Routledge
Taylor & Francis Group

Australian Psychologist

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rapy20

Psychometric properties of the Specific Phobia Dimensional Scale in an Australian adult community sample

Cameron Bartlett, Bethany M. Wootton & Karen Moses

To cite this article: Cameron Bartlett, Bethany M. Wootton & Karen Moses (14 Nov 2024): Psychometric properties of the Specific Phobia Dimensional Scale in an Australian adult community sample, Australian Psychologist, DOI: 10.1080/00050067.2024.2426463

To link to this article: https://doi.org/10.1080/00050067.2024.2426463

Published online: 14 Nov 2024.

AUSTRALIAN PSYCHOLOGIST
https://doi.org/10.1080/00050067.2024.2426463

ORIGINAL ARTICLE

OPEN ACCESS Check for updates

Psychometric properties of the Specific Phobia Dimensional Scale in an Australian adult community sample

Cameron Bartlett[a], Bethany M. Wootton[b] and Karen Moses[a]

[a]School of Psychology, Western Sydney University, Penrith, New South Wales, Australia; [b]Discipline of Psychology, Graduate School of Health, University of Technology Sydney, Ultimo, New South Wales, Australia

ABSTRACT

Objective: The Specific Phobia Dimensional Scale (SP-D) is a 10-item scale developed by the Diagnostic and Statistical Manual (Fifth Edition) Anxiety, Obsessive-Compulsive Spectrum, Posttraumatic and Dissociative Disorder work group to supplement current categorical approaches to Specific Phobia (SP) assessment. The aim of the current study was to examine the psychometric properties of the SP-D in an Australian community sample.

Method: A total of 285 participants (74% female) aged 18–76 years (M = 28.15; SD = 12.01) completed the study. A smaller subsample (n = 18) completed the measures of interest a second time to examine test–retest reliability.

Results: Confirmatory factor analysis indicated a unidimensional factor structure of the SP-D (root mean square error of approximation [RMSE] = 0.13; comparative fit index [CFI] = 0.94). Internal consistency and test–retest reliability was high (a = 0.95 and Intraclass Correlation Coefficient [ICC] = 0.95, respectively). The measure demonstrated adequate convergent validity with the Specific Phobia Questionnaire (SPQ; r_s = 0.54), and unsatisfactory divergent validity with the Fear Questionnaire – Agoraphobia subscale (FQ-Ag; r_s = 0.45).

Conclusions: Overall, the findings suggest that the SP-D can reliably and briefly measure SP symptoms within an Australian community sample. Further research is required to investigate divergent validity. Replication of the test–retest results are required due to small sample size. Overall, this study is limited in its female-biased convenience sample and recruitment methodologies.

KEY POINTS

What is already known about this topic:

(1) Specific phobia (SP) is an anxiety disorder associated with significant functional impairment.
(2) The Specific Phobia Dimensional Scale (SP-D) captures information unattainable by the traditional categorical classification system, including changes in presentation, severity and symptomology over time.
(3) It is important to validate the SP-D in more diverse samples

What this topic adds:

(1) This paper is the first to investigate the psychometric properties of the SP-D in an Australian sample.
(2) The SP-D is a reliable and valid tool to dimensionally assess SP symptomatology in Australian adults.
(3) The divergent validity of the SP-D requires further investigation

ARTICLE HISTORY
Received 5 November 2023
Accepted 30 October 2024

KEYWORDS
Specific phobia; dimensional scale; DSM-5; SP-D; psychometric properties

Introduction

Specific phobia (SP) is characterised by the experience of marked fear or anxiety in response to a particular object or situation (American Psychiatric Association, 2013). The fear or anxiety within SP typically occurs out of proportion to the actual danger posed by the object or situation, and results in significant behavioural disturbances (American Psychiatric Association, 2013). With a median onset age of 8 years, SP has a global lifetime prevalence rate of 7.4%, occurring higher in females (9.8%) than in males (4.9%), and often persisting into adulthood (Becker et al., 2007). Moreover, studies show that SP is more pervasive in higher-socioeconomic countries such as the United States of America, New Zealand and France, relative to lower-socioeconomic countries (Wardenaar et al., 2017).

SP is a pervasive condition associated with impairments to individual functioning and quality of

CONTACT Karen Moses K.Moses@westernsydney.edu.au

life (Depla et al., 2008; Lieb et al., 2016; Wardenaar et al., 2017). Surveys have shown that up to 59.2% and 43.8% of individuals with the condition report significant interference in their daily life and social functioning, respectively (Depla et al., 2008). SP commonly co-occurs with a range of other mental disorders (Magee et al., 1996), with some researchers suggesting it often precedes the development of depressive, anxiety and eating disorders (Lieb et al., 2016; Trumpf et al., 2010). Because of its high prevalence, associated impairment and frequent co-morbidity rates with other disorders (Wardenaar et al., 2017), accurate assessment and diagnosis of SP is critical from both a research and evidence-based treatments perspective (Avasthi et al., 2014; Moses et al., 2020; Wardenaar et al., 2017).

The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) is a classification system that encapsulates evidence-based criteria for defining various mental disorders (American Psychiatric Association, 2013). Current advances in the DSM-5 encourage clinicians and researchers to supplement the traditional categorical classification of disorders with dimensional ratings of severity (American Psychiatric Association, 2013; Kamphuis & Noordhof, 2009; R. LeBeau et al., 2015). Such dimensional assessments capture information unattainable by the traditional categorical classification system, including disorder severity, subclinical presentations of disorders and changes in symptom presentation over time (through repeated measure; R. T. Lebeau et al., 2012). Dimensional assessments also have the advantage of clarifying aspects of diagnostic comorbidity by identifying symptoms that persist across multiple mental disorders (R. T. Lebeau et al., 2012).

Self-report measures are a key component of evidence-based assessments of mental disorders (Moses et al., 2020). Dimensional assessments are one type of self-report measure that captures the presentation, severity and variation in disorder symptomatology (R. LeBeau et al., 2015; Moses et al., 2020). Such measures also maintain brevity and cost-effectiveness, representing an efficient means to assess and monitor symptoms (Moses et al., 2020). However, self-report measures can at times be inaccurate, encourage distortions of symptoms, lack theoretical basis and vary in the heterogeneity in format and content (Balon, 2005; R. T. Lebeau et al., 2012). Scale development procedures to minimise these threats to measurement validity include psychometric validation, simplified item wording and consistency in diagnostic assessment (e.g., emphasis on behaviours versus emphasis on symptoms; R. T. Lebeau et al., 2012). Despite the possibility of these limitations, research has shown that, when utilised as part of a multi-method, multi-information approach, such scientifically supported measures form an integral part of a comprehensive evidence-based assessment (Hunsley & Mash, 2005).

The Anxiety Disorders Subgroup of the DSM-5 Anxiety, Obsessive-Compulsive Spectrum, Post-traumatic and Dissociative Disorder Work Group developed the Specific Phobia Dimensional Scale (SP-D), a 10-item dimensional assessment of SP that correspond to the DSM-5 criteria (R. T. Lebeau et al., 2012). The psychometric properties of the SP-D were initially assessed in a sample of 57 undergraduate students, and 48 participants with a clinically significant anxiety disorder as their principal diagnosis (R. T. Lebeau et al., 2012). Results indicated excellent internal consistency ($\alpha = 0.93$), and low test–retest reliability (Intraclass Correlational Coefficient [ICC] = 0.51; R. T. Lebeau et al., 2012). Low test–retest reliability was hypothesised due to participant's inconsistent exposure to specific phobic stimulus, resulting in varied ratings (R. T. Lebeau et al., 2012). Furthermore, correlations with existing scales were not assessed due to lack of adequate existing measures, highlighting a need for further investigation. However, other Anxiety Dimensional Scales within the study showed good convergent validity ($r_s = 0.39$–0.69; R. T. Lebeau et al., 2012). As such, further evaluation of the SP-D was suggested.

The psychometric properties of the SP-D have been replicated in various languages and cultures (Beesdo-Baum & Knappe, 2012; DeSousa et al., 2017; Knappe et al., 2014; Moller & Bogels, 2016). Replication has occurred in a German student and clinical sample (Beesdo-Baum & Knappe, 2012; Knappe et al., 2014), a Dutch convenience sample (Moller & Bogels, 2016) and a Brazilian community sample (DeSousa et al., 2017). The SP-D was demonstrated to have a unidimensional structure (Beesdo-Baum & Knappe, 2012; DeSousa et al., 2017; Knappe et al., 2014), excellent internal consistency (Cronbach's α between 0.9 and 0.93; DeSousa et al., 2017; Knappe et al., 2014; Moller & Bogels, 2016) and questionable to adequate test–retest reliability (ICCs between 0.66 and 0.79; DeSousa et al., 2017; Knappe et al., 2014). Correlations of the SP-D with measures of associated convergent and divergent constructs gave evidence for convergent validity, however, not divergent validity (Beesdo-Baum & Knappe, 2012; Knappe et al., 2014; Moller & Bogels, 2016). The scales sensitivity was found to be low to adequate (Beesdo-Baum & Knappe, 2012; Moller & Bogels, 2016), and a cut-off score of 11 was suggested to identify clinically significant symptoms (Beesdo-Baum & Knappe, 2012).

Despite the promising psychometric properties of the SP-D, there are several limitations that require addressing. First, studies validating the SP-D have been conducted within restricted cultures (e.g., United States, Germany, Brazil and the Netherlands), and constrained samples (e.g., predominantly young, well-educated and single; DeSousa et al., 2017; Knappe et al., 2014; Moller & Bogels, 2016), limiting the capacity to generalise the results. As epidemiological factors of SP are known to vary across culture, country and demographics (American Psychiatric Association, 2013; Wardenaar et al., 2017), it is important to validate the SP-D in more diverse samples (R. T. Lebeau et al., 2012; Moller & Bogels, 2016). Second, there is limited data confirming whether the SP-D captures symptomatology theoretically distinct from other mental health disorders (i.e., divergent validity; DeSousa et al., 2017; R. T. Lebeau et al., 2012). Therefore, correlations of the SP-D with theoretically distinct measures is required to confirm convergent and divergent validity (Carlson & Herdman, 2012).

The aim of the current study was to investigate the psychometric properties of the SP-D within an Australian community sample. Specifically, the current study aims to investigate the measures (1) factor structure; (2) internal consistency; (3) test–retest reliability and (4) convergent and divergent validity. It is hypothesised that the SP-D will demonstrate results consistent with previous studies (Beesdo-Baum & Knappe, 2012; DeSousa et al., 2017; Knappe et al., 2014; R. T. Lebeau et al., 2012; Moller & Bogels, 2016), demonstrating a unidimensional structure and adequate reliability. In terms of validity, consistent with previous research (e.g., DeSousa et al., 2017; R. T. Lebeau et al., 2012), we hypothesise that the SP-D will demonstrate good convergent validity, demonstrating significant and large correlations with the convergent validity measure, but poor divergent validity, demonstrating significant and medium-to-large correlations with the divergent validity measure.

Materials and methods

Participants

A total of 398 participants gave consent to commence this study. The only inclusion criteria were that participants were required participants to be at least 18 years of age and living in Australia. Participants who did not meet inclusion criteria or did not provide information to assess the eligibility criteria ($N = 58$) were excluded before any data was collected. These participants were directed to an exit page that thanked them for their time. Those who met eligibility criteria but had incomplete data ($N = 55$) were excluded from all analyses. Two hundred and eighty-five participants were included in Part 1. Table 1 outlines the demographic characteristics of the sample. Participants in Part 1 were predominantly female (74%) aged 18–76 years ($M = 28.15$, $SD = 12.01$). Eighteen participants complete Part 2 of the study to examine test–retest reliability.

Table 1. Descriptive statistics for participants ($N = 285$).

Variable	*n*	%
Gender		
Male	72	25.3
Female	211	74.0
Other/prefer not to say	2	0.7
Country of Origin		
Australia	221	77.5
New Zealand	2	.7
Asia	11	3.9
Europe	4	1.4
United Kingdom	5	1.8
North America	3	1.1
South America	3	1.1
Middle East	9	3.2
Africa	2	.7
Other	25	8.8
Occupation Status		
Working part time	80	28.1
Working full time	78	27.4
Unemployed	12	4.2
Studying	99	34.7
Retired	2	.7
Full time carer	3	1.1
Other	11	3.9

Procedure

Data from this study was obtained as part of a larger program of research aimed at investigating the psychometric properties of several DSM-5 dimensional scales. Participant recruitment occurred via multiple sources, including posts on social media, advertisements on community notice boards and through email. Recruitment sources were not monitored, and each included an anonymous link to the study questionnaire. Participant engagement occurred over two parts, Part 1 and Part 2. Part 1 was an online questionnaire with an approximate completion time of 20 min. Participants were tasked with viewing an information sheet and providing informed consent before completing a demographic questionnaire, the SP-D, the SPQ and the FQ-Ag. Following completion of Part 1, participants were invited to volunteer for Part 2 of the study (test–retest reliability measure). Participants who agreed to commence Part 2 were asked to develop a unique identification code to anonymously match their responses from Part 1 and Part 2. These participants were emailed a link, 2-weeks following Part 1 and were tasked with re-completing the study

questionnaires. The study was approved by the Human Research and Ethics Committee of Western Sydney University (Approval number: H13180).

Measures

The Specific Phobia Dimensional Scale (SP-D; R. T. Lebeau et al., 2012)

The SP-D is a 10-item self-report measure examining the severity of specific phobia symptoms over the previous month (R. T. Lebeau et al., 2012). The 10-items assess cognitive and physical symptoms of fear and anxiety (e.g., "felt anxious, worried, or nervous about these situations?"), frequency of escape and avoidance behaviours (e.g., "avoid, or did not approach or enter, these situations?"), subjective emotional component of panic (e.g., "terror") and frequency of cognitive avoidance (e.g., distraction) from anxiety-producing situations (R. T. Lebeau et al., 2012). Each item is rated on a five-point Likert scale, ranging from zero ("none" or "never") to four ("extreme" or "all the time"). The SP-D has previously demonstrated adequate internal consistency ($a > 0.9$; Beesdo-Baum & Knappe, 2012; DeSousa et al., 2017; Knappe et al., 2014; R. T. Lebeau et al., 2012; Moller & Bogels, 2016). In the current sample, the SP-D had a Cronbach's a of 0.95.

Specific Phobia Questionnaire (SPQ; Ovanessian et al., 2019)

The SPQ is a 43-item self-report measure assessing the extent of fear and interference for a broad range of objects and situations (Ovanessian et al., 2019). Each item is rated on a five-point Likert scale on two dimensions: (1) level of fear, ranging from one (no fear) to five (extreme fear) and (2) extent of daily life interference, ranging from one (no interference) to five (extreme interference). A total SPQ score is generated by summing the two dimensions of fear and interference. The SPQ has previously demonstrated acceptable internal consistency (a between 0.64 and 0.92; Ovanessian et al., 2019). The SPQ was used to assess convergent validity due to its strong association with various other phobia questionnaires (Ovanessian et al., 2019). In the current sample, the SPQ had a Cronbach's a of 0.97.

Fear Questionnaire – Agoraphobia subscale (FQ-Ag; Marks & Mathews, 1979)

The FQ-Ag (Marks & Mathews, 1979) is a five-item self-report measure assessing behavioural disturbances related to agoraphobic inducing situations. On a 9-point Likert scale, each item is rated from zero ("would not avoid it") to eight ("always avoid it"). The scale has previously demonstrated acceptable internal consistency ($a = 0.76$; Cox et al., 1991). The FQ-A was used to measure divergent validity as SP-D symptoms are hypothesised to differ from symptoms of agoraphobia. However, it is acknowledged that given the high degree of comorbidity between the anxiety disorders (Kroenke et al., 2007) the relationships between these measures may be in the medium in size, rather than small range. In the current sample, the FQ-Ag has a Cronbach's a of 0.85.

Statistical analyses

All analyses were performed using IBM SPSS statistics Version 28.0.1 and IBM SPSS Amos Version 26. Assumption testing demonstrated positive skewness and multivariate nonnormality. Nonparametric tests were utilised accordingly. Sample size was considered sufficient (Lumley et al., 2002). Demographic characteristics of participants who completed the questionnaire ($N = 285$) were compared with non-completers ($N = 55$) using independent samples t-tests for continuous outcomes and chi-square tests for categorical outcomes.

Resembling previous studies, confirmatory factor analysis (CFA) was used to examine factor structure of data collected in Part 1 of this study (DeSousa et al., 2017; R. T. Lebeau et al., 2012). For fit indices, the study utilised the comparative fit index (CFI), the Tucker-Lewis Index (TLI), the standardised root mean square residual (SRMR) and the root mean square error of approximation (RMSEA). CFI and TLI values equal to or higher than 0.9 represent an acceptable fit and higher than 0.95 represent a good fit (Hu & Bentler, 1999). SRMR and RMSEA values equal to or less than 0.08 represent an acceptable fit and less than 0.05 represent a good fit (Hu & Bentler, 1999). Model parameters were estimated using the SRMR statistics due to evidence of its robustness (Shi & Maydeu-Olivares, 2020). A multigroup CFA was conducted to examine measurement invariance for gender. In this analysis, those reporting "other" gender were excluded and thus 211 females and 72 males were included in this analysis. Consistent with Cheung and Rensvold (2002), we used change in CFI to assess invariance with a CFI change equal to or greater than 0.01 suggesting differences between the groups.

The internal consistency of the SP-D in both samples (Part 1 and Part 2) was examined with Cronbach's alpha coefficients. Alpha values greater than 0.70 were considered adequate (Onwuegbuzie & Daniel, 2002). Test–retest reliability was examined by calculating the ICC between the total score on the SP-D in Part 1 and Part 2 (Koo & Li, 2016). In line with DeSousa et al.

(2017), ICCs were calculated in SPSS using Two-Way Mixed Effect Model and Absolute Agreement Type (confidence interval 95%). ICC estimates greater than 0.70 were considered adequate (Murphy & Davidshofer, 1988). Due to non-normally distributed data, convergent and divergent validity were examined with the non-parametric Spearman's rho, correlating the SP-D with the SPQ, and FQ-Ag. Correlation strength effect sizes were interpreted where 0.10 is small, 0.30 is medium and 0.50 is large (Cohen, 1992).

Results

Comparison of completers and non-completers

Participants who completed the study ($N = 285$) were compared with participants who did not complete the study ($N = 55$) on key demographic variables when available. Participants who completed the questionnaire were significantly younger ($M = 28.15$) than those who did not complete the questionnaire ($M = 34.60$) ($t_{(68.46)} = 3.05$, $p = .003$). There were no significant differences between the groups on gender (X^2 (2) = .466, $p > .05$).

Descriptive statistics

Descriptive statistics are outlined in Table 2. On the SP-D (Part 1), 53% ($n = 152$) scored between 0 and 10, 23% ($n = 65$) scored between 11 and 20, 16% ($n = 47$) scored between 21 and 30 and 8% ($n = 21$) scored between 31 and 40. One hundred and thirty-three (47%) of the sample met criteria for SP, based on cut score of 11 identified by Beesdo-Baum and Knappe (2012). Participants indicated the following categories of feared situation/objects: driving, flying, tunnels, bridges, or enclosed spaces ($n = 93$; 32.6%); animals or insects ($n = 65$; 22.8%); heights, storms, or water ($n = 52$; 18.2%); blood, needles or injections ($n = 43$; 15.1%); and choking or vomiting ($n = 32$; 11.2%).

Factor structure

Initial CFA results demonstrated good SRMR, but unacceptable CFI, TLI and RMSEA: x^2 (35) = 371.60, $p < 0.001$, CFI = 0.88, TLI = 0.84, SRMR = 0.05, RMSEA = 0.18. Post hoc analysis investigating modification indices for improving the model fit indicated strong local dependency between item sets 4 and 5 (assessing physiological anxiety symptoms) and 6 and 7 (assessing avoidance and escape behaviours related to anxiety symptoms). In line with methodology of DeSousa et al. (2017), correlations between the two sets of error terms were added (item 4 and 5 and item 6 and 7), and a new CFA was conducted. The new model's fit indices revealed adequate CFI and TLI, good SRMR, but unacceptable RMSEA: x^2 (33) = 186.76, $p < 0.001$, CFI = 0.94, TLI = 0.92, SRMR = 0.04, RMSEA = 0.13. All items loaded significantly on the single factor. Table 3 shows unstandardised and standardised factor loadings and squared multiple correlations. The correlations between SP-D items are outlined in Table 4. As outlined in Table 5, the constrained and unconstrained models resulted in fit indices that were very similar to the CFA and thus there is evidence of measurement invariance of the SP-D for males and females.

Validity

The SP-D exhibited a large positive correlation with the SPQ Total ($r_s = 0.54$, $p < 0.001$), and a medium positive correlation with the FQ-Ag ($r_s = 0.48$, $p < 0.001$).

Table 2. Descriptive statistics for measures.

Measure	*M (SD)*	Median	Observed range	Possible range
SP-D Total	12.46 (10.88)	10.00	0–40	0–40
Item 1	1.21 (1.20)	1.00	0–4	0–4
Item 2	1.44 (1.21)	1.00	0–4	0–4
Item 3	1.35 (1.24)	1.00	0–4	0–4
Item 4	1.2 (1.33)	1.00	0–4	0–4
Item 5	1.33 (1.31)	1.00	0–4	0–4
Item 6	1.46 (1.40)	1.00	0–4	0–4
Item 7	1.27 (1.35)	1.00	0–4	0–4
Item 8	1.13 (1.31)	1.00	0–4	0–4
Item 9	1.36 (1.38)	1.00	0–4	0–4
Item 10	.70 (1.17)	1.00	0–4	0–4
SPQ Fear	80.84 (29.22)	74.00	43-215	43-215
SPQ Interference	67.48 (21.76)	63.00	43-159	43-215
SPQ Total	148.33 (46.58)	138.00	86-311	86-430
FQ-Ag Total	8.5 (8.95)	6.00	0–40	0–40

SP-D = Specific Phobia Dimensional Scale; SPQ Fear = Specific Phobia Questionnaire – Fear Dimension; SPQ Interference = Specific Phobia Questionnaire – Interference Dimension; SPQ Total = Specific Phobia Questionnaire – Total; FQ-Ag = Fear Questionnaire – Agoraphobia subscale.

Table 3. Factor loadings and squared multiple correlations for SP-D.

Item	Unstandardised	Standardised	Squared multiple correlations
Item 1	1.00**	0.85**	0.47
Item 2	1.04**	0.87**	0.70
Item 3	1.01**	0.83**	0.66
Item 4	1.14**	0.88**	0.58
Item 5	1.10**	0.86**	0.64
Item 6	1.10**	0.80**	0.74
Item 7	1.00**	0.76**	0.77
Item 8	1.04**	0.81**	0.70
Item 9	1.13**	0.84**	0.76
Item 10	0.78**	0.69**	0.73

**Significant at < 0.001.

Table 4. Correlations between SP-D items ($N = 285$).

	Item									
	1	2	3	4	5	6	7	8	9	10
SP-D Item 1	–									
SP-D Item 2	0.57	–								
SP-D Item 3	0.56	0.68	–							
SP-D Item 4	0.52	0.64	0.62	–						
SP-D Item 5	0.55	0.67	0.65	0.85	–					
SP-D Item 6	0.59	0.72	0.70	0.65	0.68	–				
SP-D Item 7	0.60	0.74	0.72	0.67	0.70	0.87	–			
SP-D Item 8	0.57	0.70	0.68	0.63	0.67	0.71	0.73	–		
SP-D Item 9	0.60	0.73	0.71	0.67	0.70	0.75	0.77	0.73	–	
SP-D Item 10	0.58	0.71	0.70	0.65	0.68	0.73	0.75	0.71	0.75	–

All correlations significant at < 0.001.
SP-D = Specific Phobia Dimensional Scale.

Reliability

Cronbach's *a* for SP-D was 0.95 and 0.97 at Part 1 and Part 2, respectively. Mean scores for the SP-D total scores and item scores for Part 1 and Part 2 administration are outlined in Table 5. A strong positive correlation between total SP-D scores from Part 1 and Part 2 ($r_s = 0.95$) was found. All items of the SP-D indicated excellent test–retest reliability (ICCs between 0.74 and 0.90, $p < 0.05$), apart from Item 1, which represented moderate test–retest reliability (0.56). All item scores and intraclass coefficients are outlined in Table 6.

Table 5. Fit indices from multigroup confirmatory factor analysis.

Gender (male x female)	X^2 (df)	*p* – value	CFI	TLI	RMSEA (90%CI)	ΔCFI
Unconstrained (configural invariance)	247.06 (75)	.000	.94	.93	.09 (.08 – .10)	
Measurement weights (metric invariance)	247.06 (75)	.000	.94	.93	.09 (.08 – .10)	.00
Structural covariances (scalar invariance)	262.68 (86)	.000	.94	.93	.09 (.07 – .10)	.00

Table 6. SP-D item scores and intraclass correlation coefficients at Part 1 and Part 2 ($N = 18$).

	Part 1		Part 2		Statistic
Item	*M*	*SD*	*M*	*SD*	*r*
1	1.33	1.38	0.83	0.99	0.56*
2	1.50	1.25	0.83	0.99	0.74**
3	1.11	1.23	1.00	1.24	0.83**
4	1.00	1.19	1.00	1.50	0.90**
5	0.89	1.08	1.00	1.37	0.90**
6	1.22	1.35	1.17	1.58	0.88**
7	1.05	1.51	0.94	1.11	0.88**
8	1.17	1.42	0.67	1.14	0.80**
9	1.06	1.40	0.94	1.30	0.84**
10	1.44	0.86	0.50	1.20	0.83**
TOTAL	10.78	11.18	8.72	10.89	0.95**

**Significant as < 0.001; *Significant at < 0.05.

Discussion

The aim of this study was to investigate the psychometric properties of the SP-D within an Australian community sample. The study aimed to replicate existing literature on the SP-D by examining factor structure, internal consistency, test–retest reliability and convergent and divergent validity (R. T. Lebeau et al., 2012). Overall, results of this study indicated that the SP-D is a promising self-report measure of symptoms of specific phobia. The results are largely consistent with previous studies examining the psychometric properties of the SP-D (Beesdo-Baum & Knappe, 2012; DeSousa et al., 2017; Knappe et al., 2014; R. T. Lebeau et al., 2012; Moller & Bogels, 2016). Divergent validity of the SP-D remains questionable (Beesdo-Baum & Knappe, 2012; Knappe et al., 2014; Moller & Bogels, 2016).

Results from the current study suggest that the SP-D exhibits a unidimensional structure. This finding is consistent with previous research assessing the scales underlying factor structure (Beesdo-Baum & Knappe, 2012; DeSousa et al., 2017; Knappe et al., 2014). All SP-D items loaded significantly onto the single factor. There was evidence of high local dependency between items 4 and 5, both measuring physiological anxiety symptoms, and 6 and 7, both measuring escape and avoidance behaviours in response to anxiety symptoms, suggesting each item set may be measuring a similar concept. As shown by DeSousa et al. (2017), and the present study, there was an improvement in model fit following model adjustment to acknowledge the two local dependencies. Also, like DeSousa et al. (2017), this study failed to find acceptable RMSEA for the SP-D, suggesting further refinement of the scale may be required (i.e., adding new items or eliminating some existing items). However, it is also noteworthy that in simple models with fewer degrees of freedom the RMSEA can indicate poor fit to the data despite the model fitting well (Kenny et al., 2015), which may be the case in the current study. The results also indicated measurement invariance based on gender, however, this is only for those who identify as male or female, and thus measurement invariance in those with non-binary or other gender identities requires further examination in future research.

The SP-D demonstrated adequate internal consistency in the current sample, which is consistent with previous findings (Beesdo-Baum & Knappe, 2012; DeSousa et al., 2017; Knappe et al., 2014; R. T. Lebeau et al., 2012; Moller & Bogels, 2016). The current study also demonstrated adequate test–retest reliability of the SP-D for all items except for Item 1, which demonstrated a moderate test–retest reliability. Adequate repeat measure reliability is consistent with previous research (DeSousa et al., 2017; Knappe et al., 2014; R. T. Lebeau et al., 2012). Interestingly, this study demonstrated the strongest test–retest reliability when compared to previous investigations (DeSousa et al., 2017; Knappe et al., 2014; R. T. Lebeau et al., 2012). It is worth noting, however, that the sample of current study was considerably smaller than those of past investigations. It would be valuable for future investigations to replicate these findings with a larger sample size. Nevertheless, these results provide further support of the SP-D's ability to accurately replicate results over repeated administrations.

In assessing convergent and divergent validity, the SP-D in the current study showed a large, positive correlation with the theoretically converging SP-Q scale (Ovanessian et al., 2019), and a moderate, positive correlation with the diverging FQ-Ag scale (Marks & Mathews, 1979). These results suggest evidence of convergent validity, given the relationship between the SP-D and SP-Q was significant and a large effect size was observed. Divergent validity, however, could not be confirmed at this time due to the significant correlation (medium to large effect size) between the SP-Q and the divergent validity measure, which indicates that there may be substantial overlap with symptoms of agoraphobia. These outcomes are consistent with previous research, where convergent validity was supported, and divergent validity was not supported (Beesdo-Baum & Knappe, 2012; Knappe et al., 2014; Moller & Bogels, 2016). Previous researchers suggested these outcomes may be due to the high comorbidity present amongst anxiety disorders (Moller & Bogels, 2016), or possibly because of the development of the SP-D scale, where outcomes are influenced by the imprecise introduction and/or formulation of scale items (Knappe et al., 2014). Future studies may therefore wish to investigate the SP-D's divergent validity with scales theoretically distinct from the DSM-5's anxiety disorder cluster, or alternatively assess comorbidity in the sample to ascertain if there is significant comorbidity between the sample in terms of the divergent measure.

The current study primarily assessed the SP-D within an Australian identifying sample, which is the first of its kind. The sample characteristics were largely characterised by participants who were employed and had attained an education level equivalent to a high school certificate. Employment and education statistics from the current sample reflect similar percentages to the Australian population (Australian Bureau of Statistics, [ABS] 2021). However, the study utilised a female-biased sample, impacting generalisability of

the findings (Acharya et al., 2013). Nevertheless, the use of a female-biased sample remains consistent with previous investigations of dimensional scales (Beesdo-Baum & Knappe, 2012; DeSousa et al., 2017; Knappe et al., 2014; R. T. Lebeau et al., 2012, 2016 Macfarlane et al., 2020; Russell et al., 2020). Further research with greater male and diverse gender identities is required.

The observed mean score on the SP-D in this Australian community sample ($M = 12.46$, $SD = 10.88$) was greater than previous investigations using non-clinical (e.g., $M = 6.49$, $SD = 7.75$ in DeSousa et al., 2017; $M = 3.2$, $SD = 5.2$ in Knappe et al., 2014) and clinical ($M = 8.2$, $SD = 9$ in Knappe et al., 2014) samples. Furthermore, almost half of the participants in the study scored greater than the suggested clinically relevant cut off score of 11 (Beesdo-Baum & Knappe, 2012). This outcome may be explained in a number of ways. As this study took place during the Coronavirus (COVID-19) global pandemic, where Australian adults reported increased symptoms of psychological fear and distress (Fisher et al., 2020; Rahman et al., 2020; Rossell et al., 2021), it is possible that heightened symptoms of fear and distress subsequently led to greater SP-D mean scores. Alternatively, our sample may have been biased through participant self-selection due to personal experience with anxiety symptomology (Heckman, 1990). It is also possible that a cut score of 11 may be too low in the Australian context. As such, our results should be interpreted with these factors in mind and future research in the Australian context is required.

While the findings of the current study offer further support for the psychometric properties of the SP-D, there remain several limitations worth addressing. First, the current study did not investigate the SP-D's psychometric properties within an Australian clinical sample. While more than half the sample indicated symptoms above the cut-score on the SP-D it is important that future studies use a diagnostic interview to confirm diagnostic status of participants. Psychometric validation of the SP-D within a clinical sample may be a goal of future studies. Second, the SP-D's sensitivity and specificity regarding diagnosis of SP, and its sensitivity to treatment response, is yet to be tested in the Australian population. Such investigations have proven useful in a German clinical sample (Knappe et al., 2014). Third, the duration between Part 1 and Part 2 of the current study was not monitored. As symptoms of SP are shown to vary across time, particularly with targeted intervention and recency of exposure to feared stimulus (Barlow, 2002; Wardenaar et al., 2017), future studies should control the duration between administrations. Fourth, the sample size investigating the SP-D's test–retest reliability in our study was small ($n = 18$) and these analyses were thus underpowered (Lumley et al., 2002). Decreased motivation to complete online studies due to situational demand and anonymity may account for participant dropout (Dandurand et al., 2008). Future studies may minimise dropout rates by providing financial incentive, or informing participants the theoretical importance of the research (Clifford & Jerit, 2015; Crump et al., 2013). Fifth, the chosen measure of divergent validity (the FQ-Agoraphobia) may not have been the most appropriate measure given the high comorbidity between anxiety disorders (Kroenke et al., 2007). Future research may wish to re-examine divergent validity of the SP-D in different samples. Lastly, recruitment sources for the study were not monitored. Future studies engaging with a convenience community sample may benefit from monitoring recruitment sources to discover broader information on participant demographic composition. Such information could assist in further validating the SP-D in wider contexts (e.g., First Nations Australians relative to Anglo Australians), which is an area of particular need in Australia (Westerman & Dear, 2023).

In conclusion, this was the first study to investigate the psychometric properties of the SP-D in an Australian sample. The current study adds to the existing literature on the SP-D, and provides preliminary evidence that it is a reliable, and valid tool to dimensionally assess SP symptomology in Australian adults. In addition, this study highlighted certain areas that require further investigation, particularly divergent validity, and the psychometric validation of the measure using an Australian clinical sample. In doing so, further research can aim to strengthen the literature surrounding the SP-D, giving clinicians and researchers increased confidence in its utility within Australia, and across varying cultures.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Bethany M. Wootton http://orcid.org/0000-0001-9036-0699
Karen Moses http://orcid.org/0000-0003-0968-5432

Data availability statement

Data available from the corresponding author upon reasonable request and ethical approval.

References

Acharya, A. S., Prakash, A., Saxena, P., & Nigam, A. (2013). Sampling: Why and how of it. *Indian Journal of Medical Specialties*, *4*(2), 330–333. https://doi.org/10.7713/ijms.2013.0032

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders DSM-5* (5th ed.).

Australian Bureau of Statistics. (2021). *Education and work, Australia*. https://www.abs.gov.au/statistics/people/education/education-and-work-australia/may-2021

Avasthi, A., Sarkar, S., & Grover, S. (2014). Approaches to psychiatric nosology: A viewpoint. *Indian Journal of Psychiatry*, *56*(3), 301–304. https://doi.org/10.4103/0019-5545.120560

Balon, R. (2005). Measuring anxiety: Are we getting what we need? *Depression and Anxiety*, *22*(1), 1–10. https://doi.org/10.1002/da.20077

Barlow, D. H. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed.). Guilford Press.

Becker, E. S., Rinck, M., Türke, V., Kause, P., Goodwin, R., Neumer, S., & Margraf, J. (2007). Epidemiology of specific phobia subtypes: Findings from the Dresden mental health study. *European Psychiatry*, *22*(2), 69–74. https://doi.org/10.1016/j.eurpsy.2006.09.006

Beesdo-Baum, K. P., & Knappe, S. P. (2012). Developmental epidemiology of anxiety disorders. *Child and Adolescent Psychiatric Clinics of North America*, *21*(3), 457–478. https://doi.org/10.1016/j.chc.2012.05.001

Carlson, K. D., & Herdman, A. O. (2012). Understanding the impact of convergent validity on research results. *Organizational Research Methods*, *15*(1), 17–32. https://doi.org/10.1177/1094428110392383

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Clifford, S., & Jerit, J. (2015). Do attempts to improve respondent attention increase social desirability bias? *Public Opinion Quarterly*, *79*(3), 790–802. https://doi.org/10.1093/poq/nfv027

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin*, *112*, 1155–1159.

Cox, B. J., Swinson, R. P., & Shaw, B. F. (1991). Value of the fear questionnaire in differentiating agoraphobia and social phobia. *British Journal of Psychiatry*, *159*(6), 842–845.

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLOS ONE*, *8*(3), e57410. https://doi.org/10.1371/journal.pone.0057410

Dandurand, F., Shultz, T. R., & Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods*, *40*(2), 428–434. https://doi.org/10.3758/BRM.40.2.428

Depla, M. F., Ten Have, M. L., van Balkom, A. J., & de Graaf, R. (2008). Specific fears and phobias in the general population: Results from the Netherlands mental health survey and incidence study (NEMESIS). *Social Psychiatry & Psychiatric Epidemiology*, *43*(3), 200–208. https://doi.org/10.1007/s00127-007-0291-z

DeSousa, D. A., Moreno, A. L., Osório, F. L., Crippa, J. A. S., LeBeau, R., Manfro, G. G., Salum, G. A., & Koller, S. H. (2017). Psychometric properties of the dimensional anxiety scales for DSM-5 in a Brazilian community sample. *International Journal of Methods in Psychiatric Research*, *26*(3), n/a. https://doi.org/10.1002/mpr.1531

Fisher, J. R., Tran, T. D., Hammarberg, K., Sastry, J., Nguyen, H., Rowe, H., Popplestone, S., Stocker, R., Stubber, C., & Kirkman, M. (2020). Mental health of people in Australia in the first month of COVID-19 restrictions: A national survey. *The Medical Journal of Australia*, *213*(10), 458–464. https://doi.org/10.5694/mja2.50831

Heckman, J. J. (1990). Selection bias and self-selection. In *Econometrics* (pp. 201–224). Palgrave Macmillan UK.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hunsley, J., & Mash, E. J. (2005). Introduction to the special section on developing guidelines for the evidence-based assessment (EBA) of adult disorders. *Psychological Assessment*, *17*(3), 251. https://doi.org/10.1037/1040-3590.17.3.251

Kamphuis, J. H., & Noordhof, A. (2009). On categorical diagnoses in DSM-V: Cutting dimensions at useful points? *Psychological Assessment*, *21*(3), 294–301. https://doi.org/10.1037/a0016697

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*(3), 486–507. https://doi.org/10.1177/0049124114543236

Knappe, S., Klotsche, J., Heyde, F., Hiob, S., Siegert, J., Hoyer, J., Strobel, A., LeBeau, R. T., Craske, M. G., Wittchen, H.-U., & Beesdo-Baum, K. (2014). Test–retest reliability and sensitivity to change of the dimensional anxiety scales for DSM-5. *CNS Spectrums*, *19*(3), 256–267. https://doi.org/10.1017/S1092852913000710

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection [article]. *Annals of Internal Medicine*, *146*(5), 317–325. https://doi.org/10.7326/0003-4819-146-5-200703060-00004

LeBeau, R., Bögels, S., Möller, E., & Craske, M. (2015). Integrating dimensional assessment and categorical diagnosis in DSM-5: The benefits and challenges of the paradigm shift for the anxiety disorders. *Psychopathology Review*, *2*(1), 83–99. https://doi.org/10.5127/pr.036414

Lebeau, R. T., Glenn, D. E., Hanover, L. N., Beesdo-Baum, K., Wittchen, H.-U., & Craske, M. G. (2012). A dimensional approach to measuring anxiety for DSM-5: Dimensional measurement of anxiety for DSM-5. *International Journal of Methods in Psychiatric Research*, *21*(4), 258–272. https://doi.org/10.1002/mpr.1369

LeBeau, R. T., Mesri, B., & Craske, M. G. (2016). The DSM-5 social anxiety disorder severity scale: Evidence of validity and reliability in a clinical sample. *Psychiatry Research*, *244*, 94–96. https://doi.org/10.1016/j.psychres.2016.07.024

Lieb, R., Miché, M., Gloster, A. T., Beesdo-Baum, K., Meyer, A. H., & Wittchen, H. U. (2016). Impact of specific phobia on the risk of onset of mental disorders: A 10-year prospective-longitudinal community study of adolescents and young adults. *Depression and Anxiety*, *33*(7), 667–675. https://doi.org/10.1002/da.22487

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, *23*(1), 151–169. https://doi.org/10.1146/annurev.publhealth.23.100901.140546

Macfarlane, F., Luo, A., Moses, K., Russell, A., Cheyne, J., de Bolger, A. D. P., & Wootton, B. M. (2020). Psychometric properties of the body dysmorphic disorder-dimensional scale [article]. *The Clinical Psychologist*, *24*(2), 114–122. https://doi.org/10.1111/cp.12200

Magee, W. J., Eaton, W. W., Wittchen, H.-U., McGonagle, K. A., & Kessler, R. C. (1996). Agoraphobia, simple phobia, and social phobia in the national comorbidity survey. *Archives of General Psychiatry*, *53*(2), 159–168. https://doi.org/10.1001/archpsyc.1996.01830020077009

Marks, I. M., & Mathews, A. M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy*, *17*(3), 263–267. https://doi.org/10.1016/0005-7967(79)90041-X

Moller, E. L., & Bogels, S. M. (2016). The DSM-5 dimensional anxiety scales in a Dutch non-clinical sample: Psychometric properties including the adult separation anxiety disorder scale. *International Journal of Methods in Psychiatric Research*, *25*(3), 232–239. https://doi.org/10.1002/mpr.1515

Moses, K., Gayed, M., Chuah, S., & Wootton, B. M. (2020). The use of evidence-based assessment for anxiety disorders in an Australian sample. *Journal of Anxiety Disorders*, *75*, 102279. https://doi.org/10.1016/j.janxdis.2020.102279

Murphy, K. R., & Davidshofer, C. O. (1988). *Psychological testing. Principles, and applications.* Englewood Cliffs.

Onwuegbuzie, A. J., & Daniel, L. G. (2002). A framework for reporting and interpreting internal consistency reliability estimates. *Measurement and Evaluation in Counseling and Development*, *35*(2), 89–103. https://doi.org/10.1080/07481756.2002.12069052

Ovanessian, M. M., Fairbrother, N., Vorstenbosch, V., McCabe, R. E., Rowa, K., & Antony, M. M. (2019). Psychometric properties and clinical utility of the specific phobia questionnaire in an anxiety disorders sample. *Journal of Psychopathology & Behavioral Assessment*, *41*(1), 36–52. https://doi.org/10.1007/s10862-018-9687-1

Rahman, M. A., Hoque, N., Alif, S. M., Salehin, M., Islam, S. M. S., Banik, B., Sharif, A., Nazim, N. B., Sultana, F., & Cross, W. (2020). Factors associated with psychological distress, fear and coping strategies during the COVID-19 pandemic in Australia. *Globalization and Health*, *16*(1), 1–15. https://doi.org/10.1186/s12992-020-00624-w

Rossell, S. L., Neill, E., Phillipou, A., Tan, E. J., Toh, W. L., Van Rheenen, T. E., & Meyer, D. (2021). An overview of current mental health in the general population of Australia during the COVID-19 pandemic: Results from the COLLATE project. *Psychiatry Research*, *296*, 113660. https://doi.org/10.1016/j.psychres.2020.113660

Russell, A., Del Pozo de Bolger, A., Moses, K., Luo, A., & Wootton, B. M. (2020). Psychometric properties of the excoriation (skin-picking disorder) dimensional scale [article]. *The Clinical Psychologist*, *24*(3), 246–253. https://doi.org/10.1111/cp.12207

Shi, D., & Maydeu-Olivares, A. (2020). The effect of estimation methods on SEM fit indices. *Educational and Psychological Measurement*, *80*(3), 421–445. https://doi.org/10.1177/0013164419885164

Trumpf, J., Margraf, J., Vriends, N., Meyer, A. H., & Becker, E. S. (2010). Specific phobia predicts psychopathology in young women. *Social Psychiatry & Psychiatric Epidemiology*, *45*(12), 1161–1166. https://doi.org/10.1007/s00127-009-0159-5

Wardenaar, K. J., Lim, C. C. W., Al-Hamzawi, A. O., Alonso, J., Andrade, L. H., Benjet, C., Bunting, B., de Girolamo, G., Demyttenaere, K., Florescu, S. E., Gureje, O., Hisateru, T., Hu, C., Huang, Y., Karam, E., Kiejna, A., Lepine, J. P., Navarro-Mateu, F. … de Jonge, P. (2017). The cross-national epidemiology of specific phobia in the world mental health surveys. *Psychological Medicine*, *47*(10), 1744–1760. https://doi.org/10.1017/S0033291717000174

Westerman, T. G., & Dear, G. E. (2023). The need for culturally valid psychological assessment tools in indigenous mental health [article]. *The Clinical Psychologist*, *27*(3), 284–289. https://doi.org/10.1080/13284207.2023.2247532