Astrid Kiil, Jacob Nielsen Arendt og Michael Rosholm

Measuring Child Well-being

An overview of potential measurement instruments

TRYGFONDENS
BØRNEFORSKNINGSCENTER

Det Nationale Institut
for Kommuners og Regioners
Analyse og Forskning

Measuring Child Well-being. An overview of potential measurement instruments

The publication can be downloaded from www.kora.dk

Publisher: KORA
ISBN: 978-87-7509-963-4
Project: 10726

KORA
The Danish Institute
for Local and Regional Government Research

KORA is an independent government institution, the purpose of which is to promote quality development and better resource use and management in the public sector.

Det Nationale Institut
for Kommuners og Regioners
Analyse og Forskning

Købmagergade 22
1150 København K
E-mail: kora@kora.dk
Telefon: 444 555 00

TRYGFONDENS
BØRNEFORSKNINGSCENTER

Preface

This report addresses the members of TrygFonden's Centre for Child Research, as well as other researchers and practitioners with an interest in measuring child well-being. The overall aim is to identify a battery of instruments that are suitable as common measures of well-being across a wide range of studies of children and adolescents.

The report was prepared in the time period from December 2013 to December 2014 by Professor Michael Rosholm, Professor Jacob Nielsen Arendt and Researcher Astrid Kiil. However, information about the price and availability of the reviewed questionnaires was updated in April 2016. The work is financed by TrygFonden's Centre for Child Research.

We are fully aware that the assessment and selection of preferred instruments entail a large degree of subjectivity, and we encourage discussion of the various assessments and decisions that have been made along the way.

We would like to thank Anna Piil Damm, Carsten Obel, Dorthe Bleses, Jill Mehlbye, Johanne Jeppesen Lomholt, Mads Bonde Ubbesen, Maiken Pontoppidan, Mikael Thastum, Miriam Wüst, Per Bech, Pernille Skovbo Rasmussen, Sarah von Mastrigt, Simon Calmar Andersen and Peter Jensen for valuable input. We are especially indebted to Nina Madsen Sjö for her input and helpful comments during the working process. The report was revised in accordance with the comments from two anonymous reviewers in April 2016. However, the authors alone are responsible for the contents of the report.

May 2016,

Astrid Kiil

Table of Contents

Danish summary

For at politikere og andre beslutningstagere skal kunne bruge økonomiske evalueringer til at prioritere mellem forskellige tiltag målrettet børn og unge, er det nødvendigt, at de så vidt muligt er sammenlignelige på tværs af interventioner og aldersgrupper. Et skridt på vejen mod sammenlignelighed er, at der inkluderes sammenlignelige mål af effekter på bl.a. børns trivsel på tværs af et bredt udvalg af studier af børn og unge. Herved bliver det muligt at gennemføre cost-effectiveness-analyser, der kan sige noget om, med hvilke interventioner og for hvilke grupper af børn og unge man kan få mest trivsel for pengene. Med mange forskningsprojekter i opstartsfasen og en tværfaglig forskergruppe udgør Trygfondens Børneforskningscenter et unikt forum for at diskutere og arbejde hen imod, at der indsamles data på nogle fælles effektmål i fremtidige studier af børn og unge. Indsamling af data på nogle fælles effektmål er dog i høj grad relevant og aktuel også for andre forskere og praktikere, der er interesserede i at kunne sammenligne deres resultater med andre og derved spille ind i en bredere samfundsmæssig prioriteringsdagsorden.

Formål

I denne rapport identificeres forslag til spørgeskemaer, der er velegnede som mål af børns trivsel, på tværs af et bredt udvalg af studier af børn og unge. Rapporten er primært målrettet forskergruppen i Trygfondens Børneforskningscenter, men kan også bruges som inspiration til overvejelser omkring definition og måling af børns trivsel af andre forskere og praktikere.

Metode

Potentielle spørgeskemaer blev identificeret ved at indhente forslag fra projektlederne i Trygfondens Børneforskningscenter samt andre tilknyttede eksperter med omfattende forskningserfaring inden for børn og unge-området.

En indledningsvis screening førte til, at 27 spørgeskemaer blev udvalgt og beskrevet med hensyn til:

- Hvilke skalaer og subskalaer, de indeholder
- Svarkategorier
- Hvem der kan bruges som respondenter
- Hvilket aldersinterval, de dækker over
- Hvor lang tid, det tager at besvare dem
- Fordelingen af positivt og negativt ladede spørgsmål
- Ophavsrettigheder og pris for brug
- Hvornår de er udviklet og oversat til dansk.

Resultater

Spørgeskemaer blev udvalgt med fokus på centrale kernemål for børns udvikling givet ved subjektiv og psykosocial trivsel, inklusiv social og følelsesmæssig funktion. De udvalgte spørgeskemaer er velegnede til anvendelse på tværs af de fleste eller alle centrets studier. Projektet identificerer desuden et udvalg af supplerende mål tiltænkt at dække andre aspekter af børns trivsel, inklusiv personlighedstræk, eksekutiv funktion og andre koncepter for børns udvikling. Disse supplerende mål er tiltænkt som forslag til projektledere som leder efter instrumenter til at måle de nævnte aspekter af børns trivsel.

1 Introduction

Considering the potential of quantitative research on children and adolescents to inform policy, the comparability of effect estimates across studies and the possibility of carrying out economic evaluations that are (at least to some extent) comparable across interventions and age groups are crucial. The comparability across evaluations of different interventions targeted at children and adolescents can be improved by including similar or common outcome measures as an integrated part of the data collection across studies. The importance and potential of a higher degree of standardisation across studies in this respect are emphasised in a recent overview article, which concluded that the greatest variation in practice in cost-benefit analysis of early childhood interventions concerns the outcomes and the values attached to these (Karoly 2012). Thus, although complete standardisation of outcomes is neither possible nor desirable, it is evident that a higher degree of standardisation than the current one will improve the scope for comparability, with the ultimate goal of informing policy discussions and prioritisation.

The primary outcome of this report is identification of a battery of instruments that are suitable as common measures of well-being across a wide range of studies of children and adolescents in TrygFonden's Centre for Child Research. With numerous research projects in the start-up phase or on the drawing board and a group of cross-disciplinary researchers, the recently established TrygFonden's Centre for Child Research provides a unique forum for discussing and working towards the collection of data on some common overall outcomes in future studies of children and adolescents. In addition, we hope that the report will be useful to other researchers and practitioners by providing inspiration for considerations about how to define and measure the well-being of children and adolescents. The importance of using carefully selected common outcome measuring instruments across studies, also in the "softer" areas, is increasingly being recognized by the Danish ministries and authorities, who have initiated several systematic reviews of outcome measures in recent years (see e.g. Brauner et al. (2011), Pontoppidan & Niss (2014) and Keilow et al. (2014)).[1] Also, the Rockwool Foundation has initiated and financed a review of the measurement of social-emotional function in a Danish context, which is conducted by Nina Madsen Sjö and will be published in the course of 2016.

The remainder of this report is structured as follows: Section 1 accounts for the aim and intention of the present report and summarises the battery of instruments. Section 2 briefly outlines some different ways of thinking about child well-being and its measurement. Section 3 accounts for the methods used to identify potential instruments and the lines along which they are assessed. Section 4 contains a structured assessment of the potential instruments. Section 5 discusses the merits and limitations of a selection of the potential instruments and proposes a battery of instruments that are suitable for use across the future studies conducted in the setting of TrygFonden's Centre for Child Research.

1.1 Aim and intention of this report

The overall aim of this report is to facilitate the discussion regarding the identification of a battery of instruments which are suitable as common measures of well-being or related constructs across a wide range of studies of children and adolescents. Measures tailored to all age groups of children from birth through adolescence are considered.

The proposed battery of instruments should be considered as an input to support the work of the project managers and not as a strict guideline. Moreover, it is important to emphasise that the proposed battery

---

[1] Brauner et al. (2011) identified and assessed instruments to screen young criminals for antisocial behaviour and to measure and compare effects of various interventions targeted at this group. At the request of The National Board of Social Services, Pontoppidan & Niss (2014) identified and assessed instruments to measure the well-being of 0-3 year old children. Finally, Keilow et al. (2014) conducted a pilot study to construct, test and validate a questionnaire to measure important dimensions of the well-being and teaching environment of Danish schoolchildren. The pilot study was conducted at the request of the Danish Ministry of Education, and the resulting question frame is intended for use in national well-being measurements.

of instruments is in no way intended to replace the specific outcomes of the individual projects, but rather to facilitate the collection of data that are at least to some extent comparable across studies.

It is also essential to stress that the report is by no means intended to cover all aspects of child outcomes, or to be a review of existing measures within the domains focused upon.

1.1.1 Focus

The term well-being is generally used in the research literature as an over-arching concept regarding the quality of people's lives, but there is no agreement as to the definition of the term (Rees et al. 2010).

One way to think about well-being is to distinguish between objective and subjective well-being (Pople & Solomon 2011, Rees et al. 2013). *Objective well-being* concerns the social and economic 'objective realities' that are believed to contribute to well-being, such as wealth, absence of disease and educational attainment. Objective measures are frequently used to rank countries in international comparisons (see e.g. OECD (2009) and UNICEF (2013)), and they are also common outcomes in economic evaluations (see e.g. Karoly (2012)). *Subjective well-being* is a broad concept that both relates to how satisfied people are with their lives and mental capabilities such as personality traits and mental health.

While the Danish registries contain data in abundance on objective indicators of well-being, so that common measures to be considered across studies are readily available, data on subjective well-being must be collected using questionnaire surveys or interviews. Hence, the collection of similar or common data on subjective well-being across the various studies conducted in the setting of TrygFonden's Centre for Child Research requires some cooperation among the project managers in the planning phase. For this reason, the focus of this report is measurement of subjective well-being.[2] Section 2 contains a discussion of the theoretical frameworks that can be used to consider the well-being of children and adolescents with emphasis on subjective well-being.

In addition, we restrict the attention to questionnaire rating scales that can be completed by children or adolescents (self-report), parents, other caregivers or teachers. Thus, we exclude questionnaires requiring trained psychologists to administer them. The questionnaire rating scale is preferred over the alternatives of interview and observational formats, because it requires less training for administration and is less costly when collecting data on large samples.

1.1.1 Limitations

The method used to identify the instruments considered in this report, which is outlined in Section 3.1, is nowhere near a systematic review, and it does not claim to be so. This necessarily implies that not *all* existing instruments for measuring well-being among children and adolescents are considered. While this is of course a limitation, we nevertheless hope to have reached a sensible balance between comprehensiveness and brevity, given the aim and intention of the report.

Along a similar line, we do not cover the various aspects of how the instruments are implemented and used in practice. However, we fully acknowledge that successful implementation is crucial for the instruments to work as intended. It is thus important to be aware that the need for education and training of the administrators, as well as the handling of data, differs from instrument to instrument, and that this aspect of the measurement process requires a great deal of attention from project managers and researchers.

Finally, new instruments are continuously being developed and translated into Danish, just as new data and evidence on the properties of existing instruments continue to be published at a rapid pace. Hence,

---

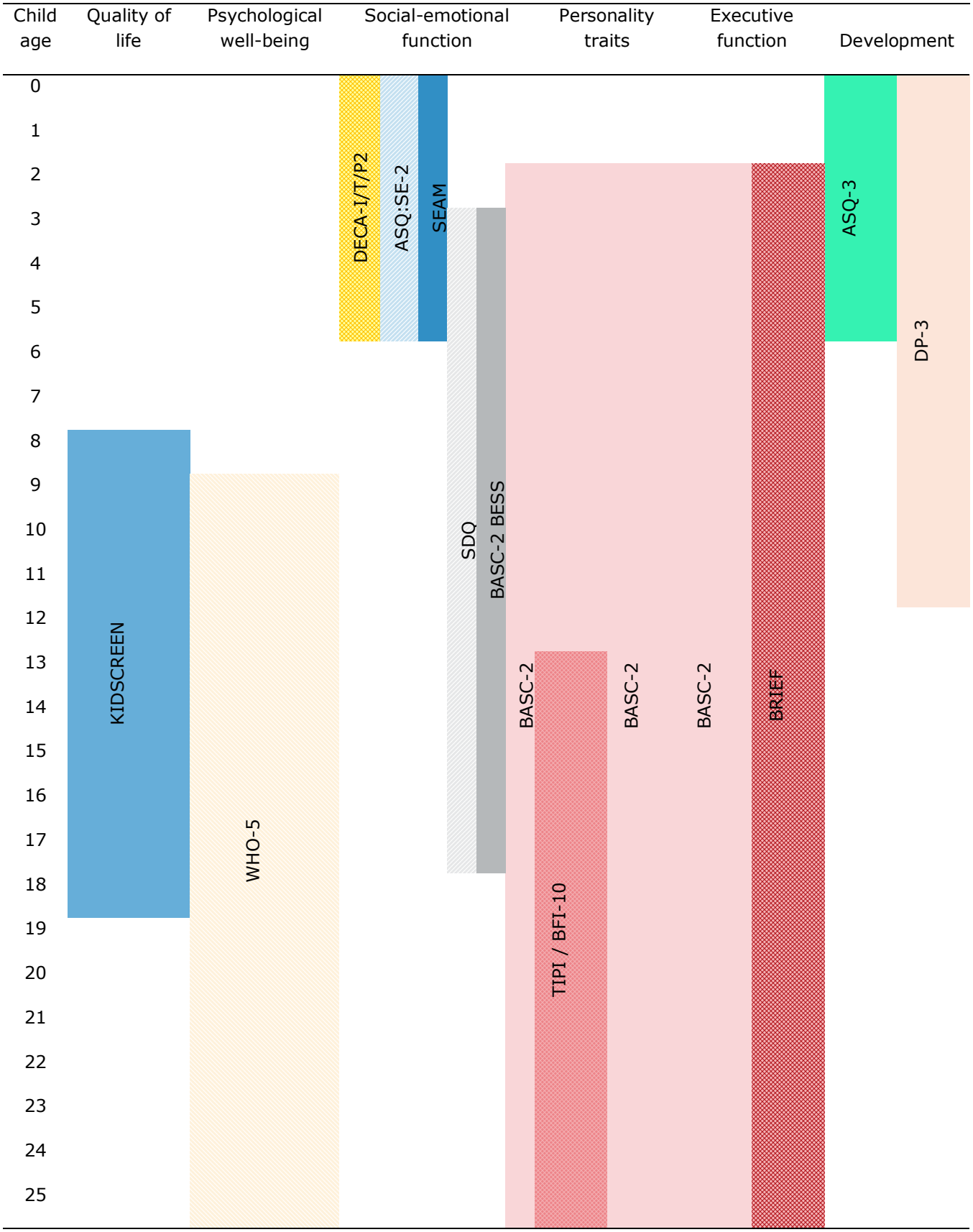
[2] In a similar manner, it may be argued that measures of child cognition are also only sparsely available in the national registries. National test scores for school children is a relevant outcome, which may capture cognitive capabilities, but it is also influenced by other factors. However, we have chosen not to focus on cognitive skills, both to keep the scope of this report at a manageable level and because most instruments require thorough testing, which is not suitable for surveys.

future readers should keep in mind that this report contains a snapshot of the information available to the authors in December 2014 (although information about price and availability was updated in April 2016).

1.2 Proposed battery of instruments (summary)

Figure 1.1 provides an overview of the various instruments selected as the best candidates to be included as common outcome measures across the studies conducted in the setting of TrygFonden's Centre for Child Research. The figure describes the focus areas of the various instruments and the age range for which they are applicable, and is elaborated on in Section 5. For all instruments answered by children or adolescents themselves, it should be kept in mind that even seemingly innocent questions can start a therapeutic process in the child. Hence, it is of crucial importance to ensure that an adult is present when the questionnaire is administered, so that the child can talk to somebody about any worries or thoughts that may arise.

Figure 1.1 Candidates for the proposed battery of instruments



It was decided to include measures of both subjective (or 'hedonic') and psychological (or 'eudaimonic') well-being, including social-emotional function, as core measures in the proposed battery of instruments. The core measures are intended for inclusion in most or all of the studies in TrygFonden's Centre for Child Research. In addition to the core measures, a selection of supplementary measures intended to cover further aspects of child well-being, including personality traits, executive function and the more comprehensive concept of psychomotor development, were included in the battery. These supplementary measures are merely intended as suggestions for project managers who are looking for instruments to measure the mentioned aspects of child well-being.

2 Frameworks for considering child well-being

Many different frameworks of self-reported well-being have been developed, and while no consensus exists some broad concepts are fairly well-established. For one thing, several strands of the literature distinguish between subjective (or 'hedonic') well-being, which focuses on happiness and life satisfaction, and psychological (or 'eudaimonic') well-being, which incorporates dynamic processes such as personal development and growth (Pople & Solomon 2011, Conti & Heckman 2012). Subjective well-being is typically measured by directly asking how the children feel, using instruments such as Cantril's ladder (Cantril 1965) and Huebner's life satisfaction scale (Huebner 1991) or by asking about the presence of positive and negative affect using for example the Positive and Negative Affect Schedule (PANAS) for children and parents (Ebesutani et al. 2012). Psychological well-being covers more complex concepts, such as self-esteem, sense of control and depression, and is typically measured using batteries of questions.

The drawback of subjective well-being measures is that it is uncertain exactly what they measure. Some scientific work partly supports this lack of reliability (Bertrand & Mullainathan 2001), while other recent evidence is more encouraging (Krueger & Schkade 2007). Ample psychological evidence, cited in Frey and Stutzer (2002), and Blanchflower and Oswald (2004), for instance, confirms that self-reported measures of happiness and satisfaction are valid and reliable, and subjective well-being data pass a variety of validation exercises. Along similar lines, there will inevitably be limitations to the usefulness of self-reported measures related to children's levels of understanding, literacy and so on. However, according to Rees et al. (2013) it is widely accepted in the field of well-being research that it is possible and valid to ask children and young people from at least age 10 upwards to report on aspects of their own well-being. In addition, some evidence suggests that children as young as 6 years old can reliably self-report, if an age appropriate measure is used, especially where measures are specifically developed for this age group (Deighton et al. 2012, Ben-Arieh 2006).

Another prevailing distinction in the literature is between the present well-being of children and a more future-oriented focus (i.e. preparing children for a productive and happy adulthood), which may be described by the term well-becoming (Ben-Arieh 2006). However, the two perspectives are not necessarily mutually exclusive.[3]

The economic framework conceptualizes child well-being in a developmental perspective (see e.g. Heckman (2007) and Conti & Heckman (2012)). This framework has adopted the concept of psychological well-being and considers the child as a work in progress. The idea is embedded in a lifecycle framework of human development that distinguishes between indicators of well-being that are amenable to policy intervention, proxies of underlying well-being and outcomes. This line of work considers indicators of well-becoming, such as the capabilities cognition, personality traits and health (Conti & Heckman 2012). There is strong evidence that both cognition and personality traits predict adult life, and some evidence indicates that production of different skills is both complementary and dynamic in nature. However, the background for identification and measurement of such capabilities is highly diverse and stretches into various areas of research.

Almlund et al. (2011) contains a discussion of the complex problems of defining and measuring personality and cognitive skills. They highlight two traditions: One focusing on personality traits and the other focusing on social cognition. The former (e.g. McCrae & Costa (2008)) holds that personality traits evolve through biological processes, so that investments and experience do not affect traits, although individuals may learn about themselves (their traits) by taking actions. Preferences or individual objectives play no part in this theory. The Big Five theories were developed as part of this tradition. The Big Five theories perhaps do not so much constitute a coherent theory – indeed they have been criticized for being atheoretical – as a

[3] In their taxonomy for child well-being indicators, Ben-Arieh & Frønes (2011) argue that the present status and position of children have to be understood within the framework of the present, as description, and within the framework of their life course and development, as predictions. The total well-being includes the well-being of the present and the predicted well-being of the future.

construct based on factor analysis, boiling the variation in personality outcomes down to the five common constructs, abbreviated OCEAN: Openness-to-experience, Conscientiousness, Extra-version, Agreeableness and Neuroticism. "Social cognitive" theories on the other hand stress the role of cognition in shaping personality, and according to these theories the role of social context in shaping actions and self-knowledge, individual goals and motives (preferences) also shape actions.

Almlund et al. (2011) highlights the potential problem of identifying and measuring personality traits through performance measures or self-reported measures, as both are influenced simultaneously by several traits, efforts and situational specificities. They also stress that standard psychometric operationalization of outcomes hinges crucially upon specific assumptions (such as linearity and exclusion restrictions, and rarely with correction for effort and environment), and that validity tests like the construct validity test entail an inherent risk of circularity. However, this is not to say that other approaches are without problems.

There are numerous other approaches to the measurement of child outcomes. Almlund et al. (2011) mention important strands of literature. Neuroscience stresses the executive functions, which overlap with both certain personality traits and more traditional measures of fluid intelligence. Diamond (2013) highlights three executive functions for children that seem to matter greatly for their future outcomes: Working memory (reasoning, planning), cognitive flexibility (e.g. ability to switch perspective) and inhibitory control (including self-control and discipline). Another strand of literature has particular focus on individuals' perception of themselves; e.g. their self-esteem and locus of control. Self-esteem refers to an individual's subjective estimation of his or her own worth. An example of a measure is the widely used Rosenberg Self-Esteem Scale (Rosenberg 1989). Locus of control refers to one's belief about whether the determinants of one's life events are largely internal or external. Those with an internal locus of control believe that life events are typically caused by their own actions. An example of a measure is the widely used Rotter Locus of Control Scale (Rotter 1966). A related concept is that of generalized self-efficacy (the belief that one can act effectively to bring about desired results).

For the most part, researchers who study self-esteem and locus of control have carried out their work isolated from each other and without reference to the Big Five taxonomy. Judge et al. (2002) and others have proposed that locus of control, self-esteem and the Big Five construct "emotional stability" (where emotional instability relates to neuroticism) are indicators of a common construct, termed core self-evaluations. Psychopathology, the study of abnormal behavior and mental illness, has also been studied more or less independently of the previous strands of literature. Recent attempts have been made to join the personality trait literature and psychopathology, in viewing mental disorders as extreme variants of personality traits. Another branch of the literature considers social-emotional function (see e.g. Haggerty et al. (2011) and Humphrey et al. (2011)). The US-based Collaborative for Academic, Social and Emotional Learning (CASEL) has identified five interrelated social-emotional competencies that are necessary for effective life functioning: Self-awareness, self-management, social awareness, relationship skills and responsible decision making (CASEL 2013). These skills have been shown to matter for both social and academic performance (Payton et al. 2008), and the importance of social-emotional function in determining an individual's income and other well-being parameters is also increasingly recognized by economists (see e.g. Cunha, Heckman & Schennach (2010), Heckman (2007) and Heckman, Yi & Zhang (2013)).[4] Finally, a separate body of literature is the developmental psychology literature that deals with child development and temperament, and bears resemblance to the Big Five theories. However, there is much less consensus on higher order factors in the child development literature, but there is evidence that they are related to adult personality, and that temperament, though established early in life, is only partly heritable and affected by environment.

Summing up, it is widely agreed that child well-being matters for a variety of reasons, and some broad concepts and distinctions, such as subjective versus psychological and present versus future well-being,

[4] In relation to social-emotional function, it is worth noting that the economic literature appears to use the terms personality traits and social-emotional traits/function interchangeably (Almlund et al. 2011, Heckman, Yi & Zhang 2013), while others consider social-emotional skills and personality as different theoretical constructs (Humphrey et al. 2011, Lopes et al. 2004).

are fairly well-established in the literature. However, when it comes to more specific definitions and measurement it is clear from the above discussion that different strands of the well-being literature have adopted different focus areas and approaches. In order to keep the options open, the remainder of this report will consider measures from various areas of research, including measures of behavioural difficulties, social-emotional function, psychological well-being, personality traits, overall quality of life etc.

3 Method

This section accounts for the method used to identify and assess potential instruments. Section 3.1 describes the approach used to identify potential instruments, Section 3.2 describes the criteria used for the initial screening, and Section 3.3 lays out the criteria used to assess the instruments that have passed the initial screening and to select those that are suitable for inclusion in the proposed battery.

3.1 Identification of possible instruments

Potential instruments were identified by asking a group of project leaders at TrygFonden's Centre for Child Research and other associated experts for suggestions. Hence, the method used to identify the instruments discussed in this report is nowhere near a systematic review, and it does not claim to be so. The expert group mainly consists of psychologists and economists, all with extensive research experience in the area of children and adolescents.

The members of the expert group suggested several potential instruments and paths. Some recommended specific standardised questionnaires, others suggested consulting the test catalogues of established psychological publishers, while others again referred to literature reviews and policy recommendations.[5] We have followed the various leads by obtaining more information from test manuals, psychometric journal articles and webpages, and by going through the suggested test catalogues, literature reviews and policy recommendations.

Overall, the proposed instruments need to be sufficiently brief to allow for routine use, yet sufficiently wide, covering broad categories of the most common issues related to child well-being.[6] In general, it is recommendable to consider outcomes that are expected to be sensitive to the intervention in question. However, looking for instruments that are suitable as common measures of well-being across a wide range of studies of varying interventions targeted at children and adolescents is a difficult task. Based on inputs from the expert group, the discussion of the various frameworks for considering child well-being in Section 2, and on considerations regarding what is feasible in practice, it was decided to include measures of both subjective (or 'hedonic') well-being and psychological (or 'eudaimonic') well-being, including social-emotional function, as core measures in the proposed battery of instruments. The overall measures of subjective well-being are expected to be able to shed light on which groups of children the various studies are dealing with. Regarding social-emotional function, it was decided to zoom in on this aspect of psychological well-being given that the most commonly reported mental health difficulties in children and adolescents are either behavioral or emotional (Deighton et al. 2012). In addition to the core measures, it was decided to also include a selection of supplementary measures in the battery. The supplementary measures are intended to cover further aspects of child well-being, including personality traits, executive function and the more comprehensive concept of psychomotor development. The focus on subjective and psychological well-being, including social-emotional function, as well as the choice of supplementary measures are subject to discussion.

3.2 Initial screening

In order to target the effort, we subjected the identified instruments to an initial screening. Four criteria for inclusion were determined as the minimum requirements instruments must meet in order to be deemed

---

[5] The following test catalogues, literature reviews and policy recommendations were examined: Hogrefe Psykologisk Forlag (2013), Williams (2008), Wolpert (2008), Humphrey (2011), Metodecentret (2013), Ringwalt (2008), Solans et al. (2008), CAMHS, Evidence Based Practice Unit (2008) and UCL & Anna Freud Centre (2011).

[6] As mentioned in Section 1.1.1, it was deliberately decided not to focus on cognitive skills, both to keep the scope of the report at a manageable level and because most instruments require thorough testing, which is not suitable for surveys.

feasible to administer on a large scale and suitable as common overall measures in future studies conducted in the setting of TrygFonden's Centre for Child Research. The criteria for inclusion are as follows:

1) Instruments should measure well-being or related constructs in children or adolescents
2) Only generic instruments are considered
3) The psychometric properties of the instruments should be validated at the individual level
4) Possible reporters should include at least one of the following: Children or adolescents, parents or other caregivers and teachers.

The suggested instruments had to meet all the criteria in order to be included for further consideration. Hence, instruments that failed to meet just one of the criteria were excluded at this point. As stated in Section 1.1.1, we limit the attention to generic instruments that measure well-being or related constructs in children and adolescents, in order to allow for comparison across different conditions and settings and between healthy and sick children. The focus on generic instruments implies that non-generic instruments, i.e. instruments limited to specific groups (not counting age groups) or conditions such as depression, ADHD, anxiety or autism, are disregarded at this point. Moreover, the inclusion criteria imply that we exclude instruments that are mainly known and used only by a Danish or Scandinavian audience. This restriction is intended to increase the probability of having the resulting research published in high impact international journals. However, it was decided to deviate from criterion 3 for instruments targeted at infants and toddlers, and also include instruments in the development phase that have not yet been assessed psychometrically, since we only identified very few instruments applicable to this age range. Finally, we excluded instruments that cannot be completed by children or adolescents themselves, parents or other caregivers, or teachers, but require reporting by professionals, such as psychologists or specially trained educators.[7]

Table 3.1 lists the identified instruments and decisions regarding exclusion or inclusion for further consideration, at this point.

[7] Hence, Bayley-III and PPVT-4 were excluded at this point, because both instruments consist of a test that has to be conducted for each individual child by a trained assessment specialist. Moreover, the focus of PPVT-4 is strictly on linguistic skills.

Table 3.1 Identified instruments and results of initial screening

Acronym	Full name	Decision
ABAS-II	Adaptive Behavior Assessment System – Second Edition	Include
ASEBA: CBCL, TRF, YSR	Achenbach System of Empirically Based Assessment: Child Behavior Checklist, Teacher Report Form, Youth Self-Report	Include
ASQ-3	Ages & Stages Questionnaires, Third Edition	Include
ASQ:SE	Ages & Stages Questionnaires: Social-Emotional	Include
BASC-2	Behavior Assessment System for Children, Second Edition	Include
BASC-2 BESS	BASC-2 Behavioral and Emotional Screening System	Include
BFI-10	Big Five Inventory – 10-item version	Include
BFQ-C	Big Five Questionnaire for Children	Include
BRIEF-F/SR/V	Behavior Rating Inventory of Executive Function	Include
BSMB	Bedre Sundhed for Mor og Barn / Danish National Birth Cohort - Questions on motor and cognitive development	Exclude on criterion 3
CDI	MacArthur-Bates Communicative Development Inventory	Include
Conners CBRS	Conners Comprehensive Behavior Rating Scales	Include
DECA -I/T/P2	Devereux Early Childhood Assessment for - Infants/Toddlers/Preschoolers	Include
DESSA	Devereux Student Strengths Assessment	Include
DP-3	Developmental Profile – Third Edition	Include
HBSC	Health Behavior in School-aged Children (Skolebørnsundersøgelsen)	Include
KIDSCREEN	KIDSCREEN Health Questionnaires for Children and Young People	Include
MACI	Millon Adolescent Clinical Inventory	Include
M&MS	Me and My School	Include
NIH Toolbox	National Institutes of Health Toolbox for the Assessment of Neurological and Behavioral Function	Include
SEAM	Social Emotional Assessment/Evaluation Measure	Include
SDQ	Strengths and Difficulties Questionnaire	Include
SFI pilot	SFI – The Danish National Centre for Social Research, pilot study: The well-being and teaching environment of Danish schoolchildren (Trivselsmålinger i folkeskolen)	Include
SSRS/SSIS(-RS)	Social Skills Rating System/Social Skills Improvement System-Rating Scales (SSIS-RS is a 2008 revision of SSRS)	Include
Termometeret	Termometeret, Dansk Center for Undervisningsmiljø	Exclude on criterion 3
TIPI	Ten Item Personality Measure	Include
WHO-5	WHO-five Well-being Index	Include

It can be seen from Table 3.1 that all of the identified instruments are deemed to be of the generic type and to measure well-being or related constructs. Three instruments (Termometeret, BSMB and M&MS) were excluded at this point, because the psychometric properties of these instruments have not (yet) been validated at the individual level, to the best knowledge of the authors. However, the newly developed Danish questionnaire and benchmarking system Termometeret, may very well turn out to be of interest to the researchers at TrygFonden's Centre for Child Research in the future when it has been implemented by the Danish schools, and hopefully also psychometrically validated at the individual level at some point. However, it is a concern that Termometeret only being available in Danish may make it more difficult to

get the resulting research published in high impact international journals. The reader is referred to the publishers' web-pages for further information on the instruments excluded at this point[8].

3.3 Assessment criteria

After the initial screening, the selected instruments are described and assessed along the dimensions given in Table 3.2. The assessment criteria include characteristics of the instruments and issues related to administration. In addition to the criteria outlined in Table 3.2, it is also assessed whether instruments include questions that are particularly likely to start a therapeutic process (such as questions on death or suicidal thoughts) and thus require immediate follow-up by a professional, or questions that may be perceived as offensive by parents. It is not ethically responsible to administer such questions in large-scale studies, where the surveyed children are not necessarily able to talk to a professional. Moreover, when administered to parents, they may result in low response rates.

[8] Termometeret: http://dcum.dk/undervisningsmiljoe/termometeret-grundskolen
BSMB: http://www.ssi.dk/Forskning/Forskningsomraader/Epidemiologi/BSMB/onsker%20du%20at%20forske/De%204%20forste%20interviews.aspx
M&MS: http://www.ucl.ac.uk/ebpu/docs/publication_files/tamhs_report (report)

Table 3.2 Assessment criteria and related consideration

Assessment criteria	Considerations
Scales and subscales	Describes what the instrument is intended to measure.
Response categories	Describes the response categories, e.g. Likert scales of varying lengths or yes/no.
Reporters	States who provides the information. As stated in Section 3.1, respondents may be children or adolescents, parents, other caregivers or teachers. Some instruments are available in several versions adapted to different types of reporters, while others are only targeted at one reporter type. For the given purpose, the more potential reporters the better, since this increases the choice and flexibility available to project managers when collecting well-being data. However, it should be noted that different respondent types might provide conflicting answers regarding the same children as child well-being is based on a subjective assessment and might vary with context.
Age range	States the age group for which the measure is applicable. For the given purpose, measures that are applicable within a reasonably wide age range are preferred.
Items/approximate completion time	Gives the number of items and approximate completion time. This is important in the current context, where the collection of data on well-being is often done in addition to the specific outcomes included in the individual projects. Hence, instruments that take no more than 10 minutes to complete are greatly preferred to more lengthy ones, as this keeps the burden on the reporters at a reasonable level.
Positive/negative items	Describes the distribution of positive and negative items in the instrument. Instruments containing mainly positive items are preferred, since this is expected to increase acceptance among teachers and caregivers. In addition, it is preferable that instruments do not contain questions that may be perceived as insulting or offensive by a majority of parents.
Danish version and norms	States whether the instrument is translated into Danish. It is considered an advantage if at least one of the questionnaires encompassed by the instrument is available in Danish. Likewise, it is noted whether there are norms for the Danish version of the questionnaire. While instruments with Danish norms are preferable, this is not considered to be of crucial importance.
Copyrights	States who holds the copyright for the instrument. In some cases the copyright agreement states that the instrument may not be modified in any way, while in other cases it is possible to opt in or out of different modules or subscales according to needs.
Price	Gives the price charged by the copyright holder for use of the instrument. This is an important criterion, because we are looking for instruments that are suitable to use in addition to the project-specific outcomes across several studies and for a large number of children. Some measures are free or subject to a symbolic one-off payment, while others require purchase of manuals, logins or software packages, possibly combined with a fee per questionnaire. In some cases, the price depends on whether the instrument is used for commercial or non-commercial purposes, and it is often possible to negotiate price reductions if buying in bulk. We take a conservative approach and state the list prices charged by the publishers. However, it should be kept in mind that especially per questionnaire fees are subject to negotiation. Some of the fees have been converted to Danish Crowns using the exchange rates at the time of writing. These fees are thus subject to some uncertainty and should be considered rough estimates.
Development	States the year of original publication and Danish translation, respectively. This information is important, because society's values and children's qualifications changes over time, and these circumstances affect which questions it is relevant to ask.

3.4 Psychometric properties

As stated in section 3.2, it is a minimum requirement that the psychometric properties of the instrument should be validated at the individual level in order for the instrument to be included for further consideration. In addition, psychometric properties have implicitly influenced the selection process, as the consulted project leaders and experts are likely to suggest using instruments with sound psychometric properties, and we have limited our attention to instruments that have received a positive evaluation in the examined literature reviews and policy recommendations. However, it is outside the scope of this report to review the psychometric properties of the identified instruments explicitly. Appendix A provides a brief introduction to the psychometric properties that can be applied to assess the performance of measurement instruments,

for readers who are not familiar with this topic. Further information about the application of psychometric properties is provided by, for instance, the European Federation of Psychologists' Associations (EFPA), who have developed a set of test review criteria (Evers et al. 2013). When deciding on whether an instrument is suitable for use in evaluations of a program or intervention, the responsiveness of the instrument, i.e. its ability to measure progression, is of particular importance. Whether an instrument appears to be able to measure progression is largely dependent on a wide range of things other than the instrument itself, such as the actual effect of the evaluated program or intervention on the dimensions measured by the instrument, the implementation of the program or intervention and its time frame, target group etc. This implies that an instrument that appears to be responsive in one context may not be so under different circumstances. Hence, researchers are encouraged to seek out the latest evidence on whether the relevant instruments are able to measure progression for programs similar to the one they are planning to evaluate, and under comparable circumstances.

4 Description of potential instruments

This section describes the instruments selected for further consideration in Section 3.2 along the dimensions outlined in Section 3.3. Table 4.1 accounts for the characteristics of the instruments and issues related to the administration. The information about the price and availability of the reviewed questionnaires was updated in April 2016. The information provided in the table is intended to guide the identification and discussion of a battery of instruments that are suitable as common measures of well-being or related constructs across a wide range of studies of children and adolescents. Information about each instrument was gathered from different sources. We relied primarily on the manuals or web-based information available from authors or publishers, supplemented by pertinent information found in research articles. For the less developed instruments, we reviewed the instrument itself and the supporting material we were able to locate, such as research reports and personal communications with authors of the instruments. Table 4.1 presents the relevant information in a compressed and somewhat simplified format, in order to provide the reader with a brief and structured overview of the many instruments. Links to sample questionnaires, when these are available, and sources of further information (such as contact persons for unofficial translations and referrals to assessments of the psychometric properties of the instruments) can be found in Appendix B. Since most of the instruments are copyrighted documents, it is not possible to include the actual questionnaires in the appendix.

The variety of scales found in Table 4.1 underlines the point made in Section 2, namely that child well-being is many things. For most of the instruments, all scales are available to all possible reporters. However, in some cases, such as BASC-2, the selection of scales differs for the different reporters. Another thing which can be learned from Table 4.1 is that the reviewed instruments typically include several questionnaires for different age groups and reporters. For obvious reasons, the selection of reporters is highly correlated with the targeted age range. Instruments that target younger children typically use parents and/or caregivers as reporters, instruments that target school-aged children typically use parents and/or teachers, and self-report can be used from age 8 and up. Some instruments place restrictions or make recommendations regarding who are suitable reporters. For example, the ASQ questionnaires recommend using reporters who spend at least 20 hours per week with the child they assess. Several instruments, especially those spanning a wide age range and those targeting infants and toddlers, include several age-specific questionnaires. Moreover, in some cases, such as DP-3, each of the age-specific questionnaires contains several blocks of questions targeted at different age intervals. Considering the number of items and approximate completion times given in Table 4.1Table 4.1, it is indisputable that some of the reviewed instruments are too lengthy for the purpose. However, these instruments are still included for further consideration because it is often possible to administer selected scales. The drawback of this approach is that the psychometric properties have usually been assessed for the full instruments, and that use of selected scales requires permission from the publisher, which may be time consuming to obtain. Regarding availability, further details about the unofficial translations can be found in Appendix B. Another thing to keep in mind when assessing the information provided in Table 4.1 is that the per questionnaire fees are often subject to negotiation, as mentioned in Table 3.2.

Table 4.1 Brief description of the selected instruments

Instrument	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. completion time	Positive/ negative items	Danish version and norms	Copyrights and approximate price	Development
ABAS-II	Adaptive behavior and skills: - Communication - Functional academics - Self-direction - Social - Leisure - Community use - Home or school living - Health and safety - Self-care - Work (adolescents and adults) - Motoric (0-5 years)	4-point Likert scale from *is not able* to *always or almost always when needed*	Parents Caregivers Teachers Self-report from age 16	0-89 years (2 age-specific questionnaires)	193-241 items 15-40 min.	All items positive	Official Danish version	Copyrighted by Pearson Assessment Danish version distributed by Hogrefe Psykologisk Forlag Manual and sample questionnaires: DKK 1550 DKK 20/additional questionnaire	First edition published in 2000. ABAS-II published in 2003 and revised in 2008 Danish translation published in 2011
ASEBA: CBCL, TRF, YSR	Behavioral and social-emotional problems: - Externalizing behavior problems - Internalizing behavior problems (the underlying subscales differ between CBCL, TRF and YSR) + DSM-oriented subscales	3-point Likert scale: *not true, somewhat or sometimes true, very true* or *often true*	Parents (CBCL) Caregivers or teachers (TRF) Self-report from age 11 (YSR)	1.5-18 years (2 age-specific questionnaires)	99-118 items Preschool: 10-15 min. School: 20-25 min.	All items negative	Official Danish version Danish norms from 2009/2010	Copyrighted by Achenbach System of Empirically Based Assessment (ASEBA) Danish version distributed by Psykiatrien i Region Syddanmark Login: DKK 5000 DKK 5/questionnaire	First editions of CBCL, TRF, YSR and preschool CBCL manuals published in 1983, 1986, 1887 and 1992, respectively, and revised in 2001 Danish translation published in 1999 and revised in 2012
ASQ-3	Psychomotor development: - Communication - Gross motor skills - Fine motor skills - Problem solving - Personal-Social	3-point Likert scale: *yes, sometimes, not yet* +indicate if overall concern	Parents (or primary caregivers)	1 month-5.5 years (21 age-specific questionnaires)	30 items 10-15 min.	Asks about specific skills	Official Danish version expected Research translations of 10, 24 and 48 month questionnaires	Copyrighted by Brookes Publishing Dansk Psykologisk Forlag has distributorship of Danish version and expects to publish this during 2016 Most likely price structure is fee/questionnaire	First edition (ASQ) published in 1995, and revised in 1999 (ASQ-2) and 2009 (ASQ-3) Danish translation of ASQ-3 will be published in 2016

Instrument	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. completion time	Positive/ negative items	Danish version and norms	Copyrights and approximate price	Development
ASQ:SE (ASQ:SE-2)	Social-emotional development: - Self-regulation - Compliance - Communication - Adaptive behaviors - Autonomy - Affect - Interaction with people	3-point Likert scale: *often or always, sometimes, rarely* +check box if concern for each item	Parents (or primary caregivers)	3 months-5.5 years (8 age-specific questionnaires)	32 items 10-15 min.	Asks about specific behaviors (high frequency of positive items)	Original version revised in 2015 Official Danish version expected Research translations of 6 and 18 month questionnaires	Copyrighted by Brookes Publishing Dansk Psykologisk Forlag has an option on distributorship of Danish version of ASQ:SE-2 and expects to publish this in late 2017 Most likely price structure is fee/questionnaire	First edition (ASQ:SE) published in 2002, and revised in 2015 (ASQ:SE-2) Danish translation of ASQ:SE-2 will be published in 2017
BASC-2 (BASC-3)	Adaptive and problem behaviors and personality: - Activities of daily living (P) - Adaptability (T/P) - Aggression (T/P) - Anxiety (T/P/S) - Attention problems (T/P/S) - Atypicality (T/P/S) - Conduct problems (T/P) - Depression (T/P/S) - Functional communication (T/P) - Hyperactivity (T/P/S) - Leadership (T/P) - Learning problems (T) - Social skills (T/P) - Somatization (T/P/S) - Study skills (T) - Withdrawal (T/P) - Alcohol abuse (S) - Attitude to school (S) - Attitude to teachers (S) - Interpersonal relations (S)	4-point Likert scale from *never* to *almost always*	Parents Teachers Self-report from age 8 (interview form available from age 6)	2-25 years (2-3 age-specific questionnaires)	100-185 items 10-30 min.	Approx. 1/3 positive items	No official Danish version Research translations	Copyrighted by Pearson Assessment Manual and sample questionnaires: DKK 1500 DKK 17/additional questionnaire	First edition (BASC) published in 1992, and revised in 2004 (BASC-2) and 2015 (BASC-3)

Instrument	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. completion time	Positive/ negative items	Danish version and norms	Copyrights and approximate price	Development
	- Locus of control (S) - Relationship with parents (S) - School maladjustment (S) - Self-esteem (S) - Self-reliance (S) - Sensation seeking (S) - Sense of inadequacy (S) - Social stress (S)								
BASC-2 BESS	Behavioral and emotional strengths and weaknesses: - Externalized behavior problems - Internalized behavior problems - School problems - Adaptive skills	4-point Likert scale from *never* to *almost always*	Parents Teachers Self-report from age 9	3-18 (2 age-specific questionnaires)	20-30 items 5-10 min.	App. 1/3 positive items	No official Danish version Research translations	Copyrighted by Pearson Assessment Manual and sample questionnaires: DKK 1300 DKK 13/additional questionnaire	BASC-2 BESS published in 2007
BFI-10	Big Five personality dimensions: - Extraversion - Agreeableness - Conscientiousness - Neuroticism - Openness	5-point Likert scale from *disagree strongly* to *agree strongly*	Self-report	Developed for adults, but used for children from age 10 and up	10 items	Asks about specific personality traits	No official Danish version Research translation	Free	Original published in 2007
BFQ-C	Big Five personality dimensions: - Energy/Extraversion - Agreeableness - Conscientiousness - Emotional instability - Intellect/Openness	5-point Likert scale from *almost never* to *almost always*	Self-report	Developed for children aged 9-13 years	65 items	Asks about specific personality traits	No official Danish version No research translations identified	Free	Original published in 2005
BRIEF BRIEF-F	Executive function behaviors: - Inhibit	3-point Likert scale	Parents or caregivers Teachers	2-90 years	63-86 items 10-15 min.	Information not identified	Official Danish version	Copyrighted by Psychological Assessment Resources (PAR)	Originals published in 2000 (BRIEF), 2003 (BRIEF-V), 2004 (BRIEF-SR) and 2005 (BRIEF-V)

Instrument	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. completion time	Positive/ negative items	Danish version and norms	Copyrights and approximate price	Development
BRIEF-SR BRIEF-V	- Shift - Emotional control - Working memory - Plan/organize - Initiate* - Organization of materials* - Monitor* *Included from age 5		Self-report from age 11					Danish version distributed by Hogrefe Psykologisk Forlag Manual and sample questionnaires: DKK 500 DKK 15-19/additional questionnaire (must be purchased separately for BRIEF, -F, -SR and -V) Electronic scoring and registration: DKK 3700	Danish translations published in 2005 (BRIEF), 2006 (BRIEF-F), 2007 (BRIEF-SR) and 2008 (BRIEF-V)
CDI	CDI-I: - Words and gestures CDI-II: - Words and sentences CDI-III: - Language and use	Word checklists (yes/no)	Parents or caregivers	CDI-I: 8-20 months CDI-II: 16-36 months CDI-III: 36-48 months	20-40 min. + short forms with app. 100 items	Asks about specific skills	Official Danish version	Copyrighted by Brookes Publishing User guide and manual: DKK 350 DKK 6/record form Contact Dorthe Bleses for information about Danish version	Originals published in 2003 and revised in 2007. Danish translations published in 2006/2007
Conners CBRS	Behavior, emotions and academic problems: - Emotional distress - Defiant/aggressive behaviors - Academic difficulties - Hyperactivity/impulsivity - Separation fears - Perfectionistic and compulsive behaviors - Violence potential indicators - Physical symptoms *+DSM-5 symptom scales*	4-point Likert scale from *not at all/never* to *very much true/very frequently*	Parents Teachers Self-report from age 8	6-18 years	CBRS: 25 min. Clinical index: 10 min.	Information not identified	Official Danish version	Copyrighted by Pearson Assessment Danish version distributed by Hogrefe Psykologisk Forlag Manual and sample questionnaires: DKK 4100 DKK 14/additional questionnaire	Original published in 2008 Danish translation published in 2013

Instrument	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. completion time	Positive/ negative items	Danish version and norms	Copyrights and approximate price	Development
	Clinical index: - Disruptive behavior disorders - Mood disorders - Anxiety disorders - Learning and language disorders - ADHD								
DECA-I/T DECA-P2	Behavior related to social-emotional health and resilience: - Initiative (from 18 months) - Self-regulation - Attachment/relationships	5-point Likert scale from *never* to *very frequently*	Parents or caregivers	Birth-5 years (3 age-specific questionnaires)	33-38 items 3-10 min.	Asks about specific behaviors (high frequency of positive items)	No official Danish version Research translations identified	Copyrighted by Kaplan Early Learning Company Manual and sample questionnaires: DKK 1500 (must be purchased separately for –I/T and -P2) DKK 5/additional questionnaire	Originals published in 2007 (DECA-I/T) and 2012 (DECA-P2) Danish research translation conducted in 2015
DESSA (-mini)	Social-emotional competencies: - Self-awareness - Social-awareness - Self-management - Goal-directed behavior - Relationship skills - Personal responsibility - Decision making - Optimistic thinking	5-point Likert scale from *never* to *very frequently*	Teachers or after-school providers	6-14 years	72 items: 5-10 min. Short-form: 8 items 1-2 min.	Asks about specific behaviors (all positive)	No official Danish version No research translations identified	Copyrighted by Kaplan Early Learning Company Manual and sample questionnaires: DKK 615 DKK 5/additional questionnaire	Original published in 2009
DP-3	Developmental strengths and weaknesses: - Physical - Adaptive behavior - Social-emotional - Cognitive - Communication	Yes/no	Parents or caregivers	Birth-12 years, 11 months (4 age-specific questionnaires)	180 items 20-40 min.	Asks about specific skills/behaviors	Official Danish version	Copyrighted by Western Psychological Services Danish version distributed by Hogrefe Psykologisk Forlag Manual: DKK 500 DKK 20/additional questionnaire	Could not retrieve information about when original first edition was published. Second edition published in 1980 (DP-2) and revised in 2007 (DP-3) Danish translation of DP-3 published in 2014

Instrument	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. completion time	Positive/ negative items	Danish version and norms	Copyrights and approximate price	Development
HBSC	Health behavior and self-reported health: - Body image - Bullying and fighting - Eating behaviors - Health complaints - Injuries - Life satisfaction - Obesity - Oral health - Physical activity and sedentary behavior - Relationships: Family and peers - School environment - Self-rated health - Sexual behavior - Socioeconomic environment - Substance use - Weight reduction behavior	Varies between items	Self-report	11, 13 and 15 years	Approximately 100 items	Asks about specific behaviors and health issues	Official Danish version	World health Organization (WHO) cross-national study Includes 44 countries and regions across Europe and North America Conducted in Denmark every fourth year since 1984	First survey conducted in 1984, subsequently revised and conducted every fourth year (most recently in 2014)
KIDSCREEN	Quality of life: - Physical well-being - Psychological well-being - Moods and emotions - Self-perception - Autonomy - Parent relation and home - Financial resources - Peers and social support - School environment - Bullying	5-point Likert scale from *never/not at all* to *always/extremely*	Parents or caregivers Self-report	8-18 years	Items/min. 52/15-20 27/10-15 10/5	High frequency of positive items	Official Danish version	Copyrighted by the KIDSCREEN Group Manual: DKK 370 Non-commercial use of paper questionnaire free Contact license holder for computer-based data collection	Original published in 2006
MACI	Personality patterns, self-reported concerns and	True/false	Self-report	13-19 years	160 items 20-25 min.	High frequency of	No official Danish version	Copyrighted by Pearson Clinical	Original published in 1993 and revised in 2006

Instrument	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. completion time	Positive/ negative items	Danish version and norms	Copyrights and approximate price	Development
	clinical symptoms measure on 27 scales + modifying indices					negative items	Research translation	Manual: DKK 460 Manual and sample questionnaires etc.: DKK 3700 DKK 13/additional questionnaire	
M&MS	Mental health: - Emotional difficulties - Behavioral difficulties	3-point Likers scale: *never, sometimes, always*	Self-report	8-15 years	16 items	High frequency of negative items	No official Danish version No research translations identified	CAMHS Evidence Based Practice Unit (part of University College London and the Anna Freud Centre) Recently developed questionnaire - fee structure unknown	Original published in 2012
NIH Toolbox	Consists of a large number of scales in the following domains: - Cognitive function - Emotional function - Motor function - Sensory function	Varies for the different domains and batteries of questions	Varies for the different domains and batteries of questions	3-85 years (not all batteries of questions applicable for all age groups)	Varies for the different domains and batteries of questions	Varies for the different domains and batteries of questions	No official Danish version No research translations identified	Free (fees may apply for use of the NIH toolbox in projects not funded by NIH)	Original published in 2012
SEAM	Social-emotional behaviors: - Interactions with others - Emotions - Social emotional responses - Empathy - Attention and engagement - Independence - Self-image	4-point Likert scale from *not true* to *very true* +check box if concern for each item	Parents or caregivers	3-66 months (3 age-specific questionnaires)	35-41 items 15-30 min.	All items positive	Official Danish version expected Research translation	Copyrighted by Brookes Publishing Hogrefe has distributorship of Danish version and expects to publish this in late 2016 Most likely price structure is fee/questionnaire	Original published in 2014 Danish research translation conducted in 2014 Official Danish translation will be published in 2016

In-stru-ment	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. comple-tion time	Positive/ negative items	Danish version and norms	Copyrights and ap-proximate price	Development
	- Attention and activity level - Compliance - Adaptive skills								
SDQ	Behavioral and emotional difficulties: - Hyperactivity - Emotional symptoms - Conduct problems - Peer relationship prob-lems - Prosocial behavior	3-point Likert scale: *not true, some-what true, cer-tainly true* +indicate if difficulties af-fect everyday life	Parents Teachers Self-report from age 11	2-17 years (5 age- and respondent-specific question-naires)	25 items 5 min.	10 posi-tive/ 15 negative items	Official Danish version Danish norms for 5, 7, 10 and 12 year old children	Copyrighted by Robert Goodman/youthinmind Non-commercial use of paper questionnaire free Use of computer-based questionnaire: DKK 500 + DKK 1/questionnaire Contact license holder for approval of computer-based data collection	Original published in 1997 Official Danish translation published in 2002 and re-vised in 2014
SFI pi-lot	Well-being and teaching environment: - Interest in school - Cognitive skills - Non-cognitive skills - Peace and order in class-room - Psychosocial teaching en-vironment - Physical and mental well-being - General well-being in school - Physical teaching envi-ronment	Varies be-tween items	Self-report	9, 11, and 14 years	Full scale: 39 items Shortlist scale: 15 items	App. 1/2 positive items	Pilot version from 2014 and official version from 2015	Developed by SFI at the request of the Danish Ministry of Education Intended for use in na-tional well-being meas-urements in schools	Pilot study conducted in 2014 National surveys con-ducted annually for all school classes starting in 2015
SSRS/ SSIS-RS	Behavior and personality: - Social skills - Problem behaviors - Academic competence	4-point Likert scale from *never/not true* to *always/very true*	Parents Teachers Self-report from age 8	3-18 years	75-83 items 15-20 min.	Social skills: Items pos-itive Problems behavior: Items neg-ative	No official Danish version Research transla-tions of SSRS	Copyrighted by Pearson Assessment Manual and sample ques-tionnaires: DKK 1840 DKK 10/additional ques-tionnaire	First edition (SSRS) pub-lished in 1990 and re-vised in 2008 (SSIS-RS)

Instrument	Scales and subscales	Response categories	Reporters	Age range	Items/ approx. completion time	Positive/ negative items	Danish version and norms	Copyrights and approximate price	Development
TIPI	Big Five personality dimensions: - Extraversion - Agreeableness - Conscientiousness - Emotional stability - Openness to experiences	10-point Likert scale from *disagree strongly* to *agree strongly*	Self-report	Developed for adults, but used for children from age 13 and up	10 items	Asks about specific personality traits	No official Danish version Research translation	Free	Original published in 2003
WHO-5	Positive psychological well-being	6-point Likert scale from *not present* to *constantly present*	Self-report	Developed for adults, but used for children from age 9 and up	5 items 2-3 min.	All items positive	Official Danish version Danish norms for sample aged 15 and up	Free	Original published in 1998 Danish translation published in 1999

5 Discussion and recommendations

This section identifies and discusses a subset of instruments that are suitable for use as common outcome measures across the future studies to be conducted in the setting of TrygFonden's Centre for Child Research. We are fully aware that the assessment of the instruments, as well as the choice of which instruments to include in the battery, entails a large degree of subjectivity and includes several judgments and decisions that are subject to discussion. Section 5.1 describes and discusses the proposed battery of instruments. Subsequently, Section 5.2 accounts for instruments that were not included in the battery and why.

In general, when deciding on which instruments to use in a given study, it is important to consider whether the questionnaire suits the purpose of the investigation and whether the dimensions covered are relevant to the context as well as the price and availability of the questionnaire for the age group of interest. Moreover, the type of respondent should be taken into account. If the aim is to evaluate the effect of an intervention, or monitor the evolution of well-being over time, particular attention should be paid to whether the instrument is able to detect changes.

5.1 Proposed battery of instruments

At the time of writing, the scores on any of the instruments included in the proposed battery have not been converted into additional life value for use in economic evaluations. However, this is undoubtedly an area for future research. As stated in Section 3.1, it was decided to include measures of both subjective ('hedonic') and psychological ('eudaimonic') well-being, including social-emotional function, as core measures in the proposed battery of instruments. The core measures are intended for inclusion in most or all of the studies in TrygFonden's Centre for Child Research. In addition to the core measures, a selection of supplementary measures intended to cover further aspects of child well-being, including personality traits, executive function and the more comprehensive concept of psychomotor development, were included in the battery. The supplementary measures are merely intended as suggestions for project managers who are looking for instruments to measure the mentioned aspects of child well-being.

Figure 5.1 provides an overview of the various instruments selected as candidates to be included as common outcome measures across the studies conducted in the setting of TrygFonden's Centre for Child Research. The table describes the focus areas of the various instruments and the age range for which they are applicable. For all instruments that are answered by children or adolescents themselves, it should be kept in mind that even seemingly innocent questions can start a therapeutic process in the child. Hence, it is of crucial importance to ensure that an adult is present when the questionnaire is administered, so that the child can talk to somebody about any worries or thoughts that may arise.

Figure 5.1 Candidates for the proposed battery of instruments

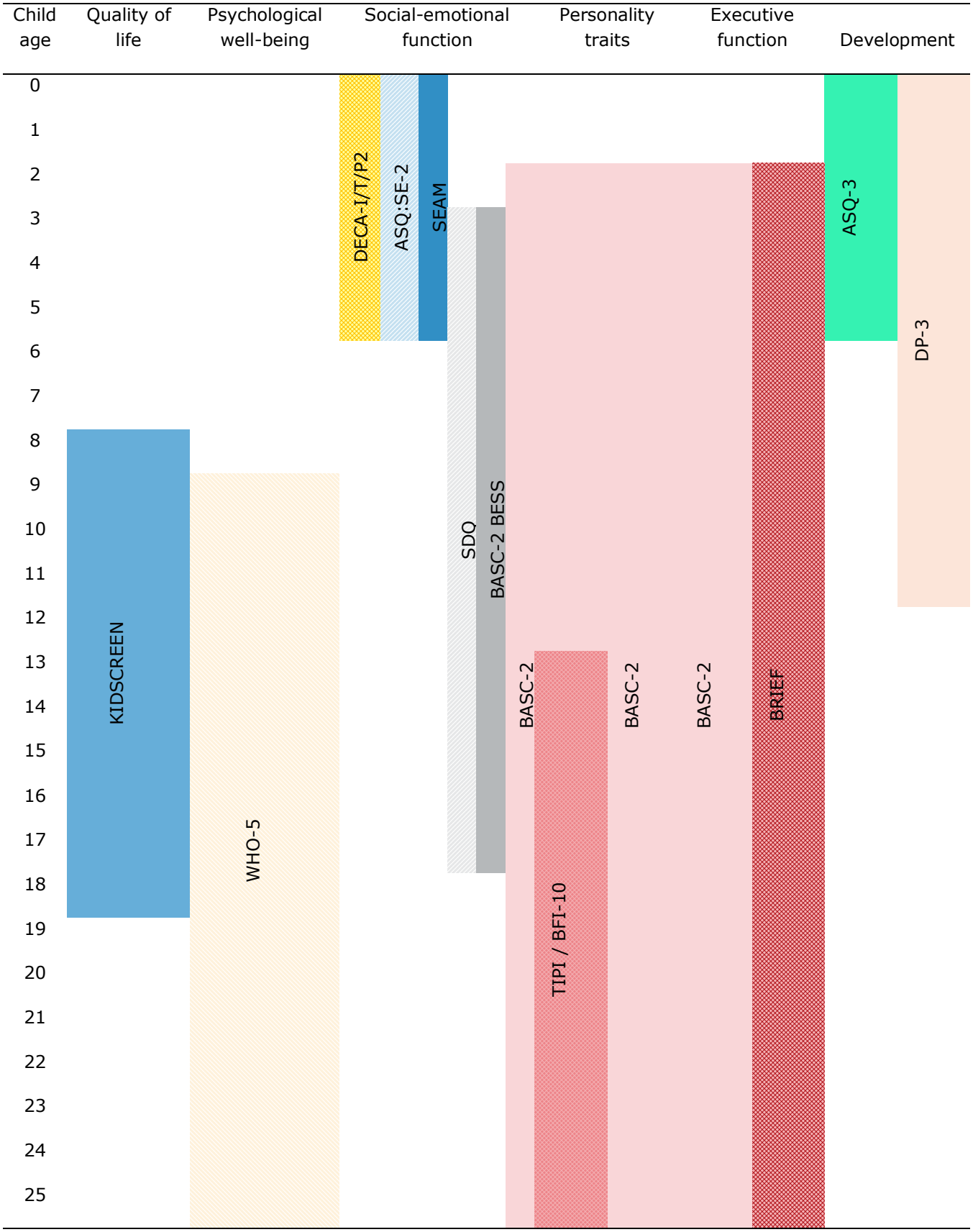


Figure 5.1 contains two candidate core measures of subjective (or 'hedonic') well-being: KIDSCREEN and WHO-5. Both of these instruments are sufficiently brief to allow for routine use across a wide range of studies, though their focus differs. WHO-5 focuses on positive psychological well-being. It is free and very quick to administer, and the author of the instrument is sympathetic towards the idea of creating a modified version that can be answered by the parents of the children who are too young to self-report. The responsiveness of the WHO-5 in controlled clinical trials and its applicability across study fields have recently been reviewed by Topp et al. (2015). A drawback of WHO-5 is that it was developed for use with adolescents and adults and has not been systematically validated for use with children. KIDSCREEN measures well-being defined as quality of life across several domains, including physical and psychological well-being, and well-being related to autonomy and parent relation, peers and school environment. KIDSCREEN comes in three different versions of varying length. Non-commercial use of paper questionnaire is free, while computer-based data collection may be subject to payment. However, the fact that KIDSCREEN covers several domains of children's lives opens up the possibility that it might be able to capture effects of many different types of interventions. On the other hand, there is a risk that many interventions will not impact any of the domains enough to have an effect on the score[9]. Hence, KIDSCREEN is proposed as an instrument to shed light on which groups of children the studies are dealing with, while it is uncertain whether it is able to capture effects of specific interventions.

In addition to the instruments shown in Figure 5.1, subjective (or 'hedonic') well-being can also be measured using life satisfaction scales (such as Cantril's ladder and Huebner's life satisfaction scale mentioned in Section 2) and measures of positive or negative affect, such as the Positive and Negative Affect Schedule (PANAS) for children and parents (Ebesutani et al. 2012). The HBSC survey includes a Danish example of such a scale, just as the NIH Toolbox contains several life satisfaction scales as well as measures of affect. This type of single question should also be kept in mind when considering measures of well-being.

Regarding social-emotional function, which is part of psychological (or 'eudaimonic') well-being, Figure 5.1 contains three candidate core measures targeted at infants and toddlers and two measures that are applicable from age three and up.

For the older children, we argue that SDQ is superior to BASC-2 BESS for practical reasons. Both questionnaires contain a mix of positive and negative items. SDQ is cheap, it is relatively quick to obtain permission to use it, and it is readily available in Danish, whereas BASC-2 BESS is copyrighted and distributed by the commercial publisher Pearson Assessment, and at the time of writing is only available in Danish in the form of an unofficial research translation. SDQ has been widely used by psychologists and economists both nationally and internationally (Heckman, Yi & Zhang 2013, Kristoffersen & Smith 2013, Niclasen et al. 2013, Niclasen et al. 2012) and thus seems a fairly safe choice. Kersten et al. (2016) have conducted a systematic review of the psychometric properties of the SDQ in children aged 3-5 years. Also, Goodman & Goodman (2011) found that mean symptom scores closely predicted the prevalence of clinician-rated mental health disorder in a sample of 18,415 British children aged 5-6 years. However, the ability of SDQ to capture changes may be argued to be questionable, and ceiling effects may occur, which makes it difficult to differentiate between children who have few difficulties. If for some reason project managers do not wish to use SDQ, BASC-2 BESS is a good alternative.

---

[9] It has been noted in the literature that there is currently a lack of evidence on whether the sort of well-being measures that covers a wide range of domains of children's lives are sensitive and reliable enough to be used as change measures in smaller samples of young people receiving project interventions (Rees et al. 2010).

Regarding the instruments targeted at infants and toddlers, it is more difficult to decide on a preferred instrument, as all three candidates involve uncertainty. The selected candidates for inclusion across studies, i.e. DECA, ASQ:SE and SEAM, are also emphasized as the preferred instruments to measure the well-being of 0-3 year old children in Pontoppidan & Niss (2014). DECA-I/T/P2 questionnaires are considered to be inferior to the alternatives, because there are no official Danish versions of the questionnaires. However, in comparison with the alternatives the 5-point response scale used in the DECA questionnaires is less likely to cause problems with limited variation in data than response scales with fewer categories. Moreover, the questions can be answered by both parents and caregivers. ASQ:SE is designed for identification of developmental delays in a normal population of children[10]. It includes 8 age-specific questionnaires, which were revised and published as ASQ:SE-2 in 2015. Dansk Psykologisk Forlag has an option on distributorship of a Danish version and expects to publish this in late 2017. The interested reader should contact the publishers for up-to-date information. SEAM has been designed by the authors of ASQ:SE, and it only contains positive items. Both instruments are universal in the sense that the same questionnaire can be completed by both parents and primary caregivers. SEAM includes 3 age-specific questionnaires, which have been translated into Danish and are currently used as outcome measure in the large-scale research project 'Fremtidens Dagtilbud' and for some of the interventions evaluated as part of 'Tidlig Indsats – Livslang Effekt'. In the SEAM questionnaires, each subsection has a title, and examples are given in relation to each question. In some respects, ASQ:SE and SEAM thus have different merits. It is expected that Dansk Psykologisk Forlag will publish ASQ:SE-2 in late 2017. Consequently, SEAM might be better suited for studies planned for the near future. However, the merits of SEAM in terms of the responsiveness of the instrument, i.e. its ability to measure progression in populations of normally developing children, hinge on the results of the ongoing large-scale projects (i.e. 'Fremtidens Dagtilbud' and 'Tidlig Indsats – Livslang Effekt'), in which it is currently being used.

Considering the supplementary measures, TIPI and BFI-10 both appear to be suitable short measure of the Big Five personality traits in research settings, where participation time is limited. Although the psychometric properties of TIPI should be noted along with the fact that both TIPI and BFI-10 were developed for use with adults and have not been validated systematically for use with children. The other identified personality measure, BFQ-C, was disregarded as a common measure due to its length (65 items) and the fact that it is not available in Danish. However, unlike the alternatives BFQ-C was developed specifically for children, and this instrument may thus be of relevance for studies with particular focus on the Big Five personality traits.

BASC-2 is the most comprehensive of the identified instruments, including scales such as social-emotional function and personality as well as executive function, self-esteem and locus of control, to mention but a few. BRIEF comes in different versions that target different age groups and may be used to measure executive function. ASQ-3 and DP-3 are examples of instruments that can be used to measure overall development for infants and toddlers, including development of cognition, executive function and social-emotional function. Of the supplementary instruments, official Danish versions of BRIEF and DP-3 are distributed by Hogrefe Psykologisk Forlag. However, it should be noted that both of these instruments have response scales with few categories (yes/no and a 3-point scale), which are more likely to cause problems with limited variation in data than response scales with more categories. Dansk Psykologisk Forlag expects to publish a

[10] Since the social-emotional development of most children is not delayed, it may be necessary to lag the questionnaires with respect to age (i.e. administer the questionnaires to children slightly younger than the questionnaire is intended for), in order to create some variation in a group of well-functioning children. However, this approach may imply that the lagged questions are seen as irrelevant or inappropriate.

Danish version of ASQ-3 in late 2016, while TIPI, BFI-10 and BASC-2 are only available in Danish in the form of an unofficial research translation at the time of writing.

Though the questionnaire developed in the SFI pilot study to measure the well-being and teaching environment of Danish Schoolchildren is not included in the battery, this instrument is expected to provide a valuable source of data for future research in the well-being of children and adolescents, as it is implemented in a large number of Danish schools. However, it is not included in the battery because, though some of the psychometric properties of the questionnaire are assessed in the pilot study, the questionnaire is not available in other languages than Danish.[11]

5.2 Rejected instruments

Table 5.1 provides an overview of which instruments were excluded from the group of candidates to be considered as common measures of well-being and the main reasons why.

Table 5.1 Instruments that were disregarded as common outcome measures and the main reasons for this

Acronym	
ABAS-II	Main focus on adaptive behavior and skills
ASEBA: CBCL, TRF, YSR	CBCL contains questions that may be perceived as insulting or offensive by parents The youth self-report (YSR) form includes questions that are particularly likely to start a therapeutic process[12] (diagnostic focus)
BFQ-C	Currently no official Danish version and no unofficial translations identified Length (65 items)
CDI	Main focus on language development
Conners CBRS	Includes questions that are particularly likely to start a therapeutic process (diagnostic focus)
DESSA	Currently no official Danish version and no unofficial translations identified
HBSC	Main focus on health (one item might be relevant – see discussion in Section 0)
MACI	Includes questions that are particularly likely to start a therapeutic process (diagnostic focus) Length
M&MS	Recently developed/pilot questionnaire
NIH Toolbox	Currently no official Danish version and no unofficial translations identified
SFI pilot	Questionnaire only available in Danish
SSIS-RS	Currently no official Danish version, but unofficial translations of SSRS identified

The main reasons (stated in Table 5.1) why some instruments were judged not to be suitable for use as common measures of well-being across a wide range of studies can be grouped into four categories. Three instruments (ASEBA, Conners CBRS and MACI) were excluded for ethical reasons, i.e. because the youth self-report versions were judged to be unsuitable for use in a context where the child does not necessarily have an adult available for support, and because it contains questions that may be perceived as insulting or offensive. These questionnaires were developed

[11] The fact that the questionnaire is only available in Danish is problematic, given that most of the research in TrygFonden's Centre for Child Research is conducted also with publication in international journals in mind.

[12] Including, for instance, "I deliberately try to hurt or kill myself".

with a more diagnostic focus and are not obvious candidates for epidemiological studies of a normal population of children.

Another reason why some instruments (ABAS-II, HBSC[13] and CDI) were not considered as candidates for the battery is that their main focus was deemed to be on the periphery of the focus of this note, as laid out in Sections 1.1.1 and 3.1. While judged not to be obvious candidates for inclusion in most or all of the studies in TrygFonden's Centre for Child Research, the instruments that were rejected on these grounds are, however, likely to be relevant as project-specific outcomes. In particular, MacArthur-Bates CDI, which exists in an official Danish version and can be completed by parents or caregivers, is an obvious candidate in studies that focus on language development.

Finally, it was decided to disregard instruments of which neither an official version nor an unofficial/research translation were identified and that consider dimensions and age groups for which alternative options are available in Danish. This decision is based on time considerations and resource constraints. However, this is definitely subject to discussion, as the instruments that were rejected on these grounds may have qualities that have not been recognized by the authors of this report[14].

---

[13] An exception to this is that one of the items in the HBSC questionnaire may be used to shed light on subjective well-being, as mentioned in Section 5.

[14] However, it should be noted that making a good translation of an instrument is not merely a linguistic exercise, but typically includes drawing upon the knowledge of professionals within the field as well as potential reporters, for example through focus group interviews. The WHO has a set of guidelines for translation of questionnaires, see http://www.who.int/substance_abuse/research_tools/translation/en/.

Appendix A: Brief introduction to psychometric properties

In the following, we provide a brief introduction to the psychometric properties that can be applied to assess the performance of measurement instruments on various dimensions (for readers who are not familiar with this topic). Further information about the application of psychometric properties is provided by, for instance, the European Federation of Psychologists' Associations (EFPA), who have developed a set of test review criteria (Evers et al. 2013).[15]

Validity

Validity concerns whether the instrument measures what it is intended to measure and how well it does so, and is thus of crucial importance to the quality of the instrument. Validity can be conceptualized as being content, criterion or construct-related.

- *Content validity* examines whether the overall domain measured by the instrument is represented by the included items. Content validity is primarily assessed using qualitative methods, and it may be subdivided into face validity and logical validity. Face validity concerns whether respondents find the items relevant and the instrument suitable for their purpose on the face of it. Logical validity is established by a professional user.
- *Criterion validity* concerns whether there is correspondence between the measurement method and another method or criterion (so-called gold standard). Criterion validity is primarily assessed using quantitative methods, and it can be subdivided into concurrent validity and predictive validity. Concurrent validity measures the degree to which the scores on the instrument correlates with scores obtained using another method, a relevant outcome, or the status or group membership of the individual, observed at the same point in time. Concurrent validity can be further divided into convergent validity (correspondence with other measures of the same characteristics) and divergent validity (lack of correspondence with irrelevant measures and criteria). Predictive validity is established by comparing the scores on the instrument with a relevant outcome, or the status or group membership of the individual, observed at a later point in time.
- *Construct validity* concerns the extent to which the instrument measures the theoretical construct of interest. It is investigated by assessing whether the score on the instrument behaves in a way that corresponds to the theoretical or conceptual expectations. Construct validity is established using quantitative methods, such as item response theory and Rasch analyses.

[15] The European Federation of Psychologists' Associations (EFPA) Test Review Criteria were largely modelled on the form and content of the British Psychological Society's (BPS) test review criteria and criteria developed by the Dutch Committee on Tests and Testing (COTAN) of the Dutch Association of Psychologists (NIP). EFPA is grateful to the BPS and the NIP for permission to build on their criteria in developing the European model. All intellectual property rights in the original BPS and NIP criteria are acknowledged and remain with those bodies.

Reliability

Reliability concerns the measurement precision and consistency of the instrument. Reliability can be assessed in different ways, depending on what the instrument measures.

- *Internal reliability (or internal consistency)* refers to the extent to which the items on a (sub)scale measure the same underlying construct. This can be assessed using Cronbach's alpha, which measures the inter-correlation between the items of a (sub)scale. Alpha values above 0.7 are usually considered satisfactory. Recent studies have used factor analysis to assess internal reliability and investigate the factor structure (see e.g. Niclasen et al. (2013, 2012), Dowdy et al. (2011a, 2011b), Dever et al. (2012) and Wiesner & Schanding (2013)).
- *Test-retest reliability* is assessed by the correlation between scores obtained for the same child on two separate occasions and is highly relevant for instruments that measure lasting psychological traits.
- *Inter-rater reliability* is assessed by the correlation between scores obtained for the same child at the same time by two different raters. Children's behaviour is generally more variable than adults' and may thus to a larger extent be affected by the company. Low correlations between the ratings from different adult raters are therefore common. While high degrees of disagreement are generally considered a weakness of the instrument, it should also be acknowledged that each rater is a potentially valid source of information and contributes to form an overall picture of the child (Trillingsgaard & Damm 2012).

Sensitivity and specificity

Sensitivity concerns the ability of the instrument to identify all children in a given group, e.g. children with a specific condition. High sensitivity implies that the instrument is able to identify all children in the group. However, it may also falsely classify some children as belonging to the group. Specificity concerns the ability of the instrument to identify the group of children with a specific condition, but avoid falsely including children without the condition in the group.

Responsivity

Responsivity concerns the ability of the instrument to detect changes in the relevant outcome measured over time, including changes (typically progression) of a size that can be expected as a consequence of programs or interventions. There are various quantitative approaches to assessing sensitivity to change, such as calculation of the standardized response mean and measurement error, effect size, minimal important difference and smallest detectable change.

Appendix B: Sample questionnaires and sources of further information

Links were updated on April 13, 2016.

ABAS-II (Adaptive Behavior Assessment System – Second Edition)

Sample questionnaire(s): No sample questionnaires identified on the Internet (using google.com)

References: Trillingsgaard & Damm (2012), pp. 168-172

ASEBA: CBCL, TRF, YSR (Achenbach System of Empirically Based Assessment: Child Behavior Checklist, Teacher Report Form, Youth Self-Report)

Sample questionnaire(s): http://www.aseba.org/forms.html

Danish publisher: http://www.psykiatrienisyddanmark.dk/wm287192

Reference(s): http://www.aseba.org/

Trillingsgaard & Damm (2012), pp. 168-172

ASQ-3 (Ages & Stages Questionnaires, Third Edition)

Sample questionnaire(s): http://agesandstages.com/pdfs/asq3_english_16_month_sample.pdf

Reference(s): http://agesandstages.com/

Klamer et al. (2005)

Østergaard et al. (2012)

ASQ:SE(-2) (Ages & Stages Questionnaires: Social-Emotional)

Sample questionnaires: http://agesandstages.com/wp-content/uploads/2015/04/ASQSE2-24-Month-Questionnaire.pdf

Danish publisher: Maiken Pontoppidan has used a Danish research translation in her PhD thesis

References: http://agesandstages.com/

Williams (2008)

BASC-2 (Behavior Assessment System for Children, Second Edition)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

Danish publisher: Contact Helene Bie Lilleør or Nina Madsen Sjö for information about Danish research translation

References: Flanagan (1995)

http://www.pearsonclinical.co.uk/Psychology/ChildMentalHealth/ChildADDADHDBehaviour/BehaviorAssessmentSystemforChildrenSecondEdition(BASC-2)/BehaviorAssessmentSystemforChildrenSecondEdition(BASC-2).aspx

BASC-2 BESS (BASC-2 Behavioral and Emotional Screening System)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

Danish publisher: Contact Helene Bie Lilleør or Nina Madsen Sjö for information about Danish research translation

References:

http://www.pearsonclinical.co.uk/Psychology/ChildMentalHealth/ChildADDADHDBehaviour/BASC-2BESS(BehavioralandEmotionalScreeningSystem)/BASC-2BESS(BehavioralandEmotionalScreeningSystem).aspx

BFI-10 (Big Five Inventory – 10-item version)

Sample questionnaires: http://www.westmont.edu/_academics/departments/psychology/documents/Rammstedt_and_John.pdf

Danish publisher: Contact Morten Hesse for information about Danish research translation

References: Rammstedt & John (2007)

BFQ-C (Big Five Questionnaire for Children)

Sample questionnaires: http://repository.asu.edu/attachments/93558/content/tmp/package-it1swu/gaio_asu_0010n_11584.pdf

References: Muris et al. (2005)

Barbaranelli et al. (2008)

BRIEF-F/SR/V (Behavior Rating Inventory of Executive Function)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

Danish publisher:

BRIEF: https://www.hogrefe.dk/shop/behaviour-rating-inventory-of-executive-function.html

BRIEF-F: https://www.hogrefe.dk/shop/behaviour-rating-inventory-of-executive-function-for-skoleborn.html

BRIEF-SR: https://www.hogrefe.dk/shop/behavior-rating-inventory-of-executive-function-selvrapportering.html

BRIEF-V: https://www.hogrefe.dk/shop/behaviour-rating-inventory-of-executive-function-vok-sne.html

BSMB (Bedre Sundhed for Mor og Barn / Danish National Birth Cohort)

Sample questionnaires: http://www.ssi.dk/Forskning/Forskningsomraader/Epidemiologi/BSMB/onsker%20du%20at%20forske/De%204%20forste%20interviews.aspx

Danish publisher: http://www.ssi.dk/English/RandD/Research%20areas/Epidemiology/DNBC/

CDI (MacArthur-Bates Communicative Development Inventory)

Sample questionnaires: http://www.uh.edu/class/psychology/dev-psych/_docs/MCDI-ShortVersion.pdf

Danish publisher: http://mb-cdi.stanford.edu/adaptations.html

References: http://mb-cdi.stanford.edu/

Conners CBRS (Conners Comprehensive Behavior Rating Scales)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

Danish publisher:

https://www.hogrefe.dk/shop/conners-comprehensive-behavior-rating-scale.html

DECA-I/T/P2 (Devereux Early Childhood Assessment for Infants/Toddlers/Preschoolers)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

References: http://www.centerforresilientchildren.org/home/about-us/summary-technical-information-assessment-tools/

Danish publisher: Contact Devereux Center for Resilient Children for information about Danish research translation

DESSA (Devereux Student Strength Assessment)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

References: http://www.centerforresilientchildren.org/home/about-us/summary-technical-information-assessment-tools/

DP-3 (Developmental Profile – Third Edition)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

Danish publisher: https://www.hogrefe.dk/shop/developmental-profile-3.html

HBSC (Health Behavior in School-aged Children / Skolebørnsundersøgelsen)

Sample questionnaires: No questionnaire identified on the Internet, but reports can give an impression of the types of questions asked

Danish publisher: http://www.hbsc.dk/

References: Danish report: http://www.hbsc.dk/rapport.php?file=HBSC-Rapport-2010.pdf

International report: http://www.euro.who.int/en/countries/armenia/publications/social-determinants-of-health-and-well-being-among-young-people.-health-behaviour-in-school-aged-children-hbsc-study

KIDSCREEN (KIDSCREEN Health Questionnaires for Children and Young People)

Sample questionnaires: http://www.kidscreen.org/english/questionnaires/example/

References: Ravens-Sieberer et al. (2013)

Erhart et al. (2009)

Ravens-Sieberer et al. (2005)

MACI (Millon Adolescent Clinical Inventory)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

Danish publisher: Mickey Toftkjær Kongerslev has used a Danish research translation in his PhD thesis http://www.regionsjaelland.dk/sundhed/geo/psykiatrien/om_psykiatrien/psykiatrisk-forskningsenhed/phd-studium/Documents/Personality-disorder-in-incarcerated_Mickey-Kongerslev.pdf)

References: http://www.millon.net/instruments/MACI.htm

M&MS (Me and My School)

Sample questionnaires:
http://www.google.dk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=14&sqi=2&ved=0CF8QFjAN&url=http%3A%2F%2Fwww.corc.uk.net%2Fwp-content%2Fuploads%2F2014%2F04%2FMMS-Questionnaire.pdf&ei=NO2TVIP6BqjCywOTto-CYBQ&usg=AFQjCNGiN6v9Ne0s_TdFiz47oKwf-t150A

References: Deighton et al. (2012)

Patalay et al. (2014)

NIH Toolbox (National Institutes of Health Toolbox for the Assessment of Neurological and Behavioral Function)

Sample questionnaires: http://www.nihtoolbox.org/WhatAndWhy/Emotion/emotionmeasures/Pages/default.aspx (emotion)

http://www.nihtoolbox.org/WhatAndWhy/Emotion/emotionmeasures/Pages/default.aspx (cognition)

SEAM (Social Emotional Assessment/Evaluation Measure)

Sample questionnaires:

http://www.brookespublishing.com/resource-center/screening-and-assessment/seam/

Danish publisher: Contact Nina Madsen Sjö for information about Danish research translation. Contact Hogrefe Psykologisk Forlag for information about forthcoming official Danish version.

References: Squires et al. (2013)

http://agesandstages.com/products-services/seam/

SDQ (Strengths and Difficulties Questionnaire)

Sample questionnaires: http://www.sdqinfo.com/py/sdqinfo/b3.py?language=Danish

Danish publisher: http://sdq-dawba.dk/

References: Goodman (1999); Goodman et al. (2010); Obel & Dalsgaard (2003); Niclasen et al. (2012); Niclasen et al. (2013); Kersten et al. (2016); Goodman & Goodman (2011)

SFI pilot (The well-being and teaching environment of Danish schoolchildren / Trivselsmålinger i folkeskolen))

Sample questionnaires and references: https://pure.sfi.dk/ws/files/239924/1424_Trivselsmaalinger_i_folkeskolen.pdf

SSRS/SSIS-RS (Social Skills Rating System / Social Skills Improvement System-Rating Scales)

Sample questionnaires: No sample questionnaires identified on the Internet (using google.com)

Danish publisher: Contact Rambøll for information about Danish research translation

References: Dyssegaard et al. (2013)

http://www.pearsonclinical.com/psychology/products/100000322/social-skills-improvement-system-ssis-rating-scales.html

Termometeret (Termometeret, Dansk Center for Undervisningsmiljø)

References: http://dcum.dk/termometeret

TIPI (Ten Item Personality Measure)

Sample questionnaires: http://www.personality-arp.org/html/newsletter07/docs/14_teaching_self_observer.pdf

Danish publisher: Contact Morten Hesse for information about Danish research translation

http://www.google.dk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CB8QFjAA&url=http%3A%2F%2Fiptp.dk%2Ffile_download%2F17%2Ftestkatalog-personlighed-og-psykopatologi.pdf&ei=PBKUVPG-IoLVau2wgaAH&usg=AFQjCNFK82d3qU645DnH1ZN7uaoPtakbAg&bvm=bv.82001339,d.bGQ

References: Gosling et al. (2003); Veselska et al. (2009)

WHO-5 (WHO-five Well-being Index)

Sample questionnaires: https://www.psykiatri-regionh.dk/who-5/Documents/WHO5_Danish.pdf

Danish publisher: www.who-5.org

References: Allgaier et al. (2012); Bech et al. (2003); Bech (2004); De Wit et al. (2007); Folker & Jensen (2001), Topp et al. (2015)

References

Allgaier, A.K., Pietsch, K., Frühe, B., Prast, E., Sigl-Glöckner, J. & Schulte-Körne, G. 2012, "Depression in pediatric care: is the WHO-Five Well-Being Index a valid screening instrument for children and adolescents?", *General Hospital Psychiatry,* vol. 34, pp. 234-241.

Almlund, M., Duckworth, A., Heckman, J.J. & Kautz, T. 2011, "Personality Psychology and Economics" in *Handbook in Economics of Education, vol. 4*, eds. E.A. Hanushek, S. Machin & L. Woessmann, North-Holland, Elsevier Science, , pp. 1-181.

Barbaranelli, C., Fida, R., Paciello, M., Giunta, L.D. & Caprara, G.V. 2008, "Assessing personality in early adolescence through self-report and other-ratings a multitrait-multimethod analysis of the BFQ-C", *Personality and Individual Differences,* vol. 44, no. 4, pp. 876-886.

Bech, P. 2004, "Measuring the Dimension of Psychological General Well-Being by the WHO-5", *QoL Newsletter,* vol. 32, pp. 15-16.

Bech, P., Olsen, L.R., Kjoller, M. & Rasmussen, N.K. 2003, "Measuring well-being rather than the absence of distress symptoms: a comparison of the SF-36 Mental Health subscale and the WHO-Five Well-Being Scale", *International Journal of Methods in Psychiatric Research,* vol. 12, no. 2, pp. 85-91.

Ben-Arieh, A. 2006, *Measuring and monitoring the well-being of young children around the world. (Paper commissioned for the EFA Global Monitoring Report 2007, Strong foundations: early childhood care and education),* UNESCO, [S.l.].

Ben-Arieh, A. & Frønes, I. 2011, "Taxonomy for child well-being indicators: A framework for the analysis of the well-being of children", *Childhood,* vol. 18, no. 4, pp. 460-476.

Bertrand, M. & Mullainathan, S. 2001, "Do People Mean What They Say? Implications for Subjective Survey Data", *American Economic Review,* vol. 91, no. 2, pp. 67-72.

Blanchflower, D. & Oswald, A. 2004, "Well-being over time in Britain and the USA", *Journal of Public Economics,* vol. 88, no. 7-8, pp. 1359-1386.

Brauner, J., Olsen, P.S. & Egelund, T. 2011, *Muligheder for dokumentation af anbringelser. En gennemgang af målemetoder (Rapport 11:49),* SFI - Det Nationale Forskningscenter for Velfærd, København.

CAMHS, Evidence Based Practice Unit 2008, *Mental Health Outcome Measures for Children and Young People (Booklet),* UCL & the Anna Freud Centre, London.

Cantril, H. 1965, *The Pattern of Human Concerns,* Rutgers University Press, New Brunswick, NJ.

CASEL 2013, *, Social and Emotional Learning Core Competencies* [Homepage of Collaborative for Academic, Social, and Emotional Learning (CASEL)]. Available: http://www.casel.org/social-and-emotional-learning/core-competencies [2013, 02/25].

Conti, G. & Heckman, J.J. 2012, *The Economics of Child Well-Being (IZA Discussion Paper No. 6930),* IZA, Bonn.

Cunha, F., Heckman, J. & Schennach, S. 2010, "Estimating the technology of cognitive and noncognitive skill formation", *Econometrica,* vol. 78, no. 3, pp. 883-931.

de Wit, M., Pouwer, F., Gemke, R.J.B.J., Delemarre-van de Waal, H. & Snoek, F.J. 2007, "Validation of the WHO-5 Well-Being Index in Adolescents With Type 1 Diabetes", *Diabetes Care,* vol. 30, pp. 2003-2006.

Deighton, J., Tymms, P., Vostanis, P., Belsky, J., Fonagy, P., Brown, A., Martin, A., Patalay, P. & Wolpert, M. 2012, "The Development of a School-Based Measure of Child Mental Health", *Journal of Psychoeducational Assessment,* vol. 31, no. 3, pp. 247-257.

Dever, B.V., Mays, K.L., Kamphaus, R.W. & Dowdy, E. 2012, "The Factor Structure of the BASC-2 Behavioral and Emotional Screening System Teacher Form, Child/Adolescent", *Journal of Psychoeducational Assessment,* vol. 30, no. 5, pp. 488-495.

Diamond, A. 2013, "Executive functions", *Annual Review of Psychology,* vol. 64, pp. 135-168.

Dowdy, E., Chin, J.K., Twyford, J.M. & Dever, B.V. 2011a, "A factor analytic investigation of the BASC-2 Behavioral and Emotional Screening System Parent Form: Psychometric properties, practical implications, and future directions", *Journal of School Psychology,* vol. 49, pp. 265-280.

Dowdy, E., Twyford, J.M., Chin, J.K. & DiStefano, C.A. 2011b, "Factor Structure of the BASC-2 Behavioral and Emotional Screening System Student Form", *Psychological Assessment,* vol. 23, no. 2, pp. 379-387.

Dyssegaard, C.B., Egeberg, J.d.H. & Steenberg, K. 2013, *Skoleparathed. Systematisk forskningskortlægning (Clearinghouse - forskningsserien 2013 nummer 18),* Dansk Clearinghouse for Uddannelsesforskning, Aarhus Universitet, København.

Ebesutani, C., Regan, J., Smith, A., Reise, S., Higa-McMillan, C. & Chorpita, B.F. 2012, "The 10-Item Positive and Negative Affect Schedule for Children, Child and Parent Shortened Versions: Application of Item Response Theory for More Efficient Assessment", *Journal of Psychopathotology and Behavioral Assessment,* vol. 34, pp. 191-203.

Erhart, M., Ottova, V., Gaspar, T., Jericek, H., Schnohr, C., Alikasifoglu, M., Morgan, A., Ravens-Sieberer, U. & the HBSC Positive Health Focus Group 2009, "Measuring mental health and well-being of school-children in 15 European countries using the KIDSCREEN-10 Index", *International Journal of Public Health,* vol. 54, pp. S160-S166.

Evers, A., Hagemeister, C., Høstmælingen, A., Lindley, P., Muñiz, J. & Sjöberg, A. 2013, *EFPA review model for the description and evaluation of psychological and educational tests. Test Review Form - Version 4.2.6 (EFPA Board of Assessment Document 110c),* European Federation of Psychologists' Associations, [S.l.].

Flanagan, R. 1995, "A Review of the Behavior Assessment System for Children (BASC): Assessment Consistent With the Requirements of the Individuals With Disabilities Education Act (IDEA)", *Journal of School Psychology,* vol. 33, no. 2, pp. 177-186.

Folker, H. & Jensen, B.M. 2001, "Undersøgelse af udvalgte metoder til selvvurdering af helbred, livskvalitet og tilfredshed med behandling", *Ugeskrift for Læger,* vol. 163, no. 24, pp. 3347-3352.

Frey, B. & Stutzer, A. 2002, "What Can Economists Learn from Happiness Research?", *Journal of Economic Literature,* vol. 40, no. 2, pp. 402-435.

Goodman, A. & Goodman, R. 2011, "Population mean scores predict child mental disorder rates: validating SDQ prevalence estimators in Britain", *The Journal of Child Psychology and Psychiatry,* vol. 52, no. 1, pp. 100-108.

Goodman, A., Lamping, D.L. & Ploubidis, G.B. 2010, "When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British parents, teachers and children", *Journal of Abnormal Child Psychology,* vol. 38, no. 8, pp. 1179-1191.

Goodman, R. 1999, "The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden", *Journal of Child Psychology and Psychiatry,* vol. 40, no. 5, pp. 791-799.

Gosling, S.D., Rentfrow, P.J. & Swann Jr., W.B. 2003, "A very brief measure of the Big-Five personality domains", *Journal of Research in Personality,* vol. 37, pp. 504-528.

Haggerty, K., Elgin, J. & Woolley, A. 2011, *Social-Emotional Learning Assessment Measures for Middle School Youth (report),* Social Development Research Group, University of Washington. Commissioned by the Raikes Foundation, Washington.

Heckman, J.J. 2007, "The Economics, Technology and Neuroscience of Human Capability Formation", *Proceedings of the National Academy of Sciences,* vol. 104, no. 3, pp. 13250-13255.

Heckman, J.J., Yi, J. & Zhang, J. 2013, *Early Health Shocks, Intrahousehold Resource Allocation, and Child Human Capital (Working paper),* [University of Chicago & Chinese University of Hong Kong], [S.l.].

Hogrefe Psykologisk Forlag 2013, *Testkatalog. Klinisk og erhverv,* Hogrefe Psykologisk Forlag, Virum.

Huebner, E.S. 1991, "Initial development of the Students' Life Satisfaction Scale", *School Psychology International,* vol. 12, pp. 231-243.

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J. & Wolpert, M. 2011, "Measures of Social and Emotional Skills for Children and Young People: A Systematic Review", *Educational and Psychological Measurement,* vol. 71, no. 4, pp. 617-637.

Judge, T.A., Erez, A., Bono, J.E. & Thoresen, C.J. 2002, "Are Measures of Self-Esteem, Neuroticism, Locus of Control, and Generalized Self-Efficacy Indicators of a Common Core Construct?", *Journal of Personality and Social Psychology,* vol. 83, no. 3, pp. 693-710.

Karoly, L. 2012, "Toward Standardization of Benefit-Cost Analysis of Early Childhood Interventions", *Journal of Benefit-Cost Analysis,* vol. 3, no. 1, pp. Article 4.

Keilow, M., Holm, A., Bagger, S. & Henze-Pedersen, S. 2014, *Udvikling af trivselsmålinger i folkeskolen. En pilotundersøgelse (Rapport 14:24),* SFI - Det Nationale Forskningscenter for Velfærd, København.

Kersten, P., Czuba, K., McPherson, K., Dudley, M., Elder, H., Tauroa, R. & Vandal, A. 2016, "A systematic review of evidence for the psychometric properties of the Strengths and Difficulties Questionnaire", *International Journal of Behavioral Development,* vol. 40, no. 1, pp. 64-75.

Klamer, A., Lando, A., Pinborg, A. & Greisen, G. 2005, "Ages and Stages Questionnaire used to measure cognitive deficit in children born extremely preterm", *Acta Pædiatrica,* vol. 94, pp. 1327-1329.

Kristoffersen, J.H.G. & Smith, N. 2013, *Gender Differences in the Effects of Behavioral Problems on School Outcomes (IZA Discussion paper No. 7410),* IZA, Bonn.

Krueger, A. & Schkade, D. 2007, *The Reliability of Subjective Well-Being Measures (NBER Working Paper No. 13027),* NBER, Cambridge, MA.

Lopes, P.N., Brackett, M.A., Nezlek, J.B., Schütz, A., Sellin, I. & Salovey, P. 2004, "Emotional Intelligence and Social Interaction", *Personality and Social Psychology Bulletin,* vol. 30, no. 8, pp. 1018-1034.

McCrae, R.R. & Costa, J.P.T. 2008, "The Five-Factor Theory of Personality" in *Handbook of Personality: Theory and Research*, eds. O.P. John, R.W. Robins & L.A. Pervin, The Guilford Press, New York, pp. 159-181.

Metodecentret 2013, *Resultatdokumentation. Inspirationskatalog med måleredskaber til det specialiserede socialområde (Rapport),* 2nd edn, Metodecentret, Center for Innovation og Metodeudvikling, Aarhus.

Muris, P., Meesters, C. & Diederen, R. 2005, "Psychometric properties of the Big Five Questionnaire for Children (BFQ-C) in a Dutch sample of young adolescents", *Personality and Individual Differences,* vol. 38, no. 8, pp. 1757-1769.

Niclasen, J., Skovgaard, A.M., Andersen, A.M., Sømhovd, M.J. & Obel, C. 2013, "A Confirmatory Approach to Examining the Factor Structure of the Strengths and Difficulties Questionnaire (SDQ): A Large Scale Cohort Study", *Journal of Abnormal Child Psychology,* vol. 41, no. 3, pp. 355-365.

Niclasen, J., Teasdale, T.W., Andersen, A.M., Skovgaard, A.M., Elberling, H. & Obel, C. 2012, "Psychometric Properties of the Danish Strength and Difficulties Questionnaire: The SDQ Assessed for More than 70,000 Raters in Four Different Cohorts", *PLoS ONE,* vol. 7, no. 2, pp. e32025.

Obel, C., Dalsgaard, S., Stax, H.P. & Bilenberg, N. 2003, "Spørgeskema om barnets styrker og vanskeligheder (SDQ-Dan)", *Ugeskrift for Læger,* vol. 65, no. 5, pp. 462-465.

OECD 2009, "Comparative Child Well-being across the OECD" in *Doing Better for Children* OECD, Paris, pp. 21-63.

Patalay, P., Deighton, J., Fonagy, P., Vostanis, P. & Wolpert, M. 2014, "Clinical validity of the Me and My School questionnaire: a self-report mental health measure for children and adolescents", *Child and Adolescent Psychiatry and Mental Health,* vol. 8, no. 17.

Payton, J., Weissberg, R.P., Durlak, J.A., Dymnicki, A.B., Taylor, R.D., Schellinger, K.B. & Pachan, M. 2008, *The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews (Report),* Collaborative for Academic, Social, and Emotional Learning, Chicago, IL.

Pontoppidan, M. & Niss, N.K. 2014, *Instrumenter til at måle små børns trivsel (Rapport 14:17),* SFI - Det Nationale Forskningscenter for Velfærd, København.

Pople, L. & Solomon, E. 2011, *How happy are our children: measuring children's well-being and exploring economic factors (report),* The Children's Society, London.

Rammstedt, B. & John, O.P. 2007, "Measuring personality in one minute or less: An 10-item short version of the Big Five Inventory in English and German", *Journal of Research in Personality,* vol. 41, pp. 203-212.

Ravens-Sieberer, U., Gosch, A., Rajmil, L., Erhart, M., Bruil, J., Duer, W., Auquier, P., Power, M., Abel, T., Czemy, L., Mazur, J., Czimbalmos, A., Tountas, Y., Hagquist, C., Kilroe, J. & European KIDSCREEN Group 2005, "KIDSCREEN-52 quality-of-life measure for children

and adolescents", *Expert Review of Pharmacoeconomics & Outcomes Research,* vol. 5, no. 3, pp. 353-364.

Ravens-Sieberer, U., Herdman, M., Devine, J., Otto, C., Bullinger, M., Rose, M. & Klasen, F. 2013, "The European KIDSCREEN approach to measure quality of life and well-being in children: development, current application, and future advances", *Quality of Life Research,* vol. Pediatric QoL Special Edition, no. DOI 10.1007/s11136-013-0428-3.

Rees, G., Bradshaw, J., Goswami, H. & Keung, A. 2010, *Understanding Children's Well-being: A national survey of young people's well-being (Report),* The Children's Society, London.

Rees, G., Goswami, H., Pople, L., Bradshaw, J., Keung, A. & Main, G. 2013, *The Good Childhood Report 2013 (Report),* The Children's Society, London.

Ringwalt, S. 2008, *Developmental screening and assessment instruments with an emphasis on social and emotional development for young children ages birth through five,* Chapel Hill: The University of North Carolina, FPG Child Development Institute, National Early Childhood Technical Assistance Center.

Rosenberg, M. 1989, *Society and the Adolescent Self-Image (Rev. ed.),* Wesleyan University Press, Middletown, CT.

Rotter, J.B. 1966, "Generalized Expectancies for Internal Versus External Control of Reinforcement", *Psychological monographs,* vol. 80, no. 1, pp. 1-28.

Solans, M., Pane, S., Estrada, M.D., Serra-Sutton, V., Berra, S., Herdman, M., Alonso, J. & Rajmil, L. 2008, "Health-Related Quality of Life Measurement in Children and Adolescents: A Systematic Review of Generic and Disease-Specific Instruments", *Value in Health,* vol. 11, no. 4, pp. 742-764.

Squires, J., Waddell, M., Clifford, J., Funk, M., Hoselton, R. & Chen, C. 2013, "Psychometric study of the infant and toddler intervals of the social emotional assessment measure", *Topics in Early Childhood Special Education,* vol. 33, no. 2, pp. 78-90.

Topp, C.W., Østergaard, S.D. & Søndergaard, S. 2015, "The WHO-5 Well-Being Index: A Systematic Review of the Literature", *Psychotherapy and Psychosomatics,* vol. 84, pp. 167-176.

Trillingsgaard, A. & Damm, D. 2012, *Spørgeskemaer i klinisk praksis med børn og unge,* Dansk Psykologisk Forlag, København.

UCL and Anna Freud Centre 2011, *Me and My School: Findings from the National Evaluation of Targeted Mental Health in Schools 2008-2011 (Research Report DFE-RR177),* Department for Education, London.

UNICEF Office of Research 2013, *Child well-being in rich countries. A comparative overview (Innocenti Report Card 11),* UNICEF Office of Research, Florence.

Veselska, Z., Geckova, A.M., Gajdosova, B., Orosova, O., van Dijk, J.P. & Reijneveld, S.A. 2009, "Socio-economic differences in self-esteem of adolescents influenced by personality, mental health and social support", *European Journal of Public Health,* vol. 20, no. 6, pp. 647-652.

Wiesner, M. & Schanding, T. 2013, "Exploratory structural equation modelling, bifactor models, and standard confirmatory factor analysis models: Application to the BASC-2 Behavioral and Emotional Screening System Teacher Form", *Journal of School Psychology,* vol. 51, pp. 751-763.

Williams, S.T. 2008, *Mental Health Screening and Assessment Tools. Literature Review (Report),* Northern California Training Academy, Center for Human Services, University of California, Davis.

Wolpert, M., Fonagy, P., Frederickson, N., Day, C., Rutter, M., Humphrey, M., Vostanis, P., Meadows, P., Croudace, P., Tymms, P., Brown, P., Aitken, J., Syrad, H. et al. 2008, *Review and recommendations for national policy for England for the use of mental health outcome measures with children and young people (DCSF Research Report 2008/56),* DCSF Publications, Nottingham, England.

Østergaard, K.K., Lando, A.V., Hansen, B.M. & Greisen, G. 2012, "A Danish reference chart for assessment of psychomotor development based on the Ages & Stages Questionnaire", *Danish Medical Journal,* vol. 59, no. 6, pp. A4429.

Det Nationale Institut
for Kommuners og Regioners
Analyse og Forskning

Købmagergade 22
1150 København K
E-mail: kora@kora.dk
Telefon: 444 555 00