

**Unfamiliar Untrained Observers' Ratings of Adolescent Safety Behaviors
within Social Interactions with Unfamiliar Peer Confederates**

Taylor Rezeppa¹

Hide Okuno¹

Noor Qasmieh¹

Sarah J. Racz¹

Jessica L. Borelli²

Andres De Los Reyes¹

¹University of Maryland, College Park

²University of California at Irvine

Author Notes

Taylor Rezeppa, Hide Okuno, Noor Qasmieh, Sarah J. Racz, and Andres De Los Reyes, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland, College Park, MD. Jessica L. Borelli, Department of Psychological Science, University of California, Irvine, CA.

Efforts by the third and sixth authors were supported by a grant from the Institute of Education Sciences (R324A180032).

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; E-mail: adlr@umd.edu; Twitter: @JCCAP_Editor

CITATION

Rezeppa, T., Okuno, H., Qasmieh, N., Racz, S.J., Borelli, J.L., & De Los Reyes, A. (2021).

Unfamiliar untrained observers' ratings of adolescent safety behaviors within social interactions with unfamiliar peer confederates. *Behavior Therapy*.

<https://doi.org/10.1016/j.beth.2020.07.006>

Abstract

Adolescents experiencing social anxiety often engage in *safety behaviors*—covert avoidance strategies for managing distress (e.g., avoiding eye contact)—that factor into the development and maintenance of their concerns. Prior work supports the psychometric properties of the Subtle Avoidance Frequency Examination (SAFE), a self-report survey of safety behaviors. Yet, we need complementary methods for assessing these behaviors within contexts where adolescents often experience concerns, namely interactions with unfamiliar peers. Recent work indicates that, based on short, direct social interactions with adolescents, individuals posing as unfamiliar peers (i.e., *peer confederates*) and without assessment training can capably report about adolescent social anxiety. We built on prior work by testing whether we could gather valid SAFE reports from *unfamiliar untrained observers* (UOs), who observed adolescents within archived recordings of these short social interactions. A mixed clinical/community sample of 105 adolescents self-reported on their functioning and participated in a series of social interaction tasks with peer confederates, who also provided social anxiety reports about the adolescent. Based on video recordings of these tasks, *trained independent observers* rated adolescents' observed social skills, and an additional set of UOs completed SAFE reports of these same adolescents. Unfamiliar untrained observers' SAFE reports (a) related to adolescents' SAFE self-reports, (b) distinguished adolescents on clinically elevated social anxiety concerns, (c) related to trained independent observers' ratings of adolescent social skills within interactions with peer confederates, and (d) related to adolescents' self-reported arousal within these same interactions. Our findings support use of unfamiliar observers' perspectives to understand socially anxious adolescents' interpersonal functioning.

Keywords: Adolescents; Multiple informants; Safety behaviors; Social anxiety; Subtle Avoidance Frequency Examination

A core feature of social anxiety involves experiencing intense fears of interacting with unfamiliar people in social settings, resulting in impaired functioning and significant life interference (Morrison & Heimberg, 2013). Social anxiety disorder has a lifetime prevalence rate of 13%, making it one of the most common mental health disorders in the United States (Kessler et al., 2012). Social anxiety manifests as not only physiological symptoms like sweating and blushing (Stein & Stein, 2008), but also fear of others noticing these external symptoms and perceiving them as an indication of inferior social performance (Rapee & Heimberg, 1997).

In an effort to minimize fearful behavioral or physiological responses and the external visibility of those responses, individuals experiencing social anxiety often engage in one or more avoidance strategies (American Psychiatric Association [APA], 2013). *Overt* strategies involve explicit, direct avoidance of participation in or escape of fear-provoking social situations (Thwaites & Freeston, 2005). By definition, overt strategies have the appeal of completely removing fear-related distress from the social environment (APA, 2013). Yet, individuals experiencing social anxiety may be unable to rely on overt strategies for *all* fear-provoking social situations. Sometimes, social situations are compulsory or unavoidable, such as mandatory performance situations at school or work. Here, socially anxious individuals may experience distress that they would otherwise overtly avoid if given the opportunity to do so, leading them to engage in *covert* strategies—what we and others (for a review, see Piccirillo, Dryman, & Heimberg, 2016) term *safety behaviors*—to regulate this distress. Examples of safety behaviors include avoiding eye contact when giving a speech, or refraining from self-disclosure in one-on-one or group conversations (Alden & Bieling, 1998).

Safety behaviors provide a means for individuals to experience short-term reductions in their anxiety and distress within the immediate social interaction, making them highly

reinforcing. Yet, these behaviors have a deleterious effect on long-term functioning in at least two ways. First, socially anxious individuals often become dependent on safety behaviors, which, in turn, serve to maintain their social anxiety (Salkovskis, 1991). Specifically, those who use safety behaviors believe that they prevent negative social outcomes (McManus, Sacadura, & Clark, 2008). Thus, after positive interactions, they credit successful social outcomes to their safety behaviors, not their social abilities (Salkovskis, Clark, & Gelder, 1996). Yet, to observers, use of safety behaviors may signal an individual's limited social abilities; more broadly, these behaviors may explain the poor social performance often observed among those experiencing social anxiety (Rowa et al., 2015; Stangier, Heidenreich, & Schermelleh-Engel, 2006).

Second, among individuals who receive exposure-based therapies for social anxiety, use of safety behaviors may prevent them from experiencing long-term therapeutic benefits. Indeed, even when clients experience symptom reductions immediately post-treatment, the degree to which these benefits stand the test of time hinges on factors beyond clients' symptom reductions (Sewart & Craske, 2020). With regard to safety behaviors, clients who display these behaviors during therapeutic exposures display poorer treatment responses, relative to clients who do not display these behaviors in-session (e.g., Hedtke, Kendall, & Tiwari, 2009). Stated another way, not only do clients vary in displays of safety behaviors, but we should also not take it as a given that treatment will result in *reductions* in clients' safety behaviors. Following treatment, if a client's life impairments stem from difficulties with initiating social relationships with unfamiliar people, then how unfamiliar people perceive clients' interpersonal functioning may be a deciding factor in whether clients experience long-term reductions in life impairments in this key domain (Cannon et al., 2020). Along these lines, not only do unfamiliar people often perceive those

displaying safety behaviors as socially unskilled, but they also endorse an unwillingness to have social contact with them in the future (for a review, see Piccirillo et al., 2016).

Taken together, the presence of safety behaviors portends lasting impairments in interpersonal functioning. This is because, by construction, a socially anxious individual's ability to interact with unfamiliar people involves both that individual's willingness to initiate contact with unfamiliar people *and* unfamiliar people wanting to maintain contact with that individual. Thus, the influence of safety behaviors on the development, maintenance, and treatment of social anxiety and associated psychosocial impairments call for accurate measures of perceived safety behaviors. In particular, we require measures of how *unfamiliar people* perceive safety behaviors. In this study, we examined how unfamiliar people perceive safety behaviors, and the links between these perceptions and key domains of interpersonal functioning (e.g., social skills).

One widely used self-report instrument shows promise for gathering unfamiliar people's reports about safety behaviors. The Subtle Avoidance Frequency Examination (SAFE) assesses the use of safety behaviors in fear-provoking social situations (Cuming et al., 2009). When administered to adolescents, scores taken from the SAFE display acceptable levels of internal consistency and evidence of convergent and criterion-related validity (Cannon et al., 2020; Qasmieh et al., 2018; Thomas, Daruwala, Goepel, & De Los Reyes, 2012). In these studies, scores on the SAFE differentiated adolescents seeking a social anxiety evaluation from community control adolescents, and predicted adolescents' self-reported arousal when interacting with personnel trained to "stand in" as a same-age, unfamiliar peer (i.e., *peer confederate*). Thus, recent work supports the psychometric properties of adolescents' self-reports on the SAFE, and the ability of these self-reports to improve our understanding of the links between adolescents' safety behaviors and their interactions with unfamiliar people.

Although there is support for the use of self-report measures of safety behaviors like the SAFE, we have a limited understanding of whether we can leverage modified, psychometrically sound versions of these instruments to understand how unfamiliar people perceive safety behaviors as displayed by adolescents. Indeed, the developmental periods encompassing adolescence represent unique periods for understanding and assessing social anxiety and related processes such as safety behaviors. Relative to earlier developmental periods, adolescents assert greater independence and gradually detach from family support systems (e.g., Alfano & Beidel, 2011). This necessitates an understanding of social anxiety concerns as they manifest in contexts outside of the home, such as interactions with unfamiliar peers. Furthermore, the median age of onset for social anxiety disorder is approximately 13 years (Kessler et al., 2005), and when left untreated these concerns pose risk into adulthood for the development of not only other internalizing disorders but also substance use disorders (Epkins & Heckler, 2011; Stein & Stein, 2008). Thus, we tested an innovative approach to taking unfamiliar people's SAFE reports about adolescents, based on tasks designed to simulate common treatment targets for adolescents, namely interactions with unfamiliar peers (e.g., Alfano & Beidel, 2011; Hofmann et al., 1999; Raggi et al., 2018).

When considering people who might provide psychometrically sound reports about adolescents' safety behaviors, a key question involves identifying people who would base their reports on observations of how adolescents interact with unfamiliar peers. Other than adolescents themselves, parents comprise the most commonly used informants in evidence-based assessments of adolescent social anxiety (De Los Reyes & Makol, 2019). Prior work indicates that parents provide psychometrically sound SAFE reports in terms of internal consistency and relations with parent report measures of adolescent social anxiety and related concerns (e.g.,

depressive symptoms; Qasmieh et al., 2018). Yet, their SAFE reports about adolescents, as well as their reports of social anxiety generally, fail to predict adolescents' perceived distress when interacting with peer confederates (Cannon et al., 2020). Thus, we require observers other than parents who can capably provide SAFE reports in reference to these peer interactions.

Use of *unfamiliar untrained observers* (UOs) might allow practitioners and researchers to simulate unfamiliar people's impressions of adolescents when they display safety behaviors within their social environment. This practice would be akin to the common use of unfamiliar personnel in a clinic as "stand-ins" for interaction partners during adolescents' therapeutic exposures (see also Raggi et al., 2018). Indeed, recruiting these unfamiliar interaction partners is an acknowledgement that therapeutic exposures require simulations of key components of fear-provoking situations, and in particular the unfamiliarity of the partner. Similarly, if untrained raters possess the ability to report about safety behaviors, then their reports may facilitate demonstrating to clients that use of safety behaviors impacts not only their social anxiety, but also how unfamiliar people perceive them within social interactions. Further, use of untrained raters may facilitate not only characterizing clients' interpersonal concerns, but also tracking changes in their safety behaviors as a marker of lasting benefits of therapy.

Incorporating UOs' reports in clinical assessments of adolescents is consistent with task-sharing models in implementation science. Briefly, task sharing involves redistributing service-related tasks to individuals who lack the extensive training of licensed practitioners, in an effort to reduce health disparities (Kazdin, 2017). These models have been applied to intervention work with clients displaying a variety of clinical concerns, including anxiety, depression, and posttraumatic stress (for a review, see Singla et al., 2017). Similarly, we tested the reports of UOs who lack extensive training, in an effort to simulate the perspectives of

unfamiliar people in the adolescent's social environment. In fact, a precedent exists for task sharing in evidence-based assessments of adolescent social anxiety, in particular using raters who lack assessment training. Specifically, in recent work researchers developed the *Unfamiliar Peer Paradigm* (for a review, see Cannon et al., 2020): a series of tasks designed to estimate adolescents' social anxiety concerns within social interactions with unfamiliar peers. These tasks included structured interpersonal contexts, like scripted scenarios or speeches, or situations that lack clear structure and performance expectations, including prompts for one-on-one informal conversation. In this work, these tasks leveraged undergraduate and post-baccalaureate research personnel who were trained to "stand in" as same-age, unfamiliar peer confederates. Following 20 minutes of observing adolescents within these tasks, peer confederates completed survey reports on widely used social anxiety scales, with no training on how to make these reports. This approach is consistent with a large body of work on *thin slice judgments* in assessments of a host of psychosocial domains including personality traits, autism, and internalizing concerns (for reviews, see Cannon et al., 2020; Slepian, Bogart, & Ambady, 2014). In essence, the peer confederates completed these reports as any collateral informant would (e.g., parents and teachers), albeit with presumably a much shorter period of observation. Peer confederate reports displayed large-magnitude relations with adolescents' self-reports of social anxiety and state arousal within the social interaction tasks (Deros et al., 2018), as well as large-magnitude relations with trained independent observers' ratings of adolescents' social anxiety as displayed within these same tasks (Glenn et al., 2019).

Purpose and Hypotheses

In this study, we built on prior work by leveraging an innovative approach to gathering unfamiliar people's SAFE reports about adolescents' safety behaviors when interacting with peer

confederates. Specifically, we leveraged a team of UOs who completed a modified version of the SAFE based on archived, videotaped observations of adolescents to whom we administered the Unfamiliar Peer Paradigm. In a mixed clinical/community sample of adolescents, we addressed five hypotheses. First, consistent with prior work on SAFE self-reports, we expected UOs' SAFE reports to demonstrate acceptable levels of internal consistency (i.e., $\alpha > .90$; Cuming et al., 2009; Qasmieh et al., 2018; Thomas et al., 2012). Second, we expected UOs' SAFE reports to positively relate to adolescent self-reports on the SAFE and on social anxiety measures, as well as positively relate to peer confederates' reports of adolescent social anxiety. Third, consistent with recent work on adolescents' SAFE self-reports (Qasmieh et al., 2018), we expected UOs' SAFE reports to distinguish adolescents who displayed clinically elevated social anxiety from those who did not. Fourth, people who display safety behaviors within social interactions tend to be seen by observers as less socially skilled (Piccirillo et al., 2016), and at the same time relatively low social skills are associated with elevated levels of depressive symptoms and observers' impressions of adolescent social anxiety (Epkins & Heckler, 2011). We wanted to confirm that any relations observed between UOs' SAFE reports and adolescents' social skills were not better accounted for by adolescents' depressive symptoms or how observers tend to view their social anxiety concerns. Thus, we hypothesized that UOs' SAFE reports would be uniquely related to trained independent observers' ratings of adolescent social skills, when taking into account adolescents' self-reports of depressive symptoms and peer confederates' reports of adolescent social anxiety. Fifth, in line with prior work on adolescents' SAFE self-reports and the notion that safety behaviors serve to regulate in-the-moment distress (Qasmieh et al., 2018), we hypothesized that UOs' SAFE reports would positively relate to adolescents' subjective experiences of arousal within the social interactions from which UOs based their reports.

Method

Participants

We recruited a sample of 105 14-15-year-old adolescents and their parents from the areas of Maryland, Washington, D.C., and Northern Virginia. Our recruitment procedures included advertisements posted online (e.g., Craigslist) and flyers posted in local businesses (e.g., cafes, libraries, doctor's offices). Participants responded to one of two posted advertisements: (a) study providing a no-cost clinical social anxiety evaluation for adolescents (i.e., *clinic-referred adolescents*) or (b) nonclinical study on family relationships (i.e., *community control adolescents*). We administered identical procedures to both groups of adolescents, which we describe below. Further, prior work indicates that this approach results in groups that significantly differ in levels of social anxiety and related processes (e.g., Deros et al., 2018; Glenn et al., 2019; Karp et al., 2018; Thomas et al., 2012).

Inclusion criteria included (a) fluency in English; (b) having a 14 to 15-year-old adolescent who could read at or above their grade-level, did not have any learning or developmental disabilities, and had not received any cognitive behavioral therapy for anxiety or any other related concerns in the three months prior to the phone screening; and (c) understanding of the consent/assent process. After completing the study, parents of clinic-referred adolescents ($n = 37$) were provided with feedback on whether their adolescent displayed clinically significant levels of social anxiety, mood levels, and/or ADHD symptoms, along with referrals to services that could address these concerns. Parents of community control participants ($n = 68$) were not given feedback about their adolescents' mental health. In total, these 105 adolescents had a mean age of 14.48 years ($SD = 0.50$) and included 68 female and 37 male participants. The participating parent identified the adolescent's racial/ethnic background as

African American or Black (64%); White, Caucasian American, or European (33%); Asian American or Asian (6%); Hispanic or Latino/a (Spanish) (11%); American Indian (1%); or “Other” (10%) (rates total above 100% because parents could select multiple response options). Parent-reported family income found that 30 of the families earned \$500 or less per week, 25 earned between \$501 and \$900 per week, and 50 earned more than \$901 in income per week. For most dyads, parents were the adolescent's biological mother/father (94%). The remaining dyads included adoptive mothers/fathers (3%), stepmothers/fathers (1%), primary caregiver's significant other (1%), or the adolescent's part-time guardian (1%). Parents reported their marital status as currently married (44%), never married (24%), divorced from a significant other (17%), separated from a significant other (9%), living with a significant other (5%), or widowed (1%). The sample's demographic figures are consistent with economic and racial/ethnic data for the geographic area of recruitment (U.S. Census Bureau, 2016).

We used an analytic approach that pooled the clinic-referred and community control groups as one sample. Prior work suggests that this approach results in a dimensionally varied sample of adolescents enriched for displays of and risk for various mental health concerns (e.g., Deros et al., 2018; Glenn et al., 2019). Further, prior work demonstrates that these clinic-referred and community control groups do not differ significantly on the demographic characteristics reported previously (see Karp et al., 2018; Rausch et al., 2017). Demographic data for the two groups are available upon request from the corresponding author.

Procedure

Before the study, the Institutional Review Board for the large, mid-Atlantic university at which the study was conducted approved all procedures. Recruitment procedures consisted of a variety of strategies including the use of advertisements online (e.g., Craigslist, the laboratory's

website), on public transportation servicing the university's surrounding area, and on local advertisement boards. We also recruited the targeted demographic through the offices of local clinicians. After initial telephone eligibility screening and assessment scheduling, we received parental consent and adolescent assent for participation in the in-person assessment. The psychosocial assessments in the laboratory began with the adolescents' and their parents' completion of a computer-administered battery of survey measures using Qualtrics data collection software. Adolescent participants subsequently completed a series of simulated social interaction scenarios with trained, gender-matched research personnel posing as same-age peer confederates. Consistent with procedures used in prior work (Deros et al., 2018; Glenn et al., 2019), these peer confederates (a) were undergraduate or post-baccalaureate research assistants receiving research training in the laboratory, (b) had no prior contact with the adolescent with whom they interacted, (c) were masked to the referral status of the adolescent, and (d) were masked to all other clinical information about the adolescent (e.g., scores on mental health measures). For their participation, families received \$100 compensation (i.e., parent: \$50; adolescent: \$50), and were debriefed on study activities, including study deception (e.g., that research personnel involved in the social interaction tasks were trained to act as same-age peers).

Survey Measures

To address our study aims, we administered a survey battery to multiple informants, which included the target adolescent themselves, peer confederates, and UUOs. As part of this battery, parents completed a demographics form to collect the adolescent, parent, and family demographic information described previously. Collectively, adolescents and peer confederates completed measures that assessed safety behaviors, anxiety symptoms, and related constructs (i.e., depressive symptoms). Based on archival videos of the 105 adolescents' participation in the

Unfamiliar Peer Paradigm, we randomly assigned 26 UOs to each observe three to five video recordings; UOs made a SAFE report about each adolescent after viewing the recording. We provide in online supplementary material a complete description of UOs' characteristics.

For each measure, adolescents made self-reports whereas peer confederates and UOs made reports about the adolescent engaging in the interaction tasks. For instances in which two informants completed the same survey (e.g., adolescent and UO reports on the SAFE), we kept all item content consistent across these reports, with minor modifications to fit the informant's perspective (e.g., "I" for self-report; "the participant" for UOs and peer confederate reports).

SAFE. We administered the SAFE to both adolescents ($M = 66.08$; $SD = 21.34$; $\alpha = .94$) and UOs ($M = 80.07$; $SD = 18.62$; $\alpha = .91$) to assess safety behaviors. Each of the 32 items describe a safety behavior that could be employed in the context of a social interaction (e.g., "Position yourself/themselves so as not to be noticed" and "Before you/they arrive, excessively rehearse what you/they might say or how you/they will behave."). Informants indicated the frequency of safety behaviors using a 5-point scale ranging from "1" (*Never*) to "5" (*Always*).

Social Phobia and Anxiety Inventory for Children (SPAIC). We assessed adolescent self-reported social anxiety using the SPAIC (Beidel, Turner, & Morris, 1995). The SPAIC is a widely used 26-item measure in which each item describes a social situation and the informant reports the frequency at which the adolescent feels nervous or scared in that situation (e.g., "I feel scared when I meet new kids"), with response choices ranging from "0" (*Never*) to "2" (*Always*). Total scores range from 0 to 52, with higher scores reflecting higher levels of social anxiety. Depending on the aim of the analysis, we examined either SPAIC continuous scores or discrete scores based on established cut scores on this measure to identify clinically elevated social anxiety (i.e., scores of 18 or above; Beidel et al., 1995). The SPAIC displayed relatively

high internal consistency in this sample, $M = 16.62$; $SD = 10.65$; $\alpha = .95$. Scores for the SPAIC varied dimensionally across the sample. Of the 40 adolescents whose scores surpassed the SPAIC cut score, 18 were community control adolescents (i.e., 26% of group), and 22 were clinic-referred adolescents (i.e., 59% of group). Further, we observed wide score ranges for both community control adolescents ($M = 13.51$; $SD = 7.79$; range: 1.33-35.47) and clinic-referred adolescents ($M = 22.32$; $SD = 12.77$; range: 1.83-45.33). Importantly, the mean score for clinic-referred adolescents surpassed the SPAIC cut score mentioned previously, and clinic-referred adolescents self-reported greater social anxiety concerns on the SPAIC, relative to self-reports from community control adolescents, t (equal variances not assumed) = 3.82; $p < .001$; $d = 0.83$.

Social Interaction Anxiety Scale (SIAS). We used the 20-item SIAS to measure adolescent social anxiety concerns displayed during direct social engagement (Deros et al., 2018). Originally developed as an adult self-report, peer confederates provided adolescent social anxiety reports using a scale ranging from “0” (*Not at all characteristic or true of the participant*) to “4” (*Extremely characteristic or true of the participant*). Total scores range from 0 to 80, with higher scores reflecting higher levels of social anxiety. Peer confederates provided reports about adolescent social anxiety using a modified form of the SIAS described previously to fit the perspective of an observer of the adolescent’s social anxiety (e.g., “The participant gets nervous ...”). Each peer confederate completed a SIAS report immediately following interactions with the adolescent. Peer confederates’ SIAS reports display high internal consistency, convergent validity, incremental validity (i.e., relative to parent SIAS reports), and criterion-related validity (i.e., predicting adolescents’ arousal within social interactions); and distinguish adolescents on referral status (Cannon et al., 2020). The SIAS displayed high internal

consistency in this sample, $M = 36.97$; $SD = 17.80$; $\alpha = .96$. Peer confederates did not provide reports for two participants, and thus we based these estimates on data for 103 participants.

Beck Depression Inventory-II (BDI-II). To assess depressive symptoms, we administered the BDI-II, a measure originally designed for use with adults and adolescents aged 13 years and older (Beck, Steer, & Brown, 1996). Respondents rated items describing depressive symptoms (e.g., sadness, guilty feelings, loss of interest) on a 4-point scale, with possible score ranges from 0 to 63 and higher scores indicating greater depressive symptoms. We excluded two items, items 9 and 21, which assess for suicidality and loss of interest in sex, respectively, based on prior work showing that parents often decline to consent to having their adolescents respond to these items due to the mature subject nature (e.g., Rausch et al., 2017; Thomas et al., 2012). Despite our exclusion of the two items, internal consistency estimates of the 19 items still maintained high internal consistency, $M = 12.64$; $SD = 11.06$; $\alpha = .93$. To ensure comparability with scoring for the full version of the measure (i.e., score ranges from 0 to 63), responses for items 9 and 21 were pro-rated or estimated for each participant, based on their mean score for the 19 remaining items. The BDI-II has been used extensively to assess adolescent depressive symptoms (e.g., Glenn et al., 2019; Qasmieh et al., 2018; Rausch et al., 2017).

Unfamiliar Peer Paradigm

Adolescent participants interacted with peer confederates in the Unfamiliar Peer Paradigm (Cannon et al., 2020), a series of counterbalanced tasks designed as ecologically valid reflections of interactions with same-age, unfamiliar peers. The tasks included a Simulated Social Interaction Test (SSIT), Unstructured Conversation Task (UCT), and Impromptu Speech Task (IST). A paper by Cannon and colleagues (2020) provides an overview of the Unfamiliar Peer Paradigm, along with detailed descriptions of each of the tasks administered within the

paradigm. Further, scripted procedures for all of the Unfamiliar Peer Paradigm's tasks exist on the Open Science Framework platform (De Los Reyes, 2020).

Trained independent observers' ratings of social skills. We wished to assess adolescent social skills within the social interaction tasks described previously, and to do so in a way that would avoid shared method bias across informants used to assess other key constructs (e.g., adolescent safety behaviors and social anxiety). Thus, we leveraged behavioral reports from trained independent observers who consisted of undergraduate and post-baccalaureate research assistants. Following the same paradigm as prior work (Glenn et al., 2019), we did not provide the trained independent observers with any information regarding the adolescents' clinical status, and they did not participate in any of the social interaction tasks as a peer confederate. Using an extensively validated behavioral coding scheme (e.g., Glenn et al. 2019), the trained independent observers made macro-level ratings, across the seven interaction tasks, of each adolescent's social skills on a 5-point scale ranging from "1" (*Not effective at all*) to "5" (*Very effective*). Greater scores indicated greater social skills. We have made available in online supplementary material additional details on training of independent observers, reliability procedures, and the rating scheme used to make social skills ratings. The *ICCs* (for average measures) testing inter-rater reliability for trained independent observers' ratings displayed an average *ICC*(1,2) of .82. This average *ICC* is considered within the "excellent" range, based on thresholds recommended by Cicchetti (1994). Since we assigned two, trained independent observers to rate each adolescent, we calculated composite scores for all seven social skills ratings by averaging the two ratings for each task. Based on these composite scores, the internal consistency estimates for the seven social skills ratings was high, $\alpha = .91$. Consequently, we aggregated the seven social skills ratings for all 105 adolescents into a single mean social skills

rating ($M = 3.43$; $SD = 0.89$) to reduce Type 1 Error. We used this composite rating for the criterion-related validity tests described below. Although we computed a composite score for all adolescents, one adolescent was missing data on one of the five SSIT role plays, and three adolescents were missing data on the IST because they declined to give a speech. Therefore, for these adolescents, we based their composite scores on six, rather than seven, social skills ratings.

Adolescent self-reported state arousal. To assess adolescents' self-perceived levels of internal arousal, we administered a paper version of the Self-Assessment Manikin (SAM; Lang, 1980). The SAM consists of a 5-level pictorial scale of affect ranging from "1" (*close-eyed/relaxed image*) to "5" (*wide-eyed/nervous image*). Adolescents completed eight SAM ratings in total: one pre-task or baseline rating ($M = 1.51$; $SD = 0.62$), and one immediately after each of the seven social interaction tasks. For analyses reported below, we sought to examine changes in SAM task ratings relative to baseline. The seven SAM social task ratings displayed high internal consistency ($\alpha = .89$) and mean inter-item correlations ($r = .55$). Thus, we calculated a mean SAM rating of these seven SAM ratings ($M = 2.36$; $SD = 0.85$).

Data Analytic Plan

We followed a six-step data-analytic plan. First, we conducted a series of preliminary analyses to determine if our data conformed to assumptions of our planned parametric analyses. These included calculating Cronbach's alpha (α) to examine the internal consistency of our scales; we interpreted α levels based on thresholds of "acceptable" (.70) and "high" (.80 and above) as recommended by others (Nunnally & Bernstein, 1994). Further, we calculated skewness and kurtosis statistics to determine whether any of our continuous measures deviated from normality (i.e., skewness/kurtosis in range of + 2.0). Beyond these tests, our preliminary analyses addressed key elements of our sample. In particular, in prior work in this sample we

determined that peer confederates' reports on the SIAS, though clustered within adolescents (i.e., confederates could rate two or more participants in our sample), did not demonstrate significant design/cluster effects (Deros et al., 2018). Similarly, 26 UOs provided SAFE reports of our 105 adolescents. Although we randomly assigned cases to UOs, they provided reports on 0-2 clinic-referred adolescents. Based on prior work in this sample examining other informants' SAFE reports (Qasmieh et al., 2018), we expected clinic-referred adolescents to contribute the most variability to UOs' SAFE reports, relative to community control adolescents. Thus, we tested for significant design/cluster effects on UOs' SAFE reports separately by adolescent referral status. Specifically, we computed ICCs via one-way random analysis of variance tests for clinic-referred and community control adolescents for whom UOs provided SAFE reports (25 for clinic-referred, 26 for community control). These tests yielded between-subject variance (1.47 and 2.60) and total variance (2.79 and 4.92) statistics for clinic-referred and community control adolescents, respectively. Average cluster sizes for the clinic-referred (1.48) and community control (2.61) groups yielded design effect statistics of 1.25 for the clinic-referred adolescents and 1.85 for the community control adolescents. These design effect statistics fell below accepted thresholds of 2 (Muthén & Satorra, 1995), indicating that no significant cluster effects existed in our data. Thus, our analyses did not account for clustering.

Second, we assessed the internal consistency (α) of UO reports on the SAFE. Third, we computed bivariate correlations among survey measures completed by peer confederates, adolescents, and UOs to test relations among their SAFE reports and reports of safety behaviors and social anxiety completed by peer confederates and adolescents. Fourth, we calculated an independent samples *t*-test to determine whether UOs' SAFE reports could distinguish adolescents who self-reported scores above the clinical cut score on the SPAIC (i.e.,

18; Beidel et al., 1995) from those below the cut score. Fifth, we tested the incremental validity of UOs' SAFE reports in predicting adolescents' observed social skills using two hierarchical multiple regression models. In these models, we entered a composite score of trained independent observers' social skills ratings as the dependent variable, either adolescents' self-reports of depressive symptoms or peer confederates' reports of adolescent social anxiety in the first step as an independent variable, and UOs' SAFE reports in the second step as an independent variable. Sixth and similar to our fifth set of tests, we used a hierarchical multiple regression model to test the ability of UOs' SAFE reports to predict adolescents' self-reported arousal within social interactions with peer confederates. Here, we entered a composite score of adolescents' SAM task self-reports as the dependent variable, their SAM baseline self-reports in the first step of the equation as an independent variable, and UOs' SAFE reports in the second step as an independent variable. For all tests, we interpreted the magnitudes of correlations based on Cohen's (1988) effect size conventions for the r metric of small (.10), moderate (.30), and large (.50), and the d metric of small (.30), moderate (.50), and large (.80).

Results

Preliminary Analyses and Internal Consistency of UOs' SAFE Reports

We conducted preliminary analyses to test if our data met assumptions for planned parametric analyses (i.e., skewness/kurtosis in range of + 2.0). We examined frequency distributions for all variables used in analyses reported below to assess normality. Scores for all measures with the exception of BDI-II reports fell within acceptable ranges of skewness and kurtosis. Specifically, adolescent self-reports on the BDI-II exhibited relatively high positive skewness and kurtosis. Thus, we applied a square-root transformation to these scores. These transformed scores displayed skewness and kurtosis statistics within acceptable levels. All

analyses reported below used these transformed BDI-II scores, $M = 3.23$; $SD = 1.50$. As mentioned previously and in support of our hypotheses, we observed high levels of internal consistency for UOs' SAFE reports, $\alpha = .91$.

Convergent Validity of UOs' SAFE Reports

We report in Table 1 correlations among all measures. Consistent with our hypotheses, UOs' SAFE reports related to adolescents' self-reports on the SAFE and SPAIC, as well as peer confederates' SIAS reports, with correlations in the moderate-magnitude range.

Links between UOs' SAFE Reports and Social Anxiety Status

We computed an independent samples t -test to compare the means of UOs' SAFE reports for adolescents scoring above ($n = 40$; $M = 87.35$; $SD = 18.60$) versus below ($n = 65$; $M = 75.58$; $SD = 17.29$) the clinical cut score on the SPAIC. This test revealed a significant difference ($t = 3.29$; $p < .01$), with a moderate-magnitude Cohen's d of 0.65.

Incremental Validity of UOs' SAFE Reports

Beyond our expectations of UOs' SAFE reports distinguishing adolescents on social anxiety status, we also expected these reports to display incremental validity in predicting a known associated feature of safety behaviors (i.e., social skills; Piccirillo et al., 2016), over-and-above adolescents' self-reported depressive symptoms and peer confederates' social anxiety reports. As mentioned previously, we tested this aim separately for each of these control variables. For self-reported depressive symptoms, in step 1, there was a significant effect of the BDI-II in relation to trained independent observers' social skills ratings ($\beta = -.25$; $\Delta R^2 = .06$; $p < .05$); an effect for which UOs' SAFE reports incrementally contributed a significant and moderate-magnitude effect in step 2 ($\beta = -.34$; $\Delta R^2 = .11$; $p < .001$). For peer confederates' reports, in step 1, there was a significant effect of the SIAS in relation to trained independent

observers' social skills ratings ($\beta = -.58$; $\Delta R^2 = .34$; $p < .001$); an effect for which UUOs' SAFE reports incrementally contributed a significant and moderate-magnitude effect in step 2 ($\beta = -.24$; $\Delta R^2 = .06$; $p < .01$). In both of these regression models, increased UUOs' reports about adolescents' safety behaviors related to decreased trained independent observers' ratings about adolescents' social skills. To facilitate interpretation of these effects, we report in online supplementary material regression plots that graphically depict these effects.

Criterion-Related Validity of UUOs' SAFE Reports

Our fifth hypothesis was that UUOs' SAFE reports would display criterion-related validity in predicting adolescents' self-reported state arousal in social interactions. In step 1 of the regression model testing this hypothesis, adolescents' self-reported state arousal at resting baseline explained significant variance in adolescents' self-reported arousal in social interactions ($\beta = .43$; $\Delta R^2 = .19$; $p < .001$). Unfamiliar untrained observers' SAFE reports explained significant variance in step 2 of the regression model predicting adolescents' self-reported state arousal in social interactions ($\beta = .30$; $\Delta R^2 = .09$; $p < .01$). In this regression model, increased UUOs' reports about adolescents' safety behaviors related to increased adolescents' self-reported state arousal in social interactions. To facilitate interpretation of this effect, we report in online supplementary material regression plots that graphically depict the effect.

Discussion

In this study, we extended previous literature on multi-informant approaches to assessing adolescent safety behaviors. We tested the psychometric properties of UUOs' reports of adolescent safety behaviors, based on their observations of adolescents when interacting in laboratory-controlled tasks with peer confederates. In a mixed-clinical/community sample of adolescents, we observed five findings. First, consistent with prior work with SAFE self-reports,

UUOs' SAFE reports demonstrated acceptable levels of internal consistency (i.e., $\alpha > .90$; Cuming et al., 2009; Qasmieh et al., 2018; Thomas et al., 2012). Second, UUOs' SAFE reports related to adolescent self-reports on the SAFE and on social anxiety measures, as well as peer confederates' reports of adolescent social anxiety. Third, consistent with recent work on adolescents' SAFE self-reports (Qasmieh et al., 2018), UUOs' SAFE reports distinguished adolescents who displayed clinically elevated social anxiety from those who did not. Fourth, UUOs' SAFE reports uniquely related to independent assessments of adolescent social skills, when taking into account adolescents' self-reports of depressive symptoms and peer confederates' reports of adolescent social anxiety. Fifth, in line with prior work on the SAFE (Qasmieh et al., 2018), UUOs' SAFE reports related to adolescents' subjective experiences with arousal within the social interactions from which UUOs based their SAFE reports.

Overall, our findings indicate that UUOs can provide psychometrically sound reports of adolescent safety behaviors, based on their observations of adolescents' social interactions with peer confederates. Yet, one key issue involves the degree to which UUOs can validly report about *all* safety behaviors. Indeed, some safety behaviors display in ways observers can readily detect (e.g., absence of eye contact) and others manifest "invisibly," such that the motivations or intentions that underlie enacting the behavior are opaque (e.g., use of baggy clothing to hide sweating, rehearsing lines to use in conversation before entering a social situation). Importantly, the SAFE includes items that assess both of these kinds of safety behaviors. Thus, future research should closely examine whether items on the SAFE vary in terms of UUOs' abilities to provide accurate responses. This work may reveal those safety behaviors that have the greatest influence on whether unfamiliar people in adolescents' social environments (e.g., same-age peers) decide to interact with the adolescents who display these behaviors.

Research and Theoretical Implications

Our findings have important implications for research on multi-informant approaches to mental health assessment broadly, and on clinical assessments of adolescent social anxiety in particular. Recent work indicates that, with limited interactions with adolescents, untrained peer confederates who directly interact with adolescents can make psychometrically sound reports about adolescent social anxiety (Deros et al., 2018). Our findings extend the psychometric rigor of reports about adolescents to UUOs who have no contact with the adolescents about whom they provide reports, and provide reports about processes related to social anxiety (i.e., safety behaviors). Further bolstering these findings is that these UUOs' SAFE reports significantly relate to ratings made by not only adolescents but also by those who directly interacted with the adolescent (i.e., peer confederates), and independent observers who received extensive training on how to rate behavior within these same interactions. These findings beg the question: What other processes that manifest within peer interactions might UUOs capably provide reports? Indeed, adolescents' exposures to aversive experiences with same-age peers portend the development and maintenance of social anxiety and a host of other mental health concerns, including attention-deficit/hyperactivity disorder, conduct problems, and depression (e.g., Cannon et al., 2020; Epkins & Heckler, 2011). In this respect, a key direction for future research involves testing the ability to leverage untrained informants to provide reports about these and other mental health domains linked to peer relations and interpersonal functioning.

Our findings also have important theoretical implications. Recent work on the active ingredients of exposure-based therapies for social anxiety is predicated on the notion that therapeutic exposures result in lasting benefits for clients, insofar as they have the "look and feel" of the social contexts where clients experience symptoms and impairments (Sewart &

Craske, 2020). Along these lines, recent work supports the idea that exposures designed to simulate adolescents' reactions to fear-provoking situations should focus on simulating social contexts germane to their clinical presentations, namely interactions with same-age, unfamiliar peers (Cannon et al., 2020; Deros et al., 2018; Glenn et al., 2019). Our findings build on these theoretical notions of exposure-based therapies and the need to simulate key elements of clients' experiences outside of therapy. That is, therapy should expose clients to *interaction partners* who share characteristics with those in their social environment (e.g., peer confederates). Additionally, therapy should gather assessments to estimate the *perspectives* that unfamiliar people might have of the adolescent, such as the degree to which they engage in safety behaviors. Indeed, in our study UOs' reports related to interpersonal domains that comprise common targets of exposure-based therapies for adolescents (e.g., improving social skills; Alfano & Beidel, 2011; Raggi et al., 2018). Further, these same safety behaviors, when displayed during therapeutic exposures, portend poor responses to exposure-based therapy (Hedtke et al., 2009; Piccirillo et al., 2016). In these respects, we recommend future theoretical work on why safety behaviors displayed during therapy impact not just outcomes immediately post-treatment, but pose risk for continued impairments in interpersonal functioning following treatment.

Clinical Implications

Our findings have important clinical implications for exposure-based therapies for adolescent social anxiety. Indeed, our findings support efforts to infuse therapeutic exposures with elements that facilitate the simulation of parameters of fear-provoking social situations present in adolescents' social environments. Recent work in implementation science calls for increased use of task-sharing models of service delivery, which involves integrating into key elements of care individuals who lack the extensive training of practitioners (Kazdin, 2017;

Singla et al., 2017). With regard to exposure-based therapies for adolescent social anxiety, our findings extend recent efforts to include relatively inexperienced personnel in exposure-based simulations of adolescents' interactions with same-age, unfamiliar peers (Cannon et al., 2020). Inclusion of these personnel seeks to address a key challenge with exposure-based therapy. Specifically, within a short period, trained professionals delivering care become *familiar* to the client. If the trained professional also “stands in” as the interaction partner—the person with whom a client participates within therapeutic exposures—their increased familiarity to the client logically results in exposures that lack the *unfamiliarity* element of fear-provoking situations in clients' social environments. Failing to induce the unfamiliarity that characterizes fear-provoking situations in “real life” poses risk that clients will experience what some call a “return of fear” following treatment (see Sewart & Craske, 2020). That is, to the degree that practitioners design exposures to provide clients with “teachable moments,” these exposures might inadequately prepare clients for situations outside of therapy if they fail to simulate the outside world.

Similarly, our findings point to the need to incorporate task-sharing models for service roles beyond interaction partners during exposures. Therapeutic exposures must also involve gathering estimates of how unfamiliar people perceive clients, particularly for domains like safety behaviors, which play a key role in interpersonal functioning. Our findings suggest that UOs provide psychometrically sound reports about adolescent safety behaviors. We encourage future work on whether use of these reports facilitates demonstrating to clients that use of safety behaviors impacts not only their social anxiety, but also how unfamiliar people perceive them. Relatedly, we encourage future work on whether use of UOs not only facilitates characterizing clients' interpersonal concerns, but also serves as a marker of long-term treatment benefits.

The clinical implications of our findings beg for questions regarding the clinical feasibility of the UOO approach. Indeed, service settings may vary as to the resources available to leverage the UOO approach as described in this paper and tested in our study. For instance, recruiting the personnel needed to gather UOO reports might only be possible within relatively large service settings that staff non-therapist personnel, such as clerical staff and research volunteers (e.g., community mental health centers, hospitals, medical centers, group practices). In this respect, we encourage future work that tests the feasibility of the UOO approach in these large service settings. Beyond the ability of large service settings to have available personnel to integrate into assessments as UOOs, recent work points to methods for service settings without these resources to make use of the approach. For example, for those service settings with the means for recording clients' exposures, use of video conferencing technology (e.g., Skype, Zoom) might allow for gathering reports from UOOs outside of the service setting. These technologies are now commonplace in current iterations of telehealth (see Comer & Myers, 2016). Leveraging these teleconferencing platforms might allow relatively small service settings (e.g., solo practitioners' offices) to partner with licensed practitioners in other service settings. Personnel across these settings might serve as UOOs to provide reports about each other's clients. Using networks of UOOs who provide ratings of adolescent clients may infuse feasibility in the UOO approach described in this paper.

Limitations

Three limitations to our study warrant comment. First, we observed relatively high levels of internal consistency for UOOs' SAFE reports, supporting one component of reliability. Further, this is an approach that has been taken to evaluate the reliability of other untrained raters' multi-item survey reports (e.g., peer confederates; Deros et al., 2018; Glenn et al., 2019).

Yet, the present design did not allow for the evaluation of other components of reliability, such as inter-rater reliability. We encourage future research testing the reliability of UOs' reports using components other than internal consistency.

Second, we leveraged UOs to simulate unfamiliar people's perceptions of adolescents' safety behaviors. As mentioned previously, a precedent exists in the literature for similar approaches yielding psychometrically sound data within assessments of adolescents' reactions to interactions with peer confederates (for a review, see Cannon et al., 2020). In fact, recent work indicates that untrained raters like our UOs provide accurate ratings of psychological phenomena that are unrelated to their personal characteristics (Borelli, Peng, Hong, Froidevaux, & Sbarra, 2019). At the same time, these UOs were not peers of adolescents selected out of adolescents' true social environments. Further, as research personnel in a laboratory focused on social anxiety, these UOs likely had relatively more experience observing or interacting with socially anxious adolescents, relative to unfamiliar people generally. Thus, these UOs may have been more attuned than unfamiliar people generally to the presence of processes linked to social anxiety, such as safety behaviors. Consequently, we encourage future work seeking to replicate and extend our findings to examine whether our findings generalize to UOs recruited outside of research settings.

Third, UOs rated adolescent safety behaviors based on adolescents' interactions with research personnel who we trained to simulate unfamiliar same-age peers. This approach is consistent with prior work with adolescents (Deros et al., 2018; Qasmieh et al., 2018; Rausch et al., 2017), particularly in terms of our focus on training confederates who could reasonably appear to adolescents as same-age, unfamiliar peers. Further, adolescents' reactions to interacting with these peer confederates predict their reactions to independent tasks where they are instructed to

interact with same-age peers (see Karp et al., 2018). Yet, our peer confederates were older than our adolescent participants. Further, although we masked UOs to key characteristics of the adolescents about whom they provided reports (see online supplementary material), UOs were not masked to the identity of the research personnel who served as peer confederates. Future work should test UOs' ratings of adolescents within interactions with age-matched peers.

Concluding Comments

Our findings build on an emerging body of work on multi-informant approaches to assessing processes relevant to understanding adolescent social anxiety. In particular, our findings elucidate the potential to leverage UOs to simulate unfamiliar people's perceptions of adolescent safety behaviors. Broadly, our findings provide evidence of the validity and psychometric utility of leveraging reports from UOs after viewing adolescents engaging in a brief set of social interaction tasks. Our findings suggest that UOs can reliably and validly report about adolescents' use of safety behaviors, as measured by the SAFE, in ways that incrementally contribute useful information in relation to clinically-relevant indices. Specifically, UOs' reports distinguished adolescents on social anxiety status and displayed incremental validity in predicting adolescents' social skills, beyond other informants' psychometrically sound reports. These findings have important implications for optimizing multi-informant approaches to the assessment of adolescent social anxiety and related processes. Specifically, we encourage future work to utilize UOs' reports in multi-informant paradigms to better understand and characterize adolescents' safety behaviors and their associations with other constructs relevant to the development, maintenance, and treatment of adolescent social anxiety. This approach to gathering unfamiliar people's impressions of safety behaviors may improve our understanding of how and why safety behaviors impact the social functioning of adolescents who display them.

References

- Alden, L. E. & Bieling, P. (1998). Interpersonal consequences of the pursuit of safety. *Behaviour Research and Therapy*, 36(1), 53-64. [https://doi.org/10.1016/S0005-7967\(97\)00072-7](https://doi.org/10.1016/S0005-7967(97)00072-7)
- Alfano, C. A., & Beidel, D. C. (2011). *Social anxiety in adolescents and young adults: Translating developmental science into practice*. Washington, D. C.: American Psychological Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory—Second Edition manual*. San Antonio, TX: The Psychological Corporation.
- Beidel, D. C., Turner, S. M. & Morris, T. L. (1995). A new inventory to assess childhood social anxiety and phobia: The Social Phobia and Anxiety Inventory for Children. *Psychological Assessment*, 7(1), 73-79. <https://doi.org/10.1037/1040-3590.7.1.73>
- Borelli, J. L., Peng, X., Hong, K., Froidevaux, N. M., & Sbarra, D. A. (2019). Accuracy and confidence in perceptions of targets' attachment to former partners: Do judges vary as a function of individual differences in attachment orientation? *Journal of Research in Personality*, 79, 1-12. <https://doi.org/10.1016/j.jrp.2019.01.001>
- Cannon, C. J., Makol, B. A., Keeley, L. M., Qasmieh, N., Okuno, H., Racz, S. J., & De Los Reyes, A. (2020). A paradigm for understanding adolescent social anxiety with unfamiliar peers: Conceptual foundations and directions for future research. *Clinical Child and Family Psychology Review*, 23(3), 338-364. <https://doi.org/10.1007/s10567-020-00314-4>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and

- standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Comer, J. S., & Myers, K. (2016). Future directions in the use of telemental health to improve the accessibility and quality of children's mental health services. *Journal of Child and Adolescent Psychopharmacology*, 26(3), 296-300. <https://doi.org/10.1089/cap.2015.0079>
- Cuming, S., Rapee, R. M., Kemp, N., Abbott, M. J., Peters, L., & Gaston, J. E. (2009). A self-report measure of subtle avoidance and safety behaviors relevant to social anxiety: Development and psychometric properties. *Journal of Anxiety Disorders*, 23(7), 879-883. <https://doi.org/10.1016/j.janxdis.2009.05.002>
- De Los Reyes, A. (2020, February 15). Materials described in the paper, "A paradigm for understanding adolescent social anxiety with unfamiliar peers: Conceptual foundations and directions for future research." Retrieved from <https://bit.ly/CCFPRPeerParadigmOSF>
- De Los Reyes, A., & Makol, B.A. (2019). Evidence-based assessment. In T.H. Ollendick, L. Farrell, and P. Muris (Eds.), *Innovations in CBT for childhood anxiety, OCD, and PTSD: Improving access and outcomes* (pp. 28-51). New York: Cambridge.
- Deros, D. E., Racz, S. J., Lipton, M. F., Augenstein, T. M., Karp, J. N., Keeley, L. M., ... De Los Reyes, A. (2018). Multi-informant assessments of adolescent social anxiety: Adding clarity by leveraging reports from unfamiliar peer confederates. *Behavior Therapy*, 49(1), 84-98. <https://doi.org/10.1016/j.beth.2017.05.001>
- Epkins, C. C., & Heckler, D. R. (2011). Integrating etiological models of social anxiety and depression in youth: Evidence for a cumulative interpersonal risk model. *Clinical Child and*

Family Psychology Review, 14(4), 329-376. <https://doi.org/10.1007/s10567-011-0101-8>

- Glenn, L. E., Keeley, L. M., Szollos, S., Okuno, H., Wang, X., Rausch, E., ... De Los Reyes, A. (2019). Trained observers' ratings of adolescents' social anxiety and social skills within controlled, cross-contextual social interactions with unfamiliar peer confederates. *Journal of Psychopathology and Behavioral Assessment*, 41(1), 1-15. <https://doi.org/10.1007/s10862-018-9676-4>
- Hedtke, K. A., Kendall, P. C., & Tiwari, S. (2009). Safety seeking and coping behavior during exposure tasks with anxious youth. *Journal of Clinical Child and Adolescent Psychology*, 38(1), 1-15. <https://doi.org/10.1080/15374410802581055>
- Hofmann, S. G., Albano, A. M., Heimberg, R. G., Tracey, S., Chorpita, B. F., & Barlow, D. H. (1999). Subtypes of social phobia in adolescents. *Depression and Anxiety*, 9(1), 15-18. [https://doi.org/10.1002/\(SICI\)1520-6394\(1999\)9:1<15::AID-DA2>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1520-6394(1999)9:1<15::AID-DA2>3.0.CO;2-6)
- Karp, J. N., Makol, B. A., Keeley, L. M., Qasmieh, N., Deros, D. E., Weeks, J. W., ... De Los Reyes, A. (2018). Convergent, incremental, and criterion-related validity of multi-informant assessments of adolescents' fears of negative and positive evaluation. *Clinical Psychology and Psychotherapy*, 25(2), 217-230. <https://doi.org/10.1002/cpp.2156>
- Kazdin, A. E. (2017). Addressing the treatment gap: A key challenge for extending evidence-based psychosocial interventions. *Behaviour Research and Therapy*, 88, 7-18. <https://doi.org/10.1016/j.brat.2016.06.004>
- 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. <https://doi.org/10.1001/archpsyc.62.6.593>
- Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. U. (2012).

- Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, 21(3), 169-184. <https://doi.org/10.1002/mpr.1359>
- Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119-137). Norwood, NJ: Ablex.
- McManus, F., Sacadura, C., & Clark, D. M. (2008). Why social anxiety persists: An experimental investigation of the role of safety behaviors as a maintaining factor. *Journal of Behavior Therapy and Experimental Psychiatry*, 39(2), 147-161. <https://doi.org/10.1016/j.jbtep.2006.12.002>
- Morrison, A. S., & Heimberg, R. G. (2013). Social anxiety and social anxiety disorder. *Annual Review of Clinical Psychology*, 9, 249-274. <https://doi.org/10.1146/annurev-clinpsy-050212-185631>
- Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, 25, 267-316. <https://doi.org/10.2307/271070>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill
- Piccirillo, M. L., Dryman, M. T., & Heimberg, R. G. (2016). Safety behaviors in adults with social anxiety: Review and future directions. *Behavior Therapy*, 47(5), 675-687. <https://doi.org/10.1016/j.beth.2015.11.005>
- Qasmieh, N., Makol, B. A., Augenstein, T. M., Lipton, M. F., Deros, D. E., Karp, J. N., ... De Los Reyes, A. (2018). A multi-informant approach to assessing safety behaviors among adolescents: Psychometric properties of the Subtle Avoidance Frequency Examination. *Journal of Child and Family Studies*, 27(6), 1830-1843. <https://doi.org/10.1007/s10826-018-1040-5>
- Raggi, V. L., Samson, J. G., Felton, J. W., Loffredo, H. R., & Berghorst, L. H. (2018). *Exposure therapy for treating anxiety in children and adolescents: A comprehensive guide*. New

Harbinger Publications.

- Rapee, R. M. & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, 35(8), 741-756. [https://doi.org/10.1016/S0005-7967\(97\)00022-3](https://doi.org/10.1016/S0005-7967(97)00022-3)
- Rausch, E., Racz, S. J., Augenstein, T. M., Keeley, L., Lipton, M. F., & Szollos, S., ... De Los Reyes (2017). A multi-informant approach to measuring depressive symptoms in clinical assessments of adolescent social anxiety using the Beck Depression Inventory-II: Convergent, incremental, and criterion-related validity. *Child and Youth Care Forum*, 46(5), 661-683. <https://doi.org/10.1007/s10566-017-9403-4>
- Rowa, K., Paulitzki, J. R., Ierullo, M. D., Chiang, B., Antony, M. M., McCabe, R. E., & Moscovitch, D. A. (2015). A false sense of security: Safety behaviors erode objective speech performance in individuals with social anxiety disorder. *Behavior Therapy*, 46(3), 304-314. <https://doi.org/10.1016/j.beth.2014.11.004>
- Salkovskis, P. M. (1991). The importance of behaviour in the maintenance of anxiety and panic: a cognitive account. *Behavioural and Cognitive Psychotherapy*, 19(1), 6-19. <https://doi.org/10.1017/S0141347300011472>
- Salkovskis, P. M., Clark, D. M. & Gelder, M. G. (1996). Cognition-behaviour links in the persistence of panic. *Behaviour Research and Therapy*, 34(5-6), 453-458. [https://doi.org/10.1016/0005-7967\(95\)00083-6](https://doi.org/10.1016/0005-7967(95)00083-6)
- Sewart, A. R., & Craske, M. G. (2020). Inhibitory learning. In J. S. Abramowitz & S. M. Blakey (Eds.), *Clinical handbook of fear and anxiety: Maintenance processes and treatment mechanisms* (pp. 265-285). Washington: American Psychological Association.
- Singla, D. R., Kohrt, B. A., Murray, L. K., Anand, A., Chorpita, B. F., & Patel, V. (2017). Psychological treatments for the world: Lessons from low-and middle-income

countries. *Annual Review of Clinical Psychology*, 13, 149-181.

<https://doi.org/10.1146/annurev-clinpsy-032816-045217>

Slepian, M. L., Bogart, K. R., & Ambady, N. (2014). Thin-slice judgments in the clinical context. *Annual Review of Clinical Psychology*, 10, 131-153.

<https://doi.org/10.1146/annurev-clinpsy-090413-123522>

Stangier, U., Heidenreich, T., & Schermelleh-Engel, K. (2006). Safety behaviors and social performance in patients with generalized social phobia. *Journal of Cognitive Psychotherapy*, 20(1), 17-31. <https://doi.org/10.1891/jcop.20.1.17>

Stein, M. B. & Stein, D. J. (2008). Social anxiety disorder. *The Lancet*, 371(9618), 1115-1125.

[https://doi.org/10.1016/S0140-6736\(08\)60488-2](https://doi.org/10.1016/S0140-6736(08)60488-2)

Thomas, S. A., Daruwala, S. E., Goepel, K. A., & De Los Reyes, A. (2012). Using the Subtle Avoidance Frequency Examination in adolescent social anxiety assessments. *Child and Youth Care Forum*, 41(6), 547-559. <https://doi.org/10.1007/s10566-012-9181-y>

Thwaites, R. & Freeston, M. (2005). Safety-seeking behaviours: Fact or function? How can we clinically differentiate between safety behaviours and adaptive coping strategies across anxiety disorders? *Behavioural and Cognitive Psychotherapy*, 33(2), 177-188.

<https://doi.org/10.1017/S1352465804001985>

U.S. Census Bureau. (2016). State and county quickfacts: Washington, DC. Retrieved from <https://www.census.gov/quickfacts/DC>

Table 1

Correlations among Study Measures

Variable	1	2	3	4	5	6	7	8
1 Subtle Avoidance Frequency Examination (UUO)		.25*	.36***	.26**	.20*	-.37***	.22*	.38***
2 Subtle Avoidance Frequency Examination (A)			.79***	.34**	.58***	-.39***	.38***	.50***
3 Social Phobia and Anxiety Inventory for Children (A)				.42***	.59***	-.41***	.48***	.67***
4 Social Interaction Anxiety Inventory (PC) ^a					.34***	-.58***	.22*	.44***
5 Beck Depression Inventory-II (A)-Square Root Transformed						-.25*	.35***	.38***
6 Adolescent Social Skills (TIO)							-.23*	-.38***
7 Self Assessment Manikin, Baseline (A)								.43***
8 Self Assessment Manikin, Tasks (A)								

Note. UUO = unfamiliar untrained observer; A = adolescent; PC = peer confederate; TIO = trained independent observer. ^a Due to peer confederates not providing reports for two adolescent participants, estimates reported for peer confederates based on data from 103 peer confederates. * $p < .05$; ** $p < .01$; *** $p < .001$.