Genetic and Environmental Contributions of Negative Valence Systems to Internalizing Pathways

Jennifer L. Cecilione,* Lance M. Rappaport,* Shannon E. Hahn, Audrey E. Anderson, Laura E. Hazlett, Jason R. Burchett, Ashlee A. Moore, Jeanne E. Savage, John M. Hettema, and Roxann Roberson-Nay
Department of Psychiatry, Virginia Commonwealth University, Richmond, Virginia, USA

The genetic and environmental contributions of negative valence systems (NVS) to internalizing pathways study (also referred to as the Adolescent and Young Adult Twin Study) was designed to examine varying constructs of the NVS as they relate to the development of internalizing disorders from a genetically informed perspective. The goal of this study was to evaluate genetic and environmental contributions to potential psychiatric endophenotypes that contribute to internalizing psychopathology by studying adolescent and young adult twins longitudinally over a 2-year period. This report details the sample characteristics, study design, and methodology of this study. The first wave of data collection (i.e., time 1) is complete; the 2-year follow-up (i.e., time 2) is currently underway. A total of 430 twin pairs ($N = 860$ individual twins; 166 monozygotic pairs; 57.2% female) and 422 parents or legal guardians participated at time 1. Twin participants completed self-report surveys and participated in experimental paradigms to assess processes within the NVS. Additionally, parents completed surveys to report on themselves and their twin children. Findings from this study will help clarify the genetic and environmental influences of the NVS and their association with internalizing risk. The goal of this line of research is to develop methods for early internalizing disorder risk detection.

■ **Keywords:** negative valence systems, internalizing disorders, twin study, endophenotypes, adolescence, genetics

The National Institute of Mental Health (NIMH) launched the Research Domain Criteria (RDoC) to advance new approaches to psychopathology in which primary, core dimensions linking mental health disorders come to serve as the basis for categorizing individuals for research purposes. This effort is in response to the disappointing findings of research studies using psychiatric phenotypes (e.g., schizophrenia and major depression) to elucidate etiological mechanisms. In its current conception, the RDoC matrix encompasses a number of domains, including the negative valence systems (NVS), which focuses on biological and psychological systems involved in the response to loss or aversive, threatening, or harmful stimuli. Within each domain, potential constructs are studied across multiple levels of analysis (e.g., neural circuits, behavior, and genes; Insel et al., 2010). The NVS currently includes the following constructs: acute threat (fear), potential threat (anxiety), sustained threat, loss, and frustrative non-reward (NIMH, 2017). The present study was sponsored through the RDoC initiative and primarily seeks to provide an estimate of genetic and environmental contributions to NVS constructs spanning multiple units of analysis and to determine shared genetic and environmental influences among NVS constructs and between them and internalizing disorders/symptoms. Evidence of shared genetic and environmental contributions to NVS constructs and psychiatric syndromes may implicate these constructs in the etiology of internalizing disorders.

The notion of examining intermediate processes derives from other areas of science where intermediate phenotypes, called *endophenotypes*, may clarify the etiology of otherwise complex syndromes (Gottesman & Gould, 2003). Gottesman and Gould (2003) outlined four criteria to identify an endophenotype, such that endophenotypes must: (1) be correlated with a medical or psychological condition in the population, (2) be heritable, (3) be

state-independent (i.e., independent of active psychiatric episode), and (4) co-segregate with the condition within a family. Cannon and Keller (2006) proposed additional qualifying criteria, such that: (1) endophenotypes must be related to causes instead of effects of disorders, (2) several endophenotypes should influence one complex disorder, (3) endophenotypes should be measured as continuous variables and at different levels of analysis, and (4) endophenotypes should be found in genetically related disorders. In summary, understanding which endophenotypes are associated with which psychiatric syndromes may provide a more useful model for the etiology of extant clinical diagnoses and help identify better prevention and intervention targets.

For the purposes of biometrical twin modeling, an unselected sample was recruited to represent an epidemiological distribution, which is needed to examine NVS constructs that are theorized to manifest in typically developing individuals. A general population sample (as opposed to a selected sample) is necessary for genetic and environmental estimates to be unbiased and maximally informative for the etiology of NVS phenotypes. Given the high prevalence rate of internalizing symptoms and conditions by late adolescence and young adulthood (Kessler et al., 2005; Kessler et al., 2012; Merikangas et al., 2010), a representative adolescent and young adult sample was expected to produce adequate numbers of individuals expressing internalizing disorder symptomatology and those at risk for developing internalizing disorders.

The behavioral genetic methods of this study allowed for an examination of genetic and environmental influences on each potential endophenotype, as well as the association of NVS-related constructs with existing clinical disorders/symptoms and risk factors. Similarly, the current study was positioned to examine the underlying structure of NVS constructs based on the correlation of observed phenotypes and the correlation between genetic and environmental influences. Thus, the current study can demonstrate heritability, examine the co-segregation of potential endophenotypes with psychiatric syndromes in families, and clarify the source of familial co-segregation. Additionally, the longitudinal design of this study will permit the examination of change in endophenotype manifestation and clinical diagnoses over a portion of adolescence and young adulthood (Khoo et al., 2006), which may shed light on an endophenotype's ability to predict risk for and expression of internalizing disorders and internalizing symptoms.

The specific aims of this study included: (1) to examine the factorial architecture and underlying latent structure (at the phenotypic level) of the suite of NVS probes and dimensional measures collected at an initial assessment (i.e., time 1; T1), (2) to determine the contributions of genetic and environmental factors to the covariance among endophenotypic measures at T1 and between endophenotypic measures and current internalizing disorder symptom expres-

TABLE 1

Sample Characteristics (430 Families; 860 Twin Adolescents/Young Adults)

	N (%)
Zygosity	
Dizygotic	264 (61.4)
Opposite-sex	122 (28.4 of full sample)
Monozygotic	166 (38.6)
Race	
Caucasian	399 (92.8)
African American	31 (7.2)
Ethnicity	
Latino/Latina	17 (3.9)
Sex	
Male	368 (42.8)
Female	492 (57.2)
DSM-5 diagnoses and phenomena (twin)	
Panic attacks	102 (11.9)
Panic disorder	10 (1.2)
Social anxiety disorder	116 (13.5)
Specific phobia	102 (11.9)
Generalized anxiety disorder	25 (2.9)
Major depressive disorder	131 (15.3)
DSM-5 diagnoses and phenomena (parent)	
Panic attacks	59 (14.0)
Panic disorder	9 (2.1)
Social anxiety disorder	18 (4.3)
Specific phobia	19 (4.5)
Generalized anxiety disorder	46 (10.9)
Major depressive disorder	129 (30.6)

sion at T1, (3) to determine the contribution of genetic and environmental factors to the covariance between endophenotypic measures at T1 and internalizing disorder symptom expression at a follow-up assessment (i.e., time 2; T2), and (4) to determine the stability of NVS endophenotypes from T1 to T2.

Material and Methods

Participants

Participants for the current study were pairs of Caucasian (including Hispanic) and African-American twins (monozygotic, dizygotic same-sex, and dizygotic opposite-sex) between the ages of 15 and 20 ($M = 17.25$ years, $SD = 1.3$) residing in the mid-Atlantic region of the United States (see Table 1). The majority of the current sample (97.7%) were recruited from the Mid-Atlantic Twin Registry (MATR), a database of twins and families who have previously expressed interest in participating in research studies (Lilley & Silberg, 2013). MATR staff informed twins of the appropriate age of the current study and provided contact information of those who expressed interest and met inclusion criteria. The study coordinator then contacted participants to describe the study in more detail and schedule an appointment. A subset of the current sample (2.3%) was recruited directly by study staff through advertisements (e.g., radio and newspaper) in the Richmond, Virginia area. Twins or parents of twins who were interested in the advertisements contacted the study coordinator

directly. The study coordinator screened twins and scheduled an appointment if twins were eligible.

The sample size of T1 ($N = 430$ twin pairs) achieved over 80% power to detect statistical significance of additive genetic effects that explain at least 40% of the variance for the proposed quantitative (endo)phenotypes (Verhulst, 2016).

Inclusion criteria included: (a) being an identical or fraternal twin, (b) being between the ages of 15 and 20, and (c) living primarily at home. Exclusion criteria were: (a) current use of psychotropic medications (e.g., antidepressants) or medications with psychotropic effects (e.g., beta-adrenergic blockers), (b) diagnosis of an autism spectrum disorder, (c) diagnosis of an intellectual disability, (d) diagnosis of a spatial learning disorder, or prior testing indicating an IQ below 70, (e) seizure without a clear and resolved etiology, (f) current or past episodes of psychosis, (g) serious, not stabilized illness (e.g., liver, kidney, gastrointestinal, respiratory, cardiovascular, endocrinologic, neurologic, immunologic, or blood disease), (h) inadequate production of human growth hormone, (i) sensory integration disorder, (j) congenital adrenal hyperplasia, (k) adrenal inefficiency, (l) deaf with bicochlear implants, (m) cancer (current or past diagnosis), and (n) pregnancy (current or lifetime). A family was excluded if either twin met any exclusionary criteria. Exclusionary criteria were established to reduce interference with physiological data (e.g., psychotropic medications) or participants' ability to complete laboratory tasks.

Measures

Zygosity. Zygosity status (i.e., monozygotic or dizygotic) for each twin pair was assessed based on parent reports about physical similarities between twins. When twin participants were over 18 and attended without a parent or legal guardian, twin participants completed the zygosity assessment. Prior research has demonstrated high validity for this zygosity assessment as compared to blood (Kasriel & Eaves, 1976) and DNA evaluations of zygosity (Jackson et al., 2001). In the present study, zygosity estimated from parent- and twin-reported physical similarity showed high concordance with zygosity estimated from assay of single nucleotide polymorphisms (SNPs) on a subset of the sample ($N = 82$ twin pairs, $\kappa = 0.95$, 95% CI [0.88, 1.0]).

DNA samples. Genetic material (i.e., blood) was collected at T1 and will be collected at T2 to leverage genetic candidates identified by large-sample gene discovery efforts (e.g., Otowa et al., 2016; Power et al., 2017). Specifically, future research will follow a deep phenotyping approach (e.g., Stepniak et al., 2014) to examine whether putative genetic correlates of internalizing psychopathology are associated with psychobiological processes assessed here (e.g., Carney et al., 2016, Goodbourn et al., 2014).

Saliva samples. Markers of neuroendocrine function were assayed from saliva collected both at a resting baseline and before and after the Trier Social Stress Test and the 7.5% carbon dioxide (CO_2) challenge (see 'Laboratory tasks' section below). Salivary cortisol (CORT) indexed neuroendocrine function of the hypothalamic-pituitary-adrenal axis (HPA-axis), while salivary alpha amylase (sAA) indexed function of the sympathetic-adrenal-medullary system (SAM).

Laboratory tasks. Several laboratory tasks, described below, were used to probe NVS-related constructs of interest. See Tables 2–4 for more detail.

Trier Social Stress Test (TSST). The TSST is a widely used task used to study stress responsivity following a laboratory-induced stressor (Kirschbaum et al., 1993). Participants were told that they would give a short speech and perform mental arithmetic in front of an audience of three people while being video-taped for a psychologist to later assess. The TSST involves four phases: a 5-min anticipatory stress phase, a 5-min speech phase, a 5-min serial subtraction phase, and a 5-min recovery phase. After 5 min of solitary preparation by the participant, three confederates unknown to the participant entered the room. The confederates maintained a neutral expression throughout the task and provided no feedback (verbal or non-verbal). After the 5-min speech portion, participants were instructed to subtract 13 successively from 1,022. When participants made an error, they were asked to restart at 1,022. After the arithmetic portion ended, the confederates exited the room. The participant remained in the room for 5 min of recovery. Participant stress response during and following the task was assessed by neuroendocrine response (i.e., SAM and HPA axis activity), participant-reported subjective distress (measured by the Subjective Units of Distress Scale [SUDS]; Wolpe, 1969), and psychophysiological response (i.e., respiratory rate [RR], heart rate [HR], heart rate variability [HRV], and skin conductance [SC]). To our knowledge, there has not been any published evidence of test–retest reliability for the TSST.

Fear generalization (FG). Previous studies have demonstrated that fear conditioning and generalization of conditioned fear play a role in the formation and maintenance of anxiety disorders (Lissek et al., 2005; 2008; Mineka & Zinbarg, 2006). Hence, the present study used a fear generalization (FG) paradigm, a classical conditioning task with a generalization phase instead of an extinction phase, to examine fear responding. The FG paradigm consisted of three phases – habituation, acquisition, and generalization, occurring in the same order for all participants. The aversive/unconditioned stimulus (US; i.e., a stimulus that elicits an automatic, unlearned response) in this task was a 100-ms electric shock, administered to the left wrist. It was paired with neutral stimuli (i.e., stimuli that do not naturally elicit a target response) until it becomes the conditioned stimulus (CS) and can elicit the target response without the presence of the US. A computer monitor presented the visual

TABLE 2

NVS Threat Constructs, Phenotypes, Units of Analysis, and Self-Report Measures

		Units of analysis		
Constructs	Phenotypes	Participant report	Paradigm	Physiological
Acute threat	Phobias	Composite international diagnostic interview[a,b]		
	Fear generalization		Fear generalization[a,b]	EMG, SC[a,b]
	Interoceptive hypersensitivity		CO_2 challenge[a,b]	HRV, SC, RR, TV, HR[a,b]
Potential threat	Anxiety	Anxiety sensitivity index[a,b]	Fear generalization[a,b]	EMG, SC[a,b]
		Depression, anxiety, stress scale-21[a,b]	Trier social stress test[a,b]	HRV, SC, RR, HR[a,b]
			CO_2 challenge[a,b]	HRV, SC, RR, HR, TV[a,b]
	Anxiety disorders*	Composite international diagnostic interview[a,b]		
	Anticipatory anxiety		Trier social stress test[a,b]	HRV, SC, RR, HR[a,b]
			Fear generalization[a,b]	EMG, SC[a,b]
			CO_2 challenge[a,b]	HRV, SC, RR, HR, TV[a,b]
Sustained threat	Emotional recognition		Facial expression labeling task[a]	
	Fear extinction		Fear generalization[a,b]	EMG, SC[a,b]

Note: [a]Included in wave 1; [b]Included in wave 2. HRV = heart rate variability; SC = skin conductance; RR = respiration rate; EMG = eye-blink electromyographic response; TV = tidal respiratory volume; HR = heart rate.
*Not formally included in RDoC matrix.

TABLE 3

NVS Loss/Frustrative Constructs, Phenotypes, Units of Analysis, and Measures

		Units of analysis	
Constructs	Phenotypes	Participant report	Paradigm
Loss	Depression*	Short mood and feelings questionnaire[a,b]	
		Depression, anxiety, stress scale – 21[a,b]	
		Composite international diagnostic interview[a,b]	
Frustrative non-reward	Irritability	Affective reactivity index[a,b]	
	Distress tolerance		Mirror tracing task[a]

Note: [a]Included in wave 1; [b]Included in wave 2.
*Not formally included in RDoC matrix.

TABLE 4

NVS Environmental Risk Factors, Personality Constructs, Alcohol and Substance Use, Phenotypes, Units of Analysis, and Measures

		Units of analysis		
Constructs	Phenotypes	Participant report	Paradigm	Physiological
Risk/protective factors	Peer victimization	Multidimensional peer victimization scale[a,b]		
		Index of peer relations[a,b]		
		Experiences with close relationships-revised[a,b]		
	Parent psychopathology	CIDI-SF[a]		
	Life events/trauma	Stress and adversity inventory[a,b]		
	Parenting	Parental bonding instrument[a]		
Alcohol use	Severity of alcohol use	Alcohol use disorder identification test[a,b]		
	Expectations of alcohol use	Drinking expectancy profile[a,b]		
Smoking	Severity of nicotine use	Fagerstrom test of nicotine dependence[a,b]		
Temperament/personality*	Neuroticism, extraversion	Eysenck personality questionnaire short form[a,b]		
	Anxiety sensitivity	Anxiety sensitivity index[a,b]	CO_2 challenge[a,b]	HRV, SC, RR, TV, HR[a,b]
	Behavioral inhibition/activation	Behavioral inhibition system/behavioral activation system scales[a,b]		
	Unemotionality	Inventory of callous-unemotional traits[a,b]		

Note: [a]Included in wave 1; [b]Included in wave 2. HRV = heart rate variability; SC = skin conductance; RR = respiration rate; TV = tidal respiratory volume; CIDI-SF= composite international diagnostic interview-short form.
*Not formally included in RDoC matrix.

conditioned stimuli with the CS+/− randomized to be either a large or small circle. During habituation, participants wore a headset and repeatedly heard a sudden white noise (i.e., startle probe). For acquisition, the CS+ was paired with a shock and the CS− was not. During the acquisition phase, 75% of the CS+ were paired with the shock. The objective of acquisition was for the participant to learn to associate either the little circle or big circle with the shock. For generalization, eight circles (generalization stimuli; GS) of intermediate size rings representing a 'continuum-of-size' between the CS+ and CS− were added. During the generalization phase, participants were randomly presented with the CS+/− and six trials from each of the eight GS sizes; 50% of the CS+ presentations were paired with a shock to prevent extinction (Lissek et al., 2008). Startle probes accompanied 50% of acquisition and generalization trials (even), and in the other 50% of trials (odd), participants rated perceived likelihood of shock (i.e., level of risk). The primary psychophysiological outcome measure was fear-potentiated startle magnitude (FPS; relative increase in the amplitude of the acoustic startle response measured by electromyography [EMG]). Additional responses to the task were assessed via subjective distress (Wolpe, 1969) and SC. To our knowledge, there has not been any published evidence of test–retest reliability for the FG task.

7.5% carbon dioxide (CO₂) challenge. Variations of the CO_2 challenge have been used to assess risk factors for panic disorder (e.g., Papp et al., 1993; Pine et al., 2005). Breathing air with elevated levels of CO_2 is known to produce distress as assessed physiologically (Papp et al., 1993) and psychologically (Roberson-Nay et al., 2010) among adults (Roberson-Nay et al., 2013) and children (Pine et al., 2005). Participants wore a face mask for the 20-min duration of this task and were informed that they would breathe both ambient room air as well as air with an elevated level of CO_2. The facemask was connected to a stopcock valve that allowed the experimenter to manually switch from ambient room air to the CO_2 air mixture. Participants breathed ambient room air for 5 min (pre-CO_2), followed by 8 min of air enriched to 7.5% CO_2, and then a 5-min recovery period (ambient room air). Participants were not told when the CO_2 air mixture was turned on or off, thus remaining unaware of the onset and discontinuation of the CO_2 air mixture. Anxious arousal during the task was assessed as neuroendocrine response (i.e., SAM and HPA activity), psychophysiological response (i.e., RR, tidal respiratory volume [TV], HR, HRV, and SC), and behavioral response (i.e., whether and when participants prematurely ended the task). Psychological response was assessed as subjective distress (Wolpe, 1969) and panic symptom severity, as measured by the Diagnostic Symptoms Questionnaire (Sanderson et al., 1989). Test–retest reliability of this CO_2 challenge task has been previously documented in a young adult sample (Roberson-Nay et al., 2017).

Mirror tracing task. Distress tolerance has been frequently associated with internalizing psychopathology risk (Leyro et al., 2010). In the present study, distress tolerance was assessed using the mirror tracing task (Strong et al., 2003). Participants were told to trace the outline of three shapes on the computer screen. They were also told that they could earn more money if they completed the task. To create psychological distress, the actions of the computer mouse and cursor were reversed. When the participant moved the mouse too slowly or moved off the line, a buzz sounded to signal them to restart at the beginning of the shape. Participants first practiced the mechanics of the task on two simple lines. During the trial stage, participants traced the outline of a star. The software was programmed so that the buzzer would sound randomly to prevent participants from completing the star shape. Distress tolerance was assessed behaviorally as the amount of time that a participant persisted on the trial stage. Additionally, subjective distress was assessed by participant self-report to ensure that persistence on the task was not conflated with distress due to the task. To our knowledge, there has not been any published evidence of test–retest reliability for the mirror tracing task.

Facial expression labeling task (FELT). Deficits in facial emotion recognition have been associated with a variety of internalizing psychiatric syndromes (e.g., Trentacosta & Fine, 2010). During the FELT, participants were asked to choose the emotion of various faces shown on the screen. All faces were Caucasian adults who expressed one of Ekman's six basic emotions (i.e., happiness, sadness, surprise, fear, disgust, or anger; Ekman & Friesen, 1976). Images were presented to participants at varying degrees of emotional expressivity from 10% to 100% expressivity of the target emotion, such that participants saw a total of 360 images (6 trials × 6 emotions × 10 expressivity levels). Images were created by morphing a picture of the actor making a neutral face with one of the actor expressing the target emotion at 100% intensity. After each face was displayed, participants were asked to choose the emotion displayed from a list of six emotions. Trials were presented in a single randomized order (Blair et al., 2001; Marsh et al., 2010) and have shown test–retest reliability among adult (Adams et al., 2016) and child samples (Cecilione et al., 2017).

Dimensional participant-report measures. Participant-report measures were chosen for inclusion in the current study to assess dimensional severity of internalizing psychopathology, related psychological constructs and behaviors (e.g., irritability, alcohol use), and environmental influences (see Tables 2–4). All questionnaires were administered via Research Electronic Data Capture (REDCap; Harris et al., 2009), except for the Stress and Adversity Inventory (STRAIN; Slavich & Epel, 2010), which was administered via its own internet-based interface.

Severity of internalizing and related syndromes was assessed based on participant-reported symptoms of depression (Short Mood and Feelings Questionnaire [SMFQ]; Angold et al., 1995; and Depression, Anxiety, Stress Scale–21 [DASS-21]; Lovibond & Lovibond, 1995), anxiety, and stress (DASS-21; Lovibond & Lovibond, 1995). Test–retest reliability and validity have been previously documented for the SMFQ (Cheng et al., 2009; Turner et al., 2014) and the DASS-21 (Mahmoud et al., 2010; Silva et al., 2016) in adolescent or young adult samples. Additionally, participants completed an expanded version of the Composite International Diagnostic Interview-Short Form (CIDI-SF; Kessler et al., 1998; Wittchen, 1994; see Appendix A in Rappaport et al., 2017). Validity and test–retest reliability of the CIDI have been established in adults (Wittchen, 1994); to our knowledge, there has not been any published evidence of test–retest reliability for the CIDI-SF in an adolescent and young adult sample. However, there is evidence of a strong relationship between diagnoses garnered from the CIDI-SF and the full-scale CIDI (Kessler et al., 1998). The present study also assessed related psychological correlates of internalizing psychopathology, including anxiety sensitivity (Anxiety Sensitivity Index [ASI]; Reiss et al., 1986), irritability (Affective Reactivity Index [ARI]; Stringaris et al., 2012), extraversion and neuroticism (modified version of the Eysenck Personality Questionnaire - Short Form [EPQ-SF]; Eysenck et al., 1985; Khan et al., 2005), psychopathy (Inventory of Callous-Unemotional Traits [ICU]; Kimonis et al., 2008), and behavioral inhibition (BIS) and behavioral activation systems (BAS) (BIS/BAS scales; Muris et al., 2005). Test–retest reliability and validity have been previously documented for the ASI (Peterson & Heilbronner, 1987; Reiss et al., 1986; Vujanovic et al., 2007), ARI (Mulraney et al., 2014; Stringaris et al., 2012), EPQ-SF (Hosokawa & Ohyama, 1993; Sato, 2005), ICU (Feilhauer et al., 2012; Kimonis et al., 2008; Kimonis et al., 2013), and the BIS/BAS scales (Carver & White, 1994) in adolescent or young adult samples.

Among potentially important environmental influences, social relationships play a critical role in the development and maintenance of internalizing disorders (Rubin et al., 2009; Semrud-Clikeman, 2007), particularly major depression (Hammen, 2005; Kendler & Gardner, 2001; Weissman et al., 1971). Social relationships were assessed as participant attachment to romantic partners among those who reported a prior or current romantic relationship (Experiences with Close Relationships-Revised [ECR-R]; Fraley et al., 2011), peer relationships (Index of Peer Relations [IPR]; Hudson, 1982), peer victimization (Multidimensional Peer Victimization Scale [MPVS]; Mynard & Joseph, 2000), and parental relationships (Parental Bonding Instrument [PBI]; Parker et al., 1979). Test–retest reliability and validity have been previously demonstrated for the ECR-R (Fairchild & Finney, 2006; Sibley et al., 2005), MPVS (Eastman et al., 2017), and PBI (Parker et al., 1979;

Parker, 1989) within adolescent or young adult samples. Validity, but not test–retest reliability, has been established for the IPR in an adolescent sample (Forte & Green, 1994). Prior research also demonstrates the importance of stressful and/or traumatic life events to internalizing disorder risk (Kendler et al., 1999; Miller et al., 2003; Nugent et al., 2011). Stressful life experiences were assessed by participant report on the STRAIN (Slavich & Epel, 2010); test–retest reliability and validity of the STRAIN have been previously demonstrated among adults but not in an adolescent sample (Slavich & Shields, 2017).

Substance use has also been correlated with internalizing disorders and their development, specifically during adolescence and young adulthood (Clark & Neighbors, 1996; Kendler et al., 2003). Participants who reported prior consumption of alcohol completed the Alcohol Use Disorders Identification Test (AUDIT; Saunders et al., 1993) and the Drinking Expectancy Profile (DEP; Young & Oei, 1996) to assess frequency, severity, and expectations regarding alcohol consumption respectively. Similarly, severity of nicotine dependence was assessed for current smokers via the Fagerstrom Test of Nicotine Dependence (FTND; Fagerstrom & Schneider, 1989). Test–retest reliability and validity of the AUDIT (de Meneses-Gaya et al., 2009; Selin, 2003) and the FTND (Haddock et al., 1999; Pomerleau et al., 1994) has been previously documented among adults but not yet in an adolescent sample. To our knowledge, there is evidence of validity (Lee et al., 2003), but presently no data on test–retest reliability for the DEP.

Since parental psychopathology and parenting styles have been shown to have an impact on children's development (e.g., Berg-Nielsen et al., 2002; Walker et al., 1989), parents reported on their lifetime experiences with psychopathology (expanded CIDI-SF), neuroticism and extraversion (modified EPQ-SF), post-traumatic stress disorder symptom severity (Post-Traumatic Stress Disorder Checklist for the DSM-5 [PCL-5]; Weathers et al., 2013), alcohol use severity (AUDIT), expectations regarding alcohol use (DEP), and severity of nicotine dependence (FTND). Test–retest reliability and validity of the PCL-5 have been previously demonstrated in young adult (Blevins et al., 2015) and veteran (Bovin et al., 2016) samples; to our knowledge, there has not been any published evidence of test–retest reliability of the PCL-5 in an adult community sample. These parent-report measures were collected to assess the role of parental psychopathology and/or substance use in the development of internalizing disorders in adolescents and young adults. Parents also reported on externalizing traits of their twin children (ICU; Kimonis et al., 2008).

Procedure

The study staff explained the study in detail to the participants and their parent(s)/legal guardian(s) and answered all questions prior to consent and assent forms being signed.

TABLE 5

Statistics for Sum Scores of Twin-Report Measures and Survey Subscales

Measure	Mean (SD)	Internal consistency (Cronbach's α)
Mood and emotion		
Short mood and feelings questionnaire	5.16 (4.44)	0.92
DASS: Depression	2.67 (3.09)	0.92
DASS: Anxiety	3.01 (2.91)	0.82
DASS: Stress	4.98 (3.42)	0.84
Anxiety sensitivity index	17.70 (9.52)	0.89
Affective reactivity index	2.65 (2.56)	0.92
Personality		
EPQ-SF: Neuroticism	4.74 (3.21)	0.90
EPQ-SF: Extraversion	5.17 (2.35)	0.90
ICU	21.07 (8.90)	0.89
BIS	11.50 (4.14)	0.83
BAS (reward responsiveness)	11.21 (2.59)	0.82
BAS (drive)	5.15 (2.64)	0.82
BAS (fun-seeking)	5.91 (2.49)	0.67
Relationships		
Index of peer relations	22.54 (14.55)	0.96
Parental bonding instrument (mom): Overprotection	13.15 (6.70)	0.87
Parental bonding instrument (mom): Care	29.21 (6.12)	0.93
Parental bonding instrument (dad): Overprotection	11.03 (6.84)	0.89
Parental bonding instrument (dad): Care	25.53 (7.95)	0.94
ECR-R: Anxiety	3.26 (1.24)	0.94
ECR-R: Avoidance	2.84 (1.08)	0.95
Multidimensional peer victimization scale: Total	5.53 (6.03)	0.95

Note: DASS = depression, anxiety, stress scale – 21 Items, EPQ-SF = eysenck personality questionnaire - short form, ECR-R = experiences with close relationships - revised, ICU = inventory of callous-unemotional traits, BIS = behavioral inhibition system, BAS = behavioral activation system.

Participants age 18 and above provided written consent. Participants under age 18 provided written assent after their parent or legal guardian provided written consent. Separate consent and assent were obtained for participation in the blood draw, which was treated as a separate, optional component of the study. The current study was reviewed and approved by the VCU Institutional Review Board.

The current study is longitudinal in design; T1 and T2 are separated by a period of approximately two years. At T1, all twin participants participated in person in a laboratory at VCU. Twin participants were separated from each other at the beginning of the appointment to complete all laboratory tasks (see 'Laboratory tasks' section). Before the twins started the laboratory tasks, their blood was drawn by a study staff member trained in phlebotomy. To reduce the time burden on twins, each completed the tasks in one of two orders. Self-report assessments (see 'Dimensional participant-report measures' section) were administered in between the laboratory tasks. Meanwhile, the accompanying parent(s) or legal guardian(s) completed questionnaires about themselves and their twin children in a separate room. Visits at T1 visits took approximately four and half to five hours to complete.

All twin participants will be invited to participate in some aspect of the T2 follow-up assessment to examine change in NVS-related constructs and internalizing disorder symptomology from T1. All participants will be asked to complete participant-report assessments at T2, and a subset of 150 twin pairs will be invited back into the lab-

oratory for T2 to complete select experimental paradigms from T1 in addition to completing questionnaire measures (see Tables 2–4). The same consent and blood procedures used in T1 will also be applied to T2 laboratory visits. Like T1, twin participants will be separated from each other to complete laboratory tasks in one of two orders. Visits at T2 will take approximately four hours to complete. Twin participants not selected to return to the laboratory for T2 will complete participant-report measures from their home using electronic data capture methods. T2 study measures will be highly consistent with the T1 assessment measures (see Tables 2–4).

Results

The full sample for T1 included 430 twin pairs ($N = 860$ individuals) and 422 parents or legal guardians. Table 1 lists the demographic information for this study's sample. Psychometric and descriptive data for dimensional participant-report assessments are provided in Tables 5–7.

Discussion

The current study aimed to identify endophenotypes that may be both risk factors for internalizing psychopathology development and intermediate mechanisms in understanding the developmental pathway from genetic loci to psychiatric syndromes. The study focused on endophenotypes within the NVS; the dimensions of the NVS describe behavioral, emotional, neurological, genetic, and cognitive

TABLE 6

Statistics for Twin-Report Substance Use Measures

Measure	Mean (SD)	Internal consistency (Cronbach's α)
Nicotine use (for those who reported smoking cigarettes)		
Fagerstrom test for nicotine dependence	2.25 (2.43)	0.59
Alcohol use (for those who reported drinking alcohol)		
DEP: Assertion	35.27 (6.73)	0.86
DEP: Affective change	21.09 (6.60)	0.87
DEP: Dependence	14.62 (4.10)	0.78
DEP: Sexual enhancement	16.69 (3.94)	0.76
DEP: Cognitive change	7.82 (2.60)	0.75
DEP: Tension reduction	10.26 (3.21)	0.63
AUDIT	6.18 (4.58)	0.79

Note: DEP = drinking expectancy profile, AUDIT = alcohol use disorders identification test.

TABLE 7

Statistics for Sum Scores of Parent-Report Measures and Survey Subscales

Measure	Mean (SD)	Internal consistency (Cronbach's α)
Personality		
EPQ-SF: Neuroticism (about self)	3.18 (3.03)	0.92
EPQ-SF: Extraversion (about self)	4.41 (2.57)	0.93
ICU (about twin)	20.03 (9.33)	0.91
Nicotine use (about use)		
Fagerstrom test for nicotine dependence	3.45 (2.66)	0.60
Alcohol use (about self)		
DEP: Assertion	28.38 (7.84)	0.88
DEP: Affective change	20.26 (6.05)	0.89
DEP: Dependence	13.50 (4.31)	0.81
DEP: Sexual enhancement	16.30 (2.89)	0.52
DEP: Cognitive change	7.25 (2.34)	0.77
DEP: Tension reduction	10.44 (3.29)	0.67
AUDIT	3.12 (2.51)	0.82
Other		
PCL-5: this month (about self)	6.46 (9.21)	0.95
PCL-5: lifetime (about self)	8.92 (11.22)	0.96

Note: DASS = depression, anxiety, stress scale - 21 Items, AUDIT = alcohol use disorders identification test, DEP = drinking expectancy profile EPQ-SF = eysenck personality questionnaire- short form, ECR-R = experiences with close relationships – revised, PCL-5 = post-traumatic stress disorder checklist for DSM-5, ICU = inventory of callous-unemotional traits.

patterns that traverse internalizing disorders (Insel et al., 2010). Therefore, the constructs of the NVS can be used to understand the development of multiple internalizing disorders and the comorbidity among them. It is theorized that NVS processes may explain genetic pleiotropy and common etiology of multiple internalizing disorders.

In keeping with the strengths of multitrait, multimethod research (Maas et al., 2009), the present study employed a range of methods from participant-report to the assessment of psychophysiological and hormonal responses. Additionally, the current study's sample and longitudinal design allowed for some examination of internalizing disorder risk and development over a critical developmental period during which many internalizing disorders manifest (Kessler et al., 2005; Merikangas et al., 2010). Of specific relevance to the psychiatric representativeness of the current twin sample, the estimated prevalence rates for panic attacks, panic disorders, generalized anxiety disorder, and major depressive disorder were similar to population estimates, while

prevalence rates of specific phobias and social phobia were slightly higher (see Table 1) (Hayward et al., 1989; Merikangas et al., 2010).

As outlined by the RDoC initiative, elucidating the genetic and environmental influences in the development of NVS phenotypes may help clarify the development of internalizing disorders (Insel et al., 2010). Ultimately, it is hoped that this approach will provide additional targets for biological and psychological interventions, which may address processes that occur prior to psychiatric illness onset. These targets may be particularly important in the identification of individuals at risk for later psychopathology, which would guide preventive intervention.

There were several limitations associated with the current study. One limitation is the small sample sizes of African-American and Hispanic twins, which precludes the examination of race/ethnic-specific differences in NVS phenotypes and internalizing disorder risk. A future study with larger samples of other races is necessary to examine

race-related differences in endophenotype heritability or internalizing disorder development. Additionally, twin participants each completed laboratory tasks in one of two orders. To account for potential task order effects, task order will be entered into relevant analyses as a covariate. Also, visits were conducted at different times of day. It is possible that participants who participated earlier in the day experienced a different level of fatigue as compared to those who participated in evening visits. The variation in time of day may also complicate analyses of neuroendocrine outcomes (i.e., cortisol). Finally, twins on antidepressants or anti-anxiety medications were excluded from study participation at T1. It may be possible that some variance of internalizing disorder symptom expression may have been lost because of this exclusionary criterion.

The goals of the current study were to (1) examine the relationship between NVS probes (i.e., laboratory tasks) and the participant-report measures collected at T1, (2) clarify the genetic and environmental contributions to endophenotypes and their relationship to internalizing disorders at T1, (3) study the development of internalizing disorders from T1 to T2 by considering the genetic and environmental factors that contribute to change in endophenotypes and internalizing disorder expression over time, and (4) to establish the temporal stability of NVS endophenotypes from T1 to T2. The host of multidimensional measures in conjunction with many units of analysis (e.g., genetic samples, psychophysiological measures, participant-reports, and laboratory paradigms), and the study's longitudinal design will allow for the thorough examination of each of the study's primary aims.

Acknowledgments

Financial Support

Conflicts of Interest

None.

Ethical Standards

References

Adams, T., Pounder, Z., Preston, S., Hanson, A., Gallagher, P., Harmer, C. J., & McAllister-Williams, R. H. (2016). Test–retest reliability and task order effects of emotional cognitive tests in healthy subjects. *Cognition and Emotion*, *30*, 1247–1259.

Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, *5*, 237–249.

Berg-Nielsen, T. S., Vikan, A., & Dahl, A. A. (2002). Parenting related to child and parental psychopathology: A descriptive review of the literature. *Clinical Child Psychology and Psychiatry*, *7*, 529–552.

Blair, R. J. R., Colledge, E., Murray, L., & Mitchell, D. G. V. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. *Journal of Abnormal Child Psychology*, *29*, 491–498.

Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., & Domino, J. L. (2015). The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*, *28*, 489–498.

Bovin, M. J., Marx, B. P., Weathers, F. W., Gallagher, M. W., Rodriguez, P., Schnurr, P. P., & Keane, T. M. (2016). Psychometric properties of the PTSD checklist for diagnostic and statistical manual of mental disorders – fifth edition (PCL-5) in veterans. *Psychological Assessment*, *28*, 1379–1391.

Cannon, T. D., & Keller, M. C. (2006). Endophenotypes in the genetic analyses of mental disorders. *Annual Review of Clinical Psychology*, *2*, 267–290.

Carney, D. M., Moroney, E., Machlin, L., Hahn, S., Savage, J. E., Lee, M., … Roberson-Nay, R. (2016). The twin study of negative valence emotional constructs. *Twin Research and Human Genetics*, *19*, 456–464.

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*(2), 319–333.

Cecilione, J. L., Rappaport, L. M., Verhulst, B., Carney, D. M., Blair, R. J. R., Brotman, M. A., … Hettema, J. M. (2017). Test–retest reliability of the facial expression labeling task. *Psychological Assessment*, *29*(12), 1537–1542.

Cheng, P.-X., Cao, F.-L., & Su, L.-Y. (2009). Reliability and validity of the short mood and feelings questionnaire in Chinese adolescents. *Chinese Mental Health Journal, 23*, 60–62, 72.

Clark, D. B., & Neighbors, B. (1996). Adolescent substance abuse and internalizing disorders. *Child and Adolescent Psychiatric Clinics of North America, 5*, 45–57.

de Meneses-Gaya, C., Zuardi, A. W., Loureiro, S. R., & Crippa, J. A. S. (2009). Alcohol use disorders identification test (AUDIT): An updated systematic review of psychometric properties. *Psychology & Neuroscience, 2*, 83–97.

Eastman, M., Moore, A., Cecilione, J., Hettema, J., & Roberson-Nay, R. (2017). *Confirmatory factor structure and psychometric properties of the multidimensional peer victimization scale.* Unpublished manuscript, Richmond, VA, USA: Department of Psychiatry, Virginia Commonwealth University.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences, 6*, 21–29.

Fagerstrom, K. O., & Schneider, N. G. (1989). Measuring nicotine dependence: A review of the fagerstrom tolerance questionnaire. *Journal of Behavioral Medicine, 12*, 159–182.

Fairchild, A. J., & Finney, S. J. (2006). Investigating validity evidence for the experiences in close relationships-revised questionnaire. *Educational and Psychological Measurement, 66*, 116–135.

Feilhauer, J., Cima, M., & Arntz, A. (2012). Assessing callous–unemotional traits across different groups of youths: Further cross-cultural validation of the Inventory of Callous–Unemotional Traits. *International Journal of Law and Psychiatry, 35*, 251–262.

Forte, J. A., & Green, R. G. (1994). The reliability and validity of the index of peer relations with a clinical and nonclinical sample of adolescents. *Journal of Social Service Research, 19*, 49–65.

Fraley, R. C., Heffernan, M. E., Vicary, A. M., & Brumbaugh, C. C. (2011). The experiences in close relationships - relationship structures questionnaire: A method for assessing attachment orientations across relationships. *Psychological Assessment, 23*, 615–625.

Goodbourn, P. T., Bosten, J. M., Bargary, G., Hogg, R. E., Lawrance-Owen, A. J., & Mollon, J. D. (2014). Variants in the 1q21 risk region are associated with a visual endophenotype of autism and schizophrenia. *Genes, Brain and Behavior, 13*, 144–151.

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

Haddock, C. K., Lando, H., Klesges, R. C., Talcott, G. W., & Renaud, E. A. (1999). A study of the psychometric and predictive properties of the fagerström test for nicotine dependence in a population of young smokers. *Nicotine & Tobacco Research, 1*, 59–66.

Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology, 1*, 293–319.

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

Hayward, C., Killen, J. D., & Taylor, C. B. (1989). Panic attacks in young adolescents. *American Journal of Psychiatry, 146*, 1061–1062.

Hosokawa, T., & Ohyama, M. (1993). Reliability and validity of a Japanese version of the short-form eysenck personality questionnaire – revised. *Psychological Reports, 72*, 823–832.

Hudson, W. W. (1982). *The clinical measurement package: A field manual.* Homewood, IL: The Dorsey Press.

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.

Jackson, R. W., Snieder, H., Davis, H., & Treiber, F. A. (2001). Determination of twin zygosity: A comparison of DNA with various questionnaire indices. *Twin Research and Human Genetics, 4*, 12–18.

Kasriel, J., & Eaves, L. (1976). The zygosity of twins: Further evidence on the agreement between diagnosis by blood groups and written questionnaires. *Journal of Biosocial Science, 8*, 263–266.

Kendler, K. S., & Gardner, C. O. (2001). Monozygotic twins discordant for major depression: A preliminary exploration of the role of environmental experiences in the aetiology and course of illness. *Psychological Medicine, 31*, 411–423.

Kendler, K. S., Karkowski, L. M., & Prescott, C. A. (1999). Causal relationship between stressful life events and the onset of major depression. *American Journal of Psychiatry, 156*, 837–841.

Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry, 60*, 929–937.

Kessler, R. C., Andrews, G., Mroczek, D., 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., Avenevoli, S., McLaughlin, K. A., Green, J. G., Lakoma, M. D., Petukhova, M., … Merikangas, K. R. (2012). Lifetime co-morbidity of DSM-IV disorders in the US national comorbidity survey replication adolescent supplement (NCS-A). *Psychological Medicine, 42*, 1997–2010.

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*(6), 593–602.

Khan, A. A., Jacobson, K. C., Gardner, C. O., Prescott, C. A., & Kendler, K. S. (2005). Personality and comorbidity of common psychiatric disorders. *The British Journal of Psychiatry, 186*, 190–196.

Khoo, S.-T., West, S., Wu, W., & Kwok, O.-M. (2006). Longitudinal methods. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 301–317). Washington, DC: American Psychological Association.

Kimonis, E. R., Branch, J., Hagman, B., Graham, N., & Miller, C. (2013). The psychometric properties of the inventory of callous–unemotional traits in an undergraduate sample. *Psychological Assessment*, *25*, 84–93.

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Munoz, L. C., ... Morris, A. S. (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.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'trier social stress test' A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*, 76–81.

Lee, N. K., Oei, T. P., Greeley, J. D., & Baglioni, A. J. Jr. (2003). Psychometric properties of the drinking expectancy questionnaire: A review of the factor structure and a proposed new scoring method. *Journal of Studies on Alcohol*, *64*, 432–436.

Leyro, T. M., Zvolensky, M. J., & Bernstein, A. (2010). Distress tolerance and psychopathological symptoms and disorders: A review of the empirical literature among adults. *Psychological Bulletin*, *136*, 576–600.

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

Lissek, S., Biggs, A. L., Rabin, S. J., Cornwell, B. R., Alvarez, R. P., Pine, D. S., & Grillon, C. (2008). Generalization of conditioned fear-potentiated startle in humans: Experimental validation and clinical relevance. *Behaviour Research and Therapy*, *46*, 678–687.

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy*, *43*, 1391–1424.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales*. Sydney, Australia: Psychology Foundation.

Maas, C. J., Lensvelt-Mulders, G. J., & Hox, J. J. (2009). A multilevel multitrait-multimethod analysis. *Methodology*, *5*, 72–77.

Mahmoud, J. S. R., Hall, L. A., & Staten, R. (2010). The psychometric properties of the 21-Item depression anxiety and stress scale (DASS-21) among a sample of young adults. *Southern Online Journal of Nursing Research*, *10*, 21–34.

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

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., … Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the national comorbidity survey replication–adolescent supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 980–989.

Miller, M. W., Greif, J. L., & Smith, A. A. (2003). Multidimensional personality questionnaire profiles of veterans with traumatic combat exposure: Externalizing and internalizing subtypes. *Psychological Assessment*, *15*(2), 205.

Mineka, S., & Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: It's not what you thought it was. *American Psychologist*, *61*, 10–26.

Mulraney, M. A., Melvin, G. A., & Tonge, B. J. (2014). Psychometric properties of the affective reactivity index in Australian adults and adolescents. *Psychological Assessment*, *26*, 148–155.

Muris, P., Meesters, C., de Kanter, E., & Timmerman, P. E. (2005). Behavioural inhibition and behavioural activation system scales for children: Relationships with Eysenck's personality traits and psychopathological symptoms. *Personality and Individual Differences*, *38*, 831–841.

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

National Institute of Mental Health (NIMH). (2017). RDoC matrix. Retrieved June 20, 2017, from https://www.nimh.nih.gov/research-priorities/rdoc/constructs/rdoc-matrix.shtml.

Nugent, N. R., Tyrka, A. R., Carpenter, L. L., & Price, L. H. (2011). Gene–environment interactions: Early life stress and risk for depressive and anxiety disorders. *Psychopharmacology*, *214*, 175–196.

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, 21*, 1391–1399.

Papp, L. A., Klein, D. F., Martinez, J., Schneier, F., Cole, R., Liebowitz, M. R., … Gorman, J. M. (1993). Diagnostic and substance specificity of carbon-dioxide-induced panic. *American Journal of Psychiatry*, *150*, 250–250.

Parker, G. (1989). The parental bonding instrument: Psychometric properties reviewed. *Psychiatric Developments*, *7*, 317–335.

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

Peterson, R. A., & Heilbronner, R. L. (1987). The anxiety sensitivity index: Construct validity and factor analytic structure. *Journal of Anxiety Disorders*, *1*, 117–121.

Pine, D. S., Klein, R. G., Roberson-Nay, R., Mannuzza, S., Moulton, J. L., Woldehawariat, G., & Guardino, M. (2005). Response to 5% carbon dioxide in children and adolescents: Relationship to panic disorder in parents and anxiety disorders in subjects. *Archives of General Psychiatry*, *62*, 73–80.

Pomerleau, C. S., Carton, S. M., Lutzke, M. L., Flessland, K. A., & Pomerleau, O. F. (1994). Reliability of the Fagerstrom tolerance questionnaire and the Fagerstrom test for nicotine dependence. *Addictive Behaviors*, *19*, 33–39.

Power, R. A., Tansey, K. E., Buttenschøn, N., Cohen-Woods, S., Bigdeli, T., Hall, L. S., … Lewis, C. M. (2017). Genome-wide association for major depression through age at onset stratification: Major depressive disorder working group of the Psychiatric Genomics Consortium. *Biological Psychiatry, 81*, 325–335.

Rappaport, L. M., Sheerin, C., Savage, J. E., Hettema, J. M., & Roberson-Nay, R. (2017). Clinical characteristics of latent classes of CO_2 hypersensitivity in adolescents and young adults. *Behaviour Research and Therapy*, *93*, 95–103.

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy*, *24*, 1–8.

Roberson-Nay, R., Gorlin, E. I., Beadel, J. R., Cash, T., Vrana, S., & Teachman, B. A. (2017). Temporal stability of multiple response systems to 7.5% carbon dioxide challenge. *Biological Psychology*, *124*, 111–118.

Roberson-Nay, R., Klein, D. F., Klein, R. G., Mannuzza, S., Moulton, J. L., Guardino, M., & Pine, D. S. (2010). Carbon dioxide hypersensitivity in separation-anxious offspring of parents with panic disorder. *Biological Psychiatry*, *67*, 1171–1177.

Roberson-Nay, R., Moruzzi, S., Ogliari, A., Pezzica, E., Tambs, K., Kendler, K. S., & Battaglia, M. (2013). Evidence for distinct genetic effects associated with response to 35% CO_2. *Depression and Anxiety*, *30*, 259–266.

Rubin, K. H., Coplan, R. J., & Bowker, J. C. (2009). Social withdrawal in childhood. *Annual Review of Psychology*, *60*, 141–171.

Sanderson, W. C., Rapee, R. M., & Barlow, D. H. (1989). The influence of an illusion of control on panic attacks induced via inhalation of 5.5% carbon dioxide-enriched air. *Archives of General Psychiatry*, *46*, 157–162.

Sato, T. (2005). The Eysenck personality questionnaire brief version: Factor structure and reliability. *Journal of Psychology*, *139*, 545–552.

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, *88*, 791–804.

Selin, K. H. (2003). Test-retest reliability of the alcohol use disorder identification test in a general population sample. *Alcoholism: Clinical and Experimental Research*, *27*, 1428–1435.

Semrud-Clikeman, M. (2007). *Social competence in children*. New York, NY: Springer.

Sibley, C. G., Fischer, R., & Liu, J. H. (2005). Reliability and validity of the revised experiences in close relationships (ECR-R) self-report measure of adult romantic attachment. *Personality and Social Psychology Bulletin*, *31*, 1524–1536.

Silva, H. A. D., Passos, M. H. P. D., Oliveira, V. M. A. D., Palmeira, A. C., Pitangui, A. C. R., & Araújo, R. C. D. (2016). Short version of the depression anxiety stress scale-21: Is it valid for Brazilian adolescents?. *Einstein*, *14*, 486–493.

Slavich, G. M., & Epel, E. S. (2010). *The stress and adversity inventory (STRAIN): An automated system for assessing cumulative stress exposure*. Los Angeles, CA: University of California, Los Angeles.

Slavich, G. M., & Shields, G. S. (2017). Assessing lifetime stress exposure using the stress and adversity inventory for adults (Adult STRAIN): An overview and initial validation. *Psychosomatic Medicine*. Advance online publication. doi:10.1097/PSY.0000000000000534.

Stepniak, B., Papiol, S., Hammer, C., Ramin, A., Everts, S., Hennig, L., … Ehrenreich, H. (2014). Accumulated environmental risk determining age at schizophrenia onset: A deep phenotyping-based study. *Lancet Psychiatry, 1*, 444–453.

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.

Strong, D. R., Lejuez, C. W., Daughters, S., Marinello, M., Kahler, C. W., & Brown, R. A. (2003). *The computerized mirror tracing task, version 1*. Unpublished manuscript.

Trentacosta, C. J., & Fine, S. E. (2010). Emotion knowledge, social competence, and behavior problems in childhood and adolescence: A meta-analytic review. *Social Development*, *19*, 1–29.

Turner, N., Joinson, C., Peters, T. J., Wiles, N., & Lewis, G. (2014). Validity of the short mood and feelings questionnaire in late adolescence. *Psychological Assessment*, *26*, 752–762.

Verhulst, B. (2016). A power calculator for the classical twin design. *Behavior Genetics*, *47*, 1–7.

Vujanovic, A. A., Arrindell, W. A., Bernstein, A., Norton, P. J., & Zvolensky, M. J. (2007). Sixteen-item anxiety sensitivity index: Confirmatory factor analytic evidence, internal consistency, and construct validity in a young adult sample from the Netherlands. *Assessment*, *14*, 129–143.

Walker, E., Downey, G., & Bergman, A. (1989). The effects of parental psychopathology and maltreatment on child behavior: A test of the diathesis-stress model. *Child Development*, *60*, 15–24.

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). The PTSD checklist for DSM-5 (PCL-5). Retrieved from http://www.ptsd.va.gov.

Weissman, M. M., Paykel, E. S., Siegel, R., & Klerman, G. L. (1971). The social role performance of depressed women: Comparisons with a normal group. *American Journal of Orthopsychiatry*, *41*, 390.

Wittchen, H.-U. (1994). Reliability and validity studies of the WHO-Composite international diagnostic interview (CIDI): A critical review. *Journal of Psychiatric Research*, *28*, 57–84.

Wolpe, J. (1969). *The practice of behavior therapy*. New York, NY: Pergamon Press.

Young, R. M., & Oei, T. P. S. (1996). *Drinking expectancy profile: Test manual*. Brisbane, Australia: Behavior Research and Therapy Centre, University of Queensland.