

Psychometric Properties of the Self-Report Version of the Strengths and Difficulties Questionnaire in Korea

Min-Hyuk Kim, Joung-Sook Ahn[✉], and Seongho Min

Department of Psychiatry, Yonsei University Wonju College of Medicine, Wonju, Republic of Korea

Objective To examine the reliability and validity of the self-report Korean version of Strengths and Difficulties Questionnaire (SDQ-Kr) in the community sample.

Methods The SDQ-Kr was administered to a large sample of school children and adolescents (n=2814) and youth attendees of a psychiatric clinic (n=385) aged 11–16 years. To examine temporal stability, the same questionnaire was administered to a subsample of 167 school youths five to six weeks after the initial assessment. To examine the reliability, we calculated Cronbach's alpha coefficients for internal consistency and Pearson's correlations for test-retest stability. In order to evaluate the factorial structure of the SDQ-Kr items, we conducted an exploratory factor analysis (EFA) with varimax rotation. Finally, discriminant validity was examined by using receiver operating characteristic (ROC) curves employing the area under the curve (AUC) as an index of discriminant ability.

Results Although the internal consistency of some subscales of the SDQ-Kr was somewhat less satisfactory (alpha coefficients; 0.28–0.54) than the English original, coefficients for the total difficulties scores approached sufficient levels (coefficients; 0.69). Other psychometric properties including discriminant validity (AUC for total difficulties and four subscales >0.7) were comparable to those obtained in other language studies.

Conclusion The self-report SDQ-Kr exhibited a low level of reliability, indicating that some items need to be further evaluated and revised to improve the psychometric properties. We suggest that the total difficulties score could be used with more confidence for screening possible mental health problems in youths.

Psychiatry Investig 2015;12(4):491-499

Key Words Strengths and difficulties questionnaire, Self-report, Validity, Reliability.

INTRODUCTION

Mental health problems in children and adolescents are common and result in significant burden on the individual child, their families, schools, and communities because they can persist into adulthood with considerable financial and sociological costs.^{1,2} A meta-analysis of 49 studies involving more than 240,000 children in numerous countries reported a 1-year prevalence rate of 12.3% for psychiatric disorders as diagnosed by interviews.³

In Korea, an increasing number of children and adolescents are being identified as suffering from a wide range of

mental health problems. According to a cross-national study on emotional and behavioral problems in primary school children in Korea, Japan, and China investigated with Rutter's questionnaires in 1993, the prevalence rates of children with deviant scores on the parents' and teachers' scales were 19.1% and 14.1% in Korea, respectively, which were higher than those of the other two countries.⁴ A more recent study showed that the most common disorder for 8- to 9-year-old aged students in Korea is ADHD, followed by tic disorder and anxiety disorders, and psychiatric comorbidity was found in 34.4% of the cases.⁵

However, no definitive nationwide statistics of how many youth being at risk of mental disorders exist so far, although there are approximately 10.3 million children and adolescents in Korea, making up 21% of the total population. In addition to lack of financial support, lack of a suitable questionnaire reported by youth themselves is one of the various reasons for it.

Since psychopathologies in youth are disguised according to the situation and therefore not revealed or expressed, in-

Received: October 14, 2014 Revised: January 19, 2015

Accepted: February 12, 2015 Available online: September 30, 2015

✉ Correspondence: Joung-Sook Ahn, MD, PhD

Department of Psychiatry, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju 26426, Republic of Korea

Tel: +82-33-741-1260, Fax: +82-33-741-0532, E-mail: jsahn@yonsei.ac.kr

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

formants such as parents or teachers cannot identify these problems accurately. Therefore a self-report questionnaire that is complementary and compatible with the informant version is needed. The self-report version of the Strengths and Difficulties Questionnaire (SDQ) for youths aged 11–16 years, like its informant (parent or teacher)-rated SDQ, only comprises 25 items and addresses a well-balanced number of positive and negative behavioral aspects, in contrast to the considerably longer Youth Self-Report (YSR), the self-report derivative of the Child Behavior Checklist (CBCL).⁶ Since its development by Goodman as a screening device to detect children with a mental problems in the UK in 1997,⁷ the SDQ has now been translated into 66 different languages and has become an internationally recognized tool for clinical settings as well as epidemiological studies.⁸

Many studies from several countries have reported psychometric properties with somewhat variable results since a pilot study of the self-rated SDQ by Goodman et al. in 1998 found that it discriminated between a community and a psychiatric sample.^{9–11} Some cross-cultural issues have been raised by several recent studies questioning the original subscale structure of the SDQ.^{12–14} In Korea, however, separate psychometric properties and normative data for the self-rated SDQ have not been assessed previously, although preliminary study results were presented by one of the author in 2002.¹⁵

The aim of the present study, therefore, is to examine the psychometric properties of the self-rated SDQ-Kr in a large community sample and to compare these results with those from other countries. For this purpose, the following issues were investigated: 1) scale means and the range of cut-off scores for banding, 2) reliability (internal consistency and temporal stability) of the questionnaire, 3) factor structure of the self-rated SDQ, and 4) discriminant validity to distinguish between community and clinic samples.

METHODS

Participants and procedure

The research was approved by the Ethics Committee of Wonju College of Medicine, Yonsei University. The study was then presented to eight schools, all of which are located in urban areas, six in a large city, and two in a metropolitan city. A total of 2866 youths from two elementary schools (11–12 years), six middle and high schools (13–16 years) participated in the present study by completing the self-rated SDQ during regular classes. The teacher and a research assistant were present during the administration of the questionnaire in order to ensure independent and confidential responding and to provide assistance when necessary.

Youths attending classes for the handicapped or mentally

retarded were excluded from the study. Most of the non-responders were absent from the respective schools on the days the questionnaires were administered. Only 52 questionnaires were discarded because they were incomplete or inappropriately filled in. A total of 2814 (1406 boys and 1408 girls) questionnaires were included in the statistical analysis. The sex distribution of respondents across the grades was even, 49.96% boys and 50.04% girls. To examine temporal stability, the self-rated SDQ was administered to a subsample of 167 youths from two middle schools five to six weeks after the initial assessment.

A clinic sample of 385 psychiatric outpatients (194 boys and 191 girls; mean age 13.8 years) enrolled from Wonju Severance Christian Hospital completed the self-rated SDQ as part of the initial evaluation of youths aged 11 to 16. Eighty-two percent of them ($n=316$) were diagnosed as emotional ($n=143$), behavioral ($n=220$), or developmental disorder ($n=62$) with a need to be treated (mean number of diagnosis was 1.3).

Questionnaires

Korean version of the Strengths and Difficulties Questionnaire

The self- and informant-report versions of the SDQ-Kr can be readily downloaded from the official website of the SDQ (<http://www.sdqinfo.com>) for non-commercial purposes.⁸ The official Korean version of the self-report SDQ was translated and back-translated by one of the authors (JS Ahn) who developed the parent version of the SDQ-Kr in 2003¹⁶ and by Su-Jin Yang.⁸

The SDQ consists of four difficulties subscales and one strength subscale of five items each: emotional symptoms, conduct problems, hyperactivity-inattention, peer problems for difficulties, and prosocial behavior for strength. Items have to be scored on a 3-point scale with 0=not true, 1=somewhat true, and 2=certainly true. Subscale scores can be computed by summing scores on relevant items (after recoding reversed items; range=0–10). A total difficulties score can also be calculated by summing scores of four difficulties subscales (range=0–40). Using cut-off scores, it is possible to build categories comprising normal, borderline, and abnormal problem scores.

Statistical analysis

Distributions of raw scores were used to determine the cut-off scores to identify normal, borderline, and abnormal bandings. To examine the reliability, we calculated Cronbach's alpha coefficients for internal consistency and Pearson's correlations for test-retest stability. As some scales were

judged to be skewed and the normality assumption was not tenable, Spearman's rank correlations were also employed. The independent sample t-test and one way ANOVA were used to compare total difficulties scores and the scales scores across gender and age groups.

In order to evaluate the factorial structure of the SDQ-Kr items, we conducted an exploratory factor analysis (EFA) with varimax rotation. Finally, discriminant validity was examined by using receiver operating characteristic (ROC) curves employing the area under the curve (AUC) as an index of discriminant ability. For this purpose, the underlying assumption was that subjects from the clinic sample were substantially more likely to have psychiatric disorders than were subjects from the community sample. There was no assumption that everyone in the clinic sample had a psychiatric disorder or that everyone in the community sample was free from psychiatric disorders. For the AUC, a score ≤ 0.6 suggests that discrimination is no better than chance; 0.6–0.75 is fair, 0.75–0.90 is good, 0.90–0.97 is very good, and 0.97–1.0 is excellent.¹⁷

Statistical analyses were carried out using the SPSS system for Windows (version 18.0; SPSS Inc., Chicago, IL, USA).

RESULTS

Scale means, gender and age effects, score bandings

The mean subscale scores for the self-rated SDQ subdivided

by gender and age-band are presented in Table 1 and 2, respectively. Boys were rated as having greater difficulties on hyperactivity/inattention ($p < 0.001$), conduct problems ($p < 0.001$), and peer problems ($p < 0.001$). On the emotional symptoms and prosocial behavior subscales, girls were rated as having higher scores ($p < 0.001$ and $p = 0.001$, respectively). There were no gender differences seen on the total difficulties score.

As the homogeneity of variance on the total difficulties and peer problems scores was not supported by Levene's test, Tamhane's test instead of Scheffe's test for post hoc analysis was applied after one way ANOVA for age differences. On the total difficulties score and emotional symptoms subscale, the main effect of age was reflected in an increase in scores as age increased. On the hyperactivity/inattention subscale, elementary school students aged 11–12 years showed lower scores. On the peer problems subscale, youths aged 13–14 years showed lesser difficulties. No main effect of age was noted on conduct problems and prosocial behavior. There were no significant interactions of gender and age for all subscales.

Bandings and cut-offs were estimated from the distributions of raw values in the manner suggested by Goodman et al.⁹ For the total difficulties and subscales scores, cut-offs were calculated by placing approximately 10% of the sample with the most extreme scores in the "abnormal" banding, the next 10% in the "borderline" banding, and the remaining 80% in the "normal" banding. These bandings are shown in

Table 1. Scale means and gender effects for the Korean self-rated SDQ scores (age range 11–16 years)

	Total sample (N=2814)		Boys (N=1406)		Girls (N=1408)		Gender effects
	Mean	SD	Mean	SD	Mean	SD	p-value*
Total difficulties	12.69	4.18	12.83	4.25	12.50	4.11	0.032
Emotional symptoms	3.56	2.09	3.17	1.99	3.93	2.13	<0.001
Conduct problems	2.56	1.32	2.71	1.35	2.40	1.28	<0.001
Hyperactivity/inattention	4.47	2.10	4.70	2.07	4.20	2.09	<0.001
Peer problems	2.11	1.40	2.25	1.47	1.97	1.31	<0.001
Prosocial behaviors	5.60	1.70	5.48	1.70	5.70	1.69	0.001

*independent sample t-test. SDQ: strengths and difficulties questionnaire

Table 2. Scale means and age effects for the Korean self-rated SDQ scores

	A: 11–12 (N=937)		B: 13–14 (N=941)		C: 15–16 (N=936)		Age effects
	Mean	SD	Mean	SD	Mean	SD	p-value
Total difficulties	11.76	4.35	12.80	4.04	13.43	3.99	(A-B, A-C)** (B-C)*
Emotional symptoms	3.04	2.01	3.59	2.08	4.02	2.07	(A-B-C)**
Conduct problems	2.56	1.36	2.56	1.27	2.53	1.34	
Hyperactivity/inattention	4.03	2.04	4.68	2.08	4.65	2.10	(A-B, A-C)**
Peer problems	2.14	1.48	1.97	1.33	2.23	1.37	(B-C)**
Prosocial behaviors	5.65	1.73	5.52	1.65	5.62	1.73	

One-way ANOVA with Scheffe's test for all except for total difficulties and peer problems, which were analyzed with Tamhane's test. * $p < 0.01$, ** $p < 0.001$. SDQ: strengths and difficulties questionnaire, ANOVA: analysis of variance

Table 3. Recommended bandings of raw scores obtained from the Korean self-rated SDQ (N=2814, aged 11–16 years)

	Normal range		Borderline range		Abnormal range		UK*
	Raw scores	Exact %	Raw scores	Exact %	Raw scores	Exact %	
Total difficulties	0–16	80.5	17–18	10.3	19–40	9.2	18
Emotional symptoms	0–5	81.4	6	9.3	7–10	9.3	7
Conduct problems	0–3	77.4	4	14.2	5–10	8.4	5
Hyperactivity/inattention	0–6	83.4	7	8.6	8–10	8.0	7
Peer problems	0–3	83.5	4	10.5	5–10	6.0	4
Prosocial behaviors	5–10	75.7	4	15.4	3–0	8.8	5

*cut-off point on UK normative sample. SDQ: strengths and difficulties questionnaire

Table 4. Internal consistency, temporal stability, and inter-subscale correlations for the Korean self-rated SDQ (Community sample N=2814, clinic sample N=385, aged 11–16 years)

	Internal consistency		Test-retest reliability (N=167)	Inter-subscale correlations Pearson's gamma* (Spearman's rho)				
	Cronbach's alpha			2	3	4	5	6
	Community	Clinic						
1. Total difficulties	0.69	0.78	0.79	0.67 (0.67)	0.51 (0.50)	0.67 (0.65)	0.50 (0.49)	-0.10 (-0.09)
2. Emotional symptoms	0.64	0.83	0.68		0.10 (0.10)	0.16 (0.16)	0.18 (0.19)	0.10 (0.10)
3. Conduct problems	0.28	0.46	0.82			0.20 (0.21)	0.14 (0.12)	-0.04 (-0.05)
4. Hyperactivity/inattention	0.68	0.71	0.72				0.07 (0.08)	-0.18 (-0.18)
5. Peer problems	0.38	0.57	0.79					-0.15 (-0.15)
6. Prosocial behaviors	0.54	0.65	0.70					

*Pearson's correlation coefficient: $p < 0.01$ except for 3×6 ($p < 0.05$), Spearman's rho: $p < 0.01$ for all. SDQ: strengths and difficulties questionnaire

Table 5. Ability of the Korean self-rated SDQ subscale scores to distinguish between community (N=2814) and clinic (N=385) samples

	AUC*	99% confidence intervals		% with abnormal score band	
		Lower bound	Upper bound	Community	Clinic
Total difficulties	0.829	0.793	0.864	9.2%	68.3%
Emotional symptoms	0.741	0.701	0.782	9.3%	47.0%
Conduct problems	0.710	0.671	0.749	8.4%	38.2%
Hyperactivity/inattention	0.753	0.716	0.790	8.0%	41.3%
Peer problems	0.702	0.663	0.741	6.0%	27.3%
Prosocial behaviors	0.438	0.395	0.481	8.8%	18.4%

*a significant AUC score < 0.5 indicates worse than chance. SDQ: strengths and difficulties questionnaire, AUC: area under the curve

Table 3 along with the actual percentage of subjects in each of the three banding categories.

Reliability

Internal consistency coefficients are displayed in Table 4. As shown, for the community sample, Cronbach's alphas approached sufficient levels for the total difficulties score and the emotional symptoms and hyperactivity-inattention subscales. However, alphas for conduct problems (0.28), peer problems (0.38), and prosocial behaviors (0.54) fell below acceptable limits. For the clinic sample, which is more homogeneous in difficulties, the reliability coefficients were higher than those found for the community sample. Nonetheless,

Cronbach's alphas for conduct problems (0.46) and peer problems (0.57) were still insufficient.

For temporal stability, 167 seventh grade students completed the self-rated SDQ for a second time five to six weeks after their first completion. Test-retest correlations of ≥ 0.7 are generally considered reliable. The correlations between these scores are also reported in Table 4. All the coefficients except for emotional symptoms were statistically significant ($p < 0.01$).

As shown in Table 4, inter-subscale correlations among SDQ difficulties scales were weak (0.07–0.20 for Pearson's gamma and 0.08–0.21 for Spearman's rho). This indicates that SDQ subscales assess relatively independent domains of difficulties.

Factor structure

Exploratory factor analysis was conducted employing varimax rotation with Kaiser Normalization. For the community sample, this analysis indeed yielded five factors with eigenvalues >1.0 , which accounted for 40.28% of the total variance. As shown in Table 6, however, inspection of this five-factor solution showed that although emotional symptoms, prosocial behavior, and hyperactivity-inattention emerged as distinct factors, peer problems and conduct problems items did not load on separate factors. Instead, these items loaded on two mixed factors; two items of peer problems loaded on emotional symptoms and two items of conduct problems loaded on peer problems. One of them (7. "I usually do as I am told") was not loaded anywhere.

Discriminant validity

ROC analysis showed that the area under the curve for total difficulties was 0.83 (99% CI 0.79–0.86), which is interpreted as good to discriminate between the community and clinic sample. Applying the recommended bandings for the total difficulties score shown in Table 3, 68.3% of the clinic samples had abnormally high scorers as compared with 9.2% of the community sample. Carrying out separate ROC analyses for four difficulties subscale scores, the AUC values ranged from 0.70 to 0.75, which is interpreted as fair to good. On the contrary, the AUC for prosocial behavior was less than 0.5, which indicates worse than chance. The results for the ROC analyses are summarized in Table 5.

DISCUSSION

In this study we presented the results of the self-rated SDQ, spanning a broad age range and both genders. This is a first attempt to examine the psychometric properties of the Korean version of the self-report SDQ for a normative sample of youths. In addition, based on the distributions of raw scores observed, we defined score bandings recommended to identify the borderline and abnormal ranges for each scale.

Gender and age differences

We noted several gender and age effects from the current results, which require further discussion. In terms of gender differences, girls reported higher levels of emotional symptoms along with higher levels of prosocial behaviors. On the other hand, boys reported higher levels of hyperactivity-inattention and conduct problems along with more difficulties in peer relations.

These gender differences have been consistently documented in a large number of reports using the self-rated SDQ from different countries.¹⁰ Girls score higher on emo-

tional symptoms and prosocial behavior and boys higher on conduct problems.¹⁸ Less consistently, some studies have also found that boys score higher on peer problems and hyperactivity-inattention.^{19,20}

With respect to age differences in this study, younger adolescents aged 11–12 years reported lower levels of hyperactivity-inattention, whereas scores for emotional symptoms and total difficulties increased as age increased. One possible explanation for higher levels of hyperactivity-inattention reported by older adolescents, which is in line with Chinese research, may be that they have to study for more extended periods of time with high pressure generated by college entrance examinations.²⁰

Significant age differences for emotional symptoms and total difficulties were largely consistent with the results of previous studies.^{18–20} Increased cognitive capacity as well as the capability of self-observation for older adolescents may influence perceiving difficulties and self-report of symptoms. Sex and age interaction, which varies across the studies, was not observed in our study findings.

Score bandings

Presenting cut-off points to establish Korean norms for the SDQ are important because even accurately translated questionnaire items to another language can yield somewhat different results.

In the current study, the cut-off points of hyperactivity-inattention and emotional symptoms scales were higher than those reported from other countries.^{18–20} One explanation for the dissimilar results in our sample can be sought by looking at cultural differences in the style of perceiving and reporting the behaviors described. Also, it is possible that actual differences in youths' behaviors are between different populations. Korea is a very competitive society where youths are expected to achieve academically with formidable extra-curricular lessons. Korean youth are likely to be sensitive to attention problems associated with academic achievement and to feel unhappy because of worrying about not satisfying adults' expectations.

On the other hand, the cut-off point of prosocial behavior, a strength domain, was lowest compared with a lot of previous reports including Goodman's UK pilot study.^{9,18–20} This finding was away from our anticipation that a collectivistic value orientation, a characteristic of Korean culture, is positively related to prosocial behavior.²¹ Korean society has been rapidly changing with reduction in the number of children, materialism, and achievement orientation, which makes it difficult to maintain a sense of belongingness and integrity with other people.²² Considering that individualism overshadows collectivism in Korea and youths invest so much of

Table 6. Factor structure and original scales of the Korean self-rated SDQ in the community (N=2814)

Extracted factor	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Initial eigenvalue	3.36	2.49	1.73	1.38	1.10
Explained variance (initial)	13.44	9.97	6.93	5.54	4.41
Explained variance (rotated)	9.51	9.27	9.24	7.19	5.08
Emotional symptoms					
3 Somatic complaints	0.468				
8 Worries	0.699				
13 Unhappy	0.653				
16 Clingy	0.628				
24 Fears	0.561				
Prosocial behavior					
1 Considerate		0.527			
4 Shares		0.478			
9 Caring		0.621			
17 Kind to Kids		0.505			
20 Helps out		0.596			
Hyperactivity/inattention					
2 Restless			0.741		
10 Fidgety			0.707		
15 Distractible			0.717		
21 Reflective*			0.456		
25 Persistent*			0.496		
Peer problems					
6 Solitary				0.429	
11 Good friend*	-0.466				
14 Popular*	-0.495				
19 Picked on, bullied				0.632	
23 Best with adults				0.453	
Conduct problems					
5 Tempers					0.486
7 Obedient*					
12 Fights					0.637
18 Lies, cheats				0.573	
22 Steals				0.517	

Five-factor solution; only rotated loadings with absolute values ≥ 0.35 are reported here. Total explained variances for the community are 40. *scores on these items were inverted before being entered in the analysis. SDQ: strengths and difficulties questionnaire

themselves emotionally, the reported lower levels of prosocial behavior might not be surprising. Further research is, however, needed to explain this contradictory finding.

Reliability

With regard to reliability, Cronbach's alpha coefficients for the total difficulties, emotional symptoms, and hyperactivity-inattention scales were higher than 0.60 which is considered sufficient. Although alpha coefficients for the conduct prob-

lems and peer problems scales were less than satisfactory, these findings are in line with previous studies using different language versions of the SDQ.^{10,19,23,24}

It is suggested that the low internal consistency of the peer problems scale may be due to many youth having a low level of awareness of their problems with peers.²⁴ Actually, we found some students endorsing or negating the items 11 ("I have one good friend or more") and 14 ("Other people my age generally like me") with an additional remark 'I am not

sure' from the current study. Another possible explanation is that the peer problems as well as conduct problems scales consist of items which measure heterogeneous properties and are too short to convey the exact meaning, which are also reflected in the factor loadings.

For the conduct problems scale, many students of our sample who negated the other four items of the scale rated 'certainly not true' on item 7 ("I usually do as I am told") which measures 'obedient', unexpectedly. They might be interpreting it as 'weak mind' or 'docility' instead of 'rule observance', which could make the internal consistency of the conduct problems scale notably lower, we assume. To confirm our assumption, further research supplemented with a second stage interview is needed.

With regard to temporal stability, the test-retest correlations over a five to six week interval were generally acceptable. Specifically, for the conduct problems and the peer problems scales, correlation coefficients were satisfactory (correlations >0.75), which suggests a consistent way of reporting on those behaviors. From these results we suggest that notably low alphas of these two scales may be taken as an indication of a low level of validity but not as an indication of low reliability in the sense of inconsistent reporting on these problems.²⁴

The pattern of intercorrelation among the Korean self-rated SDQ scales was similar to the results from some previous studies,^{10,20,24} but the correlation coefficients were much lower. As expected, the prosocial behaviors scale was inversely correlated with each of the difficulties domains except for the emotional problems scale. Taking into account that greater socialization and sensitivity to interpersonal concerns for girls increases their vulnerability to emotional problems, as mentioned above, this finding is not surprising.²⁵ Future research should, however, examine whether the assumption is applicable to all adolescents regardless of gender.

Validity

Factor analysis

Consistent with previous studies, factor analyses showed that nearly all items loaded primarily on the predicted five factors. Although the original five scales were confirmed in our data, as mentioned in the results, item 7 of the conduct problems and items 11 and 14 of the peer problems were debatable. A shared feature of these items is the positively worded question. The inclusion of these items was originally intended to increase the acceptability of the SDQ to respondents, making it particularly suitable for use in non-clinical, epidemiological studies. However, the presence of both positively and negatively worded items in a certain domain can confound the factor structure. These sorts of results have

been noticed and discussed in other studies.^{12,23,26,27} There are mixed and heterogeneous results about the factor structure of SDQ.²⁷ Some studies supported five factor model,^{19,20,26} while the other studies did support three or four factor model.^{8,27}

Goodman reported that the positive qualities of these items seemed to load highly onto the prosocial factor.²³ Several studies have also pointed to the presence of such a 'positive construal factor' in the SDQ, with substantial loadings for the positively worded items that belong to different scales.^{14,18,19,23} In the present study, however, the positive construal factor was not supported by items 11 'good friend' and 14 'popular' loaded onto the emotional factor rather than the prosocial factor. Youths responding on these items with 'not true' also tended to negate items of the emotional symptoms domain, which seems to be unique for all previous research. Item 7 'obedient', another positively worded item, was not loaded onto any factor for the community sample.

Judging from the results of factor analysis and insufficient internal consistency for the conduct problems and peer problems scales, it could be that the respondents of our sample were not interpreting these items in the way intended by the developer of the questionnaire. This may be due to some cultural differences as several recent studies with the parent-rated or self-rated SDQ have suggested.^{13,14,18} For example, an Arabic study on children within the Gaza Strip using the parent-rated SDQ reported that the emotional symptoms and peer problems scales seemed to be either more heterogeneous or more multifactorial than is typically seen in Western cultures.¹³ Therefore, it might very well be that changing the wording of these items with careful consideration for the cultural differences would increase both the factor loadings and Cronbach's alphas.

Discriminant validity

With respect to discriminant validity, the total difficulties score and five difficulties scales scores well distinguished between a community and a clinic sample. Especially on the total difficulties score, the clinic subjects were over seven times more likely than community subjects to have a score of abnormal range. If we selected clinic youths according to their presenting problems and compared the corresponding difficulties scales scores against the community sample scores, the discriminative power of those scales would be higher.

On the basis of the findings reported here, as Goodman suggested, researchers could appropriately use the self-rated SDQ-Kr to examine group differences, though they could not expect to make accurate diagnoses on individual subjects. Similarly, clinicians could sensibly use the self-report SDQ-Kr to assess youths' degree of awareness of their own problems, even though the self-rated scores by themselves would

not be an adequate basis for assigning individual diagnoses.⁹

In keeping with that the development of the self-rated SDQ was motivated to detect youths' psychopathology in community samples,²³ the present results showed that the self-rated SDQ-Kr discriminated well between youths with and without the need for clinical help.

Limitations and implications

The current study presents data from a large representative population, with a wide age range and a response rate of 98%. Data were analyzed separately for age band and gender if necessary. In addition, the assessment of clinical youths was included and compared with that of the community sample. The results suggested that the self-rated SDQ-Kr can be applied as a valuable screening instrument for a nationwide epidemiological study.

While sample size, response rate, and inclusion of the clinic sample are the major advantages of this study, there are a number of limitations that warrant discussion.

First, a major limitation of the study is the lack of cross-informant comparisons. Comprehensive assessment requires data from multiple informants when possible, as suggested by many researchers.^{19,26,28} Self-reports are more prone to distortions arising from social desirability and over- or underestimation of the frequency or severity of problems than external observers' assessments, particularly when complete anonymity is not guaranteed as in this study. Obviously, participation of at least two informants and additional use of second stage interviews would have improved the sensitivity and specificity of the self-rated SDQ-Kr as a screening instrument for the community population.

Second, a concurrent questionnaire such as the YSR, a supposed gold standard for self-appraisal of adolescents' psychopathology, was not included in order to validate the total difficulties and four difficulties scales. Goodman's study and other language studies found that the self-report SDQ and YSR are equally able to distinguish between psychiatric disorders within the clinic sample as well as behavior problems in the community sample.^{10,12,23,29} To examine concurrent validity and whether the SDQ could be an alternative for the YSR which is widely used at community mental health centers in Korea, further study is needed.

Third, the self-rated SDQ-Kr scales were only validated against the clinic sample as a whole. Comparison of the difficulties scales scores with psychiatric diagnoses as measured through standardized diagnostic interviews would certainly have strengthened this study. However, in our study the self-rated SDQ-Kr was employed as a screening device for determining the overall psychosocial conditions of a large group of youths in schools with no intention to reveal the preva-

lence of psychiatric disorders. In Korea, schools are developing various programs with the primary aim of improving the well-being of students. The use of the self-rated SDQ-Kr as a screening device is efficient and economical to gather valuable information.

In conclusion, although test-retest reliability and discriminant validity appear good, the psychometric properties of the self-report version of the SDQ-Kr are less satisfactory, especially for lower internal consistency on some subscales. Some items of SDQ-Kr need to be further evaluated and revised in order to improve the psychometric properties and to capture the originally intended constructs. In our study, the self-report SDQ-Kr was employed as a screening device for determining possible mental health problems in children, not for revealing the prevalence of psychiatric disorders. Therefore, we might say that the total difficulties score of the SDQ-Kr could be used with more confidence for screening. In Korea, schools develop several programs to improve the psychosocial conditions of youths. It is hoped that the establishment of national norms of the self-report SDQ-Kr will further facilitate and encourage its application in school mental health programs as well as research in child psychiatric settings.

REFERENCES

1. Meltzer H, Gatward R, Goodman R, Ford T. Mental health of children and adolescents in Great Britain. *Int Rev Psychiatry* 2003;15:185-187.
2. Costello EJ, Egger H, Angold A. 10-year research update review: the epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *J Am Acad Child Adolesc Psychiatry* 2005;44:972-986.
3. Graham PJ, Verhulst FC, Turk J. *Child Psychiatry: A Developmental Approach*. Oxford: Oxford University Press; 1999.
4. Matsuura M, Okubo Y, Kojima T, Takahashi R, Wang YF, Shen YC, et al. A cross-national prevalence study of children with emotional and behavioural problems—a WHO collaborative study in the Western Pacific Region. *J Child Psychol Psychiatry* 1993;34:307-315.
5. Yoo HI, Cho SC, Kim BN, Kim SY, Shin MS, Hong KE. Psychiatric morbidity of second and third grade primary school children in Korea. *Child Psychiatry Hum Dev* 2005;36:215-225.
6. Achenback TM. *Manual for the Youth Self-Report and 1991 Profile*. Burlington, VT: Dept. of Psychiatry, University of Vermont; 1991.
7. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry* 1997;38:581-586.
8. SDQ: Information for researchers and professionals about the Strengths and Difficulties Questionnaires. Available at: <http://www.SDQinfo.com>. Accessed August 1, 2014.
9. 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;7:125-130.
10. Muris P, Meesters C, van den Berg F. The Strengths and Difficulties Questionnaire (SDQ)—further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *Eur Child Adolesc Psychiatry* 2003;12:1-8.
11. Becker A, Hagenberg N, Roessner V, Woerner W, Rothenberger A. Evaluation of the self-reported SDQ in a clinical setting: do self-reports tell us more than ratings by adult informants? *Eur Child Adolesc Psychiatry* 2004;13(Suppl 2):II17- II 24.

12. Koskelainen M, Sourander A, Kaljonen A. The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. *Eur Child Adolesc Psychiatry* 2000;9:277-284.
13. Thabet AA, Stretch D, Vostanis P. Child mental health problems in Arab children: application of the strengths and difficulties questionnaire. *Int J Soc Psychiatry* 2000;46:266-280.
14. Dickey WC, Blumberg SJ. Revisiting the factor structure of the strengths and difficulties questionnaire: United States, 2001. *J Am Acad Child Adolesc Psychiatry* 2004;43:1159-1167.
15. Ahn JS, Jang HM, Kim MH, Goodman R. Development of a Korean version of the self-reported Strengths and Difficulties Questionnaire. Seoul, Korea: Korean Neuropsychiatry Association Annual Meeting; 2002.
16. Ahn JS, Jun SK, Han JK, Noh KS, Goodman R. Development of a Korean version of the Strengths and Difficulties Questionnaire. *J Korean Neuropsychiatr Assoc* 2003;42:141-147.
17. Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988;240:1285-1293.
18. Koskelainen M, Sourander A, Vauras M. Self-reported strengths and difficulties in a community sample of Finnish adolescents. *Eur Child Adolesc Psychiatry* 2001;10:180-185.
19. Ronning JA, Handegaard BH, Sourander A, Mørch WT. The Strengths and Difficulties Self-Report Questionnaire as a screening instrument in Norwegian community samples. *Eur Child Adolesc Psychiatry* 2004;13:73-82.
20. Yao S, Zhang C, Zhu X, Jing X, McWhinnie CM, Abela JR. Measuring adolescent psychopathology: psychometric properties of the self-report strengths and difficulties questionnaire in a sample of Chinese adolescents. *J Adolesc Health* 2009;45:55-62.
21. Triandis HC. Individualism-collectivism and personality. *J Pers* 2001; 69:907-924.
22. Hong KM. Impacts of rapid social and family changes on the mental health of children in Korea. *Psychiatry Investig* 2006;3:6-25.
23. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry* 2001;40:1337-1345.
24. van Widenfelt BM, Goedhart AW, Treffers PD, Goodman R. Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *Eur Child Adolesc Psychiatry* 2003;12:281-289.
25. Leadbeater BJ, Kuperminc GP, Blatt SJ, Hertzog C. A multivariate model of gender differences in adolescents' internalizing and externalizing problems. *Dev Psychol* 1999;35:1268-1282.
26. Van Roy B, Veenstra M, Clench-Aas J. Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. *J Child Psychol Psychiatry* 2008;49:1304-1312.
27. Stevanovic D, Urban R, Atilola O, Vostanis P, Singh Balhara YP, Avicenna M, et al. Does the Strengths and Difficulties Questionnaire - self report yield invariant measurements across different nations? Data from the International Child Mental Health Study Group. *Epidemiol Psychiatr Sci* 2014;1-12.
28. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Int Rev Psychiatry* 2003; 15:166-172.
29. Goodman R, Scott S. Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful? *J Abnorm Child Psychol* 1999;27:17-24.