

Community-based validation of the Social Phobia Screener (SOPHS)

Philip J. Batterham, National Institute for Mental Health Research, Research School of Population Health, The Australian National University, Canberra, Australia

Andrew J. Mackinnon, Orygen The National Centre of Excellence in Youth Mental Health, The University of Melbourne, Melbourne, Australia

Helen Christensen, Black Dog Institute, The University of New South Wales, Sydney, Australia

Corresponding author:

Philip Batterham
National Institute for Mental Health Research
Research School of Population Health
63 Eggleston Road
The Australian National University
Acton ACT 2601 AUSTRALIA
Tel.: +61 2 61251031
Fax: +61 2 61250733
Email: philip.batterham@anu.edu.au

Running title: Social Phobia Screener

Word count: 5,039

Acknowledgements: The data described were collected from two studies funded by National Health and Medical Research Council (NHMRC) Project Grants 525419 and 1043952. PB is supported NHMRC Career Development Fellowship 1083311. HC is supported by NHMRC John Cade Fellowship 1056964. Competing interests: the authors have no competing interests.

Abstract

There is a need for brief, accurate screening scales for social anxiety disorder, to enable better identification of the disorder in research and clinical settings. A 5-item social anxiety screener, the Social Phobia Screener (SOPHS), was developed to address this need. The screener was validated in two samples: (1) 12,292 Australian young adults screened for a clinical trial, including 1,687 participants who completed a phone-based clinical interview; and, (2) 4,214 population-based Australian adults recruited online. The SOPHS (78% sensitivity, 72% specificity) was found to have comparable screening performance to the SPIN (77% sensitivity, 71% specificity) and Mini-SPIN (74% sensitivity, 73% specificity) relative to clinical criteria in the trial sample. In the population-based sample, the SOPHS was also accurate (95% sensitivity, 73% specificity) in identifying DSM-5 social anxiety disorder. The SOPHS is a valid and reliable screener for social anxiety that is freely available for use in research and clinical settings.

Key words: social anxiety disorder, social phobia, anxiety, screening, measurement

Social anxiety disorder, also known as social phobia, is a highly prevalent disorder with considerable burden of disease (Kessler, Berglund, et al., 2005; Patel, Knapp, Henderson, & Baldwin, 2002). There are a number of scales designed to assess social anxiety disorder (see Clark et al., 1997; Connor et al., 2000). However, not all of these have been validated for self-report or against clinical criteria, and very few are brief screeners. Given that social anxiety disorder is treatable yet frequently goes unrecognized and untreated (Bruce & Saeed, 1999), the development of brief, accurate screening tools could facilitate redressing this situation. Furthermore, brief screening scales are highly advantageous in population-based research, where extensive questionnaires measuring a broad array of risk factors and outcomes require brief scales to reduce burden on respondents. The response burden of a scale is a particularly important consideration in screening, particularly in online settings where participant engagement is challenging (Christensen, Griffiths, & Farrer, 2009) and in primary care where resource constraints demand brevity (Kessler et al., 2002; Williams, Pignone, Ramirez, & Perez Stellato, 2002). The requirement for clinician or interviewer assessment also tends to be unfeasible in these settings, hence the need for valid self-report scales. Another barrier to use of scales is the lack of universal access to scales. Many scales require permission for use or are not publicly available. Their use may therefore be constrained in research and clinical settings.

We developed a new screening scale, the Social Phobia Screener (SOPHS), a 5-item self-report scale assessing social anxiety for use in research, online and clinical settings (Batterham et al., 2013). The scale was developed to provide brief, accurate case finding for social anxiety disorder based on DSM-IV criteria. However, case finding neglects to account for the severity of disorder or subclinical symptoms, which may be significantly debilitating. Therefore the scale was designed with the dual purpose of case finding and assessing symptom severity. The five items of the scale were chosen to briefly encapsulate symptom

severity on each of the core DSM-IV criteria for social anxiety disorder. It was important that assessment of functional impairment was incorporated into the scale, as existing social anxiety screeners do not assess impact on functioning. The scale uses a vignette to describe the typical symptoms of social anxiety, adapting an approach used previously to screen for mental disorders (Christensen, Batterham, Grant, Griffiths, & Mackinnon, 2011). This introductory vignette is followed by items covering fear/embarrassment of social situations, the degree to which the fear was excessive, avoidance of social situations, suffering in social situations and functional impairment due to social anxiety. The scale has been briefly described by Batterham et al. (2013) but the process and rationale for developing the screener were not detailed in that study, and the scale has not previously been validated against clinical criteria.

In the current study, the scale was validated against clinical criteria (Mini International Neuropsychiatric Interview; MINI; Sheehan et al., 1998) and compared to an existing scale, the Social Phobia Inventory (SPIN; Connor et al., 2000). The SPIN is a well-validated self-report scale for assessing social anxiety, commonly used in research trials. However, the SPIN includes 17 items and requires a fee for use. The short form of the SPIN, the Mini-SPIN, only contains three items so is more suitable for use as a screener (Connor, Kobak, Churchill, Katzelnick, & Davidson, 2001), although the performance of the Mini-SPIN at prescribed cut points has not always been found to be optimal (Weeks, Spokas, & Heimberg, 2007). The subset of items constituting the Mini-SPIN was used as an additional point of comparison. Administration of the Mini-SPIN requires permission from the authors. To our knowledge, no other brief screeners for social anxiety disorder are currently available.

This examination of the SOPHS was conducted within the context of a community-based survey, designed to identify young adult participants eligible for an online indicated prevention trial for Generalized Anxiety Disorder (GAD). The current study sought to

determine whether the psychometric properties and screening performance of the SOPHS were comparable to the SPIN and superior to the Mini-SPIN using a diagnostic interview as the gold standard. In addition, factors associated with social anxiety symptoms were identified. An independent population-based sample was then used to test whether the SOPHS was accurate in screening DSM-5 criteria for social anxiety disorder. These criteria differ from DSM-IV only in terms of the duration of anxiety symptoms: “typically lasting 6 or more months” (DSM-5) compared to “the duration is at least 6 months” (DSM-IV).

Method

Participants and procedure

Data for this study were taken from the screening phase of the iChill study, an intervention trial testing an online cognitive behavioural therapy intervention to prevent new cases of Generalized Anxiety Disorder. The trial itself is described by Christensen et al. (2014; 2010) and registered in Current Controlled Trials (controlled-trials.com, ISRCTN76298775). Invitations and screening surveys were mailed to 120,000 18-30 year olds selected randomly from the Australian electoral roll, with 12,430 (10.4%) responding by returning the paper-based screening survey with informed consent completed. Registration on the electoral roll and voting are compulsory in Australia. Data for both the SOPHS and the SPIN were available for 12,292 (98.9%) of respondents. The survey data were collected between May 2010 and June 2012. The age of the cohort was chosen to cover the peak period in which anxiety disorders emerge (Kessler, Berglund, et al., 2005).

Of those who returned a survey, 4,205 (33.8%) met eligibility criteria for the intervention trial, with the remainder excluded on the basis of one or more of the criteria shown in Figure 1. Of those eligible, 2,791 (66.4%) consented to consider participating in the trial, with 2,517 (59.9%) also providing sufficient information for follow-up (phone and e-

mail address). Attempts were made to contact these participants by telephone to complete a MINI interview (version 5.0; Sheehan et al., 1998) within 21 days of survey return. Twelve postgraduate clinical psychology students enrolled in a doctoral program administered the telephone MINI interviews, and were blind to the participants' responses to the survey. Interviewers were given four hours of training in administration of the MINI, including practice interviews, with oversight from a clinical psychologist and the research team. The MINI interviews included the modules on panic (lifetime, current), social anxiety disorder (current), depression (current) and generalised anxiety disorder (lifetime, current). Not all participants could be contacted within the time period, despite up to seven call attempts being made, resulting in a final sample of 1,687 (67.0% of eligible and consenting) participants with survey data and MINI clinical data available. The study was approved by the ANU Human Ethics Committee (Protocol 2008/548).

An independent adult sample was also used to test the SOPHS against DSM-5 criteria. A community-based cross-sectional survey was administered to 4,214 Australian adults aged ≥ 18 years, recruited using advertisements on the social networking website Facebook. The target population of Facebook users aged ≥ 18 years was 8.8 million, representing approximately 45% of the total Australian population aged ≥ 18 . From August to December 2014, a series of advertisements was placed on Facebook with the wording: "Assessing Mental Health Survey: Participate in a study examining your mental health by completing a 40-minute survey now". A comprehensive mental health survey was administered online using the LimeSurvey platform, with data stored on a secure server at the Australian National University (ANU), Canberra. This study had ethics approval from the ANU Human Research Ethics Committee (Protocol 2013/509). Participants received no compensation for participating in either study.

Measures

The SOPHS is 5-item screener for social anxiety, with items first described by Batterham et al (2013). The scale can be accessed at <http://www.nimhr.anu.edu.au/mental-health-measures/social-phobia-screener-sophs>. Although DSM-IV and DSM-5 criteria for social anxiety disorder require persistence over at least six months, the screener used a one month time frame to minimise recall biases. The introduction to the scale reads: “Some people feel fearful or embarrassed in social situations, such as: speaking in public, eating in public or with others, writing while someone watches, blushing or looking foolish, attending parties or meetings, speaking with someone in authority, being criticised by others”, followed by the five items:

- (1) “To what extent have you felt fearful or embarrassed of any social situations during the past month?”,
- (2) “Was the fear or embarrassment you experienced during the past month excessive or unreasonable?”,
- (3) “During the past month, have you avoided any social situations because of your fear or embarrassment?”,
- (4) “During the past month, how much have you suffered through any social situations because of your fear or embarrassment?”, and,
- (5) “During the past month, how much has your work, home or social life been disrupted because of your fear or embarrassment?”.

Each item is rated on a 5-point Likert scale: not at all (0), a little (1), moderately (2), severely (3), extremely (4). Severity scores are obtained by summing responses to all items, with scores ranging from 0-20. Respondents who respond “not at all” to the first item do not complete subsequent items as these items are not relevant, and are consequently scored zero.

Due to this gating process, 5,787 participants (46.8%) were only required to answer a single question, while the remainder were required to answer all five items.

The SPIN has 17 items rated from 0 “not at all” to 4 “extremely”, with scores ranging from 0-68. A cut-off of 19 has been identified to distinguish between individuals with and without diagnosis of social anxiety disorder with sensitivity of 72.5% and specificity of 84.3% (Connor et al., 2000), although in young people a cut-off of 24 has been found to be more appropriate (Ranta, Kaltiala-Heino, Rantanen, Tuomisto, & Marttunen, 2007). The Mini-SPIN is a short form screener using three items from the full version of the SPIN to screen for social anxiety, with a range of 0-12. Scores of six or greater have previously been reported to identify social anxiety with sensitivity of 88.7% and specificity of 90.0% relative to SCID criteria (Connor et al., 2001), although other studies have adopted a cut-off of seven (Seeley-Wait, Abbott, & Rapee, 2009).

The clinical comparator for the main study was the MINI, which has previously been demonstrated to accurate in assessing criteria for social anxiety disorder, panic disorder and Generalized Anxiety Disorder (Sheehan et al., 1998). In the follow-up population study to test the SOPHS against DSM-5 criteria, an exhaustive checklist of DSM-5 symptoms was assessed based on self-report. The checklist was developed by the authors along similar principles to the electronic version of the MINI (Zbozinek et al., 2012) in terms of structure (binary and categorical self-report items with conditional skip logic) and response burden. However, the checklist was developed independently from the MINI, is non-proprietary and is based on DSM-5 rather than DSM-IV criteria. The checklist is included in Appendix 1. The checklist had high internal consistency (Cronbach $\alpha = 0.98$). Scales used to assess the discriminant validity of the SOPHS were anxiety scores on the GAD-7 (Spitzer, Kroenke, Williams, & Lowe, 2006) and depression scores on the PHQ-9 (Spitzer, Kroenke, & Williams, 1999). These scales assess symptoms of generalized anxiety disorder and

depression, with scores ranging from 0-21 and 0-27 respectively and with high accuracy in detecting these disorders (Spitzer et al., 1999; Spitzer et al., 2006).

Analysis

The internal consistency of the SOPHS was assessed in the trial sample, followed by examination of its convergent validity of the SOPHS severity score relative to scores on the SPIN and Mini-SPIN. A two-parameter graded response model was estimated to assess the discrimination of the items across the latent continuum of social anxiety symptoms. This model estimates Item Response Theory (IRT) parameters from items with ordinal responses derived from a unidimensional CFA with a full information robust maximum likelihood estimator and a logit link function. This method relies on the formal relationship between IRT and ordinal item factor analysis to transform item factor loadings and item response category thresholds to the respective IRT slope (discrimination) and threshold (location) parameters on underlying severity dimension tapped by the items (Muthén & Muthén, 2010; Takane & De Leeuw, 1987). The accuracy (sensitivity and specificity) of the SOPHS in identifying participants who met social anxiety disorder criteria based on the MINI assessment was compared to the accuracy of the SPIN and Mini-SPIN. The area under the Receiver Operating Characteristic curve was calculated for each scale based on severity scores. Discriminant validity was also assessed by comparing SOPHS severity scores with caseness for panic disorder and generalised anxiety disorder (GAD), anxiety scores on the GAD-7 (Spitzer et al., 2006) and depression scores on the PHQ-9 (Spitzer et al., 1999). Finally, the sensitivity and specificity of the SOPHS in detecting DSM-5 social anxiety disorder at various cut points were estimated in the independent population sample. Analyses were primarily conducted in SPSS v20 (IBM Corp, Chicago Ill), although IRT analyses were conducted using Mplus version 6.1 (Muthén and Muthén, 2013). Confidence intervals for

sensitivity and specificity estimates were calculated using the software and methods described by Mackinnon (2000), and areas under the Receiver Operating Characteristic curves were compared using SAS Enterprise Guide v4.3 (SAS Institute Inc, Cary NC).

Results

Characteristics of the trial sample are shown in Table 1, separating those who only participated in the survey from those who received a MINI clinical assessment. Compared to the general population (Australian Bureau of Statistics, 2014), there was an overrepresentation of females responding to the survey, which was significantly magnified among those who received a clinical assessment. Those who received a clinical assessment were also significantly younger, less likely to be married, more likely to be in a de facto relationship, and more likely to speak English at home than those who did not. Due to the requirement of a GAD-7 score of 5 or higher to be eligible for a clinical assessment, participants who received a clinical interview had significantly higher levels of psychological symptoms including social anxiety (SOPHS and SPIN), generalized anxiety disorder (GAD-7), depression (PHQ-9) and suicidal ideation (last item of PHQ-9). However, the GAD-7 criterion was only indicative of elevated symptoms (46th percentile in the present sample) rather than presence of GAD. Only 16% of participants who met the GAD-7 cut-off and were given a MINI interview met caseness for GAD.

Excluding participants who scored zero on the SOPHS for whom consistency was guaranteed by the administration procedure ($n = 6,866$), the internal consistency was high (Cronbach $\alpha = 0.89$), similar to that of the SPIN ($\alpha = 0.92$) and superior to the Mini-SPIN ($\alpha = 0.79$). The SOPHS severity score had high convergent validity based on significant correlations with the SPIN severity score ($r = 0.73$) and with the Mini-SPIN score ($r = 0.66$). Moderate significant correlations were found between SOPHS severity and the GAD-7 ($r =$

0.47) and PHQ-9 ($r = 0.55$), similar to correlations between SPIN scores and these scales (0.46, 0.54 respectively), with these weaker effects ($Z = -31.6$ and $Z = -40.0$ respectively, $p < 0.001$) indicating discriminant validity.

Factor analysis of the SOPHS using principal axis factoring extraction yielded a single factor with eigenvalue 3.6 accounting for 72.9% of total variance, with additional factors having eigenvalues < 0.34 , indicating a strongly unidimensional structure. All items loaded on the single factor at 0.77 or higher, with loadings also displayed in Table 2. Also shown in Table 2 are IRT parameters from the two-parameter graded response model. The model had excellent fit, based on Comparative Fit Index of 0.998, Tucker-Lewis Fit of 0.997, and root mean square error of approximation of 0.037 (Hu & Bentler, 1998). The factor loadings and slopes in Table 2 reflect the discrimination of each item and were all significant, ranging from 2.9 to 4.5. By comparison, discrimination estimates for the three Mini-SPIN items ranged from 2.0 to 2.7. Test information was calculated for the SOPHS, SPIN and Mini-SPIN as a function of location on the severity dimension (θ). These IRT-derived curves show the precision with which the underlying construct is measured as a function of dimension severity. The test information curves in Figure 2 indicate that the SOPHS showed high precision across a range of severities between -1 and 3 standard deviations from the mean. In terms of total information, it performed similarly to the full SPIN and provided more information than the Mini-SPIN between -0.5 and 3 standard deviations from the mean. This is a critical region for individuals at risk of social anxiety disorder. Scales designed as screens should have highest precision at the location of likely cut points (Lord, 1980).

Relationships between sample characteristics and SOPHS severity scores are shown in Table 3. The severity of social anxiety symptoms was higher in younger, female, unmarried, lower educated, unemployed, anxious or depressed participants, and participants

reporting recent suicidal ideation. There was no difference in level of symptoms based on language spoken at home.

Comparison to DSM-IV criteria in the trial sample

Based on the MINI assessment, 163 (9.7%) of the participants who were assessed met clinical criteria for social anxiety disorder. Prevalence of social anxiety disorder was higher among lower-educated participants [$\chi^2 (2) = 12.4, p = 0.002$] and those in part-time/casual work or unemployed [$\chi^2 (4) = 16.2, p = 0.003$]. There were no differences in prevalence by gender, age group, marital status or language group.

Among participants who were clinically screened, 562 (33.6%) of the assessed participants met SPIN criteria of ≥ 24 (Ranta et al., 2007), with 436 (26.0%) false positives and 37 false negatives (2.2%). Using the cut-off of 6 on the Mini-SPIN, 534 (31.9%) screened positive, with 413 (24.7%) false positives and 42 false negatives (2.5%). For the SOPHS, the criterion for a positive screen was based on responding “a little” or higher to items 1, 2 and 5, and either item 3 or 4. This criterion was found to perform most strongly against MINI clinical criteria. Using this criterion, 559 (33.4%) screened positive for social anxiety disorder on the SOPHS, with 432 (25.8%) false positives and 36 false negatives (2.2%). Individuals who met clinical criteria had mean SOPHS scores of 8.9 (sd = 4.9), significantly higher than those who did not meet clinical criteria (mean = 3.4, sd = 3.9, $t_{184.8} = 13.8, p < 0.001$). Seven (4%) of those participants who met social anxiety disorder criteria on the MINI skipped items from the SOPHS based on a response of “not at all” to the first item. None of these seven participants met Mini-SPIN criteria (scores all < 5), and only two of the seven met SPIN criteria, suggesting that the skip-out did not adversely impact on the accuracy of the SOPHS. Removing the non-gating items from the SOPHS resulted in

significantly reduced sensitivity (items 3 and 4) or specificity (items 2 and 5), suggesting that all items contributed to the accuracy of the screener.

The sensitivity, specificity, positive predictive value and negative predictive value of the SOPHS, SPIN and Mini-SPIN, relative to MINI clinical criteria, are shown in Table 4, based on prescribed cut points. McNemar's test demonstrated no significant difference between the SOPHS and the SPIN in terms of sensitivity [$\chi^2 (1) = 0.03, p = 0.862$] or specificity [$\chi^2 (1) = 0.05, p = 0.825$] at prescribed cut-offs (Ranta et al., 2007). Using the original cut-off for the SPIN of 19 (Connor et al., 2000) resulted in inadequate specificity (57.9%), so the cut-off recommended for younger populations (Ranta et al., 2007) was retained. There were also no significant differences between the sensitivity or specificity of the SOPHS and Mini-SPIN using the prescribed cut point of 6 (Connor et al., 2001). Adopting the 7 cut-off often favoured for the Mini-SPIN (Seeley-Wait et al., 2009) resulted in significantly poorer sensitivity [65.0%, $\chi^2 (1) = 9.80, p < 0.001$] but greater specificity [79.8%, $\chi^2 (1) = 9.80, p < 0.001$] for the Mini-SPIN compared to the SOPHS. The areas under the Receiver Operating Characteristic curves (and 95% confidence intervals) for the SOPHS, SPIN and Mini-SPIN were 0.81 (0.79-0.86), 0.83 (0.78-0.85) and 0.80 (0.76-0.83) respectively, with no significant difference between the area estimates for the SPIN and SOPHS [$\chi^2 (1) = 0.75, p = 0.39$] or the Mini-SPIN and SOPHS [$\chi^2 (1) = 1.07, p = 0.30$].

Discriminant validity was further assessed by examining accuracy of the SOPHS for identifying current panic disorder, GAD and depression, as assessed by the MINI. The SOPHS had poor sensitivity (panic: 0.55, GAD: 0.60, depression: 0.58) and reduced specificity (panic: 0.68, GAD: 0.72, depression: 0.70) for all disorders. These findings indicate of discriminant validity, despite high levels of comorbidity between the disorders: among the 163 participants meeting criteria for social anxiety disorder, 18.4% also met MINI criteria for panic disorder, 45.4% for GAD and 42.3% for depression.

Within subgroups of the sample, sensitivity and specificity remained acceptable for males (sensitivity 0.75, specificity 0.67) and females (sensitivity 0.79, specificity 0.73), and among those with (sensitivity 0.79, specificity 0.72) and without (sensitivity 0.76, specificity 0.69) university education. However, there was some indication of loss of accuracy among the 291 participants who spoke a language other than English at home despite all participants understanding written English (sensitivity 0.58, specificity 0.67), compared to those who spoke only English (sensitivity 0.82, specificity 0.72). This loss of accuracy among participants who spoke a language other than English was also seen for the SPIN (sensitivity 0.77, specificity 0.52) and Mini-SPIN (sensitivity 0.58, specificity 0.68).

Comparison to DSM-5 criteria in an independent population-based sample

Finally, the accuracy of the SOPHS against DSM-5 criteria was tested in an independent population-based sample (N = 4,214), which included 79% females, 22% aged ≤ 35 years, 36% aged > 55 years, and 17% (n = 735) meeting criteria for social anxiety disorder. The item-based criterion used in the trial sample also performed well against DSM-5 criteria in this sample, with sensitivity of 0.95 (95% CI: 0.94-0.97) and specificity of 0.73 (95% CI: 0.71-0.74). In addition, it was found that a cut point of ≥ 8 on the SOPHS could also be used in a population sample to accurately identify individuals meeting social anxiety criteria (based on Youden indices), as shown in Table 5. To account for overrepresentation of females and people with a mental disorder, analyses were repeated using data weighted to reflect population estimates of disorder prevalence, age and gender (Australian Bureau of Statistics, 2014; Slade, Johnston, Oakley Browne, Andrews, & Whiteford, 2009). Using population weighted data, the SOPHS remained an accurate screener based on the item-based criterion (sensitivity 0.95, specificity 0.77) or using the cut point of ≥ 8 (sensitivity 0.90, specificity 0.88). Cut point analyses in the trial sample indicated that a cut point approach to

be less accurate than the item-based approach (Table 5), although a cut-off of ≥ 6 was supported in this sample. As in the trial sample, the internal consistency of the SOPHS in the community sample was high ($\alpha = 0.93$).

Discussion

The findings indicate that the SOPHS is a valid and reliable measure for screening social anxiety in the community. The SOPHS demonstrated acceptable convergent and discriminant reliability with respect to other severity measures. Each item in the SOPHS significantly discriminated social anxiety severity, with the total information for the screener outperforming the Mini-SPIN at critical regions on the symptom continuum. With respect to the MINI clinical assessment, the SOPHS performed equally well to the Mini-SPIN and much lengthier SPIN, although with greater sensitivity than the Mini-SPIN using the cut point recommended by Seeley-Wait et al (2009). In screening applications, a high probability of detecting true cases (high sensitivity) and a high probability that negative test results indicate non-caseness (negative predictive value) are the key tests of screener accuracy. For both of these indices, the SOPHS performed better than the Mini-SPIN at the cut point of 7. The area under the curve for the SOPHS was not significantly different to that for the SPIN or Mini-SPIN. In addition, the SOPHS was shown to be accurate in screening DSM-5 social anxiety criteria in an independent population-based sample. To our knowledge, no other social anxiety screener has been tested against DSM-5 criteria. An item-based scoring criterion provided accurate screening across samples, and is recommended for screening. In the community-based sample, a cut point of ≥ 8 on the SOPHS was also found to provide accurate screening for social anxiety, while in the trial sample, a cut point of ≥ 6 was most accurate although less accurate than the item-based criterion. The discrepancy may be due to the younger age of the trial sample.

The length of the SOPHS is comparable to the three-item Mini-SPIN. In the trial sample, 46.8% of respondents to the SOPHS were only asked one question (specifically, those least at risk of social anxiety disorder), while the remainder were asked five, resulting in an mean of 3.1 questions per respondent. The SOPHS therefore has the advantage that the nearly 50% of the population without any level of social anxiety can be screened using a single item, avoiding the necessity to ask irrelevant questions of those without symptoms. Furthermore, unlike the SPIN and Mini-SPIN, the SOPHS is available without requirement of permission or fee for use. The SOPHS includes an item assessing functional impairment, not covered by the Mini-SPIN, which may be a critical component in distinguishing both caseness and severity (Stein, Torgrud, & Walker, 2000). The use of the SOPHS or Mini-SPIN as screening scales for social anxiety was supported in the present study, while the use of the full SPIN for discriminating between community-based respondents with and without social anxiety disorder was not supported, given its considerably greater response burden with no advantage in accuracy.

As reflected in previous epidemiological research (Furmark et al., 1999; Schneier, Johnson, Hornig, Liebowitz, & Weissman, 1992), severity of social anxiety symptoms was higher in respondents who were female, younger, unmarried, less educated or unemployed. There was also indication of high comorbidity with generalized anxiety, panic and depression, indicated by significantly higher severity scores and rates of social anxiety disorder caseness among those meeting criteria for these disorders, as seen in previous population-based research (Kessler, Chiu, Demler, Merikangas, & Walters, 2005). Participants reporting suicidal ideation also had higher severity of social anxiety symptoms, consistent with previous research (Sareen et al., 2005). The consistency of the present findings with previous research on risk factors for social anxiety indicates potential targets for selective intervention studies.

The present study had a number of important strengths, including the use of two large population-based samples, and the comparison of the scale to clinical assessment and to other standard scales. Nevertheless, further validation of the screening scale in clinically-identified populations may provide additional evidence for its use more broadly in screening. The sensitivity and specificity of the SPIN and Mini-SPIN was not as high as that seen in other studies, which may reflect both the use of a community-based sample in the present study and the length of time between the screening and the clinical assessment in the trial sample (up to three weeks). Validation of the scale in primary care and other clinical settings may also be beneficial. All participants in the trial sample who received the clinical assessment had slightly elevated symptoms of generalised anxiety. This criterion may have biased detection of non-cases, potentially resulting in conservative estimates of specificity for the SOPHS, SPIN and Mini-SPIN, as reflected in the higher estimates of accuracy seen in the secondary population-based sample. Resource constraints in population-based research limit the use of structured clinical interviews administered in-person; therefore DSM-IV criteria were measured using a phone-administered MINI interview (without assessment of interrater reliability) and DSM-5 criteria were assessed using a systematic self-report checklist. Further validation based on psychiatrist diagnosis may provide a more robust test of the current findings.

There was an age restriction in the trial sample (18-30), which may reduce the applicability of the findings to older adult populations. However, the findings from the population sample, which included a broad age distribution, suggest the scale may also be accurate in general community settings. Likewise, the underrepresentation of males in the sample may limit the applicability to both genders, although estimates of sensitivity and specificity were found to be comparable in males and females. A related limitation of both samples was the low response rate, with no data available on the characteristics of non-

responders. However, the comparable outcomes in the population-weighted sample provide further evidence that the SOPHS is an accurate screener in the community. The SPIN was not administered in the community-based sample, so comparative validity could not be further assessed.

The decrement in accuracy of both the SOPHS and the SPIN among people of linguistically diverse backgrounds was not surprising, as a certain level of English literacy may be required to appropriately respond to items in the scale and the prevalence of social anxiety disorder varies across cultures (Heinrichs et al., 2006). Nevertheless, this finding may suggest a need for additional development of screeners for specific diverse populations. Finally, sensitivity to change could not be directly assessed in the current study, as the intervention did not target social anxiety symptoms and respondents meeting clinical criteria for social anxiety disorder were excluded from participation in the trial. However, the less restrictive range of the SOPHS (0-20) compared to the Mini-SPIN (0-12) may provide greater scope for assessing change in severity scores over time – future assessment of sensitivity to change is required.

In conclusion, data from two population-based studies indicate that the SOPHS is a valid and reliable screening tool for social anxiety, relative to DSM-IV and DSM-5 clinical criteria. The scale may also be used assess severity of social anxiety symptoms. The SOPHS performed as accurately as the 17-item SPIN and the 3-item Mini-SPIN, without the usage constraints of existing scales and with greater coverage of functional impairment. The SOPHS was also found to be a valid screen for DSM-5 social anxiety disorder, unlike any existing screener to date. The ability of the SOPHS to rule out social anxiety disorder using a single item for approximately half of the population is an additional advantage. The SOPHS forms a robust counterpart to the recently developed Panic Disorder Screener (Batterham, Mackinnon, & Christensen, 2015) to provide a new set of screeners for identifying anxiety

disorders in the community. The present data indicate that SOPHS may be used to screen for social anxiety in population-based settings.

References

- Australian Bureau of Statistics. (2014). *3101.0 Australian Demographic Statistics*. Canberra
- Batterham, P. J., Calear, A. L., Sunderland, M., Carragher, N., Christensen, H., & Mackinnon, A. J. (2013). Hierarchical screening for multiple mental disorders. *Journal of Affective Disorders, 151*(1), 229-236.
- Batterham, P. J., Mackinnon, A. J., & Christensen, H. (2015). The panic disorder screener (PADIS): Development of an accurate and brief population screening tool. *Psychiatry Research*.
- Bruce, T. J., & Saeed, S. A. (1999). Social anxiety disorder: a common, underrecognized mental disorder. *Am Fam Physician, 60*(8), 2311-2320, 2322.
- Christensen, H., Batterham, P., Mackinnon, A., Griffiths, K., Kalia Hehir, K., Kenardy, J., et al. (2014). Prevention of Generalized Anxiety Disorder Using a Web Intervention, iChill: Randomized Controlled Trial. *Journal of Medical Internet Research, 16*(9), e199.
- Christensen, H., Batterham, P. J., Grant, J. B., Griffiths, K. M., & Mackinnon, A. J. (2011). A population study comparing screening performance of prototypes for depression and anxiety with standard scales. *BMC Med Res Methodol, 11*, 154.
- Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in internet interventions for anxiety and depression. *J Med Internet Res, 11*(2), e13.
- Christensen, H., Griffiths, K. M., Mackinnon, A. J., Kalia, K., Batterham, P. J., Kenardy, J., et al. (2010). Protocol for a randomised controlled trial investigating the effectiveness of an online e health application for the prevention of Generalised Anxiety Disorder. *BMC Psychiatry, 10*, 25.

- Clark, D. B., Feske, U., Masia, C. L., Spaulding, S. A., Brown, C., Mammen, O., et al. (1997). Systematic assessment of social phobia in clinical practice. *Depress Anxiety*, 6(2), 47-61.
- Connor, K. M., Davidson, J. R., Churchill, L. E., Sherwood, A., Foa, E., & Weisler, R. H. (2000). Psychometric properties of the Social Phobia Inventory (SPIN). New self-rating scale. *Br J Psychiatry*, 176, 379-386.
- Connor, K. M., Kobak, K. A., Churchill, L. E., Katelnick, D., & Davidson, J. R. (2001). Mini-SPIN: A brief screening assessment for generalized social anxiety disorder. *Depress Anxiety*, 14(2), 137-140.
- Furmark, T., Tillfors, M., Everz, P., Marteinsdottir, I., Gefvert, O., & Fredrikson, M. (1999). Social phobia in the general population: prevalence and sociodemographic profile. *Soc Psychiatry Psychiatr Epidemiol*, 34(8), 416-424.
- Heinrichs, N., Rapee, R. M., Alden, L. A., Bogels, S., Hofmann, S. G., Oh, K. J., et al. (2006). Cultural differences in perceived social norms and social anxiety. *Behav Res Ther*, 44(8), 1187-1197.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, 3, 424-453.
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. L., et al. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med*, 32(6), 959-976.
- 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. *Arch Gen Psychiatry*, 62(6), 593-602.

- Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*, 62(6), 617-627.
- Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Lawrence Erlbaum.
- Mackinnon, A. (2000). A spreadsheet for the calculation of comprehensive statistics for the assessment of diagnostic tests and inter-rater agreement. *Comput Biol Med*, 30(3), 127-134.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus User's Guide. Sixth Edition*. Los Angeles, CA: Muthén & Muthén.
- Patel, A., Knapp, M., Henderson, J., & Baldwin, D. (2002). The economic consequences of social phobia. *J Affect Disord*, 68(2-3), 221-233.
- Ranta, K., Kaltiala-Heino, R., Rantanen, P., Tuomisto, M. T., & Marttunen, M. (2007). Screening social phobia in adolescents from general population: the validity of the Social Phobia Inventory (SPIN) against a clinical interview. *Eur Psychiatry*, 22(4), 244-251.
- Sareen, J., Cox, B. J., Afifi, T. O., de Graaf, R., Asmundson, G. J., ten Have, M., et al. (2005). Anxiety disorders and risk for suicidal ideation and suicide attempts: a population-based longitudinal study of adults. *Arch Gen Psychiatry*, 62(11), 1249-1257.
- Schneier, F. R., Johnson, J., Hornig, C. D., Liebowitz, M. R., & Weissman, M. M. (1992). Social phobia. Comorbidity and morbidity in an epidemiologic sample. *Arch Gen Psychiatry*, 49(4), 282-288.
- Seeley-Wait, E., Abbott, M. J., & Rapee, R. M. (2009). Psychometric properties of the mini-social phobia inventory. *Prim Care Companion J Clin Psychiatry*, 11(5), 231-236.

- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry, 59 Suppl 20*, 22-33;quiz 34-57.
- Slade, T., Johnston, A., Oakley Browne, M. A., Andrews, G., & Whiteford, H. (2009). 2007 National Survey of Mental Health and Wellbeing: methods and key findings. *Aust N Z J Psychiatry, 43(7)*, 594-605.
- Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *Jama, 282(18)*, 1737-1744.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med, 166(10)*, 1092-1097.
- Stein, M. B., Torgrud, L. J., & Walker, J. R. (2000). Social phobia symptoms, subtypes, and severity: findings from a community survey. *Arch Gen Psychiatry, 57(11)*, 1046-1052.
- Takane, Y., & De Leeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. *Psychometrika, 52(3)*, 393-408.
- Weeks, J. W., Spokas, M. E., & Heimberg, R. G. (2007). Psychometric evaluation of the mini-social phobia inventory (Mini-SPIN) in a treatment-seeking sample. *Depress Anxiety, 24(6)*, 382-391.
- Williams, J. W., Jr., Pignone, M., Ramirez, G., & Perez Stellato, C. (2002). Identifying depression in primary care: a literature synthesis of case-finding instruments. *Gen Hosp Psychiatry, 24(4)*, 225-237.

Zbozinek, T. D., Rose, R. D., Wolitzky-Taylor, K. B., Sherbourne, C., Sullivan, G., Stein, M. B., et al. (2012). Diagnostic overlap of generalized anxiety disorder and major depressive disorder in a primary care sample. *Depress Anxiety*, 29(12), 1065-1071.

Table 1: Sample characteristics for participants only in the full sample and for participants who received a clinical assessment

	Survey only sample N = 10633	Clinical assessment sample N = 1674	χ^2 / F	p
	n (%) or Mean (SD)	n (%) or Mean (SD)		
Gender = female	7428 (70%)	1270 (76%)	25.1	<0.001
Age (range 18-30)	25.7 (4.4)	25.4 (3.4)	9.7	0.002
<i>Marital status</i>			26.0	<0.001
Married	1956 (18.5%)	238 (14.3%)		
De facto	1973 (18.6%)	375 (22.5%)		
Separated / divorced / widowed	132 (1.2%)	17 (1.0%)		
Single, never married	6539 (61.7%)	1036 (62.2%)		
<i>Education</i>			3.2	0.524
< High school	226 (2.1%)	31 (1.9%)		
4 years high school	398 (3.7%)	61 (3.6%)		
6 years high school	3410 (32.1%)	532 (31.8%)		
Bachelor's degree	4981 (46.9%)	817 (48.8%)		
Higher degree	1603 (15.1%)	232 (13.9%)		
<i>Employment status</i>			6.3	0.180
Employed full-time	6147 (58.2%)	976 (58.6%)		
Employed PT/casual, seeking FT	581 (5.5%)	94 (5.6%)		
Employed part-time / casual	2346 (22.2%)	379 (22.8%)		
Unemployed looking for work	690 (6.5%)	118 (7.1%)		

Not in the labour force	795 (7.5%)	98 (5.9%)		
<i>Language spoken at home</i>			9.0	0.011
English	8478 (79.8%)	1383 (82.6%)		
English and other	1844 (17.4%)	259 (15.5%)		
Other	301 (2.8%)	32 (1.9%)		
SOPHS score	2.4 (3.6)	3.9 (4.3)	248.2	<0.001
SOPHS criteria	1988 (18.7%)	559 (33.5%)	317.6	<0.001
SPIN score	14.9 (11.8)	19.4 (12.5)	202.3	<0.001
SPIN criteria (>=19)	3224 (30.3%)	780 (46.6%)	386.8	<0.001
GAD-7 score	5.4 (4.7)	9.2 (4.0)	982.2	<0.001
PHQ-9 score	5.6 (5.3)	8.6 (5.2)	446.5	<0.001
Presence of suicidal ideation	1465 (13.8%)	404 (24.1%)	119.9	<0.001

Notes: SOPHS: Social Phobia Screener; SPIN: Social Phobia Inventory; GAD:

Generalized Anxiety Disorder; PHQ: Patient Health Questionnaire

Table 2: Rates of endorsement and estimates of Item Response Theory parameters for SOPHS items ($n = 12,361$)

Item	Factor loading	IRT parameters				
		Slope	T1	T2	T3	T4
Fearful or embarrassed	0.89	3.62	-0.09	0.94	1.75	2.36
Excessive or unreasonable	0.88	3.72	0.46	1.23	1.94	2.60
Avoided social situations	0.87	3.78	0.48	1.42	2.03	2.60
Suffered social situations	0.81	4.46	0.45	1.37	2.04	2.68
Life disrupted	0.77	2.90	0.55	1.44	2.09	2.62

Note: IRT parameters indicate Item Response Theory parameters of slope (discrimination) and response category thresholds (T1-T4) for each item based on a logistic two-parameter graded response model

Table 3: Relationship between sample characteristics and SOPHS severity scores

	Mean	SD	F	p
<i>Age</i>			16.6	<0.001
<25	2.77	3.85		
>=25	2.49	3.63		
<i>Gender</i>			22.0	<0.001
Male	2.36	3.67		
Female	2.70	3.74		
<i>Marital status</i>			38.4	<0.001
Married	1.86	3.11		
De facto	2.66	3.71		
Separated / divorced / widowed	3.56	4.53		
Single, never married	2.78	3.84		
<i>Education status</i>			41.1	<0.001
< High school	3.74	4.50		
4 years high school	3.86	4.77		
6 years high school	2.94	4.02		
Bachelor's degree	2.37	3.46		
Higher degree	2.12	3.23		
<i>Employment status</i>			38.5	<0.001
Employed full-time	2.33	3.44		
Employed PT/casual, seeking FT	2.86	3.74		
Employed part-time / casual	2.69	3.78		

Unemployed looking for work	3.81	4.60		
Not in the labour force	3.22	4.43		
<i>Language spoken at home</i>			0.6	0.576
English	2.59	3.71		
English and other	2.63	3.76		
Other	2.80	3.85		
<i>GAD-7 anxiety criteria</i>			1792.6	<0.001
Yes	5.66	5.22		
No	2.03	3.04		
<i>PHQ-9 depression criteria</i>			2304.8	<0.001
Yes	7.25	5.47		
No	2.13	3.13		
<i>Suicidal ideation</i>				
Yes	5.73	5.03	1781.7	<0.001
No	2.04	3.11		

Notes: GAD: Generalized Anxiety Disorder; PHQ: Patient Health Questionnaire

Table 4: Screening performance indices for the SOPHS, compared to the SPIN and Mini-SPIN, in identifying MINI clinical caseness for social phobia

	Sensitivity (95% CI)		Specificity (95% CI)		PPV (95% CI)		NPV (95% CI)	
SOPHS	0.779	(0.708, 0.840)	0.717	(0.693, 0.739)	0.227	(0.193, 0.264)	0.968	(0.955, 0.977)
SPIN	0.773	(0.701, 0.835)	0.711	(0.688, 0.734)	0.224	(0.190, 0.261)	0.967	(0.954, 0.977)
Mini-SPIN	0.742	(0.668, 0.808)	0.727	(0.703, 0.749)	0.227	(0.192, 0.264)	0.963	(0.951, 0.973)

Note: SOPHS: Social Phobia Screener; SPIN: Social Phobia Inventory; Mini-SPIN: short form of the Social Phobia Inventory; MINI: Mini International Neuropsychiatric Interview; PPV: positive predictive value; NPV: negative predictive value

Table 5: Performance of the SOPHS at various cut points in detecting Social Anxiety Disorder in the trial sample (N = 1674) and the independent population-based sample (N = 4,214)

	Trial sample				Community sample			
	Sensitivity (95% CI)		Specificity (95% CI)		Sensitivity (95% CI)		Specificity (95% CI)	
SOPHS >= 6	0.730	(0.655, 0.797)	0.747	(0.724, 0.769)	0.955	(0.938, 0.969)	0.750	(0.736, 0.765)
SOPHS >= 7	0.650	(0.572, 0.723)	0.796	(0.572, 0.723)	0.935	(0.914, 0.952)	0.802	(0.788, 0.815)
SOPHS >= 8	0.543	(0.463, 0.622)	0.838	(0.818, 0.856)	0.898	(0.874, 0.919)	0.845	(0.833, 0.857)
SOPHS >= 9	0.500	(0.417, 0.583)	0.879	(0.861, 0.895)	0.838	(0.809, 0.864)	0.880	(0.868, 0.890)

Figure 1: CONSORT-based flow diagram of participant inclusion

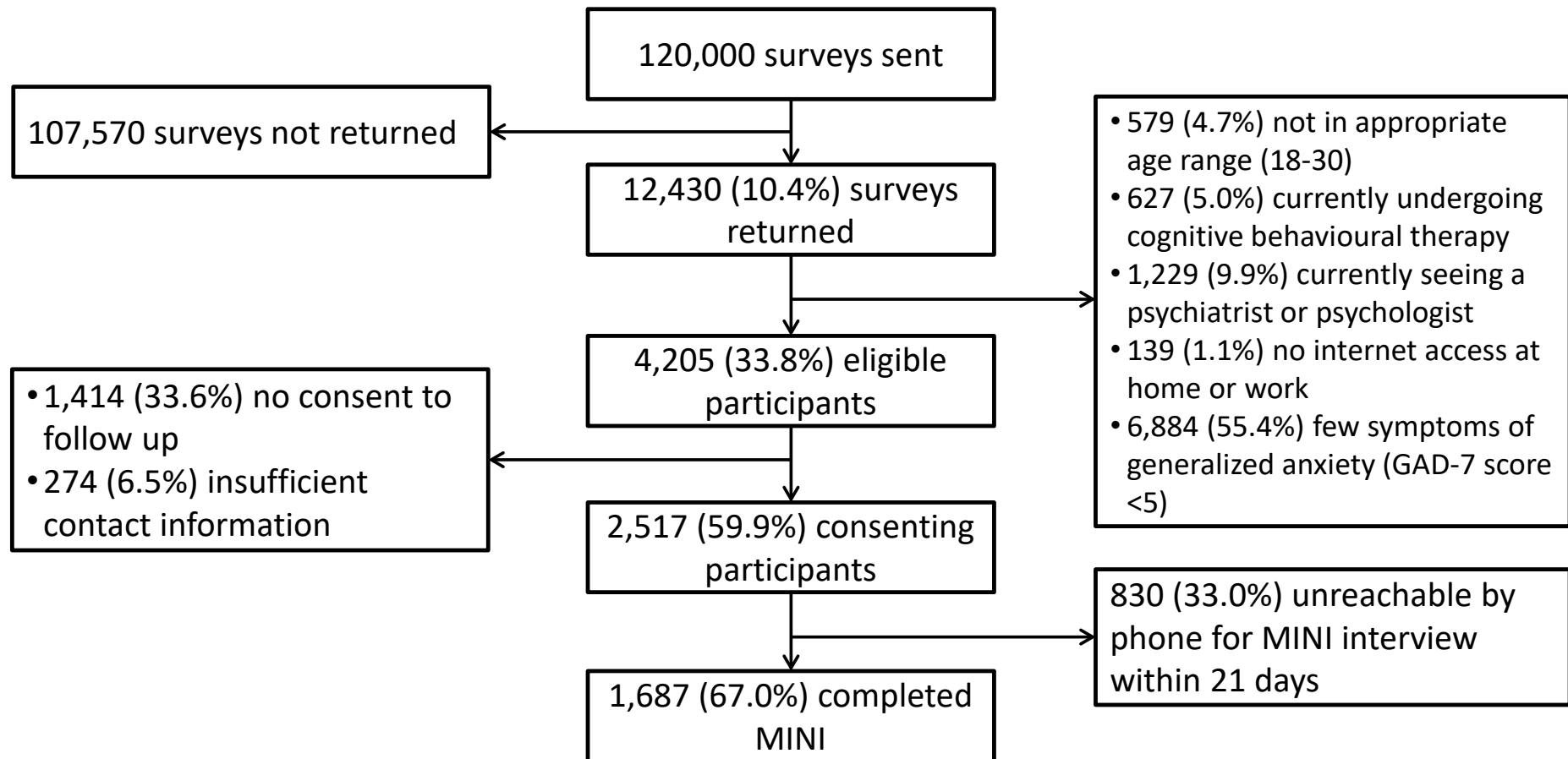
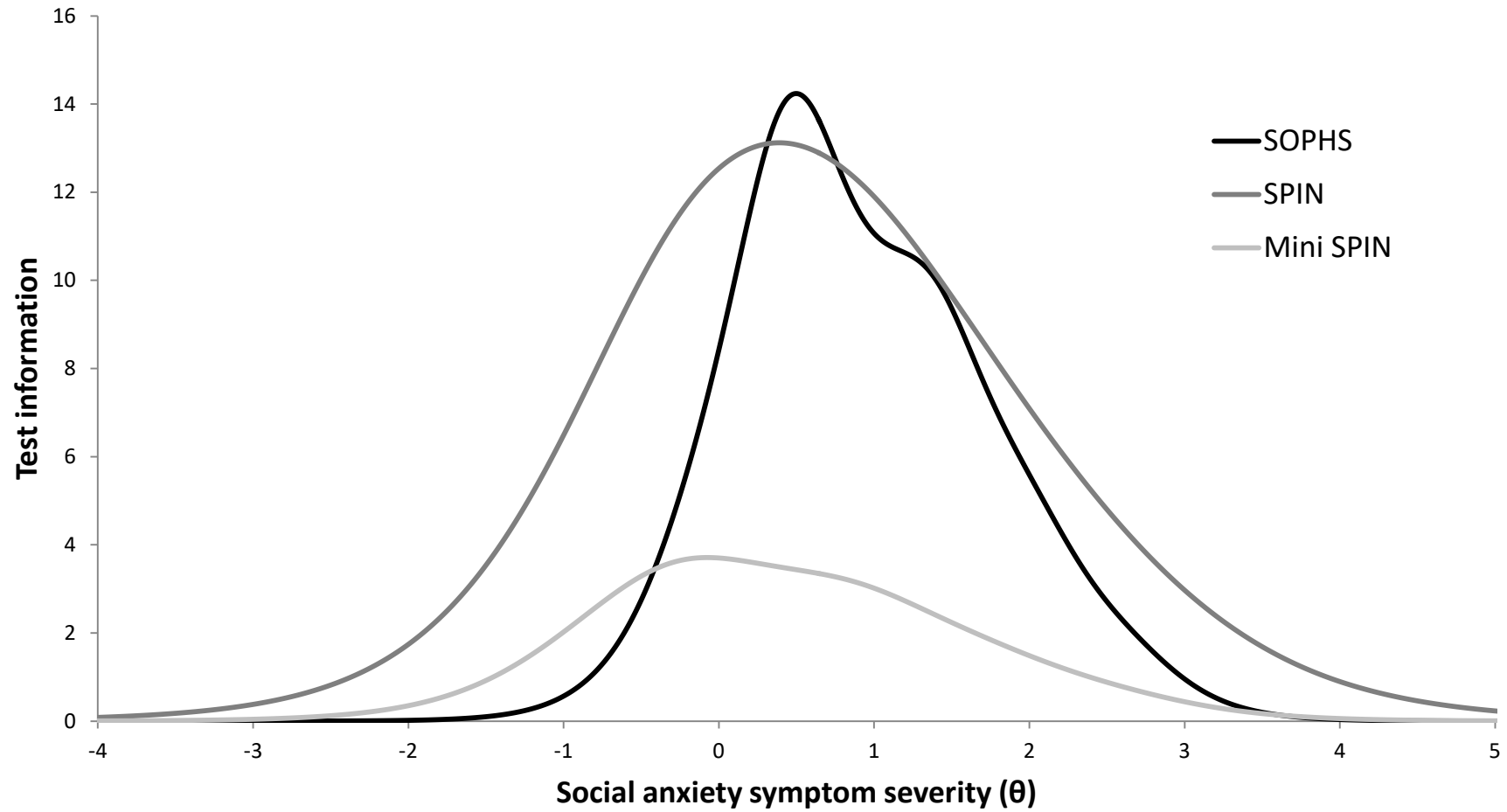


Figure 2: Test information curves for the SOPHS, SPIN and Mini-SPIN as a function of social anxiety symptom severity



Appendix 1: Social anxiety disorder (current) checklist for DSM-5

1. In the **past six months**, did you frequently experience strong fear or anxiety about social situations? Social situations might include speaking in public, eating with others, writing when someone is watching, or attending a meeting or party.
 Yes
 No [*→ Skip to end of section*]
2. In the **past six months**, did social situations nearly always make you feel frightened or anxious?
 Yes
 No [*→ Skip to end of section*]
3. In the **past six months**, when you were in social situations, did you fear that you would act in a way that would be humiliating, embarrassing or offensive?
 Yes
 No
4. In the **past six months**, when you were in social situations, did you fear that your anxiety symptoms would be humiliating, embarrassing or offensive?
 Yes
 No
5. In the **past six months**, did you avoid social situations because of your fear or anxiety of them?
 Yes
 No
6. In the **past six months**, did you participate in social situations despite experiencing strong fear or anxiety?
 Yes
 No
7. In the **past six months**, was your fear or anxiety about social situations greater than most people would experience in the same situation?
 Yes
 No [*→ Skip to end of section*]
8. In the **past six months**, did your fear or anxiety of social situations interfere with your normal routine or cause problems at home, work, school or socially?
 Yes
 No

Scoring criteria for social anxiety disorder:

All of the following conditions must be met:

- #1 = “Yes”
- #2 = “Yes”
- #3 = “Yes” or #4 = “Yes”
- #5 = “Yes” or #6 = “Yes”
- #7 = “Yes”
- #8 = “Yes”