Catalogue no. 82-003-X
ISSN 1209-1367

Health Reports

Evaluating the psychometric properties of the parent-rated Strengths and Difficulties Questionnaire in a nationally representative sample of Canadian children and adolescents aged 6 to 17 years



by Matt D. Hoffmann, Justin J. Lang, Michelle D. Guerrero, Jameason D. Cameron, Gary S. Goldfield, Heather M. Orpana and Margaret de Groh

Release date: August 19, 2020

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at STATCAN.infostats-infostats.STATCAN@canada.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© Her Majesty the Queen in Right of Canada as represented by the Minister of Industry, 2020

All rights reserved. Use of this publication is governed by the Statistics Canada Open Licence Agreement.

An HTML version is also available.

Cette publication est aussi disponible en français.

Evaluating the psychometric properties of the parent-rated Strengths and Difficulties Questionnaire in a nationally representative sample of Canadian children and adolescents aged 6 to 17 years

by Matt D. Hoffmann, Justin J. Lang, Michelle D. Guerrero, Jameason D. Cameron, Gary S. Goldfield, Heather M. Orpana and Margaret de Groh

Abstract

Background: The Strengths and Difficulties Questionnaire (SDQ) is a brief measure of children's and adolescents' mental health. There are different versions of the questionnaire: a version for children and adolescents to complete by self-reporting, a version for parents and guardians to complete ("parent-rated"), and a version for teachers to complete. The purpose of this study was to examine the psychometric properties of the parent-rated SDQ with a nationally representative sample of Canadian children and adolescents.

Data and methods: Data are from cycle 1 (2007 to 2009), cycle 2 (2009 to 2011), cycle 3 (2012 to 2013) and cycle 4 (2014 to 2015) of the Canadian Health Measures Survey. Data include 7,451 Canadian children and adolescents aged 6 to 17 years (49.3% female). Parents and guardians completed the SDQ by reflecting on their child's behaviour over the past six months. Factorial validity was examined via confirmatory factor analysis, which included testing the original five-factor SDQ model and alternative three-factor and higher-order models. Reliability was assessed through composite reliability scores. Measurement invariance across subgroups was also assessed.

Results: The original five-factor (i.e., emotional symptoms, conduct problems, peer problems, hyperactivity and prosocial behaviour) SDQ fit the data satisfactorily, demonstrated evidence of reliability, and was invariant across sex (male vs. female), age (children vs. adolescents) and survey language (English vs. French). The higher-order solution fit the data acceptably, and the three-factor solution did not fit the data well.

Interpretation: The original five-factor, parent-rated SDQ demonstrates evidence of factorial validity and reliability as a population measure of mental health difficulties among Canadian children and adolescents.

Keywords: mental health, problem behaviour, confirmatory factor analysis, measurement, youth
DOI: https://www.doi.org/10.25318/82-003-x202000800002-eng

Mental health difficulties are prevalent among children and adolescents across the world.[1] This is particularly important from a public health perspective because many psychological and behavioural disorders often begin in childhood.[2] The World Health Organization has cited poor mental health as one of the leading causes of disability and economic burden.[3] Between 1996 to 1997 and 2009 to 2010, the use of health care services for mental illness in Canada increased by 35% for children and 44% for adolescents. However, because not all children and adolescents with mental health difficulties access health care services,[4] estimates based on administrative data from health care services may underestimate the prevalence of mental health difficulties in these populations. Therefore, identifying brief measures of mental health in children and adolescents that demonstrate evidence of validity is critical to informing screening, population-level surveillance and prevention strategies.

The Strengths and Difficulties Questionnaire (SDQ) is a brief, widely used measure of children's and adolescents' social, emotional and behavioural difficulties,[5,6] or mental health difficulties.[7] The SDQ was designed to be completed by parents and teachers of children and adolescents aged 4 to 16, or by 11- to 16-year-olds themselves.[5,8] Researchers have extended this age range to include 17-year-olds.[9-11]

In Canada, the SDQ has been completed by parents or guardians of children and adolescents as part of the Canadian Health Measures Survey (CHMS). For the purposes of this study, the SDQ completed by parents or guardians as part of the CHMS is hereafter referred to as the parent-rated SDQ. Despite its inclusion in the CHMS, the psychometric properties of the parent-rated SDQ have yet to be evaluated with a nationally representative sample of Canadian children and adolescents.

The SDQ is a 25-item survey that consists of five subscales.[5] Four of the subscales (emotional symptoms, conduct problems, peer problems and hyperactivity) assess aspects of mental health difficulties. The fifth subscale, prosocial behaviour, represents one of the few available measures of positive mental health among children and adolescents. The results of several exploratory factor analyses have largely supported the five-factor structure of the parent-rated SDQ.[6,12,13] However, certain studies that used confirmatory factor analysis (CFA) have found mixed results concerning the factorial validity of the five-factor structure.[7,14-20]

Authors: Matt D. Hoffmann, Justin J. Lang (justin.lang@canada.ca), Heather M. Orpana and Margaret de Groh are with the Centre for Surveillance and Applied Research at the Public Health Agency of Canada, Ottawa, Ontario. Justin J. Lang is also with the Healthy Active Living and Obesity Research Group at the Children's Hospital of Eastern Ontario Research Institute, as are Michelle D. Guerrero and Gary S. Goldfield. Jameason D. Cameron is with the Department of Pharmacy at the Children's Hospital of Eastern Ontario. Gary S. Goldfield is also with the School of Human Kinetics and the School of Psychology at the University of Ottawa. Heather M. Orpana is also with the School of Epidemiology and Public Health at the University of Ottawa.

For instance, CFA studies using nationally representative parent-rated SDQ data have reported inadequate model fit statistics for the five-factor structure in Germany[19] and Ireland,[17] but have supported the five-factor structure in the United States[14] and (after including some correlated error terms) in Great Britain.[7] These mixed findings indicate that the factorial validity of the parent-rated SDQ may differ by country.

Building on previous research, the CFA study from Great Britain also tested an alternative, theoretically grounded three-factor structure that comprised "internalizing" (emotional symptoms and peer problems items), "externalizing" (conduct problems and hyperactivity items), and prosocial factors. This structure was found to have a generally poor model fit.[7] However, a similar model that involved higher-order internalizing and externalizing factors, alongside the prosocial factor, had a nearly identical model fit to the five-factor SDQ structure.[7] Goodman et al.[7] suggested that models that incorporate the broader internalizing and externalizing factors may be more appropriate than the five-factor structure for evaluating large samples of the general population that are at low-risk of mental health difficulties.

In Canada, researchers recently explored the factor structure of the parent-rated SDQ using a community sample of 501 children aged 6 to 9.[21] The results of the CFA indicated that the original five-factor SDQ model fit the data well. In general, the subscale scores had acceptable internal consistency when calculated using alpha coefficients, and strong internal consistency based on composite reliability scores. Although these results support the five-factor structure of the parent-rated SDQ with a Canadian sample,[21] researchers have yet to conduct a large-scale investigation of the factorial validity of this measure with Canadian children and adolescents.

To this end, the purpose of this study was to examine the psychometric properties of the parent-rated SDQ with a nationally representative sample of Canadian children and adolescents aged 6 to 17. Specific objectives included (1) assessing the factorial validity of the original five-factor SDQ structure and the Goodman et al.[7] first-order (three-factor) model and higher-order internalizing and externalizing model, (2) assessing the reliability of the SDQ's subscale scores; and (3) testing for measurement invariance across groups (i.e., male vs. female, children vs. adolescents, English questionnaire vs. French questionnaire).

Data and methods

Participants

The data used are from a subsample of participants aged 6 to 17 from cycle 1 (2007 to 2009), cycle 2 (2009 to 2011), cycle 3 (2012 to 2013) and cycle 4 (2014 to 2015) of the CHMS. The CHMS is an ongoing cross-sectional survey used to collect nationally representative health and wellness data on Canadians. CHMS data are collected from individuals living in households in the 10 provinces. The CHMS does not collect data from individuals living in the three territories, those living on reserves and Aboriginal settlements, full-time members of the Canadian Forces, institutionalized individuals, and individuals living in certain remote areas.[22] Individuals excluded from the CHMS represent roughly 4% of the target population.[22] Data collection for the CHMS consists of a household interview (demographic questions and a general health interview), followed by an in-person visit to a mobile examination centre for clinical and laboratory tests. The SDQ is completed as part of the household interview. A detailed overview of the CHMS sampling methodology and survey procedures is available elsewhere.[23] Sampling and bootstrap weights are provided by Statistics Canada.

Canadians aged 6 to 79 participated in cycle 1 of the CHMS, and Canadians aged 3 to 79 participated in cycles 2, 3 and 4. In total, 21,827 individuals (51.1% female) aged 6 to 79 participated in cycles 1 through 4 of the CHMS. The overall combined response rate for all four cycles for Canadians aged 6 to 79 was 53.1%.[24]

What is already known on this subject?

- The Strengths and Difficulties Questionnaire (SDQ) is widely used to assess children's and adolescents' mental health difficulties.
- Previous studies, including those that used nationally representative samples, have reported mixed results concerning the factorial validity of the original five-factor structure of the SDQ.

What does this study add?

- This is the first study to examine and support the factorial validity and reliability of the five-factor parent-rated SDQ in a large, nationally representative sample of Canadian children and adolescents.
- Configural, metric and scalar invariance across sex (male vs. female), age (children vs. adolescents) and survey language (English vs. French) was supported.

The SDQ was completed for children and adolescents aged 6 to 17 in cycles 1 and 2 of the CHMS, and for children and adolescents aged 4 to 17 in cycles 3 and 4. For consistency across cycles, the lower age limit for this study was 6 years. Therefore, the total sample in this study included 7,451 individuals aged 6 to 17 (49.3% female). This was a representative sample of children and adolescents based on Canadian demographics, and individuals were selected through a multistage stratified random sampling procedure.[22] Participants with missing data generally had no valid SDQ responses. Thus, the sample was further reduced using listwise deletion so that only individuals with complete data for at least one of the original five SDQ factors were retained. Further information regarding sample size retained for each of the five original SDQ factors is provided in Table 1.

All CFA-related analyses in this study included data from 6,960 individuals

Table 1
Descriptive statistics for the five Strengths and Difficulties Questionnaire factors, by sex and age

Factors	Mean (95% CI)	SEM
Total sample		
Emotional symptoms (n = 6,944)	1.88 (1.80, 1.96)	0.04
Conduct problems (n = 6,947)	1.06 (1.00, 1.12)	0.03
Peer problems (n = 6,890)	1.09 (1.03, 1.15)	0.03
Hyperactivity (n = 6,939)	2.77 (2.66, 2.88)	0.05
Prosocial behaviour (n = 6,925)	9.07 (9.01, 9.13)	0.03
Males		
Emotional symptoms (n = 3,527)	1.69 (1.58, 1.80)	0.05
Conduct problems (n = 3,531)	1.13 (1.05, 1.21)	0.04
Peer problems (n = 3,509)	1.18 (1.09, 1.27)	0.04
Hyperactivity (n = 3,524)	3.14 (2.97, 3.31)	0.09
Prosocial behaviour (n = 3,516)	8.85 (8.74, 8.96)	0.05
Females		
Emotional symptoms (n = 3,417)	2.08 (1.97, 2.19)	0.06
Conduct problems (n = 3,416)	0.98 (0.90, 1.06)	0.04
Peer problems (n = 3,381)	1.00 (0.92, 1.07)	0.04
Hyperactivity (n = 3,415)	2.36 (2.24, 2.49)	0.06
Prosocial behaviour (n = 3,409)	9.31 (9.26, 9.37)	0.03
Children (6 to 9 years)		
Emotional symptoms (n = 2,609)	1.78 (1.66, 1.90)	0.06
Conduct problems (n = 2,612)	1.24 (1.13, 1.34)	0.05
Peer problems (n = 2,588)	1.09 (0.97, 1.22)	0.06
Hyperactivity (n = 2,607)	3.24 (3.05, 3.43)	0.09
Prosocial behaviour (n = 2,606)	9.08 (8.97, 9.18)	0.05
Adolescents (10 to 17 years)		
Emotional symptoms (n = 4,335)	1.93 (1.83, 2.02)	0.05
Conduct problems (n = 4,335)	0.97 (0.89, 1.04)	0.04
Peer problems (n = 4,302)	1.09 (1.02, 1.17)	0.04
Hyperactivity (n = 4,332)	2.52 (2.38, 2.66)	0.07
Prosocial behaviour (n = 4,319)	9.06 (8.99, 9.14)	0.04
English SDQ		
Emotional symptoms (n = 5,290)	1.87 (1.77, 1.97)	0.05
Conduct problems (n = 5,293)	1.00 (0.93, 1.06)	0.03
Peer problems (n = 5,246)	1.08 (1.02, 1.14)	0.03
Hyperactivity (n = 5,287)	2.80 (2.68, 2.92)	0.06
Prosocial behaviour (n = 5,285)	9.10 (9.03, 9.16)	0.03
French SDQ		
Emotional symptoms (n = 1,600)	1.94 (1.77, 2.12)	0.09
Conduct problems (n = 1,598)	1.30 (1.16, 1.44)	0.07
Peer problems (n = 1,588)	1.17 (0.99, 1.34)	0.09
Hyperactivity (n = 1,599)	2.65 (2.42, 2.87)	0.11
Prosocial behaviour (n = 1,586)	8.95 (8.81, 9.08)	0.07

Notes: SDQ = Strengths and Difficulties Questionnaire. CI = confidence interval. SEM = standard error of the mean. Factor score range: 0 to 10.
Source: Canadian Health Measures Survey, cycles 1 to 4 (2007 to 2015).

(491 individuals were excluded, 6.6% of the total sample), with the exception of testing for survey language invariance, which comprised 6,904 individuals (547 individuals were excluded, 7.3% of the total sample). All CFA-related analyses used pairwise deletion to handle missing data.

Statistics Canada obtained ethics approval for the CHMS from the Health Canada and Public Health Agency of Canada Research Ethics Board. Participation in the CHMS was voluntary. Informed consent was provided by all respondents (i.e., parents or guardians of children and adolescents) prior to participating in the CHMS.[25]

Parent-rated Strengths and Difficulties Questionnaire

A parent or guardian answered all SDQ[5] items by reflecting on their child's behaviour over the past six months. As previously noted, the SDQ includes five subscales: emotional symptoms (five items), conduct problems (five items), peer problems (five items), hyperactivity (five items) and prosocial behaviour (five items). All items are scored on a three-point Likert-type scale: 0 (not true), 1 (somewhat true) and 2 (certainly true). Higher scores on the four difficulties subscales reflect greater difficulties in those areas, whereas higher scores on the prosocial behaviour subscale indicate a strength. Some items within the difficulties subscales are reverse-scored because of positive wording. For the Goodman et al.[7] alternative first-order and higher-order factor structures, SDQ items are allocated to the respective factors.

Statistical analyses

Descriptive statistics and polychoric correlations were calculated in SAS EG 7.1 (SAS Institute Inc., Cary, North Carolina, United States). All analyses in SAS incorporated sampling weights to account for the CHMS's complex survey design. Polychoric correlations were calculated using normalized weights. Bootstrap weights were used to estimate variance, using the balanced repeated replication method, with degrees of freedom set to 46.[24]

CFAs were conducted using the latent variable modelling program Mplus, version 7 (Muthén & Muthén, Los Angeles, California, United States). Because CFA is a construct-confirming technique, analyses were conducted using unweighted data.[26,27] The weighted least squares mean and variance-adjusted estimator was used, which is appropriate for ordered categorical (ordinal) variables, whereby normality is not assumed.[28] The following fit indices were used to evaluate model fit: comparative fit index (CFI), Tucker-Lewis index (TLI) and root mean square error of approximation (RMSEA). Model fit was deemed acceptable if CFI and TLI values were ≥0.90 and RMSEA was ≤0.08.[29,30] Model fit was considered good if CFI and TLI values were ≥0.95 and RMSEA was ≤0.06.[31]

The chi-square statistic (χ^2) was also reported, although it was not used to assess model fit because of its well-docu-

mented sensitivity to large sample sizes.[32] Factor loadings were required to be 0.32 or higher for items to be interpreted. For loadings, >0.71 was considered excellent, >0.63 was very good, >0.55 was good, >0.45 was fair, and between 0.45 and 0.32 was poor.[33,34]

The model fit statistics for the three competing SDQ models (i.e., five-factor, three-factor and higher-order model) were examined to compare the adequacy of the fit of the different models. Models can be considered indistinguishable when the change in CFI is less than 0.010 and the change in RMSEA is less than 0.015.[35]

Changes in model fit statistics were also inspected to test measurement invariance. There is support for a more constrained model when the CFI decreases by less than 0.010 and the RMSEA increases by less than 0.015.[35] Configural, metric and scalar invariance was tested to determine whether the underlying factor structure of the SDQ was the same across groups. Configural invariance assesses the fit of the factor structure when there are no invariance constraints imposed across groups. Metric invariance assesses the invariance of factor loadings across groups. Scalar invariance assesses the invariance of both factor loadings and item intercepts across groups. Metric invariance was assessed after configural invariance was established, and scalar invariance was tested after metric invariance was established. There is strong measurement invariance when all three tests demonstrate invariance across groups.[36]

In this study, these tests were performed with respect to sex (male vs. female), age (children [6 to 9 years] vs. adolescents [10 to 17 years]) and survey language (English vs. French). Age group classifications were based on guidelines from the World Health Organization,[37] which defines adolescents as individuals aged 10 to 19 years.

Finally, standardized factor loadings were used to calculate composite reliability scores for the latent factors.[38] Composite reliability was calculated because of increasing criticisms surrounding the use of the alpha coefficient as a measure of reliability.[39,40]

Results

Because the model fit statistics for the five-factor model were not inferior to those of the alternative models (see results below)—and because the five-factor model is the original, empirically derived SDQ model—descriptive statistics, correlations, internal consistencies (composite reliability), item-by-item factor loadings and invariance testing results are reported for the five-factor model only.

Descriptive statistics for the five SDQ factors are reported in Table 1. Among the four difficulties subscales, males and females scored lowest (best) on conduct problems and highest (worst) on hyperactivity. Children and adolescents scored lowest (best) on peer problems and conduct problems, respectively, while both children and adolescents scored highest (worst) on hyperactivity. Participants whose parents completed the SDQ in English scored lowest (best) on conduct problems and highest (worst) on hyperactivity, while those whose parents completed the SDQ in French scored lowest (best) on peer problems and highest (worst) on hyperactivity.

In terms of prosocial behaviour, males had a mean score of 8.85 and females had a mean score of 9.31. Children had a mean score of 9.08 and adolescents had a mean score of 9.06. Participants whose parents completed the SDQ in English had a mean score of 9.10, and those whose parents completed the SDQ in French had a mean score of 8.95.

Polychoric correlations among the five SDQ factors are reported in Table 2. Correlations were positive among the four difficulties factors, while the prosocial factor correlated negatively with the four difficulties factors.

According to the CFA results, the five-factor model fit was acceptable based on CFI (0.923) and TLI (0.913) values, and good based on the RMSEA (0.048; 90% confidence interval [CI] [0.047 to 0.049]), χ^2 (265) = 4,523.73, $p < 0.001$. Similarly, the higher-order model fit was acceptable based on CFI (0.920) and TLI (0.910) values, and good based on the RMSEA (0.049; 90% CI [0.047 to 0.050]), χ^2 (268) = 4688.34, $p < 0.001$. Although the higher-order model fit appeared to be slightly inferior to the five-factor model fit, it could not be empirically distinguished from the fit of the five-factor model (Δ CFI = -0.003; Δ TLI = -0.003; Δ RMSEA = +0.001).

Finally, CFA results indicated that the three-factor model failed to reach acceptable model fit based on CFI (0.883) and TLI (0.871) values, but that it had good model fit based on the RMSEA (0.058; 90% CI [0.057 to 0.060]), χ^2 (272) = 6718.60, $p < 0.001$. The three-factor model fit was generally inferior to the five-factor model fit (Δ CFI = -0.040; Δ TLI = -0.042; Δ RMSEA = +0.010) and the higher-order model fit (Δ CFI = -0.037; Δ TLI = -0.039; Δ RMSEA = +0.009).

All standardized factor loadings for the five-factor model were significant ($p < 0.001$) and ranged from 0.45 to 0.90 (Table 3). Composite reliability scores provided strong support for the internal consistency of all five factors (all scores ≥ 0.79; Table 3). For the higher-order model, all first-order factor loadings were significant ($p < 0.001$) and ranged from 0.45 to 0.90. All second-order factor loadings were significant ($p < 0.001$) and ranged from 0.75 to 0.98. All factor loadings for the three-factor model were significant ($p < 0.001$) and ranged from 0.40 to 0.89.

Table 2
Polychoric correlations for the five Strengths and Difficulties Questionnaire factors

Factors	F1	F2	F3	F4	F5
F1. Emotional symptoms	…	…	…	…	…
F2. Conduct problems	0.37	…	…	…	…
F3. Peer problems	0.41	0.34	…	…	…
F4. Hyperactivity	0.36	0.56	0.34	…	…
F5. Prosocial behaviour	-0.20	-0.46	-0.27	-0.36	…

… not applicable
Note: For all correlations, p < 0.001.
Source: Canadian Health Measures Survey, cycles 1 to 4 (2007 to 2015).

Table 3
Standardized factor loadings for the five-factor model

Subscale and items	Factor loading	Standard error	R-square
Emotional symptoms (CR = 0.83)			
Headaches, stomach aches	0.45	0.02	0.20
Worries	0.72	0.01	0.52
Unhappy, tearful	0.83	0.02	0.69
Nervous in new situations	0.68	0.01	0.46
Fears, scared	0.66	0.01	0.44
Conduct problems (CR = 0.81)			
Loses temper	0.68	0.01	0.46
Well behaved, does what adults request‡	0.78	0.01	0.61
Fights, bullies	0.70	0.02	0.49
Lies, cheats	0.70	0.01	0.49
Steals	0.67	0.03	0.46
Peer problems (CR = 0.79)			
Solitary, plays alone	0.47	0.02	0.22
At least one friend‡	0.69	0.03	0.48
Generally liked‡	0.90	0.02	0.81
Picked on, bullied	0.64	0.02	0.41
Gets along better with adults	0.51	0.02	0.26
Hyperactivity (CR = 0.88)			
Restless, overactive	0.75	0.01	0.57
Fidgeting, squirming	0.77	0.01	0.59
Distracted	0.81	0.01	0.66
Thinks before acting‡	0.73	0.01	0.53
Good attention span‡	0.80	0.01	0.64
Prosocial behaviour (CR = 0.83)			
Considerate of feelings	0.88	0.01	0.78
Shares with others	0.62	0.02	0.39
Helpful if someone is hurt	0.67	0.02	0.45
Kind	0.69	0.03	0.48
Offers to help	0.63	0.02	0.39

‡ Item is reverse-scored.
Notes: For all standardized factor loadings, $p < 0.001$. CR = composite reliability.
Source: Canadian Health Measures Survey, cycles 1 to 4 (2007 to 2015).

Table 4
Measurement invariance testing for the five-factor model

	CFI	TLI	RMSEA	CM	Δ CFI	Δ TLI	Δ RMSEA
Sex (49.2% female)							
1a. Configural invariance	0.923	0.913	0.047	…	…	…	…
1b. Metric invariance	0.927	0.921	0.045	1a	0.004	0.008	-0.002
1c. Scalar invariance	0.930	0.927	0.043	1b	0.003	0.006	-0.002
Age (37.6% children)							
2a. Configural invariance	0.925	0.915	0.047	…	…	…	…
2b. Metric invariance	0.927	0.921	0.045	2a	0.002	0.006	-0.002
2c. Scalar invariance	0.928	0.924	0.045	2b	0.001	0.003	0.000
Survey language (76.8% English)							
3a. Configural invariance	0.926	0.916	0.047	…	…	…	…
3b. Metric invariance	0.928	0.922	0.045	3a	0.002	0.006	-0.002
3c. Scalar invariance	0.931	0.928	0.043	3b	0.003	0.006	-0.002

… not applicable
Notes: CFI = comparative fit index, TLI = Tucker-Lewis index, RMSEA = root mean square error of approximation, CM = comparison model.
Source: Canadian Health Measures Survey, cycles 1 to 4 (2007 to 2015).

The results of measurement invariance testing for age (children vs. adolescents), sex (male vs. female) and survey language (English vs. French) provided evidence of satisfactory model fit (Table 4). All changes in fit indices pertaining to configural, metric and scalar invariance were well within acceptable ranges.[35] Therefore, there was support for strong measurement invariance across sex, age and language groups.

Discussion

The results of this study provide evidence for the factorial validity and reliability of the parent-rated SDQ with a nationally representative sample of Canadian children and adolescents. Specifically, the overall soundness of the original five-factor model[5] was supported using the parent-rated SDQ for children and adolescents aged 6 to 17. The five-factor model was also found to be invariant with respect to sex, age and survey language, highlighting the measure's robustness across different subsamples of the Canadian youth population. The findings of this study complement over two decades of international research examining the factor structure of the parent-rated SDQ.[6,13,14,16]

When examining the five-factor structure of the parent-rated SDQ, some CFA studies that used nationally representative samples reported poor model fit statistics on particular fit indices.[17,19] However, the results of all three (standard) fit indices in this study supported the five-factor structure. The five-factor SDQ structure was also confirmed in a recent CFA study with a community sample of Canadian children[21] and in other CFA studies that used nationally representative samples from the United States[14] and Great Britain.[7] Similar to this study, a study by He et al.[14] that used a sample of American adolescents found support for the five-factor model without needing to correlate errors. In contrast, Goodman et al.[7] found support for the five-factor model with a sample of children and adolescents from Great Britain, but only after correlating some error terms.

This study's findings also support the Goodman et al.[7] higher-order internalizing and externalizing model. This makes sense since the higher-order model contains the same underlying factor structure as the five-factor model. The higher-order model thus appears to be a viable alternative for researchers who want to examine parent-rated SDQ data from an internalizing and externalizing perspective. In contrast, this study found that the three-factor internalizing and externalizing model was

not a good fit to the data, according to certain fit indices. These results align with those of Goodman et al.,[7] who concluded that a first-order model that included internalizing and externalizing factors was not a viable simplification of the five-factor model. This study's results are also congruent with those of McCrory and Layte,[17] who found that the three-factor model was particularly problematic with a nationally representative sample of children from Ireland. In sum, the findings from this study suggest that the internalizing and externalizing factors demonstrate evidence of factorial validity in the context of the higher-order model, but not the first-order (three-factor) model.

Strengths and limitations

This study used data from a nationally representative sample of Canadian children and adolescents that spanned four CHMS cycles (2007 to 2015). The study included Canadian youth aged 6 to 17, resulting in an adequate representation of both children and adolescents. Parent-rated SDQ data demonstrated acceptable model fit using a stringent confirmatory technique (i.e., CFA) that allowed for no cross-loadings on unintended factors. Exploratory structural equation modelling, which is a more flexible technique because it allows for cross-loadings on unintended factors, was not required to achieve acceptable model fit in this study. Furthermore, the majority of SDQ items had factor loadings that could be considered "very good."[33,34]

Despite these strengths, the results of this study provide evidence of the SDQ's factorial validity only. Other types of construct validity were not assessed, such as criterion or concurrent validity. Moreover, test–retest reliability could not be examined because such data were not collected in the CHMS. Finally, the extent to which parent-rated SDQ scores were similar across parent raters (i.e., mother vs. father) could not be assessed because such data were not available for cycles 1 to 4 of the CHMS. However, previous research does suggest moderate to large correlations between scores for SDQs completed by mothers and fathers.[41,42]

Conclusion

Consistent with findings from several studies across the globe,[43] this study's findings support the parent-rated SDQ as a psychometrically sound tool for assessing mental health difficulties for Canadian children and adolescents in the general population. The SDQ also has the added benefit of assessing prosocial behaviour, a measure of positive mental health among children and adolescents. In addition to demonstrating evidence of factorial validity, reliability and measurement invariance, the parent-rated SDQ is also relatively brief in comparison with typical psychosocial measures. Normative parent-rated SDQ data for the general Canadian youth population are not yet available; this is an important area for future research. Establishing criterion-referenced cut points among Canadian children and adolescents would help clinicians and health care practitioners identify individuals who may be at risk for mental health difficulties. ■

References

1. Merikangas KR, Nakamura EF, Kessler RC. Epidemiology of mental disorders in children and adolescents. *Dialogues in Clinical Neuroscience* 2009; 11: 7-20.
2. Kessler RC, Amminger GP, Aguilar-Gaxiola S, et al. Age of onset of mental disorders: a review of recent literature. *Current Opinion in Psychiatry* 2007; 20: 359-64.
3. Marcus M, Yasamy MT, van Ommeren M, et al. *Depression: A Global Public Health Concern*. World Health Organization Department of Mental Health and Substance Abuse, 2012; 1: 6-8.
4. Public Health Agency of Canada. *Report from the Canadian Chronic Disease Surveillance System: Mental Illness in Canada*. Ottawa: Public Health Agency of Canada, 2015. Available at: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/report-canadian-chronic-disease-surveillance-system-mental-illness-canada-2015.html.
5. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry* 1997; 38: 581-6.
6. Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry* 2001; 40: 1337-45.
7. Goodman A, Lamping DL, Ploubidis GB. When to use broader internalising and externalising subscales instead of the hypothesized five subscales of the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *Journal of Abnormal Child Psychology* 2010; 38: 1179-91.
8. Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *European Child & Adolescent Psychiatry* 1998; 7: 125-30.
9. Youthinmind. Strengths and Difficulties Questionnaire. Available at: http://sdqinfo.org/py/sdqinfo/b0.py. Accessed July 19, 2019.
10. Siriwardhana C, Pannala G, Sirribaddana S, et al. Impact of exposure to conflict, tsunami and mental disorders on school absenteeism: findings from a national sample of Sri Lankan children aged 12-17 years. *BMC Public Health* 2013; 13: 560.
11. Hölling H, Kurth BM, Rothenberger A, et al. Assessing psychopathological problems of children and adolescents from 3-17 years in a nationwide representative sample: results of the German health interview and examination survey for children and adolescents (KiGGS). *European Child & Adolescent Psychiatry* 2008; 17(Suppl.): 34-41.
12. Becker A, Steinhausen HC, Baldursson G, et al. Psychopathological screening of children with ADHD: Strengths and Difficulties Questionnaire in a pan-European study. *European Child & Adolescent Psychiatry* 2006; 15(Suppl. 1): I/56-I/62.
13. Hawes DJ, Dadds MR. Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *Australia and New Zealand Journal of Psychiatry* 2004; 38: 644-51.
14. He JP, Burstein M, Schmitz A, et al. The Strengths and Difficulties Questionnaire (SDQ): the factor structure and scale validation in U.S. adolescents. *Journal of Abnormal Child Psychology* 2013; 41: 583-95.
15. D'Acremont M, Van der Linden M. Confirmatory factor analysis of the Strengths and Difficulties Questionnaire in a community sample of French-speaking adolescents. *European Journal of Psychological Assessment* 2008; 24: 1-8.
16. Bull R, Lee K, Koh IHC, et al. Confirmatory factor analysis of the Strengths and Difficulties Questionnaire in Singaporean kindergarteners. *Child: Care, Health and Development* 2015; 42: 109-16.
17. McCrory C, Layte R. Testing competing models of the Strengths and Difficulties Questionnaire's (SDQ's) factor structure for the parent-informant instrument. *Personality and Individual Differences* 2012; 52: 882-7.
18. Mellor D, Stokes M. The factor structure of the Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment* 2007; 23: 105-12.
19. Rothenberger A, Becker A, Erhart M, et al., the BELLA study group. Psychometric properties of the parent Strengths and Difficulties Questionnaire in the general population of German children and adolescents: results of the BELLA study. *European Child & Adolescent Psychiatry* 2008; 17(Suppl. 1): 99-105.
20. Van Leeuwen K, Meerschaert T, Bosmans G, et al. The Strengths and Difficulties Questionnaire in a community sample of young children in Flanders. *European Journal of Psychological Assessment* 2006; 22: 189-97.
21. Aitken M, Martinussen R, Wolfe RG, et al. Factor structure of the Strengths and Difficulties Questionnaire in a Canadian elementary school sample. *Assessment for Effective Intervention* 2015; 40: 155-65.
22. Statistics Canada. *Canadian Health Measures Survey (CHMS) Data User Guide: Cycle 4*. Ottawa: Statistics Canada, 2017.
23. Tremblay MS, Connor Gorber S. Canadian Health Measures Survey: brief overview. *Canadian Journal of Public Health* 2007; 98: 453-6.
24. Statistics Canada. *Canadian Health Measures Survey: Instructions for Combining Multiple Cycles of Canadian Health Measures Survey (CHMS) Data*. Ottawa: Statistics Canada, 2019.
25. Day B, Langlois R, Tremblay MS, et al. Canadian Health Measures Survey: Ethical, legal and social issues. *Health Reports* 2007; 18(Suppl.): 35-52.
26. Bougie E, Arim RG, Kohen D, et al. Validation of the 10-item Kessler Psychological Distress Scale (K10) in the 2012 Aboriginal Peoples Survey. *Health Reports* 2016; 27: 3-10.
27. Brummel BJ, Drasgow F. The effects of estimator choice and weighting strategies on confirmatory factor analysis with stratified samples. *Applied Multivariate Research* 2010; 13: 113-28.
28. Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods* 2016; 48: 936-49.
29. Browne MW, Cudeck R. Alternative ways of assessing model fit. *Sociological Methods & Research* 1992; 21: 230-58.
30. Brown TA. *Confirmatory Factor Analysis for Applied Research*. New York: Guilford Press, 2006.
31. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* 1999; 6: 1-55.
32. Kline RB. *Principles and Practice of Structural Equating Modeling*. 2nd ed. New York: Guilford Press, 2005.
33. Comrey AL, Lee HB. *A First Course in Factor Analysis*. 2nd ed. Hillsdale: Erlbaum, 1992.
34. Tabachnick, BG, Fidell, LS. *Using Multivariate Statistics*. 6th ed. Boston: Pearson, 2013.
35. Chen FF. Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling* 2007; 14: 464-504.
36. Marsh HW, Parker PD, Morin AJS. Invariance testing across samples and time: cohort-sequence analysis of perceived body composition. In: Ntoumanis N, Myers ND, eds. *An Introduction to Advanced and Intermediate Statistical Analyses for Sport and Exercise Scientists*. Hoboken: John Wiley & Sons, 2016: 101-30.

37. World Health Organization. *Adolescent Health in the South-East Asian Region*. Available at: https://www.who.int/southeastasia/health-topics/adolescent-health. Accessed August 21, 2019.

38. Raykov T. Estimation of reliability for congeneric measures. *Applied Psychological Measurement* 1997; 21: 173-84.

39. McNeish, D. Thanks coefficient alpha, we'll take it from here. *Psychological Methods* 2018; 23: 412-33.

40. Martin, JJ, Beasley VL, Guerrero MD. Sport psychology research: proper standards and limitations. In: Anshel MH, Petrie TA, Steinfeldt JA, eds. *APA Handbooks in Psychology Series, APA Handbook of Sport and Exercise Psychology*. Washington, DC: American Psychological Association, 2019: 17-40.

41. Li JB, Delvecchio E, Di Riso D, et al. Early evidence of the Italian parent-report Strengths and Difficulties Questionnaire (SDQ-P). *Child Psychiatry & Human Development* 2017; 48: 335-45.

42. Davé S, Nazareth I, Senior R, et al. A comparison of father and mother report of child behaviour on the Strengths and Difficulties Questionnaire. *Child Psychiatry & Human Development* 2008; 39: 399-413.

43. Stone LL, Otten R, Engels RCME, et al. Psychometric properties of the parent and teacher versions of the Strengths and Difficulties Questionnaire for 4- to 12-year-olds: a review. *Clinical Child and Family Psychology Review* 2010; 113: 254-74.