

Validasi Klinik *Strengths and Difficulties Questionnaire* (SDQ) sebagai Instrumen Skrining Gangguan Tingkah Laku

Mistety Oktaviana¹, Supra Wimbarti²

Fakultas Psikologi Universitas Gadjah Mada

Abstract. Conduct Disorder is one of the most common problems often found in children taken to Puskesmas. Based on SIKM (Mental Health Information System) from 2011 to May 2013, 46.3% of those taken to Puskesmas in Sleman and Yogyakarta were identified with behavioral and emotional disorder. Therefore, screening instrument is needed for early identifying conduct disorder in them. Strengths and Difficulties Questionnaire (SDQ) has been developed to meet the need. The SDQ is a 25-item questionnaire about the positive and negative attributes of children and adolescents (4-16 years old). This study aimed to adapt the English version of SDQ-Teacher Reports (TR) into Indonesian version and to examine psychometrical properties and clinical validation of the Indonesian version for screening conduct disorder. The subjects were 161 Elementary students from Sleman and Yogyakarta, boys and girls of 7-13 years old of age. The reliability coefficient using *Alpha Cronbach* was $\alpha=0,773$, and the construct validity using Principal Axis Factoring (PAF) showed that SDQ-TR had six structural factors. Clinical validation using Receiver Operating Curve (ROC) revealed that the value of sensitivity were 0.67 with spesificity 0.68 (for the optimum cut-off point ≥ 5), and using Likelihood Ratio (LR) showed LR (+) was 2.09 and LR (-) was 0.49.

Keywords: children, clinical validation, conduct disorder, SDQ-TR

Abstrak. Gangguan tingkah laku adalah salah satu gangguan yang sering dijumpai pada pasien anak di Puskesmas. Diketahui berdasarkan data SIKM, pasien anak yang mengunjungi Puskesmas wilayah Sleman dan Yogyakarta dengan gangguan emosi dan perilaku sejumlah 46,37% dari 1902 pasien, semenjak 2011 hingga Maret 2013. Guna memfasilitasi adanya kebutuhan instrumen skrining untuk deteksi dini gangguan tingkah laku, dikembangkan SDQ yang merupakan kuisisioner singkat mengenai atribut positif dan negatif anak dan remaja (usia 4-16 tahun) yang terdiri 25 aitem. Penelitian ini bertujuan untuk mengadaptasi SDQ-TR versi Inggris menjadi versi Indonesia, melakukan uji properti psikometri, dan uji kualitas skrining terhadap gangguan tingkah laku. Sampel terdiri dari 161 siswa Sekolah Dasar di wilayah Kabupaten Sleman dan Kota Yogyakarta, laki-laki dan perempuan, berusia 7-13 tahun. Uji reliabilitas dengan teknik *Alpha Cronbach* menghasilkan $\alpha=0,773$. Uji validitas konstruk menggunakan PAF menunjukkan SDQ-TR memiliki enam struktur faktor. Uji kualitas skrining menggunakan ROC menghasilkan *cut-off* ≥ 5 , nilai sensitivitas 0,67 dengan spesifisitas 0,68, sedangkan menggunakan LR menghasilkan LR(+)=2,09 dan LR(-)=0,49.

Kata kunci: anak, validasi klinik, gangguan tingkah laku, SDQ-TR

Puskesmas di wilayah Kabupaten Sleman sejak tahun 2004 telah menyedia-

kan fasilitas pelayanan psikologi, dan diikuti oleh Puskesmas di wilayah Kota Yogyakarta pada tahun 2010. Pelayanan psikologi dilakukan untuk mengurangi *treatment gap* atau kesenjangan antara tem-

¹ Korespondensi mengenai isi artikel ini dapat dilakukan melalui: mistety.oktaviana@gmail.com

² Atau melalui: supra8@gmail.com

pat pelayanan kesehatan mental lingkup primer (Puskesmas) dengan sekunder (Rumah Sakit). Pelayanan psikologi di Puskesmas diberikan untuk individu, kelompok, maupun komunitas. Secara umum kunjungan pasien dewasa lebih banyak dibanding pasien anak, remaja, maupun lansia. Pasien anak yang mendaftari poli psikologi meskipun jumlahnya tidak dominan, tetap perlu mendapat perhatian.

Berdasarkan data Sistem Informasi Kesehatan Mental (SIKM), yaitu sebuah sistem informasi hasil kerjasama antara *Center for Public Mental Health* (CPMH) Fakultas Psikologi UGM dengan Dinas Kesehatan Kabupaten Sleman dan Kota Yogyakarta, pada tahun 2011-2013 diketahui bahwa jumlah pasien anak (usia 1-19 tahun) yang berkunjung ke Puskesmas mencapai 1902 orang. Pasien yang mengalami gangguan tingkah laku dan emosi (golongan F90 – F98 berdasarkan PPDGJ III) cukup mendominasi yaitu mencapai 46,37% atau sebanyak 882 pasien. Pada golongan gangguan perilaku dan emosi yang menempati urutan tertinggi, diagnosis gangguan tingkah laku (F91) berada pada jumlah terbanyak kedua setelah gangguan hiperkinetik (F90).

Prevalensi gangguan tingkah laku di Amerika mengalami peningkatan pada beberapa dekade terakhir dan lebih banyak muncul di perkotaan daripada pedesaan. Pada laki-laki di bawah 18 tahun terdapat sekitar 6% - 16%, sedangkan pada perempuan terdapat 2% - 9% (*American Psychiatric Association*, 1994). Gangguan tingkah laku lebih banyak terjadi pada laki-laki dibanding perempuan (Kazdin dalam Conner & Lochman, 2010). Prevalensi gangguan tingkah laku di Indonesia khususnya di Provinsi Daerah Istimewa Yogyakarta (DIY) hingga kini belum diketahui.

Gangguan tingkah laku adalah gangguan yang ditandai dengan pola tingkah laku dissosial, agresif atau menentang, yang berulang dan menetap. Perilaku ini, dalam bentuk ekstremnya, berupa pelanggaran berat dari norma sosial yang terdapat pada anak seusia itu, dan karena itu pelanggarannya bersifat menetap dan lebih parah daripada kenakalan anak atau sikap memberontak remaja pada lazimnya. Penilaian tentang adanya gangguan tingkah laku perlu mempertimbangkan tingkat perkembangan anak. Contoh perilaku yang menjadi dasar diagnosisnya mencakup hal berikut: perkelahian atau pelecehan yang berlebihan, kekejaman terhadap hewan atau manusia, perusakan yang hebat atas barang milik orang lain, membakar, mencuri, melakukan kebohongan berulang, membolos dari sekolah dan lari dari rumah, *temper tantrum* yang berlebihan dan sering, perilaku provokatif yang menantang, dan sikap menentang yang hebat serta menetap. Masing-masing dari kategori tersebut apabila ditemukan, cukup untuk menjadi alasan bagi diagnosis gangguan tingkah laku. Namun, perbuatan dissosial yang tunggal bukan merupakan alasan yang kuat. Diagnosis ini tidak dianjurkan kecuali bila perilaku dissosial berlanjut selama enam bulan atau lebih (Departemen Kesehatan RI, 1993).

Penggolongan kelompok gangguan tingkah laku dalam PPDGJ III terbagi menjadi enam, yaitu: (1) gangguan tingkah laku yang terbatas lingkungan keluarga. (2) Gangguan tingkah laku tidak berkelompok, yang ditandai dengan tidak adanya keterpaduan yang efektif dengan kelompok sebaya yang merupakan perbedaan penting dengan gangguan tingkah laku berkelompok. (3) Gangguan tingkah laku berkelompok, terdapatnya ikatan persahabatan yang kuat dengan anak seusianya dimana seringkali terdiri atas

anak-anak yang juga terlibat dalam perbuatan kriminal atau dissosial. (4) Gangguan sikap menentang/ membangkang, pada gangguan ini tidak ada tindakan dissosial dan agresif yang lebih berat yang melanggar hukum ataupun melanggar hak asasi orang lain, (5) Gangguan tingkah laku lainnya, dan (6) gangguan tingkah laku YTT (Departemen Kesehatan RI, 1993).

Pola asuh yang negatif, kurang responsif, adanya penolakan, dan kurang adanya ikatan emosi mampu memprediksi gangguan tingkah laku pada balita (Lorber & Egeland, 2011). Anak dengan gangguan tingkah laku memiliki sejarah keluarga dengan gangguan halusinasi atau kekerasan fisik. Selain itu, anak dengan gangguan tingkah laku juga memiliki riwayat keluarga dengan masalah belajar dan konsumsi alkohol (Ezpeleta, Granero, & Domenech, 2005). Selain pengaruh pola asuh dan riwayat keluarga, gangguan tingkah laku lebih banyak muncul pada subjek yang memiliki riwayat malnutrisi pada masa kanak-kanak (Galler, Bryce, Waber, Hock, Harrison, Eaglesfield, & Fitzma, 2012). Temperamen juga menjadi aspek yang memprediksi gangguan tingkah laku (Rothbart dalam Lorber & Egeland, 2011).

Gangguan tingkah laku sering dikaitkan dengan inteligensi di bawah rata-rata. Gangguan tingkah laku juga diketahui berhubungan dengan gangguan kecemasan, gangguan *mood*, *Attention Deficiency and Hiperactivity Disorder* (ADHD), *Oppositional Defiant Disorder* (ODD), *Pervasif Developmental Disorder* (PDD), gangguan penyesuaian, dan gangguan terkait substansi (*American Psychiatric Association*, 2004; van Lier, van der Ende, Koot, Verhulst, 2007; Harada, Hayashida, Hikita, Imai, Sasayama, Masutani, Tomita, Saitoh, Washizuka, & Amano, 2009).

Kendala yang dihadapi saat melakukan asesmen

Metode yang digunakan oleh Psikolog dalam melakukan proses asesmen yaitu observasi, wawancara atau *significant others*, dan tes psikologi. Wawancara pada orang tua terkadang mengalami hambatan karena beberapa orang tua kurang kooperatif dan terlalu sibuk, sedangkan wawancara pada anak cenderung membutuhkan waktu yang cukup lama. Observasi pada anak pun kerap mengalami kesulitan karena terbatasnya fasilitas. Kurang tersedianya media bermain bagi anak membuat anak tidak betah berlama-lama dalam proses asesmen. Alat tes psikologi yang digunakan beragam, tetapi pengadministrasian alat tes cukup rumit dan membutuhkan waktu yang lama (Rahardian, Komunikasi Personal, Maret 2013). Berdasarkan alasan tersebut, dibutuhkan alat skrining yang sensitif, spesifik, dan sederhana untuk mempermudah proses asesmen.

Strengths and Difficulties Questionnaire (SDQ) adalah sebuah instrumen skrining perilaku singkat untuk anak dan remaja (3-17 tahun) yang memberikan gambaran singkat dari perilaku yang berfokus pada kekuatan dan juga kesulitan mereka (Black, Pulford, Christie, & Wheeler, 2010). Kuesioner singkat sangat berguna ketika digunakan dalam survei berskala besar dimana aitem sebaiknya terbatas untuk memastikan adanya respon yang dicari (Ullebo, Posserud, Heiervang, Gillberg, & Obel, 2011). Hal ini merupakan kelebihan SDQ karena jumlah aitemnya yang sedikit dan relatif sederhana.

Dahlan (2009) menyatakan instrumen skrining sebaiknya memiliki keunggulan relatif yang lebih dibanding metode asesmen lainnya, yaitu: lebih tidak invasif, tingkat risiko yang lebih rendah, tidak memerlukan keahlian khusus, lebih murah, waktu untuk memperoleh hasil lebih

cepat, lebih mudah diakses, lebih sederhana, tidak terlalu rumit, dan dapat mendeteksi gangguan lebih dini, tidak hanya mendeteksi gangguan pada tahap lanjut. SDQ memiliki beberapa poin keunggulan relatif tersebut, yaitu: dapat dilakukan tanpa memiliki keahlian khusus atau profesi tertentu, waktu yang digunakan untuk mengadminstrasikan dan melakukan skoring cukup singkat, mudah diakses, tidak harus dilakukan di pelayanan kesehatan, lebih sederhana dalam administrasi ataupun skoring, digunakan untuk melakukan deteksi dini sehingga permasalahan pada anak dapat diketahui sedini mungkin dan memperoleh intervensi secepat mungkin.

SDQ terdiri dari 25 aitem yang dialokasikan pada lima subskala. Keempat subskala termasuk ke dalam kelompok subskala kesulitan, yaitu subskala *emotional symptom*, subskala *conduct problem*, subskala *hyperactivity-inattention*, dan subskala *peer problem*. Sedangkan subskala yang kelima termasuk dalam kelompok subskala kekuatan, yaitu subskala *prosocial*. Masing-masing subskala SDQ terdiri dari lima aitem. Masing-masing aitem diskor dalam kriteria tiga poin yaitu 0=tidak benar, 1=agak benar, 2=sangat benar. Skor dari masing-masing subskala dapat dihitung dengan menjumlahkan skor dari masing-masing aitem yang relevan pada subskala tersebut. Skor tertinggi dari masing-masing subskala adalah 10 dan skor terendah adalah 0 (Goodman dalam Muris, Meesters, & van den Berg, 2003).

Studi di Belanda menyatakan kuesioner sebaiknya diberikan sesuai dengan bahasa ibu atau jika memungkinkan bahasa lain yang familiar untuk digunakan subjek (Richter, Sagatun, Heyerdahl, Oppedal, & Roysamb, 2011). Pada *Youth In Mind*, situs resmi SDQ,

diketahui bahwa SDQ telah diterjemahkan ke dalam lebih dari 40 bahasa dan bebas digunakan untuk kepentingan non-komersial (<http://www.sdqinfo.org/>).

SDQ telah dilakukan uji instrumen penelitian pada berbagai versi. Goodman mengawali penelitian pada 1997 dan menunjukkan hasil SDQ-*teacher reports* (TR) dan *parent reports* (PR) berguna sama baiknya dengan skala Rutter. Tahun 2008, Goodman menunjukkan SDQ-*self reports* (SR) berkorelasi dengan baik dengan SDQ-PR dan SDQ-TR. Tingkat *agreement* antara prediksi SDQ dan diagnosis klinis secara signifikan tinggi (Goodman, Renfrew, & Mullick, 2000). Kemampuan prediksi struktur lima faktor SDQ (masalah emosi, masalah tingkah laku, inatensi-hiperaktivitas, masalah teman sebaya, dan kemampuan prososial) dapat dikonfirmasi (Goodman, 2001). Uji reliabilitas terbukti memuaskan, dengan konsistensi internal (0,73), korelasi *cross informant* (0,34), dan stabilitas *test retest* setelah 4-6 bulan (0,62). Tahun 2003, Goodman, Ford, Simmons, Gatward, dan Meltzer menunjukkan bahwa SDQ memiliki sensitivitas 63,3%, spesifisitas 94,6%, nilai prediksi positif 52,7%, dan nilai prediksi negatif 96,4%. Sensitivitas SDQ akan signifikan lebih tinggi bila menggunakan multiinforman ketiga form SDQ dibandingkan informan tunggal. Sensitivitas berbeda pada tiap diagnosis. Sensitivitas $\geq 0,70$ untuk mengidentifikasi individu dengan gangguan perilaku, hiperaktif, depresi dan beberapa gangguan kecemasan, dan $\leq 0,50$ untuk mengidentifikasi individu dengan fobia spesifik, kecemasan akan perpisahan, gangguan makan dan gangguan panik. Uji efisiensi SDQ kembali dilakukan oleh Goodman, Ford, Corbin, dan Meltzer pada tahun 2004 dan menunjukkan hasil sensitivitas 84,8%, spesifisitas 80,1%, nilai duga positif 74,2% dan nilai duga negatif 88,7%.

Penelitian mengenai SDQ telah banyak dilakukan di berbagai negara, baik untuk pengujian properti psikometri maupun validasi terhadap gangguan tertentu. SDQ Belanda dapat mendeteksi dini masalah kesehatan mental pada anak dengan *Intellectual Disorder* (ID). Skor tingkat kesulitan menunjukkan anak dengan ID memiliki skor lebih tinggi dibandingkan tanpa ID (Emerson, 2005). SDQ Inggris mampu mendeteksi anak dengan HIV dengan menghasilkan skor kesulitan yang lebih tinggi daripada populasi umum (anak tanpa HIV) (Melvin, Krechevsky, Divac, Tacconelli, Miah, Waugh, Hekster, Byard, & Giannakopoulou, 2007).

SDQ Norwegia mendapatkan titik potong ≥ 18 untuk total kesulitan SDQ. Nilai konsistensi internal dengan *alpha cronbach* untuk total kesulitan adalah 0,74. Koefisien *alpha* untuk masing-masing subskala adalah 0,70 untuk masalah emosi, 0,50 untuk masalah perilaku, 0,66 untuk hiperaktivitas-inatensi, 0,53 untuk masalah teman sebaya, dan 0,70 untuk perilaku prososial. Hasil uji *Principal Component Analysis* (PCA) dan diagram *screen plot* menunjukkan adanya tujuh faktor yang menjelaskan varian sebesar 52,9% (Svedin & Priebe, 2008).

Reliabilitas SDQ-PR dan TR Australia pada semua subskala terbilang sedang (0,57) – tinggi (0,88) (Hawes & Dadds, 2004). Sensitivitas dari *probable diagnosis* menunjukkan skor hiperaktif=44%, gangguan tingkah laku=93%, dan gangguan emosi=36%. Sensitivitas untuk diagnosis yang tergolong mengalami gangguan atau *borderline* menunjukkan skor hiperaktif=93%, gangguan tingkah laku=100%, dan gangguan emosi=81% (Mathai, Anderson, & Bourne, 2004).

Skoring SDQ yang digunakan dapat berdasarkan skor pada masing-masing subskala ataupun skor total kesulitan yang

merupakan total dari empat skor subskala kesulitan, yaitu subskala masalah emosi, masalah perilaku, masalah dengan teman sebaya, dan hiperaktivitas-inatensi (Goodman, 1994). Penelitian ini selanjutnya menggunakan skor subskala masalah perilaku untuk melihat bagaimana efisiensinya dalam menskrining gangguan tingkah laku.

SDQ versi Inggris pernah diadaptasi menjadi versi Indonesia. Namun, proses adaptasi tersebut tidak disertai uji properti psikometri dan uji kualitas skrining. Oleh karenanya penelitian ini melakukan adaptasi ulang SDQ-TR versi Inggris, uji properti psikometri, dan uji kualitas skrining SDQ-TR versi Indonesia terhadap gangguan tingkah laku. Dahlan (2009) mengungkapkan salah satu syarat melakukan penelitian instrumen skrining maupun diagnostik yaitu harus terdapat baku emas (*gold standard*). Penelitian ini menggunakan hasil diagnosis psikolog berdasar PPDGJ III sebagai standar baku emas. Adapun pertanyaan yang hendak dijawab dari penelitian ini adalah bagaimana reliabilitas dan validitas konstruk SDQ-TR Indonesia?, bagaimana sensitivitas, spesifisitas, dan *Area Under Curve* (AUC) SDQ-TR Indonesia terhadap gangguan tingkah laku pada sampel komunitas?, serta bagaimana rasio kemungkinan positif, dan rasio kemungkinan negatif SDQ-TR Indonesia terhadap gangguan tingkah laku pada sampel komunitas?

Hasil penelitian ini secara teoritis dapat memperkaya khasanah ilmu psikologi khususnya dalam pengembangan instrumen skrining. Secara praktis, hasil penelitian ini dapat meningkatkan sensitivitas para guru dalam melakukan deteksi dini gangguan tingkah laku pada siswanya, sehingga para siswa dengan gangguan tingkah laku dapat lebih cepat memperoleh penanganan yang tepat.

Metode

Partisipan

Subjek untuk sampel komunitas dalam penelitian ini berjumlah 161 anak dengan kriteria berusia 7-13 tahun, berjenis kelamin laki-laki dan perempuan, pada komunitas sekolah dasar yang berada dalam wilayah Kabupaten Sleman dan Kota Yogyakarta. Metode *sampling* yang digunakan adalah *Multi Stage Cluster Random Sampling*, dengan memilih secara acak kecamatan, sekolah dasar, dan siswa yang dijadikan sampel dalam penelitian ini.

Pelaksanaan Penelitian

Tahap pertama adalah melakukan persiapan penelitian. Peneliti mempersiapkan sarana dan prasarana yang dibutuhkan untuk penelitian. Pertama adalah perijinan untuk melakukan penelitian di Puskesmas dan Sekolah Dasar (SD). Kedua adalah mempersiapkan sejumlah instrumen yang digunakan dalam penelitian ini meliputi SDQ-TR versi Inggris yang kemudian diadaptasi menjadi versi Indonesia, Pedoman penegakan diagnosis gangguan tingkah laku berdasarkan kriteria PPDGJ III, Modul pelatihan diagnosis gangguan tingkah laku berdasar PPDGJ III untuk para psikolog yang telah dilakukan *professional judgement* oleh psikolog klinis senior, dan video untuk pengujian kualitas diagnosis psikolog.

Tahap selanjutnya adalah melakukan

pengembangan instrumen penelitian. Adaptasi instrumen SDQ-TR telah dilakukan *back translation*, diawali dengan menerjemahkan SDQ-TR versi Inggris menjadi bahasa Indonesia oleh dua orang ahli bahasa. *Multiple translator* dilakukan untuk mencegah adanya keganjilan bahasa atau kecenderungan menggunakan bahasa yang disukai oleh *single translator* (Hambleton, Merenda, & Spielberger, 2005). Tiga orang psikolog senior kemudian melakukan *expert judgement* terhadap SDQ-TR yang telah diterjemahkan, dengan tujuan melakukan penilaian mengenai ketepatan terjemahan terhadap konsep psikologi. Proses selanjutnya adalah melakukan *cognitive debriefing* kepada 17 orang guru di SD yang terdapat di wilayah Kabupaten Sleman dan Kota Yogyakarta, dengan tujuan untuk mengetahui interpretasi responden terhadap aitem-aitem yang terdapat dalam instrumen (Plougman, Austin, Stefanelli, & Godwin, 2010). Setelah memperoleh hasil uji coba dan dilakukan penyesuaian, SDQ-TR di terjemahkan kembali ke dalam bahasa Inggris untuk mengkonfirmasi kesesuaian konsep dengan versi aslinya.

Persiapan selanjutnya adalah pelatihan kepada para psikolog pengambil data. Pelatih adalah seorang psikolog klinis senior bidang psikologi anak. Pelaksanaan pelatihan dijelaskan dalam Tabel 1.

Persiapan terakhir adalah pengujian kualitas diagnosis psikolog pengambil data dengan metode *inter-rater reliability*. Hasil *rating* psikolog dibandingkan de-

Tabel 1
Rincian Pelaksanaan Pelatihan Bagi Psikolog Calon Pengambil Data

Judul Kegiatan	Waktu	Tempat	Pelatih
Pelatihan Penegakan Diagnosis Gangguan Emosi dan Perilaku Pada Anak berdasarkan PPDGJ III	Senin-Selasa, 16-17 September 2013. Pukul 08.30-16.00 WIB	Ruang Merak, Bag. Ilmu Kesehatan Anak, RSUP Dr. Sardjito	Dr. Indria L. Gamayanti, M.Si., Psi.

ngan hasil *rating* dari psikolog anak senior sebagai *gold standard*. Fleiss (1975) mengkategorikan tingkat reliabilitas antar rater menjadi empat kategori yaitu: (a) $k < 0,4$ (buruk), (b) $k = 0,40-0,60$ (cukup), (c) $k = 0,60-0,75$ (memuaskan), dan (d) $k \geq 0,75$ (istimewa). Penelitian ini mensyaratkan untuk menggunakan psikolog dengan nilai reliabilitas antar rater minimal $k = 0,75$. Tabel 2 adalah hasil uji kualitas diagnostik psikolog.

Tabel 2
Hasil Uji *Inter-Rater Reliability* Psikolog Calon Pengambil Data

No	Nama Psikolog	Nilai k
1	AY	0,78
2	TR	0,78
3	RH	0,89
4	KW	0,89
5	RHJ	0,89
6	DNP	0,78
7	AR	0,89
8	EE	0,89
9	AU	0,78
10	NR	0,78
11	TP	0,78
12	BDA	0,89

Berdasarkan penjabaran tersebut diketahui bahwa semua psikolog memiliki tingkat reliabilitas antar-rater istimewa. Proses pengambilan data diawali dengan mengajukan *inform consent* kepada para psikolog, guru, dan orang tua subjek. Proses pengumpulan data dilakukan secara tersamar (*blinding*) dengan tujuan untuk menghindari bias pengukuran (Dahlan, 2002). Baik instrumen SDQ maupun hasil diagnosis psikolog dikumpulkan secara terpisah dan hanya boleh dibuka oleh peneliti.

Analisis Data

Analisis data yang dilakukan dalam penelitian ini secara garis besar terdiri dari

analisis uji properti psikometri yang terdiri dari uji reliabilitas menggunakan *Alpha Cronbach* dan uji validitas konstruk menggunakan *Principal Axis Factoring*, dan analisis uji kualitas skrining menggunakan *Receiver Operating Curve* (ROC) dan *Likelihood Ratio* (LR). Analisis uji asumsi dilakukan menggunakan *Kolmogorov Smirnov Test*.

Hasil

Setelah proses pengumpulan data selesai, peneliti melakukan seleksi dan skoring semua hasil yang terkumpul yaitu SDQ-TR dan diagnosis psikolog. SDQ-TR terkumpul sebanyak 181 subjek dan diagnosis psikolog hanya 161 subjek. Data 20 subjek tidak dapat dilakukan uji kualitas skrining karena tidak ada diagnosis psikolog sehingga hanya dapat dilakukan uji properti psikometri.

Data demografi menunjukkan rerata usia subjek adalah 9,74 tahun, berjenis kelamin perempuan sebanyak 40 subjek (24,8%) dan laki-laki sebanyak 121 subjek (75,2%). Berdasarkan tingkat pendidikan, kelas dengan jumlah subjek terbanyak adalah kelas III sebanyak 41 subjek (25,5%). Data demografi juga menunjukkan bahwa 33 dari 161 subjek terdiagnosis gangguan tingkah laku (20,5%), dimana sebanyak 29 subjek (87,9%) adalah laki-laki. Berdasarkan tingkat pendidikan, anak dengan gangguan tingkah laku paling banyak muncul pada anak kelas 3 yaitu sebesar 39,4%. Sedangkan berdasarkan usia, anak dengan gangguan tingkah laku paling banyak muncul pada anak dengan usia 9 tahun dan 10 tahun. Sedangkan hasil analisis data SDQ-TR menunjukkan bahwa nilai rerata skor subskala perilaku adalah 3,68, dengan skor maksimum 10 dan skor minimum 0.

Pada uji normalitas dengan metode *Kolmogorov Smirnov*, nilai signifikansi $<0,05$ mengindikasikan adanya penyimpangan dari distribusi normal (Field, 2009). Hasil menunjukkan $p=0,012$ ($p<0,05$), yang berarti penyebaran data pada SDQ-TR tidak terdistribusi normal.

Uji properti psikometri menganalisis dua hal, yaitu reliabilitas dan validitas konstruk. Hasil analisis reliabilitas dengan *Alpha Cronbach* menunjukkan bahwa subskala masalah perilaku SDQ-TR memiliki tingkat reliabilitas yang memuaskan ($r_{xx'}=0,773$). Hasil analisis validitas konstruk menggunakan *Principal Axis Factoring* (PAF) memiliki nilai Kaiser-Meyer-Olkin (KMO) sebesar 0,776 yang berarti semua aitem SDQ-TR layak untuk dilakukan analisis faktor. Tahap rotasi faktor dilakukan dengan metode *varimax*, dengan tujuan untuk memaksimalkan jumlah varians dalam faktor (Field, 2009). *Factor loading* dari setiap aitem dengan nilai yang paling besar berkelompok pada salah satu faktor. Pada SDQ-TR, pengelompokan ini membentuk enam faktor. Pada faktor 1 tampak bahwa butir aitem yang mengelompok adalah aitem 3, 6, 8, 13, 16, 19, dan 24 yang sesuai dengan konstruk masalah emosi. Faktor 2 berisi butir aitem 1, 4, 7, 9, 17, dan 20 yang sesuai dengan konstruk perilaku prososial. Faktor 3 berisi butir aitem 5, 12, 18 dan 22 yang sesuai dengan konstruk masalah perilaku. Faktor 4 berisi butir aitem 11, 14, dan 23 yang sesuai dengan konstruk masalah dengan teman sebaya. Faktor 5 berisi aitem-aitem 2, dan 10 yang sesuai dengan konstruk hiperaktivitas. Sedangkan faktor 6 berisi butir aitem 15, 21, dan 25 yang sesuai dengan konstruk inatensi. Total keseluruhan varian yang dijelaskan oleh keenam faktor tersebut adalah sebesar 62,45%.

Uji kualitas skrining SDQ-TR dilakukan menggunakan dua metode yaitu ROC

dan LR. Analisis ROC menghasilkan nilai *Area Under the Curve* (AUC) 73,4% (64,5% - 82,3%), yang berarti apabila SDQ-TR digunakan untuk menskrining gangguan tingkah laku pada 100 anak, maka kesimpulan yang tepat akan diperoleh pada 73 anak. Titik potong skor SDQ-TR dalam melakukan skrining gangguan tingkah laku adalah 4,5 dengan sensitivitas 0,67 dan spesifisitas 0,68. SDQ-TR akan mampu melakukan skrining dengan hasil nilai positif pada subjek yang menderita gangguan tingkah laku sebesar 67% dan mampu melakukan skrining dengan hasil negatif pada subjek yang tidak menderita gangguan tingkah laku sebesar 68%.

Warner (2004) menyatakan bahwa idealnya sebuah tes memiliki satu skor titik potong. Titik potong optimum berwujud angka bulat untuk dalam pengaplikasian di lapangan untuk kepentingan praktis. Oleh karenanya titik potong untuk SDQ-TR dibulatkan menjadi ≥ 5 . Berdasarkan analisis *Likelihood Ratio*, didapatkan nilai rasio kemungkinan positif (LR +) 2,09, dan rasio kemungkinan negatif (LR -) 0,49. Hal ini berarti; Subjek dengan gangguan tingkah laku memiliki kemungkinan 2,09 kali lebih tinggi terdeteksi positif SDQ-TR dibanding subjek tanpa gangguan tingkah laku, dan Subjek dengan gangguan tingkah laku memiliki kemungkinan 0,49 kali lebih rendah terdeteksi negatif oleh SDQ-TR dibanding subjek tanpa gangguan tingkah laku.

Diskusi

Hasil penelitian ini menunjukkan subjek laki-laki lebih banyak yang menderita gangguan tingkah laku daripada subjek perempuan. Hal ini sejalan dengan apa yang diungkapkan Kazdin dalam Conner dan Lochman (2010) bahwa gangguan tingkah laku lebih banyak terjadi pada

laki-laki dibanding perempuan. Prevalensi gangguan tingkah laku pada laki-laki di bawah 18 tahun terdapat sekitar 6% - 16%, sedangkan pada perempuan terdapat 2% - 9% (*American Psychiatric Association, 1994*).

Hasil analisis reliabilitas dengan metode *alpha cronbach* menunjukkan bahwa subskala masalah perilaku SDQ-TR memiliki tingkat reliabilitas yang memuaskan. Nilai koefisien reliabilitas adalah $rxr' = 0,773$. Wells dan Wollack (dalam Azwar, 2012) mengatakan bahwa tes yang digunakan di kelas oleh para guru hendaknya paling tidak memiliki koefisien reliabilitas $\geq 0,70$, dengan demikian reliabilitas SDQ-*Teacher Reports* ini dianggap memuaskan.

Uji reliabilitas yang dilakukan pertama kali pada instrumen SDQ terbukti memuaskan, baik dengan konsistensi internal (*Alpha Cronbach=0,73*), korelasi *cross informant* (0,34), maupun stabilitas *test retest* setelah 4-6 bulan (0,62) (Goodman, 2001). Nilai koefisien reliabilitas SDQ-TR Belanda rata-rata $\alpha = 0,80$ untuk semua subskala (Muris, Meesters, & van den Berg, 2003; Widenfelt, Goedhart, Treffers, & Goodman, 2003). SDQ-TR Jerman memiliki reliabilitas yang memuaskan $\alpha = 0,75-0,83$. SDQ-TR Cina memiliki konsistensi internal $\alpha = 0,55-0,84$ (Lai, Luk, Leung, Wong, Lau, & Ho, 2010). SDQ-TR Spanyol memiliki nilai konsistensi internal $\alpha = 0,64-0,83$ (Rodriguez-Hernandez dkk., 2012). Pemaparan berbagai penelitian dari berbagai populasi dan budaya, dengan hasil yang relatif sama, menguatkan bahwa SDQ-TR memang memiliki reliabilitas yang memuaskan.

Hasil analisis faktor dengan menggunakan *Principal Axis Factoring* (PAF) menemukan enam faktor yang membentuk konstruk SDQ-TR, yaitu faktor masalah emosi, perilaku prososial, masalah perilaku, masalah dengan teman sebaya, hiperaktivitas, dan inatensi. Goodman

(2001) melakukan pegujian terhadap properti psikometri SDQ dan hasilnya menunjukkan bahwa struktur prediksi lima faktor (*emotional, conduct, hyperactivity-inattention, peer problem and prosocial*) dapat dikonfirmasi. Faktor yang berbeda dalam SDQ-TR pada penelitian ini dibandingkan penelitian sebelumnya adalah terpisahnya faktor hiperaktivitas-inatensi sehingga membentuk dua faktor yang berbeda yaitu faktor hiperaktivitas dan faktor inatensi.

Analisis faktor pertama kali dilakukan oleh Goodman (2001) dan menunjukkan subskala masalah perilaku pada SDQ Inggris terdiri dari aitem 5, 7, 12, 18, dan 22. Penelitian pada beberapa negara mampu membuktikan aitem yang mengelompok pada subskala masalah perilaku terkonfirmasi sebagaimana model asli yaitu SDQ Inggris, sedangkan pada beberapa negara lain memberikan hasil yang sedikit berbeda. Berdasarkan hasil analisis faktor penelitian ini, subskala masalah perilaku SDQ-TR hanya terdiri dari aitem 5, 12, 18, dan 22. Aitem 7 memisahkan diri dan berkelompok dengan subskala perilaku prososial. Tidak terkonfirmasinya aitem 7 sesuai model asli dikarenakan adanya perbedaan budaya antara Indonesia dengan Inggris. Bagaimana masyarakat Indonesia menginterpretasi perilaku *obedience* yang disebutkan dalam aitem 7 berbeda dengan masyarakat Inggris. Di Indonesia, *obedience* atau kepatuhan lebih cenderung berkaitan dengan perilaku prososial daripada masalah perilaku. Mengelompoknya aitem 7 ke dalam subskala perilaku prososial ditemui dalam hasil penelitian di Jepang dan Swedia. Aitem 7 pada SDQ-PR Swedia dan SDQ Jepang mengelompok pada subskala perilaku prososial (Malmberg, Rydell, Smedje, 2003; Matsuishi, dkk., 2008). Hal ini menunjukkan bahwa pada kedua negara tersebut, persepsi masyarakat terhadap aitem 7

tidak berbeda dengan persepsi masyarakat Indonesia.

Terkait dengan fungsi SDQ-TR sebagai instrumen skrining, isu lain yang perlu dikaji adalah mengenai validitas klinik instrumen tersebut. Hasil analisis ROC dan LR terhadap subskala masalah perilaku SDQ-TR menunjukkan nilai AUC sebesar 73,4%, nilai batas pisah optimum ≥ 5 , sensitivitas sebesar 0,67, dan spesifisitas sebesar 0,68. Disamping itu, dihasilkan juga nilai LR + sebesar 2,09 dan nilai LR – sebesar 0,49. Berdasarkan hasil analisis tersebut, dapat disimpulkan bahwa SDQ-TR memiliki kualitas skrining yang memuaskan dalam melakukan skrining gangguan tingkah laku. Berdasar penelitian dari berbagai negara, *cut-off* total skor kesulitan yang didapat cukup beragam. Terjadinya perbedaan *cut-off* pada hampir setiap negara dalam mendeteksi suatu gangguan kemungkinan disebabkan oleh adanya pengaruh budaya dan perbedaan *gold standard* yang digunakan saat melakukan validasi (Becker, Steinhausen, Baldursson, Dalsgaard, Lorenzo, Ralston, Döpfner, Rothenberger, & ADORE Study Group, 2006).

Seorang peneliti dapat menggunakan dua jenis metode interpretasi dalam penelitian diagnostika, yaitu interpretasi klinik dan interpretasi statistik (Dahlan, 2009). Interpretasi klinik dapat sejalan dengan interpretasi statistik, tetapi dapat juga tidak sejalan. Dahlan lebih merekomendasikan interpretasi secara klinik dibandingkan interpretasi secara statistik. Interpretasi klinik tidak terlepas dari keberadaan instrumen lain yang mengukur hal serupa dan bagaimana kualitas skrining dari instrumen lain tersebut. Secara statistik, kualitas skrining SDQ-TR terhadap gangguan tingkah laku dengan AUC 73,4% terbilang sedang. Secara klinik, sejauh pengetahuan peneliti, belum ada instrumen lain yang

valid dalam menskrining gangguan tingkah laku di Indonesia. Hal ini menjadikan SDQ-TR secara klinik memiliki kualitas skrining yang baik atau memuaskan.

Analisis uji normalitas dilakukan untuk mengetahui sebaran data dari penelitian ini. Hasil analisis menunjukkan nilai $p=0,012$ ($p<0,05$) yang berarti sebaran data pada SDQ-TR tidak terdistribusi secara normal. Hal ini mengindikasikan sampel yang digunakan dalam penelitian tidak dapat mewakili populasi umum. Azwar (2013) mengatakan distribusi data yang tidak normal tidak selalu menghasilkan analisis yang tidak valid. Uji normalitas yang tidak terpenuhi tidak lantas memberikan hasil analisis yang tidak memuaskan. Hasil analisis uji properti psikometri dan uji kualitas skrining SDQ-TR dalam penelitian ini terbukti memuaskan. SDQ-TR dalam penelitian ini sensitif dan spesifik digunakan untuk menskrining gangguan tingkah laku. Azwar (2013) juga menyatakan bahwa selama hasil analisis tidak meragukan, uji asumsi tidak harus dilakukan. Dengan demikian dapat disimpulkan meski uji normalitas tidak terpenuhi, hasil penelitian tetap dapat diterima dan dapat diaplikasikan pada populasi umum.

Penelitian ini merupakan bagian dari sebuah penelitian besar yang meneliti mengenai validasi klinik SDQ-TR sebagai instrumen skrining kelompok gangguan perilaku dan emosi. Kelompok gangguan perilaku dan emosi yang diteliti dalam penelitian tersebut adalah gangguan hiperkinetik, gangguan tingkah laku, dan gangguan kecemasan pada anak. Namun, terdapat perbedaan jumlah subjek yang dilibatkan dalam penelitian ini dengan penelitian besar. Penelitian ini menggunakan jumlah subjek sebanyak 161 anak, sedangkan penelitian besar menggunakan 272 subjek. Perbedaan jumlah subjek yang

dilibatkan dalam kedua penelitian disebabkan karena perbedaan persiapan penelitian. Oleh karenanya *unpublished report* mengenai validasi klinik SDQ-TR sebagai instrumen skrining gangguan tingkah laku juga dengan 272 subjek akan dijelaskan berikut ini.

Uji normalitas subskala masalah perilaku SDQ-TR menghasilkan nilai $p=0,001$ ($p<0,05$). Hal ini berarti sebaran data subskala masalah perilaku pada 272 subjek tidak terdistribusi secara normal. Uji reliabilitas memberikan hasil yang memuaskan ($r_{xx'}=0,764$). Analisis faktor menunjukkan SDQ-TR dengan 272 subjek memiliki enam faktor yaitu faktor masalah perilaku, faktor hiperaktivitas, faktor inatensi, faktor masalah emosi, faktor masalah dengan teman sebaya, dan faktor prososial. Total varian yang dijelaskan dari enam faktor tersebut adalah 56,421%. Hasil uji kualitas skrining juga terbilang memuaskan. Nilai AUC sebesar 0,725 dengan sensitivitas 0,75 dan spesifisitas 0,60. LR (+)=1,875 dan LR (-)=0,416. Hasil tersebut menunjukkan tidak nampak perbedaan yang signifikan dengan penelitian ini, baik properti psikometri maupun kualitas skrining sama-sama memberikan hasil yang memuaskan. Namun, walaupun dengan jumlah subjek lebih besar, uji normalitas subskala masalah perilaku SDQ-TR pada penelitian besar tetap saja tidak terpenuhi. Hal ini kemudian memunculkan pertanyaan dimana letak kesalahan hingga data yang digunakan untuk penelitian besar tidak mampu menggambarkan populasi umum.

Teknik pengambilan sampel menjadi bahasan akan keterbatasan dalam penelitian ini. Jika menggunakan teknik pengambilan sampel secara acak, kemungkinan besar sampel yang terpilih akan mampu mewakili kondisi pada populasi umum (Kerlinger, 1992). Kemungkinan yang terjadi adalah adanya kesalahan pada proses

pengambilan sampel. Subjek pada masing-masing kelas dipilih oleh guru. Meskipun guru telah diberi instruksi untuk memilih secara acak, tetapi ada kecenderungan guru memilih subjek sesuai dengan kriteria yang diinginkan. Kerlinger (1992) mengatakan andai sampel sudah dikenal atau diketahui, mungkin secara tidak sengaja atau tidak disadari akan cenderung memilih sampel yang sesuai dengan kebutuhan (memiliki kriteria dari variabel yang hendak diteliti).

Asumsi adanya bias dalam pemilihan subjek ini semakin kuat dengan pernyataan dari beberapa psikolog pengambil data. Semua subjek yang dilakukan asesmen tampak memiliki permasalahan psikologi, khususnya dilihat dari keluhan orang tua dan guru (Ekawati, Komunikasi Personal, Oktober 2013). Psikolog lain mengatakan berdasar hasil wawancara dengan guru, guru sudah melabel sebagian besar subjek sebagai anak hiperaktif (Nugraheni, Komunikasi Personal, November 2013). Suatu sampel yang ditarik secara acak tidaklah mengandung bias, dalam arti bahwa tidak satupun anggota dalam populasi mempunyai peluang yang lebih besar dari anggota lain untuk terpilih menjadi sampel (Kerlinger, 1992). Metode *random* dalam melakukan seleksi pemilihan sampel tidak memperbolehkan bekerjanya bias dan campur tangan dari faktor-faktor sistematis lainnya. Prosedur ini bersifat objektif, bebas dari kecondongan serta bias yang dimiliki (Kerlinger, 1992).

Keterbatasan penelitian

Penelitian ini memiliki keterbatasan terkait metode yang digunakan, yaitu teknik *cluster random sampling* kurang objektif. Pemilihan subjek secara acak oleh guru pada masing-masing kelas memiliki kecenderungan terjadi bias. Hal ini menyebabkan uji normalitas tidak terpenuhi,

karena teknik *random sampling* yang tidak benar-benar objektif sehingga tidak benar-benar merepresentasikan populasi umum.

Kesimpulan

Berdasarkan hasil penelitian, dapat disimpulkan bahwa properti psikometri SDQ-TR tergolong memuaskan. Struktur konstruk pada SDQ-TR sedikit berbeda dengan versi Inggris, dimana faktor hiperaktivitas-inatensi pada penelitian ini terpisah menjadi faktor hiperaktivitas dan faktor inatensi. SDQ-TR memiliki kualitas skrining yang cukup memuaskan. SDQ-TR sensitif dan spesifik jika digunakan sebagai instrumen skrining gangguan tingkah laku. Titik potong SDQ-TR yang digunakan dalam menskrining gangguan tingkah laku oleh guru dengan menggunakan subskala masalah perilaku adalah ≥ 5 .

Saran

Berdasarkan hasil dan diskusi penelitian ini, maka disarankan untuk penelitian selanjutnya agar lebih berhati-hati dalam melakukan pengambilan sampel. Pilihlah teknik *sampling* sesuai dengan kebutuhan, dan lakukan sesuai metode yang disarankan. Jika hendak menggunakan *random sampling*, sebaiknya lakukan dengan objektif tanpa membedakan kemungkinan peluang antara satu individu dengan individu lainnya. Saran lain adalah pilihlah informan yang benar-benar memahami kondisi perilaku dan emosi subjek khususnya selama enam bulan terakhir, sehingga dapat memberikan informasi seakurat mungkin.

Kepustakaan

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*. Wash-

ington, DC: American Psychiatric Association.

Antony, M. M., & Barlow, D. H. (2002). *Handbook of Assessment and Treatment Planning for Psychological Disorders*. United States of America: The Guilford Press.

Azwar, S. (2008). *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar.

Azwar, S. (2012). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.

Azwar, S. (2013). Asumsi-asumsi dalam Inferensi Statistika. diunduh dari <http://www.psikologi.ugm.ac.id>.

Becker, A., Steinhausen, H., Baldursson, G., Dalsgaard, S., Lorenzo, M. J., Ralston, S. J., Döpfner, M., Rothenberger, A., & ADORE Study Group. (2006). Psychopathological screening of children with ADHD: Strengths and Difficulties Questionnaire in a pan-European study. *European Child and Adolescent Psychiatry, 15*, 56-62. doi: 10.1007/s00787-006-1008-7.

Black, S., Pulford, J., Christie, G., & Wheeler, A. (2010). Differences in New Zealand School Student's reported Strength and Difficulties. *New Zealand Journal of Psychology, 39*(9).

Conner, B., & Lochman, J. (2010). Comorbid Conduct Disorder and Substance Use Disorders. *Clinical Psychology: Science and Practice, 337* – 349.

Dahlan, M. (2009). *Penelitian Diagnostik*. Jakarta: Penerbit Salemba Medika.

Departemen Kesehatan R.I. Direktorat Jendral Pelayanan Medik. (1993). *Pedoman Penggolongan dan Diagnosis Gangguan Jiwa di Indonesia*. Jakarta: Departemen Kesehatan R.I. Direktorat Jendral Pelayanan Medik.

Emerson, E. (2005). Use of the Strengths and Difficulties Questionnaire to As-

- sess the Mental Health Needs of Children and Adolescents with Intellectual Disabilities. *Journal of Intellectual & Developmental Disability*, 30(1), 14–23. doi: 10.1080/13668250500033169.
- Ezpeleta, L., Granero, R., & Doménech, J. M. (2005). Differential Contextual Factors of Comorbid Conduct and Depressive Disorders in Spanish Children. *European Child Adolescent Psychiatry*, 14, 282–291. doi: 10.1007/s00787-005-0476-5.
- Field, A. (2009). *Discovering Statistic Using SPSS, 3rd Edition*. Dubai: SAGE Publications.
- Fleiss, J.L. (1975). Measuring Agreement between Two Judges on the presence or Absence of a Trait. *Biometrics*, 31(3), 651-659.
- Galler, J. R., Bryce, C. P., Waber, D. P., Hock, R. S., Harrison, R., Eaglesfield, G. D., & Fitzmaurice, G. (2012). Infant Malnutrition Predicts Conduct Problems in Adolescents. *Nutritional Neuroscience*, 15(4), 186 – 192.
- Goodman, R., & Scott, S. (1999). Comparing the Strength and Difficulties Questionnaire (SDQ) and Child Behavior Checklist (CBCL): Is small beautiful ?. *Journal of Abnormal child Psychology*, 27, 17-24.
- Goodman, R., Renfrew, D., & Mullick, M. (2000). Predicting type of psychiatric disorder from Strength and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European Child and Adolescent Psychiatry*, 9, 124-134.
- Goodman, R. (2001). Psychometric properties of the Strength and Difficulties Questionnaire (SDQ). *Journal of American Academy of Child and Adolescent Psychiatry*, 40, 1337-1345.
- Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2003). Using the Strength and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community samples. *British Journal of Psychiatry*, 177, 534-539. doi: 10.1080/0954026021000046128.
- Goodman, R., Ford, T., Corbin, T., & Meltzer, H. (2004). Using the Strengths and Difficulties Questionnaire (SDQ) multi-informant algorithm to screen looked-after children for psychiatric disorders. *European Child and Adolescent Psychiatry*, 13, 25-31. doi: 10.1007/s00787-004-2005-3.
- Hambleton, R., Merenda, P., Spielberger, C. (2005). *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. New Jersey: Mahwah.
- Harada, Y., Hayashida, A., Hikita, S., Imai, J., Sasayama, D., Masutani, S., Tomita, T., Saitoh, K., Washizuka, S., & Amano, N. (2009). Impact of Behavioral/Developmental Disorders Comorbid with Conduct Disorder. *Psychiatry and Clinical Neurosciences*, 63, 762–768. doi: 10.1111/j.1440-1819.2009.02029.x.
- Hawes, D. J., & Dadds, M. R. (2004). Australian Data and Psychometric Properties of the Strengths and Difficulties Questionnaire. *Australian and New Zealand Journal of Psychiatry*, 38, 644-651.
- Lai, K. Y. C., Luk, E. S. L., Leung, P. W. L., Wong, A. S. Y., Law, L., & Ho, K. (2010). Validation of the Chinese Version of the Strengths and Difficulties Questionnaire in Hong Kong. *Social Psychiatry Epidemiology*, 45, 1179–1186. doi: 10.1007/s00127-009-0152-z.
- Lorber, M., & Egeland, B. (2011). Parenting and Infant Difficulty: Testing a Mutual

- Exacerbation Hypothesis to Predict Early Onset Conduct Problems. *Child Development*, 82 (6), 2006–2020. doi: 10.1111/j.1467-8624.2011.01652.x.
- Malmberg, M., Rydell, A. M., & Smedje, H. (2003). Validity of the Swedish Version of the Strengths and Difficulties Questionnaire (SDQ). *Nordern Journal Psychiatry*, (57), 357 – 363. doi: 10.1080/08039480310002697.
- Mathai, J., Anderson, P., & Bourne, A. (2004). Comparing Psychiatric Diagnoses Generated by the Strengths and Difficulties Questionnaire with Diagnoses Made by Clinicians. *Australian and New Zealand Journal of Psychiatry*, 38, 639-643.
- Melvin, D., Krechevsky, D., Divac, A., Tacconelli, E., Miah, J., Waugh, S., Hekster, B., Byard, K., & Giannakopoulou, C. (2007). Parental Reports of Emotional and Behavioural Difficulties on the SDQ for School-Age Children with Vertically Acquired HIV Infection Living in London. *Psychology, Health & Medicine*, 12(1), 40 – 47. doi: 10.1080/13548500600731708.
- Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ): further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child and Adolescent Psychiatry*, 12, 1-8. doi: 10.1007/s00787-003-0298-2.
- Richter, J., Sagatun, A., Heyerdahl, E., Oppedal, B., & Roysamb, E. (2011). The Strengths and Difficulties Questionnaire (SDQ) – Self-Report. An analysis of Its Structure in a Multiethnic Urban Adolescent Sample. *Journal of Child Psychology and Psychiatry*, 52(9), 1002–1011. doi: 10.1111/j.1469-7610.2011.02372.x.
- Svedin, C. G., & Priebe, G. (2008). The Strengths and Difficulties Questionnaire as a Screening Instrument in a Community Sample of High School Seniors in Sweden. *Nord Journal of Psychiatry*, 62, 225-232. doi: 10.1007/s00787-011-0198-9.
- Ullebo, A. K., Posserud, M. B., Heiervang, E., Gillberg, C., & Obel, C. (2011). Screening for the Attention Deficit Hyperactivity Disorder phenotype using the Strength and Difficulties Questionnaire (SDQ). *European Child and Adolescent Psychiatry*, 20, 451-458.
- Van Lier, P. A. C., van der Ende, J., Koot, H. M., & Verhulst, F. C. (2007). Which Better Predicts Conduct Problems? The Relationship of Trajectories of Conduct Problems with ODD and ADHD Symptoms from Childhood into Adolescence. *Journal of Child Psychology and Psychiatry*, 48(6), 601–608. doi:10.1111/j.1469-7610.2006.01724.x.
- Warner, J. (2004). Clinician's Guide to Evaluating Diagnostic and Screening Tests in Psychiatry. *Journal of Continuing Professional Development*, 10, 446-445.
- Widenfelt, B. M., Goedhart, A. W., Treffers, P. D. A., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child and Adolescent Psychiatry*, 12, 281-289. doi: 10.1007/s00787-003-0341-3.