary analyses.

Tables D.1-D.6. Summary Tables for Confederate Effects: Task Type (Neutral, Comfort) x Confederate (A, B) Repeated Measures ANOVAs

1. SBP.						
Source	SS	df	MS	F	p	η^2
Confederate	60.84	1	60.84	0.32	.574	.011
Error (Confederate)	5460.49	29	188.29			
Task Type	134.19	1	134.19	3.16	.086	.098
Task Type * Confederate	265.30	1	265.30	6.25	.018	.177
Error (Task Type and TT x C)	1230.25	29	42.42			
2. DBP.						
Source	SS	df	MS	F	p	η^2
Confederate	63.01	1	63.07	0.35	.558	.012
Error (Confederate)	5209.94	29	179.65			
Task Type	3.20	1	3.20	0.04	.841	.001
Task Type * Confederate	90.28	1	90.28	1.16	.290	.038
Error (Task Type and TT x C)	2257.67	29	77.85			
3. HR.						
Source	SS	df	MS	F	p	η^2
Confederate	24.00	1	24.00	0.36	.552	.012
Error (Confederate)	1921.69	29	66.27			
Task Type	119.97	1	119.97	7.98	.008	.216
Task Type * Confederate	8.24	1	8.24	0.55	.465	.019
Error (Task Type and TT x C)	436.20	29	15.04			

4. LF-HRV

Source	SS	df	MS	F	p	η^2
Confederate	0.01	1	0.01	0.01	.912	.000
Error (Confederate)	22.64	29	0.78			
Task Type	0.24	1	0.24	1.10	.302	.037
Task Type * Confederate	0.15	1	0.15	0.69	.412	.023
Error (Task Type and TT x C)	6.33	29	0.22			

5. HF-HRV

Source	SS	df	MS	F	p	η^2
Confederate	0.12	1	0.12	0.25	.624	.008
Error (Confederate)	13.62	29	0.47			
Task Type	0.000	1	0.000	0.000	.996	.000
Task Type * Confederate	0.97	1	0.97	4.10	.052	.124
Error (Task Type and TT x C)	6.88	29	0.24			

6. RMSSD

Source	SS	df	MS	F	p	η^2
Confederate	157.93	1	157.93	1.12	.298	.037
Error (Confederate)	4081.64	29	140.75			
Task Type	149.07	1	149.07	3.77	.062	.115
Task Type * Confederate	8.39	1	8.39	0.21	.648	.007
Error (Task Type and TT xC)	1146.90	29	39.55			

Note. p-values for significant effects are bolded

Table D.7. Summary ANOVA Table for PANAS Score Comparisons between Confederate Type: Task Type (Neutral, Comfort) x Confederate (A, B) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
PANAS Positive Total						
Confederate	.03	1	.03	.00	.966	.000
Error (Confederate)	496.95	29	17.14			
Task Type	118.81	1	118.81	10.12	.003	.259
Task * Confederate	14.44	1	14.44	1.23	.277	.041
Error (Task Type)	340.53	29	11.74			
PANAS Negative Total						
Confederate	5.96	1	5.96	1.12	.299	.037
Error (Confederate)	154.24	29	5.32			
Task Type	23.78	1	23.78	4.66	.039	.138
Task * Confederate	.25	1	.25	.05	.826	.002
Error (Task Type)	148.08	29	5.11			

Note. p-values for significant effects are bolded

Table D.8. Summary ANOVA Table for Post-task Agentic and Communal Score Comparisons between Confederate Type: Task Type (Neutral, Comfort) x Confederate (A, B) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
PTQ Agentic						
Confederate	1.56	1	1.56	.220	.643	.007
Error (Confederate)	213.19	30	7.11			
Task Type	100.00	1	100.00	17.97	.000	.375
Task * Confederate	.063	1	.063	.011	.916	.000
Error (Task Type)	166.94	30	5.57			
PTQ Communal						
Confederate	2.64	1	2.64	.302	.586	.010
Error (Confederate)	261.97	30	8.73			
Task Type	83.27	1	83.27	27.88	.000	.482
Task * Confederate	5.64	1	5.64	1.89	.180	.059
Error (Task Type)	89.59	30	2.99			

Note. p-values for significant effects are bolded

Table D.9. Summary ANOVA Table for Post-task Individual Item Comparisons between Confederates: Task Type (Neutral, Comfort) x Confederate (A, B) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
How realistic?						
Confederate	3.06	1	3.06	2.10	.157	.066
Error (Confederate)	43.69	30	1.46			
Task Type	3.06	1	3.06	3.84	.059	.113
Task * Confederate	.000	1	.000	.000	1.00	.000
Error (Task Type)	23.94	30	.798			
How likely to encounter?						
Confederate	.391	1	.391	.334	.568	.011
Error (Confederate)	35.09	30	1.17			
Task Type	5.64	1	5.64	6.05	.020	.168
Task * Confederate	1.89	1	1.89	2.03	.165	.063
Error (Task Type)	27.97	30	.932			
How stressful to handle?						
Confederate	.000	1	.000	.000	1.000	.000
Error (Confederate)	36.00	30	1.53			
Task Type	10.56	1	10.56	18.24	.000	.378
Task * Confederate	.063	1	.063	.108	.745	.004
Error (Task Type)	17.38	30	.579			
How stressful to engage?						
Confederate	1.00	1	1.00	.455	.505	.015
Error (Confederate)	65.94	30	2.20			
Task Type	5.06	1	5.06	6.92	.013	.188
Task * Confederate	.000	1	.000	.000	1.00	.000
Error (Task Type)	21.94	30	.731			
How meaningful?						
Confederate	.141	1	.141	.105	.748	.003
Error (Confederate)	40.22	30	1.34			
Task Type	.016	1	.016	.028	.867	.001
Task * Confederate	.016	1	.016	.028	.867	.001
Error (Task Type)	16.47	30	.549			

How positive?

Confederate	.250	1	.250	.170	.683	.006
Error (Confederate)	44.19	30	1.47			
Task Type	2.25	1	2.25	3.82	.060	.113
Task * Confederate	.063	1	.063	.106	.747	.004
Error (Task Type)	17.69	30	.590			

How negative?

Confederate	.063	1	.063	.051	.823	.002
Error (Confederate)	36.88	30	1.23			
Task Type	1.56	1	1.56	3.95	.056	.116
Task * Confederate	.563	1	.563	1.42	.243	.045
Error (Task Type)	11.88	30	.396			

How upset?

Confederate	.141	1	.141	.089	.767	.003
Error (Confederate)	47.22	30	1.57			
Task Type	3.52	1	3.52	3.84	.059	.113
Task * Confederate	4.52	1	4.52	4.93	.034	.141
Error (Task Type)	27.47	30	.916			

How difficult?

Confederate	.563	1	.563	.237	.630	.008
Error (Confederate)	71.19	30	2.37			
Task Type	4.00	1	4.00	4.01	.054	.118
Task * Confederate	.063	1	.063	.063	.804	.002
Error (Task Type)	29.94	30	.998			

Note. p-values for significant effects are bolded

Table D.10. Summary ANOVA Table for Behavioral Parameters: Task Type (Neutral, Comfort) x Confederate (A, B) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
Information Support						
Confederate	1.00	1	1.00	.067	.798	.002
Error (Confederate)	450.00	30	15.00			
Task Type	30.25	1	30.25	4.29	.047	.125
Task * Confederate	1.00	1	1.00	.142	.709	.005
Error (Task Type)	211.75	30	7.06			
Emotional Support						
Confederate	22.56	1	22.56	2.37	.134	.073
Error (Confederate)	285.44	30	9.52			
Task Type	49.00	1	49.00	11.18	.002	.272
Task * Confederate	1.56	1	1.56	.357	.555	.012
Error (Task Type)	131.44	30	4.38			
Esteem Support						
Confederate	.135	1	.135	2.96	.096	.090
Error (Confederate)	1.37	30	.046			
Task Type	2.42	1	2.42	80.56	.000	.729
Task * Confederate	.159	1	.159	5.27	.029	.149
Error (Task Type)	.903	30	.030			
Social Network Support						
Confederate	.062	1	.062	1.36	.253	.043
Error (Confederate)	1.36	30	.045			
Task Type	.020	1	.020	.505	.483	.017
Task * Confederate	.042	1	.042	1.07	.310	.034
Error (Task Type)	1.18	30	.039			
Attentiveness						
Confederate	40.64	1	40.64	1.81	.188	.057
Error (Confederate)	671.97	30	22.40			
Task Type	3.52	1	3.52	.617	.438	.020
Task * Confederate	13.14	1	13.14	2.31	.139	.071
Error (Task Type)	170.84	30	5.70			
Positive Non-Verbal						
Confederate	.250	1	.250	.009	.924	.000
Error (Confederate)	819.69	30	27.32			
Task Type	495.06	1	495.06	51.27	.000	.631
Task * Confederate	12.25	1	12.25	1.27	.269	.041
Error (Task Type)	289.69	30	9.66			

Table D.11. Summary ANOVA Table for Individual Behavior Codes: Task Type (Neutral, Comfort) x Confederate (A, B) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
Not Tracking						
Confederate	.052	1	.052	.190	.666	.006
Error (Confederate)	8.23	30	.274			
Task Type	.006	1	.006	.069	.794	.002
Task * Confederate	.010	1	.010	.119	.733	.004
Error (Task Type)	2.63	30	.088			
Normative Verbal						
Confederate	1.27	1	1.27	.269	.608	.009
Error (Confederate)	141.34	30	4.71			
Task Type	11.39	1	11.39	4.96	.034	.142
Task * Confederate	1.27	1	1.27	.552	.463	.018
Error (Task Type)	68.84	30	2.30			
Normative Non-Verbal						
Confederate	.063	1	.063	.004	.951	.000
Error (Confederate)	487.88	30	16.26			
Task Type	6.25	1	6.25	2.30	.140	.071
Task * Confederate	2.25	1	2.25	.828	.370	.027
Error (Task Type)	81.50	30	2.72			

Note. p-values for significant effects are bolded

APPENDIX E: ANALYSES OF ORDER EFFECTS

Order Effects on Cardiovascular Measures – Rest Periods. To assess for potential task order effects on resting physiological measures, a 2 x 2 [Order (1, 2) x Task Type (Pre-Neutral Task Rest Period, Pre-Comfort Task Rest Period)] mixed factors ANOVA was conducted for each cardiovascular parameter (see Summary Table E.1). Only results pertaining to the main effect for Order and the Order x Task Type interaction are reported because results for the main effect for Task Type are reported in the primary study analyses.

Significant interactions between Order and Task Type were observed for DBP, $F(1, 30) = 4.75, p < .05, \eta_p^2 = .14$, and for HF-HRV, $F(1, 30) = 4.67, p < .05, \eta_p^2 = .14$. For those who completed the neutral task followed by the comfort task, there was no significant difference between pre-neutral ($M = 61.1$ mmHg, $SE = 1.79$) and pre-comfort ($M = 62.2$ mmHg, $SE = 2.12$) rest periods on DBP, $F(1, 15) = .38, p = .55, \eta_p^2 = .03$. For those who completed the comfort task followed by the neutral task, however, there was a significant simple main effect of Task Type, $F(1, 15) = 6.09, p < .05, \eta_p^2 = .29$, such that DBPs were higher during the pre-neutral rest period, $M = 66.2$ mm Hg, $SE = 2.72$, as compared to the pre-comfort rest period, $M = 61.9$ mm Hg, $SE = 2.70$.

Regarding the Order by Task Type interaction on HF-HRV, there was no significant difference in resting HF-HRV during the pre-neutral ($M = 7.0$ log ms², $SE = .23$) and pre-comfort ($M = 6.9$ log ms², $SE = .24$) rest periods for those who completed the neutral task followed by the comfort task, $F(1, 15) = .684, p = .42, \eta_p^2 = .04$. For those who completed the comfort task followed by the neutral task, however, there was a significant simple main effect of Task Type, $F(1, 15) = 4.59, p < .05, \eta_p^2 = .23$, such that measures of HF-HRV were higher during the pre-comfort rest period, $M = 6.8$ log ms², $SE = .31$, as compared to the pre-neutral rest period, $M =$

6.5 log ms², $SE = .35$. Together, these findings indicate that participants were still exhibiting residual arousal (increased DBP; decreased HF-HRV) from the comfort task during the pre-neutral rest period, a phenomena that was not observed when the neutral task occurred first. No other main effects of Order or Order by Task Type interactions were detected for any other resting physiological parameter.

Order Effects on Cardiovascular Measures – Task Periods. A series of 2 x 2 (Order X Task Type (Neutral, Comfort)) mixed factors ANCOVAs was conducted, covarying pre-task rest periods, to determine if there was an influence of task order on cardiovascular measures during the neutral and comfort task. (See Summary Table E.2). Main effects on Task Type are not reported here as they are reported in the primary study analyses.

For HR, the interaction between Order and Task Type was significant, $F(1, 29) = 17.85, p = <.001, \eta_p^2 = .38$. Post hoc analyses were conducted to determine simple main effects. Results revealed no significant differences in HR reactivity between the two task orders during the comfort task, $F(1, 29) = .16, p = .69, \eta_p^2 = .01$. For those who completed the neutral task followed by the comfort task, the covariance adjusted mean HR for the comfort task was $M = 85.2$ bpm, $SE = 1.72$, and for those who completed the comfort task followed by the neutral task the covariance adjusted mean HR for the comfort task was $M = 86.2$ bpm, $SE = 1.72$. In contrast, results did reveal a significant difference in HR reactivity between task orders during the neutral task, $F(1, 29) = 4.27, p = <.05, \eta_p^2 = .13$. For those who completed the neutral task followed by the comfort task, the covariance adjusted mean HR for the neutral task was $M = 90.6$ bpm, $SE = 1.48$, and for those who completed the comfort task followed by the neutral task the covariance adjusted mean for the neutral task was $M = 86.2$ bpm, $SE = 1.48$. (See Figure E.1). HR reactions

were higher during the neutral task than the comfort task, but only when the neutral task was presented first.

For SBP, DBP and all HRV measures, the main effects of task order were not significant, nor were the interactions between task order and task type significant.

Order Effects on Measures of Self-Reported Affect. Comparable mixed factors 2 (Order) x 2 (Task Type) ANCOVAs were conducted on positive and negative PANAS scores, covarying pre-task self-report scores, to assess for order effects on self-reported affect.

For the positive PANAS scores, the main effect of Order was significant, $F(1, 29) = 5.06$, $p < .05$, $\eta_p^2 = .15$. Specifically, the covariance adjusted mean score for Task Order 1 (neutral task followed by comfort) was $M = 23.5$, $SE = 1.47$, and for Task Order 2 (comfort task followed by neutral) was $M = 28.6$, $SE = 1.47$. The interaction between Order and Task Type was not significant, $F(1, 29) = 2.00$, $p = .17$, $\eta_p^2 = .06$.

For the negative PANAS scores, the main effect of Order was not significant, $F(1, 29) = 2.15$, $p = .15$, $\eta_p^2 = .07$, but there was a significant Order by Task Type interaction, $F(1, 29) = 14.53$, $p < .01$, $\eta_p^2 = .33$. For those assigned to Task Order 1 (neutral/comfort), there was no significant difference between Task Type on negative PANAS scores, $F(1, 14) = .71$, $p = .41$, $\eta_p^2 = .05$. The covariance adjusted mean negative PANAS score was $M = 13.6$, $SE = .45$ for the comfort task and $M = 14.1$, $SE = .59$ for the neutral task for those completing the study in Order 1. For those assigned to Task Order 2 (comfort/neutral), there was a significant difference between Task Type, $F(1, 14) = 11.57$, $p < .01$, $\eta_p^2 = .45$. For those assigned to Task Order 2, covariance adjusted mean negative PANAS score for the comfort task was $M = 16.4$, $SE = .83$, and for the neutral task was $M = 10.6$, $SE = .74$. (See Table E.3).

Order Effects on Self-reported Interpersonal Goals. A 2 (Order) x 2 (Task Type) mixed factors ANOVA was conducted to determine whether there were significant order effects on self-report measures of agentic and communal striving. Neither analysis revealed a significant main effect for Order or a significant Order by Task Type interaction. (See Table E.4).

Order Effects on Self-reported Task Appraisals. The rest of the post-task questionnaire items were analyzed individually with a series of 2 (Order) x 2 (Task Type) mixed factors ANOVAs, one for each item. (See Table E.5). Like prior analyses, only significant main effects for Order and Order by Task Type are reported.

For both questions regarding “how positive?” and “how negative?” the comfort and neutral tasks were rated, main effects for Order were revealed. However, in both cases, significant Order by Task Type interactions were also observed, so simple main effects were reported instead of main effects. Significant Order by Task Type interactions were observed for numerous task appraisal ratings, including: realistic, stressfulness, positive, negative, being upset, and task difficulty.

For ratings of being realistic, analysis of simple main effects showed no significant difference between Task Types for participants assigned to Order 1 (neutral/comfort), $F(1, 15) = .27, p = .61, \eta_p^2 = .02$, as realism ratings for the neutral task were $M = 3.1, SE = .24$, and for the comfort task were $M = 3.3, SE = .28$. In contrast, for those assigned to Order 2 (comfort/neutral), there was a significant difference between Task Types, $F(1, 15) = 10.00, p < .01, \eta_p^2 = .40$, such that realism ratings were higher for the neutral task, $M = 4.2, SE = .25$, as compared to the comfort task, $M = 3.2, SE = .25$.

For those who completed Task Order 1 (neutral/comfort), there was no significant difference between the two task types on ratings of stressfulness of *engaging* in the interaction,

$F(1, 15) = .055, p = .82, \eta_p^2 = .00$. Mean rating for the neutral task was $M = 2.7, SE = .34$ and mean rating for the comfort task was $M = 2.6, SE = .22$. For those who completed Task Order 2 (comfort/neutral), however, there was a significant difference on ratings of stressfulness between task types, $F(1, 15) = 23.44, p < .001, \eta_p^2 = .61$. Specifically, mean rating of stressfulness for the comfort task was $M = 3.0, SE = .33$, and mean rating for the neutral task was $M = 1.8, SE = .26$.

For ratings of positivity, post-hoc analyses showed no significant difference between Task Types for participants assigned to Order 1, $F(1, 15) = .06, p = .82, \eta_p^2 = .00$. Mean positivity rating for the neutral task was $M = 3.5, SE = .27$, and for the comfort task was $M = 3.6, SE = .22$. Conversely, for participants assigned to Order 2 (comfort/neutral), there was a significant difference between Task Types, $F(1, 15) = 12.74, p < .01, \eta_p^2 = .46$, such that ratings of positivity were higher for the neutral task, $M = 4.6, SE = .18$, as compared to the comfort task, $M = 3.8, SE = .25$.

Similarly, post-hoc analyses on ratings of negativity showed no significant difference between Task Types for those assigned to Order 1, $F(1, 15) = .00, p = 1.00, \eta_p^2 = .00$. Mean negativity rating for the neutral task was $M = 2.0, SE = .24$, and for the comfort task was $M = 2.0, SE = .20$. For those assigned to Order 2 (comfort/neutral), there was a significant difference between Task Types, $F(1, 15) = 6.82, p < .05, \eta_p^2 = .31$, such that negativity ratings were higher for the comfort task, $M = 1.9, SE = .26$, as compared to the neutral task, $M = 1.3, SE = .14$.

On ratings of being upset, simple effects analyses showed no significant difference between tasks for those in Order 1 (neutral/comfort), $F(1, 15) = .38, p = .55, \eta_p^2 = .02$. The mean rating of being upset for the neutral task was $M = 2.0, SE = .32$, and for the comfort task was $M = 1.8, SE = .25$. For those in Order 2 (comfort/neutral), there was a significant difference between

Task Types, $F(1, 15) = 10.95, p < .01, \eta_p^2 = .42$, such that ratings of being upset were higher for the comfort task, $M = 2.6, SE = .34$, as compared to the neutral task, $M = 1.4, SE = .16$.

Finally, for those in Order 1 (neutral/comfort), there was no significant difference between Task Types on the analysis of task difficulty, $F(1, 15) = 1.36, p = .26, \eta_p^2 = .08$. Mean rating of task difficulty for the neutral task was $M = 2.5, SE = .29$, and for the comfort task was $M = 2.3, SE = .27$. However, for those in Order 2 (comfort/neutral), there was a significant difference between Task Types, $F(1, 15) = 12.10, p < .01, \eta_p^2 = .45$, such that ratings of difficulty were higher for the comfort task, $M = 3.1, SE = .38$, as compared to the neutral task, $M = 1.9, SE = .30$.

Order Effects on Behavioral Measures. Analysis of behavioral parameters yielded no significant main effects for Order or Order by Task Type interactions. (See Tables E.6 and E.7) However, there was a significant Order X Task Type interaction on number of words spoken, $F(1, 30) = 7.85, p < .01, \eta^2 = .207$. For those in Order 1 (neutral/comfort), there was a significant simple main effect of Task Type, $F(1, 15) = 17.52, p < .01, \eta^2 = .539$. Specifically, mean number of words spoken for the neutral task was $M = 389.5, SE = 20.95$, and mean number of words spoken for the comfort task was $M = 314.19, SE = 23.16$. For those in Task Order 2 (comfort/neutral), there was also a significant simple main effect of Task Type, $F(1, 15) = 153.77, p < .001, \eta^2 = .911$. Specifically, mean number of words spoken for the neutral task was $M = 402.8, SE = 16.70$, and mean number of words spoken for the comfort task was $M = 268.7, SE = 17.60$. More words were spoken during neutral tasks for participants assigned to both orders, but the difference was more dramatic for participants assigned to Order 2.

Summary of Order Effects. Across all cardiovascular, self-report, and behavioral variables, several order effects were observed, most qualified by task type. Rest period analyses

revealed that for those in Task Order 2, DBP was higher during the pre-neutral rest period as compared to the pre-comfort rest period, and HF-HRV was higher during the pre-comfort rest period than the pre-neutral rest period. Task period analyses revealed that those in Task Order 1 exhibited higher HR reactivity to the neutral task than those in Task Order 2. Self-reports of positive affect revealed that those who completed Task Order 2 reported higher positive affect than those in Task Order 1, and more negative affect for the comfort task than the neutral task. Those in Order 2 also reported higher realism, less stress, more positivity, less negativity, getting less upset, and less task difficulty for the neutral task as compared to the comfort task. However, as half of the Higher and Lower Empathy participants were assigned to each of the two task orders, Order was not included as a covariate in the primary analyses.

Table E.1. Summary Table for Task Order Effects on Resting Cardiovascular Measures: Task Type (Pre-Neutral, Pre-Comfort) x Task Order (1, 2) Repeated Measures ANOVAs

Source	SS	df	MS	F	p	η^2
SBP						
Task Order	90.25	1	90.25	.418	.523	.014
Error (Task Order)	6473.69	30	215.79			
Task Type	52.56	1	52.56	1.80	.190	.057
Task * Order	16.00	1	16.00	.548	.465	.018
Error (Task Type)	875.44	30	29.18			
DBP						
Task Order	91.84	1	91.84	.590	.448	.019
Error (Task Order)	4667.35	30	4667.35			
Task Type	40.11	1	40.11	1.70	.202	.054
Task * Order	112.01	1	112.01	4.75	.037	.137
Error (Task Type)	707.77	30	23.59			
HR						
Task Order	413.95	1	413.95	1.82	.188	.057
Error (Task Order)	6834.94	30	227.83			
Task Type	1.78	1	1.78	.383	.541	.013
Task * Order	.049	1	.049	.010	.919	.000
Error (Task Type)	139.86	30	4.66			
LF-HRV						
Task Order	.964	1	.964	.638	.431	.021
Error (Task Order)	45.30	30	1.51			
Task Type	.019	1	.019	.170	.683	.006
Task * Order	.014	1	.014	.122	.730	.004
Error (Task Type)	3.43	30	.114			
HF-HRV						
Task Order	1.79	1	1.79	.711	.406	.023
Error (Task Order)	75.50	30	2.52			
Task Type	.110	1	.110	1.17	.289	.037
Task * Order	.442	1	.442	4.67	.039	.135
Error (Task Type)	2.84	30	.095			
RMSSD						
Task Order	.106	1	.106	.732	.399	.024
Error (Task Order)	4.32	30	.144			
Task Type	.002	1	.002	.492	.488	.016
Task * Order	.005	1	.005	1.68	.205	.053
Error (Task Type)	.093	30	.003			

Table E.2. Summary Table for Task Order Effects on Task Cardiovascular Measures: Task Type (Neutral, Comfort) x Task Order (1, 2) Repeated Measures ANOVAs

Source	SS	df	MS	F	p	η^2
SBP						
Task Order	21.18	1	21.18	.11	.741	.004
Error (Task Order)	5500.15	29	189.66			
Task Type	110.06	1	110.06	2.36	.136	.075
Task * Order	141.10	1	141.10	3.02	.093	.094
Error (Task Type)	1354.44	29	46.70			
DBP						
Task Order	67.16	1	67.16	.37	.546	.013
Error (Task Order)	5205.85	29	179.51			
Task Type	2.78	1	2.78	.04	.852	.001
Task * Order	86.71	1	86.71	1.11	.300	.037
Error (Task Type)	2261.23	29	77.97			
HR						
Task Order	47.66	1	47.66	.73	.400	.025
Error (Task Order)	1898.02	29	65.45			
Task Type	119.16	1	119.16	12.56	.001	.302
Task * Order	169.31	1	169.31	17.85	.000	.381
Error (Task Type)	275.13	29	9.49			
LF-HRV						
Task Order	.34	1	.34	.44	.511	.015
Error (Task Order)	22.31	29	.77			
Task Type	.25	1	.25	1.11	.301	.037
Task * Order	.06	1	.06	.27	.604	.009
Error (Task Type)	6.42	29	.22			
HF-HRV						
Task Order	.00	1	.00	.00	.951	.000
Error (Task Order)	13.74	29	.47			
Task Type	.00	1	.00	.00	.951	.000
Task * Order	.04	1	.04	.15	.700	.005
Error (Task Type)	7.82	29	.27			
RMSSD						
Task Order	.01	1	.01	.41	.529	.014
Error (Task Order)	.59	29	.02			
Task Type	.01	1	.01	1.34	.256	.044
Task * Order	.00	1	.00	.76	.390	.026
Error (Task Type)	.16	29	.01			

Table E.3. Summary ANOVA Table for PANAS Score Comparisons between Task Order: Task Type (Neutral, Comfort) x Task Order (1, 2) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
PANAS Positive Total						
Task Order	73.82	1	73.82	5.06	.032	.149
Error (Task Order)	423.16	29	14.59			
Task Type	129.68	1	129.68	11.32	.002	.281
Task * Order	22.85	1	22.85	2.00	.168	.064
Error (Task*Order)	332.13	29	11.45			
PANAS Negative Total						
Task Order	11.08	1	11.08	2.15	.153	.069
Error (Task Order)	149.12	29	5.14			
Task Type	34.49	1	34.49	10.12	.003	.259
Task * Order	49.50	1	49.50	14.53	.001	.334
Error (Task*Order)	98.82	29	3.41			

Note. p-values for significant effects are bolded

Table E.4. Summary ANOVA Table for Post-task Agentic and Communal Score Comparisons between Task Order: Task Type (Neutral, Comfort) x Task Order (1, 2) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
PTQ Agentic						
Task Order	7.56	1	7.56	1.10	.304	.035
Error (Task Order)	207.19	30	6.91			
Task Type	100.00	1	100.00	17.97	.000	.375
Task * Order	.063	1	.063	.011	.916	.000
Error (Task Type)	166.94	30	5.57			
PTQ Communal						
Task Order	8.27	1	8.27	.967	.333	.031
Error (Task Order)	256.34	30	8.55			
Task Type	83.27	1	83.27	26.34	.000	.467
Task * Order	.391	1	.391	.124	.728	.004
Error (Task Type)	94.84	30	3.16			

Note. p-values for significant effects are bolded

Table E.5. Summary ANOVA Table for Post-task Individual Item Comparisons between Task Order: Task Type (Neutral, Comfort) x Task Order (1, 2) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
How realistic?						
Task Order	4.00	1	4.00	2.81	.104	.086
Error (Task Order)	42.75	30	1.43			
Task Type	3.06	1	3.06	4.87	.035	.140
Task * Order	5.06	1	5.06	8.05	.008	.211
Error (Task Type)	18.88	30	.629			
How likely to encounter?						
Task Order	.016	1	.016	.013	.909	.000
Error (Task Order)	35.47	30	1.18			
Task Type	5.64	1	5.64	5.92	.021	.165
Task * Order	1.27	1	1.27	1.33	.258	.042
Error (Task Type)	28.60	30	.953			
How stressful to handle?						
Task Order	5.06	1	5.06	3.71	.064	.110
Error (Task Order)	40.94	30	1.37			
Task Type	10.56	1	10.56	18.17	.000	.377
Task * Order	.000	1	.000	.000	1.00	.000
Error (Task Type)	17.44	30	.581			
How stressful to engage?						
Task Order	1.00	1	1.00	.455	.505	.015
Error (Task Order)	65.94	30	2.20			
Task Type	5.06	1	5.06	9.68	.004	.244
Task * Order	6.25	1	6.25	11.95	.002	.285
Error (Task Type)	15.69	30	.523			
How meaningful?						
Task Order	3.52	1	3.52	2.86	.101	.087
Error (Task Order)	36.84	30	1.23			
Task Type	.016	1	.016	.032	.859	.001
Task * Order	1.89	1	1.89	3.89	.058	.115
Error (Task Type)	14.59	30	.486			

How positive?

Task Order	6.25	1	6.25	4.91	.034	.141
Error (Task Order)	38.19	30	1.27			
Task Type	2.25	1	2.25	4.60	.040	.133
Task * Order	3.06	1	3.06	6.26	.018	.173
Error (Task Type)	14.69	30	.490			

How negative?

Task Order	3.06	1	3.06	2.71	.110	.083
Error (Task Order)	33.88	30	1.13			
Task Type	1.56	1	1.56	4.31	.047	.126
Task * Order	1.56	1	1.56	4.31	.047	.126
Error (Task Type)	10.88	30	.363			

How upset?

Task Order	.141	1	.141	.089	.767	.003
Error (Task Order)	47.22	30	1.57			
Task Type	3.52	1	3.52	4.20	.049	.123
Task * Order	6.89	1	6.89	8.24	.007	.215
Error (Task Type)	25.09	30	.836			

How difficult?

Task Order	.250	1	.250	.105	.748	.003
Error (Task Order)	71.50	30	2.38			
Task Type	4.00	1	4.00	5.71	.023	.160
Task * Order	9.00	1	9.00	12.86	.001	.300
Error (Task Type)	21.00	30	.700			

Note. p-values for significant effects are bolded

Table E.6. Summary ANOVA Table for Behavioral Parameters: Task Type (Neutral, Comfort) x Task Order (1, 2) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
Information Support						
Task Order	7.56	1	7.56	.512	.480	.017
Error (Task Order)	443.44	30	14.78			
Task Type	30.25	1	30.25	4.66	.039	.134
Task * Task Order	18.06	1	18.06	2.78	.106	.085
Error (Task Type)	194.69	30	6.49			
Emotional Support						
Task Order	3.06	1	3.06	.301	.587	.010
Error (Task Order)	304.94	30	10.17			
Task Type	49.00	1	49.00	11.18	.002	.272
Task * Task Order	1.56	1	1.56	.357	.555	.012
Error (Task Type)	131.44	30	4.38			
Esteem Support						
Task Order	.007	1	.007	.135	.716	.004
Error (Task Order)	1.49	30	.050			
Task Type	2.42	1	2.42	69.80	.000	.699
Task * Task Order	.019	1	.019	.555	.462	.018
Error (Task Type)	1.04	30	.035			
Social Network Support						
Task Order	.039	1	.039	.844	.366	.027
Error (Task Order)	1.39	30	.046			
Task Type	.020	1	.020	.507	.482	.017
Task * Task Order	.047	1	.047	1.19	.283	.038
Error (Task Type)	1.17	30	.039			
Attentiveness						
Task Order	43.89	1	43.89	1.97	.171	.062
Error (Task Order)	668.72	30	22.29			
Task Type	3.52	1	3.52	.576	.454	.019
Task * Task Order	.766	1	.766	.125	.726	.004
Error (Task Type)	183.22	30	6.11			
Positive Non-Verbal						
Task Order	10.56	1	10.56	.392	.536	.013
Error (Task Order)	809.38	30	26.98			
Task Type	495.06	1	495.06	50.45	.000	.627
Task * Task Order	7.56	1	7.56	.771	.387	.025
Error (Task Type)	294.38	30	9.81			

Table E.7 Summary ANOVA Table for Individual Behavior Codes: Task Type (Neutral, Comfort) x Task Order (1, 2) Mixed Factor ANOVAs.

	SS	df	MS	F	p	η^2
Not Tracking						
Task Order	.738	1	.738	2.94	.097	.089
Error (Task Order)	7.54	30	.251			
Task Type	.006	1	.006	.070	.793	.002
Task * Task Order	.031	1	.031	.354	.556	.012
Error (Task Type)	2.61	30	.087			
Normative Verbal						
Task Order	13.14	1	13.14	3.05	.091	.092
Error (Task Order)	129.47	30	4.32			
Task Type	11.39	1	11.39	5.01	.033	.143
Task * Task Order	1.89	1	1.89	.831	.369	.027
Error (Task Type)	68.22	30	2.27			
Normative Non-Verbal						
Task Order	.563	1	.563	.035	.854	.001
Error (Task Order)	487.38	30	16.25			
Task Type	6.25	1	6.25	2.35	.136	.073
Task * Task Order	4.00	1	4.00	1.51	.229	.048
Error (Task Type)	79.75	30	2.66			

Note. p-values for significant effects are bolded

Figure E.1. Heart Rate (HR) Reactivity by Task Order

