

UCSF

UC San Francisco Previously Published Works

Title

Autonomic nervous system functioning assessed during the Still-Face Paradigm: A meta-analysis and systematic review of methods, approach and findings.

Permalink

<https://escholarship.org/uc/item/3bs6s73z>

Journal

Developmental review : DR, 50(Pt B)

ISSN

0273-2297

Authors

Jones-Mason, Karen
Alkon, Abbey
Coccia, Michael
et al.

Publication Date

2018-12-01

DOI

10.1016/j.dr.2018.06.002

Peer reviewed

Contents lists available at [ScienceDirect](#)

Developmental Review

journal homepage: www.elsevier.com/locate/dr

Autonomic nervous system functioning assessed during the still-face paradigm: A meta-analysis and systematic review of methods, approach and findings

Karen Jones-Mason^{a,*}, Abbey Alkon^b, Michael Coccia^a, Nicole R. Bush^{a,c}

^a Department of Psychiatry, Center for Health and Community, Weill Neurosciences Institute, University of California, San Francisco, United States

^b School of Nursing Department of Family Health Care Nursing, University of California, San Francisco, United States

^c Department of Pediatrics, University of California, San Francisco, United States

ARTICLE INFO

Keywords:

Still-face paradigm
Autonomic nervous system
Respiratory sinus arrhythmia
Meta-analysis
Systematic review

ABSTRACT

Animal and human research suggests that the development of the autonomic nervous system (ANS) is particularly sensitive to early parenting experiences. The Still-Face Paradigm (SFP), one of the most widely used measures to assess infant reactivity and emotional competence, evokes infant self-regulatory responses to parental interaction and disengagement. This systematic review of 33 peer-reviewed studies identifies patterns of parasympathetic (PNS) and sympathetic (SNS) nervous system activity demonstrated by infants under one year of age during the SFP and describes findings within the context of sample demographic characteristics, study methodologies, and analyses conducted. A meta-analysis of a subset of 14 studies with sufficient available respiratory sinus arrhythmia (RSA) data examined whether the SFP reliably elicited PNS withdrawal (RSA decrease) during parental disengagement or PNS recovery (RSA increase) during reunion, and whether results differed by socioeconomic status (SES). Across SES, the meta-analysis confirmed that RSA decreased during the still-face episode and increased during reunion. When studies were stratified by SES, low-SES or high-risk groups also showed RSA decreases during the still face episode but failed to show an increase in RSA during reunion. Few studies have examined SNS activity during the SFP to date, preventing conclusions in that domain. The review also identified multiple qualifications to patterns of SFP ANS findings, including those that differed by ethnicity, infant sex, parental sensitivity, and genetics. Strengths and weaknesses in the extant research that may explain some of the variation in findings across the literature are also discussed, and suggestions for strengthening future research are provided.

Introduction

Early childhood experiences influence health and development across the life course (Bosquet Enlow et al., 2014; Brody et al., 2013; Chen, Miller, Kobor, & Cole, 2011; Shonkoff, Boyce, & McEwen, 2009; van Lien, Neijts, Willemsen, & de Geus, 2015). Animal and human research suggests that the autonomic nervous system (ANS) is particularly sensitive to the quality of early parental care (Alkon et al., 2014; Hostinar, Sullivan, & Gunnar, 2014; McLaughlin et al., 2015; Propper, 2012; Shonkoff et al., 2009; Sroufe, 2005) and is a key factor in the prediction of mental health (Anda et al., 2006; De Bellis & Zisk, 2014; Shonkoff et al., 2009). Understanding the development of the ANS early in life advances opportunities for the prevention and treatment of future health problems, yet

* Corresponding author at: Department of Psychiatry, UCSF, 3333 California Street, Suite 465 San Francisco, CA 94118, United States.
E-mail address: Karen.Jones-Mason@ucsf.edu (K. Jones-Mason).

<https://doi.org/10.1016/j.dr.2018.06.002>

Received 13 September 2017; Received in revised form 17 May 2018
0273-2297/© 2018 Elsevier Inc. All rights reserved.

studying ANS functioning in infancy, particularly in relation to stress reactivity, provides many challenges. The Still-Face Paradigm (SFP) (Tronick, Als, Adamson, Wise, & Brazelton, 1978), one of the most widely used measures to assess infant reactivity and emotional competence, evokes infant self-regulatory responses to parental interaction and disengagement. SFP studies have provided important data that inform a broad evidence base by examining diverse developmental phenomena, including infant attachment (Braungart-Rieker et al., 2014; Holochwost, Garipey, Propper, Mills-Koonce, & Moore, 2014; Planalp & Braungart-Rieker, 2013), infant memory for social stress (Montirosso et al., 2014), infant temperament (Conradt & Ablow, 2010), infant responses to maternal sensitivity (Braungart-Rieker, Garwood, Powers, & Wang, 2001; Conradt & Ablow, 2010; Moore et al., 2009; Tarabulsky et al., 2003), the impact of parental anxiety (Grant et al., 2009), and infant sex and cultural differences (Kisilevsky et al., 1998; Weinberg, Tronick, Cohn, & Olson, 1999). The SFP may also offer some of the earliest insights into infant development and trajectories of future development and health. Although the SFP is similar to the gold-standard measure of infant attachment, the Strange Situation (SSP; Ainsworth, Blehar, Waters, & Wall, 1978) in that it involves parental engagement and disengagement and elicits similar stress in young children (Mesman, van Ijzendoorn, & Bakermans-Kranenburg, 2009), the SFP can be administered much earlier in development than the 12–18 months target age for the SSP, sometimes as early as the first few hours of life (Nagy, 2008).

The SFP has increasingly been used in studies of infant ANS function, providing needed insight into the development of this system in early life. Previous publications have reviewed the history of the SFP (Adamson & Frisk, 2003) and conducted a review and a meta-analysis of SFP findings with behavioral measures (Mesman et al., 2009). One review and meta-analysis examined differences in heart rate variability (HRV) between healthy versus at-risk infants and older children across 18 studies that used a variety of social engagement-disengagement tasks, some of which included the SFP (Shahrestani, Stewart, Quintana, Hickie, & Guastella, 2014). Other past reviews have examined prenatal and childhood proximal risk to ANS function citing some of the SFP studies (Propper & Holochwost, 2013; Propper, 2012). Another recent meta-analysis examined infant cortisol reactivity during administration of the SFP (Provenzi, Giusti, & Montirosso, 2016). These reviews have been helpful in highlighting the utility of the SFP, but ANS measurement techniques are advancing, and more recent studies are accumulating, suggesting that an up-to-date systematic review and analysis of ANS function within SFP studies, including both parasympathetic and sympathetic nervous system measures, will be informative.

Here we present a systematic review and meta-analysis of studies examining infant ANS responses demonstrated during the administration of the SFP. The review has three objectives: (1) to summarize patterns of infant ANS activity observed in the SFP across studies to date; (2) to identify different variables and methodologies used in the assessments that may, ultimately, explain divergent findings; and (3) to conduct a meta-analysis to validate the SFP as a test of the infant parasympathetic nervous system (PNS) response to a social stressor. This review starts by providing a brief operational definition of the parasympathetic and sympathetic branches of the ANS and the measures commonly used to assess both systems. The SFP is then described, followed by a description of the methodology used in this review. The review then presents a summary of the overall patterns of infant ANS activity detected in the SFP and identifies variations in administration, sampling, and analytic approach that appear relevant to variation in findings. This review is, by necessity, focused on PNS evidence but the limited findings on sympathetic nervous system (SNS) activity are also described. Meta-analysis of respiratory sinus arrhythmia (RSA) findings within the SFP are described to determine whether infant RSA consistently differentiates between periods of parental social engagement and disengagement, and whether such patterns may be modified by socioeconomic and risk status, or age. Finally, strengths and weaknesses in the literature, conclusions that can be drawn from extant research, and recommendations for future research are provided.

Background

The autonomic nervous system: constructs and measures

The ANS regulates the central nervous system (CNS), facilitates individual adjustment to changes in the internal and external environment, including adaptation to psychosocial stress, and maintains overall homeostasis (Bernston, Cacioppo, & Quigley, 1993; Mendes, 2009; Porges, 1995). The ANS consists of sensory and motor neurons that connect the CNS to a variety of internal organs including the heart, lungs, glands (e.g., endocrine) and viscera (internal organs, e.g., intestines). The ANS has two main divisions; the SNS and the PNS. The SNS activates vigilance, arousal, and mobilization in response to perceived threat (e.g., the “flight or fight” response) that stimulates the heart to beat faster and the digestive system to slow down (Alkon et al., 2014; Sapolsky, 2004; Selye, 1956). The PNS acts to slow down bodily functions such as heart rate (HR), promotes growth and generally stimulates functions that occur when the body is resting (e.g., digestion, elimination and salivation); for that reason it is known as, “the rest and digest system” (Alkon et al., 2014; Sapolsky, 2004; Selye, 1956). The PNS and SNS can have synergistic or opposing effects, or may show asymmetrical patterns of innervations (Alkon, Wolff, & Boyce, 2012; Bernston, Quigley, & Lazano, 2007).

The tenth cranial nerve, also referred to as the “vagus nerve,” is thought to regulate “homeostasis” (i.e., the resting state of the internal organs such as heart and lungs) via afferent (sensory) fibers in the vagus that carry messages from most of the internal organs, the pharynx and the larynx to the brain, and thus is considered to be a key component of the PNS (Porges, 2011; Stewart et al., 2013). The vagus also acts as an efferent nerve carrying signals from the brain to visceral organs (Cacioppo & Bernston, 2011). “Vagal tone” (VT), a commonly referenced measure of parasympathetic activity, is sometimes described as an index of neural regulation of cardiac activity by way of the vagus nerve (Mendes, 2009; Porges, 2011). Because the PNS plays such a vital role in decreasing HR and returning the body to homeostasis after exposure to stress, vagal tone is thought to be a significant indicator of self-regulatory capacity, particularly within the context of social interactions (Porges, 1995, 2007). Since vagal tone cannot be measured directly, indirect methods such as RSA are used. Fundamentally, RSA represents variations in HR that occur with respiration; HR increases with inhalation and decreases with exhalation (Ben-Tal, Shamailov, & Paton, 2012; Grossman & Taylor, 2007;

Zisner & Beauchaine, 2016). RSA decreases when the PNS withdraws which allows HR to increase. On the other hand, when RSA increases, the PNS is activated which allows HR to decrease (e.g., resting state). HR is influenced by both the SNS and PNS and may increase without any concomitant change in RSA (Moore & Calkins, 2004). Thus, RSA is the preferred index of PNS activity.

Researchers use multiple methods to calculate RSA from electrocardiogram (ECG) data, including frequency-domain, time-domain, and non-linear measures (Bernston et al., 2007; Mendes, 2009; TFESOC, 1996). The studies included in this review primarily use frequency and time domain methods. Among the time domain measures, the peak-valley statistic is a common method that determines the differences between the minimum HR during expiration and maximum HR during inspiration and is reported in units of millisecond (ms) (Grossman & Taylor, 2007; Grossman, Karemaker, & Wieling, 1991; Ritz et al., 2012). Frequency domain measures gauge the power of HRV within low, mid or high frequency bands (Zisner & Beauchaine, 2016). Variations in HR that occur in the highest frequency band occur at the same rate of respiratory inhalation and exhalation. Accordingly, RSA is considered the “high frequency (HF)” component of HRV (Grossman & Taylor, 2007). RSA indices are derived using standardized scoring programs that calculate RSA using interbeat intervals (IBI) from one R peak to another R peak on the EKG/ECG signal (sometimes referred to as heart period (HP)) and respiration on a monitor or derived from an impedance signal (e.g., dZ/dt) (Bernston et al., 2007; Mendes, 2009). Because frequency domain methods report the variation of IBI occurring within respiratory frequency, units are reported in ms squared (consistent with statistical units of variance) (Grossman & Taylor, 2007). The most common method for assessing RSA in the frequency domain is spectral analysis (Zisner & Beauchaine, 2016). Most of the studies reviewed here calculate their estimates of RSA using Porges (1985) algorithm, a method that has been widely used to assess infant RSA (Beauchaine, 2001). Porges (1985) applies a digital bandpass filtering technique to remove non-respiratory variations in the IBI (Porges, 2011) and may be used with both frequency- and time-domain methods (Grossman, van Beek, & Wientjes, 1990). Results are natural log transformed and are reported in \ln (ms squared) units. Results from the various methods have validated their derivation of RSA indices with strong, positive correlations with vagal tone indices (Goedhart, van der Sluis, Houtveen, Willemsen, & de Geus, 2007; Grossman et al., 1990); see (Lewis, Furman, McCool, & Porges, 2012) for disagreement). Skin conductance (SC), T-wave amplitude (TWA) and pre-ejection period (PEP) are measures used to index SNS activity (see Mendes (2009) for a detailed review). SC measures electrodermal activity in the skin and is associated with emotional arousal (Ham & Tronick, 2009). SC is measured with electrodes placed on the skin (frequently the hands and feet where eccrine sweat glands are most dense) while an electric current is passed between the two points; the resistance to the current is then measured, and the reciprocal of the resistance is referred to as skin conductance (Mendes, 2009). TWA attenuation is associated with SNS activation (Bosquet Enlow et al., 2014; van Lien et al., 2015). TWA is measured with an ECG (Bosquet Enlow et al., 2014) and assesses the ventricular repolarization that occurs at the end of each cardiac cycle (van Lien et al., 2015). PEP represents the period from the electrical stimulation of the heart’s left ventricle to the point at which the semilunar aortic valve opens and blood is ejected into the aorta (Cacioppo, Uchino, & Bernston, 1994). PEP is measured using impedance cardiography (Alkon et al., 2012). Shorter PEP denotes SNS activation and higher HR (El-Sheikh & Erath, 2011).

Overall, ANS measures during baseline and reactivity are commonly used to describe an infant’s physiologic response to a resting or challenging condition. Infants’ ANS measures during resting and challenging conditions are usually normally distributed within each sample, thus, showing there are individual differences in their ANS reactivity. ANS reactivity is an indication of the change from a challenging condition to a baseline or resting condition.

The Still-face Paradigm (SFP)

Introduction to the SFP

The SFP was designed to examine infant capacity for self-regulation during social interaction with their parent (Tronick et al., 1978). According to Tronick and colleagues, the infant’s reaction to the still-face episode (SF) should show the “importance of interactional reciprocity” to the infant and the infant’s “ability to regulate his/her affective displays to achieve the goals of the interaction.” (Tronick et al., 1978, p. 2). Since Tronick’s first published study in 1978, the SFP has become a common procedure in infant research (Adamson & Frisk, 2003; Mesman et al., 2009).

The standard SFP consists of a sequence of three, 2-minute episodes in which the parent and the infant are seated about one meter away from each other. During the first episode, the parent is free to play with the infant as she or he would at home. During the “still-face” episode (SF), the parent maintains a neutral face and is told not to touch or interact with the infant. The third episode is a resumption of play sometimes referred to as the “reunion” episode. Many researchers, however, adapt the SFP to fit their needs, which leads to considerable variation in SFP administration across studies. For example, some researchers add a second SF and reunion episode (Bosquet Enlow et al., 2014; Bush et al., 2017; Haley & Stansbury, 2003; Haley, Handmaker, & Lowe, 2006). Others have had the parent turn around in between episodes (Moore, Cohn & Campbell, 2001), leave the infant alone in the room (Grant et al., 2009; Grant, McMahon, Reilly, & Austin, 2010; Stoller & Field, 1982), or even substitute strangers for the parent (Bazhenova, Plonskaia, & Porges, 2001; Stewart et al., 2013). Although the SFP is frequently administered when infants are approximately 6 months of age, it has also been used with infants as young as 3 h old (Nagy, 2008) and children as old as 82 months (Ostfeld-Etzion, Golan, Hirschler-Guttenberg, Zagoory-Sharon, & Feldman, 2015). In their review, Mesman et al. (2009) reported episode durations ranging from 60 s to 180 s, although at least one study used 45-second episode intervals (Stoller & Field, 1982). The still-face “effect” on infant behavior has been well documented (Adamson & Frisk, 2003; Mesman, Linting, Joosen, Bakermans-Kranenburg, & van Ijzendoorn, 2013). One major review and meta-analysis (Mesman et al., 2009) confirmed the classic still-face effect of reduced positive affect and gaze, and increased negative affect, as well as a partial carry-over effect into the reunion episode consisting of lower positive and higher negative affect compared to baseline. Mesman et al. (2009) also concluded that the still-face effect is strong because it has been detected regardless of factors such as infant sex, ethnicity, “risk status,” and procedural differences (e.g., length of

the SFP episodes and the use of intervals between episodes). Yet, a relatively recent study showed the standard still-face effect in only approximately half of the infants assessed, while a considerable minority showed no change from SF to reunion; further, only 4–17% of infants showed the predicted patterns for negative affect and gaze from baseline to SF, and from SF to reunion (Mesman et al., 2013). Thus, although sample averages typically demonstrate the classic still face effect, there is considerable variability in individual responses. Multiple theories explaining the still face effect have been offered (Adamson & Frisk, 2003; Mesman et al., 2009). For example, originally Tronick and colleagues suggested that infants show distress within the SF episode because the infant's expectations of parental attention and responsiveness is violated (Tronick et al., 1978). Later, Tronick and colleagues formulated other models. For example, the “Mutual Regulation Model (MRM) (Tronick & Weinberg, 1997) recognized the mistakes and mismatched communications that are inherent in the parent-child relationship and emphasized the importance of relational repair to the infant's sense of self efficacy and regulation (Tronick & Beeghly, 2011). The Dyadic States of Consciousness Model (DSCM) asserts that the mutual regulation of affect between parent and infant allows them both to increase the complexity of their own state of consciousness (SOC), that is their own understanding of themselves and their “place” in the world (Tronick et al., 2005). The SF is disturbing to the infant because the parent appears to be sending a contradictory message precluding the co-construction of a coherent dyadic state of consciousness, forcing the infant to rely on their own SOC, a situation the infant could find confusing and threatening. Other researchers have focused on the failure of the parent figure to provide emotional regulation to the infant (Field, 1994). In short, multiple explanations for the still-face effect exist, with most acknowledging that although adults provide what is, in effect, scaffolding for infant self-regulation, infants also actively contribute to interactive processes (Mesman et al., 2009).

The SFP has shown validity (Braungart-Rieker et al., 2014; Hill & Braungart-Rieker, 2002; Holochwest et al., 2014; Moore et al., 2001; Yazbek & D'Entremont, 2006), and test-retest reliability (Tronick & Weinberg, 1990; Montirosso et al., 2014; Provenzi, Olson, Montirosso, & Tronick, 2016) but weak stability over time (Braungart-Rieker et al., 2014; Mesman et al., 2009; Toda & Fogel, 1993).

SFP and ANS functioning: Psychobiological model

Porges' “polyvagal theory” (Porges, 1995, 2007, 2011) is the most common psychobiological theory of early development explored within the context of SFP studies. This theory postulates that vagal functioning via the PNS plays a key role in facilitating interactive social experience and communication (Porges, 2007) via neural pathways that not only regulate vagal control of the heart but also muscles of the face and head associated with social expression (Stewart et al., 2013). Porges proposes that when the individual experiences the environment as safe, the myelinated vagus is activated first regulating the body to lower cardiac functioning, inhibiting SNS fight or flight mechanisms and HPA (hypothalamic-pituitaryadrenal) functioning, and even decreasing inflammation. If the environment is perceived as threatening the individual may resort to two other, older subsystems (the “sympathetic-adrenal system” related to active avoidance, and the “immobilization system” associated with passive avoidance) consecutively. Porges (2007) asserts that the “vagal brake” supports individual efforts to interact or disengage with others and fosters the ability to self-soothe and calm, making it particularly relevant for studies of infant regulation in the SFP.

In line with polyvagal theory, higher resting RSA and consistent RSA suppression in response to challenge are considered “positive” indicators of social and emotional regulation, while lower resting RSA and inconsistent RSA suppression are “risk” indicators for problems in social and emotion regulation (Beauchaine, 2007; Graziano & Derefinko, 2013; Zisner & Beauchaine, 2016). Indeed, high resting RSA and low RSA during challenge in children has been associated with positive emotions, social outcomes and effective regulation (Bazhenova et al., 2001; El-Sheikh, Harger, & Whitson, 2001; Kogan et al., 2014; Propper, 2012). Some researchers have nevertheless reported problematic outcomes in children with both high and low levels of RSA. Maintaining high levels of vagal tone during stress has been associated with regulatory dysfunction in children (Beauchaine, Gatzke-Kopp, & Mead, 2007; Calkins, Graziano, & Keane, 2007) while high vagal tone during non-challenging circumstances has also been associated with reduced emotion regulation in adolescents, young children and infants (Dietrich et al., 2007; Eisenberg et al., 1995). Other researchers have failed to find that either baseline or RSA suppression in response to challenge predicts self-regulatory behaviors in young children (Calkins et al., 2007; Eisenberg et al., 2012; Stevenson-Hinde & Marshall, 1999). Examining ANS function in the SFP in young infants, when the brain is particularly malleable and open to environmental influence, and self-regulatory behaviors are developing, may help inform this dynamic area of research.

Review: Method

Literature search

The literature search for this review was conducted between October 2014 and January 2017. The search strategy was informed by guidelines set forth in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher, Liberati, Tetzlaff, Altman, & Group, 2009). The studies were identified using electronic databases including PubMed/MED-LINE, PsychINFO and Web of Science. Search term combinations included the phrase “still-face” (“paradigm” wasn't included in order to ensure that we maximized the number of studies detected) to identify studies that administered the task to infants, and studies were examined for simultaneous use of physiologic measures of the ANS, including “respiratory sinus arrhythmia”, “vagal tone,” “pre-ejection period,” “heart rate” and “skin conductance.” “Autonomic nervous system,” “parasympathetic nervous system” and “sympathetic nervous system” were also used in the search combination to capture any remaining ANS measures. Citations appearing within the studies collected as well as references in reviews were also considered for inclusion. Although international studies were included, most of those studies were conducted on European samples. A flowchart with our selection steps is provided in Fig. 1. Inclusion criteria included that research was submitted in English, that studies actually used the SFP data in their analyses, that studies used valid and

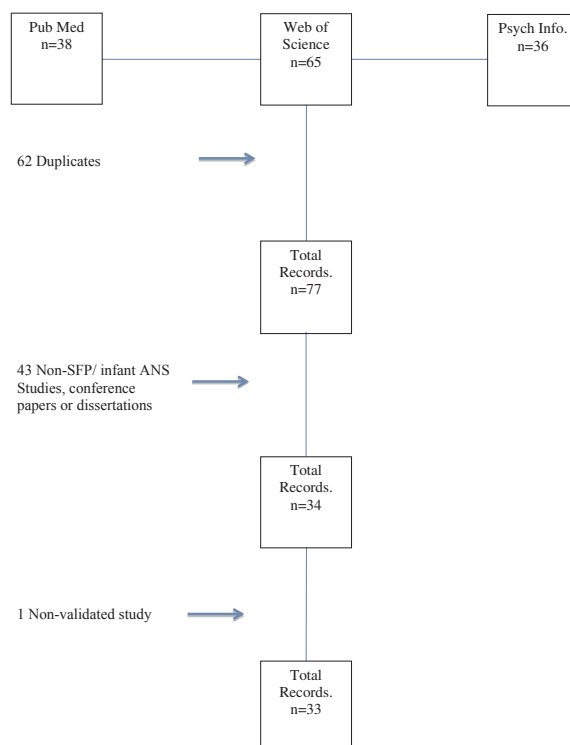


Fig. 1. Flowchart.

reliable measures to assess ANS in the SFP, and that the studies appeared in peer reviewed journals (e.g. Grossman & Taylor, 2007; Porges, 1985; Zisner & Beauchaine, 2016). The studies included and the data extracted for this review are presented in Tables 1, 2 and Supplementary Table S1. Data extracted included sex, infant age, ethnicity, sample size, SES, RSA means/SD, whether the SFP was conducted in the home or in a lab, clinical versus non-clinical samples, analytic methods and outlier removal criteria, and selected measures of interest and findings. All authors participated in the determination that the SFP was being administered and that valid measures were used to assess ANS function. Two authors participated in extracting data values from the publications. All studies measuring ANS function during administration of the SFP were included except one that used an ANS measure that has not yet been fully validated. In all, 33 studies met our criteria for the review.

Results of SFP and ANS studies

Stoller and Field (1982), one of the first published ANS findings within the context of the SFP, found an initial deceleration of infant HR at the start of the SF episode, but by 7 s into the task, HR accelerated producing the classic negative response. Since that time, over 30 other published studies have measured some facet of ANS activity in the SFP. Early research studies suggested that infants showed an ANS response in the SFP consisting of a decrease in RSA during the SF episode and a return to baseline levels (i.e., higher RSA) during reunion (Bazhenova et al., 2001; Ham & Tronick, 2006; Weinberg & Tronick, 1996). In their analyses of 18 studies of infants and older children, Shahrestani et al. (2014) found that RSA did not change between baseline and various social engagement tasks (e.g., Strange Situation, teaching and play tasks, and SFP) but decreased between baseline and disengagement tasks. In the meta-analysis involving only the SFP ($n = 7-8$) Shahrestani and colleagues found that RSA was lower in the SF compared to baseline and higher during reunion compared to SF, with no significant differences between baseline and reunion. When one study with a high-risk sample was added to their analysis, the high-risk children did not show a difference in RSA between the SF and reunion episodes. This review builds upon prior research by incorporating additional and more recent SFP studies, including studies using measures of the SNS, and critically examines a range of factors that can lead to differing RSA patterns across SFP studies, including infant age, sex, ethnicity, SES, parental behavior and psychopathology.

The following section reviews general patterns in parasympathetic and sympathetic activity across the studies that met inclusion criteria. Although we often highlight when two published studies drew from the same participant sample, the reader is directed to Table S1 for details on which studies share at least some participants, as well as studies' sample sizes (note that publications drawn from the same sample sometimes had notably different sample sizes). Table 1 provides the SFP means and standard deviations of RSA, HR/HP and PEP from publications included in this review or that were provided to us through personal communication by those study authors. When possible, in order to standardize the findings and improve comparability across studies, we included our calculated effect sizes using Hedges' g (Hedges & Olkin, 1985) reflecting the size of the change in the physiological variables between

Table 1
Means (SD) of HR and RSA in SFP and effect sizes.

Author, Year	HR Baseline	HR Play #1	HR SF #1	HR Play #2	HR SF#2	HR Play#3	RSA Baseline	RSA Play#1	RSA SF #1	RSA Play#2	RSA SF#2	RSA Play#3
Bush et al. (2017)							4.26 (1.07)	4.27 (1.04) ES: 0.01	3.99 (1.21) ES: -0.25	4.21 (1.21) ES: 0.18	3.73 (1.16) ES: -0.40	4.10 (1.15) ES: 0.32
Busuito and Moore (2017)								3.65 (1.07)	3.40 (0.93) ES: -0.25	3.61 (1.09) ES: 0.21		
Quigley et al. (2016)							3.68 (0.87)	3.57 (0.97) ES: -0.12	3.40 (1.00) ES: -0.17	3.47 (0.99) ES: 0.07		
Suurland et al. (2016)	Low risk: 135.30 (13.51) High risk: 134.74 (12.74)	Low risk: 140.21 (11.40) ES: 0.39 High risk: 139.72 (12.02) ES: 0.40	Low risk: 150.40 (14.31) ES: 0.78 High risk: 143.96 (14.73) ES: 0.31	Low risk: 146.39 (16.50) ES: -0.26 High risk: 148.38 (18.05) ES: 0.27			Low risk: 3.35 (0.45) High risk: 3.41 (0.36)	Low risk: 3.36 (0.38) ES: 0.02 High risk: 3.42 (0.33) ES: 0.03	Low risk: 3.14 (0.56) ES: -0.46 High risk: 3.27 (0.44) ES: -0.38	Low risk: 3.28 (0.53) ES: 0.26 High risk: 3.13 (0.57) ES: -0.27		
Pratt et al. (2015)									Change in RSA from P to SF: 0.4 (0.73)	2.83 (1.01)		
Provenzi et al. (2015)								Suppressors: 55.30 (74.04) Non- suppressors: 21.24 (18.12)	Suppressors: 25.65 (23.09) ES: -0.54 Non- suppressors: 49.00 (62.18) ES: 0.60	Suppressors: 64.50 (132.83) ES: 0.41 Non- suppressors: 48.80 (50.86) ES: 0.0		
Montirosso et al. (2014)								62.10 (124.25)	54.14 (74.69) ES: -0.08	144.09 (519.52) ES: 0.24		
Holochwost, Garipey, Propper, Mills- Koonce, W.R., Moore (2014), 6 month olds							3.72 (0.90)	3.57 (1) ES: 0.15	3.46 (1.09) ES: 0.11	3.51(1.08) ES: 0.10		
Holochwost et al. (2014)							3.72 (0.909)	3.57 (1) ES: -0.16	3.46 (1.09) ES: -0.10	3.51(1.08) ES: 0.05		

(continued on next page)

Table 1 (continued)

Author, Year	HR Baseline	HR Play #1	HR SF #1	HR Play #2	HR SF#2	HR Play#3	RSA Baseline	RSA Play#1	RSA SF #1	RSA Play#2	RSA SF#2	RSA Play#3
Tibu et al. (2014)							Helper-Hinderer: Boys: 3.14 (0.86) Girls: 3.21 (0.84) Novel Toy: Boys: 2.93 (0.75), Girls: 3.01 (0.86) ES (boys): -0.26 ES (girls): -0.23	Social Engage: Boys: 3.38 (0.77), ES: 0.59 (from NT) Girls: 3.27 (0.89) ES: 0.30	Still Face: Boys: 2.80 (0.73), ES = -0.77 Girls: 2.86 (0.79) ES = -0.49	Reunion: Boys: 3.40 (0.91), ES = 0.73 Girls: 3.30 (0.99) ES = 0.49		
Mattson et al. (2013)		151.47 SE = 1.26	157.78 SE = 1.41	158.07, SE = 1.79								
Gunning et al. (2013)		152.2 (9.8)	156.3 (11.3) ES: 0.39	155.6(12.0) ES: -0.06								
Stewart et al. (2013)		HP-441.78 (33.97)	HP: 409.96 (39.54) ES: -0.86	Not stated				3.25(0.74)	3.14 (0.91) ES = -0.13	Not stated		
Weisman et al. (2012)								3.45 (.77)	3.25 (.77) ES = -0.26	3.63 (1.14) ES: 0.39		
Ritz et al. (2012) (RSA adjusted for respiration)		(n = 23) 135.06 (11.5) (n = 16) 137.4 (10.5)	146.66 (12.6) ES: 0.95 147.1 (12.5) ES = 0.82	144.56 (11.2) ES: -0.17 142.3 (10.2) ES = -0.41	152.1 (15.5) ES = 0.73	149.3 (16.4) ES = -0.17	n = 23 0.139 (0.020) n = 16 0.125 (0.021)	0.108 (0.022) ES: -1.45 0.095 (0.022) ES: -1.36	0.112 (0.020) ES: 0.019 0.105 (0.016) ES: 0.51	0.068 (0.037) ES: 0.59 -1.3		
Sharp et al. (2012)							#Helper-Hinderer (HH): 3.18 (0.85) Novel Toy (NT): 2.97(0.81) ES: -0.25	3.32 (0.84) ES: 0.42 (from NT)	2.83 (0.76) ES = -0.61	3.35 (0.95) ES = 0.60		
Conradt and Ablow (2010)		137.95 (9.74)	149.10 (13.25) ES: 0.89	150.11 (17.32) ES: 0.07				.0170 (.013) ES: 0.18	.0141 (0.0067) ES: -0.28	.0141 (0.0064) ES: 0		
Moore (2010)								3.68 (1.08)	3.41 (0.96) ES: -0.26	3.61 (1.09) ES: 0.19		
Feldman et al. (2010)							T: 3.56 (0.85) NT: 3.65 (0.74)	T: 3.37 (0.56) ES: -0.26 NT: 2.74 (0.68) ES: -1.27	T: 3.51 (0.87) ES: 0.19 NT: 2.83(0.75) ES: 0.12			
Bosquet Enlow et al. (2009)		135.9 (11.5)	147 (12.4) ES: 0.91	145.3 (10.7) ES: -0.14								

(continued on next page)

Table 1 (continued)

Author, Year	HR Baseline	HR Play #1	HR SF #1	HR Play #2	HR SF#2	HR Play#3	RSA Baseline	RSA Play#1	RSA SF #1	RSA Play#2	RSA SF#2	RSA Play#3
Ham and Tronick (2009)		143.57 (9.97)	151.67 (11.93) ES: 0.71	Not printed			3.68 (0.85) 3 Months: AFA: 3.42(0.95) EA: 3.43(1.03) 6 Months: AFA: 3.78(0.95) EA: 3.57 (0.80) 12 Months AFA:4.00 (1.00) EA: 3.41 (0.88) 2.82 (.75)	3.50 (0.93) ES: -0.20	3.38 (0.96) ES: -0.13 Change in RSA from baseline to SF episode: 3 months: AFA: 0.35 (0.90) EA: 0.66 (0.75) 6 months: AFA: -0.19 (0.39) EA: 0.39 (0.76) 12 months: AFA: 0.19 (0.60) EA: 0.12 (0.70)	3.44 (0.95) ES: 0.06		
Moore et al. (2009)												
Propper et al. (2008)												
Moore and Calkins (2004)	HP: 410.69 (30.64)	HP: 409.33 (28.76) ES: -0.05	HP: 393.35 (31.54) ES: -0.53	HP: 405.43(32.60) ES: 0.37						2.84 (0.76) ES: 0.03	2.64 (0.71) ES: -0.27	2.92 (0.89) ES: 0.34
Hailey and Stansbury (2003)		146.60 (11.92)	148.79 (11.60) ES: 0.18	146.55 (12.07) ES: -0.19	151.15 (12.21) ES: 0.38	149.11 (13.15) ES: -0.16						
Bazhenova et al. (2001)	HP:450 (28)	HP: 421 (30) ES: -0.99	HP: 446 (40) ES: 0.70				Baseline: 3.26 (-9) Picture Attention Task (PA): 3.77 (1) ES: 0.53	Toy Attention Task: 3.73 (1) ES from PA: -0.04	3.28 (0.8) ES: -0.49	3.77 (1) ES: 0.54		
Weinberg and Tronick (1996) 6 months		138.20 (no SDs)	143.88	139.91				3.165 (no SDs)	3.034			3.237

Table 2
Description and statistics of studies included in meta-analysis.

Study and measurement categories				Play/Pre-SF		SF		Post-SF		Hedges g		
Publication	UOM#	SES	RSA method	Age +	N	mean (SD)	N	mean (SD)	N	mean (SD)	SF	Re-union
Bazhenova et al. (2001)	lnRSA	middle class	Porges	5	41	3.73 (.1)	41	3.28 (.8)	41	3.77 (.1)	-0.49	0.54
Moore and Calkins (2004)	lnRSA	middle class	Porges	3	61	2.84 (.8)	60	2.64 (.7)	62	2.92 (.9)	-0.27	0.35
Feldman et al. (2010)	lnRSA	middle class	Porges	6	33	3.65 (.7)	33	2.74 (.7)	33	2.83 (.8)	-1.27	0.12
Weisman et al. (2012)	lnRSA	middle class	Porges	5	35	3.45 (.8)	35	3.25 (.8)	35	3.63 (1.1)	-0.26	0.39
Stewart et al. (2013)	lnRSA	middle class	Porges	6	75	3.25 (.7)	75	3.14 (.9)	75		-0.13	
Montirosso et al. 2014	RSA	middle class	Porges	4	78	62.1 (124.3)	78	54.14 (74.7)	78	144.09 (519.5)	-0.08	0.24
Suurland et al. (2016) (low-risk)**	lnRSA	middle class	Grossman	6	69	3.36 (.4)	68	3.14 (.6)	67	3.28 (.5)	-0.46	0.26
Sharp et al. (2012)	lnRSA	mixed SES	Porges	7.25	257	3.32 (.8)	253	2.83 (.8)	247	3.35 (1)	-0.61	0.6
Ritz et al. (2012)* (uncorrected)	lnRSA	mixed SES	Grossman	6	23	1.47 (.2)	23	1.26 (.3)	23	1.35 (.3)	-.80	0.31
Quigley et al. (2016)	lnRSA	mixed SES	Porges	6	94	3.57 (1)	91	3.4 (1)	89	3.47 (1)	-0.17	0.07
Busuito and Moore (2017)	lnRSA	mixed SES	Porges	7	50	3.65 (1.1)	51	3.4 (.9)	48	3.61 (1.1)	-0.25	0.21
Bush et al. (2017)	lnRSA	high risk	MindWare	6	63	4.27 (1)	60	3.99 (1.2)	44	4.21 (1.2)	-0.25	0.18
Suurland et al. (2016) (high-risk)**	lnRSA	high risk	Grossman	6	40	3.42 (.3)	38	3.27 (.4)	38	3.13 (.6)	-0.38	0.27
Conradt and Ablow (2010)	RSA	high risk/low SES	Grossman	5	91	0.02 (.01)	91	0.01 (.01)	91	0.01 (.01)	-0.28	0

#UOM; Units of Measurements. +Age is in months.

* Standard unadjusted values were used for comparison.

** Low-risk and high-risk samples used in Suurland et al. (2016) are presented.

preceding episodes of the SFP. Hedges' g generally uses the same definition of effect sizes as Cohen's d, with 0.2 considered a small effect, 0.5 a medium effect, and 0.8 a large effect (Cohen, 1988).

RSA responses differ across studies

As Table S1 indicates, multiple studies showed that RSA decreases from baseline/play to the SF episode (Bazhenova et al., 2001; Bosquet Enlow et al., 2014; Bush et al., 2017; Busuito & Moore, 2017; Conradt & Ablow, 2010; Ham & Tronick, 2006; Moore & Calkins, 2004; Moore et al., 2009; Suurland et al., 2016; Weinberg & Tronick, 1996), and although RSA may increase after the SF to baseline/play levels of RSA (e.g., Ham & Tronick, 2006; Moore & Calkins, 2004; Weinberg & Tronick, 1996; Weisman, Zagoory-Sharon, & Feldman, 2012), some studies found that RSA did not differ between the SF episode and reunion (Bush et al., 2017; Conradt & Ablow, 2010; Montirosso et al., 2014; Suurland et al., 2016) or rises in reunion but does not reach the level of RSA in play/baseline (Bosquet Enlow et al., 2014). At least four studies reported that roughly half of the infants decreased RSA during the SF episode (called "suppressors" (S)) while others failed to decrease or actually increased RSA during the SF episode (called "non-suppressors" (NS)) (Bazhenova et al., 2001; Montirosso et al., 2014; Moore & Calkins, 2004; Provenzi et al., 2015; note that Provenzi et al., 2015; Montirosso et al., 2014 are drawn from/use the same sample although sample size does differ). Two other studies, (also drawn from the same sample) found that although RSA decreased in the SF episode, RSA was not significantly different in reunion from play or the SF episode suggesting heterogeneity in results (Busuito & Moore, 2017; Moore, 2010). Accordingly, PNS results have not all been consistent.

Heart Rate/Heart Period patterns differ across the SFP studies

Generally, studies report that infant HR increases (or HP decreases) during the SF episode (e.g., Bazhenova et al., 2001; Bosquet Enlow et al., 2009; Gunning, Halligan, & Murray, 2013; Haley et al., 2006; Haley & Stansbury, 2003; Ham & Tronick, 2006; Moore & Calkins, 2004; Moore et al., 2009; Suurland et al., 2016). Some studies reported that HR decreased during the reunion (Bazhenova et al., 2001 (HP increases during "social interaction"); Gunning, Halligan & Murray, 2015 (for non-irritable infants); Haley et al., 2006 (during reunion-1); Haley & Stansbury, 2003 (during reunion-1); Moore & Calkins, 2004; Weinberg & Tronick, 1996). Other studies reported no decrease in infant HR during reunion (Bosquet Enlow et al., 2014; Conradt & Ablow, 2010; Gunning et al., 2013; Ham & Tronick, 2006; Moore et al., 2009). In Moore et al. (2009) overall, HP decreased for all infants from baseline to reunion.

Four of the studies that measured HR/HP did so in a modified SFP involving two repeated still-face exposures. Ritz et al. (2012) reported that HR increased during the SF episodes but did not return to play levels during either reunion episode. Drawing from the same participant sample but with a larger sample size, Bosquet Enlow et al. (2014) found that HR remained elevated after the first SF episode. Haley and Stansbury (2003) found that some infants actually increased HR from SF-2 to reunion 2. Haley et al. (2006) found that HR increased from play to SF-I, decreased from SF-I to reunion I, and increased from reunion I to SF-II.

In sum, HR usually increased during the SF episode but did not always decrease during reunion.

Exploring differences in measuring ANS within SFP studies

Next, we review a number of reasons why different studies may produce different results.

Various calculations of RSA are used in the SFP.

As indicated by Table S1, most of the studies in this review used Porges' (1985, 1995) algorithm to calculate RSA. Several other studies used a time-domain peak-valley approach thought to be correlated with the frequency domain methods (Goedhart et al., 2007; Grossman et al., 1990). One study used spectral analysis with an algorithm by Berntson, and another reported using spectral analysis with an algorithm in Chart (version 4.2) software. We examined whether these different calculations of RSA functioning might produce particularly strong or weak effect sizes for the change in ANS response between SFP episodes. As indicated in Table 1, using the peak-valley method, Ritz et al. (2012) produced an exceptionally high effect size for change from play to the SF ($g = -1.45$ (controlled for respiration)). The other two studies that used the peak-valley method reported more modest effect sizes (Suurland et al., 2016 ($g = -0.38$ for high-risk sample/ -0.46 for low-risk); Conrads & Ablow, 2010 ($g = -0.28$)), which may be related to their larger sample sizes ($n = 121$ and $n = 91$, respectively), relative to Ritz et al. (2012) ($n = 23$). Moreover, at least one study that used the Porges method, Feldman, Singer, and Zagoory (2010), reported an effect size of $g = -1.27$, comparable to Ritz et al. (2012). A number of other studies that used Porges method also produced moderate to high-moderate effect sizes, (Tibu et al., 2014 ($g = -.77$); Bazhenova et al., 2001 ($g = -0.49$)). Accordingly, both Porges method and the peak-valley method have produced large and moderate effect sizes.

Note that studies also use different time epochs, ranging from 5 to 30 s, and variations in how they combined epochs for calculating RSA baseline and reactivity (see Table S1). It was not possible to determine if epoch length made any difference for study findings because some studies did not report epoch lengths for scoring, and those that did had a variety of other differences in methodology and sample size, limiting comparability.

Different studies use different methods of calculating RSA reactivity.

A variety of different methods were used to calculate RSA reactivity (see Table S1). For example, some studies computed reactivity scores by comparing mean RSA in baseline to mean RSA during each SFP episode (e.g., Busuito & Moore, 2017; Moore & Calkins, 2004; Moore et al., 2009). Acknowledging that most studies compared baseline RSA to RSA during subsequent SFP episodes, Moore (2010) argued that "...because the SFP presents a series of distinct but contiguous social contexts, the most relevant measure of RSA reactivity was change from the preceding episode" (Moore, 2010, p. 6). Some studies compared mean RSA in baseline with the mean RSA in each SFP episode, as well as mean RSA in preceding episodes (e.g., Bazhenova et al., 2001; Busuito & Moore, 2017). Others computed a RSA reactivity ratio of mean episode RSA/mean baseline RSA (Montirosso et al., 2014; Provenzi et al., 2015 (same sample)). Suurland et al. (2016) computed reactivity by subtracting the second minute of preceding episodes (i.e., play – still-face). Two studies from the same sample also calculated a latent variable that constituted baseline RSA (Sharp et al., 2012; Tibu et al., 2014). To leverage their modified SFP that included two SF episodes, but adjust for the fact that some infants terminated after the first SF episode, Bush et al. (2017) calculated both a "First SF RSA reactivity score" as well as a "Last SF RSA reactivity score" by subtracting the average response during the last available of the two SF episodes in which the infant had three or more scoreable 30-second epochs from the play episode. Accordingly, although reactivity was most often calculated by comparing baseline measures to the remaining SFP episodes, some studies used different methods, not only comparing preceding episodes, but also using more novel approaches. These differences in calculations might be important because, for example, if baseline and play differ significantly, reactivity scores and effect sizes could also differ depending on whether RSA in the SF is compared to RSA in baseline or play.

Baseline measures lack standardization

Baseline RSA has been defined as a quiet alert state (Bar-Haim, Marshall, & Fox, 2000; Porges, 2007) and is viewed as an indicator of the infant's ability to engage with the environment, positive or negative (Beauchaine, Katkin, Strassberg, & Snarr, 2001; Conrads, Measelle, & Ablow, 2013; Propper, 2012). Yet, the studies examined herein acquired baseline measures under a variety of conditions. Some studies used the play episode before the SF as a baseline (Feldman et al., 2010; Ham & Tronick, 2006, 2009; Pratt, Singer, Kanat-Maymon, & Feldman, 2015; Stewart et al., 2013; Weinberg & Tronick, 1996). For the play episode, the mother is instructed to play with the child as she would at home, but some mothers play with their child quietly and gently while others are quite activating and use vigorous physical play. Accordingly, it is unclear how "quiet" various children could be when playing with their mothers, suggesting that infant movement and verbalization may affect or confound the RSA baseline measure.

Other studies, however, used a pre-SFP episode as a baseline measure. These baseline paradigms differed considerably. As indicated in Table S1, baseline RSA measures appear to range from one to four minutes, with the majority using two minutes. In terms of activity, baseline protocols may, for example, instruct mothers to place the infant in a seat, sit in the chair in front of the infant and read instructions for 3 min while infant baseline measures are taken (Busuito & Moore, 2017; Moore, 2009, 2010). Other studies allow mothers to actively play with the infant just before taking baseline measures (Ritz et al., 2012). Some studies have the infant watch a video while sitting in their mother's lap (Conrads & Ablow, 2010; Ostlund, Measelle, Laurent, Conrads, & Ablow, 2017 (same sample)) or while lying down on a blanket (Suurland et al., 2016). Some studies provide the infant with a three-minute period of "adaptation" to the environment during which baseline measurements are taken for varying amounts of time (Moore, 2009, 2010; Provenzi, Giusti et al., 2016). In a series of studies from the same research group, the infant was allowed to calm after application of the electrodes (approximately 5 min), then seated in the mother's lap (facing away from the mother (Holochwost et al., 2014) or in a seat (Propper et al., 2008)), and the mother was instructed not to interact (Gueron-Sela et al., 2017; Moore et al., 2009; Quigley,

Moore, Propper, Goldman, & Cox, 2016) or to provide toys to the infant (Moore et al., 2009; Propper et al., 2008; Quigley et al., 2016) to minimize stimulation. One research group, Sharp et al. (2012); and Tibu et al. (2014), administered two measures prior to the SFP, the “helper-hinderer” (HH; over an average of 3.74 min infants are assessed to determine if they prefer toys that are “helpers” or “hinderers”) and the “novel toy task” (NT; the infant sits on the mother’s lap for 2 min and given a toy), to capture the infant’s attention in a non-stressful manner; they then calculated a latent baseline RSA variable with all episodes (HH, NT, play, still-face and reunion) in their model.

In sum, baseline RSA measures vary considerably in this literature with little understanding of how those variations impact findings. In particular, studies that permit infants to sit on their mothers’ laps prior to the SF episode introduce confounds in light of evidence that sitting on mother’s lap enables contagion of affect that could impact infant RSA (Waters, West, & Mendes, 2014).

Instructions to parents or researchers in play and reunion episodes might differ

Instructions to parents in the play or reunion episodes differ with respect to touch or interactions with the infant. For example, one study prohibited the use of toys and instructed parents to simply play with their babies (Conradt & Ablow, 2010), whereas another study gave parents a toy with which to play with their infants but directed parents not to touch the babies (Haley & Stansbury, 2003). Some studies told parents to engage in “interactive play” with their infants (Ham & Tronick, 2006); or to just play with their babies as they “normally” would without further instruction (e.g., no mention of toys) (e.g. Moore, 2010; Moore & Calkins, 2004). One study expressly told researchers used in the SFP not to touch the infant in reunion (Bazhenova et al., 2001). In summary, the extent to which parents touched their infants in most of these studies is inconsistent or unknown, and the use of toys versus only parent interactions during play varied. As with lap-sitting, parental touching could also enable contagion of affect (Waters et al., 2014) or other factors that affect infant ANS (Feldman et al., 2010).

Analyzing infants by individual or dyadic characteristics may influence ANS outcomes

Infants may be categorized, organized or assessed using different criteria across studies, as well. For example, Ham and Tronick (2006) divided infants into four groups based on behavior: (1) Recovered ($n = 4$) infants who protested during SF > 25% of time but reduced protesting during the reunion episode; (2) Stably Low ($n = 5$) infants who did not protest for > 25% of time in SF or reunion; (3) Dysregulated ($n = 2$) infants who protested during SF > 25% time and continued to increase in the reunion; and (4) Cry in reunion only ($n = 1$) infants who only protested during the reunion. Although infants in the “recovered” group showed the greatest increase in RSA from SF to reunion, the authors suggested those in the “stably low” group may represent the most “resilient” response because they remained calm through the SFP, SC measurements suggested that their mothers were also calm, and mothers in these dyads were responsive to their infants. Two major limitations of the study, however, were the small sample size and that the paper did not report actual RSA values. The following section presents some of the other ways infants are categorized, organized or assessed in the SFP/ANS literature.

Not all infants suppress RSA during challenge: Suppressor (S) vs. non-suppressor infants (NS)

As noted earlier, some studies distinguish between infants based on RSA reactivity and recovery and categorized them as either “suppressors” (S) who demonstrate PNS withdrawal during challenges, suggesting a stress response, or “non-suppressors” (NS). This distinction is of importance as some studies found that approximately half or more than half of the infants were NS (Bazhenova et al., 2001 (45%); Montirosso et al., 2014 (50%) and Provenzi et al., 2015 (55.3%) (shared sample); Moore & Calkins, 2004 (53%)). Categorizing infants across the SFP in this manner, however, did not provide strictly consistent patterns of PNS findings. Provenzi et al. (2015) and Montirosso et al. (2014) did not find significant differences in RSA across the SFP overall. Provenzi et al. (2015) did report that S infants increased RSA in reunion while NS infants showed no change in RSA from the SF to the reunion episode. Moore and Calkins (2004) found significant differences in RSA between episodes in their entire sample (e.g., lower RSA in SF compared to play) but also found significant differences in RSA when the infants were divided into S and NS groups. S infants showed an increase in RSA from baseline to play, a decrease in RSA from play to SF episode, and an increase in RSA from play to reunion. The NS infants, however, showed a significant decrease in RSA during play and reunion, but showed no differences in RSA during the SF episode compared to S infants. The S infants had higher RSA during play and reunion than NS infants. Although Bazhenova et al. (2001) subjected infants to five different episodic conditions (baseline, toy attention, picture attention, SF episode, social interaction and a second picture attention) and involved a stranger instead of a parent, they also reported that S infants increased RSA in the social interaction episode, while NS infants failed to increase RSA between the SF and social interaction episodes.

Parental behavior may influence ANS function

A number of SFP studies examined the association between infant ANS activity and a range of measures of parental behaviors variously referred to as parental “sensitivity” (Bosquet Enlow et al., 2014; Conradt & Ablow, 2010; Gunning et al., 2013; Holochwost et al., 2014; Moore et al., 2009; Propper et al., 2008); parental “responsiveness” (Haley & Stansbury, 2003) “dyadic” or mother-infant synchrony (Moore & Calkins, 2004; Provenzi et al., 2015), and maternal touch (Feldman et al., 2010; Sharp et al., 2012). Most studies evaluated these behavioral constructs within the SFP (Bosquet Enlow et al., 2014; Conradt & Ablow, 2010; Feldman et al., 2010; Gunning et al., 2013; Haley & Stansbury, 2003; Ham & Tronick, 2006, 2009; Moore & Calkins, 2004; Provenzi et al., 2015) but at least one set of studies drawing from the same sample evaluated parental sensitivity outside of the SFP (Holochwost et al., 2014; Moore et al., 2009; Propper et al., 2008; Quigley et al., 2016). The most common divisions found in the reviewed studies were to separate infants into those with “sensitive” versus “insensitive” parents using a variety of approaches to determine sensitivity (see

Table S1). For example, [Bosquet Enlow et al. \(2014\)](#) found that infants of mothers who were insensitive during play showed higher distress and lower levels of RSA through the SFP. [Conradt and Ablow \(2010\)](#) found that higher sensitivity in reunion predicted lower HR in all three SFP episodes, even after adjusting for both infant temperament and movement in the SFP. [Moore et al. \(2009\)](#) reported that maternal sensitivity was associated with slower infant HR during the SF episode. Although [Moore et al. \(2009\)](#) found no main effects of sensitivity, they did report a significant interaction between sensitivity and SF episode; the infants of mothers classified as *highly* sensitive showed lower RSA in the reunion than other infants and a decrease in RSA from baseline to reunion, suggesting that these infants had a more difficult time recovering. Later, the same research group reported that breastfeeding infants also showed lower RSA from baseline to reunion compared to non-breastfed babies, but maternal sensitivity was ultimately insignificant in their model suggesting that breastfeeding may exert its own independent effect on infant ANS function ([Quigley et al., 2016](#)).

Other studies examined constructs related to parental sensitivity. For example, [Moore and Calkins \(2004\)](#) reported that infants who did not show RSA withdrawal in the SF episode also showed lower infant-mother synchrony (i.e., the correlation between mother and infant affect without regard to whether affect matches). In addition, they found that infants in dyads showing lower levels of matched affect had greater decreases in RSA in reunion. Two studies drawing from the same sample ([Busuito & Moore, 2017](#); [Moore, 2010](#)) found that infant RSA was not associated with maternal-infant synchrony. Moore and colleagues did report, however, that infants exposed to parental conflict showed lower mean RSA across the SFP, and that these infants actually withdrew RSA during the *play* episode ([Moore, 2010](#)). [Busuito and Moore \(2017\)](#) later reported that the association between high parental conflict and lower RSA reactivity was mediated through lower parental “flexibility” (i.e., balanced variability in dyadic states) assessed in the reunion episode. In a recent study, [Ostlund et al. \(2017\)](#) examined the physiological synchrony, called “attunement,” between maternal and infant RSA in the SFP. Although the study detected no mother–infant attunement in play, they did find that during the first half of the reunion episode maternal RSA increased while infant RSA decreased. During the second half of reunion, however, maternal RSA decreased and infant RSA increased suggesting that mothers were preparing physiologically to provide support to the stressed infant. [Provenzi et al. \(2015\)](#) found that reparation rate (i.e., the extent to which dyads repair mismatched states) during play was lower in dyads with NS infants. Higher reparation rates during the play episode were associated with less negative emotionality for S infants. Interestingly, there were no main effects of reparation rate or RSA classification on infant negative emotionality during the SF episode. The amount of time mothers and infants spent in matched states overall was highest in dyads with S infants compared to dyads with NS infants. Finally, [Haley and Stansbury \(2003\)](#) found that infants of parents who showed higher levels of contingent responsiveness to infant vocalizations or facial expressions decreased their HR from SF 2 to reunion 2 compared to infants of less responsive parents.

Three studies assessed the importance of some form of maternal touch. [Sharp et al. \(2012\)](#) found that higher maternal depression was associated with decreasing infant vagal withdrawal and with increasing infant negative emotionality only in infants whose mothers self-reported low levels of maternal stroking of infants. [Feldman et al. \(2010\)](#) found that infants whose mothers were not allowed to touch them during the SF showed higher RSA suppression than infants who were touched. Moreover, among infants in the touch condition, RSA almost returned to baseline levels in reunion, whereas RSA remained the same among infants in the no-touch condition. Higher touch frequency was also related to higher infant RSA during play, whereas touch “mysynchrony” (mother touches while infant looks away) was associated with lower infant RSA during play. The same research lab later confirmed the finding that infants showed higher RSA suppression when mothers are not allowed to touch them ([Pratt et al., 2015](#)).

In sum, these studies appear to identify some relationship between SFP, RSA function, and parental behaviors. Most of the studies found that infants of insensitive or nonresponsive parents had lower RSA during reunion than the infants of more sensitive or responsive parents suggesting poorer vagal regulation. Differing outcomes might be explained by a number of factors including different ways studies assessed sensitivity and other, related parental behaviors (e.g., parental touch) in the play episode, the reunion episode or both, whether parental behavior was assessed within or outside of the SFP (see Table S1), and whether a study had two SF episodes instead of one and/or had different baseline measures.

Environmental sensitivity to context

Developmental-evolutionary theories have suggested that children differ in their sensitivity to both positive and negative environmental experiences ([Boyce & Ellis, 2005](#); [Bush & Boyce, 2016](#); [Ellis, Boyce, Belsky, Bakermans-Kranenburg, & van Ijzendoorn, 2011](#)), and that individual variation in genotype is a key factor in determining susceptibility to environmental effects. Multiple studies in this review suggest that gene-environment interactions influence RSA reactivity or that RSA levels may be associated with environmental sensitivity. For example, hypothesizing that high RSA may be associated with greater environmental sensitivity, [Holochwost et al. \(2014\)](#) compared infants that showed high RSA throughout the entire SFP to those who showed lower RSA. They found that infants with high RSA in play and reunion at 6 months of age, and who had mothers who showed high levels of negative intrusiveness, were more likely to be classified as disorganized during the strange situation (SS) administered at 12 months of age. No relationship between maternal intrusiveness and disorganization was detected for infants with low RSA in play and reunion. It may also be of interest to note, however, that high indices of RSA and maternal sensitivity did not predict secure attachment at 12 months of age. Similarly, [Gueron-Sela et al. \(2017\)](#) reported that maternal depression was positively associated with infant sleep problems at 18 months only for those infants who showed high baseline RSA during the SFP at 3 and 6 months of age compared to infants with low RSA-B. Contrary to the BSC hypothesis, however, infants with high baseline RSA did not have lower levels of sleep problems when mothers reported low maternal depression suggesting that low maternal depression did not necessarily mean the infant was living in a positive environment.

[Pratt et al. \(2015\)](#) reported that highly negative infants with poor dyadic synchrony showed poor vagal recovery in reunion, but highly negative infants with high dyadic synchrony showed high levels of RSA during reunion, comparable with calm infants

suggesting that infant negativity in that study may reflect sensitivity to the environment.

Finally, [Propper et al. \(2008\)](#) reported that infants who carried the dopamine receptor's risk allele (DRD2) showed lower RSA withdrawal at 3 and 6 months of age during the SF than those infants who were not DRD2 carriers. By 12 months of age, however, the infants who carried DRD2 but who also had sensitive mothers displayed the same level of RSA withdrawal during the SSP as those infants without DRD2. [Propper et al.'s \(2008\)](#) finding is in line with other evidence that the DRD2 polymorphism may convey "sensitivity to context" rather than merely "risk" for problems ([Belsky & Beaver, 2011](#); [Belsky, Suppli, & Israel, 2014](#)). Further investigation of this type of "biological sensitivity to context" (BSC) within studies of ANS regulation in the SFP may be fruitful.

Demographic characteristics may influence infant ANS function

Different ANS responses may be associated with SES

Despite increasing research suggesting that children living in poverty, low SES or high-risk households suffer negative health outcomes, there are few studies of ANS in the SFP for infants living under these conditions ([Propper, 2012](#)). In this review, the majority of studies included middle class samples (see Table S1 for details on SES/risk). Some of the studies with mixed SES samples did not find ANS differences based on SES (e.g., [Moore, 2010](#); [Moore et al., 2009](#) (different samples)). In fact, [Moore et al. \(2009\)](#) noted that infant's vagal tone patterns in her diverse sample (approximately half was low-SES) were the same as those identified in studies with infants from middle class families (e.g., [Bazhenova et al., 2001](#); [Ham & Tronick, 2006](#); [Moore & Calkins, 2004](#); [Weinberg & Tronick, 1996](#)). [Quigley et al. \(2016\)](#), on the other hand, reported that infants of higher income families had lower RSA in reunion.

Two studies comprised of primarily low-SES samples experiencing adversity generally found that although RSA decreased in the SF episode, RSA failed to rise in reunion ([Bush et al., 2017](#); [Conradt & Ablow, 2010](#)). A recent study with a high-risk (a composite measure of neuropsychiatric and psychosocial measures) and low-risk sample, reported that RSA decreased during the SF episode and increased during reunion among low-risk infants; high-risk infants showed a decrease in RSA between SF and reunion ([Suurland et al., 2016](#)).

Even among studies that involved primarily low-risk, middle class participants, results differed. For example, [Weinberg and Tronick \(1996\)](#) reported that infants showed lower RSA during the SF episode with no significant differences between play and reunion. [Feldman et al. \(2010\)](#) confirmed the classic SF effects of lower RSA in the SF episode, but found RSA only increased during reunion among those infants whose mothers were allowed to touch them during the SF.

In sum, in the middle-class samples, unless the study divided infants between S and NS, researchers generally reported a decrease in RSA in the SF episode, and usually an increase in reunion, although not always. High-risk or low-SES studies generally found that RSA and HR in reunion did not differ from that of the SF episode or RSA decreased even further.

Few studies report sex differences in infant ANS function

It is unclear from the broader ANS literature whether there are sex differences in ANS functioning. Some studies report that girls have lower resting RSA and suppression than boys ([El-Sheikh, 2005](#); [van Dijk, van Eijsden, Stronks, Gemke, & Vrijkotte, 2012](#)), or that girls have higher resting RSA than boys ([Fabes, Eisenberg, Karbon, Troyer, & Switzer, 1994](#); [Gordis, Feres, Oleszki, Rabkin, & Trickett, 2010](#)), or that girls and boys show no differences ([Alkon et al., 2003](#); [Wagner, Propper, Gueron-Sela, & Mills-Koonce, 2015](#)). Accordingly, it was of interest to examine the patterns of findings for sex differences for ANS during the SFP for this review.

Reportedly, there were no sex differences in vagal regulation in the SFP before 2009 when [Moore and colleagues](#) found that 6-month old male infants demonstrated a higher baseline RSA than female infants ([Moore, 2009](#)). Since then one study found that that lower birth weight was associated with higher vagal reactivity during the SFP in girls but not boys, and that prenatal maternal anxiety was associated with less vagal withdrawal in boys but not girls ([Tibu et al., 2014](#)), suggesting infant sex may moderate associations with infant ANS. [Haley and Stansbury \(2003\)](#), however, found that boys had slower HR than girls, and girls had faster HR than boys during SF II and reunion II. Among mothers who were classified as high frequency drinkers, girls also showed higher HR than boys in the SFP ([Haley et al., 2006](#)). Studies have also reported no significant effects for infant sex on HP ([Moore et al., 2009](#)). Though requiring replication, these sets of findings point the possibility that infant sex may moderate associations of other factors with infant ANS during the SFP.

ANS patterns may differ with age during the first year of life

Researchers have reported that resting levels of RSA increase during the first year of life and HR decreases with age ([Alkon et al., 2006](#); [Alkon, Boyce, Davis, & Eskenazi, 2011](#); [Bar-Haim et al., 2000](#); [Propper & Moore, 2006](#)). Accordingly, investigating whether age might be associated with RSA was of some interest. Unfortunately, age variability across the studies was limited; although infants ranged from eight weeks of age ([Stoller & Field, 1982](#)) to 8.5 months of age ([Moore, 2010](#)), most conducted the SFP at approximately six months of age.

The S versus NS studies generally involved infants under five months of age (Table S1), which raises the question of whether it is developmentally normative for approximately half of infants under the age of five months to fail to show RSA suppression during the SF episode. Unfortunately, studies involving older children did not report on infants RSA suppression rates making it difficult to address these questions. One other cross-sectional study reported that infant age was positively associated with RSA change from baseline to the play episode in 6–8.5 month-old infants (older infants showed less RSA withdrawal) ([Moore, 2009](#)).

[Propper et al. \(2008\)](#) measured the stability of infant RSA during the SFP at ages 3 months and 6 months, and at 12 months in the Strange Situation (SS). They reported that baseline RSA, but not HP, at 3 and 6 and 12 months were correlated, but RSA reactivity was not correlated across time. Their age-related SFP findings are somewhat consistent with those of [Alkon et al. \(2011\)](#), which found

stability in the HR, RSA and PEP assessed under resting and challenge conditions (not the SFP), but no stability in ANS reactivity among Latino children from ages 6 to 60 months.

Though rare, some studies do detect racial/ethnic differences in ANS

Outside of the SFP literature, some studies report no difference in ANS function between racial/ethnic groups among infants and children (Alkon et al., 2011; Wagner et al., 2015). Yet, a meta-analysis of 17 studies (33% of which included child samples), reported that African Americans had higher resting HRV than European Americans (Hill et al., 2015). Accordingly, it was of some interest to determine whether the SFP studies reported ethnic differences in ANS function.

Table S1 details the race/ethnicity of samples within this review. Only a few of the studies reviewed reported significant differences in ANS by race/ethnicity. For example, in one sample used across several publications, African American infants were found to have higher RSA than European American infants in the SFP (Holochwost et al., 2014; Moore et al., 2009; Quigley et al., 2016), showed reduced RSA withdrawal at 3 months and 6 months during the SF episode, and had higher average baseline RSA at 12 months than European American infants (Propper et al., 2008). It may be of some interest to note that the effect of race/ethnicity on RSA withdrawal disappeared at 12 months once the study controlled for maternal sensitivity and a genetic risk allele. Other studies found no difference in RSA reactivity scores between African-American and European American infants (Moore, 2009, 2010 (same sample)).

Maternal clinical conditions may be associated with infant ANS

Maternal depression, anxiety, stress and trauma may impact infant ANS. Research has shown an association between maternal anxiety and lower infant resting RSA (Field et al., 2002, 2003; Propper & Holochwost, 2013; Propper, 2012). Among the studies reviewed here, Ostlund et al. (2017) reported that mothers with higher anxiety had infants with higher RSA during reunion. Tibu et al. (2014) found that higher levels of prenatal maternal anxiety at 32 weeks of pregnancy, but not contemporaneous maternal anxiety, were associated with less vagal withdrawal in boys during SF.

Maternal “stress” has also been studied. One study found that high maternal lifetime trauma exposure and elevated perinatal trauma (over the course of pregnancy and the postnatal period) were associated with reduced infant HR recovery in reunion (Bosquet Enlow et al., 2009). More recently, Bush et al. (2017) reported that maternal reports of higher counts of stressful life events (SLE) during pregnancy and higher postnatal perceived stress were associated with higher infant RSA reactivity to the first SF episode. Moreover, higher prenatal SLE and both prenatal and postnatal perceived stress were associated with higher infant reactivity in whichever SF episode the infant completed last (of two possible episodes; some infants terminated prior to the second SF due to distress).

Maternal depression has also shown some association with infant ANS function (Propper & Holochwost, 2013). One study in this review found that infants of mothers placed into a depressed category had longer HP in the reunion episode and less change in HP from baseline to reunion than infants of non-depressed mothers (Moore & Calkins, 2004). Dyads in the depressed group also showed less synchrony in play and higher matched affect in reunion than the remaining dyads. Another study found that increasing maternal prenatal depression was associated with decreasing vagal withdrawal and increasing infant negative emotionality only in the infants of mother who engaged in low “maternal stroking” (Sharp et al., 2012). Of note, maternal depression may be less salient than other clinical diagnoses, given that Ostlund et al. (2017) found maternal depression was not associated with infant RSA after adjusting for the significant effect of maternal anxiety, and Bush et al. (2017) found maternal depression was not associated after adjusting for the effects of stress and SLE (depression was dropped from their models).

In summary, there were few and mixed findings with respect to the impact of maternal depression or anxiety on ANS functioning during the SFP. There was some indication among the studies that higher levels of prenatal or postnatal maternal stress or trauma is associated with higher infant ANS reactivity and a failure to recover in reunion.

Maternal substance abuse may impact infant ANS. Few studies have examined the impact of maternal substance abuse on infant ANS, but some studies have reported that infants exposed to alcohol or opioids in utero demonstrate higher HR and/or lower vagal tone, or fail to show RSA withdrawal during stress (Fifer, Fingers, Youngman, Gomez-Gribben, & Myers, 2009; Jansson, Dipietro, Elko, & Velez, 2010; Oberlander et al., 2010; Propper & Holochwost, 2013). Although there were very few studies in this review that examined the association between substance use during pregnancy and infant ANS, one study did find that infants of mothers who reported “high drinking frequency” (HDF) showed higher HR, and increased negative affect in the SFP than infants of mothers in the “low drinking frequency” group (Haley et al., 2006). Mattson et al. (2013) found no differences in HR between cocaine exposed and non-exposed infants; however, the authors posited that the level of cocaine exposure may have been insufficient to detect differences between exposed (n = 44; only 8 had been heavily exposed to prenatal cocaine use) and non-exposed (n = 49) infants. Note that all the infants in Mattson et al. (2013) reported pre-natal exposure to tobacco, marijuana and alcohol; cocaine exposed infants did, however, have significantly higher exposure to marijuana and tobacco. Mattson et al. (2013) confirmed that HR increased during the SF episode and remained elevated from the SF to reunion.

Fathers are rarely represented in the literature. Unfortunately, only one study assessed fathers in the SFP. Weisman et al. (2012) administered oxytocin (OT) to fathers alone prior to the SFP and found that OT administration was associated with higher infant RSA during play, as well as increased infant salivary OT levels and social behavior, compared to infants in the placebo condition.

*Design and methodology: Comparing effect sizes**Factors such as sample size, maternal touch, and controlling for respiration may impact effect sizes*

Overall, there was a wide range of sample sizes in this review ranging from 12 to 270 dyads, with most studies having less than 100 (see Table S1). As Table 1 indicates, the two publications (using the same sample of participants) that showed some of the largest effect sizes also had the largest sample sizes, Tibu et al. (2014) ($g = -0.77$ (boys; play to SF)) and Sharp et al. (2012) ($g = -0.61$) (play to SF). Of note, those researchers subjected infants to two other tasks before the SFP, suggesting that infants might have been more stimulated before experiencing the SF episode. Bazhenova et al. (2001) had a smaller sample size but produced medium effect sizes ($g = -0.49$ (toy attention to SF); $g = 0.54$ (SF to reunion)); they also subjected infants to multiple measures before the SFP, and, as noted earlier, involved a stranger in the SFP. The other study that used strangers in the SFP produced smaller effect sizes even though it had a larger sample size than Bazhenova et al. (2001) (Stewart et al., 2013). Note that Stewart et al. (2013) did not indicate that infants were subjected to prior tasks just prior to the SFP.

Although Conradt and Ablow (2010), Ritz et al. (2012), and Suurland et al. (2016) all used the peak-valley method, Table 1 reveals that the largest effect size across those 3 studies was found for the smaller sample of Ritz et al. (2012) ($g = -1.45$ (play to SF1) (RSA corrected for respiration)). Overall, there were too few peak-valley studies to provide a meaningful comparison with the studies that used Porges (1985) algorithm, particularly considering that both Conradt and Ablow (2010) and Suurland et al. (2016) included high risk/low SES samples, which might confound the findings. Accordingly, it is not clear that either method produces larger effect sizes for change in RSA from play to the SF episode, and results across methods may be comparable.

As Table 1 shows, some of the largest effect sizes relating to RSA changes in the SFP was found in the study that explicitly directed some mothers to touch their infants during the SF episode and other mothers not to touch the infant (NT) (Feldman et al., 2010). The high effect size was found among the NT infants ($g = -1.27$). It is not clear why the no-touch condition would produce such a high effect size given that parents are usually told not to touch their infants in the SF episode and other studies have not produced such a large effect size. Note also that the study appeared to involve a low-risk sample and calculated RSA using the standard Porges (1985) algorithm. One difference between this study and others is that Feldman et al. permitted mothers to play with their infants for 3 min instead of the usual 2 min, which might have allowed the infants longer to become calmer in the lab environment creating a greater contrast to the stress reaction during the SF episode. Other possibilities include sample differences in culture, parenting, or other environmental influences. Additional research on the role of touch and play episode duration will be clarifying.

Table 1 also illustrates that some of the most consistently high effect sizes across studies were produced by the change in HR (or heart period) from baseline/play to the SF episode, (Bazhenova et al., 2001; Bosquet Enlow et al., 2009; Conradt & Ablow, 2010; Ritz et al., 2012; Stewart et al., 2013; Suurland et al., 2016), possibly because both the SNS and PNS contribute to HR and SNS effects were not assessed. The difference in effect size could be quite stark when comparing HR and RSA reactivity within the same study (e.g., Stewart et al., 2013 (effect size for: HP reactivity; $g = -0.86$; RSA reactivity; $g = -0.13$)).

The higher effect size found in Provenzi et al. (2015) among NS infants ($g = 0.60$; play to SF), is likely accounted for by the stratification of the sample by S and NS. The remainder of the studies essentially produced between small and medium effect sizes. Within these studies, the larger effect sizes were found within larger sample sizes ranging from 63 to 151 (Bush et al., 2017; Moore & Calkins, 2004; Provenzi et al., 2015; Suurland et al., 2016; Weisman et al., 2012) or in studies examining HR (Haley & Stansbury, 2003).

Although comparing effect sizes is challenging because of differences in methodology, in summary, increasing sample size, using strangers in the SFP instead of parents, subjecting infants to prior tasks before administering the SFP or controlling for respiration may result in greater effect sizes.

Repeated SF episodes increase reactivity and contribute to effect size

Five studies used a modified version of the SFP, with two SF and two reunion episodes, to enhance ANS reactivity (Bosquet Enlow et al., 2014; Bush et al., 2017; Haley & Stansbury, 2003; Haley et al., 2006; Ritz et al., 2012) (note that Bosquet Enlow et al. (2014) and Ritz et al. (2012) draw from the same sample but differ in terms of sample size). For example, Haley and Stansbury (2003) found a relationship between negative affect and HR in the second SF episode, but not during the first one, suggesting that behavioral and physiological systems may become more tightly coupled under conditions of greater and repeated stress. Bosquet Enlow et al. (2014) found that HR increased during the first SF episode but remained elevated during the entire SFP. Ritz et al. (2012) and Bush et al. (2017) found RSA progressively decreased in SF 1 and SF 2 for RSA, and Bush et al. (2017) reported that RSA in reunion never returned to play levels. Ritz and colleagues suggested that the extended reductions in RSA due to repeated exposure to stress may influence the infant's ability to recover from stress and may have clinical implications for infants in high risk environments. As noted above, Ritz et al. (2012) produced some of the largest effect sizes but whether that was because the study controlled for respiration, subjected infants to two SF episodes or was influenced by the small sample size is unclear (note that high effect sizes were produced during both SF episodes). Finally, researchers should use some caution when using the modified SFP because studies report losing more than a quarter of the participants by the time SF 2 is administered due to infant distress or fatigue (Bosquet Enlow et al., 2014; Bush et al., 2017; Haley & Stansbury, 2003; Haley et al., 2006; Ritz et al., 2012).

In sum, the modified SFP with two SF and reunion episodes has advantages and disadvantages. Researchers must be prepared to terminate the measure when infants become too distressed and lose possibly over a quarter of the sample. On the other hand, the second SF does seem to increase stress and provide higher reactivity which may help differentiate the infants at risk for negative outcomes later in life.

Similar outcomes in the SFP are produced whether researchers or parents are used.

Some studies used stranger/infant dyads rather than parent/infant dyads to test various aspects of Porges Polyvagal theory (Bazhenova et al., 2001; Stewart et al., 2013). These studies reported the same overall pattern of RSA withdrawal in the SF compared to play or preceding episodes. Similarly, the SF behavioral effect is found regardless of the identity of the adult participant (parent or stranger) (Mesman et al., 2009). Since Porges' theory proposes that the ANS provides support for the affective adjustments individuals must make when interacting with environmental stimuli broadly, it is not necessarily surprising that infants would show similar responses to the SFP with a stranger as with a parent. But these parallel results with both parents and strangers raise the question of whether and how the SFP informs researchers about specific parent-infant relationships or whether the classic SF effect reflects more general inborn infant interactional responses. Although it is possible that the relationship between parents and their six-month-old infants could influence the SFP response with strangers, accounting for this shared pattern of findings across types of individuals (Mesman et al., 2009), it is less likely that attachment relationships could be influencing the behavior of infants only hours old. Indeed, even hours-old neonates exposed to the SFP exhibit decreased eye contact, greater distress, and self-regulatory behaviors (Nagy, 2008), suggesting the existence of early inborn behavioral systems designed to eventually facilitate interaction with an attachment figure (Bowlby, 1968/1982). More likely, the behaviors assessed in the SFP reflect the emerging confluence of biological, genetic, epigenetic and environmental influences that remain to be understood.

Administering SFP at home vs. lab may produce similar results

The Mesman et al. (2009) meta-analysis excluded studies that administered the SFP in the home on the grounds that the familiar environment might influence infant behavior. Researchers working with high-risk or difficult-to-engage populations, however, may have to administer some measures at home, because travel to a hospital or university lab could be too challenging or prohibitive for some participants. Accordingly, a few studies reviewed here conducted the SFP in the home. Haley et al. (2006) conducted the SFP in the home (n = 38) or in a hospital lab (n = 17) and found no differences between the two groups in infant HR or cortisol reactivity. Bush et al. (2017) used a standardized blank visual screen to surround infants in both home and lab administrations and found no differences in infant RSA between the two settings. Despite the findings of these two studies additional research is necessary before concluding that administering the SFP in the home does not significantly impact ANS outcome, relative to administration in the lab.

Studies handle missing data in a variety of ways

Table S1 provides examples of some of the statistical approaches taken to deal with missing data resulting from equipment malfunctions, ECG artifacts, early termination due to infant distress, noncompliance or outliers. Most studies report excluding outliers, though determination of outliers was variable across studies, ranging between scores higher than 1 SD to those higher than 3 SD. Some studies excluded cases that required over a certain threshold of editing within the ANS scoring, ranging from those requiring more than 2 or 3% editing to data files that required more than 10% editing. A number of studies had missing data due to terminating due to infant protest (Bosquet Enlow et al., 2009, 2014; Bush et al., 2017; Conradt & Ablow, 2010; Haley & Stansbury, 2003; Haley et al., 2006; Moore et al., 2009; Moore, 2009, 2010; Ritz et al., 2012). For example, Moore et al. (2009) stated that it is "common" for 20% of cardiac data to be missing from each of the episode primarily because infants "...tend to be noncompliant" with the recording measures and opined that loss of this data is most likely due to "movement artifact." In Bosquet Enlow et al. (2014) 26% of infants terminated after reunion 1 and 11% terminated after SF 2, and maternal sensitivity during reunion 1 was found to be significantly greater among dyads who completed the entire SFP than those who failed to finish. These differences in outcomes between those infants who terminate early and those who complete the SFP may bias the study findings and underestimate the results since highly reactive infants are more likely to disengage and not participate in the full SFP. Researchers should be aware that as many as 30% of the sample might be lost because of these various factors (Bosquet Enlow et al., 2009, 2014; Holochwest et al., 2014; Moore, 2010) which should be accounted for and addressed in future research. As Table S1 indicates, the studies reviewed also adopted a broad array of statistical methods to deal with missing data, which may influence results of models.

Other considerations

Movement

Because physical movement may modify the ANS response to challenge, many researchers have emphasized the importance of controlling for motor activity (Bazhenova et al., 2001; Bosquet Enlow et al., 2014; Bush, Obradovic, Stamperdahl, & Boyce, 2011; Grossman & Taylor, 2007; Laborde, Mosley, & Thayer, 2017). One study reviewed here reported that movement was positively correlated with HR throughout the SFP (Conradt & Ablow, 2010), yet controlling for movement may be particularly challenging when working with infants. The literature suggests that, to date, there is no widely-accepted standardized strategy to parse out the effects of movement in models predicting infant RSA (Laborde et al., 2017).

Studies measuring SNS activity during SFP are limited: Skin Conductance (SC), T-wave amplitude (TWA) and pre-ejection period (PEP)

Our review found few studies that examined SNS activity during the SFP. For example, one study reported that SC increased across the SFP (Ham & Tronick, 2006), suggesting SNS activation. Another study from the same research team found that concordance between mother and infant SC in the reunion episode positively correlated with mother-infant behavioral synchrony (Ham & Tronick, 2009). A third study found that infant TWA did not vary across episodes of the SFP, but higher T-wave attenuation (SNS activation) during the SF episode was associated with maternal insensitivity during play and reunion (Bosquet Enlow et al., 2014).

Finally, a fourth study detected no differences in infant PEP across the SFP in the sample as a whole, but authors did find SFP effects on PEP depended upon risk status (Suurland et al., 2016); low risk infants demonstrated an increase in PEP during reunion while high risk infants showed a decrease in PEP. Moreover, the researchers reported that within the high-risk group an increase in risk factors was associated with greater decreases in PEP (reactivity) from play to the SF episode. Suurland et al. also reported that boys had lower PEP (SNS activation) than girls in all episodes of the SFP. In sum, this small set of SNS results is somewhat consistent with those assessing PNS response in the sense that they indicate that the SFP may provoke a stress reaction in infants, but that those reactions are inconsistently detected and may be modified by risk status or parental behaviors.

Meta-analysis

Hypotheses

We hypothesized, based on the literature set forth above, that across the set of studies analyzed, infants would show a withdrawal of the PNS during the SF episode. Based upon the variation in the pattern of findings for recovery by SES, however, we hypothesized that in middle and upper-class samples infants would respond with an increase in RSA during reunion, whereas in low SES/high risk samples, infants would show no change in RSA from SF to reunion. Finally, although most of the infants in this review were approximately on average 6 months of age, the samples in the meta-analysis ranged in age between 3 and 7.5 months. Although we expected that the samples lacked sufficient variability in age to detect any differences, we did examine whether age was associated with reactivity. We further hypothesized that age would not be significantly associated with reactivity because theoretically, reactivity reflects the impact of differing experiences on genetics and biological inheritance and is not expected to be stable or necessarily linear in its age trajectory (Alkon et al., 2011). Although we would have liked to test other potential modifiers of infant ANS such as sex, race, and SFP procedural differences, either there were too few appropriate studies to conduct such an examination or the variable-specific M/SD required to compare those effects were not provided in publications or by request.

Study data

Stricter criteria for selection of studies were used for the meta-analysis than for the review. First, as with the review, only studies that examined PNS activity during administration of the SFP using validated methods were included; however, too few studies of SNS activity were available for a valid analysis of SFP effects on the SNS. Second, studies that examined exclusively HR were excluded because both the PNS and SNS, as well as other systems, influence HR. Third, inclusion was limited to one published study per participant sample (multiple studies/publications drew on the same sample). Fourth, only studies with the necessary data either published or available upon request (note that data from four older studies were no longer available) were included. Fifth, we excluded a cohort of mothers who were told to touch their infants during the SF episode for two reasons: such instructions are in violation of the standard SFP procedure, and maternal touch could alter infant reactivity. Accordingly, the meta-analysis was restricted to 14 studies (see Table 2). For each study we extracted sample size, age, SES, mean and standard deviation statistics on RSA from each of three SFP episodes: (1) Free play/face-to-face immediately preceding still-face (except in the case of Bazhenova et al. (2001) where we use the Toy Attention episode that precedes SF); (2) still-face; and (3) reunion/play immediately following still-face (one study did not include reunion data (Stewart et al., 2013)). Data was extracted and checked independently by two authors. SES data was coded into three categories: middle/upper SES, mixed SES and lower SES samples based on data provided in the publications. When additional data (e.g., means/SD, outcomes, etc.) were necessary, authors were contacted up to two times with requests.

Statistical methods

In line with the literature reviewed above, we calculated the SFP meta-analytic RSA change scores for reactivity using episodes 1 and 2 (the difference between SF and play) and then for recovery using episodes 2 and 3 (the difference between reunion and SF). Reactivity and recovery scores were converted into standardized mean differences (SMD) and recorded as Hedge's g , which was then used as the basis for comparison across studies (Borenstein, Hedges, Higgins, & Rothstein, 2009). Effects of SES were tested by comparing a subset of studies ($n = 10$) where relative homogeneity suggested classification as either high-risk ($n = 3$) or predominantly middle class ($n = 7$). The remaining studies ($n = 4$) included within-sample heterogeneity of mixed SES and were excluded from this follow-up analysis. Although there were 5 studies that produced mixed SES samples, Suurland et al. (2016) also divided their sample into high-risk and low-risk samples, which allowed us test SES effects using their data as well.

Results

Analyses of SMDreactivity and SMDrecovery were performed in R using the metafor package (Viechtbauer, 2010) with a random effects model. Meta-analytic results showed a significant overall effect of RSA suppression, SMDreactivity as Hedge's g , of $-.35$ (with 95% CI $-.48$ to $-.22$, $p < .001$) and significant RSA recovery, SMDrecovery, of $.24$ (with 95% CI $.10$ to $.30$, $p < .001$). Significant heterogeneity was demonstrated for suppression ($I^2 = 66.5\%$, $p < .01$) and for recovery ($I^2 = 49.4\%$, $p < .01$). Forest plots for each outcome are shown in Figs. 2 and 3; these plots provide effect sizes adjusted for sample size, and thus present a more nuanced view of outcomes.

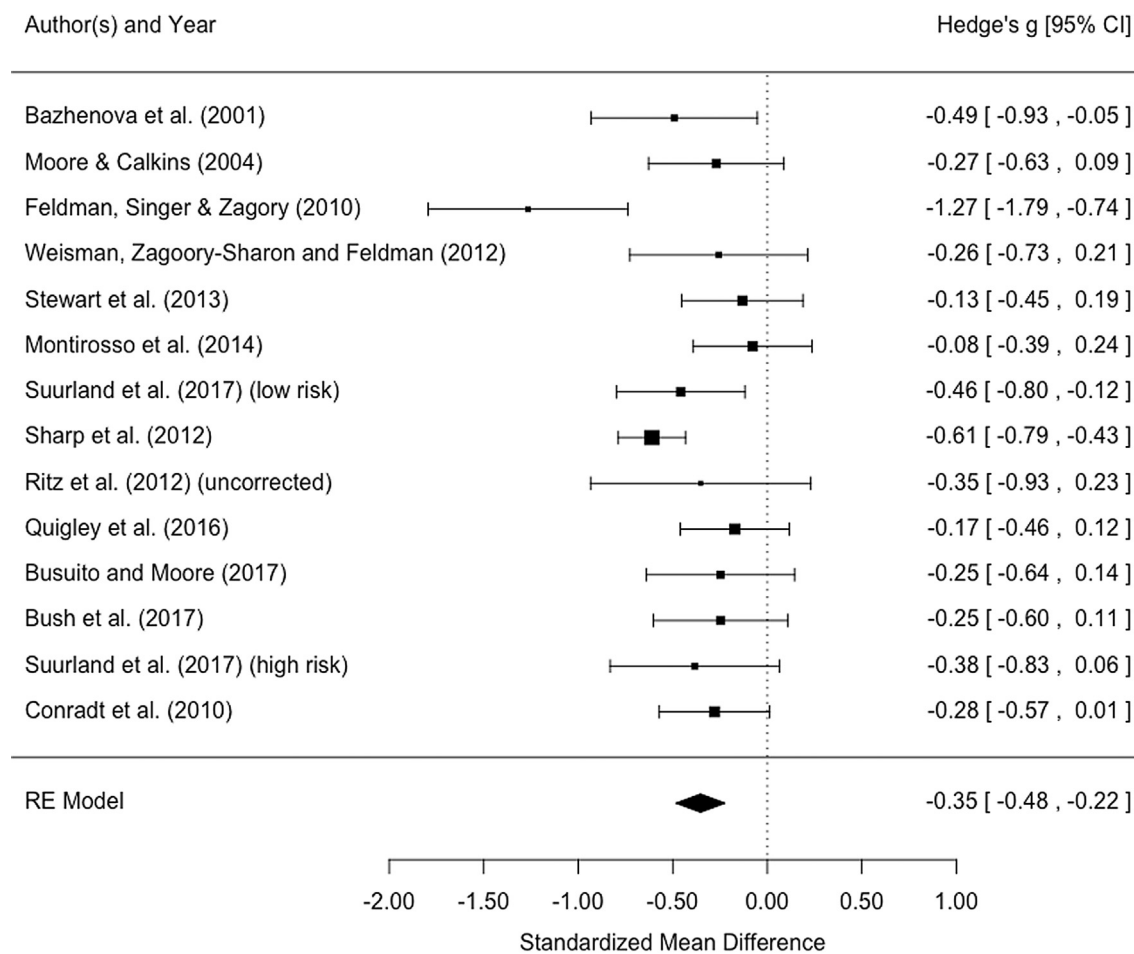


Fig. 2. Change in RSA from play to still-face.

Examination of sample SES as a moderator of associations was conducted by subgroup analyses. In studies with “mixed SES” samples ($n = 4$), analyses revealed heterogeneity estimates that were large for suppression ($I^2 = 89.2\%$, $p < .01$) and recovery ($I^2 = 61.2\%$, $p = .02$). Among studies of predominantly middle-class samples ($N = 7$), heterogeneity of suppression was smaller than “mixed SES”, though still large ($I^2 = 68.6$, $p < .01$); for recovery, heterogeneity was not observed ($I^2 = 0\%$, $p = .85$). Studies with samples characterized as “high-risk” ($n = 3$) did not demonstrate heterogeneity in either suppression ($I^2 = 0\%$, $p = .89$) or recovery ($I^2 = 0\%$, $p = .33$). Comparison between SES groups (high-risk vs. mid/upper class) revealed no significant differences for SMDreactivity ($Q = .15$, $p = .15$). However, for SMDrecovery the two groups were significantly different ($Q = 5.52$, $p = .02$), with the high-risk group showing no overall change in RSA (effect = $-.01$, $p = .95$) while the mid/upper class group showed significant recovery (effect = 0.31 , $p < .001$). As expected, tests examining infant age as a moderator were not significant for either SMDreactivity ($B = -.08$, $p = .13$) or SMDrecovery ($B = -.02$, $p = .77$).

Publication bias and sensitivity analysis

To assess publication bias we calculated the number of unpublished negative studies (fail-safe N) that would be required to: (1) increase the p-value of the meta-analysis above .05 (Sterne & Egger, 2005). For SMDreactivity the fail-safe n was 289 ($p > .05$), and for SMDrecovery it was 105. To assess bias in the estimated size of effect we inspected the funnel plot and tested for plot asymmetry using random effect version of Egger’s regression test (Egger, Smith, & Minder, 1997). Tests for asymmetry were not significant for either SMDreactivity ($z = -.71$, $p = .48$) or SMDrecovery ($z = -1.25$, $p = .21$). Visual inspection of funnel plots however indicated that individual studies might be influential so “leave-one-out” sensitivity analyses were performed (Viechtbauer, 2010). No one study changed the significance of overall effects. However, the exclusion of Sharp et al. (2012) substantially altered the estimated effect; for SMDreactivity, effect changes from $-.37$ to $-.30$ and for SMDrecovery from $.24$ to $.18$. Taken together, the results across studies appear robust, yet the estimation of effect size is somewhat challenged in part by the largest study relative to the collected other smaller studies in the meta-analysis.

As noted in Table 1, Ritz et al. (2012) showed much larger effect sizes after adjusting for respiration and making other statistical

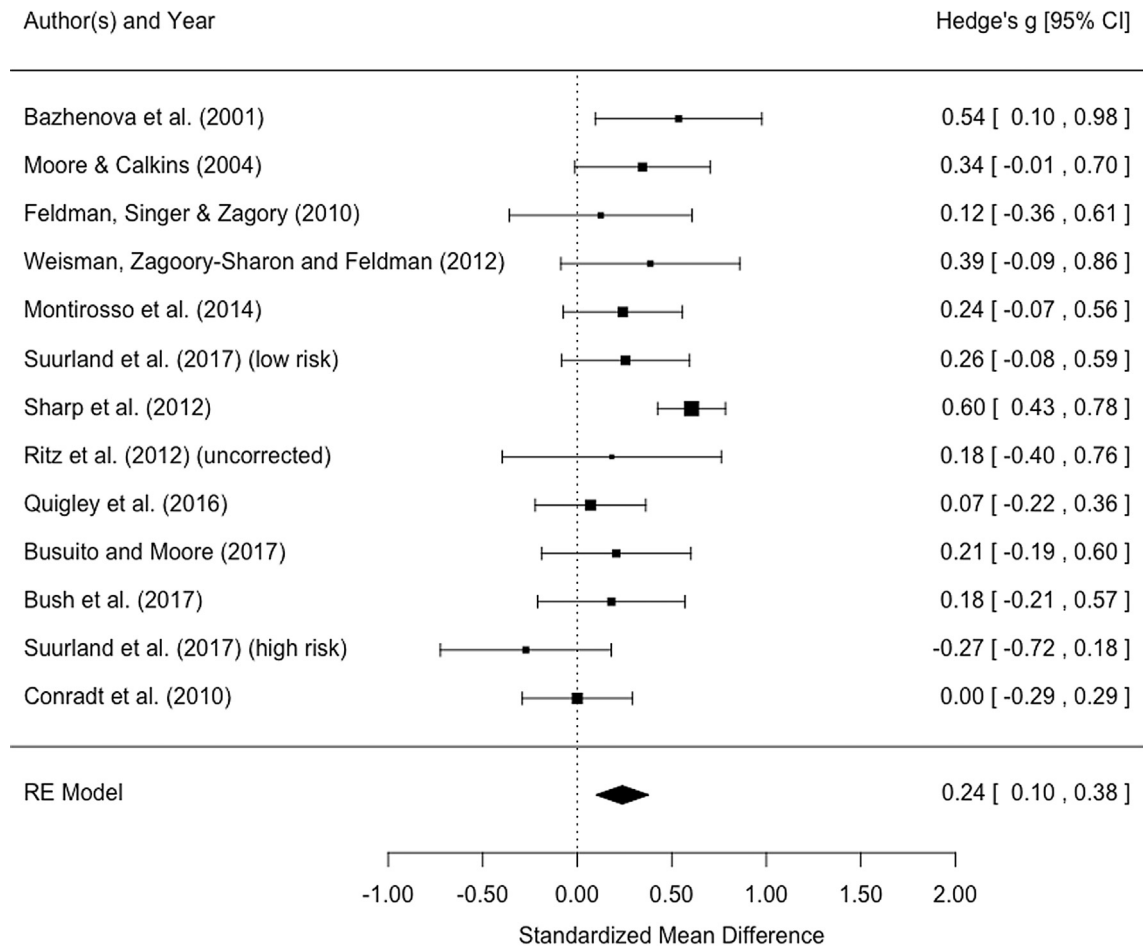


Fig. 3. RSA recovery in reunion episode.

adjustments. Using the respiration-adjusted values instead of the standard log-transformed values reported in our meta-analysis resulted in a significant asymmetry test. We chose to present our meta-analysis using the standard unadjusted values in order to present a more homogenous comparison. However, the overall results of the meta-analysis were not appreciably different using either calculation.

Discussion and conclusion

Patterns of ANS activity for infants administered the SFP differ

Providing a comprehensive review and a meta-analysis helped us tell a more nuanced story about infant stress reactivity and recovery than a review or meta-analysis alone could have provided. The review confirmed that the majority of studies showed that infants exhibit a decrease in RSA (PNS withdrawal) during the SF episode compared to baseline/play suggesting that the SFP effectively induces stress reactions in infants. Fewer studies, however, found that RSA returned to baseline levels (recovery) after the SF episode, with some reporting that RSA increased in reunion but not to baseline levels and others reporting that RSA did not change from SF to reunion. Moreover, a few studies reported that, for some infants, RSA actually decreased across the SFP. In three study samples, authors reported that approximately half of infants decrease RSA during the SF episode while the remaining infants either do not decrease or show increased RSA. Studies also showed that HR increased during the SF episode compared to baseline/play and decreased during reunion; in some studies, however, infant HR remained elevated from SF to reunion and/or did not return to play levels. In terms of SNS measures, although an increase in SC was triggered during the SFP suggesting SNS activation, no effects were found for TWA or PEP unless infants were classified by maternal sensitivity or risk status. Despite the range of findings identified by the review, the meta-analytic results supported our hypothesis that, overall, the SF episode reliably induces a decrease in infant RSA (PNS withdrawal).

Our second hypothesis was also supported: middle/upper class/low-risk infants showed an increase in RSA (PNS activation) during reunion, whereas low-SES/high-risk infants did not. These results suggest that infants coping with the stresses associated with

poverty and/or maternal psychopathy and risk have more difficulty reinstating the vagal brake than those babies raised in a low risk/middle class environment. This review indicates that a number of factors may be associated with infants' failure to recover in the SFP including differences in parental behaviors, SES and high-risk factors such as maternal trauma (Bosquet Enlow et al., 2009; Bush et al., 2017; Busuito & Moore, 2017; Conradt & Ablow, 2010; Haley & Stansbury, 2003; Mattson et al., 2013; Suurland et al., 2016). Integrating across these findings, it makes sense that infants with sensitive parents would show vagal recovery upon reunion as their source of emotional regulation has re-engaged. Overall, these studies suggest that it is important to assess parental behavior because reliable parenting measures may help identify and explain different infant ANS responses. There were a number of concerns that arose from this review, however;

Study outcomes relating to parental behavior were not all consistent.

For example, a number of studies found that parental sensitivity or related concepts were associated with ANS recovery in reunion (although not always back to baseline levels) (Haley & Stansbury, 2003; Moore & Calkins, 2004; Provenzi et al., 2015). There was one notable exception, however: Moore et al. (2009) reported that the infants of the most sensitive mothers during play showed lower RSA in the reunion than other infants, and a decrease in RSA from baseline to reunion. Noting that these sensitive mothers also showed a decrease in RSA during reunion, Moore et al. (2009) suggested that the lower RSA shown by infants may have reflected "mutual responsiveness" between mother and child. In this review, we found that lower RSA in reunion, or even a failure to increase RSA in reunion was more common for high-risk infants or infants living in high conflict environments (Busuito & Moore, 2017; Bush et al., 2017; Conradt & Ablow, 2010; Suurland et al., 2016). Why were the results in Moore et al. (2009) different? There are a number of possibilities. For example, perhaps the infants in Moore et al. (2009) were more environmentally sensitive; they may have had genetic or epigenetic differences that impacted their ability to recover. It is also possible that the outcome differences reflect differences in measures of parental sensitivity: Moore et al. (2009) was one of the few studies that assessed maternal sensitivity outside of the SFP. A later study that drew from the same sample may also offer some hints; Quigley et al. (2016) found that babies who were breastfed had lower RSA from baseline to reunion than non-breastfed babies, and that the babies of high SES parents had lower RSA in reunion. Perhaps the findings in Moore et al. (2009) reflect the presence of a significant number of breastfed babies although Quigley et al. (2016) reported that in their model maternal sensitivity was insignificant (there were some sample differences between the two studies). It is not clear why the babies of high SES parents would have lower RSA in reunion. Additional research is needed to unravel all these complex variables.

Although the meta-analysis confirmed that infant RSA decreases in the SF episode, approximately half of the infants in the S-NS studies failed to show a decrease in RSA during the SF episode.

There are multiple reasons why infants might not show RSA withdrawal during challenge. First, these infants may simply not be stressed. But NS infants appeared to show similar levels of negative behaviors and elevated HR as S infants in the SF at least suggesting that they were stressed. Second, these infants may not have developed the ability to engage in emotional self-control yet (Moore et al., 2009; Porges, 1996). Third, multiple findings suggested an association between the failure to withdraw during the SF episode and insensitive or unresponsive parenting (Moore & Calkins, 2004; Provenzi et al., 2015), and one found that infants from high conflict families actually showed PNS withdrawal during the play episode (Moore, 2010). One reason for such patterning might be that interaction with the parent is more stressful for the infant than parental disengagement, particularly if the parent is aggressive or intrusive. Moore also suggests that since the NS infants do not appear to receive the parental support necessary to regulate emotional arousal, they may have to habitually self-regulate contributing to the failure to exhibit PNS withdrawal during the SF episode. As suggested by the BSC theory, the failure to withdraw RSA during times of stress might be adaptive in a high-risk environment where parental disengagement is a consistent phenomenon. Another possibility is that infants may simply show "innate" differences in ANS function (Moore & Calkins, 2004). The attachment literature also gives some clues about what might be happening with NS infants; among toddlers and older children, those who do not withdraw PNS during challenge are classified with ambivalent or disordered attachment (Abtahi & Kerns, 2017; Oosterman, De Schipper, Fisher, Dozier, & Schuengel, 2010; Paret, Bailey, Roche, Bureau, & Moran, 2015), while toddlers who showed PNS withdrawal from baseline to a stressor are classified as securely attached (Paret et al., 2015). Insecurity or even attachment disorganization would be expected in an environment in which the child coped by suppressing normal stress reactions in response to parental disengagement or frightening parental behavior. On the other hand, parental sensitivity may support flexible PNS function. Given the linkage between attachment and self-regulation, longitudinal follow-up of infants who fail to suppress RSA in the SF may confirm that ANS response during the SFP provides an early glimpse into future emotional regulation.

Other factors that might impact ANS.

Subjecting infants to more than one task prior to administering the SFP (Bazhenova et al., 2001; Sharp et al., 2012; Tibu et al., 2014), a large sample size (Sharp et al., 2012; Tibu et al., 2014), controlling for respiration (Ritz et al., 2012), subjecting infants to two SF episodes (Ritz et al., 2012) and dividing infants into S and NS groups (Provenzi et al., 2015) appeared to produce some of the moderate and high effect sizes for RSA reactivity. We did not perform a meta-analysis comparing studies that used the peak-valley method versus Porges (1985) algorithm because of the low number of peak-valley studies and the potentially confounding factor of SES (two of the three peak valley studies used a low SES/high risk sample). Nevertheless, we note that both methods produced, with few exceptions, between small and medium effect sizes. We were also not able to discern any difference in effect sizes based on scoring epochs or handling of missing data. Finally, we were not able to ascertain the impact of parental touch in the SFP in our meta-analysis, because many studies simply told parents to play with their infants as they normally would, which might or might not

include touching. Given the studies that suggest parental touch may impact infant RSA (Feldman et al., 2010; Sharp et al., 2012) additional research should be conducted to investigate this phenomenon.

Strength and weaknesses in the SFP literature and proposals for future research

The meta-analysis findings suggest one major strength in the literature: the SFP reliably produces a PNS reaction. Moreover, the SFP appears to differentiate PNS reactivity between low-risk/mid/high-SES infants and high-risk/low-SES infants. Yet there are many weaknesses in the literature that must be addressed before further conclusions can be made. Below is a list of limitations and proposed solutions to guide future research:

- (1) Limitation: The most persistent weakness across studies is a lack of standardization of RSA measures and analyses (e.g., Frequency vs. peak-trough methods; measures of RSA at 5, 10, 15, 20 or 30 s epochs; different definitions of outliers; different techniques for dealing with missing data; different levels of data editing; different methods of calculating ANS reactivity (e.g., SFP episode minus baseline vs. differences between preceding episodes); different baseline conditions (e.g., play episodes vs. another condition before onset of the SF episode); different criteria to determine when to terminate an episode; and different instructions to parents/researchers during play and reunion (e.g., touch vs. no touch; toys vs. no toys, etc.). This prevents researchers from meaningfully comparing across studies or determining whether these differences impacts infant ANS measurements.

*Proposal: While we understand the need to modify the SFP to test specific hypotheses, some standardization across these realms may be appropriate to improve reliability and generalization of findings. For example, given the importance of parental touch we suggest that studies standardize instructions on parental touch, and/or explicitly code instances of parental touch at any point during the SFP (Mantis, Stack, Ng, Serbin, & Schwartzman, 2014), and report these details in publications. We also urge researchers to adopt a standardized method to calculate ANS reactivity to allow for easier comparisons between studies. Researchers could then modify reactivity calculations in secondary analyses, to assess the impact of such innovation. Standardizing baseline measures (and limiting challenging tasks administered prior to the SFP) is also important as infants could start the SFP at different levels of activation or relaxation depending on the nature of the pre-task activity. Finally, because differences in scoring and cleaning ANS data can impact RSA, we suggest that both be reported and that the field endeavors toward uniform approaches. Researchers might work together to achieve greater standardization by organizing specialized pre-conference seminars at conferences specializing in child development and publishing suggestions for future studies.

- (2) Limitation: Differences in methodology and protocol design (e.g., use parents vs. strangers, length of episodes, one SF vs. two SF episodes, etc.).

*Proposal: Again, although we acknowledge the potential utility of variation in study design, we suggest some basic standardization is possible here. For example, although the length of the SFP episodes may be adjusted in the case of very young infants (e.g. neonates) most of the infants in studies reviewed were 6 months of age. Accordingly, standardization of episode duration would be feasible for similarly-aged infants, unless other study elements required minimizing task length.

- (3) Limitation: Lack of standard methods to control for influence of movement on ANS measures.

*Proposal: We recommend that the community adopt replicable methods to assess the role of infant movement in SFP ANS responses.

- (4) Limitation: Lack of standardization when measuring infant and maternal behaviors during the SFP (maternal sensitivity, parental responsiveness, “matching,” reparation rates, different child behavior codes and the array of “sensitivity” measures). Moreover, it is difficult to measure such a complex phenomenon as “maternal sensitivity”, particularly based upon 2–4 min of play behavior within the SFP. These difficulties could explain why studies produce mixed results.

*Proposal: Because the attachment relationship provides the primary environment in which the highly-plastic infant brain develops (Jones-Mason, Allen, Bush, & Hamilton, 2016; Kolb & Gibb, 2011), the most salient modifier for infant ANS function is likely the infant – parent relationship. Parental sensitivity is a key starting point for studies seeking to understand differences ANS function. We suggest that investing the time and resources necessary to conduct reliable and valid measures of parental sensitivity, as noted by others (Lindhiem, Bernard, & Dozier, 2011), would advance SFP research. For example, the Maternal Behavior Q Sort (MBQSPederson & Moran, 1995) is a highly respected method of assessing maternal sensitivity, has an abbreviated 25 item version that can be used relatively quickly in the lab (Behrens, Haltigan, & Bahm, 2016). Increased collaboration across labs might facilitate the use of the more validated parental sensitivity measures and lower barriers to their use. Otherwise, researchers are encouraged to clarify that they are assessing parental behavior within the context of the SFP rather than assessing parental sensitivity more broadly.

- (5) Limitation: Few studies were found that involved a preponderance of participants from low-SES families, limiting the ability to understand the role of SES on infant ANS response to the SFPO. Further, although there were few studies of sex and ethnic/race differences available for review, there was some evidence for such. There is need for more investigation into sex, racial/ethnic, and SES differences.

*Proposal: We advise the use of representative, diverse sample populations, when feasible, and/or increase the number of studies focusing on samples of infants living in low SES households. Additional research on sex and ethnicity/race differences would clarify whether differences are reliably identified. In addition, despite the fact that rural areas have higher levels of poverty and health disparities, such populations in this research were rarely examined (HAC, 2012; Hartley, 2004). Accordingly, we suggest that research include more participants living in the rural areas.

- (6) Limitation: A number of researchers have noted the lack of research into infant SNS reactivity (Moore et al., 2009; Suurland et al., 2016) despite evidence for variability in ANS branch reactivity among children and adults –some show PNS withdrawal while others show SNS activation and still others show both (Bernston et al., 2007). Yet, as this review shows, there is a dearth of research in this area, with only four published studies attempting to assess SNS indicia during the SFP (Bosquet Enlow et al., 2014; Ham & Tronick, 2006, 2009; Suurland et al., 2016).
 *Proposal: Suurland et al. (2016) and Caron et al. (2017) are two of the newest studies that demonstrate it is possible to measure SNS functioning, via PEP, in infants. Accordingly, we recommend that SFP researchers include measures of SNS. Moreover, we endorse the suggestion made by Moore et al. (2009) for more integrated studies that incorporate not only PNS and SNS indicia, but also variables that may add a more multi-dimensional aspect to understanding infant ANS development including infant and maternal behavioral responsiveness and other physiological variables relevant to self-regulation (e.g., cortisol or salivary alpha amylase). Multiple researchers have proposed cross-system coordination of stress responses (Bernston et al., 1993; McEwen, 2006; Quas et al., 2014; Sapolsky, 2004), and some studies have already examined multi-system relationships in infants, but more research is needed (Bauer, Quas, & Boyce, 2002; Quas et al., 2014; Rash et al., 2016).
- (7) Limitation: There are still other important individuals and communities that have not been involved in SFP-ANS research. For example, one of the starkest weaknesses in these studies is the almost total lack of SFP- infant ANS data with fathers. Weisman et al. (2012) was the only study found that focused on fathers, demonstrating that fathers too may have a profound impact on attachment related infant biological function. Fathers are also important to infant ANS function through the role they play in their relationships with mothers (Moore, 2010). To fully understand infant ANS function, research must include fathers to the same extent as mothers. Moreover, the composition of families is growing more diverse and there are few, if any, studies that examine the ANS function in the SFP of infants living in non-traditional configurations (e.g., infants raised by grandparents or other relatives, adoptive parents, or infants with same sex parents, etc.).
 *Proposal: Include fathers in future research to expand the breadth of understanding of the home environment and its influence on infant-parent responsiveness. Further extend samples to include infants living in non-traditional households.
- (8) Limitation: Investigation into genetic and epigenetic differences that impact ANS has barely begun.
 *Proposal: Researchers have already begun to demonstrate interest in examining epigenetic correlates of SFP response in healthy (Conradt et al., 2015) and at-risk infants (Montirosso et al., 2016). We suggest expansion to include consideration of genetic and epigenetic influences and search for interactions with environmental experience (e.g., SES, parental sensitivity, parental psychopathology).
- (9) Limitation: There is limited longitudinal research.
 *Proposal: Longitudinal research would allow for determination of, for example, the implications of being a S versus a NS infant over time and under what conditions those RSA profiles change. Is it possible that ANS category is influenced by genes or epigenetic processes or maternal prenatal experience? We also urge researchers to report the precise numbers of infants that are S and NS to help the field definitively determine the normative distribution of reactivity among infants and to determine if the distribution of S vs. NS infants changes over time.
- (10) Limitation: Few studies examine whether outcomes differ depending on whether measures are administered in the home or the lab.
 *Proposal: Many families, particularly high-risk families, are not able or willing to travel to a lab to participate in research. Although no differences between results when the SFP was administered in a lab vs. the home were described in this review, the evidence was limited. Further, researchers could explore opportunities for standardization across settings; for example, mobile vans equipped with EKG/ECG equipment have been used successfully with older children.
- (11) Limitation: Sample sizes may vary depending on drop-out rate.
 *Proposal: Researchers should conduct a power analysis to determine a sufficient starting sample size, considering the effect sizes in their population of interest and the likelihood of a significant drop out rate. Note that Moore (2010) indicated that it is typical to lose about a quarter of data due to infant movement and noncompliance, and even more because of infant protest and technical problems.

Conclusion

In conclusion, several questions and overall issues arise from this review. First, this review shows that many researchers are trying to demonstrate that the parent-child relationship may influence ANS function. As alluded to earlier, if infants show the same overall ANS responses in the SFP with strangers as with parents, however, to what extent is the SFP revealing something about the history of the parent-child relationship as opposed to, perhaps, evidence of a bio-behavioral system present at birth? More research is necessary to determine to what extent ANS function measured in the SFP may reveal something about the quality of *specific* relationships.

Second, can researchers be certain about what the behaviors and physiological reactions demonstrated by infants mean in the SFP? For example, as noted in this review, NS infants may be at risk of developmental or behavioral problems, but they may also not suppress simply because they are not stressed (Ham & Tronick, 2006). Administering other measures of infant stress reactivity (e.g., salivary alpha amylase, PEP, etc.) may aid interpretation of findings.

Third, while we acknowledge study heterogeneity, the results did confirm that the SFP appears to create a stress reaction in infants resulting in PNS withdrawal. We remain concerned, however, that the lack of standardization and other factors compromise the ability to understand what is driving variation in findings. The S-NS studies suggest that the “normal” distribution of RSA responses among infants to the SF episodes is that half of infants withdraw and half do not. If that is the case, many studies might

produce null findings, and positive finding publication bias may be skewing the knowledge base. That said, our tests for publication bias in the meta-analyses suggest this is not the case although we acknowledge that tests of publication bias are not perfect and are, ultimately, based on published studies. Collectively, although our hypothesis of a SFP effect on RSA was confirmed, we believe considerable additional work is required to derive definitive conclusions about factors that predict infant ANS responses in the SFP.

Does the review suggest any clinical implications for this field of research? The lack of standardization in the field, some inconsistent findings, particularly around parental behavior, and limited samples of high-risk/low SES communities prevent conclusive answers. Nevertheless, results of the meta-analysis do confirm that sustained parental disengagement is associated with infant PNS withdrawal, on average, and that re-engagement with the infants in play typically results in recovery. Clinically, this affirms the key role parenting behaviors play in activating infant stress responses and in supporting their return to “non-stressed” levels of physiology. Results also suggest that infants raised in high-risk environments take longer to recover from stressful challenges. On the other hand, as the review indicated, infants exposed to insensitive parenting or high levels of parental conflict within the household may show lower RSA than other infants, or may not suppress RSA during challenge at all (Moore, 2010; Provenzi et al., 2015). These results are consistent with the conclusion of other research reviews finding that factors such as prenatal stress, and post-natal environmental exposures such as marital conflict, maternal psychopathology and insensitive parenting are associated with low vagal tone and inconsistent vagal withdrawal (Graziano & Derefinko, 2013; Propper and Holochwost, 2013). Given the neuroplasticity of the infant brain, prolonged stress reactions could have significant consequences to both brain development and future stress reactivity (e.g., heightened or blunted) (Kolb & Gibb, 2011; McEwen, 2015). Failure to exhibit PNS withdrawal during times of challenge has also been associated with less optimal outcomes such as insecure or disorganized attachment (Oosterman et al., 2010). If these results are confirmed in larger meta-analytic samples, consideration of therapeutic intervention may be advisable. For example, interventions such as “The Attachment and Biobehavioral Catch-up (ABC) have been shown to improve parenting and the parent-child relationship, and appears to contribute to sustained normalization of children’s HPA-axis functioning—such work may be useful for supporting healthy ANS physiology as well (Bernard, Simons, & Dozier, 2015). We emphasize three points, however. First, the intervention must take into account the context in which the infant lives. Prolonged protest and PNS withdrawal (low RSA) may be adaptive in a high-risk household; the infant may have to protest longer in order to capture the attention of a parent who may be distracted or unavailable. Second, as alluded to earlier, the failure to withdraw RSA may be adaptive; infants may be self-regulating because the parent is too stressed or not available for other reasons to help the baby regulate. Third, some forms of maternal “psychopathology” such as anxiety may actually have some beneficial effects for the infant. Recall that Ostlund et al. (2017) found that infants of anxious mothers had the highest RSA in reunion; the authors opined that anxious mothers may be more responsive to the infants resulting in quick recovery. In short, any interventions considered in the future will have to thoughtfully consider the biological and social complexities, particularly for families living in high-risk environments.

Ultimately, the challenge facing researchers in the future is to untangle a web of complex variables without assuming that any one pattern is optimal. As the BSC literature suggests, optimal outcomes may depend on the challenges facing the infant in life (Bush & Boyce, 2016). The only way researchers can meaningfully answer these questions is through longitudinal studies of diverse populations that employ rigorous standardized conditions and include measures of family demographics and parenting.

Acknowledgements

This research was supported, in part, by NHLBI 5 R01 HL116511.

The authors thank Zoe Caron for her assistance with the literature review and careful reading of the manuscript.

The authors are also grateful to the following researchers for sharing their data with us for the meta-analysis and/or, in many cases, generously addressing questions and concerns: Dr. Ruth Feldman, Dr. Omri Weisman, Dr. Livio Provenzi, Dr. Rosario Montirosso, Dr. Steven Holochwost, Dr. Kelsey Quigley, Dr. Ginger Moore, Dr. Michelle Bosquet Enlow, Dr. Keri Heilman.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.dr.2018.06.002>.

References

- Abtahi, M. M., & Kerns, K. A. (2017). Attachment and emotion regulation in middle childhood: Changes in affect and vagal tone during a social stress task. *Attachment & Human Development, 19*(3), 221–242.
- Adamson, L. B., & Frisk, J. E. (2003). The still face: A history of a shared experimental paradigm. *Infancy, 4*(4), 451–473.
- Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation* (Erlbaum, Ed.). NJ: Hillsdale.
- Alkon, A., Boyce, W. T., Davis, N. V., & Eskenazi, B. (2011). Developmental changes in autonomic nervous system resting and reactivity measures in Latino children from 6 to 60 months of age. *Journal of Developmental and Behavioral Pediatrics, 32*(9), 668–677. <http://dx.doi.org/10.1097/DBP.0b013e3182331fa6>.
- Alkon, A., Boyce, W. T., Tran, L., Harley, K. G., Neuhaus, J., & Eskenazi, B. (2014). Prenatal adversities and Latino children’s autonomic nervous system reactivity trajectories from 6 months to 5 years of age. *PLoS One, 9*(1), e86283. <http://dx.doi.org/10.1371/journal.pone.0086283>.
- Alkon, A., Goldstein, L. H., Smider, N., Essex, M. J., Kupfer, D. J., & Boyce, W. T. (2003). Developmental and contextual influences on autonomic reactivity in young children. *Developmental Psychobiology, 42*(1), 64–78. <http://dx.doi.org/10.1002/dev.10082>.
- Alkon, A., Lippert, S., Vujan, N., Rodriguez, M. E., Boyce, W. T., & Eskenazi, B. (2006). The ontogeny of autonomic measures in 6- and 12-month-old infants. *Developmental Psychobiology, 48*(3), 197–208. <http://dx.doi.org/10.1002/dev.20129>.
- Alkon, A., Wolff, B., & Boyce, W. T. (2012). Poverty, stress, and autonomic reactivity. In V. Maholmes, & R. B. King (Eds.). *The Oxford handbook of poverty and child development* (pp. 221–239). New York: Oxford University Press.
- Anda, R. F., Felitti, V. J., Bremner, J. D., Walker, J. D., Whitfield, C., Perry, B. D., ... Giles, W. H. (2006). The enduring effects of abuse and related adverse experiences

- in childhood. A convergence of evidence from neurobiology and epidemiology. *European Archives of Psychiatry and Clinical Neuroscience*, 256(3), 174–186. <http://dx.doi.org/10.1007/s00406-005-0624-4>.
- Bar-Haim, Y., Marshall, P. J., & Fox, N. A. (2000). Developmental changes in heart period and high-frequency heart period variability from 4 months to 4 years of age. *Developmental Psychobiology*, 37(1), 44–56.
- Bauer, A. M., Quas, J. A., & Boyce, W. T. (2002). Associations between physiological reactivity and children's behavior: Advantages of a multisystem approach. *Journal of Developmental and Behavioral Pediatrics*, 23(2), 102–113.
- Bazhenova, O. V., Plonskaia, O., & Porges, S. W. (2001). Vagal reactivity and affective adjustment in infants during interaction challenges. *Child Development*, 72(5), 1314–1326.
- Beauchaine, T. (2001). Vagal tone, development, and Gray's motivational theory: Toward an integrated model of autonomic nervous system functioning in psychopathology. *Development and Psychopathology*, 13(2), 183–214.
- Beauchaine, T. (2007). A brief taxometrics primer. *Journal of Clinical Child and Adolescent Psychology*, 36(4), 654–676. <http://dx.doi.org/10.1080/15374410701662840>.
- Beauchaine, T., Gatzke-Kopp, L., & Mead, H. K. (2007). Polyvagal Theory and developmental psychopathology: Emotion dysregulation and conduct problems from preschool to adolescence. *Biological Psychology*, 74(2), 174–184. <http://dx.doi.org/10.1016/j.biopsycho.2005.08.008>.
- Beauchaine, T., Katkin, E. S., Strassberg, Z., & Snarr, J. (2001). Disinhibitory psychopathology in male adolescents: Discriminating conduct disorder from attention-deficit/hyperactivity disorder through concurrent assessment of multiple autonomic states. *Journal of Abnormal Psychology*, 110(4), 610–624.
- Behrens, K. Y., Haltigan, J. D., & Bahm, N. I. (2016). Infant attachment, adult attachment, and maternal sensitivity: Revisiting the intergenerational transmission gap. *Attachment & Human Development*, 18(4), 337–353. <http://dx.doi.org/10.1080/14616734.2016.1167095>.
- Belsky, J., & Beaver, K. M. (2011). Cumulative-genetic plasticity, parenting and adolescent self-regulation. *Journal of Child Psychology and Psychiatry*, 52(5), 619–626. <http://dx.doi.org/10.1111/j.1469-7610.2010.02327.x>.
- Belsky, D. W., Suppli, N. P., & Israel, S. (2014). Gene-environment interaction research in psychiatric epidemiology: A framework and implications for study design. *Social Psychiatry and Psychiatric Epidemiology*, 49(10), 1525–1529. <http://dx.doi.org/10.1007/s00127-014-0954-5>.
- Ben-Tal, A., Shamailov, S. S., & Paton, J. F. (2012). Evaluating the physiological significance of respiratory sinus arrhythmia: Looking beyond ventilation-perfusion efficiency. *Journal of Physiology*, 590(Pt 8), 1989–2008. <http://dx.doi.org/10.1113/jphysiol.2011.222422>.
- Bernard, K., Simons, R., & Dozier, M. (2015). Effects of an Attachment-Based Intervention on Child Protective Services-Referred Mothers' Event-Related Potentials to Children's Emotions. *Child Development*, 86(6), 1673–1684.
- Bernston, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Respiratory sinus arrhythmia: Autonomic origins, physiological mechanisms, and psychophysiological implications. *Psychophysiology*, 30, 183–196.
- Bernston, G. G., Quigley, K. S., & Lazano, D. (2007). Cardiovascular psychophysiology. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology*, 3rd ed. (pp. 182–210). Cambridge England, New York: Cambridge University Press.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to meta-analysis. The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK: John Wiley & Sons, Ltd.
- Bosquet Enlow, M., King, L., Schreier, H. M., Howard, J. M., Rosenfield, D., Ritz, T., & Wright, R. J. (2014). Maternal sensitivity and infant autonomic and endocrine stress responses. *Early Human Development*, 90(7), 377–385. <http://dx.doi.org/10.1016/j.earlhumdev.2014.04.007>.
- Bosquet Enlow, M., Kullowatz, A., Staudenmayer, J., Spasojevic, J., Ritz, T., & Wright, R. J. (2009). Associations of maternal lifetime trauma and perinatal traumatic stress symptoms with infant cardiorespiratory reactivity to psychological challenge. *Psychosomatic Medicine*, 71(6), 607–614. <http://dx.doi.org/10.1097/PSY.0b013e3181ad1c8b>.
- Bowlby, J. (1968). *Attachment and Loss: Vol. I, Attachment, Vol. I*, New York, NY: Basic Books 1982.
- Boyce, W. T., & Ellis, B. (2005). Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity. *Development and Psychopathology*, 17(2), 271–301. Retrieved from < <http://www.ncbi.nlm.nih.gov/pubmed/16761546> > .
- Braungart-Rieker, J. M., Garwood, M. M., Powers, B. P., & Wang, X. (2001). Parental sensitivity, infant affect, and affect regulation: Predictors of later attachment. *Child Development*, 72(1), 252–270.
- Braungart-Rieker, J. M., Zentall, S., Lickenbrock, D. M., Ekas, N. V., Oshio, T., & Planalp, E. (2014). Attachment in the making: Mother and father sensitivity and infants' responses during the Still-Face Paradigm. *Journal of Experimental Child Psychology*, 125, 63–84. <http://dx.doi.org/10.1016/j.jecp.2014.02.007>.
- Brody, G. H., Yu, T., Chen, Y. F., Kogan, S. M., Evans, G. W., Windle, M., ... Philibert, R. A. (2013). Supportive family environments, genes that confer sensitivity, and allostatic load among rural African American emerging adults: A prospective analysis. *Journal of Family Psychology*, 27(1), 22–29. <http://dx.doi.org/10.1037/a0027829>.
- Bush, N., & Boyce, T. (2016). Differential sensitivity to context: Implications for developmental psychopathology. In D. Cicchetti (Ed.), *Developmental psychopathology*. Hoboken, NJ: John Wiley & Sons.
- Bush, N. R., Jones-Mason, K., Coccia, M., Caron, Z., Alkon, A., Thomas, M., ... Epel, E. S. (2017). Effects of pre- and postnatal maternal stress on infant temperament and autonomic nervous system reactivity and regulation in a diverse, low-income population. *Development and Psychopathology*, 29(5), 1553–1571. <http://dx.doi.org/10.1017/S0954579417001237>.
- Bush, N., Obradovic, J., Stamplerdahl, J., & Boyce, W. T. (2011). Differentiating challenge reactivity from psychomotor activity in studies of children's psychophysiology: Considerations for theory and measurement. *Journal of Experimental Child Psychology*, 110(1), 62–79. <http://dx.doi.org/10.1016/j.jecp.2011.03.004>.
- Busuito, A., & Moore, G. A. (2017). Dyadic flexibility mediates the relation between parent conflict and infants' vagal reactivity during the Face-to-Face Still-Face. *Developmental Psychobiology*. <http://dx.doi.org/10.1002/dev.21508>.
- Cacioppo, J. T., & Bernston, G. G. (2011). The brain, homeostatis and health: Balancing demands of the internal and external milieu. In H. S. Friedman (Ed.), *The Oxford handbook of health psychology* (pp. 121–137). New York, NY: Oxford University Press.
- Cacioppo, J. T., Uchino, B. N., & Bernston, G. G. (1994). Individual differences in the autonomic origins of heart rate reactivity: The psychometrics of respiratory sinus arrhythmia and preejection period. *Psychophysiology*, 31, 412–419.
- Calkins, S. D., Graziano, P. A., & Keane, S. P. (2007). Cardiac vagal regulation differentiates among children at risk for behavior problems. *Biological Psychology*, 74(2), 144–153. <http://dx.doi.org/10.1016/j.biopsycho.2006.09.005>.
- Caron, Z., Jones-Mason, K., Coccia, M., Chu, W., Alkon, A., & Bush, N. (2017). The prism of reactivity: Concordance between behavioral, questionnaire, and physiological domains of stress reactivity in infants. Paper presented at the Society for Research in Child Development, Austin, TX.
- Chen, E., Miller, G. E., Kober, M. S., & Cole, S. W. (2011). Maternal warmth buffers the effects of low early-life socioeconomic status on pro-inflammatory signaling in adulthood. *Molecular Psychiatry*, 16(7), 729–737. <http://dx.doi.org/10.1038/mp.2010.53>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Conradt, E., & Ablow, J. (2010). Infant physiological response to the still-face paradigm: Contributions of maternal sensitivity and infants' early regulatory behavior. *Infant Behavior and Development*, 33(3), 251–265.
- Conradt, E., Fei, M., LaGasse, L., Tronick, E., Guerin, D., Gorman, D., ... Lester, B. M. (2015). Prenatal predictors of infant self-regulation: The contributions of placental DNA methylation of NR3C1 and neuroendocrine activity. *Frontiers in Behavioral Neuroscience*, 9, 130. <http://dx.doi.org/10.3389/fnbeh.2015.00130>.
- Conradt, E., Measelle, J., & Ablow, J. C. (2013). Poverty, problem behavior, and promise: Differential susceptibility among infants reared in poverty. *Psychological Science*, 24(3), 235–242. <http://dx.doi.org/10.1177/0956797612457381>.
- De Bellis, M. D., & Zisk, A. (2014). The biological effects of childhood trauma. *Child and Adolescent Psychiatric Clinics of North America*, 23(2), 185–222, vii. doi: 10.1016/j.chc.2014.01.002.
- Dietrich, A., Riese, H., Sondejker, F. E., Greaves-Lord, K., van Roon, A. M., Ormel, J., ... Rosmalen, J. G. (2007). Externalizing and internalizing problems in relation to autonomic function: A population-based study in preadolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46(3), 378–386. <http://dx.doi.org/10.1097/CHI.0b013e31802b91ea>.
- Egger, M., Smith, G. D., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629–634.

- Eisenberg, N., Fabes, R. A., Murphy, B., Maszk, P., Smith, M., & Karbon, M. (1995). The role of emotionality and regulation in children's functioning: A longitudinal Study. *Child Development*, 66, 1360–1384.
- Eisenberg, N., Sulik, M. J., Spinrad, T. L., Edwards, A., Eggum, N. D., Liew, J., ... Hart, D. (2012). Differential susceptibility and the early development of aggression: Interactive effects of respiratory sinus arrhythmia and environmental quality. *Developmental Psychology*, 48(3), 755–768. <http://dx.doi.org/10.1037/a0026518>.
- Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2011). Differential susceptibility to the environment: An evolutionary-neurodevelopmental theory. *Development and Psychopathology*, 23(1), 7–28. <http://dx.doi.org/10.1017/S0954579410000611>.
- El-Sheikh, M. (2005). Stability of respiratory sinus arrhythmia in children and young adolescents: A longitudinal examination. *Developmental Psychobiology*, 46(1), 66–74. <http://dx.doi.org/10.1002/dev.20036>.
- El-Sheikh, M., & Erath, S. A. (2011). Family conflict, autonomic nervous system functioning, and child adaptation: State of the science and future directions. *Development and Psychopathology*, 23(2), 703–721. <http://dx.doi.org/10.1017/S0954579411000034>.
- El-Sheikh, M., Harger, J., & Whitson, S. M. (2001). Exposure to interparental conflict and children's adjustment and physical health: The moderating role of vagal tone. *Child Development*, 72(6), 1617–1636.
- Fabes, R. A., Eisenberg, N., Karbon, M., Troyer, D., & Switzer, G. (1994). The relations of children's emotion regulation to their vicarious emotional responses and comforting behaviors. *Child Development*, 65(6), 1678–1693.
- Feldman, R., Singer, M., & Zagoory, O. (2010). Touch attenuates infants' physiological reactivity to stress. *Developmental Science*, 13(2), 271–278. <http://dx.doi.org/10.1111/j.1467-7687.2009.00890.x>.
- Field, T. (1994). The effects of mother's physical and emotional unavailability on emotion regulation. *Monographs of the Society for Research in Child Development*, 59(2–3), 208–227.
- Field, T., Diego, M., Hernandez-Reif, M., Salman, F., Schanberg, S., Kuhn, C., ... Bendell, D. (2002). Prenatal anger effects on the fetus and neonate. *Journal of Obstetrics and Gynaecology*, 22(3), 260–266. <http://dx.doi.org/10.1080/01443610220130526>.
- Field, T., Diego, M., Hernandez-Reif, M., Schanberg, S., Kuhn, C., Yando, R., & Bendell, D. (2003). Pregnancy anxiety and comorbid depression and anger: Effects on the fetus and neonate. *Depress Anxiety*, 17(3), 140–151. <http://dx.doi.org/10.1002/da.10071>.
- Fifer, W. P., Fingers, S. T., Youngman, M., Gomez-Gribben, E., & Myers, M. M. (2009). Effects of alcohol and smoking during pregnancy on infant autonomic control. *Developmental Psychobiology*, 51(3), 234–242. <http://dx.doi.org/10.1002/dev.20366>.
- Goedhart, A. D., van der Sluis, S., Houtveen, J. H., Willemsen, G., & de Geus, E. J. (2007). Comparison of time and frequency domain measures of RSA in ambulatory recordings. *Psychophysiology*, 44(2), 203–215. <http://dx.doi.org/10.1111/j.1469-8986.2006.00490.x>.
- Gordis, E. B., Feres, N., Olezeski, C. L., Rabkin, A. N., & Trickett, P. K. (2010). Skin conductance reactivity and respiratory sinus arrhythmia among maltreated and comparison youth: Relations with aggressive behavior. *Journal of Pediatric Psychology*, 35(5), 547–558. <http://dx.doi.org/10.1093/jpepsy/jsp113>.
- Grant, K. A., McMahon, C., Austin, M. P., Reilly, N., Leader, L., & Ali, S. (2009). Maternal prenatal anxiety, postnatal caregiving and infants' cortisol responses to the still-face procedure. *Developmental Psychobiology*, 51(8), 625–637. <http://dx.doi.org/10.1002/dev.20397>.
- Grant, K. A., McMahon, C., Reilly, N., & Austin, M. P. (2010). Maternal sensitivity moderates the impact of prenatal anxiety disorder on infant mental development. *Early Human Development*, 86(9), 551–556. <http://dx.doi.org/10.1016/j.earlhumdev.2010.07.004>.
- Graziano, P., & Derefinko, K. (2013). Cardiac vagal control and children's adaptive functioning: A meta-analysis. *Biological Psychology*, 94(1), 22–37. <http://dx.doi.org/10.1016/j.biopsycho.2013.04.011>.
- Grossman, P., Karemaker, J., & Wieling, W. (1991). Prediction of tonic parasympathetic cardiac control using respiratory sinus arrhythmia: The need for respiratory control. *Psychophysiology*, 28(2), 201–216.
- Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, 74(2), 263–285. <http://dx.doi.org/10.1016/j.biopsycho.2005.11.014>.
- Grossman, P., van Beek, J., & Wientjes, C. (1990). A comparison of three quantification methods for estimation of respiratory sinus arrhythmia. *Psychophysiology*, 27(6), 702–714.
- Gueron-Sela, N., Propper, C. B., Wagner, N. J., Camerota, M., Tully, K. P., & Moore, G. A. (2017). Infant Respiratory Sinus Arrhythmia and Maternal Depressive Symptoms Predict Toddler Sleep Problems. *Developmental Psychobiology*, 59(2), 261–267.
- Gunning, M., Halligan, S. L., & Murray, L. (2013). Contributions of maternal and infant factors to infant responding to the still face paradigm: A longitudinal study. *Infant Behavior and Development*, 36(3), 319–328. <http://dx.doi.org/10.1016/j.infbeh.2013.02.003>.
- HAC. (2012). Rural research brief; Poverty in rural America. Housing Assistance Council. Retrieved from < http://www.ruralhome.org/storage/research_notes/rrn_poverty.pdf > .
- Haley, D. W., Handmaker, N. S., & Lowe, J. (2006). Infant stress reactivity and prenatal alcohol exposure. *Alcoholism, Clinical and Experimental Research*, 30(12), 2055–2064. <http://dx.doi.org/10.1111/j.1530-0277.2006.00251.x>.
- Haley, D. W., & Stansbury, K. (2003). Infant stress and parent responsiveness: Regulation of physiology and behavior during still-face and reunion. *Child Development*, 74(5), 1534–1546.
- Ham, J., & Tronick, E. (2006). Infant resilience to the stress of the still-face: Infant and maternal psychophysiology are related. *Annals of the New York Academy of Sciences*, 1094, 297–302. <http://dx.doi.org/10.1196/annals.1376.038>.
- Ham, J., & Tronick, E. (2009). Relational psychophysiology: Lessons from mother-infant physiology research on dyadically expanded states of consciousness. *Psychotherapy Research*, 19(6), 619–632. <http://dx.doi.org/10.1080/10503300802609672>.
- Hartley, D. (2004). Rural health disparities, population health, and rural culture. *American Journal of Public Health*, 94(10), 1675–1678.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hill, A., & Braungart-Rieker, J. M. (2002). Four-month attentional regulation and its prediction of three-year compliance. *Infancy*, 3, 261–273.
- Hill, L. K., Hu, D. D., Koenig, J., Sollers, J. J., 3rd, Kapuku, G., Wang, X., ... Thayer, J. F. (2015). Ethnic Differences in Resting Heart Rate Variability: A Systematic Review and Meta-Analysis. *Psychosomatic Medicine*, 77(1), 16–25.
- Holochwost, S. J., Garipey, J. L., Propper, C. B., Mills-Koonce, W. R., & Moore, G. A. (2014). Parenting behaviors and vagal tone at six months predict attachment disorganization at twelve months. *Developmental Psychobiology*, 56(6), 1423–1430. <http://dx.doi.org/10.1002/dev.21221>.
- Hostinar, C. E., Sullivan, R. M., & Gunnar, M. R. (2014). Psychobiological mechanisms underlying the social buffering of the hypothalamic-pituitary-adrenocortical axis: A review of animal models and human studies across development. *Psychological Bulletin*, 140(1), 256–282. <http://dx.doi.org/10.1037/a0032671>.
- Jansson, L. M., Dipietro, J. A., Elko, A., & Velez, M. (2010). Infant autonomic functioning and neonatal abstinence syndrome. *Drug and Alcohol Dependence*, 109(1–3), 198–204. <http://dx.doi.org/10.1016/j.drugalcdep.2010.01.004>.
- Jones-Mason, K., Allen, I. E., Bush, N., & Hamilton, S. (2016). Epigenetic marks as the link between environment and development: Examination of the associations between attachment, socioeconomic status, and methylation of the SLC6A4 gene. *Brain and Behavior*, 6(7), e00480. <http://dx.doi.org/10.1002/brb3.480>.
- Kisilevsky, B. S., Hains, S. M., Lee, K., Muir, D. W., Xu, F., Fu, G., ... Yang, R. L. (1998). The still-face effect in Chinese and Canadian 3- to 6-month-old infants. *Developmental Psychology*, 34(4), 629–639.
- Kogan, A., Oveis, C., Carr, E. W., Gruber, J., Mauss, I. B., Shallcross, A., ... Keltner, D. (2014). Vagal activity is quadratically related to prosocial traits, prosocial emotions, and observer perceptions of prosociality. *Journal of Personality and Social Psychology*, 107(6), 1051–1063. <http://dx.doi.org/10.1037/a0037509>.
- Kolb, B., & Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 20(4), 265–276.
- Laborde, S., Mosley, E., & Thayer, J. F. (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. <http://dx.doi.org/10.3389/fpsyg.2017.00213>.
- Lewis, G. F., Furman, S. A., McCool, M. F., & Porges, S. W. (2012). Statistical strategies to quantify respiratory sinus arrhythmia: Are commonly used metrics equivalent? *Biological Psychology*, 89(2), 349–364. <http://dx.doi.org/10.1016/j.biopsycho.2011.11.009>.
- Lindhiem, O., Bernard, K., & Dozier, M. (2011). Maternal sensitivity: Within-person variability and the utility of multiple assessments. *Child Maltreatment*, 16(1), 41–50. <http://dx.doi.org/10.1177/1077559510387662>.
- Mantis, I., Stack, D. M., Ng, L., Serbin, L. A., & Schwartzman, A. E. (2014). Mutual touch during mother-infant face-to-face still-face interactions: Influences of

- interaction period and infant birth status. *Infant Behavior and Development*, 37(3), 258–267. <http://dx.doi.org/10.1016/j.infbeh.2014.04.005>.
- Mattson, W. I., Ekas, N. V., Lambert, B., Tronick, E., Lester, B. M., & Messinger, D. S. (2013). Emotional expression and heart rate in high-risk infants during the face-to-face/still-face. *Infant Behavior and Development*, 36(4), 776–785. <http://dx.doi.org/10.1016/j.infbeh.2012.11.009>.
- McEwen, B. S. (2006). Protective and damaging effects of stress mediators: Central role of the brain. *Dialogues in Clinical Neuroscience*, 8(4), 367–381.
- McEwen, B. S. (2015). Biomarkers for Assessing Population and Individual Health and Disease Related to Stress and Adaptation. *Metabolism*, 64(3 Suppl 1), S2–S10.
- McLaughlin, K. A., Sheridan, M. A., Tibu, F., Fox, N. A., Zeanah, C. H., & Nelson, C. A., 3rd. (2015). Causal effects of the early caregiving environment on development of stress response systems in children. *Proceedings of the National Academy of Sciences of the United States of America*, 112(18), 5637–5642. <http://dx.doi.org/10.1073/pnas.1423363112>.
- Mendes, W. (2009). Assessing the autonomic nervous system. In E. Harmon-Jones, J. S. Beer (Eds.), *Methods in social neuroscience* (pp. 118–147). New York, NY: Guilford Press.
- Mesman, J., Linting, M., Joosen, K. J., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2013). Robust patterns and individual variations: Stability and predictors of infant behavior in the still-face paradigm. *Infant Behavior and Development*, 36(4), 587–598. <http://dx.doi.org/10.1016/j.infbeh.2013.06.004>.
- Mesman, J., van Ijzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2009). The many faces of the Still-Face Paradigm: A review and meta-analysis. *Developmental Review*, 29(2), 120–162. <http://dx.doi.org/10.1016/j.dr.2009.02.001>.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Journal of Clinical Epidemiology*, 62(10), 1006–1012. <http://dx.doi.org/10.1016/j.jclinepi.2009.06.005>.
- Montirosso, R., Provenzi, L., Giorda, R., Fumagalli, M., Morandi, F., Sirgiiovanni, I., ... Borgatti, R. (2016). SLC6A4 promoter region methylation and socio-emotional stress response in very preterm and full-term infants. *Epigenomics*, 8(7), 895–907. <http://dx.doi.org/10.2217/epi-2016-0010>.
- Montirosso, R., Provenzi, L., Tronick, E., Morandi, F., Reni, G., & Borgatti, R. (2014). Vagal tone as a biomarker of long-term memory for a stressful social event at 4 months. *Developmental Psychobiology*, 56(7), 1564–1574. <http://dx.doi.org/10.1002/dev.21251>.
- Moore, G. A. (2009). Infants' and mothers' vagal reactivity in response to anger. *Journal of Child Psychology and Psychiatry*, 50(11), 1392–1400. <http://dx.doi.org/10.1111/j.1469-7610.2009.02171.x>.
- Moore, G. A. (2010). Parent conflict predicts infants' vagal regulation in social interaction. *Development and Psychopathology*, 22(1), 23–33. <http://dx.doi.org/10.1017/S095457940999023X>.
- Moore, G. A., & Calkins, S. D. (2004). Infants' vagal regulation in the still-face paradigm is related to dyadic coordination of mother-infant interaction. *Developmental Psychology*, 40(6), 1068–1080. <http://dx.doi.org/10.1037/0012-1649.40.6.1068>.
- Moore, G. A., Cohn, J. F., & Campbell, S. B. (2001). Infant affective responses to mother's still face at 6 months differentially predict externalizing and internalizing behaviors at 18 months. *Developmental Psychology*, 37(5), 706–714. <http://dx.doi.org/10.1037/0012-1649.37.5.706>.
- Moore, G. A., Hill-Soderlund, A. L., Propper, C. B., Calkins, S. D., Mills-Koonce, W. R., & Cox, M. J. (2009). Mother-infant vagal regulation in the face-to-face still-face paradigm is moderated by maternal sensitivity. *Child Development*, 80(1), 209–223. <http://dx.doi.org/10.1111/j.1467-8624.2008.01255.x>.
- Nagy, E. (2008). Innate intersubjectivity: Newborns' sensitivity to communication disturbance. *Developmental Psychology*, 44(6), 1779–1784. <http://dx.doi.org/10.1037/a0012665>.
- Oberlander, T. F., Jacobson, S. W., Weinberg, J., Grunau, R. E., Molteno, C. D., & Jacobson, J. L. (2010). Prenatal alcohol exposure alters biobehavioral reactivity to pain in newborns. *Alcoholism, Clinical and Experimental Research*, 34(4), 681–692. <http://dx.doi.org/10.1111/j.1530-0277.2009.01137.x>.
- Oosterman, M., De Schipper, J. C., Fisher, P., Dozier, M., & Schuengel, C. (2010). Autonomic reactivity in relation to attachment and early adversity among foster children. *Development and Psychopathology*, 22(1), 109–118. <http://dx.doi.org/10.1017/S0954579409990290>.
- Ostfeld-Etzion, S., Golan, O., Hirschler-Guttenberg, Y., Zagoory-Sharon, O., & Feldman, R. (2015). Neuroendocrine and behavioral response to social rupture and repair in preschoolers with autism spectrum disorders interacting with mother and father. *Molecular Autism*, 6, 11. <http://dx.doi.org/10.1186/s13229-015-0007-2>.
- Ostlund, B. D., Measelle, J. R., Laurent, H. K., Conradt, E., & Ablow, J. C. (2017). Shaping emotion regulation: Attunement, symptomatology, and stress recovery within mother-infant dyads. *Developmental Psychobiology*, 59(1), 15–25. <http://dx.doi.org/10.1002/dev.21448>.
- Paret, L., Bailey, H. N., Roche, J. F., Bureau, J. F., & Moran, G. (2015). Preschool ambivalent attachment associated with a lack of vagal withdrawal in response to stress. *Attachment & Human Development*, 17(1), 65–82. <http://dx.doi.org/10.1080/14616734.2014.967786>.
- Pederson, D. R., & Moran, G. (1995). Appendix B. Maternal behavior Q-set (E. Waters, B. E. Vaughn, G. Poseda, & K. Kondo-Ikemura Eds. Vol. 60 (2–3, Serial No. 244)).
- Planalp, E. M., & Braungart-Rieker, J. M. (2013). Temperamental precursors of infant attachment with mothers and fathers. *Infant Behavior and Development*, 36, 686–808.
- Porges, S. W. (1995). Cardiac vagal tone: A physiological index of stress. *Neuroscience & Biobehavioral Reviews*, 19(2), 225–233.
- Porges, S. W. (1996). Physiological regulation in high-risk infants: A model for assessment and potential intervention. *Development and Psychopathology*, 8, 43–58.
- Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology*, 74(2), 116–143. <http://dx.doi.org/10.1016/j.biopsycho.2006.06.009>.
- Porges, S. W. (1985). Washington, DC: U.S. Patent No. 4,510,944. U. S. P. Office.
- Porges, S. W. (2011). The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation. (W. W. N. Company, Ed.). New York, NY.
- Pratt, M., Singer, M., Kanat-Maymon, Y., & Feldman, R. (2015). Infant negative reactivity defines the effects of parent-child synchrony on physiological and behavioral regulation of social stress. *Development and Psychopathology*, 27(4 Pt 1), 1191–1204. <http://dx.doi.org/10.1017/S0954579415000760>.
- Propper, C. (2012). *The early development of vagal tone: Effects of poverty and elevated contextual risk*. New York, NY: Oxford University Press.
- Propper, C., & Holochowst, S. J. (2013). The influence of proximal risk on the early development of the autonomic nervous system. *Developmental Review*, 33, 151–167.
- Propper, C., & Moore, G. A. (2006). The influence of parenting on infant emotionality: A multi-level psychobiological perspective. *Developmental Review*, 26, 427–460.
- Propper, C., Moore, G. A., Mills-Koonce, W. R., Halpern, C. T., Hill-Soderlund, A. L., Calkins, S. D., ... Cox, M. (2008). Gene-environment contributions to the development of infant vagal reactivity: The interaction of dopamine and maternal sensitivity. *Child Development*, 79(5), 1377–1394. <http://dx.doi.org/10.1111/j.1467-8624.2008.01194.x>.
- Provenzi, L., Casini, E., de Simone, P., Reni, G., Borgatti, R., & Montirosso, R. (2015). Mother-infant dyadic reparation and individual differences in vagal tone affect 4-month-old infants' social stress regulation. *Journal of Experimental Child Psychology*, 140, 158–170. <http://dx.doi.org/10.1016/j.jecp.2015.07.003>.
- Provenzi, L., Giusti, L., & Montirosso, R. (2016). Do infants exhibit significant cortisol reactivity to the Face-to-Face Still-Face paradigm? A narrative review and meta-analysis. *Developmental Review*, 42, 34–55. <http://dx.doi.org/10.1016/j.dr.2016.07.001>.
- Provenzi, L., Olson, K. L., Montirosso, R., & Tronick, E. (2016). Infants, Mothers, and Dyadic Contributions to Stability and Prediction of Social Stress Response at 6 Months. *Developmental Psychology*, 52(1), 1–8.
- Quas, J. A., Yim, I. S., Oberlander, T. F., Nordstokke, D., Essex, M. J., Armstrong, J. M., ... Boyce, W. T. (2014). The symphonic structure of childhood stress reactivity: Patterns of sympathetic, parasympathetic, and adrenocortical responses to psychological challenge. *Development and Psychopathology*, 26(4 Pt 1), 963–982. <http://dx.doi.org/10.1017/S0954579414000480>.
- Quigley, K. M., Moore, G. A., Propper, C. B., Goldman, B. D., & Cox, M. J. (2016). Vagal regulation in breastfeeding infants and their mothers. *Child Development*. <http://dx.doi.org/10.1111/cdev.12641>.
- Rash, J. A., Thomas, J. C., Campbell, T. S., Letourneau, N., Granger, D. A., Giesbrecht, G. F., & Team, A. P. S. (2016). Developmental origins of infant stress reactivity profiles: A multi-system approach. *Developmental Psychobiology*. <http://dx.doi.org/10.1002/dev.21403>.
- Ritz, T., Bosquet Enlow, M., Schulz, S. M., Kitts, R., Staudenmayer, J., & Wright, R. J. (2012). Respiratory sinus arrhythmia as an index of vagal activity during stress in infants: Respiratory influences and their control. *PLoS One*, 7(12), e52729. <http://dx.doi.org/10.1371/journal.pone.0052729>.
- Sapolsky, R. (2004). *Why zebras don't get ulcers*. New York, NY: Henry Holt & Co.
- Selye, H. (1956). *The stress of life*. New York, NY: McGraw-Hill.
- Shahrestani, S., Stewart, E. M., Quintana, D. S., Hickie, I. B., & Guastella, A. J. (2014). Heart rate variability during social interactions in children with and without psychopathology: A meta-analysis. *Journal of Child Psychology and Psychiatry*, 55(9), 981–989. <http://dx.doi.org/10.1111/jcpp.12226>.
- Sharp, H., Pickles, A., Meaney, M., Marshall, K., Tibu, F., & Hill, J. (2012). Frequency of infant stroking reported by mothers moderates the effect of prenatal

- depression on infant behavioural and physiological outcomes. *PLoS One*, 7(10), e45446. <http://dx.doi.org/10.1371/journal.pone.0045446>.
- Shonkoff, J. P., Boyce, W. T., & McEwen, B. S. (2009). Neuroscience, molecular biology, and the childhood roots of health disparities: Building a new framework for health promotion and disease prevention. *JAMA*, 301(21), 2252–2259. <http://dx.doi.org/10.1001/jama.2009.754>.
- Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment & Human Development*, 7(4), 349–367. <http://dx.doi.org/10.1080/14616730500365928>.
- Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.). *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 99–110). Chichester, England: Wiley.
- Stevenson-Hinde, J., & Marshall, P. J. (1999). Behavioral inhibition, heart period, and respiratory sinus arrhythmia: An attachment perspective. *Child Development*, 70(4), 805–816.
- Stewart, A. M., Lewis, G. F., Heilman, K. J., Davila, M. I., Coleman, D. D., Aylward, S. A., & Porges, S. W. (2013). The covariation of acoustic features of infant cries and autonomic state. *Physiology & Behavior*, 120, 203–210. <http://dx.doi.org/10.1016/j.physbeh.2013.07.003>.
- Stoller, S., & Field, T. (1982). Alteration of mother and infant behavior and heart rate during a still-face perturbation of face-to-face interaction. In T. Field, & A. Fogel (Eds.). *Emotion and early interaction* (pp. 57–82). Hillsdale, NJ: Lawrence Erlbaum.
- Suurland, J., van der Heijden, K. B., Smaling, H. J., Huijbregts, S. C., van Goozen, S. H., & Swaab, H. (2016). Infant autonomic nervous system response and recovery: Associations with maternal risk status and infant emotion regulation. *Development and Psychopathology*, 1–15. <http://dx.doi.org/10.1017/S0954579416000456>.
- Tarabulsy, G. M., Provost, M. A., Deslandes, J., St. Laurent, D., Moss, E., Lemelin, J., ... Dassylva, J. (2003). Individual differences in still-face response at 6 months. *Infant Behavior & Development*, 26(3), 421–438.
- TFESOC. (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. *European Heart Journal*, 17(3), 354–381.
- Tibu, F., Hill, J., Sharp, H., Marshall, K., Glover, V., & Pickles, A. (2014). Evidence for sex differences in fetal programming of physiological stress reactivity in infancy. *Development and Psychopathology*, 26(4 Pt 1), 879–888. <http://dx.doi.org/10.1017/S0954579414000194>.
- Toda, S., & Fogel, A. (1993). Infant response to the still-face situation at 3 and 6 months. *Developmental Psychology*, 29(3), 532–538.
- Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, 17(1), 1–13.
- Tronick, E., & Beeghly, M. (2011). Infants' meaning-making and the development of mental health problems. *American Psychologist*, 66(2), 107–119. <http://dx.doi.org/10.1037/a0021631>.
- Tronick, E., Messinger, D. S., Weinberg, M. K., Lester, B. M., Lagasse, L., Seifer, R., ... Liu, J. (2005). Cocaine exposure is associated with subtle compromises of infants' and mothers' social-emotional behavior and dyadic features of their interaction in the face-to-face still-face paradigm. *Developmental Psychology*, 41(5), 711–722. <http://dx.doi.org/10.1037/0012-1649.41.5.711>.
- Tronick, E., & Weinberg, M. K. (1990). Emotion Regulation in Infancy: Stability of Regulatory Behaviors. In: *Paper presented at the International Conference on Infant Studies*, Montreal, Quebec, Canada.
- Tronick, E., & Weinberg, M. K. (1997). *Depressed mothers and infants: Failure to form dyadic states of consciousness*. New York, NY: Guilford Press.
- van Dijk, A. E., van Eijsden, M., Stronks, K., Gemke, R. J., & Vrijlkotte, T. G. (2012). Prenatal stress and balance of the child's cardiac autonomic nervous system at age 5–6 years. *PLoS One*, 7(1), e30413. <http://dx.doi.org/10.1371/journal.pone.0030413>.
- van Lien, R., Neijts, M., Willemsen, G., & de Geus, E. J. (2015). Ambulatory measurement of the ECG T-wave amplitude. *Psychophysiology*, 52(2), 225–237. <http://dx.doi.org/10.1111/psyp.12300>.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. Retrieved from *Journal of Statistical Software*, 36(3), 1–48. <http://www.jstatsoft.org/v36/i03/>.
- Wagner, N. J., Propper, C., Gueron-Sela, N., & Mills-Koonce, W. R. (2015). Dimensions of maternal parenting and infants' autonomic functioning interactively predict early internalizing behavior problems. *Journal of Abnormal Child Psychology*. <http://dx.doi.org/10.1007/s10802-015-0039-2>.
- Waters, S. F., West, T. V., & Mendes, W. B. (2014). Stress contagion: Physiological covariation between mothers and infants. *Psychological Science*, 25(4), 934–942. <http://dx.doi.org/10.1177/0956797613518352>.
- Weinberg, M. K., & Tronick, E. Z. (1996). Infant affective reactions to the resumption of maternal interaction after the still-face. *Child Development*, 67(3), 905–914.
- Weinberg, M. K., Tronick, E., Cohn, J. F., & Olson, K. L. (1999). Gender differences in emotional expressivity and self-regulation during early infancy. *Developmental Psychology*, 35(1), 175–188.
- Weisman, O., Zagoory-Sharon, O., & Feldman, R. (2012). Oxytocin administration to parent enhances infant physiological and behavioral readiness for social engagement. *Biological Psychiatry*, 72(12), 982–989. <http://dx.doi.org/10.1016/j.biopsych.2012.06.011>.
- Yazbek, A., & D'Entremont, B. (2006). A longitudinal investigation of the still-face effect at 6 months and joint attention at 12 months. *British Journal of Developmental Psychology*, 24(3), 589–601.
- Zisner, A. R., & Beauchaine, T. P. (2016). Psychophysiological methods and developmental psychopathology. In *Developmental psychopathology* (Vol. 2, pp. 1–53).