Personality and Individual Differences 52 (2012) 882–887

Contents lists available at SciVerse ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

# Testing competing models of the Strengths and Difficulties Questionnaire's (SDQ's) factor structure for the parent-informant instrument

Cathal McCrory *, Richard Layte

*The Economic and Social Research Institute, Whitaker Square, Sir John Rogerson's Quay, Dublin 2, Ireland*

ARTICLE INFO

*Article history:*
Received 4 August 2011
Received in revised form 1 February 2012
Accepted 10 February 2012

*Keywords:*
SDQ
Construct validity
Confirmatory Factor Analysis
Cohort study
Ireland

ABSTRACT

The Strengths and Difficulties Questionnaire (SDQ) is a brief 25-item instrument that has been widely employed in clinical and epidemiological studies to assess children's psychological adjustment. Despite its widespread application in child and adolescent research, concerns have been expressed regarding the construct validity of the instrument and whether it might be tainted by a method factor that may undermine its utility as a diagnostic tool. We employed a confirmatory factor analytic approach to compare the goodness of fit of four competing models suggested by the extant literature for the parent-informant version of the questionnaire using data for 8514 nine-year-old children participating in the Growing Up in Ireland Study – a large population based cohort study in the Republic of Ireland. While analysis of the data provided support for the traditional five-factor conceptualisation of the instrument, a six-factor model which incorporated a method factor was found to fit the data marginally better. Nevertheless, we conclude that the existence of method effects does not present any great threat to the structural validity of the instrument taking account of patterns in the data and model parsimony.



## 1. Introduction

The Strengths and Difficulties Questionnaire (SDQ: Goodman, 1997) was designed as a relatively brief (25-item) behavioural screening questionnaire for measuring psychological adjustment in children aged 3–16 years. The instrument produces scores for five subscales: emotional symptoms, conduct problems, hyperactivity/inattention, peer problems and prosocial behaviour. A total difficulties score is obtained by summing scores across the four-deficit focused scales (i.e. all except the prosocial behaviour scale).

The SDQ has increased in popularity in recent years because it has a number of desirable qualities from a test administrator's perspective. It has the obvious advantage of brevity as the instrument takes less than 10 min to complete. A second well regarded feature is its deviation from a narrow deficit-focused approach to encompass facets of children's prosocial behaviour. Thirdly, a recent review of 48 studies attests to the robust psychometric properties of the instrument including good concurrent, predictive and discriminant validity, and its ability to differentiate between clinic and community based samples (see Stone, Otten, Engels, Vermulst, & Janssens, 2010 for a review).

Although the five factor structure of the SDQ has been affirmed in a number of studies in the US (Bourdon, Goodman, Rae, Simpson, & Koretz, 2005), UK (e.g. Goodman, 2001), Germany (Woerner, Becker, & Rothenberger, 2004), the Netherlands (Muris, Meesters, & van den Berg, 2003) Australia (Hawes & Dadds, 2004), Sweden (Smedje, Broman, Hetta, & von Knorring, 1999), and France (Capron, Therond, & Duyme, 2007) these validation studies have used exploratory rather than confirmatory techniques. This is important because a number of more recent studies employing Confirmatory Factor Analytic (CFA) techniques have tended to be more equivocal in their support for the putative factor structure.

Dickey and Blumberg (2004), for example, failed to replicate the postulated factor structure of the SDQ in a large nationally representative sample of 9574 respondents to the US National Health Interview Survey. Specifically, they found rather poor support for the factorial independence of two of the hypothesised dimensions (peer problems and conduct problems) when subjected to principal components analysis. A subsequent exploratory factor analysis of the data suggested that a 3-factor solution represented the best structural configuration of the data. Moreover, the authors argued that the three extracted factors were conceptually meaningful. The five hyperactivity items and four of the conduct items loaded on a factor which they labelled externalising problems. The five emotionality items and three of the peer problem items loaded on a second factor which they interpreted as an internalising problems dimension. Finally, the five prosocial items, two of the positively worded peer problem items ('has at least one good friend' and 'popular with other children') and one of the positively worded conduct items ('generally obedient') loaded on a factor which they tentatively suggested could represent a 'positive construal' or method factor.

* Corresponding author. Tel.: +353 1 863 2027; fax: +353 1 863 2100.
*E-mail address:* cathal.mccrory@esri.ie (C. McCrory).

0191-8869/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.paid.2012.02.011

This tendency for the reverse-scored problem-oriented items to load on the prosocial factor had previously been noted by Goodman (2001) in relation to the teacher and self-reported, but not the parent-informant versions of the instrument. Dickey and Blumberg (2004) undertook a CFA using the same dataset and confirmed that a three factor model did indeed provide a good fit to the data: however, they acknowledged that changes to the SDQ's item wording, which were designed to aid content understanding among American respondents, may have altered item interpretation and confounded the factor structure.

Palmieri and Smith (2007) undertook a CFA exercise in which they tested three different structural models of the SDQ for a sample of 733 custodial grandparents. The first of these was a hierarchical model in which the four deficit-focused scales defined a second-order difficulties factor and a separate first-order strengths factor. The second of these was a modification of the first model in which the five factors were independent but correlated to a modest extent. Finally, they also tested a six-factor model which comprised five correlated factors, and a separate uncorrelated 'method factor' which consisted of all positively worded items (the five prosocial items, and the five reverse-scored problem-oriented items). Although models one and two both demonstrated acceptable fit to the data, model two was found to fit the data significantly better. Nevertheless, these authors concluded that the extent of the scale inter-relationships was such as to justify the continued use of a total difficulties score. Dickey and Blumberg's (2004) premise that the SDQ might be tainted by a method factor was corroborated by Palmieri and Smith (2007) who found that nine of the ten positively worded items loaded greater than 0.30 on the positive construal factor stipulated in model 3. However, they surmised that this did not represent any great threat to the construct validity of the problem-oriented scales as the symptom factor loadings were higher than the method factor loadings.

While Van Roy and colleagues (Van Roy, Veenstra, & Clench-Aas, 2008) observed that the five-factor model demonstrated acceptable model fit in a large population based study of 26,269 Norwegian children aged 10–19 years in a multi-method (parent and self-report) multi-trait examination of the SDQ, they too noted that the introduction of a positive construal factor – defined by the five positively worded reverse coded items – significantly improved the fit of the model. However, this factor was found to be highly correlated with the prosocial factor ($r = 0.63–0.70$) and examination of the pattern of factor loadings revealed that only two of the items were more heavily determined by the positive construal factor rather than they were with their original trait. The authors also noted considerable conceptual overlap between some of the subscales. High correlations between the hyperactive and conduct problems subscales, and between the peer problems and emotional problems subscales suggested that there may be some substance in Dickey and Blumberg's alternative conceptualisation of the SDQ along internalising and externalising dimensions. These investigators acknowledged that "there may be alternative models that fit equally well or perhaps even better" (p. 1310).

Goodman, Lamping, and Plobidis (2010) explored whether a second-order model, which specified additional internalising and externalising factors in addition to the first order factors, might provide a better conceptualisation. However, they found that it fit the data less well across separate parent, teacher and self-report versions of the instrument than the model specifying five first order factors. Mellor and Stokes (2007) used confirmatory methods to examine the factor structure of proxy and self-report versions of the SDQ in a sample of 914 Australian children aged 7–17 years and found that the hypothesised factor structure failed to emerge across any of the different informant versions of the instrument and item analysis indicating that none of the putative subscales were unidimensional.

In addition to concerns surrounding the structural validity of the SDQ, investigators have also identified a number of other potential difficulties. These include a number of items that do not load on their hypothesised scales (e.g. Hawes & Dadds, 2004; Smedje et al., 1999) and low scale reliabilities for at least two of the subscales (e.g. Goodman, 2001; Smedje et al., 1999; Stone et al., 2010). This discrepant pattern of results highlights the need for continued evaluation of the SDQ. Mellor and Stokes (2007) remarked that "further large factor analytic studies that ask whether or not it is feasible to find an alternative structure that would force a rearrangement of items onto alternative subscales are warranted" (p. 111).

The present study therefore proposes to examine the structural characteristics of the SDQ within a large nationally representative sample of nine-year old children. This will involve testing competing models suggested by the literature to ascertain which structural model provides the most parsimonious fit to the data using a range of fit measures and CFA.

## 2. Method

### 2.1. Sample

The study comprised a nationally representative sample of 8568 nine-year-old school children resident in Ireland who were participating in the first wave of the Growing Up in Ireland project, a national longitudinal study of Irish children. The sample was selected through the national school system using a probability proportionate to size (PPS) sampling method with schools serving as the primary sampling units. 1105 schools from the national total of 3200 Primary schools were selected for inclusion in the study and the sample was randomly generated from within those schools. At the school level, a response rate of 82% was achieved, while at the level of the household (i.e. eligible child selected within the school) a total of 57% of children and their families participated in the study.

### 2.2. Materials

The English language (UK) informant rated version of the SDQ for parents of children aged 4–16 years was administered to parents on a computer assisted personal interviewing format as part of the standard household interview ($n = 8568$). The SDQ produces scores for each of five subscales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and prosocial behaviour, with each subscale comprising 5 items. A total difficulties score is obtained by summing scores across the four deficit-focused scales (i.e. all except the prosocial behaviour scale). Respondents are required to indicate their level of agreement to each item on a three-point scale with 0 = 'not true', 1 = 'somewhat true' and 2 = 'certainly true'. Subscale scores vary from 0 to 10 and the total difficulties score ranges from 0 to 40. Higher scores on the problem-oriented scales are indicative of more problems. Analysis revealed that the incidence of missing cases was small with only 53 cases or 0.6% of the sample having missing data across any of the items. Thus the effective case base for the CFA analysis was 8514.

## 3. Results

### 3.1. Confirmatory Factor Analysis (CFA)

CFA is a widely used technique for examining hypothesised relations among variables and comparing the goodness of fit of competing models. CFA was undertaken in Mplus (Muthen &