



Published in final edited form as:

Psychophysiology. 2019 May ; 56(5): e13329. doi:10.1111/psyp.13329.

Respiratory Sinus Arrhythmia Reactivity across Empirically-Based Structural Categories of Psychopathology Among Adults: A Meta-Analysis

Theodore P. Beauchaine, Ziv Bell, Erin Knapton, Heather McDonough-Caplan, Tiffany Shader, and Aimee Zisner

The Ohio State University

Abstract

Low resting respiratory sinus arrhythmia (RSA) is observed in many mental health conditions, including anxiety disorders, mood disorders, schizophrenia spectrum disorders, disruptive behavior disorders, and nonsuicidal self-injury, among others. Findings for RSA reactivity are more mixed. We evaluate associations between RSA reactivity and empirically-derived structural categories of psychopathology—including internalizing, externalizing, and thought problems—among physically healthy adults. We searched multiple electronic databases for studies of RSA among participants who were assessed either dimensionally using well-validated measures or diagnostically using structured interviews. Strict inclusion criteria were used to screen 3,605 published reports, which yielded 37 studies including 2,347 participants and 76 effect sizes. We performed a meta-analysis, with meta-analytic regressions of potential moderators, including psychopathology subtypes. The sample-wide meta-analytic association between RSA reactivity and psychopathology was quite small, but heterogeneity was considerable. Moderation analyses revealed significant RSA reactivity (withdrawal) specifically in externalizing samples. Additional moderators included (a) stimulus conditions used to elicit RSA reactivity (only negative emotion inductions were effective), (b) sex (women showed greater RSA reactivity than men), and (c) adherence to established methodological guidelines (e.g., higher electrocardiographic sampling rates yielded greater RSA reactivity). These findings indicate that associations between RSA reactivity and psychopathology are complex, and suggest that future studies should include more standardized RSA assessments to increase external validity and decrease measurement error.

Keywords

RSA; HRV; reactivity; meta-analysis; psychopathology

1 INTRODUCTION

Respiratory sinus arrhythmia (RSA) is a periodic cardiorespiratory phenomenon characterized by heart rate acceleration during inhalation and heart rate deceleration during

Address correspondence and reprint requests to Theodore P. Beauchaine, Department of Psychology, The Ohio State University, 1835 Neil Avenue, Columbus, OH 43210; beauchaine.1@osu.edu.

No previous publications have appeared using these metadata.

exhalation. When appropriate quantification methods are used, RSA indexes parasympathetic nervous system (PNS)-linked cardiac activity (Berntson et al., 1997; Task Force of the European Society of Cardiology, 1996). RSA is of clinical interest given consistent associations with a variety of adverse physical and mental health outcomes (see Beauchaine & Thayer, 2015; Laborde, Mosley, & Thayer, 2017; Thayer & Lane, 2009). As outlined in several recent literature reviews and meta-analyses, low *resting* RSA is observed across transdiagnostic psychiatric conditions, including internalizing disorders (e.g., depression), externalizing disorders (e.g., conduct disorder), schizophrenia, autism spectrum disorder, and nonsuicidal self-injury, among others (e.g., Beauchaine, 2015a; Beauchaine, Gatzke-Kopp, & Mead, 2007; Chalmers, Quintana, Abbott, & Kemp, 2014; Clamor, Lincoln, Thayer, & Koenig, 2016; Koenig, Kemp, Feeling, Thayer, & Kaess, 2016; Neuhaus, Bernier, & Beauchaine, 2016; Shader et al., 2018). These transdiagnostic associations between resting RSA and psychopathology are found among children, adolescents, and adults (Beauchaine, 2001; Beauchaine & Thayer, 2015; Shader et al., 2018). Although some have attributed low RSA among depressed samples to antidepressant treatment including tricyclics and SSRIs (O'Regan, Kenny, Cronin, Finucane, & Kearney, 2015), unmedicated, physically healthy individuals with major depression also exhibit lower than normal RSA (Kemp, Quintana, Felmingham, Matthews, & Jelinek, 2012). Emerging evidence suggests that for at least some disorders, low resting RSA is state-dependent, and often improves between illness episodes and following effective psychosocial interventions (e.g., Bell, Shader, Webster-Stratton, Reid, & Beauchaine, 2018; Bylsma, Salomon, Taylor-Clift, Morris, & Rottenberg, 2014).

Given that low resting RSA is observed across diverse forms of psychopathology, obvious questions emerge over the precise vulnerability it marks among otherwise healthy individuals. Several authors have suggested that low RSA marks deficiencies in top-down self- and emotion-regulatory processes, which are compromised in many forms of psychopathology (Beauchaine, 2015b; Caspi et al., 2014; Porges, 2007; Thayer, Hansen, Saus-Rose, & Johnsen, 2009). Although full articulation of such views is beyond the scope of this meta-analysis, low resting RSA is observed across the lifespan among those with personality traits and clinical disorders characterized by self- and emotion dysregulation (see e.g., Beauchaine, 2015a, 2015b; Beauchaine & Thayer, 2015; Koenig, Kemp, Beauchaine, Thayer, & Kaess, 2016; Kuo, & Linehan, 2009; Shader et al., 2018). A handful of null outcomes notwithstanding (e.g., Graziano Derefinko, 2013), such findings are impressively consistent—especially in an era of frequent non-replication (see Lilienfeld, 2017; Tackett et al., 2018).

In contrast to resting RSA, associations between psychopathology and RSA *reactivity* to eliciting events are far from consistent (Balzarotti, Biassoni, Colombo, & Ciceri, 2017; Shader et al., 2018). Even when considering only internalizing disorders, some studies find excessive RSA reactivity (withdrawal) to emotion-eliciting stimuli (e.g., Crowell et al., 2005), others find blunted RSA reactivity (e.g., Bylsma et al., 2014), and still others find heterogeneity in RSA reactivity (e.g., Panaite et al., 2016). There are several possible explanations for such discrepancies. Foremost among these may be the diverse nature of laboratory tasks used, including negative emotion inductions, positive emotion inductions, attention allocation tasks, and social stressors, among others (see Fortunato, Gatzke-Kopp, &

Ram, 2013). Given that the PNS helps regulate cardiac output to facilitate immediate active coping and meet associated metabolic demands, negative emotion inductions (e.g., anger, threat) should yield larger effect sizes than many other lab tasks (Zisner & Beauchaine, 2016). Notably, however, RSA reactivity is often interpreted as a biomarker of emotion regulation or dysregulation regardless of stimulus conditions. Yet inferences about emotional states are best derived from inductions that are most likely evoke those states. For this reason, the NIMH National Advisory Mental Health Council Workgroup on Tasks and Measures for Research Domain Criteria (2016) recently recommended that standardized negative emotion inductions be used when assessing RSA reactivity in studies of arousal and regulatory processes implicated in psychopathology.

The Workgroup's (2016) recommendations point to several other possible sources of heterogeneity in studies of RSA reactivity in psychopathology. For example, baseline recording conditions often differ across studies. Some use stimulus-free baselines whereas others use baselines during which participants watch videos, listen to music, etc. (cf. Jennings, Kamarck, Stewart, Eddy, & Johnson, 1992). These "vanilla" baselines may themselves induce RSA reactivity through attention allocation mechanisms (Laborde et al. 2017). This affects estimates of RSA reactivity, which are computed as the difference between resting and task responding.

Careful baseline recording is but one aspect of well-established standards for RSA assessment. Guidelines for data collection, instrumentation precision, data reduction, and data analysis are outlined in two foundational papers, one from the Society for Psychophysiological Research (Berntson et al., 1997), and the other from the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Only a subset of published studies adheres to these standards, with possible implications for accuracy and precision of RSA estimation. Among other important guidelines, these documents provide recommendations for minimum sampling rates of electrocardiograph (ECG) equipment, minimum epoch lengths of ECG recordings, and preferred algorithms for quantifying RSA.

In addition, RSA reactivity-psychopathology relations may differ depending on whether symptoms are assessed dimensionally, or diagnostic cutoffs are used. For example, negative correlations between *resting* RSA and externalizing symptom dimensions are small and non-significant, even though those with clinical levels of externalizing behavior exhibit significantly lower resting RSA than age-matched, healthy controls (Beauchaine et al., 2007; Shader et al., 2018). Because analyses of dimensional scores carry more statistical power than analyses of dichotomous diagnoses, such findings run counter to expectation, and may suggest stronger physiology-behavior mechanisms at externalizing extremes. To date, few such comparisons have been made for RSA reactivity.

Finally, the literature on sex differences in RSA reactivity is mixed. Some studies report greater RSA reactivity among women (e.g., Yaroslavsky, Rottenberg, & Kovacs, 2013), some report greater RSA reactivity among men (e.g. Jönsson, & Sonnby-Borgström, 2003), and still others report no sex differences (e.g., Hamilton, & Alloy, 2016). Notably, however, many studies in the psychopathology literature do not evaluate sex effects, and many are

underpowered statistically for doing so. Meta-analysis presents an opportunity to assess sex effects.

Following from this discussion, our objectives in conducting this meta-analysis were to evaluate the strength of association between RSA reactivity and vulnerability to psychopathology, and to assess likely moderators of heterogeneity in RSA reactivity-psychopathology relations. We assessed several specific moderators. First, we evaluated RSA reactivity across broadband, empirically-derived psychopathology categories—including internalizing, externalizing, and thought problems—consistent with the well-replicated latent structure of psychopathology in which these three latent dimensions invariably emerge (e.g., Beauchaine & Zisner, 2017; Caspi et al., 2014; Kotov et al., 2017). Despite an adequate number of studies for such a meta-analysis (see below), too few studies were available to assess more specific psychopathology diagnoses.

In addition, we evaluated possible moderating effects of stimulus conditions (negative emotion induction vs. other), likely adherence to European Task Force (1996) data collection and analysis standards, dimensional vs. clinical assessment, and sex. As detailed below, studies were excluded if they (1) included speech tasks, which interfere with periodicity of the cardio-respiratory cycle, thereby violating the stationarity assumption and artificially attenuating RSA; (2) used low ECG sampling rates by modern standards; and/or (3) required participant movement, which induces metabolic demands and confounding reductions in RSA (see Zisner & Beauchaine, 2016).

2 METHOD

2.1 Literature search

Systematic literature searches were conducted in PyscINFO, PubMed, Web of Science, and OpenGrey. Search terms were designed to be comprehensive and crossed terms related to RSA reactivity (*heart rate variability, HRV, respiratory sinus arrhythmia, RSA withdrawal, RSA suppression, RSA reactivity, vagal reactivity, vagal regulation, vagal suppression, vagal control, vagal tone, vagal withdrawal, vagal modulation, cardiac autonomic regulation, physiological regulation, parasympathetic flexibility*) AND psychopathology (*psychopatholog*, externalizing, aggress*, delinquen*, ADHD, attention deficit hyperactivity, oppositional, conduct, substance use, substance abuse, substance dependen*, internalizing, anxi*, depress*, MDD, panic, GAD, generalized anxiety, fear, phobi*, PTSD, posttraumatic stress, OCD, obsessive-compulsive, autis*, schizophreni*, bipolar, mani*, personality disorder*). We searched English-language article abstracts of research with humans. References from included studies were surveyed. Given the scope of the search, dissertations were excluded.

2.2 Inclusion and exclusion criteria

Eligible studies were required to measure psychopathology and RSA reactivity to a specific task, compared with a baseline recording. Measures of psychopathology were required to be well-established and validated for approximating/characterizing *DSM-III-R, DSM-IV*, and *DSM-5* disorders (American Psychiatric Association, 1987, 1994, 2013). We required RSA

to be measured using ECG sampling rates greater 512Hz. Quantification methods included spectral analysis (fast-Fourier transform or auto-regressive), RMSSD, or peak-valley (see Berntson et al., 1997; Grossman, Beek, & Wientjes, 1990; Shader et al., 2018). Studies were excluded if participants with psychopathology were screened out (which would leave only healthy controls), had medical conditions that affect RSA (obesity, cardiovascular disease), or were younger than age 18 years. Participants with eating disorders were also excluded, given effects of excessively low and high body mass on RSA (Fraleigh, Birch, Senkottaiyan, & Alpert, 2005; Galetta et al., 2003). Although not all eating disorders are defined by altered BMI, all eating-disordered samples were excluded to avoid potential confounds. Tasks were included if they used a specific stimulus that (1) followed a baseline recording, (2) did not require participant movement or vocalization, and (3) did not involve paced breathing or relaxation. All baseline periods were at least 1 minute—a minimum duration for reliable RSA quantification using spectral analysis according to 1996 Task Force guidelines (see also Berntson et al., 1997).

Our initial search returned 4,505 results. After removing duplicates, 3,605 abstracts were screened jointly by several raters (ZB, TS, AZ, HM-C, EK). Among these, 2,587 articles were excluded (see Figure 1). The remaining 1,018 were reviewed in full-text by ZB. Among these, 981 were excluded. An independent rater (EK) reviewed the same 1,018 articles. Disagreements, which were few, were resolved through discussion. The final set of articles included 37 studies with 2,347 participants and 76 effect sizes. Among these, 36 effect sizes were comparisons of RSA reactivity between healthy controls and those with clinical levels of psychopathology, and 42 were correlations between RSA reactivity and continuous measures of psychopathology.

2.3 Quantifying RSA reactivity

As is customary in the psychophysiology literature, RSA reactivity was indexed as task minus baseline. Negative scores therefore indicate reduced RSA. In many instances, authors computed RSA as baseline minus task, in which case we converted. Hedges' g was calculated across studies that included a clinical and control group, and r was calculated for studies with a single sample in which psychopathology was measured continuously. Other statistics (e.g., F , t) were converted where possible. Hedges' g was selected over Cohen's d given the number of studies included in final analyses (Hedges & Olkin, 1985).

2.4 Moderators

2.4.1 Overview—Primary moderators included (1) broadband psychopathology subtype (internalizing, externalizing, thought problems), (2) task type (negative emotion induction vs. other), (3) documented adherence to European Task Force standards (yes, no), (4) nature of the sample (clinical-control, at-risk, community) and (5) sex (percentage of males). Additional moderators, which are unrelated to study hypotheses, included (6) mean age of clinical and control samples (when available), and (7) race (percentage of Caucasians). A summary of included studies appears in Table 1. A full reference list is provided in the online supplement. Clinical samples included participants who met criteria for *DSM-III-R*, *DSM-IV*, or *DSM-5* diagnoses, or exceeded clinical cutoffs on validated rating scales. At-risk samples included participants who were vulnerable to psychopathology, but did not

meet diagnostic criteria or exceed clinical cutoffs. Examples include samples comprised of relatives of diagnosed individuals, samples comprised of those who were previously diagnosed with psychiatric disorders but were currently in remission, and samples exposed to significant trauma (e.g., war, natural disaster) without meeting full criteria for PTSD. Community participants included those recruited from the general population or college samples. Assessment of psychopathology was categorized by diagnostic interview vs. self-report rating scale. Articles with multiple assessments yielded multiple effect sizes when sufficient statistics were reported. Diagnostic interviews were coded as structured (e.g., Mini International Neuropsychiatric Interview; Lecrubier et al., 1997), semi-structured (e.g., Structured Clinical Interview for DSM; First, Spitzer, Gibbon, & Williams, 1997), or unstructured. Only rating scales with validated diagnostic specificity (e.g., PTSD Checklist; Weathers, Litz, Herman, Huska, & Keane, 1994) or clinical cut-off scores (e.g., Beck Depression Inventory-II; Beck, Steer, & Brown, 1996) were allowed. Trait measures were not included.

2.4.2 Psychopathology subtypes—For diagnostic samples, specific psychopathology subtypes were sometimes unclear. Given our objectives of comparing broadband structural categories of psychopathology, this could usually be resolved. For example, some articles assessed clinical levels of anxiety, but did not indicate a specific disorder. These samples were categorized as internalizing (i.e., anxiety and unipolar depressive disorders). Externalizing disorders included ADHD, substance use disorders, antisocial personality, and borderline personality disorder. Thought disorders included bipolar, schizophrenia spectrum, and unspecified psychotic disorders, consistent with the well-replicated latent structure of psychopathology derived from bifactor models (see Beauchaine & Thayer, 2015; Caspi et al., 2014; Kotov et al., 2017).

Antisocial and borderline personality disorders were included in the externalizing category because (1) symptoms of both disorders load on the externalizing dimension in hierarchical structural models of psychopathology, (2) males with ASPD and females with BPD score very high on externalizing measures, (3) the disorders share genetic liability and affected individuals are often reared in the same families, and (4) males with ASPD and females with BPD report similar extra-familial environmental risk exposures including affiliations with delinquent peers (see Beauchaine, in press; Crowell, Beauchaine, & Linehan, 2009; Goldman, D'Angelo, & DeMaso, 1993). Comorbid diagnoses were noted, yet very few studies reported percentages of participants with comorbid psychopathology. Studies that specified exclusion of participants for medications that alter RSA were identified as medication free. All others, including those with no mention of medication, were identified as allowing medications. Some studies specified lengths of time prior to physiological recordings for which participants were alcohol, caffeine, nicotine, and/or other medication free. These times were recorded in hours, when reported.

2.4.3 Baseline and task conditions—Length of baselines in seconds was recorded. In many cases, recordings were from the end of a longer baseline. We also coded whether the baseline was stimulus-free or minimally attention-demanding (e.g., neutral video, recurring instructions). When baselines were at least 10 min without attention allocation, they were

coded as stimulus-free (cf. Jennings et al., 1992). Tasks were coded as negative emotion inductions if they evoked sadness, fear, anxiety, anger, embarrassment, or frustration. This included watching film clips, listening to trauma scripts, rumination/worry inductions, and preparation for the Trier Social Stress Task (Kirschbaum, Pirke, & Hellhammer, 1993). Attention tasks included continuous performance tasks, oddball tasks, cognitive, and non-emotional tasks such as the Wisconsin Card Sort (Heaton, 1981). Positive emotion tasks included positively valenced images, positive film clips, and pleasant memory recall tasks. Neutral tasks included emotionally neutral film clips, imagery, and scripts. Thermoregulatory tasks, which were few but needed to be evaluated given their valence, involved cold pressors applied to the hand or forehead. Finally, we coded task length in seconds given how common habituation effects on RSA are in the psychophysiology literature (e.g., Brenner, Beauchaine, & Sylvers, 2005). For studies with multiple tasks (see Table 1), each was coded with a dependent effect size.

2.4.4 Methodological rigor—Several approximations of rigor of data collection were coded. These included (1) ECG sampling rate; (2) high-frequency band used to define RSA (for studies that used frequency-domain quantification [FFT, AR]); (3) percentage missing data; and (4) whether 1996 Task Force guidelines were referenced. Although mere reference to Taskforce guidelines provides only a loose index of adherence to psychophysiological standards, failure to follow such guidelines is a significant problem in the psychopathology literature. Coding this variable captures some recognition by authors that psychophysiological standards exist and should be followed.

2.5 Analyses

Meta-analyses, meta-regressions of putative moderators, and tests of heterogeneity were conducted following practices set forth in recently published meta-analyses (e.g., Graziano & Derefinko, 2013; Holzman & Bridgett, 2017; Koenig et al., 2016). Analyses were conducted in the Metafor package in *R* (Viechtbauer, 2010). Multivariate meta-analyses using random effects models were conducted to account for dependent effect sizes (Gleser & Olkin, 2009). For each multivariate meta-analysis, the QE statistic for quantifying residual heterogeneity was calculated. First, effect size estimates of differences in RSA reactivity between clinical and control groups included in the same study were meta-analyzed ($n = 17$). Second, effect size estimates of correlations between RSA reactivity and psychopathology were estimated among all studies, including clinical groups included in the first meta-analysis ($N = 37$). Finally, to confirm results from the second meta-analysis were not unduly influenced by clinical samples from studies that recruited both clinical and control groups, we estimated correlations between RSA reactivity and psychopathology among studies that included no control group ($n = 20$).

3 RESULTS

3.1 Main effects of psychopathology

As described immediately below, several meta-analytic effects were found. A file drawer analysis revealed that the fail-safe N (Rosenthal, 1979) would require 112 additional studies with an average effect size of zero to reduce the overall significance level of this meta-

analysis to $p > .05$. Among studies that included clinical and control groups, there was no group difference in RSA reactivity, Hedges' $g = .09$, 95% CI $(-.16, .33)$. Figure 2 displays the forest plot of 36 effect sizes from these 17 studies. Next, all studies were meta-analyzed together. Among 76 effect sizes from 37 studies, RSA withdrawal (task minus baseline) was associated with higher psychopathology scores (Figure 3), $r = -.06$, 95% CI $(-.12, -.001)$. Among 40 effect sizes from 20 studies that assessed RSA reactivity within psychopathological samples, there was no association between RSA reactivity and symptoms (Figure 4) $r = -.07$, 95% CI $(-.15, .01)$. Of note, substantial heterogeneity was observed among all studies, $Q(75) = 177.33$, $p < .001$, among studies with clinical and control groups, $Q(35) = 122.87$, $p < .001$, and among community samples, $Q(39) = 72.91$, $p < .001$. These findings suggest the possibility of moderation.

3.2 Moderating effects

3.2.1 Psychopathology subtypes—Several patterns of moderation emerged. As shown in Figure 5, significant RSA withdrawal was observed only for externalizing, $r = -.06$, 95% CI $(-.11, -.02)$. Effects for internalizing, $r = -.04$, 95% CI $(-.12, .03)$ and thought problems, $r = -.07$, 95% CI $(-.19, .05)$ were non-significant. In addition, as shown in Figure 6, significant RSA withdrawal was confined to negative emotion induction tasks, $r = -.12$, 95% CI $(-.21, -.08)$. Neutral, thermoregulatory, positive, and attention tasks all yielded non-significant changes in RSA, $r = -.02$ to $.19$, all 95% CIs include 0.

3.2.2 Other moderators—Sex and ECG sampling rate also moderated outcomes (see online supplement Figure 1S). Studies with higher proportions of women yielded greater RSA reactivity than samples with higher proportions of men, $QM(1) = 4.81$, $p = .03$, and higher ECG sampling rates were associated with greater RSA reactivity, $QM(1) = 5.42$, $p = .02$. Two other methodological variables approached significance in tests of moderation, including baseline conditions (stimulus-free yielded less RSA reactivity than vanilla), $QM(1) = 3.32$, $p = .08$ (see online supplement Figure 1S), and task length (longer tasks yielded less RSA reactivity), $QM(1) = 3.07$, $p = .08$. Although we apprise readers of these findings, we do not interpret them further given that they failed to reach traditional significance thresholds. Age, race, psychopathology assessment method, and medication status were all non-significant.

4 Discussion

4.1 Summary

To date, this is the first meta-analysis to assess associations between RSA reactivity across empirically-derived structural dimensions of psychopathology, including internalizing, externalizing, and thought problems. We took great care to exclude studies with known confounding effects on RSA quantification (e.g., speech, movement). Although this constrained the scope of our meta-analysis, we nevertheless included 37 studies, 2,347 participants, and 76 effect sizes. Across studies, task types, and methods, associations between psychopathology and RSA reactivity were small, and differences in RSA reactivity between clinical and control groups were non-significant. However, there was considerable heterogeneity in findings, and several important moderators emerged.

These findings suggest that at least some inconsistencies and failures to replicate in the literature on RSA reactivity as a marker of emotion dysregulation in psychopathology may be attributable to systematic differences in sample compositions, selection of stimulus conditions, and measurement precision (see Zisner & Beauchaine, 2016; Shader et al., 2018). Thus, consistent with recent calls to standardize assessment methods, as exemplified in the NIH Toolbox® (National Institutes of Health, 2017) and Behavioral Assessment Methods for RDoC Constructs (National Advisory Mental Health Council Workgroup on Tasks and Measures for Research Domain Criteria, 2016), the field may benefit from (1) using tasks that reliably elicit negative emotions, (2) invoking well-established methods for quantifying RSA as precisely as possible, and (3) recruiting more participants so sex effects can be analyzed. Only then can research provide stronger inferences about the validity of RSA reactivity as a peripheral marker of *in vivo* emotional processes (see Beauchaine, 2015b). In sections to follow, we discuss each moderator in turn, including more specific recommendations for future research.

4.2 Structural dimensions of psychopathology

The finding that RSA reactivity is moderated by structural dimensions of psychopathology may help to explain some longstanding discrepancies in the literature. RSA withdrawal was confined to studies of externalizing behaviors; no such associations were found for internalizing or thought problem dimensions. This contrasts with *resting* RSA, which, as reviewed above, is consistently low across structural dimensional of psychopathology. RSA reactivity to strong emotion inductions among those with externalizing disorders may reflect propensities toward fight/flight responding in certain environmental contexts (Beauchaine et al., 2007; Porges, 2007). Although a similar argument may apply to anxiety, we did not have adequate power to subdivide the internalizing spectrum into depressive vs. anxiety disorders. Authors of future studies may wish to do so. Regardless, our findings suggest that structural dimensions of psychopathology should be considered when planning studies to measure RSA reactivity.

4.3 Sex

Studies with higher proportions of female participants yielded greater RSA reactivity (withdrawal), consistent with some though not all previous research (e.g., Yaroslavsky et al., 2013). Notably, men are more likely to exhibit externalizing psychopathology, so this finding is somewhat perplexing given that externalizing was associated with greater RSA withdrawal. Such findings may suggest complex interactive effects of sex and structural dimensions of psychopathology on RSA reactivity. Future research may offer clarification. It should also be noted that many men exhibit lower resting RSA than women (Koenig & Thayer, 2016), which can create floor effects on reactivity. Hormonal effects and sex differences in brain function are also possible explanations (Du, Fang, & Kiriazis, 2006; Koenig & Thayer, 2016), but fall beyond the scope of this article.

4.4 Baseline conditions

Perhaps the most straightforward recommendation for future research is to standardize baseline assessments. This will almost certainly reduce literature-wide noise in both resting RSA and RSA reactivity. Type of baseline (stimulus-free vs. vanilla) moderated effects, and

1996 Taskforce Guidelines were often not followed. Many studies used baselines shorter than three minutes, some combined baseline periods of spontaneous and paced breathing, and others recorded baselines while participants completed questionnaires or engaged in other tasks. Noise associated with such variability may obscure real, consistent effects of psychopathology on RSA reactivity and/or produce spurious false positive results in some studies.

4.5 Stimulus conditions

Consistent with theoretical perspectives linking excessive RSA withdrawal specifically to negative and not positive emotional reactivity (e.g., Beauchaine & Thayer, 2015; Obradovi & Finch, 2017), only negative emotion inductions yielded significant phasic changes in RSA. In contrast, attention demanding tasks, which are often used in the literature, yielded a near zero meta-analytic effect, $r < -0.01$, 95% CI (-0.09, 0.08). This is consistent with previous reviews suggesting that attention-allocation and cognitive control tasks have unpredictable effects on RSA (see Beauchaine, 2001; Obradovi & Finch, 2017). Future evaluations comparing different types of attention tasks on RSA reactivity may be useful in disentangling this unpredictability (cf., Suess, Porges, & Plude, 1994).

Following from contemporary theory in psychopathology research, many authors construed RSA reactivity as a peripheral index of emotional reactivity, emotion regulation, or emotion dysregulation. Notably, however, tasks were often not designed to elicit strong emotions or emotion-regulatory mechanisms. As we note in the introduction, excessive RSA withdrawal may be specific to negative emotion inductions among those who suffer from emotion dysregulation (e.g., Beauchaine & Thayer, 2015; Beauchaine et al., 2007; Crowell et al., 2005; Kuo, & Linehan, 2009). Given a primary role of the PNS in modulating emotional responding (Beauchaine, 2001; Porges, 2007), tasks that elicit negative affect may be best suited for evaluating individual differences in (a) how readily emotions are evoked and (b) how long negative emotions persist, which are consistent problems across subtypes of psychopathology. Ordinarily, one would not evaluate attentional processes with an emotion-induction task, yet it remains common among psychopathologists to evaluate emotional processes with attention-induction (and other) tasks. Matching stimulus conditions more closely with the constructs we seek to evaluate will improve external validity. Although there is clear pressure for research to be novel (see Lilienfeld, 2017), using previously untested and un-validated tasks makes direct comparisons across studies difficult (National Advisory Mental Health Council Workgroup on Tasks and Measures for RDoC, 2016; NIH, 2017).

4.6 Methodological rigor

Relatedly, although National Advisory Mental Health Council Workgroup (2016) guidelines call for standardized assessments, there are currently no established tasks for RSA reactivity research. We found extensive variability in methods even after setting strict inclusion thresholds that allowed only 37 (1.03%) of the 3,605 studies that we screened into the meta-analysis. In most cases, confounds including vocal responding and movement rendered RSA uninterpretable. Moreover, differences in vocalization and movement as a function of psychopathology were rarely assessed or controlled. All of these issues affect the validity and precision of RSA assessment—more so for change scores than for baseline scores. This

may help explain why findings regarding RSA reactivity and psychopathology are so inconsistent. It is also worth noting that existing guidelines are most appropriate for resting RSA—not RSA reactivity.

ECG sampling rate also moderated outcomes. Even though we included only studies with sampling rates ≤ 512 Hz, sampling rates of 1024Hz yielded greater RSA reactivity. Of note, both European Task Force (1996) guidelines and Berntson et al. (1997) suggest that 256Hz (which we did not evaluate) may be adequate for RSA assessment under some circumstances. However, most modern psychophysiological systems sample at 1024Hz. Our findings suggest that 1024Hz sampling is preferred, and might be adopted as a recommended minimum.

4.7 Null findings

Several other factors, including age, race, and whether medications that affect RSA were allowed, did not moderate associations between RSA reactivity and psychopathology. Although resting RSA increases and RSA reactivity decreases across childhood and adolescence (Shader et al., 2018), both remain stable throughout most of adulthood among healthy individuals, before declining in late-middle to early-old age (e.g., De Meersman & Stein, 2007; Holzman & Bridgett, 2017). Given that we did not include children and that we screened samples with poor health out, lack of an age effect is unsurprising.

Several researchers have noted effects of psychotropic medications on RSA—particularly SSRIs (e.g., O'Regan et al., 2015). As noted above, however, low RSA-depression relations are observed over-and-above antidepressant effects (Kemp et al., 2012). Many of the studies we included did not control for medications, and many did not specify in enough detail what medications participants used. Many others did not mention medication status at all, in which case we presumed medications were allowed. Thus, medication status could not be coded with precision, so it is possible we did not detect effects. This may have added noise to our analyses and reduced effect sizes.

4.8 Reporting inconsistencies

Although not an assessed moderator, we noted considerable inconsistency in reporting. According to established guidelines (e.g., Berntson, 1997), RSA reactivity is calculated as task minus baseline. Thus, negative scores reflect RSA withdrawal. However, it was not uncommon for researchers report baseline minus task. Although it is straightforward to convert in such cases, which we did, many studies did not indicate which approach they used and instead referred simply to “reactivity” or “change”.

In addition, although some authors reported residualized change scores, simple difference scores were used most often to compute RSA reactivity. The primary advantage of residualized change scores is that they reduce additive unreliability of simple difference scores when imprecise measures are used (see Burt & Obradovi, 2013). Unreliability of change scores is especially problematic with certain types of measures, including rating scales, surveys, and repeated assessments collected across long time spans. In contrast, most psychophysiological measures, including RSA, are measured with excellent precision, and change is typically evaluated over short intervals. In such situations, residualized change

scores are unnecessary and raw change scores may be preferred (see Zisner & Beauchaine, 2016; Shader et al., 2018).

4.9 Limitations and future directions

The relatively limited number of studies available constrained our analyses in a number of ways. First, several newer methods for calculating RSA (e.g., Lyapunov exponents, approximate entropy, detrended fluctuation analysis) were too few in number to include (Acharya, Joseph, Kannathal, Lim, & Suri, 2006; Laborde et al., 2017). Notably, however, additional research on these methods is needed to clarify their validity, interpretation, and clinical utility.

In addition, nonlinear models of associations between RSA reactivity and psychopathology may be informative. As described above, RSA withdrawal facilitates adaptive emotional responses to threat (Beauchaine et al., 2007; Porges, 2007). Moderate withdrawal to negative emotion evocation may be adaptive, whereas extreme RSA reactivity may mark pathology. One method of evaluating nonlinearity involves comparisons between normative variation and variation at the extremes of continuous distributions. Such nonlinearities may apply to resting RSA across the externalizing spectrum (Shader et al., 2018). Another method of evaluating nonlinearity uses piecewise analyses that allow for changes in directions of reactivity across time and levels of traits (e.g., Obradovi & Finch, 2017). At present, too few such analyses exist for inclusion in meta-analysis.

We also focused on bivariate associations between RSA reactivity and psychopathology. Some studies have examined interactions between psychopathology and (a) resting RSA and RSA reactivity (Hinnant & El-Sheikh, 2013; Yaroslavsky et al., 2014), and (b) both RSA and measures of SNS responding (Giuliano, Gatzke-Kopp, Roos, & Skowron, 2017). Given few examples of such studies, we could not include them. Future research focused on interactive effects on psychopathology of RSA reactivity and its interactions with the SNS is warranted.

A literature also exists on psychopathology and recovery of RSA following task cessation. For example, timely recovery to baseline levels of RSA following negatively emotion induction is associated with effective emotion regulation (Rottenberg, Clift, Bolden, & Salomon, 2007). In contrast, prolonged RSA recovery is associated with poor functional outcomes (Gordon, Ditto, & D'Antono, 2012). Future reviews and/or meta-analyses of RSA recovery may therefore be useful when an adequate literature base evolves.

Given the current state of the literature and relatively limited number of studies, we were unable to evaluate interactions among moderators because of inadequate statistical power. For example, differential sex effects across latent dimensions of psychopathology could not be assessed, even though sex differences in both prevalence and presentation of internalizing and externalizing disorders are well known. Moreover, although medication status did not emerge as a significant moderator, different medications alter RSA in different ways, and such effects could be obscured. In the future, authors may wish to provide more detailed information about classes and doses of medications used.

In addition, longstanding debate exists in the literature regarding the need to control for respiration when computing RSA—an issue we have addressed in our own work (see Denver, Reed, & Porges, 2007; Grossman, Karemaker, & Wieling, 1991; Ritz, 2009; Shader et al., 2018). Although we cannot review this complex literature, we purposefully eliminated many studies, such as those that used vocalization tasks, because of their dramatic effects on respiration and therefore RSA. In addition, we included only studies of healthy adults, thereby reducing age-related respiratory variability (children breath much faster than adults) and pathophysiology (obesity alters respiration rate and tidal volume). Nevertheless, if one seeks to quantify vagal efference to the heart with precision, respiration should be controlled (Grossman et al., Ritz).

In considering respiration, two additional points are important. First, had we included only the few studies that covaried respiration rate or quantified tidal volume, no meta-analysis could have been conducted. Perhaps more importantly, RSA reactivity to negative emotion induction may be a valid biomarker of affect dysregulation and vulnerability to psychopathology *even if vagal efference is not captured precisely*. Put another way, RSA reactivity might have predictive validity (to emotion dysregulation or other aspects of psychiatric vulnerability) without perfect construct validity (to efferent vagal neural traffic). Although we understand readers of *Psychophysiology* are concerned with construct validity, results of our meta-analysis may nevertheless be useful in clarifying potential biomarkers of vulnerability to psychopathology.

Recently, it has also been suggested that corrections for cardiac chronotropy (heart period or heart rate) be used when computing RSA (de Geus, Gianaros, Brindle, Jennings, & Berntson, 2018). Although the rationale for this recommendation is also beyond the scope of this article, it stems from positive associations between heart period HRV metrics, including RSA. In fact, individual and group differences in RSA are driven in part by factors that are associated with heart rate, such as age and, as outlined above, respiratory tidal volume.

As also outlined above, our inclusion criteria were strict. Notably, omitting studies can have dramatic effects on outcomes of meta-analyses (Licht, Penninx, & de Geus, 2011). Inclusion criteria were selected to maximize validity and measurement precision. Unavoidably, some large, likely well-conducted studies were excluded due to particular aspects of tasks and/or quantification. In addition, physical fitness—which was not reported for a vast preponderance of studies—affects RSA strongly (Aubert, Seps, & Beckers, 2003). Although some studies specified BMI as an inclusion criterion, BMI provides only a rough approximation of fitness.

Finally, studies with relatively homogenous samples of specific psychiatric diagnoses were too few in number to yield valid meta-analyses of RSA reactivity for specific *DSM* diagnoses. For example, among 19 effect sizes for participants with clinical levels of anxiety, only 5 included a specific *DSM*-defined disorder.

As reviewed above, associations between RSA reactivity and psychopathology are inconsistent across the literature. Our findings may shed light on at least some inconsistencies. RSA withdrawal is most pronounced among externalizing samples, is

elicited best by negative emotion inductions, is more pronounced among women, and is best captured by adhering to methodological guidelines. We hope these findings are useful as researchers plan future experiments on peripheral correlates of emotional lability and dysregulation in psychopathology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

Research reported in this article was supported by Grant UH2DE025980 from the National Institute of Dental and Craniofacial Research, and by the National Institutes of Health Science of Behavior Change (SoBC) Common Fund.

REFERENCES

- Acharya UR, Joseph KP, Kannathal N, Lim CM, & Suri JS (2006). Heart rate variability: A review. *Medical and Biological Engineering and Computing*, 44, 1031–1051. 10.1007/s11517-006-0119-0 [PubMed: 17111118]
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA; Author.
- Aubert AE, Seps B, & Beckers F (2003). Heart rate variability in athletes. *Sports Medicine*, 33, 889–919. 10.2165/00007256-200333120-00003 [PubMed: 12974657]
- Balzarotti S, Biassoni F, Colombo B, & Ciceri MR (2017). Cardiac vagal control as a marker of emotion regulation in healthy adults: A review. *Biological Psychology*, 130, 54–66. 10.1016/j.biopsycho.2017.10.008 [PubMed: 29079304]
- Beauchaine TP (2001). Vagal tone, development, and Gray's motivational theory: Toward an integrated model of autonomic nervous system functioning in psychopathology. *Development and Psychopathology*, 13, 183–214. 10.1017/S0954579401002012 [PubMed: 11393643]
- Beauchaine TP (2015a). Respiratory sinus arrhythmia: A transdiagnostic biomarker of emotion dysregulation and psychopathology. *Current Opinion in Psychology*, 3, 43–47. 10.1016/j.copsyc.2015.01.017 [PubMed: 25866835]
- Beauchaine TP (2015b). Future directions in emotion dysregulation and youth psychopathology. *Journal of Clinical Child and Adolescent Psychology*, 44, 875–896. 10.1080/15374416.2015.1038827. [PubMed: 26016727]
- Beauchaine TP (in press). A developmental psychopathology perspective on the emergence of antisocial and borderline personality pathologies across the lifespan In Lejuez CW & Gratz KL (Eds.), *Handbook of personality disorders*. New York: Cambridge University Press.
- Beauchaine TP, Gatzke-Kopp LM, & Mead HK (2007). Polyvagal theory and developmental psychopathology: Emotion dysregulation and conduct problems from preschool to adolescence. *Biological Psychology*, 74, 174–184. 10.1016/j.biopsycho.2005.08.008 [PubMed: 17045726]
- Beauchaine TP, & Thayer JF (2015). Heart rate variability as a transdiagnostic biomarker of psychopathology. *International Journal of Psychophysiology*, 98, 338–350. 10.1016/j.ijpsycho.2015.08.004 [PubMed: 26272488]
- Beauchaine TP, & Zisner A (2017). Motivation, emotion regulation, and the latent structure of psychopathology: An integrative and convergent historical perspective. *International Journal of Psychophysiology*, 119, 108–118. 10.1016/j.ijpsycho.2016.12.014 [PubMed: 28057475]

- Beck AT, Steer RA, & Brown GK (1996). Beck Depression Inventory-II. San Antonio, TX: Psychological Corporation.
- Bell Z, Shader TM, Webster-Stratton C, Reid MJ, & Beauchaine TP (2018). Improvements in negative parenting mediate changes in children's autonomic responding following a preschool intervention for ADHD. *Clinical Psychological Science*, 6, 134–144. 10.1177/2167702617727559 [PubMed: 29545976]
- Berntson GG, Bigger JT, Eckberg DL, Grossman P, Kaufmann PG, Malik M, ... van der Molen M (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*, 34, 623–648. 10.1111/j.1469-8986.1997.tb02140.x [PubMed: 9401419]
- Brenner SL, Beauchaine TP, & Sylvers PD (2005). A comparison of psychophysiological and self-report measures of BAS and BIS activation. *Psychophysiology*, 42, 108–115. 10.1111/j.1469-8986.2005.00261.x [PubMed: 15720586]
- Burt KB, & Obradovi J (2013). The construct of psychophysiological reactivity: Statistical and psychometric issues. *Developmental Review*, 33, 29–57. 10.1016/j.dr.2012.10.002
- Bylsma LM, Salomon K, Taylor-Clift A, Morris BH, & Rottenberg J (2014). RSA reactivity in current and remitted major depressive disorder. *Psychosomatic Medicine*, 76, 66–73. 10.1097/PSY.000000000000019 [PubMed: 24367127]
- Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, ... Moffitt TE (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, 2, 119–137. 10.1177/2167702613497473 [PubMed: 25360393]
- Chalmers JA, Quintana DS, Abbott MJ, & Kemp AH (2014). Anxiety disorders are associated with reduced heart rate variability: A meta-analysis. *Frontiers in Psychiatry*, 5, 80 10.3389/fpsy.2014.00080 [PubMed: 25071612]
- Clamor A, Lincoln TM, Thayer JF, & Koenig J (2016). Resting vagal activity in schizophrenia: Meta-analysis of heart rate variability as a potential endophenotype. *British Journal of Psychiatry*, 208, 9–16. 10.1192/bjp.bp.114.160762 [PubMed: 26729841]
- Crowell SE, Beauchaine TP, & Linehan M (2009). A biosocial developmental model of borderline personality: Elaborating and extending Linehan's theory. *Psychological Bulletin*, 135, 495–510. 10.1037/a0015616 [PubMed: 19379027]
- Crowell SE, Beauchaine TP, McCauley E, Smith C, Stevens AL, & Sylvers P (2005). Psychological, autonomic, and serotonergic correlates of parasuicidal behavior in adolescent girls. *Development and Psychopathology*, 17, 1105–1127. <https://doi.org/10.1017/S0954579405050522> [PubMed: 16613433]
- de Geus EJC, Gianaros PJ, Brindle RC, Jennings JR, & Berntson GG (2018). Should heart rate variability be “corrected” for heart rate? Biological, quantitative, and interpretive considerations. *Psychophysiology*. Published ahead of print.
- De Meersman RE, & Stein PK (2007). Vagal modulation and aging. *Biological Psychology*, 74, 165–173. 10.1016/j.biopsycho.2006.04.008 [PubMed: 17045727]
- Denver JW, Reed SF, & Porges SW (2007). Methodological issues in the quantification of respiratory sinus arrhythmia. *Biological Psychology*, 74, 286–294. 10.1016/j.biopsycho.2005.09.00 [PubMed: 17067734]
- Du XJ, Fang L, & Kiriazis H (2006). Sex dimorphism in cardiac pathophysiology: Experimental findings, hormonal mechanisms, and molecular mechanisms. *Pharmacology and Therapeutics*, 111, 434–475. 10.1016/j.pharmthera.2005.10.016 [PubMed: 16439025]
- First MB, Spitzer RL, Gibbon M, & Williams JB (1997). *Structured Clinical Interview for DSM-IV*. Washington, DC: American Psychiatric Press.
- Fortunato CK, Gatzke-Kopp LM, & Ram N (2013). Associations between respiratory sinus arrhythmia reactivity and internalizing and externalizing symptoms are emotion specific. *Cognitive Affective and Behavioral Neuroscience*, 13, 238–251. 10.3758/s13415-012-0136-4.
- Fraleigh MA, Birchem JA, Senkottaiyan N, & Alpert MA (2005). Obesity and the electrocardiogram. *Obesity Reviews*, 6, 275–281. 10.1111/j.1467-789X.2005.00199.x [PubMed: 16246213]
- Galetta F, Franzoni F, Praticchizzo F, Rolla M, Santoro G, & Pentimone F (2003). Heart rate variability and left ventricular diastolic function in anorexia nervosa. *Journal of Adolescent Health*, 32, 416–421. 10.1016/S1054-139X(03)00048-X [PubMed: 12782452]

- Giuliano RJ, Gatzke-Kopp LM, Roos LE, & Skowron EA (2017) Resting sympathetic arousal moderates the association between parasympathetic reactivity and working memory performance in adults reporting high levels of life stress. *Psychophysiology*, 54, 1195–1208. 10.1111/psyp.12872. [PubMed: 28449242]
- Gleser LJ, & Olkin I (2009). Stochastically dependent effect sizes In Cooper H, Hedges LV, & Valentine JC (Eds.), *Handbook of research synthesis and meta-analysis* (2nd ed., pp. 357–376). New York: Russell Sage Foundation.
- Goldman SJ, D'Angelo EJ, & DeMaso DR (1993). Psychopathology in the families of children and adolescents with borderline personality disorder. *American Journal of Psychiatry*, 150, 1832–1835. 10.1176/ajp.150.12.1832 [PubMed: 8238638]
- Gordon JL, Ditto B, & D'Antono B (2012). Cognitive depressive symptoms associated with delayed heart rate recovery following interpersonal stress in healthy men and women. *Psychophysiology*, 49, 1082–1089. 10.1111/j.1469-8986.2012.01397.x [PubMed: 22725718]
- Graziano P, & Derefinko K (2013). Cardiac vagal regulation and children's adaptive functioning outcomes: A meta-analysis. *Biological Psychology*, 94, 22–37. 10.1016/j.biopsycho.2013.04.011 [PubMed: 23648264]
- Grossman P, Beek JV, & Wientjes C (1990). A comparison of three quantification methods for estimation of respiratory sinus arrhythmia. *Psychophysiology*, 27, 702–714. 10.1111/j.1469-8986.1990.tb03198.x [PubMed: 2100356]
- Grossman P, Karemaker J, & Wieling W (1991). Prediction of tonic parasympathetic cardiac control using respiratory sinus arrhythmia: The need for respiratory control. *Psychophysiology*, 28, 201–216. 10.1111/j.1469-8986.1991.tb00412.x [PubMed: 1946886]
- Hamilton JL, & Alloy LB (2016). Atypical reactivity of heart rate variability to stress and depression across development: Systematic review of the literature and directions for future research. *Clinical Psychology Review*, 50, 67–79. 10.1016/j.cpr.2016.09.003 [PubMed: 27697746]
- Heaton R (1981). *A manual for the Wisconsin Card Sorting Test*. Odessa, FL: Psychological Assessment Resources.
- Hedges L, & Olkin I (1985). *Statistical methods for meta-analysis*. London, UK: Academic Press.
- Hinnant JB, & El-Sheikh M (2013). Codevelopment of externalizing and internalizing symptoms in middle to late childhood: Sex, baseline respiratory sinus arrhythmia, and respiratory sinus arrhythmia reactivity as predictors. *Development and Psychopathology*, 25, 419–436. 10.1017/S0954579412001150 [PubMed: 23627954]
- Holzman JB, & Bridgett DJ (2017). Heart rate variability indices as bio-markers of top-down self-regulatory mechanisms: A meta-analytic review. *Neuroscience and Biobehavioral Reviews*, 74, 233–255. 10.1016/j.neubiorev.2016.12.032 [PubMed: 28057463]
- Jennings JR, Kamarck T, Stewart C, Eddy M, & Johnson P (1992). Alternate cardiovascular baseline assessment techniques: Vanilla or resting baseline. *Psychophysiology*, 29, 742–750. 10.1111/j.1469-8986.1992.tb02052.x [PubMed: 1461961]
- Jönsson P, & Sonny-Borgström M (2003). The effects of pictures of emotional faces on tonic and phasic autonomic cardiac control in women and men. *Biological Psychology*, 62, 157–173. 10.1016/S0301-0511(02)00114-X [PubMed: 12581690]
- Kemp AH, Quintana DS, Felmingham KL, Matthews S, & Jelinek HF (2012). Depression, comorbid anxiety disorders, and heart rate variability in physically healthy, unmedicated patients: Implications for cardiovascular risk. *PLoS One*, 7, e30777 10.1371/journal.pone.0030777 [PubMed: 22355326]
- Kirschbaum C, Pirke KM, & Hellhammer DH (1993). The 'Trier Social Stress Test'—a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, 28, 76–81. 10.1159/000119004 [PubMed: 8255414]
- Koenig J, Kemp A, Beauchaine TP, Thayer JF, & Kaess M (2016). Depression and resting state heart rate variability in children and adolescents: A systematic review and meta-analysis. *Clinical Psychology Review*, 46, 136–150. 10.1016/j.cpr.2016.04.013 [PubMed: 27185312]
- Koenig J, Kemp AH, Feeling NR, Thayer JF, & Kaess M (2016). Resting state vagal tone in borderline personality disorder: A meta-analysis. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 64, 18–26. 10.1016/j.pnpbp.2015.07.002 [PubMed: 26169575]

- Koenig J, & Thayer JF (2016). Sex differences in healthy human heart rate variability: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, 64, 288–310. 10.1016/j.neubiorev.2016.03.007 [PubMed: 26964804]
- Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, ...Eaton NR (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126, 454–477. 10.1037/abn0000258 [PubMed: 28333488]
- Kuo JR, & Linehan MM (2009). Disentangling emotion processes in borderline personality disorder: Physiological and self-reported assessment of biological vulnerability, baseline intensity, and reactivity to emotionally evocative stimuli. *Journal of Abnormal Psychology*, 118, 531–544. 10.1037/a0016392 [PubMed: 19685950]
- Laborde S, Mosley E, & Thayer JF (2017). Heart rate variability and cardiac vagal tone in psychophysiological research—Recommendations for experiment planning, data analysis, and data reporting. *Frontiers in Psychology*, 8, 213. 10.3389/fpsyg.2017.00213 [PubMed: 28265249]
- Leclercq Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Sheehan KH, ...Dunbar GC (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry*, 12, 224–231. 10.1016/S0924-9338(97)83296-8
- Licht CM, Penninx BW, & de Geus EJ (2011). To include or not to include? A response to the meta-analysis of heart rate variability and depression. *Biological Psychiatry*, 69, e1. 10.1016/j.biopsych.2010.06.034 [PubMed: 20926065]
- Lilienfeld SO (2017). Psychology's replication crisis and the grant culture: Righting the ship. *Perspectives on Psychological Science*, 12, 660–664. 10.1177/1745691616687745 [PubMed: 28727961]
- National Advisory Mental Health Council Workgroup on Tasks and Measures for Research Domain Criteria (2016). Behavioral assessment methods for RDoC constructs. Bethesda, MD: National Institute of Mental Health.
- National Institutes of Health (2017). NIH Toolbox for assessment of neurological and behavioral function. Downloaded on 5/4/2018 from <https://www.neuroscienceblueprint.nih.gov/factSheet/toolbox.htm>
- Neuhaus E, Bernier RA, & Beauchaine TP (2016). Children with autism show altered autonomic adaptation to novel and familiar social partners. *Autism Research*, 9, 579–591. 10.1002/aur.1543 [PubMed: 26305051]
- Obradovi J, & Finch JE (2017). Linking executive function skills and physiological challenge response: Piecewise growth curve modeling. *Developmental Science*. Published ahead of print. 10.1111/desc.12476
- O'Regan C, Kenny RA, Cronin H, Finucane C, & Kearney PM (2015). Antidepressants strongly influence the relationship between depression and heart rate variability: Findings from the Irish Longitudinal Study on Ageing. *Psychological Medicine*, 45, 623–636. 10.1017/S0033291714001767 [PubMed: 25075912]
- Panaite V, Hindash AC, Bylsma LM, Small BJ, Salomon K, & Rottenberg J (2016). Respiratory sinus arrhythmia reactivity to a sad film predicts depression symptom improvement and symptomatic trajectory. *International Journal of Psychophysiology*, 99, 108–113. 10.1016/j.ijpsycho.2015.12.002 [PubMed: 26681648]
- Porges SW (2007). The polyvagal perspective. *Biological Psychology*, 74, 116–143. 10.1016/j.biopsycho.2006.06.009 [PubMed: 17049418]
- Ritz T (2009). Studying noninvasive indices of vagal control: The need for respiratory control and the problem of target specificity. *Biological Psychology*, 80, 158–168. 10.1016/j.biopsycho.2008.08.003 [PubMed: 18775468]
- Rosenthal R (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86, 638–641. 10.1037/0033-2909.86.3.638
- Rottenberg J, Clift A, Bolden S, & Salomon K (2007). RSA fluctuation in major depressive disorder. *Psychophysiology*, 44, 450–458. 10.1111/j.1469-8986.2007.00509.x [PubMed: 17371497]

- Shader TM, Gatzke-Kopp LM, Crowell SE, Reid MJ, Thayer JF, Vasey MW, ...Beauchaine TP (2018). Quantifying respiratory sinus arrhythmia: Effects of misspecifying breathing frequencies across development. *Development and Psychopathology*, 30, 351–366. 10.1017/S0954579417000669 [PubMed: 28554343]
- Suess PE, Porges SW, & Plude DJ (1994). Cardiac vagal tone and sustained attention in school-age children. *Psychophysiology*, 31, 17–22. 10.1111/j.1469-8986.1994.tb01020.x [PubMed: 8146250]
- Tackett JL, Lilienfeld SO, Patrick C, Johnson S, Krueger R, Miller J, ... Shrout PE (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science*, 12, 742–756. 10.1177/1745691617690042 [PubMed: 28972844]
- Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*, 93, 1043–1065. 10.1161/01.CIR.93.5.1043 [PubMed: 8598068]
- Thayer JF, Hansen AL, Saus-Rose E, & Johnsen BH (2009). Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. *Annals of Behavioral Medicine*, 37, 141–153. 10.1007/s12160-009-9101-z [PubMed: 19424767]
- Thayer JF, & Lane RD (2009). Claude Bernard and the heart–brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience and Biobehavioral Reviews*, 33, 81–88. 10.1016/j.neubiorev.2008.08.004 [PubMed: 18771686]
- Viechtbauer W (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36, 1–48. 10.18637/jss.v036.i03
- Weathers FW, Litz BT, Herman D, Huska J, & Keane T (1994). *The PTSD checklist-civilian version (PCL-C)*. Boston, MA: National Center for PTSD.
- Yaroslavsky I, Rottenberg J, & Kovacs M (2013). The utility of combining RSA indices in depression prediction. *Journal of Abnormal Psychology*, 122, 314–321. 10.1037/a0032385. [PubMed: 23713496]
- Zisner A, & Beauchaine TP (2016). Psychophysiological methods and developmental psychopathology In Cicchetti D (Ed.), *Developmental psychopathology*. Vol. 2: Developmental neuroscience (3rd ed., pp. 832–884). Hoboken, NJ: Wiley.

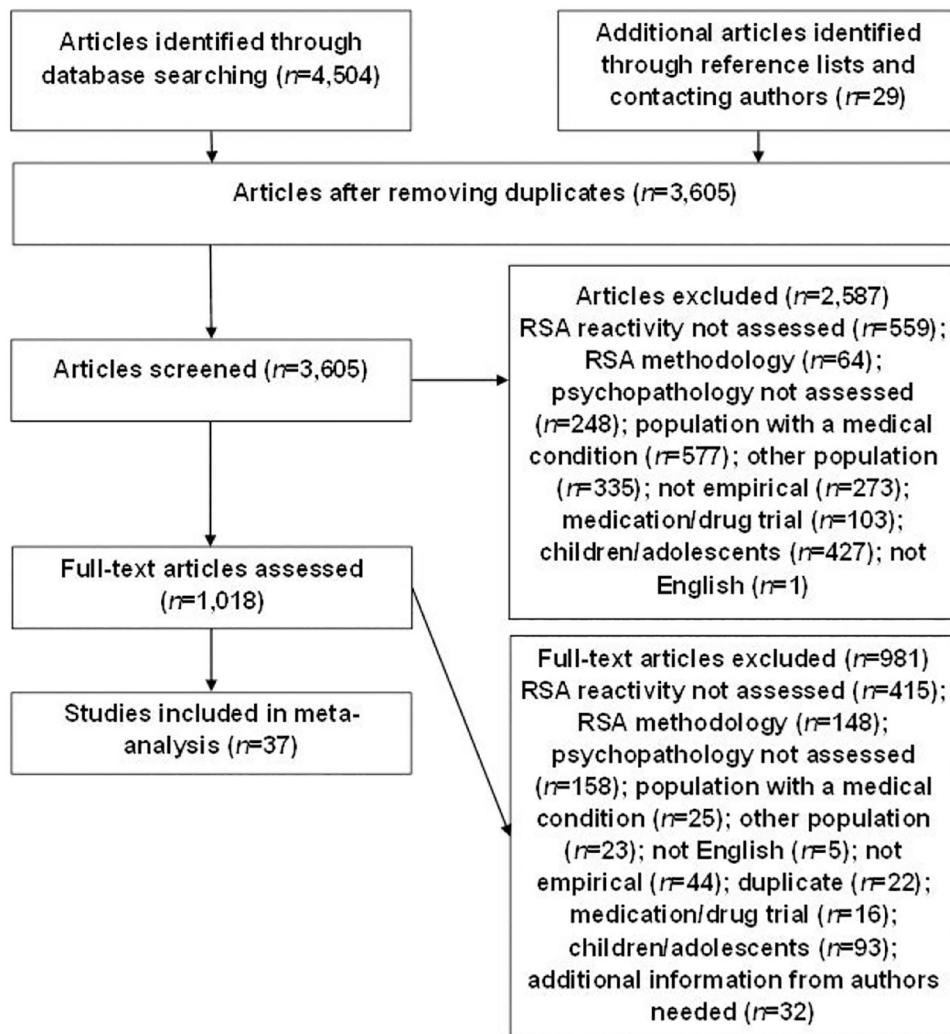


Figure 1.
PRISMA flowchart for study selection.

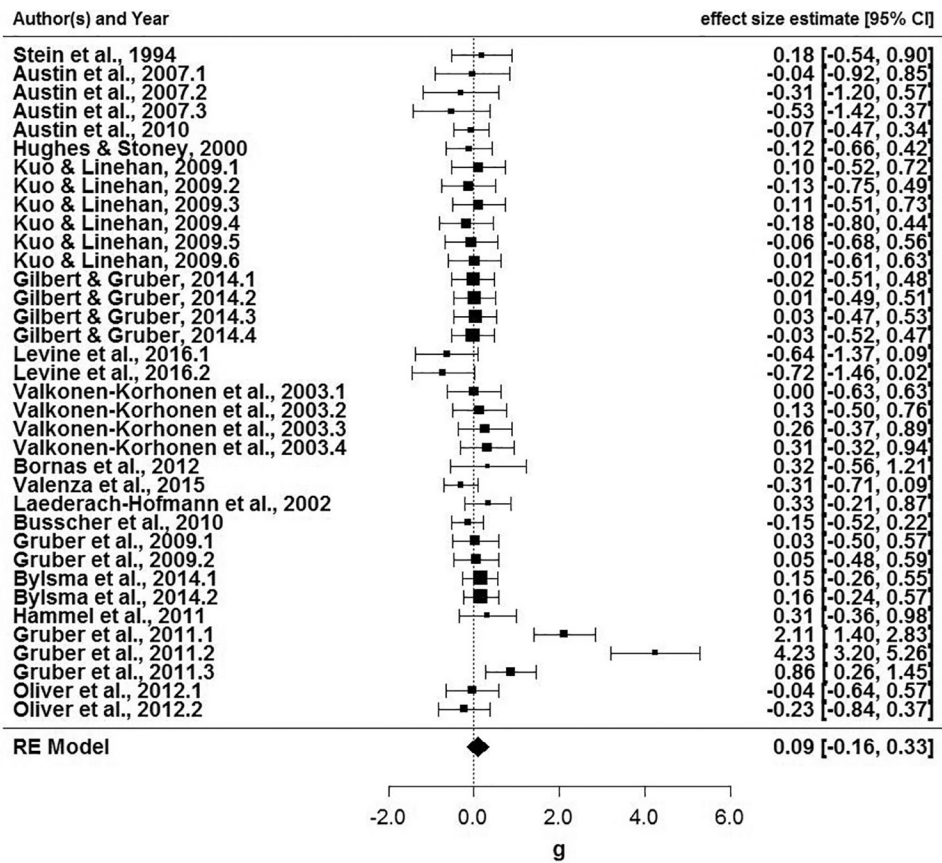


Figure 2. Forest plot for standardized mean differences (*d*) in RSA reactivity in studies that included clinical and control groups.

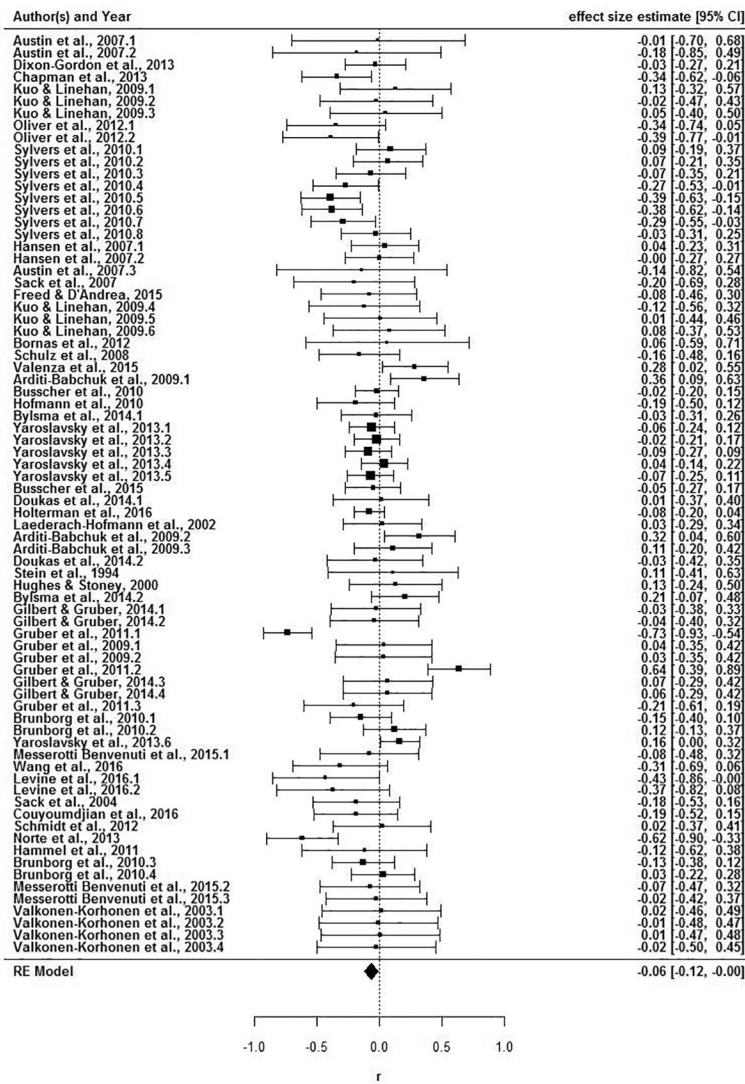


Figure 3. Forest plot of correlations between RSA reactivity and psychopathology (all studies).

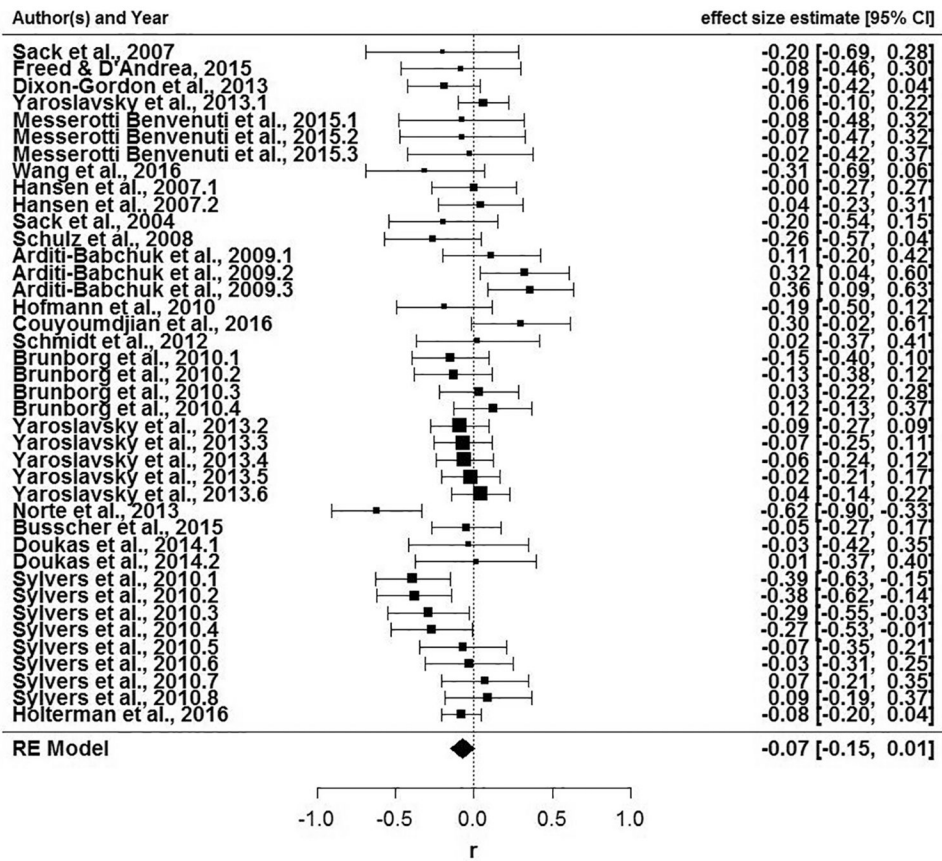


Figure 4. Forest plot of correlations (r) between RSA reactivity and symptoms in studies of psychopathology-only samples.

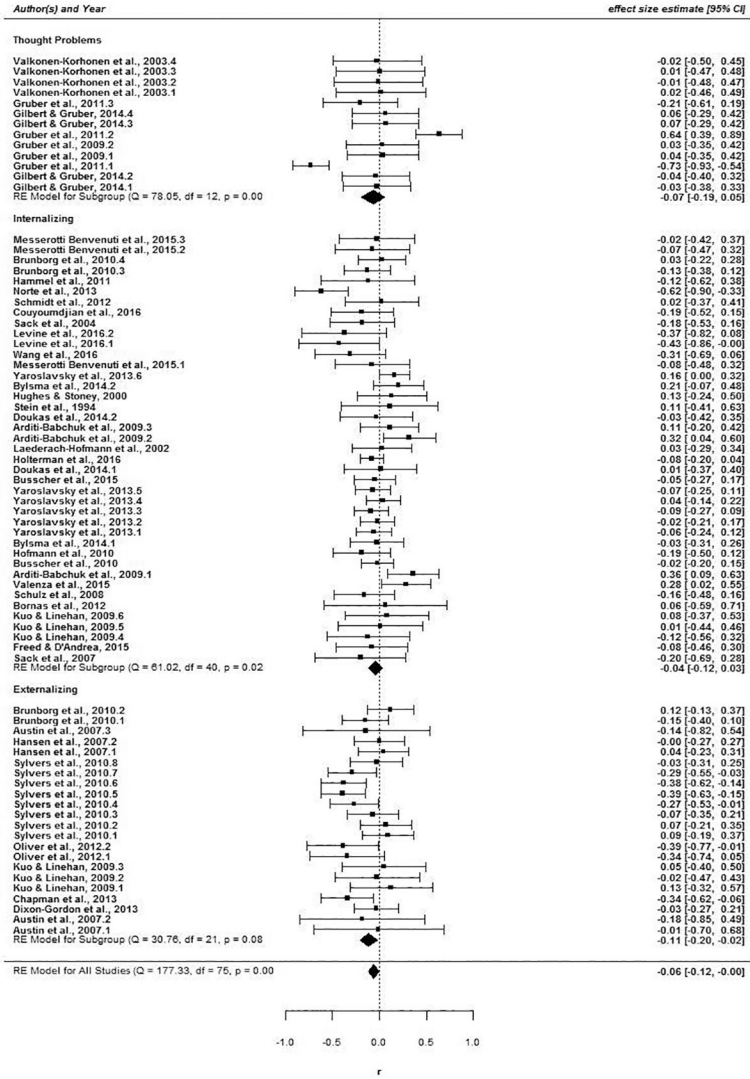


Figure 5. Forest plot of correlations between RSA reactivity and broadband psychopathology subtypes.

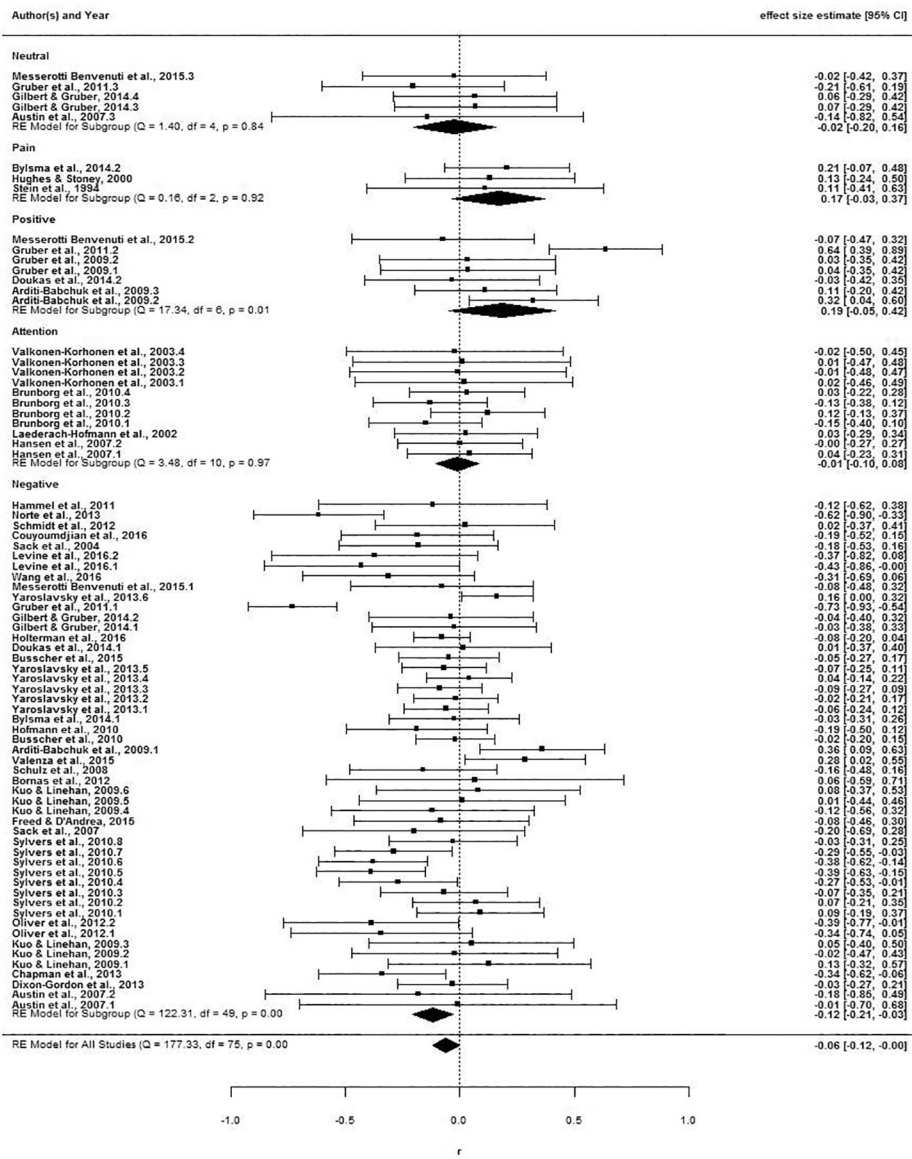


Figure 6. Forest plot of correlations between RSA reactivity and psychopathology by task type.

Table 1.

Summary of included studies and moderators.

Author(s)	Year	Mean age in years (SD)	% male	% Cauc.	Clinical sample N	Control sample N	Sample type ^a	Primary diagnosis	Assessment instrument	Psychopathology spectrum type	Medication status	Baseline type	Baseline length (s)	Task	Task category	Task length (s)	Sampling rate (Hz)	Task force referenced?
Suck et al.	2007	40.5 (8.4)	37.5	100	16	0	Clinical only	PTSD	SCID	Internalizing	Medications allowed	Vanilla	120	Trauma script	Negative Emotion Evocation	120	1024	No reference
Freed & D'Andrea	2015	37.5 (13.9)	0	74	27	0	At-risk	PTSD	PCL	Internalizing	Medications allowed	Vanilla	120	Trauma related images	Negative Emotion Evocation	50	.	No reference
Stern et al.	1994	36.6 (8.9)	53.3	.	15	15	Clinical and control	Social anxiety disorder	SCID	Internalizing	Medication free	Stimulus-free	600	Cold pressor	Thermoregulatory	180	1024	No reference
Austin et al.	2007	.	0	.	9	11	Clinical only	Borderline personality disorder	SCID	Externalizing	Medication free	Stimulus-free	300	Conflict films, neutral film	Negative Emotion Evocation, Neutral Emotion Task	300	512	No reference
Dixon-Gordon et al.	2013	21.68 (6.2)	0	38.2	68	0	Clinical and control	Borderline personality disorder	PAI-BOR	Externalizing	Medications allowed	Vanilla	300	Social rejection recording	Negative Emotion Evocation	300	.	No reference
Chapman et al.	2013	.	.	10.3	40	57	Clinical and control	Borderline personality disorder	PAI-BOR	Externalizing	Medications allowed	Vanilla	300	Fear film	Negative Emotion Evocation	209	.	No reference
Yaroslavsky et al.	2013	27.95 (5.17)	0	76	149	0	At-risk	Major depressive disorder	SCID	Internalizing	Medications allowed	Vanilla	360	Sad film	Negative Emotion Evocation	172	512	Referenced Task Force (1996)
Hughes & Shony	2000	18.4 (0.71)	50	35	28	25	Community	Major depressive disorder	BDI	Internalizing	Medication free	Stimulus-free	300	Cold pressor	Thermoregulatory	180	512	No reference
Kuro & Laibson	2009	23.6 (.)	0	65	20	20	Clinical and control	Borderline personality disorder	SCID	Externalizing	Medications allowed	Vanilla	300	Sad, fear and anger films	Negative Emotion Evocation	171	1024	No reference
Messner-Barnum et al.	2015	21.2 (1.9)	8	.	25	29	At-risk	Depression	BDI	Internalizing	Medication free	Vanilla	300	Phasant, unpleasant, and neutral emotional imagery	Negative Emotion Evocation, Pleasant Emotion Evocation, Neutral Emotion Task	180	512	Referenced Task Force (1996)
Wang et al.	2016	43.78 (9.9)	52.17	.	25	0	Clinical only	Panic disorder	SCID	Internalizing	Medications allowed	Vanilla	300	Aversive images	Negative Emotion Evocation	300	.	Referenced Task Force (1996)
Gilbert & Gamber	2014	30.94 (9.82)	35	90	31	31	Clinical and control	Bipolar disorder	SCID	Psychotic	Medications allowed	Vanilla	60	Remission, mindfulness	Negative Emotion Evocation, Neutral Emotion Task	300	1024	No reference
Hansen et al.	2007	32.07 (.)	100	.	53	0	At-risk	Psychopathy	Psychopathy Checklist	Externalizing	Medications allowed	Vanilla	300	Continuous performance task, working memory task	Attention Task	.	.	No reference
Lewis et al.	2016	21.6 (3.92)	0	.	15	15	Community	Generalized anxiety disorder	ADIS	Internalizing	Medications allowed	Vanilla	300	Worry tasks	Negative Emotion Evocation	300	1024	Referenced Task Force (1996)
Naltonen-Korhonen et al.	2003	27 (12.3)	44.44	.	18	21	Clinical and control	Psychotic disorder	SCID	Psychotic	Medication free	Vanilla	90	Oddball task, Wisconsin card sorting task	Attention Task	.	.	Referenced Task Force (1996)
Bernis et al.	2012	38.2 (10.36)	.	.	10	10	Clinical and control	Specific phobia	ADIS	Internalizing	Medications allowed	Vanilla	180	Flight film	Negative Emotion Evocation	180	512	No reference
Suck et al.	2004	38.9 (9.2)	32.26	.	31	0	Clinical only	PTSD	SCID	Internalizing	Medications allowed	Vanilla	120	Trauma script	Negative Emotion Evocation	120	1024	Referenced Task Force (1996)
Schult et al.	2008	18.83 (0.94)	36.11	72.22	36	0	Community	Social anxiety disorder	Social Phobia and Anxiety Inventory	Internalizing	Medications allowed	Vanilla	120	Anticipation of speech	Negative Emotion Evocation	720	1024	No reference
Valenza et al.	2015	.	.	.	48	48	Clinical and control	Major depressive disorder	SCID	Internalizing	Medication free	Stimulus-free	600	Emotional narration	Negative Emotion Evocation	.	.	No reference
Authi-Bakshuk et al.	2009	25 (4)	35	.	40	0	Clinical only	PTSD	PDS	Internalizing	Medications allowed	Vanilla	300	Trauma script, phasant script	Negative Emotion Evocation	300	512	No reference
Leibovich-Hoffmann et al.	2002	32 (9)	55	.	40	20	Clinical only	Specific phobia	.	Internalizing	Medications allowed	Stimulus-free	300	Stroop task	Attention Task	.	.	No reference
Busecher et al.	2010	40.5 (11)	44.88	.	127	36	Clinical and control	Specific phobia	Flight Anxiety Situations	Internalizing	Medication free	Vanilla	360	Flight film	Negative Emotion Evocation	360	1024	No reference
Hoffmann et al.	2010	28 (9.92)	28.21	10.26	39	0	Clinical only	Generalized anxiety disorder	ADIS	Internalizing	Medications allowed	Vanilla	300	Worry task	Negative Emotion Evocation	300	512	No reference
Coyanadjan et al.	2016	28.8 (9.8)	18.18	100	33	0	Clinical only	Specific phobia	SCID	Internalizing	Medication free	Vanilla	300	Phobic animal film	Negative Emotion Evocation	30	.	Referenced Task Force (1996)
Gamber et al.	2009	34.37 (15.78)	22.2	55.6	27	27	Clinical and control	Bipolar disorder	SCID	Psychotic	Medications allowed	Vanilla	60	Happy memory recalls	Pleasant Emotion Evocation	60	1024	No reference
Byluma et al.	2014	31.2 (11.46)	16.3	99.2	49	45	Clinical and control	Major depressive disorder	SCID	Internalizing	Medications allowed	Vanilla	600	Speech preparation, cold pressor	Negative Emotion Evocation	180	1024	No reference
Schmidt et al.	2012	33.3 (11)	54.54	.	26	0	Clinical only	Social anxiety disorder	SCID	Internalizing	Medications allowed	Vanilla	300	Speech preparation	Negative Emotion Evocation	180	512	Referenced Task Force (1996)
Bramberg et al.	2010	.	49.2	.	61	0	Clinical only	Anxiety	Hospital Anxiety and Depression Scale, Daily Tobacco Use, South Oaks Gambling Screen	Internalizing	Medications allowed	Vanilla	300	Gambling task	Attention Task	.	.	Referenced Task Force (1996)
Yaroslavsky et al.	2013	(-3.82)	49.2	.	113	93	Clinical and control	Major depressive disorder	SCID, BDI, EAI, Follow-Up Depressive Scale (FDS), Follow-Up Anxiety Scale (PAS)	Internalizing	Medications allowed	Vanilla	384	Sad film	Negative Emotion Evocation	172	512	No reference
Neve et al.	2013	25.97 (7.53)	26	78	19	16	Clinical and control	PTSD	SCID	Internalizing	Medications allowed	Vanilla	300	Trauma script	Negative Emotion Evocation	40	1024	Referenced Task Force (1996)
Hammad et al.	2011	20.8 (2.9)	18.75	69	16	19	Clinical and control	Generalized anxiety disorder	ADIS	Internalizing	Medication free	Vanilla	420	Worry induction	Negative Emotion Evocation	630	1024	Referenced Task Force (1996)

Author(s)	Year	Mean age in years (SD)	% male	% Cauc.	Clinical sample N	Control sample N	Sample type ^a	Primary diagnosis	Assessment instrument	Psycho-pathologic spectrum type	Medications status	Baseline type	Baseline length (s)	Task	Task category	Task length (s)	sampling rate (Hz)	Task power relevance?
Gehrer et al.	2011	39.13 (2.61)	22.2	77.8	25	24	Clinical and control	Bipolar disorder	SCID	Psychotic	Medications allowed	Vanilla	60	Happy film, sad film, neutral film	Negative Emotion Evocation, Pleasant Emotion Evocation, Neutral Emotion Task	331	1024	No reference
Bauscher et al.	2015	40.4 (11)	46.84	.	79	0	Clinical only	Specific phobia	self-referat	Internalizing	Medication free	Vanilla	300	Flight film	Negative Emotion Evocation	1024	No reference	
Doukas et al.	2014	38.77 (13.34)	0	74	27	0	Clinical only	PTSD	PCL	Internalizing	Medications allowed	Vanilla	120	Trauma images, positive images	Negative Emotion Evocation, Pleasant Emotion Evocation	512	No reference	
Oliver et al.	2012	. (.)	25	75	20	22	Clinical and control	ADHD	ADHD Diagnostic Scale	Externalizing	Medications allowed	Vanilla	60	Frustration tasks	Negative Emotion Evocation	1024	No reference	
Slyvers et al.	2010	. (.)	100	81	50	0	Community	Antisocial personality disorder	SCID, Psychopathic Personality Inventory	Externalizing	Medications allowed	Vanilla	300	Threat images	Negative Emotion Evocation	.	No reference	
Hohmann et al.	2016	18.77 (0.97)	26	84	247	0	Community	Major depressive disorder	CIESD	Internalizing	Medications allowed	Vanilla	240	Cyberball	Negative Emotion Evocation	1024	No reference	

Notes. Some studies, such as those with multiple psychopathological groups, provided more than one effect size.

^aHedges' g was calculated across studies that included a clinical and control group, and r was calculated for studies with a single sample in which psychopathology was measured continuously.