



Published in final edited form as:

J Clin Child Adolesc Psychol. 2015 ; 44(2): 264–279. doi:10.1080/15374416.2013.859080.

Implementing Psychophysiology in Clinical Assessments of Adolescent Social Anxiety: Use of Rater Judgments Based on Graphical Representations of Psychophysiology

Andres De Los Reyes¹, Tara M. Augenstein¹, Amelia Aldao², Sarah A. Thomas¹, Samantha Daruwala¹, Kathryn Kline¹, and Timothy Regan¹

Andres De Los Reyes: adlr@umd.edu; Tara M. Augenstein: taugenst@umd.edu; Amelia Aldao: aldao.1@osu.edu; Sarah A. Thomas: thomas82@umd.edu; Samantha Daruwala: sdaruwal@gmail.com; Kathryn Kline: kkline1@umd.edu; Timothy Regan: tregan2149@gmail.com

¹University of Maryland, College Park, MD 20742

²The Ohio State University, Columbus, OH

Abstract

OBJECTIVE—Social stressor tasks induce adolescents’ social distress as indexed by low-cost psychophysiological methods. Unknown is how to incorporate these methods within clinical assessments. Having assessors judge graphical depictions of psychophysiological data may facilitate detections of data patterns that may be difficult to identify using judgments about numerical depictions of psychophysiological data. Specifically, the *Chernoff Face* method involves graphically representing data using features on the human face (eyes, nose, mouth, and face shape). This method capitalizes on humans’ abilities to discern subtle variations in facial features. Using adolescent heart rate norms and Chernoff Faces, we illustrated a method for implementing psychophysiology within clinical assessments of adolescent social anxiety.

METHOD—Twenty-two clinic-referred adolescents completed a social anxiety self-report and provided psychophysiological data using wireless heart rate monitors during a social stressor task. We graphically represented participants’ psychophysiological data and normative adolescent heart rates. For each participant, two undergraduate coders made comparative judgments between the dimensions (eyes, nose, mouth, and face shape) of two Chernoff Faces. One Chernoff Face represented a participant’s heart rate within a context (baseline, speech preparation, or speech-giving). The second Chernoff Face represented normative heart rate data matched to the participant’s age.

RESULTS—Using Chernoff faces, coders reliably and accurately identified contextual variation in participants’ heart rate responses to social stress. Further, adolescents’ self-reported social anxiety symptoms predicted Chernoff Face judgments, and judgments could be differentiated by social stress context.

Correspondence regarding this manuscript should be addressed to Andres De Los Reyes, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland, Biology/Psychology Building, Room 3123H, College Park, MD 20742; Office: 301-405-7049; Fax: 301-314-9566; adlr@umd.edu.

Andres De Los Reyes, Tara M. Augenstein, Sarah A. Thomas, Samantha Daruwala, Kathryn Kline, Timothy Regan, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland, College Park, MD. Amelia Aldao, Department of Psychology, The Ohio State University, Columbus, OH.

CONCLUSIONS—Our findings have important implications for implementing psychophysiology within clinical assessments of adolescent social anxiety.

Keywords

adolescents; Chernoff Face; psychophysiology; social anxiety; social stress

Social anxiety is highly prevalent, characterized by intense fear and avoidance of social situations that involve the potential for evaluation or rejection by others, and may significantly impair social, educational, and/or work functioning (Kessler et al., 2005). Adolescents evidence a spike in social anxiety relative to earlier and later developmental periods (Grant, et al., 2005), indicating that this is a key period for assessment and intervention. As with other mental health concerns (e.g., aggression, depression, and hyperactivity), “best practices” when assessing adolescent social anxiety involve gathering clinical information using multiple sources (e.g., parents, adolescents, and clinicians or trained raters) and methods (e.g., surveys, interviews, and behavioral observations; see Hunsley & Mash, 2007; Silverman & Ollendick, 2005). Yet, researchers and practitioners rarely incorporate psychophysiological measures in clinical assessments of child and adolescent social anxiety, and treatment studies rarely test efficacy using psychophysiological measures (e.g., Davis & Ollendick, 2005). This is an important limitation to address in light of physiological arousal’s role as a key component of clinical models of social anxiety (for a review, see Barlow, 2002). Clinical models of social anxiety often inform the techniques used in social anxiety treatments, and thus the absence of psychophysiological measures limits our ability to examine whether treatments enact changes in this crucial element of patients’ responses to social situations (Davis, May, & Whiting, 2011).

In fact, the absence of psychophysiological measures in clinical work and research is an important challenge facing mental health research and practice generally. Specifically, recently the National Institutes of Mental Health (NIMH) launched a research initiative to promote advancements in clinical neuroscience for the purposes of improving patient care (i.e., Research Domain Criteria; Insel et al., 2010). To this end, biological measures such as psychophysiology have the potential to improve methods of care for adolescent patients. However, mental health professionals will likely require tools for facilitating the use of biological data. Indeed, if biological measures such as psychophysiology historically have been underutilized in clinical decision-making and care for adolescent social anxiety patients (e.g., Davis et al., 2011), then the likelihood is quite high that few mental health professionals have the expertise to interpret biological data, thus limiting the potential of these tools for clinical use. If researchers and practitioners do not have the tools available to facilitate use of biological measures within clinical assessments, it is reasonable to expect the NIMH to encounter challenges to seeing its research priorities lead to innovations in patient care and improvements in public health. In this study, we illustrate the implementation of psychophysiological measures in clinical assessments of adolescent social anxiety. Specifically, we describe and test a paradigm for interpreting psychophysiological data, using methods designed to facilitate interpreting these data, namely graphical representations.

Two recent technological advancements bring us closer to integrating psychophysiology in clinical assessments of adolescent social anxiety. First, social stressor tasks induce adolescent social stress as indexed by psychophysiology (e.g., heart rate during tasks; Bouma, Riese, Ormel, Verhulst, & Oldehinkel, 2009). In fact, tasks that elicit biological reactions to social stress appear to be tasks: (a) that include unpredictable social scenarios, (b) whose outcomes cannot be controlled by the participant, and (c) that expose the participant to having key facets of their self-identity negatively evaluated by others (Gunnar, Talge, & Herrera, 2009). For instance, the Groningen Social Stress Task (GSST; Bouma et al., 2009) elicits social stress using contexts relevant to public speaking and other evaluative contexts (i.e., preparing and giving a speech and performing mental arithmetic in front of unfamiliar confederates). Noteworthy among available tasks, the GSST was recently implemented in a large-scale epidemiological study of adolescent stress, in a sample of over 600 adolescents aged 15–17 years (Bouma et al., 2009). Thus, the GSST shows much promise for improving adolescent social anxiety assessments and can be administered in a large scale. Second, recent advancements allow for the low-cost and non-invasive implementation of wireless heart rate monitors that collect multiple psychophysiological metrics (for a review, see Thomas, Aldao, & De Los Reyes, 2012). In fact, these heart rate monitors have been used to assess whether psychophysiology predicts clinical indices (e.g., clinic referral status and diagnostic status; Anderson & Hope, 2009; De Los Reyes et al., 2012a).

Researchers and practitioners also encounter challenges to using psychophysiological measures when clinically assessing adolescent social anxiety. First, the data are equivocal as to whether it is valid to interpret adolescents' psychophysiological reactions to social stressor tasks as indicators of social anxiety symptoms. For example, in a community sample of 49 children and adolescents ages 6–17 years, heart rate but not skin conductance responses to viewing a video of a large dog significantly related to children's and adolescents' self-reported anxiety symptoms, but not parent-reported child and adolescent anxiety symptoms (Weems, Zakem, Costa, Cannon, & Watts, 2005). Conversely, studies comparing the responses to laboratory stressor tasks of anxious and non-anxious children and adolescents largely find no significant between-group differences in responses taken from objective psychophysiological indicators (e.g., Anderson & Hope, 2009; Gonzalez, Moore, Garcia, Thienemann, & Huffman, 2011; Miers, Blote, Sumter, Kallen, & Westenberg, 2011). Indeed, two of these studies found that anxious children and adolescents could be differentiated from non-anxious controls based on subjective self-reports of psychophysiological arousal during laboratory stressor tasks (Anderson & Hope, 2009; Miers et al., 2011). Overall, a key challenge involves identifying social stressor tasks for which psychophysiological reactions relate to clinical indicators of social anxiety.

Second, the GSST currently does not have normative data available by which researchers and practitioners can compare patients' responses to these tasks. That is, researchers developed the GSST to elicit social stress in the general population by creating laboratory simulations of social situations that most adolescents would find anxiety provoking (e.g., public speaking; see Anderson & Hope, 2009; Steiner, Ryst, Berkowitz, Gschwendt, & Koopman, 2002). Thus, the GSST was designed to elicit social stress regardless of social anxiety status, making it difficult to compare patients' reactions to the task to control

participants' reactions to the same task. Third, perhaps in light of the first two challenges, it should not be surprising that researchers and practitioners rarely incorporate psychophysiological assessments within clinical assessments (e.g., Davis et al., 2011; Davis & Ollendick, 2005). Consequently, researchers and practitioners might benefit from methods for assessing and interpreting adolescent patients' psychophysiology that: (a) relate to clinical reports of social anxiety symptoms, (b) incorporate normative physiological data to interpret patients' task performance, (c) do not require an extensive background in psychophysiology, and (d) can be administered within a reasonable timeframe.

Methods exist for addressing challenges to implementing psychophysiological methods when clinically assessing adolescent social anxiety. Specifically, clinical norms exist with which to compare indices of adolescent patients' psychophysiology (i.e., heart rate) to the normative resting physiological functioning of same-age adolescents (Siegfried & Henderson, 2002). However, a key issue is that social stressor tasks such as the GSST include multiple contexts that might elicit individual differences in reactivity. This is particularly important given the role that contextual factors play in emotional expression and regulation (for reviews, see Aldao, 2013; Chaplin & Aldao, 2012). Indeed, a patient's heart rate during a social stressor task might be elevated relative to clinical norms or not, depending on the context. Patients might experience stress when preparing a speech, giving a speech, within a resting baseline assessment, or none of these contexts. Evidence of these contextual variations comes from a recent study in which socially anxious adult patients could be differentiated based on whether they evidenced social skills deficits across social contexts (e.g., impromptu speech, structured or unstructured one-on-one social interactions) or only within specific contexts (De Los Reyes, Bunnell, & Beidel, 2013). Thus, social stressor tasks may elicit psychophysiological reactivity differently across contexts, yielding a multivariate profile of psychophysiology for which some, none, or all of the contexts elicit an elevated reaction to social stress relative to clinical norms. Therefore, researchers and practitioners would benefit from a method that facilitates interpretations of adolescent patients' psychophysiology.

Interpretations of psychophysiological assessments can be augmented by methods for graphically representing psychophysiological data. Decades of research indicate that graphically depicting data facilitates reliable and valid decision-making (e.g., Lee, Butavicius, & Reilly, 2003). Specifically, graphical depictions allow users to detect patterns in data that may go undetected when interpreting data in numerical form (e.g., Cleveland & McGill, 1984). Relative to alternative forms of data presentation, graphically depicting data may enhance user interest, increase memory retention of information depicted, and decrease decision-making time (i.e., graphics allow for large amounts of data to be presented and viewed all at once; Feinberg, 1979). Thus, graphically depicting psychophysiological data may allow an assessor to use and interpret such data more efficiently, reliably, and validly than if she or he were tasked with interpreting the data as presented numerically (see also Lipkus & Hollands, 1999).

One method for graphically representing data involves creating graphics that treat each individual data point in a multivariate profile (e.g., heart rate measurements taken during different contexts of a social stressor task) as a feature on the human face (e.g., eyes, nose,

mouth, and face shape), a method known as the Chernoff Face (Chernoff, 1973). Using Chernoff Faces, one might create a facial graphic to represent patients' heart rates, a comparison graphic to represent clinical heart rate norms, and have people (e.g., research personnel or clinicians) judge differences between characteristics of two graphics (e.g., narrowness of the eyes, nose, or mouth; see Figure 1, 2, 3, Table 3). The value in using Chernoff Faces is that they capitalize on humans' abilities to discern subtle variations in human facial features (Chernoff, 1973). Importantly, Chernoff Faces have been used extensively in a variety of professional contexts, including business (Huff, Mahajan, & Black, 1981), environmental policy (Apaiwongse, 1995), and medical decision-making settings (S.Y. Lee, S.X. Lee, Decker, & Roberts, 2012). Therefore, prior implementations of Chernoff Faces speak to the likelihood of its utility generalizing to other contexts, such as within clinical assessments of adolescent social anxiety. Specifically, graphically representing psychophysiological data using Chernoff Faces may allow an assessor to make comparative judgments between adolescent patients' psychophysiological functioning during social stressor tasks and clinical norms of physiological functioning. Using these comparative judgments, one might identify whether a given adolescent patient, relative to clinical norms, expresses elevated psychophysiological reactivity within and across the contexts represented by a social stressor task (i.e., baseline vs. speech preparation vs. speech-giving).

Purpose and Hypotheses

The purpose of this paper was to extend the literature on implementing psychophysiology in clinical assessments of adolescent social anxiety. In a sample of clinic-referred adolescent social anxiety patients who provided psychophysiological data within multiple contexts (i.e., baseline, speech preparation, and speech-giving during a social stressor task), we extended the literature in three ways. First, we examined whether laboratory coders could reliably and validly make Chernoff Face judgments that identified instances in which adolescents' heart rates were higher or lower than published clinical norms. Consistent with prior work using Chernoff Faces (Apaiwongse, 1995; Huff et al., 1981; S.Y. Lee et al., 2012), we expected to find that coders would reliably and accurately judge relative differences between Chernoff Face representations of adolescent patients' heart rates and normative heart rates.

Second, we examined whether we could detect differences between adolescent patients' heart rates across laboratory social stressor contexts. In line with recent work with adults (De Los Reyes et al., 2013), we expected to observe contextual variation in adolescent patients' heart rates, as represented by Chernoff Faces. Specifically, we expected coders to differentiate adolescents' heart rates from normative heart rates at different frequencies as adolescents transitioned from the baseline, to speech preparation, to speech-giving contexts.

Third, we examined whether changes in adolescent patients' heart rates could be predicted by adolescents' self-reported social anxiety and adolescent gender. As in prior work in community samples (Weems et al., 2005), we expected greater adolescents' self-reported social anxiety symptoms to predict Chernoff Face judgments of adolescent patients' heart rates that indicated greater heart rates relative to clinical norms.

Method

Study Procedures

The Internal Review Board of the large Mid-Atlantic university at which we conducted the study approved all procedures. After we screened respondents for eligibility over the telephone, we scheduled them for an assessment. Participants completed all measures in this study as part of a single visit. To be eligible to participate, families needed to have an adolescent and parent present for the assessment. After parents provided written consent and adolescent participants provided written assent, we provided participants with training materials to familiarize them to wearing heart rate monitors. Following training, gender-matched research personnel led participants to a private and secure area to apply their heart rate monitor. Following application of the heart rate monitor, research personnel assessed each participant's baseline psychophysiology, which we examined in this study. Research personnel then led participants to a separate room to complete assessments via individual computer-based questionnaires. We then administered the social stressor task to the adolescent. For survey assessments, participants provided computer-based responses to items via IBM SPSS Data Collection survey administration software (Version 5.6; IBM Corporation, 2009). Upon study completion, we debriefed participants as to study goals and provided monetary compensation.

Participants

We recruited participants through community agencies and events, online advertisements (e.g., Craigslist), and newspaper advertisements in qualifying neighborhoods (i.e., neighborhoods targeted for demographic variability). Additional recruitment sources included pediatricians, mental health professionals, and other health care providers. Adolescents' parents contacted the laboratory in response to an advertisement for a social anxiety screening evaluation offering feedback on adolescents' social anxiety and referrals for diagnostic testing. In order to participate, families had to: (a) speak English, (b) understand the consent and assent process, (c) have an adolescent currently living in the home who the parent did not report as having a history of learning or developmental disabilities, and (d) complete information on all constructs.

Participants included 22 adolescent patients from a larger study of adolescent social anxiety (De Los Reyes, Aldao et al., 2012). In previous research, we have demonstrated that these adolescent patients can be significantly differentiated from age- and gender-matched community control adolescents on social anxiety symptoms, psychophysiology, and associated features of social anxiety (i.e., safety-seeking behaviors; De Los Reyes, Aldao et al. 2012; Thomas, Daruwala, Goepel, & De Los Reyes, 2012). Thus, patients in this study were clinically referred for a social anxiety evaluation, and their clinical referral was corroborated by significant differences between control adolescents on social anxiety and related concerns.

The sample included families with an adolescent aged 14–to-17 years (10 boys and 12 girls; $M=15.23$ years; $SD=1.06$) who lived in a large metropolitan area in the Mid-Atlantic United States. Parents consisted of 4 (18.2%) male and 18 (81.8%) female caregivers. Parental

relationship status to the adolescent varied and consisted of 20 biological parents, 1 adopted parent, and 1 parent reported relationship status as “other”, but declined to specify. Parental marital status also varied, with 15 (68.2%) married or cohabitating, 2 (9.1%) divorced, 2 (9.1%) widowed, and 3 (13.6%) never married. Regarding parental education history, 3 (13.6%) did not complete high school, 1 (4.5%) received a high school diploma, 6 (27.3%) completed some college, 2 (9.1%) received an associate’s degree, 4 (18.2%) received a bachelor’s degree, and 6 (27.3%) received an advanced degree (e.g., master’s degree, Ph.D., MD, JD).

The parent identified family ethnicity/race as African American or Black (59.1%), White, Caucasian American, or European (31.8%), Asian or Asian American (9.1%), Hispanic or Latino/a (4.5%), or “Other” (4.5%; 1 participant entered “Indian”).¹ Parents reported weekly household income across 10 categories varying by \$100 increments (i.e., *Less than \$100 per week* through *901+ per week*). Based on this scale, nearly one third (31.8%) of the families had a weekly household income of \$500 or less, 40.9% had a weekly income between \$501 and \$900, and 27.3% earned \$901 or more per week. These figures are consistent with the economic and ethnic representation of the geographic area of recruitment (U.S. Census Bureau, 2010).

Measures and Behavioral Tasks

Adolescent survey measures—Families completed measures assessing domains of adolescent and family demographics and adolescent social anxiety:

Adolescent and family demographics: We obtained demographic data through parent reports of child age and gender, family ethnicity/race, and family income.

Self-reports of adolescent social anxiety symptoms: We assessed adolescent self-reported social anxiety using the Multidimensional Anxiety Scale for Children (MASC; March, 1997), a 39-item scale that assesses various domains of anxiety functioning in youths: physical symptoms, harm avoidance, social anxiety, and separation anxiety/panic. Adolescents rated each symptom on a scale from “0” (*Never true about me*) to “3” (*Often true about me*). Total scores could range from 0 to 117, with higher scores reflecting greater anxiety. Much research attests to the MASC’s reliability and validity (March, 1997; Silverman & Ollendick, 2005). For the current study, we examined the nine-item Social Anxiety subscale, which yielded high internal consistency estimates, $\alpha=0.86$.

Adolescent exposure to social stress

To assess adolescents’ psychophysiology, we administered the GSST (Bouma et al., 2009). The GSST elicits behavioral and biological responses to social stress via settings typified by social evaluation by authority figures (Gunnar et al., 2009). Specifically, the task began with a participant being led into an unoccupied room where a research assistant instructed her or him to spend ten minutes preparing a speech on their favorite hobby or activity. We then led

¹The composition of family ethnicity/race exceeds 100% because there was overlap among the ethnic/racial categories (i.e., participants could select more than one category).

the participant to a second room, and instructed the participant to give a six-minute speech in front of three trained confederates with whom the participant did not previously have contact. After six minutes had passed, the research assistant who originally led the participant into the speech room informed the participant that the video camera taping their speech malfunctioned, and the research assistant asked the participant to wait until personnel resolved the problem. After three minutes, the research assistant returned to instruct the participant to complete a numerical task that involved the participant performing mental subtraction (i.e., subtracting 17 from a large number) for six minutes, again in front of the trained confederates. For this study, we examined adolescents' mean heart rates within the speech preparation and speech-giving contexts, and a baseline period taken before the GSST.

Adolescents' heart rate during baseline and GSST tasks

Consistent with recent work, adolescents wore ambulatory heart rate monitors (i.e., Polar Electro RS800CX; Anderson & Hope, 2009; De Los Reyes, Aldao et al., 2012). Specifically, gender-matched research personnel assisted each adolescent in applying the heart rate monitor, which consisted of a wristwatch and a dampened elastic chest band worn underneath clothing. After applying the monitor, we instructed adolescents to sit quietly during a five-minute baseline period. Due to variability in start and stop times, we calculated adolescents' mean heart rates based on the first four minutes of their baseline. Further, to equate time ranges across contexts, we took mean heart rates for adolescents during the speech preparation and speech-giving contexts of the GSST based on the first four minutes of these tasks. From each four-minute period, we calculated four one-minute mean heart rates for each adolescent (i.e., 12 one-minute mean heart rates per adolescent). Specifically, we exported the interbeat interval (IBI) series, that is, the distance in milliseconds (ms) elapsed between one heart beat and the next one, from the heart rate monitors into an electronic spreadsheet for semi-automated identification and mean replacement of outliers (i.e., defined here as a high-IBI value of "1200 and above" or low-IBI value of "200 and below"). We then imported this cleaned IBI series into a free and publically available software program, CmetX (Allen, Chambers, & Powers, 2007), for automatic beat detection and calculation adolescents' mean heart rates. We used these one-minute mean heart rates to create Chernoff Faces of patient and normative heart rate data (Figures 1–3).

We focused on adolescents' mean heart rates as opposed to other metrics (e.g., HRV or skin conductance) for three reasons. First, recent work indicates that heart rate metrics vary in their ability to distinguish between participants' performance in resting baseline versus stressor contexts, with mean heart rate yielding larger-magnitude effects than other metrics such as HRV (Allen et al., 2007). Further, prior work indicates that mean heart rate responses to laboratory threat tasks relate to child and adolescent self-reported anxiety, whereas skin conductance responses do not (Weems et al., 2005). Thus, using mean heart rate data afforded us both greater statistical power to detect hypothesized effects and a way to compare our findings to that of prior work. Second, we wanted to illustrate the Chernoff Face method using heart rate metrics that could be extracted from even relatively low-cost heart rate monitors (e.g., under \$100 US). Third, to our knowledge heart rate is the only metric for which we had available age-matched normative comparison data (Siegfried &

Henderson, 2002). Although HRV norms are available, they are quite preliminary and focus on resting HRV rather than phasic HRV in response to emotional stimuli (see Task Force, 1996). Indeed, the literature on phasic HRV in response to emotion-eliciting tasks reveals challenges in implementing HRV norms for clinical use. Specifically, emotional reactivity sometimes relates to decreases in HRV (i.e., vagal withdrawal; e.g., Rottenberg, Salomon, Gross, & Gotlib, 2005) and other times relates to increases in HRV (e.g., Aldao & Mennin, 2012). Consequently, much work remains in delineating the precise contexts under which increases and/or decreases in HRV can be considered normative.

Procedures for Applying Chernoff Face Methods to Adolescent Heart Rate Data

Creating patient Chernoff Faces—We organized every participant’s mean heart rates into single-minute segments for each context (e.g., four one-minute segments for their baseline). For each one-minute segment, we identified the highest mean heart rate in the sample to create mean proportion values for each participant. For each participant, a mean proportion value represented how high the participant’s heart rate was during a specific minute of the task (e.g., minute 1 of baseline assessment), relative to the highest mean heart rate observed in the sample on that same minute of the task. To illustrate, we report in Table 1 the highest mean heart rate value in the sample for each one-minute segment of the baseline assessment. Specifically, for the first minute of the baseline period, we identified the highest mean heart rate value in the entire sample. To create a mean proportion value for a participant’s first-minute baseline data, we divided that participant’s mean heart rate for the first minute by the highest heart rate value observed among all participants in the sample for that first minute of data. We subsequently applied this mean proportion value procedure to a participant’s remaining minutes of data across the three contexts, using the highest mean heart rate for the entire sample, for each minute (see Appendix for highest mean heart rates for the sample). The highest proportion value for any one feature (e.g., eyes) within a single participant’s Chernoff Face never exceeded “1”. This is because for any one minute of heart rate data used to create a Chernoff Face feature, the denominator we relied on was the one patient in the entire sample who had the highest heart rate for that particular minute of heart rate data (i.e., ceiling mean proportion value was “1”; Table 1).

We created a single Chernoff Face per context (i.e. baseline, speech preparation, and speech-giving) and per participant, resulting in 66 faces. We collected four minutes of heart rate data per context. Thus, we identified four features on the Chernoff Faces to correspond with the underlying heart rate data. These predetermined features included: width of eyes, width of nose, width of mouth, and width of face. We chose these features because of their visual clarity within the faces. Specifically, we surmised that subtle variations within other possible features (e.g., height of ear) might be much more difficult for coders to reliably discern than the eye, nose, mouth, and face shape features.

Once we created the proportion values, we randomly assigned each value (i.e., mean proportion value for a single minute of a task) to one of the four Chernoff Face features. We carried out this random assignment to account for coders possibly finding some facial features more discernible than others (e.g., eyes vs. nose; see Chernoff, 1973). We counterbalanced the positioning of each value within each Chernoff Face, such that the

“width of eyes” feature for one participant may represent the mean proportion value for that participant’s first minute of the baseline assessment, whereas the “width of eyes” feature for another participant might represent the second minute of her/his baseline assessment. Once we counterbalanced positioning of the four minutes across the 3 tasks, we created the faces using a specialized, free and publically available Chernoff Face package available in the statistical software package RStudio.

Creating normative control Chernoff Faces—We created normative control Chernoff Faces using the procedures described above; however, the underlying mean heart rate values were representative of clinical norm heart rates within nonclinical populations. Specifically, the clinical norms we used were based on large samples of infants, children, and adolescents recruited to collect representative data on resting heart rates (i.e., in beats per minute) for different developmental periods (i.e., Davignon et al., 1980; Park, 1996; see Table 6-4 in Siegfried & Henderson, 2002). To this end, to create normative control Chernoff Faces to compare against the heart rates of participants aged 14–16 years, we applied a mean heart rate value of “85,” which corresponded to the mean heart rate value in beats per minute for children and adolescents aged 12–16 years. Further, we applied a mean heart rate value of “80” to create normative control Chernoff Faces to compare against the heart rate values of participants aged 17 years; a value of “80” corresponded to the mean heart rate value for adolescents aged above 16 years. Thus, we matched the normative mean heart rate values to participants by age, and we calculated the mean proportion values needed to create the Chernoff Faces using the same method we used to create the patient Chernoff Faces (i.e., normative mean heart rate divided by the highest mean heart rate). Similar to Table 1, we illustrate in Table 2 the method used to create mean proportion values for normative heart rate data. Importantly, the highest mean heart rate used as the denominator for creating a patient’s mean proportion value for a one-minute segment is the same denominator used to create the mean proportion value for the normative heart rate data used to compare against the patient data for that one-minute segment (cf. Tables 1 and 2). Similar to the patient Chernoff Faces, the highest proportion value for any one feature within a normative Chernoff Face never exceeded “1”. This is because the highest mean heart rate within any one-minute segment of heart rate in our sample always exceeded the normative heart rates used to create normative control Chernoff Faces.

To ensure adequately paired comparisons, we mirrored the randomizations used to match the participant values to specific features within the Chernoff Face with the corresponding normative control Chernoff Face. For instance, if the eyes of a patient’s Chernoff Face represented the first minute of their baseline assessment, we based the eye region of the normative control Chernoff Face on a normative heart rate divided by the corresponding highest first-minute mean value of the baseline assessment. In Table 3, we illustrate comparisons between mean proportion values for a patient and normative heart rate data from one minute of each task. To further clarify these comparisons, in Figures 1, 2, and 3 we illustrate how the data provided in Table 3 map onto actual Chernoff Faces generated for the current study.

Comparative judgments between Chernoff Faces—Two senior undergraduate research assistants provided comparative judgments for all trials. Specifically, we instructed the raters to compare a specific feature between the Chernoff Faces for relative width (e.g., wider or narrower; see Figures 1–3). Coders made their judgments on stimuli presented using E-Prime Professional 2.0 software (Psychology Software Tools, Inc., Sharpsburg, PA). We placed each participant Chernoff Face side-by-side with a matched normative control Chernoff Face. Coders made four judgments per pair of Chernoff Faces (i.e., eyes, nose, mouth, and face shape), and coders made each judgment within a unique trial (Figures 1–3). Consequently, coders made 264 judgments (i.e., four judgments per pair of Chernoff Faces across 66 pairs).

We constructed our stimuli presentation methods to reduce the likelihood of our unduly influencing coder judgments. First, we blinded both coders to the study hypotheses and all data from the present study (e.g., heart rate data and adolescents' survey reports). By construction, the Chernoff Faces only included information regarding adolescents' heart rates. Thus, coders could not discern from the Chernoff Faces identifying information about the adolescents represented by the Chernoff Faces (e.g., gender, age, or ethnic/racial background; clinical data).

Second, we counterbalanced presentations of participant Chernoff Faces within pairs (i.e., presentation on left vs. right side of screen) to reduce effects of any side bias a coder may have, such as always judging features of the Chernoff Face on the left side of the screen as wider. Third, we randomized the positioning of judgments for heart rates from different contexts (and Chernoff Face features within these contexts) across the 264 trials, in order to reduce the likelihood of order effects in coder judgments. Fourth, the look and appearance of all stimuli presentations (i.e., judgments of Chernoff Face pairs) were identical and as represented in Figures 1–3, regardless of the context being represented (i.e., baseline, speech preparation, or speech-giving). In this way, coders could not be influenced simply by the look and formatting of the Chernoff Face stimuli to make judgments differently for each of the contexts.

Comparison ratings and composite scores—Each rating yielded a response indicating whether coders perceived the participant's Chernoff Face feature (e.g., eyes) as wider or narrower than the control Chernoff Face feature. We coded these responses to facilitate the creation of composite scores for each participant Chernoff Face (0=narrower, 1=wider). Specifically, for each context (i.e., baseline, speech preparation, and speech-giving), we collapsed the four unique comparisons per participant's Chernoff Face (i.e. eyes, nose, mouth, and face shape). This resulted in a composite score of perceived width of a participant's Chernoff Face features relative to the normative Chernoff Face features; scores ranged from 0–4.

Data-Analytic Plan

To assess reliability of coders' Chernoff Face judgments, we computed kappa coefficients. We also estimated percentage accuracy rates, or whether coders' Chernoff Face judgments accurately reflected numerical differences between adolescents' heart rates and normative

heart rates. Further, we calculated frequencies of the instances in which Chernoff Face judgments reflected greater adolescent heart rates relative to normative heart rates.

Tests of our hypotheses involved examining multiple Chernoff Face judgments, based on adolescents' heart rates across contexts (baseline, speech preparation, and speech-giving). These non-independent observations and correlated data structure violated assumptions underlying the general linear model (GLM). Thus, we tested our main hypothesis using generalized estimating equations (GEE): an extension of the GLM that assumes correlated observations of dependent and/or independent variables (Hanley, Negassa, Edwardes, & Forrester, 2003).

For GEE modeling, we used a binary logistic link function with an unstructured correlation matrix. Our binary logistic link function reflected the dichotomous repeated-measures dependent variable, which we describe below. We used an unstructured correlation matrix in light of the small number of dependent variables and the fact that we had complete data on all constructs for the 22 adolescents we examined. Specifically, we statistically modeled Chernoff Face judgments as a nested, repeated-measures (baseline, speech preparation, and speech-giving) dichotomous dependent variable. We statistically modeled the dependent variable as a function of three variables, and we tested codes within our nominal factor (i.e., Context) in descending order, with "baseline" as the reference code. First, we entered as an independent variable one within-subjects "Context" factor to account for the assessment period at which coders made Chernoff Face judgments (coded in ascending order of baseline, speech preparation, and then speech-giving). Second, the normative heart rate data used to create Chernoff Faces to compare against heart rates for adolescents in our sample were matched by age, so that creation of our dependent variable accounted for adolescent age. However, these data did not match on adolescent gender, and thus we entered a between-subjects "Adolescent Gender" factor (coded in ascending order of male and then female). Third, we entered as a continuous, independent between-subjects variable adolescent self-reported social anxiety symptoms via the MASC Social Anxiety subscale (centered). Statistical tests interpreted as indicating significant effects were in reference to tests statistics below the .05 *p*-value threshold.

Results

Preliminary Analyses

Scores taken from adolescent survey reports of social anxiety ($M=15$; $SD=6.12$) met the statistical assumptions for our analyses (i.e., acceptable ranges of skewness statistics [$\approx +/ -1.0$]; see Tabachnick & Fidell, 2001). This mean score for adolescent self-reports approximated the mean on this subscale found in other clinic samples (e.g., Wood, Piacentini, Bergman, McCracken, & Barrios, 2010). Further, the mean *T* score conversion of our raw scores (i.e., normative values for each patient in reference to adolescents of the same age and gender; March, 1997) was 60.45, or roughly one standard deviation above the mean of the *T* score distribution (i.e., 50).

Two of the three scores taken from the Chernoff Face comparisons between adolescents' baseline ($M=0.5$; $SD=1.3$), speech preparation ($M=0.73$; $SD=1.24$), and speech-giving

($M=1.36$; $SD=1.29$) heart rates and normative heart rates exhibited significant skewness (i.e., skewness ranged from 0.84 to 2.36). Thus, for analyses reported below, we transformed each of the three scores into dichotomous scores. Specifically, a code of “0” indicated that all Chernoff Face features for the four one-minute heart rates of an adolescent’s context (e.g., speech-giving) were judged as “narrower” (i.e., lower mean heart rate) relative to features for the normative heart rates. A code of “1” indicated that at least one of the Chernoff Face features among the four one-minute heart rates of an adolescent’s context were judged as “wider” (i.e., greater mean heart rate) relative to features for the normative heart rates.

Reliability of Chernoff Face Codes

Two coders made 264 Chernoff Face judgments (i.e., 12 judgments per adolescent). They exhibited almost perfect agreement in their Chernoff Face judgments, $\kappa=.87$, as per the kappa conventions outlined by Landis and Koch (1977). They also exhibited almost perfect agreement in the accuracy of their judgments, $\kappa=.82$. With a toss of a fair coin, we selected a single coder as the “master coder,” whose Chernoff Face data we report below.²

Accuracy of Chernoff Face Codes

Across the 264 Chernoff Face judgments, the accuracy rate (e.g., correctly judging that the eye region of an adolescent patient’s Chernoff Face was wider than the eye region of a normative comparison Chernoff Face) was 86%. This accuracy rate was significantly greater than chance (i.e., 50%) via one-sample t test, $t=16.81$ (263), $p<.001$.

Frequencies of Chernoff Face Codes by Social Stress Laboratory Context

We computed composite scores of Chernoff Face judgments within the baseline, speech preparation, and speech-giving contexts, resulting in three dichotomous scores (Table 4). Coders judged a larger proportion of Chernoff Face features for adolescents’ heart rates as greater than the Chernoff Face features of normative heart rates, as adolescents progressed from the baseline context, to speech preparation context, and then to speech-giving context. Thus, we tested the variance explained by this “Context” effect in analyses reported below.³

Relations among Changes in Chernoff Face Codes Across Periods of Social Stress, Adolescent Self-Reports of Social Anxiety, and Adolescent Gender

Relative to the differences we observed between adolescents and normative heart rates during the baseline context, we hypothesized that Chernoff Face judgments would reveal greater heart rate responses for adolescents than normative heart rates during the speech

²Originally, coders judged features on patients’ Chernoff Faces as wider, narrower, or equal, relative to normative control Chernoff Faces. This resulted in at least one coder (i.e., master or reliability coder; or both coders) making 32 “equal” judgments. In order to have complete data on all 264 Chernoff Face judgments, we had coders return and make “wider” or “narrower” judgments on the 32 comparisons originally judged as equal. Thus, our data is based on the 264 wider/narrower judgments. Importantly, the accuracy rate for the master coder’s 232 judgments (i.e., all judgments originally judged wider/narrower) was only slightly higher (i.e., 89%) than the final rate of 86% (i.e., after coders re-coded judgments originally coded as equal).

³For comparative purposes, we calculated the mean heart rates (in beats per minute) for the three assessment contexts, using the mean heart rates across the four minutes used to create patient Chernoff Faces. These mean heart rates were 73.94 for baseline, 77.06 for speech preparation, and 89.33 for speech-giving. These rates approximated the ranges of heart rates for assessments taken during baseline and stressful assessment contexts in prior work (e.g., Allen et al., 2007; Gonzalez et al., 2011; Miers et al., 2011; Weems et al., 2005).

preparation and speech-giving contexts. We also hypothesized that greater adolescent self-reports of social anxiety symptoms would predict an increased likelihood for Chernoff Face judgments to reflect greater heart rate responses for adolescents relative to normative heart rates. To test this, we conducted a GEE analysis using the analytic plan that we described previously. We report findings from this GEE analysis in Table 5. Consistent with our hypotheses, we observed a significant main effect for Context, as well as a significant main effect for adolescents' MASC Social Anxiety subscale scores, and a non-significant main effect for adolescent gender. The significant Context effect indicated that both the speech preparation and speech-giving contexts were more likely than the baseline context to elicit at least one judgment of a Chernoff Face for an adolescent's heart rate as greater than a Chernoff Face for normative heart rates. The significant main effect of the adolescent MASC Social Anxiety subscale score indicated a positive relation between adolescents' self-reported social anxiety symptoms and Chernoff Face judgments of differences between adolescents' heart rates and normative heart rates. Specifically, we observed greater adolescent self-reported social anxiety symptoms predicting an increased risk of Chernoff Face judgments to indicate that adolescents' heart rates were greater than normative heart rates. Overall, Chernoff Face coders judged adolescent heart rates to increase relative to normative heart rates as a function of social stress context, and these judgments were corroborated by adolescents' self-reported social anxiety symptoms.

Discussion

Main Findings

In this paper, we advanced the literature on implementing psychophysiological measures in clinical assessments of adolescent social anxiety. In a sample of clinic-referred adolescents, we used: (a) a laboratory task that elicits social stress in adolescents (Bouma et al., 2009), (b) low-cost wireless heart rate monitors to assess psychophysiology (Thomas et al., 2012), and (c) methods for graphically representing psychophysiological data for use by assessors to make judgments about patients' responses to social stress (Chernoff, 1973).

We made three findings. First, consistent with prior work in other professional contexts (e.g., Apaiwongse, 1995; Huff et al., 1981; S.Y. Lee et al., 2012), coders reliably and accurately judged relative differences between Chernoff Face representations of adolescent patients' heart rates and normative heart rates. Second, similar to recent work with adult social anxiety patients (De Los Reyes et al., 2013), coders identified contextual variation in adolescent patients' heart rates. Specifically, heart rates were judged by coders as greater relative to normative heart rates at varying frequencies, as a function of social stress context (i.e., baseline, speech preparation, and speech-giving contexts). Third, greater adolescents' self-reported social anxiety symptoms predicted Chernoff Face judgments of adolescent patients' heart rates that indicated greater heart rates relative to normative heart rates. Of particular interest is that this third finding extends prior work in community samples examining links between child and adolescent heart rates in response to laboratory stressor tasks and child and adolescent self-reported anxiety (Weems et al., 2005). That is, we identified evidence supporting the validity of psychophysiological reactions to social stress to indicate adolescent social anxiety symptoms.

Importantly, our findings are inconsistent with prior work finding no significant differences between anxious and non-anxious children's and adolescents' psychophysiological reactions to stressor tasks (e.g., Anderson & Hope, 2009; Gonzalez et al., 2011; Miers et al., 2011). These inconsistencies may be due to our comparing psychophysiological reactions in adolescent patients to normative heart rate data, whereas previous work has taken a case-control approach in which anxious and non-anxious children and adolescents were all exposed to the same tasks. Alternatively, because we broke down heart rate assessments for our social stressor tasks into both (a) one-minute segments and (b) social stress context, we were able to identify more "peaks" in heart rate reactivity than if we averaged heart rate responses either within four-minute chunks or across all laboratory contexts. However, these interpretations are speculative. Overall, our findings support our method for graphically representing psychophysiology when clinically assessing adolescent social anxiety.

Implications for Clinical Research and Practice

Our findings have important implications for clinical research and practice. First, we successfully used undergraduate coders to make judgments about psychophysiological data using Chernoff Face representations of these data. We consider this a key first step in implementing psychophysiological methods in clinical assessments. Indeed, our study is an important "proof-of-concept": We demonstrated a way for assessors without a background in psychophysiology to integrate psychophysiological data into interpretations of adolescents' performance during clinically meaningful social stressor tasks. Further, the method we used (i.e., Chernoff Face) has been implemented in a variety of professional contexts including financial decision-making (Huff et al., 1981), judgments about environmental policy (Apaiwongse, 1995), and medical decision-making (S.Y. Lee et al., 2012). Our work and prior use of the Chernoff Face in other professional contexts support future research seeking to examine if this method generalizes to judgments made by clinically trained personnel in applied research and practice settings.

Our data indicate that using Chernoff Faces to represent psychophysiological data, assessors can judge whether adolescent patients' socially anxious responses vary within and across contexts. If future research identifies strategies for translating our paradigm for clinical use, then clinicians may use Chernoff Faces to incorporate psychophysiology into assessments of adolescent patients, and in a way that is consistent with NIMH's efforts to integrate clinical neuroscience into models of patient care (e.g., Insel et al., 2010). For example, future research may involve testing whether Chernoff Faces can be represented and used by clinicians using modalities for integrating wireless technology (e.g., integrating data from heart rate monitors with software on mobile computing devices such as tablets or smartphones). These issues merit further study.

A key goal of our paper was to report important "proof-of-concept" data on use of Chernoff Faces for interpreting psychophysiological data within a clinic-referred sample of adolescents. However, we also wanted to provide researchers and practitioners with a preliminary guide on how to implement these methods in their own work. Specifically, Tables 1–3 and the Appendix include all of the sample-level data we implemented to create the Chernoff faces for individual patients in our sample. In turn, researchers and

practitioners may use the data from our study to calculate values to construct Chernoff Faces for both patients and normative heart rate data. Additionally, we included in an Appendix three pieces of information to further facilitate the process of constructing Chernoff Faces. First, we reported all of the highest mean heart rate values used in our study to create the proportion mean heart rate values that we implemented to create Chernoff Faces. Second, we inserted a column and row template that one could use to represent on a spreadsheet program (e.g., Excel) the heart rate data on which the Chernoff Face program relies to construct Chernoff Faces (i.e., Chernoff Face package available in the statistical software package RStudio). Third, we inserted the R-programming syntax used by the Chernoff Face program to construct the faces, along with step-by-step instructions for setting up the program commands. In sum, we provided information to assist future researchers seeking to translate our paradigm for use in research and practice settings.

Limitations

We see four limitations of this study. First, we based our use of Chernoff Face judgments to interpret psychophysiological data on adolescents' heart rates during the social stressor task. Importantly, different indices of psychophysiological functioning, such as heart rate and HRV, may yield different conclusions (e.g., Allen et al., 2007). Further, it would be premature to develop clinical norms using other metrics (e.g., HRV), in light of the difficulty in deciphering normative from clinically relevant functioning in reference to clinical tasks (Aldao & Mennin, 2012; Rottenberg et al., 2005; Task Force, 1996). Nevertheless, we encourage researchers seeking to replicate and extend our findings to use metrics other than heart rate.

Second, the Chernoff Face judgments made in this study were conducted by undergraduate research assistants. Thus, we do not know if these same judgments can be accurately made by mental health professionals. Therefore, future research ought to be conducted to examine if our findings generalize to Chernoff Face judgments made by mental health professionals in applied research and practice settings.

Third, we were statistically underpowered to test interaction effects. In fact, we tested for both Context \times MASC Social Anxiety subscale and Context \times adolescent gender interaction effects, and the models testing these interaction effects did not converge. Consequently, our findings need to be followed up in studies of samples capable of testing relatively more complex analytic models than those tested in this study. Two modeling techniques may prove particularly informative. Specifically, researchers might use traditional multiple regression and post-hoc probing approaches that test for interactions between physiology within social stress contexts and adolescents' self-reported social anxiety levels (e.g., Holmbeck, 2002). Alternatively, researchers might implement person-centered statistical models to uncover individual differences in our observed effects. For example, recent work with adult patients has incorporated latent classification analyses (LCA; McCutcheon, 1987) to identify patient subgroups that can be distinguished by observed behavioral expressions of social skills deficits within laboratory-based social interactions (De Los Reyes et al., 2013). Further, LCA has been successfully used within studies examining contextual variations in children's observed disruptive behavior within controlled laboratory

interactions between children and parental and non-parental adults (De Los Reyes, Henry, Tolan, & Wakschlag, 2009). Similarly, future research may involve using LCA to examine whether relatively larger adolescent patient samples can be subgrouped in terms of individual differences in physiological reactivity to social stressor tasks. For example, researchers might apply LCA to identifying subgroups of adolescents, or patients (a) who experience elevated heart rates relative to clinical norms within baseline, speech preparation, and speech-giving contexts or (b) whose heart rates only surpass clinical norms within the speech-giving context. Researchers could then examine the characteristics of these two groups, or whether they vary as a function of demographic characteristics (e.g., adolescent gender) or patients' self-reported social anxiety symptoms. Thus, prior work points to methods (i.e., LCA) that may inform future research seeking to uncover individual differences in our effects. Needless to say, these issues merit further study.

Fourth, we compared adolescent participants' heart rates to heart rate clinical norms that may have been limited in their validity. For example, the clinical norms we used were based on large samples of adolescents that may have differed from our current sample in terms of method of gathering heart rate data and the demographic composition of the sample (i.e., Davignon et al., 1980; Park, 1996; see Table 6-4 in Siegfried & Henderson, 2002). Yet, we were only able to adjust for normative variation by a single demographic characteristic (i.e., age). Perhaps as a consequence of both variations in assessment methodology and our ability to only account for adolescent age in our norms, we encountered some unexpected findings, namely that for many of our patients, their baseline heart rate values fell below the normative values we used (see Footnote 3). Importantly, a variety of demographic characteristics may contribute to variation in psychophysiological responses to social stressor tasks, including gender (Bouma et al., 2009). We statistically controlled for gender and observed null effects in relation to Chernoff Face judgments. However, future work ought to focus on identifying heart rate clinical norms for gender and other potentially meaningful demographic correlates of stress response, such as ethnic or racial background. Indeed, calibrating these clinical norms relative to sources of variation beyond age may enhance the validity of these norms and thus their utility when assessing adolescent patients.

An additional consideration is that we observed significant effects for adolescent self-reported social anxiety symptoms. This may indicate a potential benefit in creating adolescent heart rate norms for each of the contexts within the GSST, based on whether the adolescent self-reports social anxiety above clinical cutoffs on widely used screening instruments, such as the MASC. In this way, researchers and practitioners may have access to clinical norm data that more closely correspond to the social situations approximated within the GSST contexts (i.e., social evaluative threat).

Concluding Comments

Our findings support the reliability and validity of use of graphical methods for representing psychophysiological data in clinical assessments of adolescent social anxiety. Using the Chernoff Face method for graphically representing data, coders with minimal backgrounds in psychophysiology could reliably and accurately distinguish adolescent patients' heart rates during laboratory tasks from age-matched normative heart rates. These heart rate

judgments indicated that adolescent patients' heart rates could be differentiated from normative heart rates as a function of social stress contexts. Additionally, adolescent patients' self-reported social anxiety symptoms predicted coders' Chernoff Face judgments, thus providing corroborating evidence of the validity of these judgments. These findings have important implications for implementing psychophysiological measures in applied research and practice settings. Notably lacking from outcome assessments within studies of treatments for child and adolescent anxiety are indices of psychophysiology, and thus we have limited knowledge of treatment response on this important domain. Thus, we encourage researchers to use the methods we illustrated in this paper as a guide for implementing psychophysiological assessments in applied research and practice settings. An increased focus on incorporating psychophysiological measurements in clinical assessments may improve the reliable and valid detection of clinically relevant levels of social anxiety in adolescent patients.

Acknowledgments

This work was supported, in part, by an internal grant from the University of Maryland (College of Behavioral and Social Sciences Dean's Research Initiative) awarded to Andres De Los Reyes. This work was also partially supported by an NRSA Predoctoral Award to Sarah Thomas from the National Institute on Drug Abuse (F31-DA033913). We thank Anna Swan, Michael Van Wie, Ho-Man Yeung, and William Lechner for assisting with data collection. We also thank Martin Buschkuehl and Jeffrey S. Chrabaszcz for assistance with Chernoff Face and coding procedures.

References

- Aldao A. The future of emotion regulation research: Capturing context. *Perspectives on Psychological Science*. 2013; 8:155–172.
- Aldao A, Mennin DS. Paradoxical cardiovascular effects of implementing adaptive emotion regulation strategies in generalized anxiety disorder. *Behaviour Research and Therapy*. 2012; 50:122–130. [PubMed: 22218164]
- Allen JJB, Chambers AS, Towers DN. The many metrics of cardiac chronotropy: A pragmatic primer and a brief comparison of metrics. *Biological Psychology*. 2007; 74:243–262. [PubMed: 17070982]
- Anderson ER, Hope DA. The relationship among social phobia, objective and perceived physiological reactivity, and anxiety sensitivity in an adolescent population. *Journal of Anxiety Disorders*. 2009; 23:18–26. [PubMed: 18436426]
- Apaiwongse TS. Facial display of environmental policy uncertainty. *Journal of Business and Psychology*. 1995; 10:65–74.
- Barlow, DH. *Anxiety and its disorders: The nature and treatment of anxiety and panic*. 2nd ed.. New York, NY: Guilford Press; 2002.
- Bouma EMC, Riese H, Ormel J, Verhulst FC, Oldehinkel AK. Adolescents' cortisol responses to awakening and social stress: Effects of gender, menstrual phase and oral contraceptives. The TRAILS study. *Psychoneuroendocrinology*. 2009; 34:884–893. [PubMed: 19195792]
- Chapin T, Aldao A. Gender differences in emotion expression in children: A meta analytic review. *Psychological Bulletin*. 2012 Advance online publication.
- Chernoff H. Using faces to represent points in k -dimensional space graphically. *Journal of the American Statistical Association*. 1973; 68:361–368. from <http://www.jstor.org/stable/2284077>.
- Cleveland WS, McGill R. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*. 1984; 79:531–554.
- Davignon A, Rautaharju P, Boisselle E, Soumis F, Mégélas M, Choquette A. Normal ECG standards for infants and children. *Pediatric Cardiology*. 1980; 1:123–131.

- Davis TE III, May A, Whiting SE. Evidence-based treatment of anxiety and phobia in children and adolescents: Current status and effects on the emotional response. *Clinical Psychology Review*. 2011; 31:592–602. [PubMed: 21482318]
- Davis TE III, Ollendick TH. A critical review of empirically supported treatments for specific phobia in children: Do efficacious treatments address the components of a phobic response? *Clinical Psychology: Science and Practice*. 2005; 12:144–160.
- De Los Reyes A, Aldao A, Thomas SA, Daruwala SE, Swan AJ, Van Wie M, et al. Adolescent self-reports of social anxiety: Can they disagree with objective psychophysiological measures and still be valid? *Journal of Psychopathology and Behavioral Assessment*. 2012; 34:308–322.
- De Los Reyes A, Alfano CA, Beidel DC. Are clinicians' assessments of improvements in children's functioning "global"? *Journal of Clinical Child and Adolescent Psychology*. 2011; 40:281–294. [PubMed: 21391024]
- De Los Reyes A, Bunnell BE, Beidel DC. Informant discrepancies in adult social anxiety disorder assessments: Links with contextual variations in observed behavior. *Journal of Abnormal Psychology*. 2013; 122:376–386. [PubMed: 23421526]
- De Los Reyes A, Henry DB, Tolan PH, Wakschlag LS. Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology*. 2009; 37:637–652. [PubMed: 19247829]
- Feinberg SE. Graphical methods in statistics. *The American Statistician*. 1979; 33:165–178.
- Gonzalez A, Moore PS, Garcia AM, Thienemann M, Huffman L. Activation during observed parent-child interactions with anxious youths: a pilot study. *Journal of Psychopathology and Behavioral Assessment*. 2011; 33:159–170. [PubMed: 21765594]
- Grant BF, Hasin DS, Blanco C, Stinson FS, Chou SP, Goldstein RB, Huang B. The epidemiology of social anxiety disorder in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry*. 2005; 66:1351–1361. [PubMed: 16420070]
- Gunnar MR, Talge NM, Herrera A. Stressor paradigms in developmental studies: What does and does not work to produce mean increases in salivary cortisol. *Psychoneuroendocrinology*. 2009; 34:953–967.
- Hanley JA, Negassa A, Edwardes DB, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: An orientation. *American Journal of Epidemiology*. 2003; 157:364–375. [PubMed: 12578807]
- Holmbeck GN. Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. *Journal of Pediatric Psychology*. 2002; 27:87–96. [PubMed: 11726683]
- Huff DL, Mahajan V, Black WC. Facial representation of multivariate data. *Journal of Marketing*. 1981; 45:53–59. from <http://www.jstor.org/stable/1251471>.
- Hunsley J, Mash EJ. Evidence-based assessment. *Annual Review of Clinical Psychology*. 2007; 3:29–51.
- IBM Corporation. IBM SPSS Data Collection (version 5, 6). Somers, NY: IBM Corporation; 2009. [Computer Software].
- Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Wang P. Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry*. 2010; 167:748–757. [PubMed: 20595427]
- Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*. 2005; 62:593–602. [PubMed: 15939837]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
- Lee MD, Butavicius MA, Reilly RE. Visualizations of binary data: A comparative evaluation. *International Journal of Human-Computer Studies*. 2003; 59:569–602.
- Lee, SY.; Lee, SX.; Dekker, GA.; Roberts, CT. Proceedings of the Annual Applied Statistics Education and Research Collaboration (ASEARC) Conference. Vol. 5. New South Wales, Australia: ASEARC; 2012. Multivariate visual clustering of single nucleotide polymorphisms and

clinical predictors using Chernoff faces; p. 56-59. from <http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1001&context=asearc> [Retrieved February 26, 2013]

- Lipkus IM, Hollands JG. The visual communication of risk. *Journal of the National Cancer Institute Monographs*. 1999; 25:149–163. from <http://jncimono.oxfordjournals.org/content/1999/25/149.short>. [PubMed: 10854471]
- March, JS. *Manual for the Multidimensional Anxiety Scale for Children (MASC)*. Toronto: Multi-Health Systems; 1997.
- McCutcheon, AL. *Latent class analysis*. Newbury Park, CA: Sage Publications; 1987.
- Miers AC, Blote AW, Sumter SR, Kallen VL, Westenberg PM. Subjective and objective arousal correspondence and the role of self-monitoring processes in high and low socially anxious youth. *Journal of Experimental Psychopathology*. 2011; 2:531–550.
- Nunnally, JC.; Bernstein, IH. *Psychometric theory*. 3rd ed.. New York: McGraw-Hill; 1994.
- Park, MK. *Pediatric cardiology for practitioners*. 3rd ed.. St Louis: Mosby; 1996.
- Rottenberg J, Salomon K, Gross JJ, Gotlib IH. Vagal withdrawal to a sad film predicts subsequent recovery from depression. *Psychophysiology*. 2005; 42:277–281. [PubMed: 15943681]
- Siegfried, BH.; Henderson, TO. *Cardiology*. In: Gunn, VL.; Nechyba, C., editors. *The Harriet Lane handbook: A manual for pediatric house officers*. 16th ed.. Philadelphia: Mosby; 2002. p. 123-168.
- Silverman WK, Ollendick TH. Evidence-based assessment of anxiety and its disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*. 2005; 34:380–411. [PubMed: 16026211]
- Steiner H, Ryst E, Berkowitz J, Gschwendt MA, Koopman C. Boys' and girls' responses to stress: affect and heart rate during a speech task. *Journal of Adolescent Health*. 2002; 30S:14–21. [PubMed: 11943570]
- Tabachnick, B.; Fidell, LS. *Using multivariate statistics*. 4th ed.. Boston: Allyn and Bacon; 2001.
- Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation*. 1996; 93:1043–1065. [PubMed: 8598068]
- Thomas SA, Aldao A, De Los Reyes A. Implementing clinically feasible psychophysiological measures in evidence-based assessments of adolescent social anxiety. *Professional Psychology: Research and Practice*. 2012; 43:510–519.
- Thomas SA, Daruwala SE, Goepel KA, De Los Reyes A. Using the Subtle Avoidance Frequency Examination in adolescent social anxiety assessments. *Child and Youth Care Forum*. 2012; 41:547–559.
- U.S. Census Bureau. *State and county quickfacts*: Washington, DC. 2010 from <http://quickfacts.census.gov/qfd/states/11000.html>.
- Weems CF, Zakem A, Costa NM, Cannon MF, Watts SE. Physiological response and childhood anxiety: Association with symptoms of anxiety disorders and cognitive bias. *Journal of Clinical Child and Adolescent Psychology*. 2005; 34:712–723. [PubMed: 16232068]
- Wood JJ, Piacentini JC, Bergman R, McCracken J, Barrios V. Concurrent validity of the anxiety disorders section of the Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions. *Journal of Clinical Child and Adolescent Psychology*. 2002; 31:335–342. [PubMed: 12149971]

Appendix A

Sample Highest Mean Heart Rate Values for Creating Proportion Values

	Minute 1	Minute 2	Minute 3	Minute 4
Baseline	98.73	97.85	96.35	97.89
Speech Preparation	118.29	106.18	115.33	106.47
Speech-Giving	150.24	133.19	133.39	123.42

Sample Excel Spreadsheet Data Template to Generate Chernoff Faces

case.ID	bl	face	Bl	bl	mouth	bl	Bl	eyes	Bl	bl	bl	bl	nose	bl	bl
001	1	0.86	1	1	0.86	1	1	0.86	1	1	1	1	0.88	1	1

Note. bl=blank. Blank cells correspond to other Chernoff Face feature options in R. The template above assigns patient values to the 4 facial features used for the current study (i.e., width of face, width of mouth, width of eyes, width of nose). However, additional features are available.

Steps Used to Generate Chernoff Faces in R (with Accompanying R Code in Italics)

1. Install Chernoff Face Application:

```
library(aplpack)
```

2. Upload patient data in an Excel spreadsheet file:

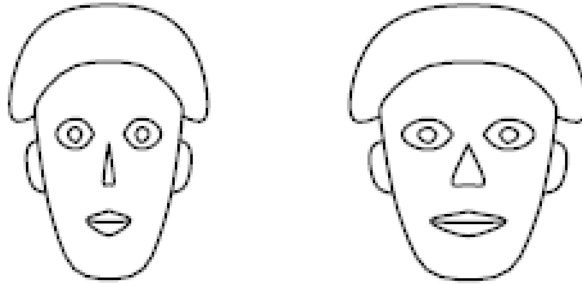
```
data <-read.csv("INSERT_FILE_NAME.csv")
```

3. 3. Run Chernoff Face Code in R:

```
faces(data[,2:16], labels=data$case.ID, face.type = 0, scale = TRUE, main = "Patient Data", col.nose = "black", col.eyes = "skyblue", col.hair = "gray", col.face = "white", col.lips = "brickred", col.ears = "white")
```

Note. The labels and main title generated for the pdf output of the Chernoff Face images may be modified from the code above.

For the faces below, is the nose for the LEFT face narrower or wider than the nose for the RIGHT face?



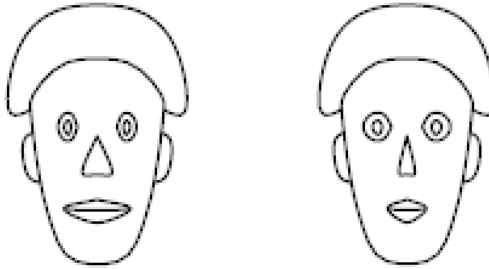
1 – Narrower

0 – Wider

Figure 1.

Sample Chernoff Face comparison trial between a patient face (left) and a normative control face (right) based on the baseline data provided in Table 3. For the patient face (left), the underlying mean proportion value for the nose=0.72. For the normative control face (right), the underlying mean proportion value for the nose=0.87.

For the faces below, is the mouth for the RIGHT face narrower or wider than the mouth for the LEFT face?



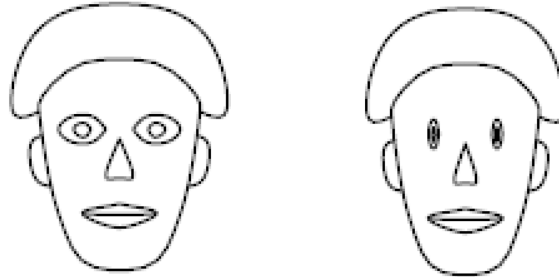
1 – Narrower

0 – Wider

Figure 2.

Sample Chernoff Face comparison trial between a patient face (right) and a normative control face (left) based on the speech preparation data provided in Table 3. For the patient face (right), the underlying mean proportion value for the mouth=0.68. For the normative control face (left), the underlying mean proportion value for the mouth=0.80.

For the faces below, are the eyes for the LEFT face narrower or wider than the eyes for the RIGHT face?



1 – Narrower

0 – Wider

Figure 3.

Sample Chernoff Face comparison trial between a patient face (left) and a normative control face (right) based on the speech-giving task data provided in Table 3. For the patient face (left), the underlying mean proportion value for the eyes=0.83. For the normative control face (right), the underlying mean proportion value for the eyes=0.56.

Table 1

Creating Proportion Mean Heart Rate Values from the Baseline Assessment for a Single Patient

Segment of Heart Rate Data	Patient's Mean Heart Rate	Highest Mean Heart Rate Across Participants	Proportion Value
Baseline: Minute 1	81.44	98.73	$(81.44/98.73)=0.82$
Baseline: Minute 2	70.93	97.85	$(70.93/97.85)=0.72$
Baseline: Minute 3	78.18	96.35	$(78.18/96.35)=0.81$
Baseline: Minute 4	76.48	97.89	$(76.48/97.89)=0.78$

Note. All heart rates that we report in this table are in beats per minute. The highest mean heart rate values used to create the proportion values represent the largest mean heart rate value within the corresponding segment of the task (e.g., baseline minute 1) across all patients included in the sample.

Table 2

Creating Proportion Mean Heart Rate Values from the Baseline Assessment for an Age-Matched Control Comparison

Segment of Heart Rate Data	Normative Mean Heart Rate	Highest Mean Heart Rate Across Participants	Proportion Value
Baseline: Minute 1	85	98.73	$(85/98.73)=0.86$
Baseline: Minute 2	85	97.85	$(85/97.85)=0.87$
Baseline: Minute 3	85	96.35	$(85/96.35)=0.88$
Baseline: Minute 4	85	97.89	$(85/97.89)=0.87$

Note. All heart rates that we report in this table are in beats per minute. The highest mean heart rate values used to create the proportion values represent the largest mean heart rate value within the corresponding segment of the task (e.g., baseline minute 1) across all patients included in the sample.

Table 3

Mean Proportion Values for Patient and Control Data for One-Minute Segments of Each Task and the Corresponding Chernoff Face Feature

Segment of Heart Rate Data	Patient Mean Proportion Value	Control Mean Proportion Value	Chernoff Face Feature
Baseline: Minute 2	0.72	0.87	Nose
Speech Preparation: Minute 4	0.68	0.80	Mouth
Speech-Giving: Minute 1	0.83	0.56	Eyes

Note. These values represent mean proportion values for a single patient and the normative control values to which we compared patient values, based on one-minute segments of each context (i.e., baseline, speech preparation, and speech-giving). See Figures 1, 2, and 3 for illustrations of how these values map onto actual Chernoff Face comparisons.

Table 4

Frequencies of Dichotomous Chernoff Face Judgments by Context

Context	a. All of Patient's Chernoff Face Features Judged as "Narrower" (i.e., Lower Mean Heart Rate) Than the Normative Chernoff Face Features	a. At Least One of Patient's Chernoff Face Features Judged as "Wider" (i.e., Greater Mean Heart Rate) Than the Normative Chernoff Face Features
Baseline	19	3
Speech Preparation	14	8
Speech-Giving	6	16

Table 5

Generalized Estimating Equations Predicting Changes in Chernoff Face Codes as a Function of Context, Adolescent Self-Reported Social Anxiety Symptoms, and Adolescent Gender

Factor	OR	B (SE)	Wald χ^2	95% CI	<i>P</i>
Intercept	0.18	-1.69 (0.87)	3.78	[-3.39, 0.01]	0.052
Context			16.64		<i>p</i> <0.001
Speech-Giving	22.77	3.12 (0.76)	16.62	[1.62, 4.63]	<i>p</i> <0.001
Speech Preparation	3.99	1.38 (0.57)	5.90	[0.26, 2.50]	<i>p</i> <0.05
Adolescent gender	0.48	-0.72 (0.82)	0.76	[-2.34, 0.90]	0.38
Adolescent self-reported social anxiety symptoms	1.18	0.17 (0.08)	3.97	[0.003, 0.34]	<i>p</i> <0.05

Note. OR=Odds Ratio; B=Unstandardized beta; SE=Standard error; 95% CI=95% Wald confidence interval. The reference code for the dependent variable (i.e., dichotomous Chernoff Face codes for adolescent heart rates at baseline, speech preparation, and speech-giving contexts) was “0”. Context factor contrasts based on comparisons of factors in descending order, with “Baseline” serving as the reference. The Context factor (coded in ascending order) was coded Baseline, Speech Preparation, and then Speech-Giving. The adolescent gender factor (coded in ascending order) was coded Male and then Female.