Are male and female responses to social phobia diagnostic criteria comparable?

ERICA CROME,¹ ANDREW BAILLIE¹ & ALAN TAYLOR²

1 Centre for Emotional Health, Psychology Department, Macquarie University, NSW, Australia

2 Psychology Department, Macquarie University, NSW, Australia

Key words

social phobia, measurement invariance, confirmatory factor analysis

Correspondence

Erica Crome, Psychology Department, Macquarie University, NSW 2109, Australia.
Telephone (+61) 2-9850-9436
Fax (+61) 2-9850-8062
Email: erica.crome@mq.edu.au

Received 28 September 2010;
revised 11 April 2011;
accepted 18 May 2011

Abstract

Females typically report higher social phobia levels than males in community samples, and this may be due to sex bias in assessment measures. This study aims to establish whether patterns of responding to social phobia diagnostic criteria in the Composite International Diagnostic Interview (CIDI) are comparable across males and females. A subsample of participants in the Australian National Survey of Mental Health and Wellbeing (1997) reporting at least one social fear were selected ($n = 1755$). Analyses were conducted using a series of multi-group confirmatory factor analyses for categorical data, with unique steps to model invariance of residual variances. Partial, but not full, invariance was established, as males and females differed in their responses to items assessing physical anxiety symptoms at low levels of social fear. Whilst these differences were statistically significant, they are likely not to affect clinical practice or rates of social phobia diagnosis. This supports differences on this measure being interpreted as genuine, and strengthens findings females are more vulnerable to social phobia than males. Copyright © 2012 John Wiley & Sons, Ltd.

Introduction

Social phobia, or social anxiety disorder, is characterized by anxiety regarding social situations and potential negative evaluation [American Psychiatric Association (APA), 2000]. It is one of the most common mental disorders in epidemiological surveys, with estimates of lifetime prevalence in Western communities as high as 16% (Hidalgo *et al.*, 2001). Social phobia typically emerges in adolescence and is associated with chronic distress, impairment and comorbid disorder (Furmark, 2002). Females in community samples are more likely than males to meet social phobia criteria, with differential neurobiology, socialization and life experiences all presented as possible contributors to this difference (Rapee and Spence, 2004).

However, before these avenues are researched further, a more basic explanation must be ruled out. That is, that differential sex prevalences are due to differences in how males and females respond to diagnostic measures. If such a measurement bias exists, it is possible males and females would appear differentially affected, despite actually having comparable levels of underlying social phobia. Therefore, it is important that, before male and female responses on any diagnostic measure can be meaningfully compared, this underlying measurement assumption must be tested.

There is growing evidence males and females do differ in their patterns of responses to psychological assessments. For example, females have a generally tendency to use extreme response options (e.g. “strongly agree”, “strongly disagree”) more frequently than males (Ritvo *et al.*, 2008). There are

also important sex differences in recall, willingness to self-report distress, interpretation of symptomatology, interviewer effects and treatment seeking behaviour (Grant and Weissman, 2007). These measurement biases can even occur in diagnostic criteria themselves, with Agrawal and Lynskey (2007) reporting the wording of “hazardous use” criteria for cannabis abuse biased diagnosis towards males. Whilst differences associated with measurement variance may appear trivial, potential consequences include misdiagnosis and inappropriate allocation of resources, as well as misdirecting research into etiology and treatment differences which may not exist (Wisner and Dolan-Sewell, 2007). Assessing the accuracy of assessment tools is especially timely given both international mental health assessment systems International Statistical Classification of Diseases, 10th edition [ICD-10: World Health Organization (WHO), 1992] and Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR: APA, 2000) are under review (Bögels *et al.*, 2010).

The assumption people from different groups respond to assessment items in the same way is referred to as measurement invariance. In order for groups, such as males and females, to be meaningfully compared this assumption must be met. If measures are not invariant, systematic differences between group responses may affect the comparability of outcomes, and group differences may be more indicative of another confounding factor (Meredith and Teresi, 2006). An example of this can be found in Chung and Breslau's (2008) finding that whether or not a person experienced an assaultive or non-assaultive trauma, in addition to their level of post-traumatic stress, determined self-reported distress and emotional numbing. Despite commonly researched sex differences in social phobia, there is minimal research supporting the invariance of common diagnostic measures. The most common measure of mental disorder prevalence is the Composite International Diagnostic Interview (CIDI: Kessler and Ustun, 2004; WHO, 1997). In Australia, variants of the CIDI have been used to assess the prevalence of mental disorders, assess the burden of disease and guide both mental health care policy and research (Andrews *et al.*, 2004; Issakidis *et al.*, 2004). Given the gravity of outcomes from these surveys, it is essential any potential measurement biases are detected. This paper aims to test the underlying measurement invariance assumption of the social phobia diagnostic section of the CIDI used in Australian surveys to determine whether outcomes for males and females can be meaningfully compared.

Material and methods

Participants

Australian National Survey of Mental Health and Well-being (NSMHWB) 1997

The Australian National Survey of Mental Health and Well-being (NSMHWB) 1997 provides the most recent Australian data suitable for measurement invariance analyses of a social phobia diagnostic measure. It was conducted by the Australian Bureau of Statistics (ABS) within a multi-stage stratified random sample of adults in private dwellings representative of the Australian population. In total, 10,641 people aged 18 and over participated in the voluntary face-to-face survey (response rate of 78.1%; Andrews *et al.*, 2001). Of the 10,641 respondents, 1755 (16.5%) endorsed items related to an unusually strong fear and/or avoidance of at least one of seven possible social scenarios (e.g. giving a speech or speaking in public) and thus screened into the social phobia diagnostic section. These responders were used in current analyses. Corresponding population weights for this subset were used to increase generalization to Australians experiencing social fears as per ABS (1997) guidelines. Sample demographics and proportions of symptom endorsement are detailed in Table 1.

Measure

Composite International Diagnostic Interview (CIDI)

The NSMHWB 1997 survey was based on the CIDI v2.1 (WHO, 1997), a fully structured, standardized diagnostic interview covering both ICD-10 and DSM-IV diagnoses. The CIDI has strong psychometric properties (Andrews and Peters, 1998; Pennell *et al.*, 2004) and was administered by trained lay interviewers via computer-assisted personal interviews. The CIDI v2.1 assesses the 12 month prevalence of DSM-IV social phobia using diagnostic criteria used to structure analyses. These were: (A) marked fear of social situations with potential for humiliation or embarrassment (*fear*), (B) exposure to the situation invokes sensations of anxiety (*physical anxiety*), (C) recognition the fear is unreasonable or excessive (*excessive*), (D) feared situations are avoided or endured with distress (*avoid/endure*), (E) social fears cause significant interference (*interference*) and (G&H) social fears are not due to other conditions such as substance use, medical illness or another mental disorder (*not due to other*: APA, 2000). Criterion F was excluded as it applies to people under the age of 18 who were not included in the survey. Criteria

Table 1. Demographics and diagnostic status of the National Survey of Mental Health and Well-being (1997) participants screening into the social phobia section of the Composite International Diagnostic Interview ($n=1755$)

	Males ($n=733$)		Females ($n=1022$)	
<i>Age</i>		Percentage		Percentage
18–35	283	38.6	388	38.0
35–55	307	41.9	476	46.6
55–> 75	143	19.5	158	15.4
<i>Marital status</i>				
Married/ <i>de facto</i>	383	52.3	553	54.1
Separated/divorced/widowed	111	15.1	261	25.5
Never married	239	32.6	208	20.4
<i>Employment</i>				
Full-time	472	64.4	306	29.9
Part-time	64	8.7	298	29.2
Unemployed	45	6.1	60	5.9
Not in labour force	152	20.8	358	35.0
<i>Primary language</i>				
English	706	96.3	980	95.9
Other	27	3.7	42	4.1
<i>DSM-IV diagnosis</i>				
12 month social phobia	62	8.5	83	8.1
(A) <i>Fear</i> present	352	44.3	443	55.7
(B) <i>Physical anxiety</i> present	401	41.6	563	58.4
(C) <i>Excessive</i> present	413	44.9	507	55.1
(D) <i>Avoid/endure</i> present	422	44.7	523	55.3
(E) <i>Interference</i> present	351	42.6	472	57.4
(G&H) <i>Not due to other</i> present	163	42.3	222	57.7

DSM-IV criteria are: (A) marked fear of social situations with potential for humiliation or embarrassment (*fear*); (B) exposure to the situation invokes sensations of anxiety (*physical anxiety*); (C) recognition the fear is unreasonable or excessive (*excessive*); (D) feared situations are avoided or endured with distress (*avoid/endure*); (E) social fears cause significant interference (*interference*); (G&H) social fears are not due to other conditions such as substance use, medical illness or another mental disorder (*not due to other*)

were dichotomously coded as either “not present” or “present” and further information on questions and criteria determination can be found in Robins *et al.* (1989).

Statistical analyses

The measurement invariance hypothesis was explored using a series of multi-group confirmatory factor analyses (MGCFA). These analyses test whether differences in the pattern of male and female responses emerge when restrictions to model comparability are imposed at each step (Steenkamp and Baumgartner, 1998; Byrne, 2004). This method was based on a two-step procedure recommended for categorical data outlined by Muthén

and Muthén (2008). These models are based on the assumption that if a construct is invariant, all people experiencing similar levels of an underlying construct should report a similar pattern of symptoms. When using categorical data, the underlying construct is modelled by a continuous and normally distributed latent response variable, based on the assumption each categorical indicator could have been measured in a more precise manner (Millsap and Yun-Tein, 2004). Firstly, a Baseline Model was created to model similarity in factor structure between males and females (freely estimated factor loadings and thresholds; Figure 1A). Then in step two, this model was compared to a second Measurement Invariance Model, modelling similarities between males and females

Figure 1. Three step measurement invariance analysis based on a one-factor model of social phobia: (A) Step One (Baseline Model) with factor loadings, thresholds and residual variance freely estimated between males and females; (B) Step Two (Measurement Invariance Model) factor loadings and thresholds restrained to equality; (C) Step Three (Strict Measurement Invariance Model) residual variance restrained to equality.

on both the factor loading (how strongly each item related to underlying social phobia) and threshold (the level of social phobia at which a response changed from a “no” to a “yes”) of each item (Figure 1B). If there are no significant

differences between these two models, it is interpreted as a sign of no systematic differences in how males and females approach each assessment item. However, significant differences between the models suggests males and females systematically differ on some questionnaire items, whether this be due to real differences or measurement bias.

Whilst this two-step approach overcomes common modelling limitations associated with categorical data, it may ignore important sex differences in residual variance. Residual, or error, variance refers to the variance in each assessment item not accounted for by the underlying construct. For example, males and females may respond in an equivalent manner to items worded in a positive manner, yet differ in their reactions to negatively worded items, despite both assessing the same underlying construct (e.g. social phobia). Assessing potential differences in residual variance is important to ensure the accuracy of invariance analyses (Steenkamp and Baumgartner, 1998; Meredith and Teresi, 2006). To address this, a third step was included which compared an additional model, the Strict Invariance Model (Figure 1C), with the Measurement Invariance Model to explore equality of residual variance. If significant differences emerged between consecutive models at any step, the measurement invariance hypothesis would be rejected and partial invariance would be explored. This determines whether any items can be used to provide meaningful comparisons between the two groups (Steenkamp and Baumgartner, 1998). Modification indices estimating changes in model fit with specific changes can be used to guide the formation of partial invariance models (Schreiber, 2008).

The current analyses were conducted in MPlus (v 5.21: Muthén and Muthén, 2009) using a robust weighted least squares (WLSMV) estimator for categorical data. Theta parametrization was selected to allow the specification of residual variances (Millsap and Yun-Tein, 2004; Muthén and Muthén, 2008). Model fit was assessed via multiple fit indexes, including a robust chi-squared test (χ^2) for WLSMV estimators, Comparative Fit Index (CFI), Tucker–Lewis Index (TLI) and root mean square error of approximation (RMSEA) and the weighted root mean square residual (WRMR: Brown, 2006; Schreiber, 2008). Non-significant χ^2 values suggest model retention, and cutoff values for other fit indexes in categorical data were ≥ 0.95 for CFI, ≥ 0.96 for TLI, < 0.90 for WRMR and < 0.06 for RMSEA (Schreiber, 2008).

Two indexes were used to assess differences between the models, the first being difference in χ^2 adjusted for nested models using MPlus DIFFTEST. Non-significant differences on this index supports the assumption of comparability between male and female responses, as adding increased

restrictions did not significantly alter model fit (Muthén and Muthén, 2008). Changes in CFI values were the second change index, with differences less than 0.01 suggesting no significant difference between models (Cheung and Rensvold, 2002). Current diagnostic conceptualizations of social phobia guided model specification, with all criteria loading onto a single social phobia factor. This was supported by evidence of an underlying social phobia dimension (Crome *et al.*, 2010). Using MPlus defaults, the factor metric was set by the first indicator [(A) *fear*: Muthén and Muthén, 2008], and the baseline model was tested in males and females separately prior to analysis (Byrne, 2004). To conserve space, parameters from only the initial and final standardized solutions are reported and other estimates are available on request.

Results

The one-factor model appeared more applicable to females; however, fit indexes (detailed earlier) largely supported combining male and female responses to create the Baseline Model (Table 2). Fit indexes for the Baseline Model supported model retention, so this model was then compared to the Measurement Invariance Model in step two.

Change of fit indexes provided mixed support for equivalence of these two models, with $\Delta\chi^2$, but not ΔCFI , reflecting significant differences between models. Evidence for the increased ability of χ^2 over CFI to detect significant differences between models guided decisions to reject the

more restrictive model and thus the measurement invariance hypothesis (French and Finch, 2006). As full measurement invariance was not established, partial invariance was explored. Modification indices suggested model improvement if males and females were allowed to differ in their report of physical anxiety symptoms [(B) *physical anxiety*] and physical anxiety symptoms were allowed to vary as a function of social fears [(A) *fear*] for males only. These relationships were adopted to form the revised Partial Invariance Model detailed in Figure 2.

Figure 2. Partial Invariance Model based on a one-factor model of social phobia. Factor loadings and thresholds for criterion B (*physical anxiety*) freely estimated for males and females; residual variance between criteria A (*fear*) and criteria B (*physical anxiety*) allowed to co-vary for males only.

Table 2. Fit indexes and change in fit indexes between models for individual groups, full measurement invariance and partial invariance analyses

	χ^2	df	$\Delta\chi^2$	df	RMSEA	WRMR	TLI	CFI	ΔCFI
<i>Individual group analyses</i>									
Females ($n = 1022$)	10.743	9			0.014	0.576	0.998	0.998	
Males ($n = 733$)	23.417*	8			0.051	0.938	0.968	0.974	
<i>Measurement invariance analyses (Figure 1)</i>									
Baseline Model	33.499*	16			0.035	1.101	0.986	0.989	
Measurement Invariance Model	43.433*	20	9.95*	4	0.037	1.252	0.985	0.985	0.004
<i>Partial invariance analyses (Figure 2)</i>									
Original Baseline Model	33.499*	16			0.035	1.101	0.986	0.989	
Partial Invariance Baseline Model	17.827	15			0.015	0.800	0.998	0.998	
Partial Measurement Invariance Model	19.065	18			0.008	0.825	0.999	0.999	-0.01
Stricter Partial Measurement Invariance Model	18.224	22	1.062	5	0.000	0.856	1.002	1	-0.001

Note: χ^2 , chi-square goodness-of-fit index; df, degrees of freedom; $\Delta\chi^2$, adjusted change in χ^2 for nested models; RMSEA, root mean square error of approximation; WRMR, weighted root mean square residual; TLI, Tucker–Lewis Index; CFI, comparative fit index; ΔCFI , change in CFI

*Significance at $p < 0.05$; non-significance ($p > 0.05$) indicates better model fit.

Whilst not directly compared, these modifications appeared to provide a better picture of the performance of diagnostic items for the overall sample. As seen by $\Delta\chi^2$ and ΔCFI in Table 2, increasing model restrictions in Step Two (Partial Invariance Model) or Step Three (Strict Measurement Invariance Model) did not significantly affect model fit. This provides support for invariance between male and female responses on all diagnostic criteria aside from criteria B exposure to the situation invokes sensations of anxiety (*physical anxiety*) (Table 3). This difference was explored using logistic regression and cross-tabulation in SPSS (2006). Logistic regression identified a significant interaction between sex and social fears in the report of physical anxiety symptoms (interaction $\beta = -0.575$ (0.210), $p < 0.05$: Males = 0, Females = 1).

Cross-tabulation clarified this interaction, with evidence of minimal difference in the proportion of males and females reporting physical anxiety symptoms in the presence of a significantly high social fear (70.4% and 73.9%, respectively), yet a much higher proportions of females (73.9%) over males (32.8%) reporting physical anxiety symptoms at low levels of social fear. This highlighted differences in the specificity (number of people without social phobia who also report no physical symptoms) of this item for males and females. A higher specificity for males [0.55: 95% confidence interval (CI) = 0.51–0.58] compared with females (0.43: 95% CI = 0.40–0.47: $\chi^2 = 10.233$ (1), $p = 0.001$) meaning this item may be more informative of actual social phobia status for males. Differences in the sensitivity (true positives) could not be explored as every diagnosis of social phobia requires this criterion to be met. The clinical significance of this variance is unclear as, in practice, social phobia criteria are rarely considered in isolation. This effect on diagnosis would also only be observed in the 1.5% of females and 1.4% of males in the sample meeting all other diagnostic criteria aside from criterion B.

Discussion

The social phobia diagnostic section of the CIDI used in the NSMHWB 1997 meets criteria for partial measurement invariance. This means people with the same level of underlying social phobia will respond to most diagnostic items in comparable manner, regardless of whether they are male or female. Full invariance was not established as males tended to report fewer physical anxiety symptoms than females at low social fear levels. Whilst criterion B (physical anxiety) appears more informative for males, it is unlikely this has any real significance for clinicians and researchers as it is unlikely to affect rates of diagnosis. Even within the

small populations of males and females with potential for misdiagnosis based on this one criterion, given they also meet all other diagnostic criteria, they are unlikely to also experience the low levels of social fears where this relationship is observed. Overall, findings of strong partial invariance suggests the higher proportion of females in this sample being classified as having social phobia (Lampe *et al.*, 2003) reflects a genuine difference in the experience and incidence of social phobia between males and females.

This finding is consistent with an emerging measurement invariance literature demonstrating comparability between male and female responses to other clinical social anxiety measures including the Social Phobia and Anxiety Inventory (SPAI: Roberson-Nay *et al.*, 2007) and Goldberg Anxiety and Depression Scales (Leach *et al.*, 2008). The precise cause of the sex difference in physical symptoms at low levels of social anxiety is unclear, however it is consistent with findings socially anxious females are hyper-responsive on self-report, hemodynamic and autonomic measures. In contrast, males tend to display minimal physiological changes despite self-reported psychological symptoms (Grossman *et al.*, 2001). There is also evidence that the experience of social phobia may be qualitatively different for males and females based on factors not assessed during diagnosis, such as age of onset, severity of initial symptoms and rates of comorbidity (Bezerra De Menezes *et al.*, 2008). Subtle differences in the report of physical anxiety symptoms may be one of these differences actually captured during diagnosis. Alternatively, differences in physical symptoms may be due to findings females generally report more general physical symptoms regardless of pathology, measure or response format used or population sampled (Gijsbers van Wijk *et al.*, 1999; Barsky *et al.*, 2001). Interestingly, there is also a large overlap between general physical symptoms reported by females in community samples and those used to assess anxiety or distress (e.g. fatigue, dizziness, difficulty concentrating and sleeping, nausea and dyspnoea: Barsky *et al.*, 2001). Differences in socialization may also contribute to this difference, with males typically encouraged to rely on other forms of expression such as anger, over involvement in work or isolation rather than reporting symptoms of anxiety or distress (Rochlen *et al.*, 2010).

This research has many strengths, including using a large epidemiological sample, and methods and software best suited to the available data. Few researchers have applied Muthen's methods (e.g. Batinic *et al.*, 2008; Hutchinson *et al.*, 2008; Leach *et al.*, 2008) despite the potential for erroneous results associated with applying common measurement invariance methods designed for continuous data to ordinal or Likert-type data (Brown, 2006; Schreiber, 2008).

Table 3. Standardized factor loadings, thresholds and variance explained (R^2) for Baseline and Strict Measurement Invariance Models presented separately for males and females

	Males			Females		
	Standardized factor loading (SE)	Standardized threshold (SE)	R^2 (SE)	Standardized factor loading (SE)	Standardized threshold (SE)	R^2 (SE)
<i>Initial Multi-group Model</i>						
(A) Fear ¹	0.788* (0.095)	0.232* (0.068)	0.383* (0.057)	0.617* (0.071)	0.071 (0.052)	0.275* (0.046)
(B) Physical anxiety	0.807* (0.105)	0.040 (0.068)	0.394* (0.062)	0.637* (0.075)	-0.314* (0.054)	0.289* (0.049)
(C) Excessive	1.101* (0.141)	-0.006 (0.079)	0.548* (0.063)	1.143* (0.125)	-0.168* (0.068)	0.566* (0.054)
(D) Avoid/endure	0.708* (0.087)	-0.040 (0.065)	0.334* (0.055)	0.747* (0.083)	-0.192* (0.056)	0.358* (0.051)
(E) Interference	1.094* (0.142)	0.276* (0.082)	0.545* (0.064)	1.289* (0.141)	-0.028 (0.072)	0.624* (0.051)
(G&H) Not due to other	1.112* (0.173)	1.268* (0.143)	0.553* (0.077)	1.343* (0.164)	1.194* (0.118)	0.643* (0.056)
<i>Final Partial Measurement Invariance Model</i>						
(A) Fear ¹	0.618* (0.063)	0.203* (0.047)	0.276* (0.041)	0.636* (0.059)	0.203* (0.047)	0.288* (0.038)
(B) Physical anxiety	0.658* (0.091)	0.037 (0.063)	0.302* (0.058)	0.639* (0.076)	-0.191* (0.060)	0.290* (0.049)
(C) Excessive	1.127* (0.107)	0.029 (0.067)	0.559* (0.047)	1.160* (0.106)	0.029 (0.067)	0.574* (0.045)
(D) Avoid/endure	0.727* (0.070)	-0.045 (0.051)	0.346* (0.043)	0.749* (0.067)	-0.045 (0.051)	0.359* (0.041)
(E) Interference	1.222* (0.123)	0.248* (0.074)	0.599* (0.048)	1.259* (0.113)	0.248* (0.074)	0.613* (0.043)
(G&H) Not due to other	1.251* (0.142)	1.390* (0.117)	0.610* (0.054)	1.288* (0.128)	1.390* (0.117)	0.624* (0.047)

¹Marker variable.*Significance at $p < 0.05$.

This research has the additional strength of extending Muthen's methods to test invariance in residual variances, claimed to be essential for fair comparisons between groups (Meredith and Teresi, 2006). Whilst using retrospective self-report data is typically a limitation to psychiatric research, as patterns of responding were of interest in this study, it is not a concern in the current paper. Similarly, whilst no data is available for participants reporting no social fears given they screened out of the diagnostic section, as this measure is not used for assessment in this population, relying on a subsample will not likely affect generalization.

Using categorical data within confirmatory factor analysis (CFA) does have limitations, such as reduced accuracy of fit indexes (French and Finch, 2006) and difficulty ensuring relationships between the latent factor and latent response variable extend to observed variables (Millsap and Yun-Tein, 2004). However, the impact of this is likely to be reduced given the current sample was large and the model tested was simple (French and Finch, 2006). These analyses also do not provide a definitive answer as to whether any group differences are due to measurement bias or real differences in the experience of that particular symptom. Other concerns about error in hierarchical rules for diagnosis in the CIDI v2.1 social phobia section (Means-Christensen *et al.*, 2003) did not apply to the current study, as diagnostic criteria were assessed without applying hierarchy rules. This also means sex differences may still emerge if exclusion criteria differentially affect the sexes. The CIDI used in the NSMHWB 1997 has a narrower range of social situations used as screening items than versions of the CIDI used in more recent WHO World Mental Health (WMH) surveys (Kessler and Ustun, 2004). This potentially means not all participants with social fears screened into the diagnostic section.

Whilst this study assesses measurement invariance in the diagnostic criteria of the CIDI, it is important to acknowledge this does not necessarily precisely extend to social phobia diagnostic criteria prescribed by the DSM-IV. For example, relying primarily on physical anxiety symptoms may not fully represent criterion B ("exposure to the feared social situation almost invariably provokes anxiety, which may take the form of a situationally bound or situationally predisposed panic attack", APA, 2000, p. 456). It may be

physical symptoms are non-specific markers of distress rather than specific symptoms of social phobia (Clark *et al.*, 1994; Baillie and Rapee, 2005). The opportunity to explore differences in the report of physical symptoms between males and females is further limited by screening points within the CIDI designed to reduce participant load. Once participants report a sufficient number of physical anxiety symptoms to meet criterion B, no further physical anxiety symptoms are elicited, which creates a potential ceiling effect in the data.

Overall this research supports the invariance of the social phobia diagnostic section of the CIDI used in the NSMHWB 1997. This strengthens previous findings Australian females are more vulnerable to social phobia and supports research into differential causal and maintaining factors (Grant and Weissman, 2007). This research also identifies differences in how males and females may respond to items regarding physical anxiety symptoms at low levels of social fears. However, this difference does not appear likely to affect rates of social phobia diagnosis or responses of people experiencing clinical levels of social phobia. This research establishes the importance of explicitly testing underlying measurement invariance assumptions prior to interpreting group differences. Comparable analyses should be conducted in other versions of the CIDI, such as that used in the National Comorbidity Survey – Replication (Kessler *et al.*, 2004), to confirm that slight variations between surveys do not affect measurement invariance.

Acknowledgements

Both National Survey of Mental Health and Well-being were funded by the National Health Branch of the Commonwealth Department of Health and Aged Care, under the National Mental Health Strategy. They were conducted by the Australian Bureau of Statistics. The current study was supported by funding from the Macquarie University Research Excellence Scholarship to the first author. The authors also acknowledge the productive feedback provided by four anonymous reviewers.

Declaration of interest statement

The authors have no competing interests.

References

- Agrawal A., Lynskey M.T. (2007) Does gender contribute to heterogeneity in criteria for cannabis abuse and dependence? Results from the national epidemiological survey on alcohol and related conditions. *Drug and Alcohol Dependence*, **88**(2–3), 300–307, DOI: 10.1016/j.drugalcdep.2006.10.003
- American Psychiatric Association (APA). (2000) *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*, Washington, DC, APA.
- Andrews G., Henderson S., Hall W. (2001) Prevalence, comorbidity, disability and service utilization: overview of the Australian National Mental Health Survey. *The British Journal of Psychiatry*, **178**, 145–153, DOI: 10.1192/bjp.178.2.145
- Andrews G., Issakidis C., Sanderson K., Corry J., Lapsley H. (2004) Utilising survey data to

- inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. *The British Journal of Psychiatry*, **184**(6), 526–533, DOI: 10.1192/bjp.184.6.526
- Andrews G., Peters L. (1998) The psychometric properties of the Composite International Diagnostic Interview. *Social Psychiatry and Psychiatric Epidemiology*, **33**(2), 80–88, DOI: 10.1007/s001270050026
- Australian Bureau of Statistics (ABS). (1997) 4329.0 – Information Paper: Mental Health and Wellbeing of Adults, Confidentialized Unit Record File. <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4329.0Main+Features11997?OpenDocument> [15 August 2008].
- Baillie A.J., Rapee R.M. (2005) Panic attacks as risk markers for mental disorders. *Social Psychiatry and Psychiatric Epidemiology*, **40**(3), 240–244, DOI: 10.1007/s00127-005-0892-3
- Barsky A.J., Peekna H.M., Borus J.F. (2001) Somatic symptom reporting in women and men. *Journal of General Internal Medicine*, **16**(4), 266–275, DOI: 10.1046/j.1525-1497.2001.016004266.x
- Batinic B., Wolff H., Haupt C.M. (2008) Construcion and factorial structure of a short version of the Trendsetting Questionnaire (TDS-K): a cross-validation using multigroup confirmatory factor analyses. *European Journal of Psychological Assessment*, **24**(2), 88–94, DOI 10.1027/1015-5759.24.2.88
- Bezerra De Menezes G., Fontenelle L., Marcio V. (2008) Gender effect on clinical features and drug treatment response in social anxiety disorder (social phobia). *International Journal of Psychiatry in Clinical Practice*, **12**(2), 151–155, DOI: 10.1080/13651500701568317
- Bögels S.M., Alden L., Clark L.A., Pine D.S., Stein M.B., Voncken M. (2010) Social anxiety disorder: questions and answers for the DSM-V. *Depression and Anxiety*, **27**(2), 168–189, DOI: 10.1002/da.20670
- Brown T.A. (2006) *Confirmatory Factor Analysis for Applied Research*, New York, The Guilford Press.
- Byrne B.M. (2004) Testing for multigroup invariance using AMOS graphics: a road less traveled. *Structural Equation Modeling*, **11**(2), 272–300, DOI: 10.1207/s15328007sem1102_8
- Cheung G.W., Rensvold R.B. (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, **9**(2), 233–255, DOI: 10.1207/S15328007SEM0902
- Chung, H., Breslau, N. (2008) The latent structure of post-traumatic stress disorder: tests of invariance by gender and trauma type. *Psychological Medicine*, **38**(4), 563–573, DOI: 10.1017/S0033291707002589
- Clark D.A., Steer R.A., Beck A.T. (1994) Common and specific dimensions of self-reported anxiety and depression: implications for the cognitive and tripartite models. *Journal of Abnormal Psychology*, **103**(4), 645–645, DOI: 10.1016/S0005-7967(98)00087-4
- Crome E., Baillie A., Slade T., Ruscio A.M. (2010) Social phobia: further evidence of dimensional structure. *The Australian and New Zealand Journal of Psychiatry*, **44**(11), 1012–1020, DOI: 10.3109/00048674.2010.507544
- French B.F., Finch W.H. (2006) Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling*, **13**(3), 387–402, DOI: 10.1207/s15328007sem1303_3
- Furmark T. (2002) Social phobia: overview of community surveys. *Acta Psychiatrica Scandinavica*, **105**(2), 84–93, DOI: 10.1034/j.1600-0447.2002.1r103.x
- Gijsbers van Wijk C.M.T., Huisman H., Kolk A.M. (1999) Gender differences in physical symptoms and illness behavior. A health diary study. *Social Science & Medicine*, **49**(8), 1061–1074, DOI: 10.1016/S0277-9536(99)00196-3
- Grant B.F., Weissman M.M. (2007) Gender and the prevalence of psychiatric disorders. In Narrow W.E., First M.B., Sirovatka P.J., Regier D.A. (eds) *Age and Gender Considerations in Psychiatric Diagnosis: A Research Agenda for DSM-V*, pp. 31–45, Arlington, VA, APA.
- Grossman P., Wilhelm F.H., Kawachi I., Sparrow D. (2001) Gender differences in psychophysiological responses to speech stress among older social phobics: congruence and incongruence between self-evaluative and cardiovascular reactions. *Psychosomatic Medicine*, **63**(5), 765–777, DOI: 10.1111.117.6914
- Hidalgo R., Barnett S., Davidson, J.R.T. (2001) Social anxiety disorder in review: two decades of progress. *The International Journal of Neuropsychopharmacology*, **4**(03), 279–298, DOI: 10.1017/S1461145701002504
- Hutchinson S.R., Raymond K.J., Black K.R. (2008) Factorial invariance of a campus climate measure across race, gender, and student classification. *Journal of Diversity in Higher Education*, **1**(4), 235–250, DOI: 10.1037/a0014419
- Issakidis C., Sanderson K., Corry J., Andrews G., Lapsley H. (2004) Modelling the population cost-effectiveness of current and evidence-based optimal treatment for anxiety disorders. *Psychological Medicine*, **34**(1), 19–35, DOI: 10.1017/S003329170300881X
- Kessler R.C., Berglund P., Chiu W., Demler O., Heeringa S., Hiripi E., Jin R. (2004) The US National Comorbidity Survey Replication (NCS-R): design and field procedures. *International Journal of Methods in Psychiatric Research*, **13**(2), 69–92, DOI: 10.1002/mpr.167
- Kessler R.C., Ustun T.B. (2004) The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*, **13**(2), 93 – 121, DOI: 10.1002/mpr.168
- Lampe L., Slade T., Issakidis C., Andrews G. (2003) Social phobia in the Australian National Survey of Mental Health and Well-Being (NSMHWB). *Psychological Medicine*, **33**(4), 637–646, DOI: 10.1017/S0033291703007621
- Leach L.S., Christensen H., Mackinnon A.J. (2008) Gender differences in the endorsement of symptoms for depression and anxiety: are gender-biased items responsible? *The Journal of Nervous and Mental Disease*, **196**(2), 128–135, DOI: 10.1097/NMD.0b013e318162aa63
- Means-Christensen A., Sherbourne C.D., Roy-Byrne P., Craske M.G., Bystritsky A., Stein M.B. (2003) The Composite International Diagnostic Interview (CIDI-Auto): problems and remedies for diagnosing panic disorder and social phobia. *International Journal of Methods in Psychiatric Research*, **12**(4), 167–181, DOI: 10.1002/mpr.154
- Meredith W., Teresi J.A. (2006) An essay on measurement and factorial invariance. *Medical Care*, **44**, S69–S77, DOI: 10.1097/01.mlr.0000245438.73837.89
- Millsap R.E., Yun-Tein J. (2004) Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, **39**(3), 479–515, DOI: 10.1207/S15327906MBR3903_4
- Muthén L.K., Muthén B.O. (2008) *Mplus User's Guide*, 5th edition, Los Angeles, CA, Muthén & Muthén.
- Muthén L.K., Muthén B.O. (2009) *Mplus* (Version 5.21), Los Angeles, CA, Muthén & Muthén.
- Pennell B.E., Bowers A., Carr D., Chardoul S., Cheung G.Q., Dinkelmann K., Gebler N., Hansen S.E., Pennell S., Torres M. (2004) The development and implementation of the National Comorbidity Survey Replication, the National Survey of American Life, and the National Latino and Asian American Survey. *International Journal*

- of Methods in Psychiatric Research*, **13**(4), 241–269, DOI: 10.1002/mpr.180
- Rapee R.M., Spence S.H. (2004) The etiology of social phobia: empirical evidence and an initial model. *Clinical Psychology Review*, **24** (7), 737–767, DOI: 10.1016/j.cpr.2004.06.004
- Ritvo P., Myers R., Del Giudice M.L., Pazsat L., Campbell P.T., Howlett R.I., Mai V., Sullivan T., Tiro J., Rabeneck L. (2008) Factorial validity and invariance of a survey measuring psychosocial correlates of colorectal cancer screening in Ontario, Canada – a replication study. *Cancer Epidemiology, Biomarkers & Prevention*, **17**(11), 3279–3283, DOI: 10.1158/1055-9965.EPI-08-0241
- Roberson-Nay R., Strong D.R., Nay W.T., Beidel D. C., Turner S.M. (2007) Development of an abbreviated Social Phobia and Anxiety Inventory (SPAI) using item response theory: the SPAI-23. *Psychological Assessment*, **19**(1), 133–144, DOI: 10.1037/1040-3590.19.1.133
- Robins L.N., Wing J., Wittchen H.-U., Helzer J.E., Babor T.F., Burke J., Farmer A., Jablensky A., Pickens R., Regier D.A., Sartorius N., Towle L.H. (1989) The Composite International Diagnostic Interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*, **45**(12), 1069–1077.
- Rochlen A.B., Paterniti D.A., Epstein R.M., Duberstein P., Willeford L., Kravitz R.L. (2010) Barriers in diagnosing and treating men with depression: a focus group report. *American Journal of Mens Health*, **4**(2), 167–175, DOI: 10.1177/1557988309335823
- Schreiber J.B. (2008) Core reporting practices in structural equation modeling. *Research in Social and Administrative Pharmacy*, **4**(2), 83–97, DOI: 10.1016/j.sapharm.2007.04.003
- SPSS Inc. (2006) *SPSS for Windows v.15*, Chicago, IL, SPSS Inc.
- Steenkamp J.B.E.M., Baumgartner H. (1998) Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, **25**(1), 78–90.
- Wisner K.L., Dolan-Sewell R. (2007) Why gender matters. In Narrow W.E., First M.B., Sirovatka P.J., Regier D.A. (eds) *Age and Gender Considerations in Psychiatric Diagnosis: A Research Agenda for DSM-V*, pp. 7–18, Arlington, VA, APA.
- World Health Organization (WHO). (1992) *International Classification of Diseases and Related Disorders (ICD-10)*, Geneva, WHO.
- World Health Organization (WHO). (1997) *Composite International Diagnostic Interview Version 2.1*, Geneva, WHO.