Twin Research and Human Genetics
Volume 19 ▌ Number 5 ▌ pp. 456–464 ▌ © The Author(s) 2016 ▌ doi:10.1017/thg.2016.59

The Twin Study of Negative Valence Emotional Constructs

Dever M. Carney,[1] Elizabeth Moroney,[2] Laura Machlin,[2] Shannon Hahn,[1] Jeanne E. Savage,[1] Minyoung Lee,[1] Kenneth A. Towbin,[2] Melissa A. Brotman,[2] Daniel S. Pine,[2] Ellen Leibenluft,[2] Roxann Roberson-Nay,[1] and John M. Hettema[1]

[1]Virginia Institute for Psychiatric and Behavioral Genetics and Department of Psychiatry, Virginia Commonwealth University, Richmond, Virginia, USA
[2]Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland, USA

The Twin Study of Negative Valence Emotional Constructs is a multi-site study designed to examine the relationship between a broad selection of potential measures designed to assess putative endophenotypes for negative valence systems (NVS) and early symptoms of internalizing disorders (IDs). In this article, we describe the sample characteristics, data collection protocols, and measures used. Pre-adolescent Caucasian twin pairs were recruited through the Mid-Atlantic Twin Registry; data collection began in February of 2013. Enrolled twins completed various dimensional self-report measures along with cognitive, emotional, and psychophysiological tasks designed to assess NVS function. Parents also completed surveys about their twins and themselves. In addition, a subset of the twins also participated in a neuroimaging protocols. Data collection is in the final stages, and preliminary analyses are underway. The findings will potentially expand our understanding of the mechanisms by which genetic and environmental factors contribute to individual differences in NVS phenotypes and provide new insights into underlying risk factors for IDs.

■ **Keywords:** negative valence, anxiety, internalizing, twin study

The Research Domain Criteria initiative (RDoC) is part of the National Institutes of Mental Health (NIMH) plan to provide a new framework for conceptualizing mental disorders. As such, projects funded under RDoC are intended to develop 'new ways of classifying disorders based on dimensions of observable behaviors and brain functions' (Insel et al., 2010). The current study was sponsored through the first RDoC funding request (RFA-MH-12-100) and aims to provide an integrated perspective of genetic and environmental aspects of negative valence systems (NVS) in relation to internalizing disorders (IDs), which include mood and anxiety disorders. As characterized by the RDoC theoretical framework, the NVS domain encompasses biological and psychological systems involved in the response to aversive, threatening, or harmful stimuli, including systems that have a putative relationship to IDs.

Neuroscientists and psychologists have developed dimensional assessment instruments and laboratory tasks that probe cognitive, emotional, biological, and/or behavioral aspects of NVS as phenotypes for study as alternatives to ID clinical symptoms. Such intermediate or 'endo' phenotypes might reflect processes more proximal to gene expression than clinical symptoms while also playing an important role in the development of pathological emotional states. Gottesman and Gould (2003) suggested criteria for a putative endophenotype, including (a) association with illness, (b) heritability, and (c) co-segregation with the illness in families.

Early studies suggest that various putative endophenotypic measures might index shared or specific components of ID risk. One hypothesis is that rather than showing associations with specific diagnoses, such measures might be more likely to align with broader constructs proposed in the RDoC matrix; for example, acute threat (fear), potential threat (anxiety), and elements of negative affect (loss, anhedonia, frustrative non-reward). However, many such endophenotypic measures have been examined only in relationship to clinical diagnoses via case-control studies, rather than using more broadly informative, unselected samples. Furthermore, some progress has been made in

RECEIVED 5 May 2016; ACCEPTED 30 May 2016. First published online 26 July 2016.

ADDRESS FOR CORRESPONDENCE: John M. Hettema, MD, PhD, VCU Department of Psychiatry, PO Box 980126, Richmond, VA 23298-0126, USA. E-mail: john.hettema@vcuhealth.org

establishing that such measures satisfy some version of criterion (a), with much less data available regarding criteria (b) and (c) (Savage et al., 2016). Knowledge about the external predictors and underlying sources of variance of endophenotypes, as gained via epidemiologically and genetically informative studies, is still in progress. In particular, it is not clear how these constructs will map onto the genetic risk structure of IDs in developing children.

Twins are ideal for determining the differential effects of genes and environment and their shared and specific contributions across phenotypes (Kendler, 2001). Under RDoC's specified 'Units of Analysis', the Twin Study of Negative Valence Emotional Constructs, also referred to as the Virginia Commonwealth University Juvenile Anxiety Study (VCU-JAS), measures genetic risk factors, self-report, behavior, physiology, and paradigms, as well as specific childhood environmental risk factors in relation to NVS. The measures and experimental tasks were chosen for inclusion in this study in support of its six major aims:

1. Via subject level phenotypic analyses, estimate the response structure and underlying latent constructs of a broad suite of possible endophenotypes, including dimensional measures and psychological tasks that probe negative valence emotional states in developing children.
2. Examine the relationships between the higher order constructs identified in Aim 1 and internalizing symptom clusters.
3. Apply twin modeling approaches to determine the genetic and environmental sources of each putative endophenotypic measure as well as their composite multivariate risk structure.
4. Estimate the degree to which common and specific genetic factors identified in Aim 3 overlap the genetic risk structure of IDs.
5. Estimate the role of well-established developmental risk factors for IDs in predicting endophenotypic construct measures and how they moderate the relationships with internalizing symptoms.
6. Analyze associations and genetic correlations between brain-based phenotypes (structure and neural activity) and the self-report and lab-based measures studied in prior aims.

Materials and Methods

Sample

Recruitment of twin pairs occurred through the Mid-Atlantic Twin Registry (MATR), a Virginia Commonwealth University (VCU) database comprised of twins, other higher order multiples, and their family members who were willing to consider participating in research (Lilley & Silberg, 2013). MATR personnel contacted the parents of potentially eligible twins in accordance with their standard operating procedures. They gathered basic demographic information and conducted a brief screen of exclusionary criteria and baseline data that could be used for study assignments. If twins were eligible and a parent agreed to participation of their children, they were enrolled in the study and their contact information was given to the study's project coordinator. In addition to inviting twins already registered in MATR, IRB-approved advertisements and other communications helped ascertain additional twins for this study. Twin pairs were considered ineligible for the study if either child met any of the following exclusion criteria: severe or unstable medical or neurological illness, past seizures without a clear or resolved etiology, intellectual disabilities, substance abuse, recent thoughts of suicide or homicide, episodes of psychosis, or currently taking psychotropic medications besides stimulants for ADHD. These exclusions aimed to (1) decrease the risk of exacerbation of possible medical conditions during the potentially stressful laboratory tasks and (2) minimize the likelihood that physiological responses recorded during the tasks were confounded by the potential effects of these variables. This study was approved by the VCU Institutional Review Board.

Recruitment targeted a sample of 450 twin pairs aged 9–13 years residing in the Mid-Atlantic region around Virginia and Maryland. The transition between childhood and adolescence is a key period for the emergence and progression of IDs, allowing the examination of their developmental aspects to infer underlying mechanisms (Kessler et al., 2005). For the purposes of biometrical twin modeling, an unselected sample was recruited that should produce an epidemiological distribution of internalizing syndromes. Given the high population prevalence of internalizing symptoms in children, this strategy should produce substantial numbers of children at risk for the target ID symptomatology: generalized anxiety (GAD), social anxiety (SOC), separation anxiety (SAD), panic/somatization (PAN), phobias (PHOB), and depression (DEP). This design provides adequate power to detect significant additive genetic effects explaining around 40% of the variance (heritability) for the proposed quantitative phenotypes (Visscher, 2004).

The sample was limited to Caucasian families, the largest available racial group in the local region. The authors are aware of no twin studies that have examined racial differences in heritability of internalizing phenotypes or their genetic risk structure. Preliminary analyses thus require stratification by race, resulting in the dilution of power to detect genetic effects that might not be shared across races. The inclusion of multiple racial groups would require a substantially expanded timeline and additional resources in order to recruit a sufficiently large, diverse sample. Future molecular analyses would also need to be performed separately by race due to differing allele frequencies across ancestral populations that can induce spurious association results. The investigators believed it prudent to stick with a more

homogeneous sample for this study funded early under the RDoC initiative and, if indicated, plan for a replication study in a more racially diverse sample.

Protocols

After obtaining written informed consent from parents and assent from minor children, research protocols were conducted by trained research assistants in laboratory settings at VCU in Richmond, Virginia and at the Intramural Program of the National Institute of Mental Health (NIMH-IRP), part of the National Institutes of Health in Bethesda, Maryland. Involvement of the NIMH-IRP allowed the study to recruit subjects from areas in close proximity to Washington, DC. All data collected during these assessments were de-identified and given a study-generated participant identification number. Twins and willing parents underwent the full assessment of primary study measures during Visit 1. Willing families were brought back to complete one of several possible protocols during Visit 2: (1) participants were administered a reduced set of the same assessments from Visit 1 to examine their test–retest reliability (approximately 2–4 weeks after Visit 1); (2) neuroimaging data were collected from eligible twin participants. Parents and children were financially compensated for their participation.

Measures

A broad suite of NVS-relevant measures were included to examine their reliability, heritability, covariance structure, and how individual differences in these measures contribute to ID symptoms and outcomes. They are described in detail below and listed in Table 1 together with the NVS constructs and phenotypes they assess. The selected measures are multimethod–multitrait and included self-report surveys completed by the twins, parent report measures on the children about ID symptoms and risk factors, parental self-report on their own psychiatric history, and laboratory and neuroimaging paradigms.

In addition to these phenotypic measures available for biometrical twin modeling, salivary DNA samples were collected using Oragene kits (DNA Genotek). DNA was extracted and stored for future genetic association studies that explore the relationship between NVS phenotypes derived from this study and novel genetic loci identified by ongoing genome-wide association studies of IDs (Otowa et al., 2016; Ripke et al., 2013).

Survey Measures

Several extant survey instruments were selected that cover a broad range of NVS phenotypic domains available by child self-report, parent report, or both. Survey data were collected and managed using REDCap electronic data capture tools hosted at VCU (Harris et al., 2009). Among the child report surveys, clinical symptoms were assessed using the Short Mood and Feelings Questionnaire (SMFQ; Angold et al., 1995), Screening for Childhood Anxiety Related Disorders (SCARED; Birmaher et al., 1997), Affective Reactivity Index (ARI; Stringaris et al., 2012), and the Fear Survey Schedule for Children Revised (FSSC-R; Ollendick, 1983). Normal personality and various indices of anxious temperament were assessed using the Junior Eysenck Personality Questionnaire (JEPQ; Eysenck, 1965), Childhood Anxiety Sensitivity Index (CASI; Silverman et al., 1991), Behavioral Inhibition System/Behavioral Activation System (BIS-BAS; Carver & White, 1994), and the Threat and Fearlessness Questionnaire (TF-20, a shorter version of the TF-55 described in Kramer et al., 2012). In addition, the Multidimensional Peer Victimization Scale (MPVS; Mynard & Joseph, 2000) and sex-specific assessments of pubertal stage were administered.

Parents completed the following surveys about each of their twin children: SMFQ-Parent, SCARED-Parent, ARI-Parent, Child Behavior Checklist (CBCL; Achenbach, 1991), Retrospective Behavioral Inhibition Questionnaire (RBIQ; Reznick et al., 1992) modified for this age group, Inventory of Callous-Unemotional Traits (ICU; Kimonis et al., 2008), Parental Bonding Instrument (PBI; Parker et al., 1979), Traumatic Events Screening Inventory – Parent Report Revised (TESI-PRR; http://www.ptsd.va.gov/professional/assessment/child/tesi.asp). A masters- or doctoral-level trained clinician administered the Kiddie Schedule for Affective Disorders and Schizophrenia-Present & Lifetime Version (KSADS-PL; Kaufman et al., 1997) about each twin during a face-to-face interview with the parent(s); the children were also assessed with the KSADS-PL at the NIMH as part of their standard protocol. In addition, parents answered questions about their twins shown to be informative for estimating zygosity in prior twin studies, such as frequency with which the twins are confused by others, degree of physical similarity, and blood types or DNA tests (if available). A zygosity algorithm was developed for the current study to create a pair (dis)similarity score from those questions, with each item weighted by its relative discriminant predictive ability (Nichols & Bilbro, Jr., 1966; Peeters et al., 1998). Zygosity assignments derived from this algorithm showed high levels of concordance with zygosity determined by DNA testing ($\kappa = 1.0$, $n = 13$) or a medical diagnosis from placental sharing, blood tests, or in-vitro fertilization information ($\kappa = 0.91$, $n = 112$) per maternal report.

Parents completed the following surveys about themselves: short form of the Eysenck Personality Questionnaire (EPQ-SF; Eysenck & Eysenck, 1975), stressful personal and network life events that occurred in the prior 12 months (adapted from an adult twin study; Kendler et al., 1998), and questions about medically relevant events that occurred during the pregnancy and birth of the children. Willing parents also completed an online clinical interview about their own psychiatric history adapted from the CIDI-SF (Kessler et al., 1998).

TABLE 1

NVS Constructs, Phenotypes, Units of Analysis, and Measures

Construct	Phenotype	Units of analysis	Measure
Acute threat	Fears	Self-reports	FSSC-R
			TF-20
	Phobias*	Self-reports	KSADS-PL
	Fear acquisition	Physiology	FPS
			SCR
	Interoceptive hypersensitivity	Physiology	CO_2 inhalation
	Amygdala Reactivity	Circuits	EFMT
Potential threat	Anxiety	Self-reports	SCARED
	Anxiety disorders*	Self-reports	KSADS-PL
	Anticipatory anxiety	Physiology	Baseline startle
Sustained threat	PTSD*	Self-reports	KSADS-PL
	Attentional bias to threat	Paradigms	FEP
	Emotion labeling	Paradigms	FELT
	Fear extinction	Paradigms	ER
		Circuits	
Loss	Sadness; anhedonia	Self-reports	SMFQ
	Depression*	Self-reports	KSADS-PL
	Attentional bias to sadness	Paradigms	FEP
Frustrative Non-reward	Irritability	Self-reports	ARI
	Frustration	Paradigms	AP2
		Circuits	
Temperament*	Neuroticism, extraversion	Self-reports	JEPQ
	Anxiety sensitivity	Self-reports	CASI
	Behavioral inhibition/activation	Behaviors	RBIQ
			BIS/BAS
	Unemotionality	Self-reports	ICU
Risk/protective factors *	Parenting	Self-reports	PBI
	Peer victimization		MPVS
	Life events/trauma		TESI
	Sex	Physiology	
	Age		
	Pubertal status		

Note: *Not formally included in RDoC matrix. SCARED = screening for childhood anxiety related disorders; SMFQ = short mood and feelings questionnaire; JEPQ = junior eysenck personality questionnaire; CASI = childhood anxiety sensitivity index; ARI = affective reactivity index; BIS/BAS = behavioral inhibition system behavioral activation system; FSSC-R = fear survey schedule for children revised; TF-20 = threat and fearlessness questionnaire - 20-item version; ICU = inventory of callous-unemotional Traits; RBIQ=retrospective behavioral inhibition questionnaire; PBI = parental bonding instrument- Authoritarianism/Coldness/Protectiveness; MPVS = multidimensional peer victimization scale; TESI-PRR = traumatic events screening inventory - parent report revised; KSADS-PL = kiddie schedule for affective disorders and schizophrenia - present & lifetime Version; FEP = face-emotion processing; FELT = facial expression labeling task; FPS = fear-potentiated startle; SCR = skin conductance response; AP2 = affective posner 2; EFMT = emotional face matching task; ER = extinction recall.

Laboratory Paradigms

Twin participants completed five experimental tasks during Visit 1 that probe cognitive, emotional, and psychophysiological components of NVS. Some twins repeated these during Visit 2 to estimate test–retest reliability. Reliability is important for accurately assessing heritability of a measure and has not yet been examined for most of these paradigms. Participants completed tasks in a sequence according to one of four experimental schedules; this was determined randomly before the participants arrived and conserved across twins in a pair. The four conditions attempted to control for order effects while preserving the chronology of certain tasks mandated by the study protocols.

The Face-Emotion Processing (FEP) task utilizes the fact that the processing of facial stimuli is one of the most fundamental human capabilities and an essential part of social communication (Philippot & Feldman, 1990). Various FEP tasks have been used to study psychopathology and associated traits, each attempting to capitalize on the finding that facial expressions engage a neural circuit involved in a core, evolutionarily based system independent of cultures (Haxby et al., 2002). The version included in this study engages two cognitive processes, attention, and recognition memory, each of which exhibits somewhat distinct associations with psychopathology (Guyer et al., 2011; Pine et al., 2005). During the FEP task, participants viewed a series of standardized photographs of 32 actors expressing one of four emotions (angry, fearful, happy, sad) displayed for 4 s. During these viewings, participants made one of three ratings: degree of emotion expressed (How sad is the face?), the subject's internal emotional response (How sad does this person make you feel?), and the size of a non-emotional facial feature (How wide is the nose?). Thirty minutes after the FEP task, a surprise recognition memory test was administered. Pairs of two faces displaying neutral expressions, one of which had appeared in the FEP task and one of

which was novel, were shown side by side, and participants identified which of the two faces they had previously seen.

The Facial Expression Labeling Task (FELT) assesses the participant's ability to read emotions in others. Youth with psychiatric syndromes often experience social impairment indexed by their ability to read and interpret emotions (Geller et al., 2000; Rich et al., 2008), with information-processing biases that may vary across diagnostic groups (Dalgleish et al., 2003). The current study included a FELT paradigm based on a study by Marsh et al. (2010). Participants were shown faces expressing six basic emotions: anger, disgust, fear, happiness, sadness, or surprise. Participants viewed faces for 500 ms morphed at 10% increments of emotional intensity, ranging from a neutral face (0% intensity) to full emotional (100%) intensity. They were then asked to label the emotion from the six possible choices. Response selections and response latencies were recorded, and percentage accuracy scores were calculated for each emotion at each intensity level.

A fear-potentiated startle (FPS) paradigm was used to assess fear conditioning and extinction. The startle reflex is a psychophysiological measure putatively sensitive to individual differences in emotional reactivity; it is readily measured in humans by recording the eye-blink electromyographic (EMG) response (Landis, 1939). Startle is potentiated when elicited in the presence of a stimulus that signals an aversive stimulus like shock (Davis, 1986; Grillon & Baas, 2003). The current study employed a FPS paradigm developed for use with children in which participants viewed photographs of two women: one woman serving as the positive conditioned stimulus (CS^+) and the other as the CS^- (Britton et al., 2011). The CS^+ is paired with a loud, piercing scream as the unconditioned stimulus (UCS). Twins were conditioned to one of the women counterbalanced across twins. Not all CS^+ presentations were reinforced with the UCS, thus creating a level of threat unpredictability. Predictable and unpredictable threats putatively distinguish between phasic fear and sustained anxiety (Schmitz et al., 2011) with differential association to anxiety disorders (Grillon et al., 2008). The experimental phases were: habituation, pre-acquisition, acquisition, and extinction. In addition to startle EMG and self-reported distress, skin conductance response (SCR), and electrocardiogram (ECG) were recorded throughout the paradigm.

The carbon dioxide (CO_2) inhalation task assesses physiologic and emotional responses when breathing air containing increased concentrations of CO_2. The CO_2 inhalation procedure has been used extensively by clinical researchers to study individual variation in response to the experience of bodily sensations (interoception), particularly those sensations tied to anxiety sensitivity and panic attacks (Bailey et al., 2005; Papp et al., 1993). CO_2 offers several advantages over other 'panicogens', including safety, ease of administration, tolerability, and reliability. The two most commonly applied CO_2 administration protocols are sustained inhalation of air enriched with lower concentrations of CO_2 (5% or 7.5%) versus one or two vital capacity inhalations of air containing high CO_2 concentration (35%). Because the latter is substantially more aversive, the milder respiratory stimulating effects of sustained administration of low-concentration CO_2 are preferred with children (Pine et al., 2000). This study used a three-phase protocol: a 5-minute baseline breathing room air, 8-minutes breathing 7.5% CO_2 enriched air, and a 5-minute recovery period. Participants were unaware of when the CO_2 was turned on and off. The Diagnostic Symptom Questionnaire (DSQ) and Subjective Units of Distress (SUDS) ratings were administered at multiple time points to track anxiety and symptomatic reactivity before, during, and after the CO_2 challenge. Physiologic indices of respiratory rate, tidal volume, and end-tidal CO_2, as well as SCR, heart rate, and heart-rate variability were continuously monitored throughout the paradigm.

TABLE 2

Sample Characteristics (n = 398 Families, 796 Twin Children)

Demographic	Value
Age in years: mean (*SD*)	11.20 (1.42)
Female sex: *n* (%)	419 (52.6%)
Zygosity groups: *n* (%)	
MZ male	56 (14.1%)
DZ male	66 (16.6%)
MZ female	72 (18.1%)
DZ female	68 (17.1%)
DZ opposite sex	111 (27.9%)
Unknown/missing	21 (5.4%)

Note: *SD* = standard deviation; MZ = monozygotic; DZ = dizygotic.

The Affective Posner 2 (AP2) task was adapted from a task used in previous studies (Deveney et al., 2013; Rich et al., 2010). It is designed to elicit feelings of frustration in pediatric samples with a particular focus on studying irritability in children. The task provides rigged feedback in the context of a reward task, thereby provoking feelings of frustration. Deception was involved, and the children were debriefed at the end of the task.

Neuroimaging Paradigms

A subset of eligible twin pairs underwent neuroimaging protocols using 3T magnetic resonance imaging (MRI) during Visit 2 at NIMH or VCU. High-resolution structural and resting state functional scans were conducted during each session. Depending on the site and timing of the visit, some participants also performed a mix of the following three tasks while functional MRI (fMRI) data was acquired: the AP2 task, an emotional face-matching task (EFMT; Hariri et al., 2000), modified according to Swartz et al. (2014), and the extinction recall (ER) portion of the previously described FPS fear-conditioning task (Britton et al., 2013).

Results

Table 2 lists key demographics for the sample. To date, the study has acquired data from 398 twin pairs (796

TABLE 3
Statistics for Sum Scores of Survey Measures By Visit

		Visit 1			Visit 2			Visit 1–Visit 2
Reporter	Measure	*n*	Mean (*SD*)	Inter-item reliability (Cronbach's α)	*n*	Mean (*SD*)	Inter-item reliability (Cronbach's α)	Test–retest reliability (correlation)
Child	SCARED	775	24.1 (11.8)	0.90	266	20.7 (12.8)	0.93	0.78
	SMFQ	781	5.23 (4.17)	0.81	266	4.98 (4.34)	0.84	0.61
	JEPQ-E	775	17.3 (4.14)	0.77	267	17.3 (4.69)	0.83	0.78
	JEPQ-N	775	8.38 (4.83)	0.84	267	8.16 (5.11)	0.86	0.78
	CASI	776	10.3 (5.96)	0.84	266	8.69 (6.07)	0.87	0.67
	ARI	718	3.47 (2.98)	0.82	265	3.25 (3.23)	0.86	0.78
	BISBAS-A	762	19.4 (6.09)	0.80	265	18.5 (6.59)	0.83	0.66
	BISBAS-I	762	9.0 (4.11)	0.77	265	8.44 (4.35)	0.78	0.70
	FSSC-R	773	53.9 (26.5)	0.96	261	47.1 (27.5)	0.96	0.83
	TF-20	748	29.4 (10.5)	0.87	251	28.5 (12.1)	0.91	0.81
Parent	SCARED	769	12.5 (11.0)	0.94	265	11.4 (11.7)	0.96	0.89
	SMFQ	767	1.92 (3.07)	0.85	265	1.80 (3.13)	0.85	0.54
	ARI	722	1.76 (2.61)	0.88	265	1.67 (2.48)	0.87	0.75
	ICU	741	17.6 (7.59)	0.82	255	16.9 (7.92)	0.84	0.80
	RBIQ	767	93.1 (35.6)	0.96	265	90.4 (39.1)	0.97	0.90
	PBI-A	741	3.96 (2.15)	0.63	255	3.62 (2.11)	0.69	0.57
	PBI-C	741	19.3 (2.3)	0.74	255	19.2 (2.5)	0.80	0.62
	PBI-P	741	3.34 (2.46)	0.66	255	2.81 (2.29)	0.66	0.68

Note: *SD* = standard deviation; SCARED = screening for childhood anxiety related disorders; SMFQ = short mood and feelings questionnaire; JEPQ-E/N = junior eysenck personality questionnaire - Extraversion/Neuroticism; CASI = childhood anxiety sensitivity index; ARI = affective reactivity index; BISBAS-I/A = behavioral inhibition system behavioral activation system - Inhibition/Activation; FSSC-R = fear survey schedule for children revised; TF-20 = threat and fearlessness questionnaire - 20-item version; ICU = inventory of callous-unemotional traits; RBIQ = retrospective behavioral inhibition questionnaire; PBI-A/C/P = Parental Bonding Instrument - Authoritarianism/Coldness/Protectiveness.

individuals) plus one or both parents (514 parents in all). Sixty twin pairs have thus far participated in MRI scans. The zygosity estimation procedure we employed estimated 132 MZ pairs (34.6%), 245 DZ pairs (64.3%), four pairs of unknown zygosity (1.0%), and 35 pairs missing zygosity information. An additional single-item index of parent-rated zygosity was used to assign zygosity to 17 of the 35 same-sex pairs for whom the zygosity questionnaire was not completed; this item showed high concordance ($\kappa = 0.94$, $n = 252$) with the algorithm-assigned zygosity in twin pairs for whom both measures were collected. Table 3 lists basic statistics for the child self-report and parent report survey measures. Within measure and between visit test–retest reliabilities were good to excellent. A wide array of analyses using these and the other assessed measures are underway to address the aims of this study.

Discussion

The breadth and depth of this study promises an innovative and integrated perspective on the genetic and environmental aspects of NVS constructs. This perspective is not available in studies focused solely on specific psychological paradigms or individual disorder symptomatology. This study is unique in its opportunity to characterize the inter-individual and genetic relationships of a broad suite of dimensional measures and laboratory tasks that probe NVS. It can potentially determine to what extent the genetic factors underlying individual differences in responses to these phenotypes index genetic risk for IDs. Furthermore, paradigms measuring peripheral physiology and brain circuits can provide insights into biological mechanisms of risk. This study also examines the role of specific measures of childhood environment with known, potent

effects on IDs (pubertal stage, peer victimization, parenting style, parental psychopathology) on these responses and how they moderate the development of internalizing symptomatology in the context of genetic risk.

As mentioned previously, a primary limitation to this study is the inclusion of Caucasian families only. This precludes the ability to generalize our findings to children of other racial and ethnic groups. A replication study in a more racially diverse sample is needed. The current study was also limited to one assessment in pre-adolescent children. Follow-up with the current cohort in the context of a longitudinal study would estimate the stability versus change of the measures across development, replicability of findings, and predictive value of the endophenotypes for emerging IDs.

Acknowledgments

The authors are very grateful to the twins and parents who have participated in the study over the past several years. JAS was extremely fortunate to have a dedicated team of scientists and research staff who contributed to the success of this study: Andrea Molzhon, Shannon Hahn, Laura Hazlett, Audrey Anderson, Oumaima Kaabi, Elizabeth Long, Diti Sheth, Lisa Ulmer, Lance Rappaport, Jeremy Cornelissen, Aditya Devineni, Lisa Straub, Daniel Deaton, Kevin Kim, Abigail Welch, Alexis Exum, Hiren Kolli, Hannah Mayberry, Tulsi Shah, Casey Guerra, Divya Patel, Ian Muse, Mosa Shahzada, Alana Nichols, Mazin Elmubarak, Pascaline Ezouah, Angela Chung, Lakshmi Ravindra, Hannah Donnelly, Jun Qi, Usama Nasir, Katelyn (Haejung) Shin, Chetna Jhurani, Essam Elrazaq, Jessica Chrisinger, Harper Lorencki, Johnnie Mortensen, Kalina Michalska, Christian Grillon, and Scott Vrana. The authors are also very appreciative of the excellent staff at the Mid-Atlantic Twin Registry, especially Anne Taylor-Morris, Carol Williams, Renolda Gelzinis, Emily Lilley, and Dr Judy Silberg.

Financial Support

This study was supported by the National Institutes of Health (R01MH098055 to JMH and NIMH-IRP-ziamh002781 to DSP) and the Brain and Behavior Research Foundation (BBRF Grant 21984 to JMH). REDCap and the Mid-Atlantic Twin Registry were supported through the NIH Center for Advancing Translational Research Grant Number UL1TR000058.

References

Achenbach, T. M. (1991). *Manual for the child behavior checklist/4-18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. (1995). The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research, 5*, 1–12.

Bailey, J. E., Argyropoulos, S. V., Kendrick, A. H., & Nutt, D. J. (2005). Behavioral and cardiovascular effects of 7.5% CO2 in human volunteers. *Depression and Anxiety, 21*, 18–25.

Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J. A., & Neer, S. M. (1997). The screen for child anxiety related emotional disorders (SCARED): Scale construction and psychometric characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 545–553.

Britton, J. C., Grillon, C., Lissek, S., Norcross, M. A., Szuhany, K. L., Chen, G., … Pine, D. S. (2013). Response to learned threat: An FMRI study in adolescent and adult anxiety. *American Journal of Psychiatry, 170*, 1195–1204.

Britton, J. C., Lissek, S., Grillon, C., Norcross, M. A., & Pine, D. S. (2011). Development of anxiety: The role of threat appraisal and fear learning. *Depression and Anxiety, 28*, 5–17.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*, 319–333.

Dalgleish, T., Taghavi, R., Neshat-Doost, H., Moradi, A., Canterbury, R., & Yule, W. (2003). Patterns of processing bias for emotional information across clinical disorders: A comparison of attention, memory, and prospective cognition in children and adolescents with depression, generalized anxiety, and posttraumatic stress disorder. *Journal of Clinical Child & Adolescent Psychology, 32*, 10–21.

Davis, M. (1986). Pharmacological and anatomical analysis of fear conditioning using the fear-potentiated startle paradigm. *Behavioral Neuroscience, 100*, 814–824.

Deveney, C. M., Connolly, M. E., Haring, C. T., Bones, B. L., Reynolds, R. C., Kim, P., … Leibenluft, E. (2013). Neural mechanisms of frustration in chronically irritable children. *American Journal of Psychiatry, 170*, 1186–1194.

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Manual of the eysenck personality questionnaire*. London: Hodder and Stoughton.

Eysenck, S. B. (1965). A new scale for personality measurements in children. *British Journal of Educational Psychology, 35*, 362–367.

Geller, B., Bolhofner, K., Craney, J. L., Williams, M., DelBello, M. P., & Gundersen, K. (2000). Psychosocial functioning in a prepubertal and early adolescent bipolar disorder phenotype. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*, 1543–1548.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry, 160*, 636–645.

Grillon, C., & Baas, J. (2003). A review of the modulation of the startle reflex by affective states and its application in psychiatry. *Clinical Neurophysiology, 114*, 1557–1579.

Grillon, C., Lissek, S., Rabin, S., McDowell, D., Dvir, S., & Pine, D. S. (2008). Increased anxiety during anticipation of unpredictable but not predictable aversive stimuli as a psychophysiologic marker of panic disorder. *American Journal of Psychiatry, 165*, 898–904.

Guyer, A. E., Choate, V. R., Grimm, K. J., Pine, D. S., & Keenan, K. (2011). Emerging depression is associated with face memory deficits in adolescent girls. *Journal of the American Academy of Child & Adolescent Psychiatry, 50*, 180–190.

Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional responses: Effects of a neocortical network on the limbic system. *Neuroreport, 11*, 43–48.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*, 377–381.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51*, 59–67.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., … Wang, P. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*, 748–751.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., … Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 980–988.

Kendler, K. S. (2001). Twin studies of psychiatric illness: An update. *Archives of General Psychiatry, 58*, 1005–1014.

Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1998). Stressful life events and major depression: Risk period, long-term contextual threat, and diagnostic specificity. *Journal of Nervous and Mental Disease, 186*, 661–669.

Kessler, R. C., Andrews, G., Mroczek, D. K., Ustun, B., & Wittchen, H.-U. (1998). The World health organization composite international diagnostic interview short form (CIDI-SF). *International Journal of Methods in Psychiatric Research, 7*, 171–185.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*, 593–602.

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., … Morris, A. (2008). Assessing callous-unemotional traits in adolescent offenders: Validation of the Inventory of Callous-Unemotional Traits. *International Journal of Law and Psychiatry, 31*, 241–252.

Kramer, M. D., Patrick, C. J., Krueger, R. F., & Gasperi, M. (2012). Delineating physiologic defensive reactivity in the domain of self-report: Phenotypic and etiologic structure of dispositional fear. *Psychological Medicine, 42*, 1305–1320.

Landis, C. W. (1939). *The startle pattern*. Oxford: Farrah & Rinehart.

Lilley, E. C., & Silberg, J. L. (2013). The mid-atlantic twin registry, revisited. *Twin Research and Human Genetics, 16*, 424–428.

Marsh, A. A., Yu, H. H., Pine, D. S., & Blair, R. J. (2010). Oxytocin improves specific recognition of positive facial expressions. *Psychopharmacology (Berl), 209*, 225–232.

Mynard, H., & Joseph, S. (2000). Development of the multidimensional peer-victimization scale. *Aggressive Behavior, 26*, 169–178.

Nichols, R. C., & Bilbro, W. C. Jr. (1966). The diagnosis of twin zygosity. *Acta Genetica et Statistica Medica, 16*, 265–275.

Ollendick, T. H. (1983). Reliability and validity of the revised fear surgery schedule for children (FSSC-R). *Behaviour Research and Therapy, 21*, 685–692.

Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., … Hettema, J. M. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Molecular Psychiatry*. Advance online publication.

Papp, L. A., Klein, D. F., & Gorman, J. M. (1993). Carbon dioxide hypersensitivity, hyperventilation, and panic disorder. *American Journal of Psychiatry, 150*, 1149–1157.

Parker, G., Tupling, H., & Brown, L. (1979). A parental bonding instrument. *British Journal of Medical Psychology, 52*, 1–10.

Peeters, H., Van, G. S., Vlietinck, R., Derom, C., & Derom, R. (1998). Validation of a telephone zygosity questionnaire in twins of known zygosity. *Behavior Genetics, 28*, 159–163.

Philippot, P., & Feldman, R. S. (1990). Age and social competence in preschoolers' decoding of facial expression. *British Journal of Social Psychology, 29*, 43–54.

Pine, D. S., Klein, R. G., Coplan, J. D., Papp, L. A., Hoven, C. W., Martinez, J., … Gorman, J. M. (2000). Differential carbon dioxide sensitivity in childhood anxiety disorders and nonill comparison group. *Archives of General Psychiatry, 57*, 960–967.

Pine, D. S., Klein, R. G., Mannuzza, S., Moulton, J. L. III, Lissek, S., Guardino, M., & Woldehawariat, G. (2005). Face-emotion processing in offspring at risk for panic disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*, 664–672.

Reznick, J. S., Hegeman, I. N., Kaufman, E. R., Woods, S. W., & Jacobs, M. (1992). Retrospective and concurrent self-report of behavioral inhibition and their relation to adult mental health development and psychopathology. *Development and Psychopathology, 4*, 301–321.

Rich, B. A., Grimley, M. E., Schmajuk, M., Blair, K. S., Blair, R. J., & Leibenluft, E. (2008). Face emotion labeling deficits in children with bipolar disorder and severe mood dysregulation. *Development and Psychopathology, 20*, 529–546.

Rich, B. A., Holroyd, T., Carver, F. W., Onelio, L. M., Mendoza, J. K., Cornwell, B. R., … Leibenluft, E. (2010). A preliminary study of the neural mechanisms of frustration in pediatric bipolar disorder using magnetoencephalography. *Depression and Anxiety, 27*, 276–286.

Ripke, S., Wray, N. R., Lewis, C. M., Hamilton, S. P., Weissman, M. M., Breen, G., … Sullivan, P. F. (2013). A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry, 18*, 497–511.

Savage, J. E., Sawyers, C., Roberson-Nay, R., Hettema, J. M. (2016). The genetics of anxiety-related negative valence system traits. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. Advance online publication.

Schmitz, A., Merikangas, K., Swendsen, H., Cui, L., Heaton, L., & Grillon, C. (2011). Measuring anxious responses to predictable and unpredictable threat in children and adolescents. *Journal of Experimental Child Psychology, 110*, 159–170.

Silverman, W. K., Fleisig, W., Rabian, B., & Peterson, R. A. (1991). Childhood anxiety sensitivity index. *Journal of Clinical Child Psychology, 20*, 162–168.

Stringaris, A., Goodman, R., Ferdinando, S., Razdan, V., Muhrer, E., Leibenluft, E., & Brotman, M. A. (2012). The Affective Reactivity Index: A concise irritability scale for clinical and research settings. *Journal of Child Psychology and Psychiatry, 53*, 1109–1117.

Swartz, J. R., Phan, K. L., Angstadt, M., Fitzgerald, K. D., & Monk, C. S. (2014). Dynamic changes in amygdala activation and functional connectivity in children and adolescents with anxiety disorders. *Development and Psychopathology, 26*, 1305–1319.

Visscher, P. M. (2004). Power of the classical twin design revisited. *Twin Research, 7*, 505–512.