Validity of the self-report version of the strengths and difficulties questionnaire in Yemen

Mohammed H. Almagrami, MSc, MD, Abdullah Y. Shuwail, MBChB, MD.

ABSTRACT

Objective: The main objective of the study was to determine the validity of the Arabic version of the Strengths and Difficulties Questionnaire (SDQ) –Self-Report version in Yemen.

Methods: Following translation and back translation of the scale into the Arabic language by a panel of experts, psychometric properties were assessed using 600 students from schools (low risk group) and 57 patients attending the Mental Health Hospital in Sana'a, Republic of Yemen (high risk group) from January through to April 2002. The age range of the 2 groups was from 12-17-years-old. Discriminative validity, concurrent validity and factorial validity were studied.

Results: The difference between the means of the "total difficulties" scores in both groups was highly significant

(*p*<0.001); the area under the curve (AUC) of the total scores and subscales scores were ranged from 0.77-0.89. The chance corrected agreements between the clinical diagnosis and SRQ subscales' prediction were significant. The sensitivity and specificity were 72% and 55%. Factor analysis yielded 5 dimensions of emotional symptoms, conduct disorders, hyperactivity, peer problems, and pro-social items.

Conclusion: Results of this validation study suggest that the Arabic version of the SRQ of SDQ is valid in Yemen; and it can be a useful tool for investigating childhood behavioral and emotional disorders at clinical settings.

Saudi Med J 2004; Vol. 25 (5): 592-601

In psychiatric and psychological researches and clinical practice especially in developing countries, there is a great need for brief, cheap, validated and effective instrument for detecting emotional and behavioral disorders in children and adolescents. Although parent and teacher ratings have provided extensive data on common childhood disorders, there are several reasons that parent and teacher ratings must be supplemented with self report ratings as children enter adolescence.1 These self reports instruments are advantageous in that they are inexpensive and easy to administer. development Moreover, of self-report the

questionnaires can lead to greater description and accuracy in identification, diagnosis, and treatment of children. The Strengths and Difficulties Questionnaire (SDQ) is a recent and brief behavioral screening instrument developed by Goodman² in 1997 to meet the needs of researchers, clinicians, as well as educationalists. It provides balanced coverage of children and young people's behaviors, emotions, and relationships. There are two versions of SDQ: an informant rated version that could be completed by one of the parents or the teacher of 4-16-years-old; as well as a self-report version which could be filled in by the child himself

From the Department of Psychiatry (Almaqrami), Faculty of Medicine, University of Dhmar, Department of Psychiatry (Shuwail), Faculty of Medicine, University of Sana'a, Republic of Yemen.

Received 7th October 2003. Accepted for publication in final form 9th February 2004.

Address correspondence and reprint request to: Dr. Mohammed H. Almaqrami, Assistant Professor of Psychiatry, Faculty of Medicine, University of Dhmar, PO Box 20908, Republic of Yemen. Tel. +967 (1) 240368. Fax. +967 (1) 262426

from 11-16-years-old. The performance of the SDQ could potentially have been undermined by 3 of its design features: inclusion of strengths as well as difficulties, a compact presentation on just one side of the paper, and the use of an identical questionnaire for both parents, teachers, and self-reports. The 25 items are divided between 5 scales of 5 items each, generating scores for problems, emotional symptoms, conduct hyperactivity inattention, peer problems, and pro-social behavior; all but the last are summed to generate a total difficulty score. The extended versions of SDQ include the 25 core items plus a brief impact supplement that asks whether the respondent thinks that the child or youth has a problem, and if so, inquires further regarding overall distress, social impairment, burden, and chronicity. The validity of the SDQ has been investigated since 1997 in developed as well as developing countries.²⁻⁴ However, only one pilot study has confirmed the ability of the self-report version of the SDQ to discriminate satisfactorily between the high risk and the low risk samples in the British community.5

Although the Self-Report SDQ was proved to be valid, easy to administer, score, and analyze, its validation has not been studied yet in Yemen. Hence, it may work in Yemen where 54% of the total population is under 15-years of age, the level of literacy is low, and child mental health professionals are very limited or even not existed. The aim of this study is to test the validity and of the Arabic version of Self-Report version of the Strengths and Difficulties Questionnaire (SDQ) in Yemen.

Methods. The original English Self-Report Questionnaire (SRQ) of the extended version SDQ was given to a panel of 6 Yemeni and Arabic psychiatrists to be translated into Arabic language. (Appendix 1) After that, a bilingual Yemeni journalist back translated the SRQ into English. (Appendix 2) The semi final draft was prepared by the principle investigator and revised with the original author of the SRQ. The final draft have been just completed in Yemen after conducting a pilot testing into groups of 12-17-years-old Yemeni students to test the suitability of the translation to their level of understanding. Fifty-seven patients recruited from the Mental Health Hospital (MHH), which is the only mental hospital in Sana'a city (capital of Yemen). Subjects were consecutive new patients of both sexes, between the age-range of 12-17-years, who were referred to the hospital during the period of the study. The study was conducted from January through to 30 April 2002. Although the sample was not intended to be representative, it was selected from 4 (2 states and 2 privates) urban schools inside Sana'a city, 2 semiurban schools (near Sana'a), and 4 rural schools (outside Sana'a). Urban and semiurban schools were obtained through a multistage sampling while rural schools were obtained according to the accessibility and coordination of their administrators. Six hundred equal equal numbers of boys and girls from 6th grades (6th, 7th, 8th and 9th in the primary level and first and second in the secondary level) were randomly selected from these 10 schools. The Arabic translation of the Self-Report SDQ was administered to all children from the high risk sample. The questionnaire was scored according to the standardized instructions. Clinicians assigned the patients an International Classification Data (ICD) 10 diagnosis depending on detailed clinical assessments. In the low risk sample, the translated Self-Report SDQ was administered to the all students. Data were entered to computer using the SDQ computer program and the Statistical Package for Social Sciences version 11, was used for statistical analysis.

Results. It was found that the mean (SD) of the total difficulties scores was 12.0 (5.0) for the community sample, and 18.4 (6.7) for the clinic sample. The difference was highly significant (t= 8.9, p < 0.001). It was found that the total difficulty scores were reasonably normally distributed. The ability of the SRO total scores and subscales' scores to distinguish between community and clinic subjects were examined using receiver operating curves (ROC), employing the areas under the curves (AUC) as the index of discriminant validity. Figure 1 presents AUC for the total difficulty and subscales' scores of SRQ. The AUC is the most commonly used measure of overall fit in the ROC analysis. It is interpreted as the probability of correctly classifying a randomly selected pair of subjects where one is normal and one is the case. A scale with an AUC of 0.8 or more is considered to be useful for clinical prediction. In this study, the AUC for the total SRQ total scores and its subscales were ranged from 0.77-0.89. Another aim for using the ROC is to find the point at which sensitivity and specificity is optimal in relation to each other. Sixteen to seventeen were found to be cut-points, which were the best. These points provide sensitivity of 61% and specificity of 81% for the total SRQ scale. Therefore, 112 (18.9%) children in the community sample were found to be likely psychological clinically having significant problems. However, when the predictive algorithm of the SRQ was used and impact scores, therefore, have been taken into account, 7% was found to be 'probable', 13% was 'possible' and 80% was unlikely to have clinically significant mental health problems. Table 1 presents the cross-tabulations of

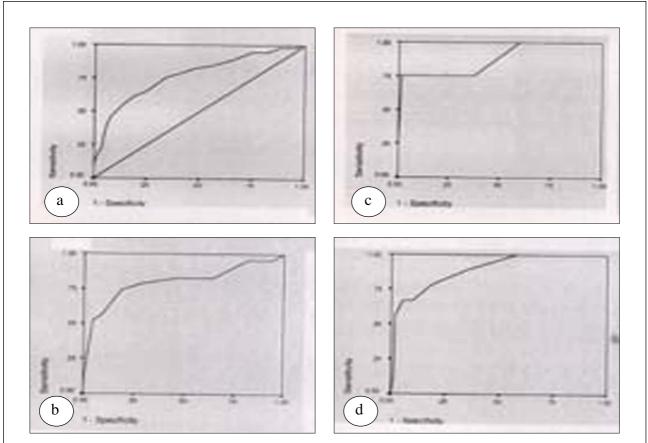


Figure 1 - Receiver operating characteristic curve showing (a) Total difficulty scores. (b) Conduct subscale. (c) Hyperactivity subscale. (d) Emotional subscale.

clinical diagnoses and SRQ predictions of conduct, and emotional disorders in the clinic sample. Each clinically into 'absent', disorder was rated 'borderline', or 'definite', while SRQ predictive 'unlikely', 'possible', or algorithm generates 'probable' classification. For conduct and emotional disorders the level of chance corrected agreement between clinical diagnosis and SRQ prediction were statistically significant (the Kendall's tau b=0.39, p < 0.001 and 0.27, p = 0.02). For example, when a child had a 'definite' clinical diagnosis of conduct disorder, the SRQ prediction for the conduct disorder was 'probable' for nearly all of them (8/9). Although, false negative was only one of 9, false positives were high (15/23). By contrast, in emotional disorder over half of children who were diagnosed clinically as having 'definite' disorder had been predicted as 'probable' by the SRQ (12/23). However, if we combine those who were predicted as 'probable' and 'possible' by the SRQ, they will represent 83% (19/23) of the definitely diagnosed patients. In addition, there were more false positives than false negatives. Hyperactivity

subscale' scores were not included in this part of validation due to that, both self and teacher reports were needed for the hyperactivity disorder to predicted as 'probable' by SDQ. In order to estimate sensitivity, specificity, positive predictive and negative predictive values, both predictions of both the SRQ algorithm and the clinical diagnosis were modified into 2x2 table. In clinical diagnosis, definite disorders were considered positives while absent and borderline disorders were considered negatives. Similarly, in SRQ predictions' algorithm, unlikely and possible were evaluated as negatives whereas probable predictions were counted as positives. It was found that the SRQ can identify more than 70% of patients with any diagnosis and conduct disorders, but only 52% of emotional disorders. Furthermore, except for the total SRQ, the negative predictive values were greater than positive predictive values. In addition, the sensitivity for the total SRQ scale and conduct disorder subscale were higher than specificity, while the reverse was true for the emotional subscale. A principal component factor analysis of the 25 items

Table 1 - Cross tabulations of clinical diagnosis and self-report questionnaire predictions.

conduct disorder	Absent	Borderline	Definite	Total
SRQ prediction of conduct disorder				
Unlikey	22	2	1	25
Possible	8	1	-	9
Probable	11	4	8	23
Total	41	7	9	57
Clinical diagnosis of emotional disorder				
emotional disorder SRQ prediction of	15	0	4	19
emotional disorder SRQ prediction of emotional disorder	15 6	0 0	4 7	19 14
emotional disorder SRQ prediction of emotional disorder Unlikely				

of the SRQ extracted 5 factors with total explained variance 41.1%. Factor one was the major factor and account for 12.7% of the total variance. It was defined by the 5 items of emotional symptoms and by 2 more items related to them. Factor II accounted for 7.4% of the total variance and included 4 of the 5 of pro-social items together with 4 of the positively framed items. Factor III accounted for 5.5% of the total variance was related to the conduct disorders. Factor IV accounted for 5.1% of the total variance and was predominantly related to peer problems. Factor V was related the hyperactivity and account for 4.9% of the total variance.

Discussion. There were 2 approaches to test the discriminative ability of SRQ of SDQ. First, the highly significant differences on means (SD) of the total difficulty scores between clinic and community samples of 12-17-year-olds. Thus on the total difficulty scores, the clinic sample had a mean that was 1.2 standard deviations higher than the community sample. These findings support the discriminative validity of the SRQ. The second approach was the use of the area under the ROC. As shown in Figure 1, the total difficulties, conduct, hyperactivity, and emotional sub-scales of the SRQ can distinguish well between the clinic and the community samples. For each mentioned subscale the AUC was highly significant (p=0.001, 0.001,0.009, and 0.001). For the sake of trans-cultural comparisons, the AUCs for the total difficulties and conduct problems subscale in Yemeni population were greater than that for the corresponding subscales in Bangladeshi population (0.77 and 0.89 in Yemen compared to 0.54 and 0.72 in Bangladesh).6 By contrast, the AUCs for the hyperactivity and emotional subscales were less in Yemen than in Bangladesh (0.80, and 0.88 in contrast to 0.87, and 0.92). In other words, the SRQ can detect total psychiatric difficulties and conduct problems in Yemen better than in Bangladesh and the reverse is true regarding hyperactivity and emotional problems. According to the ROC, 19% (above the cut offs) of the community sample may have clinically significant mental health problems. Although no one can on the basis of this type of study determine the prevalence rates of childhood psychiatric disorders, it can be considered as a useful pointer for further systematic and intensive work. Although, this finding is broadly in line with results obtained from regional international studies,7 SRQ prediction reduce the percentage to only 7%. This reduction was reported in several cross cultural studies.8 Case definition in psychiatry depends currently on social impairment or severity of distress. According to the results of this study, the concurrent validity of the self report SDQ was adequate. The whole SRQ was able to predict a broad spectrum of child psychiatric morbidity in metal health hospital with a whole sample sensitivity of 72% and specificity of 55%. Despite the very good sensitivity (89%), the conduct disorder subscale showed a weak positive predictive value (35%), which could be considered rather a substantial finding in this study. Yemeni adolescents may prefer self report questionnaire to display their conduct problems rather than face to face interview. Such reluctance is expected in a society where an authorized figure is highly respected. Furthermore, this finding could imply an important clinical implication in terms of the best way to gather information regarding conduct behaviors in our society. By contrast, emotional subscale revealed rather a fair sensitivity, specificity, and positive predictive value.

On direct Oblimin rotation, a 5 factor solution yielded the best description of the 25 items SRQ in this Yemeni sample. Generally, findings in this study were similar to other factor analytic studies of SRQ. The 5 factor solutions which were extracted by the factor analysis are predominantly: emotional, pro-social, conduct, peer, and hyperactivity factors. The 5 items of emotional symptoms which loaded in factor one were exactly similar to those found by Goodman.³ However, The loading of the item "I get very angry and often lose temper"with the emotional symptoms might be explained by that, anger reaction could be considered within the emotional profile in the Yemeni society rather than as a sign of aggression. Close finding was also

found in the Swedish study.8 Similarly, the item "I am usually on my own. I generally play alone or keep to my self" (solitary) could be considered as one of the internalizing symptoms. In factor II which appears to identify pro-social items, the most striking feature is that, 4 of the 5 positively phrased items such as "I think before I do things" (Reflective), "I finish the work I'm doing" (Persistent), "I usually do as I am told" (Obedient) and "I have one good friend" (Good friend) are loaded together with the pro-social items. pro-social items are "I try to be nice to other people. I care regarding their feelings" (Considerate), I am helpful if someone is hurt, upset or feeling ill" (Caring), "I am kind to younger children" (Kinds to kids), and "I often volunteer to help others" (Helps out). These items were added to the questionnaire in order to address the strengths of a child. It has been reported in previous studies that adding some positively phrased items to the questionnaire can increase its acceptance to the respondents and consequently, its response rates. On the other hand, the focus of the questionnaire only on negative aspects of the child was found to be rather disappointing to respondents.9 Moreover, 5 more positively framed items addressing some aspects of child psychopathology (the reversibly coded items) were incorporated also into the items of the questionnaire in an attempt to avoid the halo effects and respondent bias. However, this rephrasing of child psychopathology items in a positive way may led Yemeni adolescents to misunderstand these items. Especially in schools where desirability is expected to be high, adolescents who read positively framed statements might not be able to distinguish well between those items related to pro-social activities from those addressing his/her psychopathology. In other words, adolescents may think that these positively stated items are addressing also pro-social activities rather than his/her psychological problems. Whether this was the case in the present study, so they are loaded together in the same factor, it needs really further study. Interestingly, 4 of these 5 positively framed items were also loaded together with the pro-social factor in the previous British study.³ In factor III, the item regarding being bullied or picked was included among the symptoms of conduct disorder (lies, steals, and fights) rather than among the peer problem. However, It is rather expected to find the concept of bullying among the spectrum of conduct symptoms. Similarly, the item "I usually share with others (food, games, pens, for example) was loaded in factor IV which is the factor of peer items. Unfortunately, the lines between items and factors are always not magic and some statements may be interchangeable. Finally, the items of restlessness, fidgety, and distraction were predominantly loaded on the factor of hyperactivity. To sum up, factor

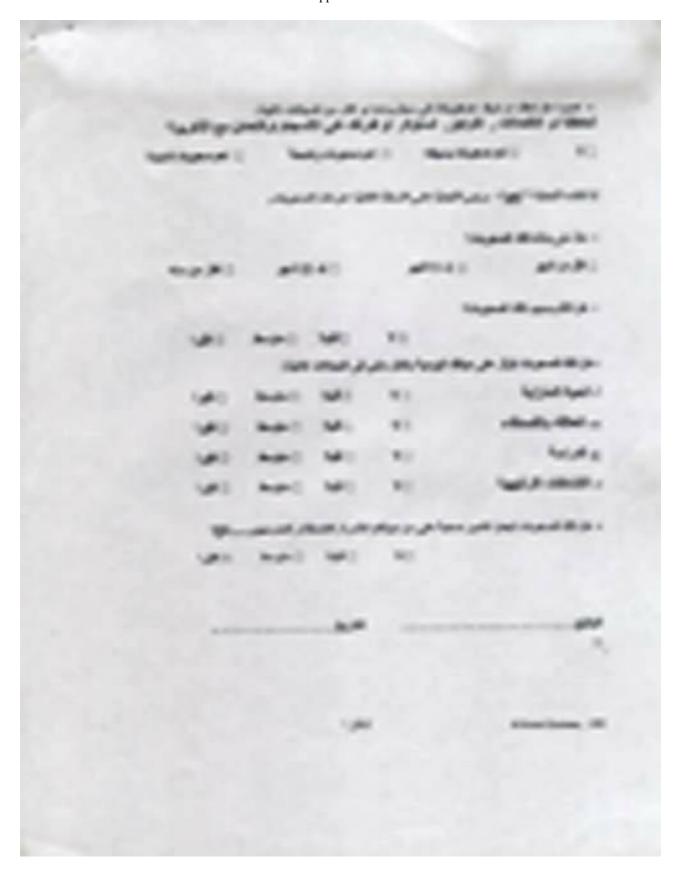
structure for the SRQ in this study could be considered as similar but not identical to that established in the United Kingdom. Further, study will be required using confirmatory factor analysis in order to examine the nature of differences of item loadings among Yemeni and British adolescents. The results of this study are limited by several factors. The first, clinical assessment and diagnosis of children in the clinic sample were dependent on clinical experience of the psychiatrist rather than on standardized interview. The second, sample size for the clinic sample was low. The third, children in the community sample were not clinically assessed in order to examine the predictive validity of the SRQ in a community setting. Final limitation is that, psychosocial risk factors are not included in this study. One way of evaluating construct validity is by looking to how far psychopathology correlated to various risk factors between cultures. The current study showed that, SRQ of SDQ would be a useful tool in the clinical practice of child and adolescent mental health services in Yemen. It can predict efficiently the presence of childhood psychopathology especially conduct problems, hyperactivity, and emotional disorders. Moreover, this prediction would be useful in the first assessment of a child to highlight his need, direct clinician to the probable diagnosis, and guide him to the effect of intervention(s) in the follow up process. The results of this study suggest that the Arabic version of SRQ is appropriate for use in screening purposes such as 2 stage surveys in Yemen. However, further researches on the use of diagnostic instruments in different settings are required. Large standardized interview, clinic sample, conjunction with multiple informants should be Direct in particular, work might useful used. towards investigating the differential effect of different risk factors on validity. To sum up, results of this validation study suggest that the SRQ of SDQ is promising in Yemeni community: It discriminates appropriately between referred and non-referred samples. Moreover, It has been demonstrated that the SRQ is capable in detecting childhood emotional and behavioral disorders in clinical settings. Generally, 5 factors structure corresponding to the original 5 subscales were found.

References

- Conners CK, Wells KC, Parker JDA, Sitarenios G, Diamond JM, Powell JW. A new Self-Report Scale for assessment of adolescent psychopathology: factor structure, reliability, validity, and diagnostic sensitivity. *J Abnorm Child Psychol* 1997; 25: 487-497.
- Goodman R. The Strengths and Difficulties Questionnaire: a research note. J Child Psychol Psychiatry 1997; 38: 581-586.
- Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. J Am Acad Child Adolesc Psychiatry 2001; 40: 1337-1345.

- 4. Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The Development and Well-Being Assessment: description and initial validation of an integrated assessment psychopathology. J Child Psychol Psychiatry 2000a; 41: 645-655.
- 5. 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 7: 125-130.
- 6. Goodman R, Renfrew D, Mullick M. Predicting type of Psychiatric disorder from Strengths and Difficulties Questionnaire scores in child mental health clinics in London and Dhaka. Eur Child Adolesc Psychiatry 2000; 9: 129-134.
- 7. Bird HR, Canino G, Rupio-Stipec M, Gould MS, Ribera J, Sesman M et al. Estimates of the prevalence of childhood maladjustment in a community survey in Puerto Rico. Arch Gen Psychiatry 1988; 45: 1120-1126.
- Goodman R. The extended version of the Strengths and Difficulties Questionnaire as a guide the child psychiatric caseness and consequent burden. *J Child Psychol* Psychiatry 1999; 40: 791-801.
- 9. Goodman R. A modified version of the Rutter parent questionnaire, including extra items on children's strengths. **Ĵ Child Psychol Psychiatry** 1994; 35: 1483-1494.
- 10. Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and difficulties Questionnaire to screen for child psychiatric disorders in a community sample. Br J Psychiatry 2000; 177: 534-539.





VN			
Your Name			Male/Female
Date of Birth	Not True	Somewhat True	Certainly True
I try to be nice to other people. I care about their feelings			
I am restless, I cannot stay still for long			
I get a lot of headaches, stomach-aches or sickness			
I usually share with others (food, games, pens etc.)			
I get very angry and often lose my temper			
I am usually on my own. I generally play alone or keep to myself			
I usually do as I am told			
I worry a lot			
I am helpful if someone is hurt, upset or feeling ill			
I am constantly fidgeting or squirming			
I have one good friend or more			
I fight a lot. I can make other people do what I want			
I am often unhappy, down-hearted or tearful			
Other people my age generally like me			
I am easily distracted, I find it difficult to concentrate			
I am nervous in new situations. I easily lose confidence			
I am kind to younger children			
I am often accused of lying or cheating			
Other children or young people pick on me or bully me			
I often volunteer to help others (parents, teachers, children)			
I think before I do things			
I take things that are not mine from home, school or elsewhere			
I get on better with adults than with people my own age			D-
I have many fears, I am easily scared			
I finish the work I'm doing. My attention is good			
Do you have any other comments or concerns?			
Do you have any other comments or concerns?			

