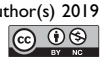
Original Article

DEVELOPMENTAL CHILD WELFARE

Developmental Child Welfare
2019, Vol. 1(2) 177–196



DOI: 10.1177/2516103219829756
journals.sagepub.com/home/dcw

Mental health screening for children in care using the Strengths and Difficulties Questionnaire and the Brief Assessment Checklists: Guidance from three national studies

Michael Tarren-Sweeney[1,2], Anouk Goemans[3], Anna Sophie Hahne[4], and Matthew Gieve[4]

Abstract

Although children residing in statutory out-of-home care and those adopted from care are more likely than not to have mental health difficulties requiring clinical intervention or support, their difficulties often remain undetected. Children's agencies have a duty of care to identify those child clients who require therapeutic and other support services, without regard to the availability of such services. The present article proposes a first-stage mental health screening procedure (calibrated for high sensitivity) for children and adolescents (ages 4–17) in alternative care, which children's agencies can implement without clinical oversight using the Strengths and Difficulties Questionnaire (SDQ) and Brief Assessment Checklists (BAC). The screening procedure was derived from analyses of BAC, SDQ, and "proxy SDQ" scores obtained in three national studies of children and adolescents residing in alternative care (Australia, the Netherlands, and England). The SDQ and BAC demonstrated moderate to high screening accuracy across a range of clinical case criteria—the SDQ being slightly better at predicting general mental health problems and the BAC

[1] University of Canterbury, New Zealand
[2] University of Newcastle, Australia
[3] Leiden University, the Netherlands
[4] The Tavistock Institute, UK

Submitted: 16 March 2018; Accepted: 12 January 2019

Corresponding author:
Anouk Goemans, Institute of Education and Child Studies, Leiden University, Wassenaarseweg 52, Leiden 2333 AK, the Netherlands.
Email: a.goemans@fsw.leidenuniv.nl

slightly better at predicting attachment- and trauma-related problems. Accurate first-stage screening is achieved using either the SDQ or the BAC alone, with recommended cut points of 10 (i.e., positive screen is 10 or higher) for the SDQ and 7 for the BAC. Greater accuracy is gained from using the SDQ and BAC in parallel, with positive screens defined by an SDQ score of 11 or higher *or* a BAC score of 8 or higher. Agencies and post-adoption support services should refer positive screens for comprehensive mental health assessment by clinical services.

Keywords
Brief Assessment Checklist, children in care, mental health screening, screening protocol, Strengths and Difficulties Questionnaire

Introduction

Children in statutory alternative care are more likely than not to have psychological difficulties of sufficient scale or severity to require clinical intervention and guidance, or further monitoring. Between a third and half of children in foster care have clinical-level mental health difficulties (around four times the rate among children at large), and another 15–25% have difficulties approaching clinical significance (Tarren-Sweeney, 2008). Two U.K. studies have measured comparable rates among children adopted from care (DeJong, Hodges, & Malik, 2016; Selwyn, Wijedasa, & Meakings, 2014), while rates are a little lower for children in kinship care (Winokur, Holtan, & Batchelder, 2018) and considerably higher for those in residential care (Hukkanen, Sourander, Bergroth, & Piha, 1999; Li, Chng, & Chu, 2017; Meltzer, Corbin, Gatward, Goodman, & Ford, 2003).

While governments are cognizant of the extent of these children's difficulties, many children (and their carers) who require clinical interventions and guidance are not seen by mental health services. This is despite governments holding parental responsibility obligations for children in statutory care. Among a nationally representative sample of U.S. children in foster care, less than half of those with clinical-level difficulties received a mental health service in the year leading up to the survey (National Survey of Child and Adolescent Well-Being, 2003). Whereas 61% of a survey sample of South Australian children in foster care ($N = 326$) had clinical-level difficulties, and 53% were identified by their foster carers as needing clinical intervention, only 27% had received a service in the previous 6 months (Sawyer, Carbone, Searle, & Robinson, 2007). A survey of 182 Scottish children in foster care found that while their overall mental health difficulties predicted ongoing involvement with social workers, they did not predict children's access to public Child and Adolescent Mental Health Services (CAMHS) (Minnis, Everett, Pelosi, Dunn, & Knapp, 2006).

Various factors account for this shortfall in service delivery, including CAMHS not having sufficient capacity to respond to the high numbers of children in care requiring their service, and children being excluded from services by way of discriminatory intake criteria (such as a requirement that children be in stable placements) (House of Commons Education Committee, 2016). However, an important reason that children in care do not receive the clinical interventions and guidance that they and their carers need is their difficulties are not detected. Sizable proportions of children residing in care are "undetected clinical cases"—that is, they have clinical-level mental health difficulties that have not been identified through screening or formal assessment (Romanelli et al., 2009). For example, among a large Australian population sample of children in out-of-home

care ($N = 347$), between 22% and 25% had not been referred for clinical assessment and intervention despite having clinical-level mental health difficulties (Tarren-Sweeney, 2010a). This highlights a lack of universal mental health screening or assessment in many jurisdictions. One exception is the U.K., which mandates annual mental health screening of children in statutory care using carer-report Strengths and Difficulties Questionnaires (SDQs) (Department for Education & Department of Health, 2015). Given their exceptionally high prevalence of clinical-level difficulties, children in alternative care would be best served by *universal*, *comprehensive assessment* by clinicians with specialized knowledge and skills (Leslie et al., 2003; Tarren-Sweeney, 2010b). In most western jurisdictions, this could not be implemented without building more specialized capacity within their clinical workforce. Until this can be realized, the next best alternative is to institute systematic, universal mental health screening, coupled with comprehensive assessment of positive screens.

What is the meaning and purpose of mental health "screening" for children in care?

The word "screening" is often loosely applied in child and adolescent mental health practice. It has at least three different connotations. It can refer to screening in the proper or classic use of the word—that is, employing a brief assessment tool or procedure that has good sensitivity for detecting clinical cases, and which is typically followed by further assessment of positive screens. "Screening" is also used colloquially as a synonym for "brief," "short," or "quick"—as in, a screening assessment is one that is brief. Thirdly, the word is commonly used to describe initial intake assessments in CAMHS. These assessments typically follow a prescribed format; are applied to all new referrals; they tend to be questionnaire-driven; and they tend not to require high-level clinical reasoning. They are brief, routine initial assessment protocols that are designed to act as a gateway to the service.

In practice, this means that children's agencies and clinical services don't have a common understanding of the meaning and purpose of screening. While children's agencies have a moral and legal duty of care to monitor and ensure the psychological well-being of *all* children residing in their statutory care, many governments allocate responsibilities for children's *social care* and their *clinical assessment and therapeutic services* to different agencies. Several Australian state child welfare departments maintain their own psychological services, such that they can deliver both social work and psychological services to children that are in their statutory care. But in most western jurisdictions, public provision of clinical assessment and therapeutic services for children in statutory care is the responsibility of health ministries (i.e., provided through CAMHS, or primary health care, or contracted services). Yet, initial "screening" or "intake assessments" carried out by clinical services such as CAMHS are designed to address the service's intake criteria, which in turn are governed to a large extent by the service's intake capacity. In this context, the word "screening" is misapplied. It is not screening in the classic sense of the word. Rather, it is a form of *triage*.

The argument for employing classical screening with this population. The classic approach to screening is to employ it as the first of a two-stage or multistage assessment strategy. The aim of the first (i.e., screening) stage is to locate as many cases as possible. The aim of the second and subsequent assessment stages is to detect (i.e., rule out) as many noncases as possible. In the first stage, sensitivity is thus more critical than specificity, except where there are seriously adverse social or emotional consequences for false positive screening. This is particularly true for children in

alternative care. Firstly, the implications for children growing in care with undetected mental health problems are more serious than they are for children who do not have such problems being subjected to a second-stage assessment. Second, while false positive screens are likely to be correctly identified as noncases in second-stage assessment, there is no such remedy for false negative screens—they stay undetected! And, because as little as a third of children in care have normative mental health (i.e., neither elevated nor clinical-level difficulties), loss of specificity within this population translates as fewer false positive screens and higher positive predictive value (the proportion of positive screens who are true cases) than occurs with mental health screening of children and young people in the general community. For these reasons, we would argue that a first-stage screening procedure for application with children in care should have at least 85% sensitivity, with specificity above 50% being acceptable.

Designing a mental health screening procedure for children's agencies

There are several considerations for designing a first-stage mental health screening procedure that can be implemented by children's agencies. The first is that it can be performed safely and accurately without clinical oversight, using tools that have an easy-to-calculate score and a simple screening decision. The second is that screening decisions are derived from caregivers' reports of children's difficulties. Relying on other informants poses both ethical and practical challenges. A third consideration is that in addition to common mental health difficulties, the screening procedure should detect clinical-level attachment- and trauma-related difficulties. Finally, screening decisions (based on score cut points) should be highly sensitive (>85%) in identifying clinical cases, while retaining acceptable (>50%) specificity.

Screening measures that are highly sensitive while retaining acceptable specificity have high overall screening accuracy. Screening accuracy is conventionally reported as the "area under the Receiver Operating Characteristics (ROC) curve" (AUC). The ROC curve plots the true positive (sensitivity) rate against the false positive (1-specificity) rate for each score cut point. A "perfect" screening instrument has an AUC of 1.0, while a measure that performs no better than chance has an AUC of 0.5. Screening procedures with AUC ≤ 0.70 are not sufficiently accurate for use in the field, while procedures with AUC ≥ 0.80 have good to excellent screening properties.

The Strengths and Difficulties Questionnaire. The SDQ is the only caregiver-report screening measure of *common* child and adolescent mental health problems for which screening accuracy has been tested with children in alternative care, and which can be used by children's agencies *without* clinical oversight (Tarren-Sweeney, 2018). The SDQ is a 25-item measure of children's general mental health difficulties and prosocial behavior (Goodman, 2001) that includes parent-report forms for ages 3–4 and 4–16. It consists of 20 symptom items that constitute a "total difficulties scale" and four clinical subscales (emotional symptoms, conduct problems, hyperactivity inattention, and peer problems) and five prosocial behavior items. A longer version of the SDQ also includes an impact scale that measures subjective distress and the functional impact that the child's symptoms exerts over his or her daily life. A multi-informant screening algorithm was devised for the SDQ, which combines symptom and impact scores from parent-report, teacher-report, and youth self-report measures (Goodman, Renfrew, & Mullick, 2000). As stated previously, we do not believe that a multi-informant screening procedure can be feasibly implemented by children's agencies.

The parent-report SDQ total difficulties score has reasonable screening accuracy when used with children at large. ROC analyses in eight screening studies yielded AUC values between 0.64 and 0.91 (i.e., ranging from poor to excellent), with a mean weighted AUC of 0.87 (Stone, Otten, Engels, Vermulst, & Janssens, 2010). To date, ROC analyses of SDQ screening accuracy with *children in care* have been reported from four studies. In a survey of 1,028 5- to 17-year-old English children in statutory care, the multi-informant algorithm (combined parent- and teacher-reported SDQ problem and impact scores) had poor sensitivity when screening decisions were based solely on caregiver report or teacher report (Goodman, Ford, Corbin, & Meltzer, 2004). Among a sample of 223 Norwegian children in foster care, the accuracy with which the SDQ total difficulties score screened for children with one or more mental disorders (as identified by the Development and Well-Being Assessment) was AUC = 0.83, with sensitivity of 0.81 and specificity of 0.75 using the 14+ borderline range cut point (Lehmann, Heiervang, Havik, & Havik, 2014). Furthermore, in this study, the multi-informant algorithm performed poorly in comparison with simpler screening procedures. A recent population mental health survey of Dutch 4- to 17-year-olds in foster care ($N = 219$) found the SDQ to be moderately accurate in predicting children's receipt of clinical interventions, as well as high foster carer stress (AUCs ranged from 0.72 to 0.83) (Goemans, Tarren-Sweeney, van Geel, & Vedder, 2018). Finally, a recent English study of 144 clinic-referred children and adolescents in out-of-home care found that carer-report ($N = 97$, AUC = 0.61) and self-report ($N = 41$, AUC = 0.57) total difficulties scores poorly predicted children's assessed treatment needs, while teacher-report ($N = 41$, AUC = 0.74) scores performed a little better (Wright, Wellsted, Gratton, Besser, & Midgley, 2019). This study identified 10+ as the optimal carer-report SDQ cut point, where the goal is accurate identification of cases (high sensitivity).

The Brief Assessment Checklists. The Brief Assessment Checklist for Children (BAC-C, ages 4–11) and the Brief Assessment Checklist for Adolescents (BAC-A, ages 12–17) are caregiver-report, 20-item measures designed to screen for clinically meaningful attachment- and trauma-related mental health difficulties (Tarren-Sweeney, 2013a). They were derived from the Assessment Checklist for Children (ACC, 120 items) (Tarren-Sweeney, 2007) and Assessment Checklist for Adolescents (ACA, 105 items) (Tarren-Sweeney, 2013b), respectively. The ACC and ACA are caregiver-report psychiatric rating scales that were designed to measure a broad range of mental health difficulties observed among children and young people in care that are not adequately measured by standard rating instruments, such as the Child Behavior Checklist (CBCL), SDQ, and Conners scales. These consist of a number of interpersonal, attachment-related difficulties; insecure relating; social, behavioral, and emotional dysregulation; trauma-related anxiety and dissociation; abnormal responses to pain; overeating and related food maintenance behaviors; sexual behavior problems; self-injury; and suicidal behaviors and discourse.

In the New South Wales (NSW) Children in Care Study (CICS)—the surveys in which the measures were developed—the BAC-C and BAC-A were highly accurate in predicting attachment- and trauma-related difficulties, as indicated by ACC and ACA scores in the clinical and "elevated or clinical" ranges (AUC = 0.96–0.99). They were also moderately accurate in predicting common mental health difficulties, as indicated by CBCL scores in the clinical and "borderline or clinical" ranges (AUC = 0.89–0.94) (Tarren-Sweeney, 2013a).

Rationale for the further analysis of SDQ and BAC screening performance

The aforementioned Dutch survey was the first to compare the screening properties of the SDQ and BAC. The survey found that the SDQ and BAC were comparably accurate at predicting which children were receiving clinical intervention or support services and which of their foster carers were experiencing high levels of stress (Goemans et al., 2018). There are several other results of the Dutch survey that are relevant to the present discussion but have not been previously published. In that study, the SDQ 14+ borderline range cut point had insufficient sensitivity for a first-stage screening procedure, while the recommended BAC 5+ screening cut point had insufficient specificity. Also, sizable proportions of children had simultaneously low SDQ and high BAC scores and vice versa. Furthermore, 86% of the child sample were positive screens on the BAC-C (5+ cut point), as compared to only 40% on the SDQ (14+ cut point), which equates to a 54% screening decision concordance rate. The equivalent percentages for adolescents were 78%, 37%, and 58%. Together these findings translate as low interscale agreement on screening decisions, adding weight to the possibility that overall screening accuracy is increased by using both measures, in comparison to using the SDQ or BAC as sole screening measures.

These findings highlight a need (1) to identify optimal BAC and SDQ screening cut points for children in alternative care, for use in a classic, first-stage screening procedure, (2) to investigate whether meaningful additional screening accuracy is gained by using *both* the BAC and SDQ in a screening procedure, as compared to using *either* of the measures, and (3) to validate any proposed screening procedures across multiple populations of children in care.

The present article reports new analyses of BAC, SDQ, and CBCL scores obtained in three national studies: the Australian CICS (the surveys in which the BAC-C and BAC-A were originally developed), the aforementioned Dutch study, and a recent survey of clinic-referred English children adopted from care (an evaluation of the new Adoption Support Fund [ASF]).

Method

The present article compares mental health screening properties of the SDQ, BAC, and a "SDQ proxy" score (generated from a set of CBCL items approximating the SDQ total difficulties scale) in relation to various clinical case reference criteria, across three national studies (Australia, the Netherlands, and England) of children residing in alternative care. Given the BAC has separate forms for school-aged preadolescent children (ages 4–11, BAC-C) and for adolescents (ages 12–17, BAC-A), all of the present analyses were done separately for these two age ranges. The surveys were approved by the Human Ethics Committees of Newcastle University and Canterbury University (Australian study); Leiden University (Dutch study); and the U.K. Department for Education (English study). In all three studies, caregivers provided informed consent to participate. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

The three studies

The respective study methodologies of the Australian, Dutch, and English studies are reported elsewhere in greater detail (Goemans et al., 2018; King, Gieve, Iacopini, Hahne, & Stradling, 2017; Tarren-Sweeney, 2007; Tarren-Sweeney, 2013a). While general child protection practice in

the three study locations are likely to be quite similar, there are some important differences in their care systems. Neither the Netherlands nor NSW had formalized adoption of children from statutory care at the time the relevant surveys were carried out (although NSW has since legislated for adoption from care). Another key difference is that NSW had closed down all residential care at the time the child survey was carried out (a small number of residential places was reinstated by the time of the adolescent survey), such that almost all children and young people were residing in family-based care.

Australian sample. The CICS was a prospective, statewide study of the psychosocial development and risk exposures of children residing in long-term foster and kinship care in NSW, Australia. The child (conducted from 1999 to 2003) and adolescent (conducted 2009 and 2011) survey samples were broadly representative of their respective sampling frames. Children's mental health was measured from the CBCL in both surveys and from the ACC in the child survey and the ACA in the adolescent surveys. The ACC, ACA, and BAC measures were developed from these survey data, and sample BAC-C and BAC-A scores have been previously published (Tarren-Sweeney, 2013a). The child sample was 347 4- to 11-year-olds, with even gender ratio, and a mean age of 7.8 years, though most (>60%) were 6- to 8-year olds. Their mean time in care was 4.3 years, and their mean (median) number of placements was 3.1 (2). Rates of children residing in nonkinship foster care and kinship care were 86% and 14%, respectively. The adolescent sample was 230 12- to 17-year-olds, 85 of whom had been in the child survey sample and 145 being newly recruited to the study. Their mean (median) age was 15.3 (15.0) years, their mean time in care was 10.4. years, and boys (54%) were slightly more represented than girls (46%). Rates of young people residing in foster and kinship care were 85% and 15%, respectively.

Dutch sample. The Dutch sample was derived from the third wave (October 2015) of a longitudinal study on the development of children in full-time nonkinship foster care and kinship care in the Netherlands. Foster parents from seven foster care agencies (25% of Dutch agencies) were invited to complete a questionnaire about the foster placement. The questionnaire consisted of questions regarding the foster child, the foster family, and the foster placement. The Dutch versions of the SDQ (Van Widenfelt, Goedhart, Treffers, & Goodman, 2003) and the BAC (Goemans et al., 2018; King et al., 2017) were used to measure foster children's mental health. The child sample (4- to 11-year-old) consisted of 118 children, with a mean age of 7.80 ($SD = 2.15$) and slightly more boys (56.4%) than girls (43.6%). The majority of the children were living in non-kinship foster families (71.0%). The children resided on average 48 months in their current foster placement ($SD = 31.7$) and experienced on average 1.17 previous placements ($SD = 1.7$). The adolescent sample contained 101 adolescents (12- to 17-year-olds), their mean age being 14.25 ($SD = 1.7$). Again, there were slightly more boys (52.3%) than girls (47.7%). Almost two thirds resided in kinship foster care (64.3%). The mean time in the current placement for adolescents was almost 67 months ($SD = 51.3$), with a placement history of 1.04 previous placements ($SD = 1.0$).

English sample. The longitudinal survey of adoptive parents in England was part of the evaluation of the ASF commissioned by the Department for Education (King et al., 2017). The ASF is an initiative introduced in 2015 aimed at enabling adopted children and their families to access therapeutic support. The two-wave longitudinal survey was conducted between 2015 and 2017 as a self-completion postal questionnaire. The study used a volunteer sampling approach of all adoptive parents that had successfully applied to the ASF for therapeutic support and that had

given consent to take part in the evaluation. The first questionnaire was completed soon after approval of the application to the ASF and a follow-up questionnaire was completed 7 months later. The first survey was open for 1 year from May 2015 to May 2016. Children's behavior, development, and well-being was measured with the BAC and the SDQ. The first survey sample consisted of 499 children between 4 years and 11 years and 271 adolescents between 12 years and 17 years. The mean age of the children was 7.8 ($SD = 2.17$) and the gender ratio was approximately even. Most adolescents (69%) were between 12 years and 14 years with a mean age of 13.8 ($SD = 1.40$). The sample had an even gender ratio. For the purposes of the present analyses, all of the English sample are considered to be clinical cases, given they had all been assessed to be valid recipients of therapeutic services.

Approximating the SDQ total difficulties score from CBCL items

Although the SDQ and CBCL were developed separately and without reference to each other, they have a high degree of overlap in their symptom coverage. Parent-report SDQ total difficulties scores for ages 4–12 have an average correlation of $r = 0.76$ (range 0.70–0.87) with CBCL total problems scores, across a large number of studies of children in the general community (Stone et al., 2010), while a correlation of $r = 0.81$ was reported in a study of 292 Belgian children in out-of-home care (Janssens & Deboutte, 2010). The two measures also employ a 3-point response scale using very similar wording (CBCL scores 0-1-2 = "not true," "somewhat or sometimes true," "very true or often true"; SDQ 0-1-2 scores = "not true," "somewhat true," "certainly true"). The CBCL's very large item pool provides considerable scope to approximate the SDQ total difficulties scale. Table 1 lists the SDQ items that constitute the total difficulties scale, and the corresponding CBCL items that were selected to constitute a "SDQ proxy" scale.

BAC and SDQ/SDQ-proxy mean scores, internal consistency (Cronbach's α), and interscale score correlations are listed in Table 2. The SDQ-proxy performed similarly to the SDQ in terms of score distributions, internal consistency, and correlations with the BAC, suggesting it is a reasonable approximation of the SDQ. As expected, the English sample mean scores on both measures were around double that of the Dutch and Australian mean scores, due to it being a clinic-referred sample.

Estimating screening accuracy

The Australian and Dutch surveys included various dichotomous outcome data that are suggestive of either the need for clinical services or support for caregivers, including foster-carer-reports of whether or not children currently use clinical services (Australian and Dutch), scores falling above or below clinical range cut-points on long-form assessment measures (Australian), and high caregiver stress (Dutch). The latter was defined as scores ranging from 62 to 150 on the abbreviated version of the Nijmeegse Ouderlijke Stress Index (De Brock, Vermulst, Gerris, & Abidin, 1992). The screening accuracy of BAC and SDQ/SDQ-proxy scores for predicting these various clinical reference criteria in the Australian and Dutch samples was estimated through ROC analyses. There was no need to define additional clinical reference criteria for the English sample, as they were all children receiving clinical and related support services—that is, they are *all* defined as clinical cases. Given the English sample consisted solely of clinical cases, we were able to calculate the sensitivity of BAC and SDQ cut points, but not their specificity.

Table 1. CBCL items selected to calculate proxy SDQ total difficulties score.

SDQ Emotional Problems subscale	
SDQ item	Equivalent CBCL item
3. Often complains of headaches, stomachaches, or sickness	56b. Headaches; 56c. Nausea, feels sick; 56f. Stomachaches[a]
8. Many worries, often seems worried	112. Worries
13. Often unhappy, downhearted, or tearful	103. Unhappy, sad or depressed (1991 profile); 103. Sad (2001 profile)
16. Nervous or clingy in new situations, easily loses confidence	45. Nervous, high-strung, or tense
24. Many fears, easily scared	50. Too fearful or anxious
SDQ Conduct Problems subscale	
SDQ item	Equivalent CBCL item
5. Often has temper tantrums or hot tempers	95. Temper tantrums or hot temper
7. Generally obedient, usually does what adults request (*reverse*)	22. Disobedient at home; 23. Disobedient at school[a]
12. Often fights with other children or bullies them	37. Gets in many fights
18. Often lies or cheats	43. Lying or cheating
22. Steals from home, school, or elsewhere	81. Steals at home; 82. Steals outside the home[a]
SDQ hyperactivity subscale	
SDQ item	Equivalent CBCL item
2. Restless, overactive, cannot stay still for long	10. Can't sit still, restless or hyperactive
10. Constantly fidgeting or squirming	10. Can't sit still, restless or hyperactive
15. Easily distracted, concentration wanders	8. Can't concentrate, can't pay attention for long; 78. Inattentive, easily distracted[b]
21. Thinks things out before acting (*reverse*)	41. Impulsive or acts without thinking
25. Sees tasks through to the end, good attention span (*reverse*)	8. Can't concentrate, can't pay attention for long; 4. Fails to finish things he/she starts[c]
SDQ Peer Problems subscale	
SDQ item	Equivalent CBCL item
6. Rather solitary, tends to play alone	111. Withdrawn, doesn't get involved with others
11. Has at least one good friend (*reverse*)	12. Complains of loneliness
14. Generally liked by other children (*reverse*)	48. Not liked by other kids
19. Picked on or bullied by other children	38. Gets teased a lot
23. Gets on better with adults than with other children	25. Doesn't get along with other kids

Note. SDQ = Strengths and Difficulties Questionnaire; CBCL = Child Behavior Checklist.
[a]Highest scoring item.
[b]1991 profile = Item 8; 2001 profile = highest scoring of Item 8 and Item 78.
[c]1991 profile = Item 8; 2001 profile = Item 4.

Identifying optimal cut points

Sensitivity and specificity rates of BAC cut points 5+ to 8+ and AUC values are listed in Table 3. The BAC was highly accurate in predicting CBCL and ACC/ACA clinical range scores (AUC = 0.94–0.99) and moderately accurate in predicting treatment status and high carer stress (AUC = 0.72–0.87). The latter results are not unexpected, given that many children in foster care with

Table 2. BAC and SDQ/SDQ-proxy mean scores, internal consistency, and interscale correlation.

	Children (4–11 years)			Adolescents (12–17 years)		
	Netherlands $N = 118$	England $N = 493$	Australia $N = 347$	Netherlands $N = 101$	England $N = 257$	Australia $N = 230$
Mean BAC score (*SD*)	12.1 (8.2)	21.2 (7.7)	10.0 (8.0)	11.50 (7.8)	22.9 (6.3)	8.5 (7.6)
BAC internal consistency	$\alpha = .89$	$\alpha = .84$	$\alpha = .89$	$\alpha = .84$	$\alpha = .78$	$\alpha = .87$
Mean SDQ/SDQ-proxy score (*SD*)	12.9 (7.0)	22.8 (6.8)	12.9 (8.4)	11.7 (6.9)	23.8 (5.8)	10.9 (8.0)
SDQ/SDQ-proxy internal consistency	$\alpha = .85$	$\alpha = .82$	$\alpha = .89$	$\alpha = .84$	$\alpha = .73$	$\alpha = .90$
BAC–SDQ/BAC–SDQ-proxy correlation	$r = .83$	$r = .67$	$r = .76$	$r = .80$	$r = .54$	$r = .84$

Note. BAC = Brief Assessment Checklist; SDQ = Strengths and Difficulties Questionnaire.

clinical-level difficulties (and their carers) are not receiving clinical interventions or professional guidance (National Survey of Child and Adolescent Well-Being, 2003), and research indicating that foster carer stress is partly determined by child welfare systemic pressures (Murray, Tarren-Sweeney, & France, 2011). These data confirm that the presently recommended BAC cut point (5+) is set too low, resulting in unacceptably low specificity (ranging from 22% to 67%). Furthermore, the table shows that higher cut points offer better trade-off between sensitivity and specificity, while remaining highly sensitive.

Sensitivity and specificity rates of SDQ/SDQ-proxy cut points 9+ to 14+ and AUC values are listed in Table 4. The SDQ/SDQ-proxy performed very similarly to the BAC across the three studies. It was also highly accurate in predicting CBCL and ACC/ACA clinical range scores (AUC = 0.90–0.97) and was moderately accurate in predicting present treatment status and high carer stress (AUC = 0.72–0.83). These data confirm that the SDQ 14+ cut point is too high—it favors specificity over sensitivity and incurs unacceptably low sensitivity for a first-stage screening procedure.

For each measure, we compared trade-off between sensitivity and specificity across the range of cut points across the three studies. We sought to identify cut points that best met the following specifications:

1. highly sensitive prediction of CBCL and ACC/ACA clinical range scores;
2. greater sensitivity than specificity in predicting treatment status and caregiver stress;
3. retains acceptable specificity for this population;
4. If all of the above specifications can be met, we also strove to select common cut-point scores for the child and adolescent age range, with a view to simplifying the screening procedure.

For the BAC, the best options were the 6+ and 7+ cut points. While the 8+ cut point performed well for adolescents, it was notably less sensitive than the 7+ cut point for children. The 7+ cut point had higher specificity than the 6+ cut point with little loss of sensitivity. We thus propose 7+ as the optimal BAC cut point for use across the two age ranges. It identified 94% of children and adolescents with clinical range scores on either the CBCL or ACC/ACA, with a manageable specificity (for this population) of 80%. The 7+ cut point also correctly identified 97% and 99% of English children and adolescents as clinical cases.

Table 3. Sensitivity and specificity of BAC-C and BAC-A cut points in three national studies.

		BAC-C, ages 4–11					BAC-A, ages 12–17				
Study	Case criterion	Case prevalence (%)	AUC [95% CI]	Screening cut point	Sensitivity (%)	Specificity (%)	Case prevalence (%)	AUC [95% CI]	Screening cut point	Sensitivity (%)	Specificity (%)
Dutch	Intervention[a]	51	.72	≥5	98.3	27.6	42.6	.76	≥5	90.7	31.0
			[.63, .81]	≥6	90.0	34.5		[.66, .86]	≥6	88.4	39.7
			N = 118	≥7	85.0	44.8		*N* = 101	≥7	86.0	44.8
				≥8	76.7	58.6			≥8	86.0	51.7
	High carer stress[b]	41	.79	≥5	95.8	22.1	40	.87	≥5	97.5	33.3
			[.71, .88]	≥6	93.8	33.8		[.80, .95]	≥6	95.0	41.7
				≥7	91.7	45.6			≥7	95.0	48.3
				≥8	83.3	58.8			≥8	95.0	55.0
NSW	Diagnosis or intervention[c]	69	.75	≥5	78.8	58.9	67	.78	≥5	73.9	64.9
			[.70, .81]	≥6	73.8	63.6		[.71, .84]	≥6	71.2	70.1
			N = 347	≥7	72.1	72.0		*N* = 230	≥7	62.8	70.1
				≥8	65.8	74.8			≥8	58.8	80.5
	ACC/ACA clinical score[d]	45	.99	≥5	100.0	59.4	37	.99	≥5	100.0	61.6
			[.98, 1.00]	≥6	100.0	68.2		[.98, 1.00]	≥6	100.0	67.1
				≥7	100.0	75.0			≥7	98.8	75.3
				≥8	100.0	84.4			≥8	98.8	84.9
	CBCL clinical score[e]	46	.92	≥5	96.9	58.6	38	.94	≥5	96.6	60.8
			[.89, .95]	≥6	93.8	65.1		[.91, .97]	≥6	95.4	65.7
				≥7	93.2	71.5			≥7	94.3	74.1
				≥8	90.7	79.0			≥8	90.8	81.8
	ACC/ACA or CBCL clinical score[f]	52	.96	≥5	97.2	65.3	43	.97	≥5	97.0	66.9
			[.94, .98]	≥6	94.4	72.5		[.95, .99]	≥6	96.0	72.3
				≥7	93.9	79.6			≥7	94.0	80.8
				≥8	91.7	88.0			≥8	91.0	89.2
English	N/A (all cases)	100	N/A	≥5	99.0	X	100	N/A	≥5	99.2	X
			N = 494	≥6	98.4	X		*N* = 258	≥6	99.2	X
				≥7	97.4	X			≥7	99.2	X
				≥8	96.4	X			≥8	99.2	X

Note. BAC-C = Brief Assessment Checklist for Children; BAC-A = Brief Assessment Checklist for Adolescents; CBCL = Child Behavior Checklist; NSW = New South Wales; ACC = Assessment Checklist for Children; ACA = Assessment Checklist for Adolescents.

[a]Child and/or carer receiving therapeutic intervention or behavioral support.

[b]Abbreviated Nijmeegse Ouderlijke Stress Index, based on the Parenting Stress Index.

[c]Carer-reported diagnosis, or receiving intervention (including carer intervention), or seeking intervention.

[d]ACC (child) or ACA (adolescent) total clinical score in clinical range.

[e]CBCL total problem score in clinical range.

[f]ACC (child) or ACA (adolescent) total clinical score *and/or* CBCL total problem score, in clinical range.

Table 4. Sensitivity and specificity of SDQ and "CBCL SDQ-proxy" cut points in three national studies.

		Children, ages 4–11					Adolescents, ages 12–17				
Study	Case criterion	Case prevalence (%)	AUC [95% CI]	Screening cut point	Sensitivity (%)	Specificity (%)	Case prevalence (%)	AUC [95% CI]	Screening cut point	Sensitivity (%)	Specificity (%)
Dutch SDQ											
	Intervention[a]	51	.72	≥9	83.3	53.4	42.6	.77	≥9	83.7	56.9
			[.63, .81]	≥10	80.0	55.2		[.67, .86]	≥10	83.7	62.1
			$N = 118$	≥11	73.3	58.6		$N = 101$	≥11	79.1	67.2
				≥12	70.0	63.8			≥12	79.1	74.1
				≥13	65.0	65.5			≥13	72.1	77.6
				≥14	58.3	79.3			≥14	58.1	79.3
	High carer stress[b]	41	.79	≥9	87.5	48.5	40	.83	≥9	90.0	58.3
			[.70, .88]	≥10	85.4	51.5		[.75, .92]	≥10	90.0	63.3
				≥11	83.3	58.8			≥11	82.5	66.7
				≥12	81.3	64.7			≥12	82.5	73.3
				≥13	77.1	67.6			≥13	80.0	80.0
				≥14	66.7	77.9			≥14	67.5	83.3
NSW "CBCL SDQ-proxy"											
	Diagnosis or intervention[c]	69	.80	≥9	80.4	66.4	67	.79	≥9	69.9	75.3
			[.75, .85]	≥10	77.5	71.0		[.73, .85]	≥10	65.4	80.5
			$N = 347$	≥11	72.9	73.8		$N = 230$	≥11	60.1	81.8
				≥12	70.0	79.4			≥12	56.2	84.4
				≥13	64.2	82.2			≥13	49.7	87.0
				≥14	56.3	86.9			≥14	48.4	87.0
	ACC/ACA clinical score[d]	45	.90	≥9	96.1	58.3	37	.93	≥9	98.9	72.0
			[.87, .93]	≥10	94.2	63.0		[.90, .96]	≥10	96.6	78.3
				≥11	92.3	68.8			≥11	96.6	84.6
				≥12	86.5	71.4			≥12	92.0	87.4
				≥13	82.6	76.6			≥13	86.2	92.3
				≥14	76.1	83.9			≥14	86.2	93.7
	CBCL clinical score[e]	46	.96	≥9	99.4	62.9	38	.97	≥9	92.9	82.1
			[.94, .98]	≥10	98.8	68.8		[.95, .99]	≥10	88.5	87.2
				≥11	96.9	74.7			≥11	85.8	92.3
				≥12	93.2	79.0			≥12	80.5	94.0
				≥13	89.4	84.4			≥13	72.6	96.6
				≥14	82.6	91.4			≥14	71.7	97.4

(continued)

Table 4. (continued)

Study	Case criterion	Children, ages 4–11 Case prevalence (%)	AUC [95% CI]	Screening cut point	Sensitivity (%)	Specificity (%)	Adolescents, ages 12–17 Case prevalence (%)	AUC [95% CI]	Screening cut point	Sensitivity (%)	Specificity (%)
	ACC/ACA or CBCL clinical score[f]	52	.94 [.92, .96]	≥9	96.1	66.5	43	.97 [.95, .99]	≥9	97.0	77.7
				≥10	94.4	71.9			≥10	95.0	84.6
				≥11	92.8	78.4			≥11	92.0	89.2
				≥12	87.8	81.4			≥12	87.0	91.5
				≥13	83.3	86.2			≥13	80.0	95.4
				≥14	76.1	92.8			≥14	80.0	96.9
English SDQ											
	N/A (all cases)	100	N/A $N = 498$	≥9	98.4	X	100	N/A $N = 270$	≥9	99.6	X
				≥10	97.0	X			≥10	98.5	X
				≥11	95.4	X			≥11	97.4	X
				≥12	93.8	X			≥12	96.7	X
				≥13	91.4	X			≥13	96.3	X
				≥14	89.8	X			≥14	94.1	X

Note. SDQ = Strengths and Difficulties Questionnaire; CBCL = Child Behavior Checklist; NSW = New South Wales; ACC = Assessment Checklist for Children; ACA = Assessment Checklist for Adolescents.

[a]Child and/or carer receiving therapeutic intervention or behavioral support.

[b]Abbreviated Nijmeegse Ouderlijke Stress Index (NOSI-K), based on the Parenting Stress Index.

[c]Carer-reported diagnosis, or receiving intervention (including carer intervention), or seeking intervention.

[d]ACC (child) or ACA (adolescent) total clinical score in clinical range.

[e]CBCL total problem score in clinical range.

[f]ACC (child) or ACA (adolescent) total clinical score *and/or* CBCL total problem score, in clinical range.

Table 5. Positive screening rates for the SDQ10+ and BAC7+ cut points.

		Child			Adolescent		
		Dutch (%)	NSW (%)	English (%)	Dutch (%)	NSW (%)	English (%)
BAC	≥7	70	58	97	68	52	99
SDQ	≥10	63	63	97	57	50	99

Note. BAC = Brief Assessment Checklist; SDQ = Strengths and Difficulties Questionnaire; NSW = New South Wales.

For the SDQ/SDQ-proxy, the best options were the 10+ and 11+ cut points. These two cut points performed very similarly, and both are acceptable choices. However, *we identified the 10+ cut point as the best solution* for use across the two age ranges. It identified 95% of children and adolescents with clinical range scores on either the CBCL or ACC/ACA, with acceptable specificity of 72% for children and 85% for adolescents. The 10+ cut point also correctly identified 97% and 99% of English children and adolescents as clinical cases. The positive screening rates for the proposed SDQ10+ and BAC7+ cut points are listed in Table 5. The recent study by Wright, Wellsted, Gratton, Besser, and Midgley (2019) separately identified 10+ as the optimal carer-report SDQ cut point, where the goal is accurate identification of cases (high sensitivity).

Is one screening measure better than the other? Are two measures better than one?

The present analyses suggest the SDQ and BAC are similarly accurate mental health screeners for children in foster care. When using the new cut points, they demonstrate comparable sensitivity and specificity across three different national samples of children in alternative care. However, we also need to consider to what extent the measures identify the same children. Table 6 lists the rates of agreement between SDQ/SDQ-proxy and BAC screening decisions in the three samples.

Screening agreement is accentuated in the English sample because the children were all clinical cases. For the other samples, BAC and SDQ screening decisions had only 80–84% agreement. While screening disagreement is partly a consequence of measurement error, the SDQ and BAC may each be more accurate (than the other) at detecting particular types of psychopathology. In the Australian study, the BAC was slightly more accurate at identifying children who have clinically significant attachment- and trauma-related difficulties, as measured by the ACC and ACA total scores, while the SDQ proxy was slightly more accurate at identifying more common internalizing, externalizing, and attentional difficulties, as encapsulated within the CBCL total problems score (Tarren-Sweeney, 2013a). This suggests that overall screening accuracy might be improved by employing *both* screening measures within a single screening procedure. To investigate this possibility, we estimated the screening accuracy of two methods that employ both screening measures, which we refer to as *combination* and *parallel* screening methods. The first method combines the BAC and SDQ total difficulties scale items into a single, 40-item screening scale. The second method employs both measures in parallel, such that a positive screen is defined as a positive score *on either* measure.

Table 7 compares the sensitivity and specificity of recommended screening measure cut points for the three screening methods: (1) BAC or SDQ used as sole measure, (2) combined BAC-SDQ score, and (3) BAC and SDQ used in parallel. The SDQ-BAC *combined* score was indeed found to be more accurate than either the SDQ or BAC. In the ROC analyses, AUCs for predicting CBCL

Table 6. Screening agreement using the SDQ10+ and BAC7+ cut points.

Dutch children		BAC-C		Dutch adolescents		BAC-A	
		Negative screen	Positive screen			Negative screen	Positive screen
SDQ	Negative screen	24.6%	12.7%	SDQ	Negative screen	27.7%	14.9%
	Positive screen	5.1%	57.6%		Positive screen	4.0%	53.5%
English children		BAC-C		English adolescents		BAC-A	
		Negative screen	Positive screen			Negative screen	Positive screen
SDQ	Negative screen	1.6%	1.4%	SDQ	Negative screen	0.4%	1.2%
	Positive screen	1.0%	95.9%		Positive screen	0.4%	98.1%
Australian children		BAC-C		Australian adolescents		BAC-A	
		Negative screen	Positive screen			Negative screen	Positive screen
SDQ proxy	Negative screen	30.3%	11.2%	SDQ proxy	Negative screen	40.9%	7.4%
	Positive screen	7.2%	51.3%		Positive screen	9.1%	42.6%

Note. BAC-C = Brief Assessment Checklist for Children; BAC-A = Brief Assessment Checklist for Adolescents; SDQ = Strengths and Difficulties Questionnaire.

and/or ACC/ACA clinical range scores ranged from 0.97 to 0.99, while AUCs for predicting present treatment status and high carer stress ranged from 0.73 to 0.87 (see Table 7).

Employing two screening measures in parallel has the effect of increasing sensitivity at the cost of specificity. Since our newly proposed SDQ and BAC screening cut points already favor sensitivity over specificity, it was apparent that a parallel screening method would require different cut points to avoid further loss of specificity. We compared sensitivity and specificity rates for various combinations of SDQ and BAC cut points and identified 8+ and 11+ as the optimal BAC and SDQ cut points, respectively, when used in a parallel screening procedure. As revealed in Table 7, the highest screening accuracy is attained from employing the two measures in parallel, using cut points that are one point higher than the single measure cut points.

Study limitations and further research

The present analyses have two major limitations that should be addressed in further research. The first is that neither the Australian or Dutch surveys included a gold standard measure of clinical mental health status. Ordinarily, a standard child clinical assessment would represent the gold standard, but with this population, the validity of clinical assessments varies somewhat depending on the level of specialist expertise that the clinician has for assessing attachment- and trauma-related psychopathology. While the English sample's clinical status was known, a clinic-referred sample provides limited opportunity for estimating screening accuracy (as there are no noncases). The second major limitation was including a proxy estimation of SDQ scores for the Australian samples. While we have attempted to show that the measure we derived from CBCL items provides a good approximation of the SDQ total difficulties score, it is nevertheless a suboptimal

Table 7. Sensitivity and specificity of recommended screening measure cut points using three screening methods.[a]

		Dutch sample						NSW sample												English sample
		Intervention			High carer stress			Intervention			ACC/ACA clinical range			CBCL clinical range			ACC/ACA or CBCL clinical range			
	Cut point	AUC	Sens. (%)	Spec. (%)	AUC	Sens. (%)	Spec. (%)	AUC	Sens. (%)	Spec. (%)	AUC	Sens. (%)	Spec. (%)	AUC	Sens. (%)	Spec. (%)	AUC	Sens. (%)	Spec. (%)	Sens. (%)
Children																				
BAC-C	7+	.72	85	45	.79	92	46	.75	72	72	.99	100	75	.92	93	72	.96	94	80	97
SDQ	10+	.72	80	55	.79	85	52	.80	78	71	.90	94	63	.96	99	69	.94	94	72	97
Combined BAC-C and SDQ score	17+	.73	83	53	.81	85	50	.80	76	76	.97	99	71	.97	98	73	.98	97	80	98
BAC-C *or* SDQ parallel screening	BAC 8+ / SDQ 11+	N/A	85	52	N/A	92	52	N/A	79	66	N/A	100	63	N/A	99	65	N/A	99	72	98
Adolescents																				
BAC-A	7+	.76	86	45	.87	95	48	.78	63	70	.99	99	75	.94	94	74	.97	94	81	99
SDQ	10+	.77	84	62	.83	90	63	.79	65	81	.93	97	78	.97	89	87	.97	95	85	99
Combined BAC-C and SDQ score	17+	.78	84	60	.87	90	62	.79	63	77	.98	96	77	.97	97	79	.99	95	85	99
BAC-C *or* SDQ parallel screening	BAC 8+ / SDQ 11+	N/A	88	48	N/A	98	52	N/A	69	75	N/A	99	71	N/A	100	73	N/A	99	80	100

Note: BAC-C = Brief Assessment Checklist for Children; BAC-A = Brief Assessment Checklist for Adolescents; SDQ = Strengths and Difficulties Questionnaire; CBCL = Child Behavior Checklist; ACC = Assessment Checklist for Children; ACA = Assessment Checklist for Adolescents.

[a]1. BAC or SDQ used as sole measure; 2. combined BAC-SDQ score; 3. BAC and SDQ used in parallel.

solution. While we are confident that the present analyses have increased our understanding of the relative screening properties of the BAC and SDQ when applied to children in care, they clearly need to be replicated in further population studies that include both of the screening measures, as well as comprehensive clinical assessments.

Screening procedure for children's agencies for ages 4–17

The present analyses of SDQ, BAC, and CBCL scores in three national studies yielded sufficient information about BAC and SDQ screening properties, for us to propose a suitable first-stage mental health screening procedure (for ages 4–17) for use by children's agencies without clinical oversight. The SDQ (25 items) and BAC (20 items) are each typically completed by caregivers in less than 10 min. The BAC-C and BAC-A questionnaires can be downloaded at www.childpsych.org.uk and the SDQ questionnaire and instructions for scoring the total difficulties score can be downloaded at www.sdqinfo.com.

Recommended procedure—SDQ and *BAC administered together*

The recommended procedure (see Figure 1) is to have caregivers complete the SDQ and BAC measures simultaneously and to refer children for a comprehensive mental health assessment if *either* their SDQ total difficulties score is *11* or *higher or* their BAC score is *8 or higher*. The present analyses indicate that this procedure is highly accurate at identifying children in alternative care with clinical-level mental health difficulties (including attachment- and trauma-related difficulties), with very few such children remaining undetected. This procedure also achieves very high sensitivity (true positive detection rate) while retaining acceptable specificity.

Alternative procedure—SDQ or *BAC used as sole screening measure*

Alternatively, where children's agencies are unable or unwilling to employ our recommended parallel screening procedure, and only use a single measure, *acceptable* screening accuracy is achieved when caregivers complete either the SDQ or BAC measures (see Figure 2). If the SDQ is used, agencies should refer children for a comprehensive mental health assessment if their SDQ total difficulties score is *10 or higher*. If the BAC is used, agencies should refer children for a comprehensive mental health assessment if their BAC-C or BAC-A score is *7 or higher*.

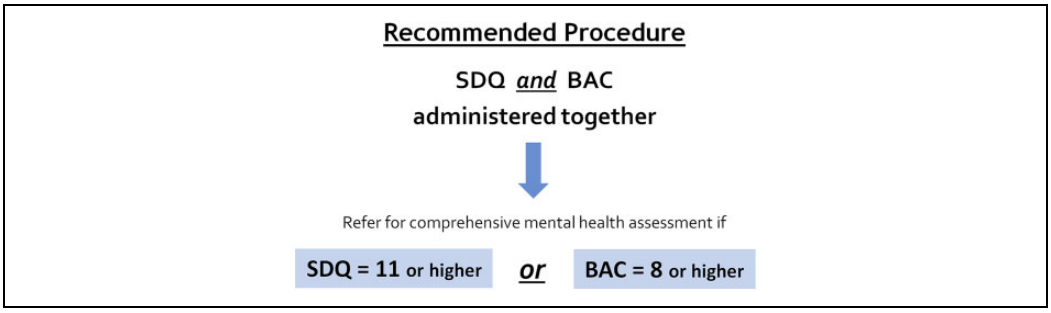


Figure 1. Recommended procedure—Strengths and Difficulties Questionnaire and Brief Assessment Checklist administered together.

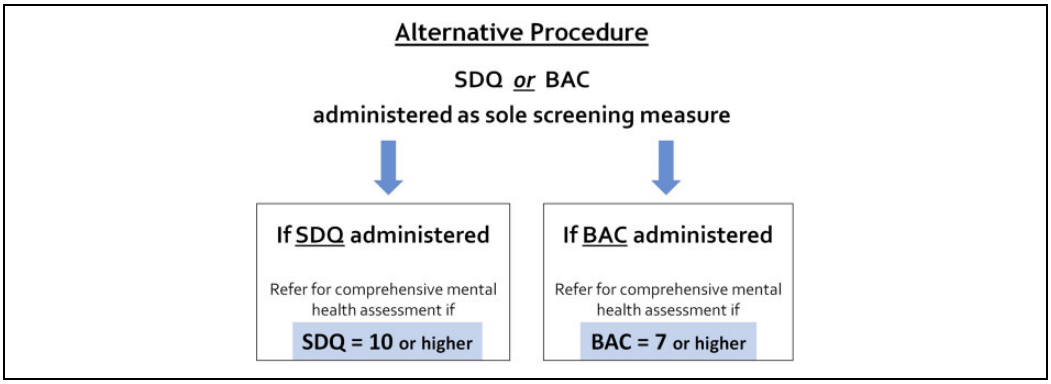


Figure 2. Alternative procedure—Strengths and Difficulties Questionnaire or Brief Assessment Checklist used as sole screening measure.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The English study was part of the evaluation of the Adoption Support Fund, commissioned and funded by the Department for Education. The Dutch study was partly funded by the Leiden University Fund/Van Steeden. The Australian study was funded and facilitated by the New South Wales Department of Family and Community Services.

References

De Brock, A., Vermulst, A., Gerris, J., & Abidin, R. (1992). *NOSI-Nijmeegse Ouderlijke Stress Index, Handleiding experimentele versie* [NOSI-Nijmegen Parenting Stress Index, Manual experimental version]. Lisse, the Netherlands: Swets en Zeitlinger.

DeJong, M., Hodges, J., & Malik, O. (2016). Children after adoption: Exploring their psychological needs. *Clinical Child Psychology & Psychiatry*, *21*, 536–550.

Department for Education & Department of Health. (2015). *Promoting the health and well-being of looked-after children: Statutory guidance for local authorities, clinical commissioning groups and NHS England*. London, England: HM Government.

Goemans, A., Tarren-Sweeney, M., van Geel, M., & Vedder, P. (2018). Psychosocial screening and monitoring for children in foster care: Psychometric properties of the Brief Assessment Checklists in a Dutch population study. *Clinical Child Psychology & Psychiatry*, *23*, 1–16.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*, 1337–1345.

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.

Goodman, R., Renfrew, D., & Mullick, M. (2000). Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European Child & Adolescent Psychiatry*, *9*, 129–134.

House of Commons Education Committee. (2016). *Mental health and well-being of looked-after children: Fourth Report of Session 2015–16*. London, England: UK Parliament.

Hukkanen, R., Sourander, A., Bergroth, L., & Piha, J. (1999). Psychosocial factors and adequacy of services for children in children's homes. *European Child and Adolescent Psychiatry*, *8*, 268–275.

Janssens, A., & Deboutte, D. (2010). Psychopathology among children and adolescents in child welfare: A comparison across different types of placement in Flanders, Belgium. *Journal of Epidemiology & Community Health*, *64*, 353–359.

King, S., Gieve, M., Iacopini, G., Hahne, A., & Stradling, H. (2017). *The evaluation of the Adoption Support Fund*. London: Department for Education.

Lehmann, S., Heiervang, E., Havik, T., & Havik, O. (2014). Screening foster children for mental disorders: Properties of the Strengths and Difficulties Questionnaire. *PLoS One*, *9*, e102134. doi:10.1371/journal.pone.0102134

Leslie, L., Hurlburt, M., Landsverk, J., Rolls, J., Wood, P., & Kelleher, K. (2003). Comprehensive assessments for children entering foster care: A national perspective. *Pediatrics*, *112*, 134–142.

Li, D., Chng, G., & Chu, C. (2017). Comparing long-term placement outcomes of residential and family foster care: A meta-analysis. *Trauma, Violence and Abuse*. doi:10.1177/1524838017726427

Meltzer, H., Corbin, T., Gatward, R., Goodman, R., & Ford, T. (2003). *The mental health of young people looked after by local authorities in England*. London, England: Office for National Statistics, The Stationery Office.

Minnis, H., Everett, K., Pelosi, A., Dunn, J., & Knapp, M. (2006). Children in foster care: Mental health, service use and costs. *European Child and Adolescent Psychiatry*, *15*, 63–70.

Murray, L., Tarren-Sweeney, M., & France, K. (2011). Foster carer perceptions of support and training in the context of high burden of care. *Child and Family Social Work*, *16*, 149–158.

National Survey of Child and Adolescent Well-Being. (2003). *One year in foster care: Wave 1 data analysis report*. Retrieved August 2008, from http://www.acf.hhs.gov/programs/opre/abuse_neglect/nscaw/reports/nscaw_oyfc/oyfc_report.pdf

Romanelli, L., Landsverk, J., Levitt, J., Leslie, L. K., Hurley, M., Bellonci, C. . . . Pecora, P. J. (2009). Best practices for mental health in child welfare: Screening, assessment, and treatment guidelines. *Child Welfare*, *88*, 163–188.

Sawyer, M., Carbone, J. A., Searle, A. K., & Robinson, P. (2007). The mental health and well-being of children and adolescents in home-based foster care. *Medical Journal of Australia*, *186*, 181–184.

Selwyn, J., Wijedasa, D., & Meakings, S. (2014). *Beyond the adoption order: Challenges, interventions and adoption disruption*. London, England: Department for Education. Retrieved from https://www.gov.uk/government/publications/beyond-the-adoption-order-challenges-intervention-disruption

Stone, L., Otten, R., Engels, R., Vermulst, A., & Janssens, J. (2010). Psychometric properties of the parent and teacher versions of the Strengths and Difficulties Questionnaire for 4- to 12-year olds: A review. *Clinical Child and Family Psychology Review*, *13*, 254–274.

Tarren-Sweeney, M. (2007). The Assessment Checklist for Children - ACC: A behavioral rating scale for children in foster, kinship and residential care. *Children & Youth Services Review*, *29*, 672–691.

Tarren-Sweeney, M. (2008). The mental health of children in out-of-home care. *Current Opinion in Psychiatry*, *21*, 345–349.

Tarren-Sweeney, M. (2010a). Concordance of mental health impairment and service utilisation among children in care. *Clinical Child Psychology and Psychiatry*, *15*, 481–495.

Tarren-Sweeney, M. (2010b). It's time to re-think mental health services for children in care, and those adopted from care. *Clinical Child Psychology and Psychiatry*, *15*, 613–626.

Tarren-Sweeney, M. (2013a). The Brief Assessment Checklists (BAC-C, BAC-A): Mental health screening measures for school-aged children and adolescents in foster, kinship, residential and adoptive care. *Children and Youth Services Review*, *35*, 771–779.

Tarren-Sweeney, M. (2013b). The Assessment Checklist for Adolescents – ACA: A scale for measuring the mental health of young people in foster, kinship, residential and adoptive care. *Children and Youth Services Review*, *35*, 384–393.

Tarren-Sweeney, M. (2018). *Mental health screening and monitoring for children in care: A short guide for children's agencies and post-adoption services*. London, England: Routledge.

Van Widenfelt, B., Goedhart, A., Treffers, P., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child & Adolescent Psychiatry*, *12*, 281–289.

Winokur, M., Holtan, A., & Batchelder, K. (2018). Systematic review of kinship care effects on safety, permanency, and well-being outcomes. *Research on Social Work Practice*, *28*, 19–32.

Wright, H., Wellsted, D., Gratton, J., Besser, S., & Midgley, N. (2019). Use of the Strengths & Difficulties Questionnaire to identify treatment needs in looked after children referred to CAMHS. *Developmental Child Welfare*. doi:10.1177/2516103218817555