

## Short report

## Assessing the readability of the self-reported Strengths and Difficulties Questionnaire

Praveetha Patalay, Daniel Hayes and Miranda Wolpert

The Strengths and Difficulties Questionnaire (SDQ) is one of the most widely used measures in child and adolescent mental health in clinical practice, community-based screening and research. Assessing the readability of such questionnaires is important as young people may not comprehend items above their reading ability when self-reporting. Analyses of readability in the present study indicate that the self-report SDQ might not be suitable for young people with a reading age below 13–14 years and highlight differences in readability between subscales. The findings suggest a need for caution in using the SDQ as a self-report measure for children below the age of 13, and highlight considerations of readability in measure development, selection and interpretation.

**Declaration of interest**

None.

**Copyright and usage**

© The Royal College of Psychiatrists 2018. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work.

Self-reported perspectives of mental health symptoms are being used more widely in screening, research and practice in child and adolescent mental health.<sup>1–3</sup> Measures have been developed and tested for use by young people to report on their mental health symptoms, and several such measures that are well validated in terms of psychometric properties exist.<sup>4</sup> Measure development has become increasingly sophisticated, and a range of psychometric and other properties of questionnaires are assessed before deciding they are fit for purpose.<sup>5,6</sup> However, while observing the criteria used to assess measure suitability,<sup>6</sup> we noted that a key criterion that is specifically relevant to child measures, but is routinely not accounted for, is the readability of the items in the scale. Readability has been defined in many ways, but in simple terms refers to the ease with which a reader can read and understand text.<sup>7,8</sup> The concern regarding the readability of psychological questionnaires is not new, with the readability of some adult psychopathology questionnaires<sup>9,10</sup> and some measures of child psychopathology, as reported by parents and children, having been assessed for their readability.<sup>11</sup> However, the latter investigation does not include a readability assessment of one of the most widely used child mental health measures in the UK, the Strengths and Difficulties Questionnaire (SDQ). The self-report SDQ is used extensively in the UK in research, community settings and clinical practice<sup>12,13</sup> and is available to complete in different modes.<sup>14</sup> The developers recommend this self-report measure for young people aged 11–17 years.<sup>15</sup> However, it has since been psychometrically validated for use in younger children, including from 8 years<sup>16</sup> and as young as 6 years.<sup>17</sup> In this short paper, we investigate the reading age suitability of the self-report SDQ.

**Method****Strengths and Difficulties Questionnaire**

The SDQ is a 25-item questionnaire that comprises five five-item subscales: emotional symptoms, conduct problems, hyperactivity, peer problems and prosocial behaviour.<sup>18</sup> Participants respond to each item by selecting one of three responses: not true, somewhat true and certainly true.

**Readability**

In the current study, we used four standard methods that can be used to examine the readability of text.

## Flesch–Kincaid reading grade (FK)

This method,<sup>19</sup> adapted from the Flesch Reading Ease score,<sup>19</sup> is one of the oldest and most widely used readability indices, and is based on average numbers of syllables per word and sentence lengths. Scores are estimated as a US grade level.

$$FK = (0.39 \times ASL) + (11.8 \times ASW) - 15.59$$

where *ASL* = average words per sentence; *ASW* = average syllables per word

## Gunning Fog Index (GFI)

The GFI corresponds to the number of years of formal education required to understand text<sup>20</sup> and uses the numbers of words, sentences and complex words, which are defined as having three or more syllables.

$$FI = 2/5 \times [(A/N) + (100L/A)]$$

where *A* = number of words; *N* = number of sentences; *L* = number of words with three or more syllables (excluding -ing and -ed ends)

## Coleman Liau Index (CLI)

The CLI<sup>21</sup> differs from the FK and GFI tests by focusing on the number of letters (rather than syllables) per word. It also results in a US grade-level score. The formula for estimating the CLI is:

$$CLI = (0.0588 \times L) - (0.296 \times S) - 15.8$$

where *L* = average letters/100 words; *S* = average sentences/100 words

## Dale–Chall Readability Formula (DC)

The DC differs from the previous three indices by incorporating the level of difficulty of the words in the text into the formula for its

estimation.<sup>7,22</sup> A list of words that up to 80% of fourth graders (children aged around 10) know is used as the basis for identifying words that can be considered difficult.

$$DC = 0.1579(DW/A \times 100) + 0.0496 (ASL) + 3.6365$$

where *DW* = difficult words (i.e. words not in the list), *A* = number of words, *ASL* = average words per sentence

These indices mainly estimate readability as a US grade-level score. Grade levels can be translated to age levels by adding 6 to a grade-level score (children in US grade 1 are aged 6–7 years), and this was done in this study. The four methods were chosen not only because they are well established and widely used measures of readability, but also because they have differences in focus when estimating readability, which can lead to varying readability estimates. The unique elements as part of their estimation include that the FK focuses on syllables in words, the GFI on words with more than three syllables, the CLI on the number of letters per word and the DC on the presence of difficult words.

### Procedure

The readability formulae described above were applied to the instruction section and the items of the SDQ (for the full set of items in the measure and for each of the subscales separately), resulting in four readability estimates for each of the examined components (Table 1). The estimates across the four readability estimates were averaged to provide a single readability estimate.

## Results

For the full measure, age estimates for readability ranged from 10.94 to 12.74, with a mean estimate of 11.75 years (Table 1). For the subscales, the conduct problems subscale had the lowest mean readability age of 10.46 years, followed by peer problems (*M* = 11.83), prosocial behaviour (*M* = 12.84), hyperactivity (*M* = 13.56), with the emotional symptoms subscale (*M* = 13.85) having the highest average readability age. For the instructions, the average readability estimate was 13.41 years.

## Discussion

The results indicate that while some of the SDQ subscales have a readability of around 11–12 years (peer and conduct problems), the instructions and some of the subscales (notably emotional difficulties and hyperactivity) have average readability estimates that are

substantially higher (ranging up to 13.9 years). On the basis of these readability estimates, the SDQ would be considered unsuitable for 6- and 8-year-olds (despite psychometric validation studies at these ages)<sup>16,17</sup>. Moreover, these findings suggest that it might be difficult to understand overall for 11-year-olds (the recommended starting age for this measure). This difficulty may be further compounded when young people have lower reading ages relative to their developmental age. This might be of particular relevance in clinical settings, given that many children with mental health difficulties also have learning difficulties and special educational needs.<sup>23</sup> It is important to note that this issue of unsuitably high readability is not unique to this self-report measure; for example, the Youth Self Report version of the Achenbach assessments of child mental health has a readability estimate of 12.5 years,<sup>11</sup> although, as with the SDQ, it is meant to be suitable from age 11 years onwards.

We present the results from a range of readability indices to highlight the variation in the age estimates they provide. The inclusion of the DC is especially relevant here, as it is the only index to include word complexity in its estimate, an aspect that provides insight into possible difficulty in understanding the specific content of the items. Previous attempts to map the readability of psychopathology measures<sup>9</sup> have been criticised for not taking word complexity into account in their estimates.<sup>24</sup> In this example, the hyperactivity and emotional symptoms subscales have eight and seven difficult words, respectively (e.g. squirming, fidgeting, down-hearted). This highlights the importance of also considering specific words and their suitability for the age group in question when designing questionnaires. These findings raise issues for interpretation of the SDQ self-report data derived from younger children. If younger children with a reading age lower than around 13 who are completing the SDQ are not comprehending items, key words or instructions, this may affect their responses and subsequently the derived scores used in analysis or to inform treatment.

Psychometric properties are not the only criteria that determine whether or not a measure is fit for purpose. There is a crucial prior step: assessing whether the target population can read and understand the items in a questionnaire comfortably. In this paper we describe and apply a non-resource-intensive approach to assess the content of measures using four standardised measures of readability. Estimates of readability provide an overarching view of the complexity of the language used in a questionnaire and help identify words that might be difficult to understand. In addition, more intensive approaches to investigating measure comprehension, such as cognitive interviewing, can help illuminate how items are understood and interpreted by respondents.<sup>25</sup>

Advice in relation to measures and health-related materials for adults is that they should have a readability of around 12 years, although the average reading age of adults is around 14 years.<sup>26</sup> Extrapolating from this advice, we recommend that child self-report measures should aim for a readability age of around 2 years below the target minimum age for the measure.

The implications of these findings for the use of the SDQ include reconsidering the age of the target group for the questionnaire, developing accompanying support materials and explanations to aid completion of the survey or developing alternative questionnaires with a lower readability age. The wider implications of the results are that they highlight that alongside psychometric properties, a key consideration for the selection and reliable use of any self-report measure with children (or adults) should be: can the target user understand this?

### Ethical approval

Given that no data from human participants were used in this report, no ethical approvals were necessary.

**Table 1** Readability estimates (in years) made using the four different approaches and the average estimate for the full measure, instructions and subscales of the Strengths and Difficulties Questionnaire

	FK	GFI	CLI	DC	Average readability age across the four indices
Full measure	10.94	12.74	11.63	11.69	<b>11.75</b>
Instructions	12.88	15.22	13.33	12.22	<b>13.41</b>
<i>Subscales</i>					
Conduct disorder	9.40	10.74	11.27	10.44	<b>10.46</b>
Emotional symptoms	12.01	14.30	16.16	12.92	<b>13.85</b>
Hyperactivity	11.16	14.16	15.79	13.12	<b>13.56</b>
Peer problems	10.82	12.43	12.38	11.69	<b>11.83</b>
Prosocial behaviour	11.85	12.43	16.38	10.77	<b>12.84</b>

FK, Flesch–Kincaid reading grade; GFI, Gunning Fog Index; CLI, Coleman Liau Index; DC, Dale–Chall Readability Formula.

**Praveetha Patalay**, PhD, University of Liverpool, UK, Evidence Based Practice Unit, UCL, London, UK, and Anna Freud National Centre for Children and Families, London, UK; **Daniel Hayes**, MSc, Evidence Based Practice Unit, UCL, London, UK and Anna Freud National Centre for Children and Families, London, UK; **Miranda Wolpert**, DClinPsy, Evidence Based Practice Unit, UCL, London, UK, Anna Freud National Centre for Children and Families, London, UK, and Child Outcomes Research Consortium, London, UK

**Correspondence:** Praveetha Patalay, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool, L69 7ZA, UK. Email: [ppatalay@liverpool.ac.uk](mailto:ppatalay@liverpool.ac.uk)

First received 23 Oct 2017, accepted 23 Nov 2017

## References

- Green H, McGinnity A, Meltzer H, Ford T, Goodman R. *Mental Health of Children and Young People in Great Britain, 2004*. Palgrave Macmillan, 2005.
- Deighton J, Lereya ST, Morgan E, Breedvelt H, Martin K, Feltham A, et al. Measuring and monitoring children and young people's mental wellbeing: a toolkit for schools and colleges. Public Health England and the Evidence Based Practice Unit 2016.
- Department of Health, Department for Education, NHS England. *Future in Mind: Promoting, Protecting and Improving Our Children and Young People's Mental Health and Wellbeing*. Department of Health, 2015.
- Deighton J, Croudace T, Fonagy P, Brown J, Patalay P, Wolpert M. Measuring mental health and wellbeing outcomes for children and adolescents to inform practice and policy: a review of child self-report measures. *Child Adolesc Psychiatry Ment Health* 2014; **8**: 1.
- Rust J, Golombok S. *Modern Psychometrics: the Science of Psychological Assessment*. 3rd ed. Routledge, 2009.
- Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality of life instruments: attributes and review criteria. *Qual Life Res* 2002; **11**: 193–205.
- Dale E, Chall JS. A formula for predicting readability. *Educ Res Bull* 1948; **27**: 11–28.
- Badgett BA. Toward the development of a model to estimate the readability of credentialing-examination materials. University of Nevada, Texas 2010.
- McHugh RK, Behar E. Readability of self-report measures of depression and anxiety. *J Consult Clin Psychol* 2009; **77**: 1100–12.
- McHugh RK, Rasmussen JL, Otto MW. Comprehension of self-report evidence-based measures of anxiety. *Depress Anxiety* 2011; **28**: 607–14.
- Jensen SA, Fabiano GA, Lopez-Williams A, Chacko A. The reading grade level of common measures in child and adolescent clinical psychology. *Psychol Assess* 2006; **18**: 346–52.
- Wolpert M, Jacob J, Napoleone E, Whale A, Calderon A, Edbrooke-Childs J. *Child- and Parent-reported Outcomes and Experience from Child and Young People's Mental Health Services 2011–2015*. CAMHS Press, 2016.
- Fink E, Patalay P, Sharpe H, Holley S, Deighton J, Wolpert M. Mental health difficulties in early adolescence: a comparison of two cross-sectional studies in England from 2009 to 2014. *J Adolesc Health* 2015; **56**(5): 502–7.
- Patalay P, Hayes D, Deighton J, Wolpert M. A comparison of paper and computer administered strengths and difficulties questionnaire. *J Psychopathol Behav Assess* 2016; **38**: 242–50.
- Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *Eur Child Adolesc Psychiatry* 1998; **7**: 125–30.
- Muris P, Meesters C, Eijkelenboom A, Vincken M. The self-report version of the Strengths and Difficulties Questionnaire: its psychometric properties in 8- to 13-year-old non-clinical children. *Br J Clin Psychol* 2004; **43**: 437–48.
- Curvis W, McNulty S, Qualter P. The validation of the self-report Strengths and Difficulties Questionnaire for use by 6- to 10-year-old children in the UK. *Br J Clin Psychol*. 2014; **53**: 131–7.
- Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *Int Rev Psychiatry* 1998; **15**: 173–7.
- Kincaid JP, Fishburne Jr RP, Rogers RL, Chissom BS. *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel*. Naval Technical Training Command, 1975.
- Gunning R. *The Technique of Clear Writing*. McGraw-Hill, 1952.
- Coleman M, Liau TL. A computer readability formula designed for machine scoring. *J Appl Psychol* 1975; **60**: 283–4.
- Chall JS, Dale E. *Readability Revisited: The New Dale-Chall Readability Formula*. Brookline Books, 1995.
- Emerson E, Hatton C. Mental health of children and adolescents with intellectual disabilities in Britain. *Br J Psychiatry* 2007; **191**: 493–9.
- Schinka JA. Further issues in determining the readability of self-report items: comment on McHugh and Behar (2009). *J Consult Clin Psychol* 2012; **80**: 952–5.
- Willis GB. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Sage Publications, 2004.
- Parker RM, Williams MV, Weiss BD, Baker DW, Davis TC, Doak CC, et al. Health literacy: report of the council on scientific affairs. *JAMA* 1999; **281**: 552–7.

