



# Assessment Tools for Executive Function and Adaptive Function Following Brain Pathology Among Children in Developing Country Contexts: a Scoping Review of Current Tools

Kwabena Kusi-Mensah<sup>1,2</sup> · Nana Dansoah Nuamah<sup>3</sup> · Stephen Wemakor<sup>6</sup> · Joel Agorinya<sup>4</sup> · Ramata Seidu<sup>4</sup> · Charles Martyn-Dickens<sup>2</sup> · Andrew Bateman<sup>1,5</sup>

Received: 8 June 2021 / Accepted: 21 September 2021 / Published online: 6 December 2021  
© The Author(s) 2021

## Abstract

Several tools have been developed to assess executive function (EFs) and adaptive functioning, although in mainly Western populations. Information on tools for low-and-middle-income country children is scanty. A scoping review of such instruments was therefore undertaken.

We followed the Preferred Reporting Items for Systematic Review and Meta-Analysis- Scoping Review extension (PRISMA-ScR) checklist (Tricco et al., in *Annals of Internal Medicine* 169(7), 467–473, 2018). A search was made for primary research papers of all study designs that focused on development or adaptation of EF or adaptive function tools in low-and-middle-income countries, published between 1<sup>st</sup> January 1894 to 15<sup>th</sup> September 2020. 14 bibliographic databases were searched, including several non-English databases and the data were independently charted by at least 2 reviewers.

The search strategy identified 5675 eligible abstracts, which was pruned down to 570 full text articles. These full-text articles were then manually screened for eligibility with 51 being eligible. 41 unique tools coming in 49 versions were reviewed. Of these, the Behaviour Rating Inventory of Executive Functioning (BRIEF- multiple versions), Wisconsin Card Sorting Test (WCST), Go/No-go and the Rey-Osterrieth complex figure (ROCF) had the most validations undertaken for EF tests. For adaptive functions, the tools with the most validation studies were the Vineland Adaptive Behaviour Scales (VABS- multiple versions) and the Child Function Impairment Rating Scale (CFIRS- first edition).

There is a fair assortment of tests available that have either been developed or adapted for use among children in developing countries but with limited range of validation studies. However, their psychometric adequacy for this population was beyond the scope of this paper.

**Keywords** Executive function · Adaptive function · Assessment · Psychometrics · Developing countries · Children

✉ Kwabena Kusi-Mensah  
kpk27@cam.ac.uk

<sup>1</sup> Department of Psychiatry, University of Cambridge, Clifford Allbutt Building Cambridge Biomedical Campus CB2 OAH, Cambridge, UK

<sup>2</sup> Komfo Anokye Teaching Hospital, P. O. Box 1934, Kumasi, Ghana

<sup>3</sup> Pantang Hospital, Accra, Ghana

<sup>4</sup> Accra Psychiatric Hospital, Accra, Ghana

<sup>5</sup> School of Health and Social Care, University of Essex, Colchester, UK

<sup>6</sup> Department of Psychiatry, University of Michigan Health System, 1500 E Medical Center Dr, Ann Arbor, 48109 MI, USA

## Introduction

### Background

Executive function is a modern complex concept that has only recently begun to find a common consensus of what exactly it means. Since Lezak first coined the term “executive functions” (EF) (Lezak, 1983) and subsequently defined it as “those capacities that enable a person to engage successfully in independent, purposive, self-serving behaviour” (Lezak et al., 2004), the concept has undergone several evolutions to its current understanding. In the early 2010’s, Diamond defined executive function as “top-down control processes” of human behaviour (Diamond, 2013) whose primary function was “supervisory control” as

posited by earlier researchers (Stuss & Alexander, 2000). Diamond went on to describe three core or basic executive functions: **Inhibition**, **Working Memory** and **Cognitive Flexibility** (also called ‘set/task shifting’). They (Diamond, 2013) and others then went on to posit that these three core functions are combined in different ways to achieve what they described as the “higher executive functions”: **Planning** (referring to the ability to identify and organize the steps and elements needed to achieve a goal; Diamond, 2013; Lezak et al., 2004), **Problem-solving**, and **Logical and abstract reasoning**. To back up their model, Diamond refers to work done by Miyake and colleagues (Miyake et al., 2000) and by Lehto et al. (Lehto et al., 2003) who both independently showed by factor analysis that three latent factors emerged (Diamond, 2013). This 3-core-factor model has been fairly well replicated (Collins & Koechlin, 2012; Hall & Marteau, 2014; Karbach & Unger, 2014; Lunt et al., 2012; Miyake & Friedman, 2012).

Adaptive function is a broad concept that covers an array of physical and social functioning- including communication, motor skills and daily living skills. But in the context of cognitive functioning, adaptive function may be viewed as the practical expression of executive functions in an everyday functional context. It may be defined as the ability to carry out everyday tasks within age and context appropriate constraints (World Health Organization, 2001), which may also be impaired following brain injury (Simblett et al., 2012). In the broader context of physical functioning, many chronic conditions can result in impairment of adaptive function. However, in this present study, the term shall be restricted to the narrow scope of adaptive function following brain injury or brain pathology.

### Measuring Executive and Adaptive Functions in The Context of Developing Countries: The Effect of Culture and Socioeconomic Status

Until recently, many researchers seemed to take for granted that the concept of “executive functions” *would* be understood in largely the same way by all people-groups. The seminal work of Lev Vygotsky in the 1930’s (Vygotsky, 1986 English translation) makes a profound point about the dangers of simply “translating” a term meant to convey a specific concept developed in a particular culture into another language and culture without first checking for conceptual equivalence in that recipient language’s culture when working cross-culturally. To that point, several effects of socio-cultural differences on EF have been noted in the literature such as age-matched children and adolescents in Hong Kong out-performing their UK counterparts on all EF functions when controlling for all other relevant factors (Ellefson et al., 2017), with similar differences among immigrants in Denmark (Al-Jawahiri & Nielsen, 2020), and

among an indigenous Mayan community in Mexico compared to urban controls (Ostrosky-Solís et al., 2004) plausibly attributable to cultural differences.

Similarly, regarding socioeconomic status (SES) and EFs, a recently published study among pre-schoolers from South Africa and Australia reported surprisingly that a highly disadvantaged South African subsample from Soweto out-performed middle- and high-SES Australian pre-schoolers on two of three EFs (Howard et al., 2020), suggesting the possibility of EF-protective practices within low-and-middle-income countries. In contrast, in a study from the United States, chronic exposure to poverty was predictive of young children’s poor performance on measures of executive function (Raver et al., 2013) suggesting that the impact of SES on EFs is complex and may depend on the stage of development among other factors. Even more so than with EF, social contexts and cultural expectations affect adaptive function because they shape whatever children learn and perform (Law, 2002; Poulsen & Ziviani, 2006).

But in all these speculations of plausible explanations for these observed differences, one cannot draw the conclusions typically drawn on the effect of culture, SES and other factors on the true underlying latent variable of EF or adaptive function if the thorny issue of using tools developed in foreign cultures to obtain cross-cultural measurement is not adequately addressed (Gannotti & Handwerker, 2002). Thus, assessment of item bias and measurement invariance of any tool used in a cross-cultural context is a crucial but oft-ignored step in the comparison of any two groups, for meaningful conclusions to be made (Fischer & Karl, 2019).

Therefore, a review of assessment tools that, rather than simply assuming universal applicability of any tool developed anywhere, have specifically been either developed or purposefully adapted to a developing country context using scientifically robust methods, is a worthwhile endeavour for practitioners and researchers in such contexts.

### Rationale for Present Study

Improved healthcare has led to reduced mortality rates among children under five years in developing countries (Bakare et al., 2014). However, whether these children who survive are thriving adequately is in some doubt. Increasingly, the disease burden among children in low-and-middle-income countries is shifting from the so-called “childhood killer diseases” to other chronic conditions which may lead to significant impairment and morbidity but not outright mortality (Abubakar et al., 2016; Kieling et al., 2011). Several of these conditions may lead to neurobehavioral difficulties which affects brain function as well as the mental health and well-being of these children. These are often described under the catch-all term of “Acquired Brain Injury” (Bennett et al., 2005; Stuss, 1983). These conditions can lead to

frontal lobe dysfunction which encompasses EF and adaptive function but is often neglected (Simblett et al., 2012).

Better EF is linked to many positive outcomes (Diamond, 2013) such as greater success in school (Duncan et al., 2007; St Clair-Thompson & Gathercole, 2006), while deficits in EF are associated with slow school progress (Morgan et al., 2017) difficulties in peer relationships (Tseng & Gau, 2013) and poor employment prospects (Bailey, 2007). Behaviourally, EF deficits may manifest as distractibility, fidgetiness, poor concentration, chaotic organization of materials, and trouble completing work (Bathelt et al., 2018). Given the difficulties seen, it is therefore important that mental health and rehabilitation services are able to pinpoint areas of greatest difficulty and target interventions appropriately and cost effectively through accurate assessments (Simblett et al., 2012).

Several tools have been developed to assess these areas of frontal lobe functioning in various populations. However, most of these assessment tools have been developed for mainly Western or high-income country populations with not much being known about the tools available for assessing children from low-and-middle-income countries. A literature search revealed only one recently published scoping review on the subject (Nyongesa et al., 2019), which while being very commendable only focused on tools for adolescents (excluding school-age children), searched a very limited scope of databases (only 3) and did not particularly focus on tools developed or adapted specifically for low-and-middle-income countries. Given the high burden of infections and neurodevelopmental conditions in low-and-middle-income countries (Bitta et al., 2017; Merikangas et al., 2009) which are known causes of acquired brain injury which affects EF and adaptive function, awareness of appropriate assessment tools for EF and adaptive function in this specific context and among a wide age-range of children will be highly desirable for clinicians and researchers in these settings.

A Scoping Review approach was chosen in this study because unlike a typical systematic review (which aims to answer a specific question about a specific population according to a rigid set of a priori delimiting factors detailed in a protocol), a scoping review has a broader interest, with the general aim of mapping out the literature and addressing a broader research question, but with the same level of rigour as a systematic review (Shamseer et al., 2015). A scoping review is also unlike a traditional systematic review in that a critical appraisal or risk of bias assessment would be done in a systematic review, while not done in a scoping review (Tricco et al., 2018). Therefore, because the question of interest here was not about the rigour of evidence of one specific tool, or class of tools, used in one specific population within a specific context, but rather just a broad overview of all potential tools used in a broadly defined context, a scoping review was better suited than a systematic review for the research questions specified below.

## Objectives

The present study seeks to undertake a scoping review of published literature to systematically map out whether there are purpose-built assessment tools for executive and adaptive functioning among children (including adolescents) in low-and-middle-income country contexts, and if not, which developed-country tools have been adapted or validated for use among the population of interest, as well as document any knowledge gaps that may exist. The following research questions were therefore formulated:

1. What tools for executive function and adaptive functioning following brain pathology have been adapted or developed or validated for use among children in low-and-middle-income country contexts?
2. Which of these tools have the most literature published supporting their validation for use among children and adolescents in low-and-middle-income countries?

In this paper, we do not aim to critically appraise and summarise the evidence for the scientific rigour of the methodologies used, or report on the psychometric measurement properties established for the identified instruments, as this is beyond the scope of a scoping review. A systematic review of EF and adaptive function measurement tools in low-and-middle-income countries will however be reported on in a subsequent paper.

## Methods

This scoping review was conducted following the Preferred Reporting Items for Systematic Review and meta-analysis-Scoping Review extension (PRISMA-ScR) reporting checklist (Tricco et al., 2018), with our protocol also being drafted using the PRISMA- Protocol extension (PRISMA-P) 2015 guideline (Shamseer et al., 2015). While a specific protocol for this scoping review could not be pre-registered on the dedicated systematic reviews protocols repository PROSPERO (it was rejected on grounds that PROSPERO only registers systematic reviews and not scoping reviews), the protocols for the subsequent systematic reviews of EF and adaptive function measurement tools that proceeded from this scoping review *have* been successfully registered on the PROSPERO website (see here: <https://www.crd.york.ac.uk/prospéro/>) with registration numbers CRD42020202190 and CRD42020203968 for the EF tools systematic review and adaptive function tools systematic review respectively. Since the papers that were critically appraised in these subsequent systematic reviews were initially identified following essentially the same protocol used for this scoping review, the reader's attention is drawn to the pre-registration details of

these systematic reviews as an accurate documentation of the pre-registered methodology that was followed for this scoping review as well.

### Eligibility Criteria

We considered papers of all study designs (qualitative, quantitative, and mixed methods papers) that focused mainly on the target outcomes (executive functioning or adaptive functioning following brain pathology) among children in low-and-middle-income countries. The paper also had to primarily be concerned with developing, adapting or assessing the validity of the instrument of choice as one of its main stated study aims (if not the main), and not just as an incidental concern, to be eligible. Peer-reviewed articles, as well as expert opinions and published guidelines with provided rationales were considered. Participants considered were children aged 5 years and above to 18 years, including both healthy and clinical populations.

All studies recorded in the databases searched that were published at any time were considered, with no year limitations being placed on the search. In operational terms, this meant that papers published from 1<sup>st</sup> January 1894 (the earliest date in PsychINFO, one of the databases searched) to 15<sup>th</sup> September 2020 (the last day of update of the search strategy) were included. With respect to settings, studies conducted in developing country settings were selected. “Developing country” was defined using the World Bank list of lower-income countries (LIC), and middle-income countries (including both lower-middle-income countries-LMIC and upper-middle-income countries-UMIC) defined as of 2012 (Cochrane Library, 2012; World Bank Group, 2019) which are collectively referred to as “low-and-middle-income countries”. In terms of language, no a priori language limitations were placed on the search. Full articles written in English, and those in other languages that could reasonably be translated (either by using Google Translate or by finding native speakers willing to volunteer their services) were included. An appendix of potentially relevant articles in other languages that could not be translated is provided in Appendix II.

Specifically excluded, apart from those which did not generally meet the inclusion criteria above, were: (a) studies that only used the instrument as an outcome measurement instrument (for instance in randomized controlled trials) rather than specifically evaluating its psychometric properties or reporting on its local adaptation, (b) studies in which the EF or adaptive function instrument was used in a validation study of another (non-EF or adaptive function) instrument (i.e., validation was **not** of the EF or adaptive function tool, but rather another instrument for another construct such as say “long term memory”), (c) news articles, blog posts and other such mass media outlet writings, (d) studies on animals, and (e) studies focused on high income countries.

### Information Sources

The following databases were searched for the following reasons:

1. MEDLINE (OVID interface, 1946 onwards), because this is one of the largest databases for health and medical literature (also known as “pubmed”) maintained by the United States government.
2. EMBASE (OVID interface, 1974 onwards), because it is complementary to MEDLINE and also because it is mainly a pharmacology and pharmaceuticals database it might have had validation papers related to drugs targeting frontal lobe dysfunction.
3. Cochrane library (current issues), since it is the main database of systematic reviews. Also, this was included because both Cochrane library and EMBASE include not only just published journal articles but also unpublished data (“grey literature”) like conference proceedings, and drug repository databases etc.
4. PsychINFO (1894 onwards), because it is the main database for psychiatry and psychology related research.
5. Global health (1973 onwards), this is a database for public health focused articles which might include eligible validation papers in the context of epidemiological surveys.
6. Scopus, since this is a database dedicated to multidisciplinary research (including qualitative research), given that we were interested in capturing research of a multidisciplinary nature.
7. Web of Science, since this is a database dedicated to multidisciplinary research
8. SciELO, this is Latin America focused database providing scholarly literature in sciences, social sciences, and arts and humanities published in leading open access journals from Latin America, Portugal, Spain, and South Africa; this was an important source of non-English language studies from low-and-middle-income countries in Latin America, particularly from Brazil.
9. Education Resources Information Centre (ERIC, 1966 onwards), important for papers on cognitive assessment published in the special education literature; included theses, dissertations, and teaching guides.
10. British Education Index (BEI, 1996 onwards), because it covers research done in education on evaluation and assessment, technology, and special educational needs.
11. Child Development & adolescent studies (CDAS, 1927 onwards), for same reason as above. Included theses, dissertations, and teaching guides.
12. Applied Social Sciences Index and Abstracts (ASSIA), because it is an important source for multidisciplinary papers. Includes social work, nursing, mental health and education journals.

## Gray Literature Data Sources

13. Open grey (1992 onwards). Includes theses, dissertations, and teaching guides
14. PROSPERO. Repository of pre-registered study protocols for systematic reviews for trial protocols for similar scoping reviews through PROSPERO.
15. Cochrane library (see above)
16. EMBASE (see above)
17. ERIC (see above)
18. CDAS (see above)

Thus 14 unique databases were searched. The initial search was done by 20<sup>th</sup> March 2020 while the final updated search was completed by 15<sup>th</sup> September 2020. Further, to make sure we were thorough in our literature search, we also scanned the reference list of selected papers for other papers of possible interest which might have been missed in the literature search, particularly so for systematic and scoping review papers we found in our search. The search in the ‘grey literature’ was included to enable us capture unpublished and non-peer-reviewed data on the matter. This was to help mitigate the risk of publication bias and other meta-biases.

## Search Strategy

We developed literature search strategies using text words and medical subject headings (MeSH terms- including using the ‘explode’ function to get all related sub-categories of the MeSH terms) related to the following themes:

- Executive function/Frontal lobe function/Frontal lobe damage/Adaptive Function and their variants using truncation
- Assessments/Validation/reliability/norms/reproducibility/standardization of instruments and their variants using truncation
- Children/adolescents and their variants using truncation
- Developing countries/lower-middle-income-countries and their variants using truncation

The search strategy was developed by a member of the study team (KKM) who had undergone extensive training in conducting Systematic Reviews and in using search strategies in all the above-named databases from the Medical Library Services. The search strategy was also reviewed by an experienced Medical Librarian who has extensive expertise in systematic review searching. The full search strategy for MEDLINE is re-produced in Appendix I (see Supplemental Material).

## Selection of Sources of Evidence

For the selection process six reviewers worked on all abstracts and full papers. At the screening phase, at least two independent reviewers screened each title and abstract obtained from the search and compared their results with each other. Where there were disagreements on eligibility based on abstract alone, the full text article was retrieved and reviewed, and all discrepancies discussed and resolved. Where resolution was not possible after discussion, a third independent reviewer was brought in as arbiter, and as a last resort, the guarantor was consulted as a final arbiter. Further, the pre-resolution inter-rater agreement was calculated and reported as ranging between 81.6%–88.9%, which was above the recommended minimum 80% agreement. In accordance with PRISMA recommendations, the selection process was documented in a flow diagram (see Fig. 1 below).

## Data Charting Process

At least two reviewers independently extracted data from the screened articles using a purpose-made data extraction chart, designed using items from the PRISMA-ScR and PRISMA-P checklists, to select suitable articles meeting the inclusion criteria. Table 1 in Appendix II shows a sample of this chart. First a calibration exercise was done for all 6 reviewers using a sample of 100 abstracts and 10 full paper articles to ensure uniform use of the screening criteria and the charting forms. The search results were uploaded and saved into Mendeley using the ‘groups’ function, which allowed online collaboration and discussion among the reviewers. Study abstracts and full texts were uploaded based on the screening criteria. Duplication were minimized by employing the ‘merge duplicates’ function on Mendeley. In spite of this, manual de-duplication still had to be done for a few abstracts. The Data were charted using Excel spreadsheets.

## Data Items

The following data points were collected using the following definitions:

**Instrument Reference** This referred to the name of lead author and publication year of the paper.

**Instrument Name** This referred to the instrument name and version under consideration.

**Outcome Variable** The outcome variables of interest were executive functioning and adaptive functioning. We collected data on which outcome of interest was considered in the



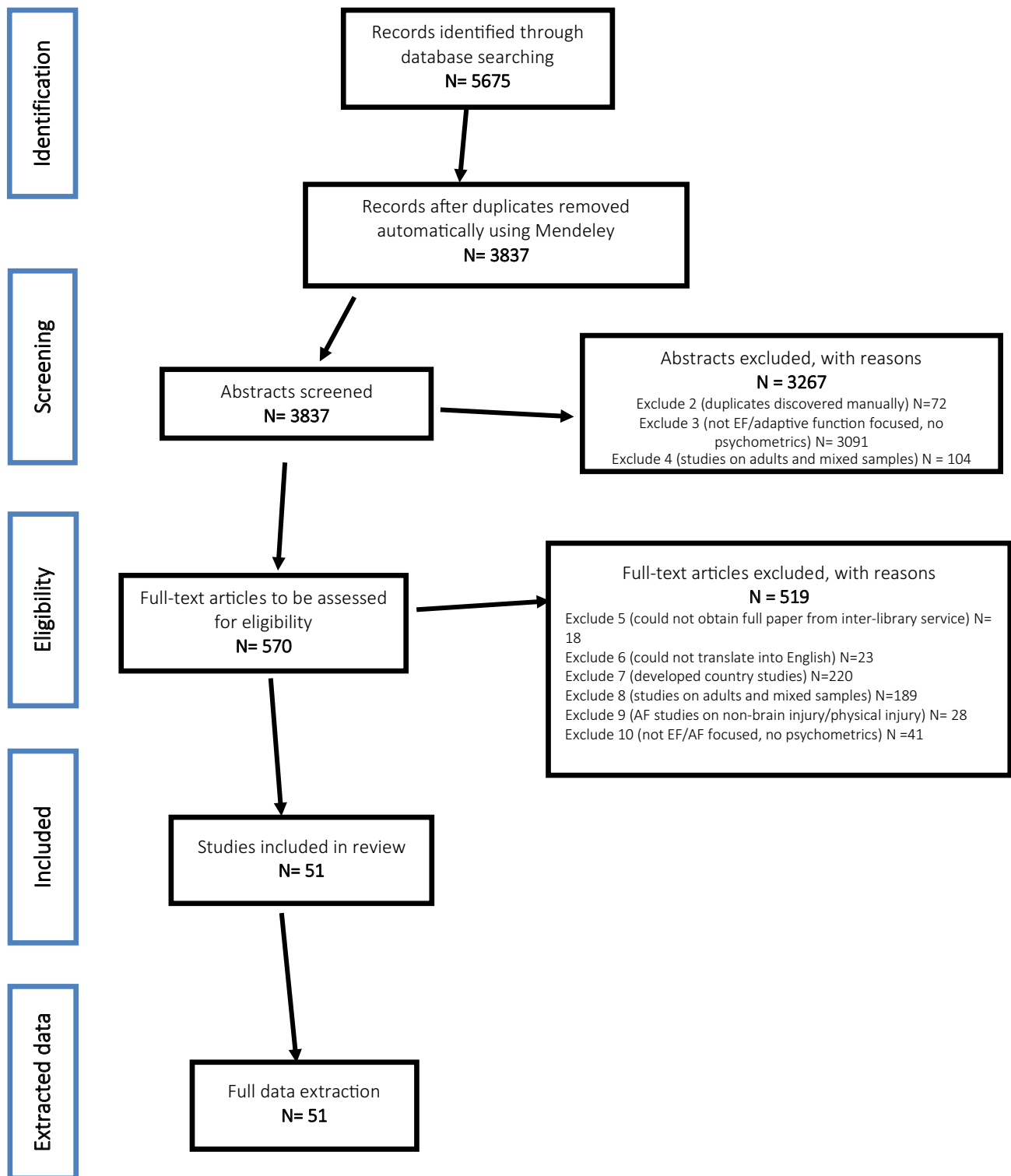


Fig. 1 PRISMA flow diagram of study selection

paper. Executive functioning was defined as “those capacities that enable a person to engage successfully in independent, purposive, self-serving behaviour, with specific executive functions consisting of initiation, planning, purposive action, self-monitoring, self-regulation, decision-making

or flexibility, inhibition and volition” as reported by Stuss (Stuss, 2011). Adaptive functioning was defined as behaviours necessary for age-appropriate, independent functioning in social, communication, daily living or motor areas (Matson et al., 2009).

**Table 1** Database Results and Dates

Database	Date of search coverage	Initial search by 20/03/20	Initial de- duplication	Updated search by 15/09/20	Updated de- duplication
MEDLINE via OVID	1946—05/09/20	803	662	939	760
EMBASE <sup>a</sup>	1974—05/09/20	27	26	29	27
Cochrane Library <sup>a</sup>	* 11/09/20	67	49	68	48
PsychINFO	1894—11/09/20	1314	985	1490	812
Global Health	1973—11/09/20	38	14	38	13
Scopus	1970—11/09/20	1093	1080	1134	910
Web of Science	1900 – 11/09/20	1730	1254	1820	1184
SciELO	2002 – 15/09/20	29	28	29	26
ERIC <sup>a</sup>	1966—15/09/20	19	2	19	2
BEI	1996—15/09/20	2	2	3	1
CDAS <sup>a</sup>	1927—15/09/20	13	3	13	2
ASSIA	1987—15/09/20	84	81	87	48
Open Gray <sup>a</sup>	1992—15/09/20	4	4	4	4
PROSPERO <sup>a</sup>	2011—15/09/20	1	0	2	0
Other (reference lists etc.)	-	-	-	0	0
<b>TOTAL ABSTRACTS</b>				<b>5675</b>	<b>3837</b>

ERIC: Education Resources Information Centre, ASSIA: Applied Social Sciences Index and Abstracts, CDAS: Child Development & Adolescent studies, BEI: British Education Index

\*Although first established in 1996, composite nature of Cochrane library means it does not have a “start date” as other biomedical databases do

<sup>a</sup>includes theses, dissertations, teaching guides and other such non- peer-reviewed “grey” literature

**Country Settings** The desired setting was ‘developing country’ setting which was defined according to the World Bank list of lower income (LIC) (GNI per capita less than \$1025), Lower-middle-Income Country (LMIC) (GNI per capita between \$1026 to \$3995) and upper-middle-income country list (UMIC) (GNI per capita between \$3,996 TO \$12,375) (Cochrane Library, 2012; World Bank Group, 2019), which were collectively referred to as “low-and-middle-income countries” in this paper. This broad approach to defining “developing country” was taken because we wanted to only exclude high-income countries, since the goal of the review was to capture tools that had been adapted for use in relatively low-resource cross-cultural settings. In that regard, we felt using as broad a definition as possible would make the findings more relevant to a lot more readers working in non-high-income country settings. Secondly, the decision to use the 2012 data was guided by the fact that we used an extremely detailed country list of low-and-middle-income countries developed by the Cochrane Library (Cochrane Library, 2012), in which rather than simply using only terms like “developing country” or “LAMIC” in the search term, the names of all 160+ non-high-income country countries (including their former names and abbreviated names- see Appendix I) could be directly entered into the search strategy. This Cochrane list was based on 2012 world bank GNI levels and was employed to be extremely thorough in our search and to improve the accuracy of the results. Having said that, when an abstract was selected using this method, the full paper was further screened to ensure

that at the time the data was collected, the country was still on the World Bank List of LIC, LMIC or UMIC countries. This was done by cross-referencing data collection dates in the paper with historical data obtained from the World Bank (World Bank Group, 2019).

**Type of Study** Data on whether the study was an adaptation or validation of an existing instrument, or a development of a new instrument was extracted, usually from the aims and objectives, methods, or results section, and documented. ‘Validation of an assessment tool’ was defined according to specific items or criteria used for reliability and validity according to the **CO**nsensus-based **S**tandards for the selection of health status **M**easurement **I**Nstruments) checklist items (COSMIN) guidelines (Mokkink et al., 2018; Prinsen et al., 2018). Specific items (including their taxonomy and definitions) that were included as part of validation if they were reported upon were defined as follows:

- **Reliability:** The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions. This comprises of the following subsets:
  - **Internal consistency:** The degree of the interrelatedness among the items. In other words, internal consistency is the maintenance of the same score for the same patient when different sets of items from the same instrument are used.

- *Measurement error*: The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured.
- *Reliability*: The proportion of the total variance in the measurements which is due to 'true' differences between patients.
- **Content Validity (including face validity)**: The degree to which the content or items of an instrument is an adequate reflection of the construct to be measured.
- **Construct Validity**: The degree to which the scores of the instrument are consistent with hypotheses (for instance regarding internal relationships, relationships to scores of other instruments, or differences between relevant groups) based on the assumption that the instrument validly measures the construct to be measured. This will comprise of the following subsets:
  - *Structural Validity*: The degree to which the scores of an instrument are an adequate reflection of the dimensionality of the construct to be measured.
  - *Cross-cultural Validity*: The degree to which the performance of the items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version of the instrument.
  - Hypothesis-testing: How well an expected hypothesis of how the tool is expected to behave, is fulfilled (i.e., how well the tool behaves as expected). Specific examples of hypothesis-testing assessed here were “discriminant validity” (how well the tool discriminates between expected population groups, such as autism patients versus non-patients) and “convergent validity” (how well the assessed tool converges with another similar tool in terms of their expected scores).
- **Criterion Validity**: The degree to which the scores of an instrument are an adequate reflection of a 'gold standard'. For outcome measurement instruments the 'gold standard' is usually taken as the original full version of an instrument where a shortened version is being evaluated.
- **Responsiveness**: The ability of an instrument to detect change over time in the construct to be measured.
- **Interpretability**: Interpretability is the degree to which one can assign qualitative meaning - that is, clinical or commonly understood connotations - to an instrument's quantitative scores or change in scores. Although not a measurement property, it is an important characteristic of a measurement instrument.

Any paper that reported information on any of the above was considered as eligible for having included an eligible outcome measure. In accordance with the COSMIN guideline (Mokkink et al., 2018; Prinsen et al., 2018), “study”

was defined as any individual validation conducted in any given research project or paper. For example, a given paper might report the conduct of construct validation, cross cultural validation, and structural validation of one instrument all within the same paper. This was thus reported as three studies reported within one paper.

**Target Population** The exact age-category of children that were included.

**Mode of Administration** Whether the instrument was a performance-based task, or an informant-based tool (i.e., self-reported or parent or proxy-based questionnaire etc.).

**Sub-domains** The number of sub-scales or sub-domains or items of interest of the tool in question.

**Language of Publication** Which language the paper was originally published in.

**Sample Size** Number of participants used.

**Demographics** Mean age and gender percentages of sample.

**Local Settings** Whether study was predominantly set in rural or urban settings (or both).

**Condition** Whether study was conducted among a healthy sample or clinical sample, and if so, what clinical condition.

**Language of population** What local language-group was the study conducted among.

### Analysis and Synthesis of Results

A synthesis of all data to summarise findings of included studies was done independently by at least 2 reviewers, compared and consensus reached. First the number of individual instruments reported on in all eligible studies found in our search were documented. Then we grouped the instruments according to the construct they measured- EF versus adaptive function. Where we encountered a systematic or scoping review, we retrieved and screened the original papers reviewed by that systematic review according to our eligibility criteria and included any that had been missed by our search for our own independent evaluation. The extracted data were then re-categorized according to each individual instrument reported and summarised. Percentage frequencies and other statistics were calculated in Microsoft Excel spreadsheets. Results were displayed using the Data Extraction Chart as shown in Table 2 below. The results are presented below in the narrative and tabular formats, and reported according to the PRISMA-ScR checklist (Tricco et al., 2018).



## Results

### Selection of Sources of Evidence

Table 1 summarises the results of the initial search in each individual data source, along with the dates of coverage of the search in each database. The databases with the most hits given the search criteria were Web of Science, PsychINFO, Scopus and MEDLINE, which was expected given their respective scopes of subject matter as discussed above (see *Information Sources*).

After an automatic de-duplication using Mendeley, the scoping review identified 3837 potentially eligible abstracts for manual screening. Of these a further 72 were excluded for being duplicates (missed by the automatic de-duplication), 3091 were excluded for either not reporting any psychometric data (i.e., not being development or validation or adaptation studies) or not being about EF or adaptive functions at all, while 104 were studies focused wholly or mostly on adult populations. 570 full articles were thus further screened for eligibility, with 519 being excluded for such reasons as failure to translate into English (23 papers), country settings being predominantly high-income (220 papers), and the study being about an adaptive function tool being used in a non-brain pathology or physical injury context such as limb amputation (28 papers), among other reasons (see Fig. 1 for summary). Ultimately, 51 full papers were found to be eligible for full data extraction and review.

### PRISMA Flowchart for Study Selection

#### Characteristics and Results of Individual Sources of Evidence

Table 2 presents the characteristics of the data extracted from all eligible papers (sources of evidence) in the data extraction chart. As can be seen from the ‘‘Type of Study’’ column of this table, several papers reported multiple ‘‘studies’’ in a single paper, where ‘‘study’’ was defined as an individual validation as recommended by the COSMIN guideline. For example, in Senturk et al. (2014) paper on the Junior Brixton Test, they reported on two studies- both structural validity and construct validity- for the JBT. By this count, a total of 163 studies were reported in 51 papers. When disaggregated into individual studies, the most frequently conducted type of study (in descending order) were structural validity and construct validity or hypothesis testing studies at 38 studies each (23.3% each of total individual studies), followed in order by internal consistency studies at 27 (16.6%), reliability studies at 23 (14.1%), cross cultural validity studies at 14 (8.6%), adaptation or content validity studies at 13 (7.9%), instrument development at 6 (3.7%), responsiveness 3 (1.8%) and measurement error 1

(0.6%). Tables 3 and 4 in Appendix III shows the breakdown of these by EF individual instruments and adaptive function individual instruments, and are reported on in ‘‘Synthesis of Results’’ below. Informant-based instruments (either by self, parents, or another proxy) were slightly more frequent at 29 instances (53.7%) compared to performance-based measures at 25 instances (46.3%). The top 3 low-and-middle-income countries in which these validation studies were conducted were Iran (7 papers), Brazil (6 papers) and Colombia and Argentina (4 papers each). But when categorized in terms of world regions (i.e., regions with broadly similar cultural or linguistic environments), the top performing regions with the most papers reporting validation studies from them were Latin America (Central and South America) with 17 papers (30.4% of instances), Sub-Saharan Africa 12 papers (21.4%), the Middle East 10 papers (17.9%) and South-East Asia (including the Indian sub-continent) with 8 papers (14.3%). Most studies conducted involved urban populations at 46 instances (76.7%), and exclusively healthy populations at 31 instances (60.8%) as opposed to clinical populations (with healthy controls) at 20 instances.

### Synthesis of Results

In this sub-section, the results are disaggregated according to the individual instruments and reported on. In this scoping review, 40 unique tools, including 49 version or variants, were identified as having been either developed or adapted or validated for use among children in low-and-middle-income countries from the 51 papers reviewed. A total of 130 individual studies were done for the EF instruments reported on in this review (see Table 3 in Appendix III). Figure 2 shows the top 5 EF instruments that reported any validation study. BRIEF (Gioia et al., 2000), in all its various versions, was by far the most validated instrument in terms of numbers of reported studies with 26 validation studies in total (20%), followed by DKEFS (Delis et al., 2001) (9 studies- 6.9%), WCST (Berg, 1948) (8 studies- 6.2%), Go/No-go (Luria, 1973) (7 studies- 5.4%) and NEPSY (Korkman, 1998), and ROCF (Rey, 1941) (6 studies each- 4.6%). This is a remarkable performance for the BRIEF as it was developed only in the last 20 years compared to the legacy tests WCST and Go/No-go which have been existence for over 70 years, hence it is not surprising that several studies have been done on those tests. In Fig. 3, these results were broken down by types of validation studies and again summarised by top-performing instruments (Table 3 in Appendix III shows the full results of the frequencies of specific types of validation studies done for all individual Executive Function instruments). In this figure, only instruments with 2 or more validation studies were individually named, with all instruments that had only one validation study in any particular validation study type being grouped under ‘‘other instruments’’. When these results were thus broken-down

**Table 2** Characteristics of Sources of Evidence

Instrument Reference (authors, year)	Name and Version of Instrument	Outcome Measure Assessed	Country/ Study Sites	Type of Study*; Development, validation	Target Population	Mode of Administration	Sub-scales/ number of items of interest	Sample size	Age range and mean	Local Setting	Condition (clinical/ healthy)	Language(s) of Population
(Baker et al., 2011)	Behavior Rating Inventory of Executive Function BRIEF (Gioia et al., 2000)	EF	Turkey	construct validity (convergent and discriminant)	school-age children	parent report	8 sub-scales; as above for BRIEF	80; 61 ADHD <sup>a</sup> ; 19 controls	6–11 years;	urban	ADHD <sup>a</sup> and healthy	Turkish
(de Bustamante Carm et al., 2012)	Behavior Rating Inventory of Executive Function BRIEF (Gioia et al., 2000)	EF	Brazil	structural validity and internal consistency	children and adolescents	self-report, parent, and proxy (teacher) report	8 sub-scales: inhibition, flexibility, and emotional control (behavioral regulation index), initiative, working memory, planning / organization, material organization and monitoring (metacognition index)	277 parents; 282 teachers; 112 adolescents	9, 8 (3.4)	urban	healthy	Portuguese
(Burley et al., 2015)	Behavior Rating Inventory of Executive Function BRIEF (Gioia et al., 2000)	EF	Uganda	Internal consistency	school-age children	parent report	8 sub-scales; as above for BRIEF	185 children; 28 ADHD <sup>a</sup> ; 157 healthy	8, 5 years	urban	clinical (ADHD <sup>a</sup> ) and healthy	Luganda
(Zarabi et al., 2015)	BRIEF (Gioia et al., 2000)	EF	Iran	construct validity (discriminant)	school-age children	parent report	8 sub-scales; as above for BRIEF	60; 30 cases, 30 controls	range: 7–12 years; mean: 9yrs (1.41)	urban	ADHD <sup>a</sup> and healthy	Persian
(Rincon Diaz & Rey Anacona, 2017)	BRIEF- pres-school version (Gioia et al., 2000)	EF	Colombia	adaptation, structural validity, internal consistency, test-retest reliability, construct validity (convergent)	pre-school children	parent report	5 sub-scales: inhibition, change, emotional control, working memory and planning / organization	125	range: 2–5 years; mean 4.3	urban and rural	healthy	Spanish
(Chernoff et al., 2018)	BRIEF (Gioia et al., 2000)	EF	South Africa, Malawi, Zimbabwe, Uganda	adaptation, reliability (test-retest), measurement error, convergent validity	school-age children	parent report	8 sub-scales; as above for BRIEF	603; 244 HIV, 179 HIV-exposed, 180 no HIV)	range: 5–10	rural and urban	HIV, HIV exposed but healthy, no HIV exposure	Luganda, Afrikaans, Xhosa, Shona, Zulu, Chichewa, Sesotho, and Setswana
(Selvam et al., 2018)	BRIEF- pres-school version (Gioia et al., 2000)	EF	India	structural validity, internal consistency, reliability (inter rater), cross cultural validity, construct validity (discriminant)	pre-school children	parent report	5 sub-scales: inhibition, change, emotional control, working memory and planning / organization	412	2–5 years	urban	healthy	Kannada
(Amami et al., 2018)	BRIEF- teacher (Gioia et al., 2000)	EF	Iran	structural validity, internal consistency, construct validity (convergence)	school-age children	proxy (teacher) report	8 sub-scales; as above for BRIEF	360 children, 12 teachers	range: NR	urban	healthy	Persian
(Rosetti et al., 2018)	Ball Search Field Task- BSFT (Rosetti et al., 2016), BRIEF	EF	Mexico	construct validity (convergent and discriminant (age)), internal consistency (for BRIEF)	children and adolescents	interview-based task; parent report (BRIEF)	single task	106	range: 6–16; mean: 10.2 (2.7)	urban	ADHD <sup>a</sup> only	Spanish

**Table 2** (continued)

Instrument Reference (authors, year)	Name and Version of Instrument	Outcome Measure Assessed	Country/ Study Sites	Type of Study; Development, validation	Target Population	Mode of Administration	Sub-scales/ number of items of interest	Sample size	Age range and mean	Local Setting	Condition (clinical/ healthy)	Language(s) of Population
(Rosete, 2018)	Delis Kaplan Executive Function System DKEFS Children's Color Trails Test (CCTT); Color Trails Test (CTT) (Delis et al., 2001)	EF	Thailand	Structural validity	adolescents	interview-based tasks	Design Fluency and Verbal fluency subscales of DKEFS	156	16.44 (2.04)	urban	healthy	Thai
(Pluck et al., 2019)	Tower Test of DKEFS (Delis et al., 2001)	EF: planning, working memory	Ecuador	Structural validity, internal consistency, reliability (test-retest)	children and adolescents	interview-based task	1 task	264	range 10–20, mean 15.05	urban	healthy, but some street children and foster children with controls	Spanish
(Tageu et al., 2012)	Wisconsin Card Sorting Test (WCST) (Berg, 1948)	EF	Romania	construct validity (discriminant)	Mostly children and few adults	interview-based task	single task	1101	26.7	urban	clinical and healthy controls	Romanian
(Pineda et al., 2007)	bespoke battery including: WCST- abbreviated, verbal fluency test, token test, Rey-Osterrieth complex figure ROCF (Rey, 1941)	EF in ADHD <sup>1</sup>	Colombia	structural validity, construct validity (discriminant)	school-age children	interview-based task	3 tasks	621 total: 249 cases, 372 controls	range: 6–11,	urban	ADHD <sup>2</sup> and healthy	Spanish
(Nampijja et al., 2010)	Wisconsin Card Sorting Test (WCST); Knock-Tap Game (Go/No-Go) (Luria, 1973)	EF: cognitive flexibility, planning	Uganda	adaptation, reliability (test-retest), construct validity (discriminant)	pre-school children	interview-based task	single task (each)	64	range: 4–6; mean: 5.2	rural and urban	healthy	Luganda
(Sarconi et al., 2020)	Go/No-go (Luria, 1973) app version	EF: inhibition	Brazil	content, structural validity, internal consistency, cross-cultural, and construct (convergent and discriminant)	school-age children	interview-based task	4 tasks of inhibition	306: 253 healthy, 53 DCD <sup>3</sup>	range: 8–10 years	urban	Development Coordination Disorder (DCD <sup>3</sup> ), healthy	Portuguese
(Holding et al., 2018)	Rey-Osterrieth complex figure-ROCF (Rey, 1941), Go/No-Go (Luria, 1973), Shift (Holding et al., 2018)	EF: Working Memory, Inhibition, cognitive flexibility	Bangladesh, Ghana, Tanzania	adaptation, internal consistency, cross cultural validity, reliability (test-retest, inter rater), responsiveness	children	interview-based task	3 scales: each 1 item/ task given multiple times	786 total: 166 Ghana, 323 Tanzania, 297 Bangladesh	range: 7–18; mean: 13	rural	healthy	Kasem and Nankam (Ghana), Swahili (Tanzania), Chitagonian (Bangladesh)
(Kashala et al., 2005)	Design Copying, Tower Test (from NEPSY <sup>2</sup> ) (Korkman, 1998), Digit Span (backwards) (Blackburn & Benton, 1957)	EF: planning, working memory	Democratic Republic of Congo	Structural validity, construct validity (discriminant)	children	interview-based task	Design Copying, Tower Test (from NEPSY <sup>2</sup> ), Digit Span (backwards)	185 (28 cases, 157 controls)	7–9.9 years, mean 8.5 years	urban	mixed: 15% clinical (ADHD <sup>3</sup> ), 85% healthy	French; assorted African dialects
(Trevisan et al., 2017)	Children's Executive Function Inventory CHEXI (Thorell & Nyberg, 2008)	EF in ADHD <sup>1</sup>	Brazil	adaptation (development), structural validity, internal consistency, construct validity (convergence or concurrent)	pre-school children	parent report/ proxy report	4 subscales: working memory, inhibition, self-regulation, and planning	408 parents and teachers	range 4–7, M=5.51, SD=0.59	urban	ADHD <sup>3</sup> and healthy controls	Portuguese
(Thorell et al., 2020)	Teenage Executive Functioning Inventory TEFI (Thorell et al., 2020)	EF: Working memory, inhibition	Serbia	adaptation, structural validity, internal consistency, reliability (inter-rater)	adolescents	self-report, parent, and proxy (teacher) report	20 items, 2 subscales: inhibition and working memory	302 adolescents	range: 13–19; mean 16.4 (1.8)	rural and urban	healthy	Serbian

Table 2 (continued)

Instrument Reference (authors, year)	Name and Version of Instrument	Outcome Measure Assessed	Country/ Study Sites	Type of Study: Development, validation	Target Population	Mode of Administration	Sub-scales/ number of items of interest	Sample size	Age range and mean	Local Setting	Condition (clinical/ healthy)	Language(s) of Population
(Ruffieux et al., 2010)	Assessment Battery containing: Color Trail Test, Hand movements, Verbal Fluency, letter numbering sequence (Ruffieux et al., 2010)	EF in Sickle Cell Disease	Cameroun	Structural validity, inter rater reliability	children and adolescents	interview-based task	14 tasks; but only 5 relevant for EFs	125	range: 6–20, mean 11.4 (4.2)	urban	healthy	French, unnamed local dialects
(Injoque-Riele et al., 2011)	The Automated Working Memory Assessment AWMA (Alloway, 2007)	EF: working Memory	Argentina	adaptation (development), internal consistency, construct validity (convergent)	school-age children	interview-based task	12 tasks: Digit Recall, Word Recall, Non-word Recall, Dot Matrix, Block Recall, Mazes Memory, Listening Recall, Counting Recall, Backward Digit Recall, Odd One Out, Mr. X Spatial Span	26 adaptation, 210 validation	range 6–11	urban	healthy	Spanish
(Wong et al., 2012)	Ballet Executive Scale BES (Wong et al., 2012)	EF in dance/ ballet	Cuba	structural validity, internal consistency, concurrent validity	adolescents	self-report	5 subscales: Strategic Planning, Organization of Dance Behavior, Motivation to Dance, Impulse Control in Dance Behavior, Empathy toward Other Dancers	149	range: 11–18	urban	healthy	Spanish
(Sobeh & Spijkers, 2012)	Testbatterie zur Aufmerksamkeitsprüfung für Kinder KITAP (Zimmermann et al., 2002)	EF: Attention, inhibition, cognitive flexibility	Syria	Cross cultural validity, Construct validity (discriminant)	school-age children	performance-based task	Flexibility (The Dragon's House), Go/No-Go subscales	143	8	urban	healthy	Syrian Arabic
(Malek et al., 2013)	Stroop Colour Word Test (Stroop, 1935) Victoria Edition	EF: inhibition	Iran	construct validity (discriminant), reliability (test-retest)	adolescents	interview-based task	single task	180: 150 healthy, 30 ADHD*	range: 12–17 years	urban	ADHD* and healthy	Persian and Turkish
(Senturk et al., 2014)	Junior Brixton Test JBT (Shallice et al., 2002)	EF	Turkey	Internal structure (structural validity), construct validity (convergent)	School-age children	interview-based task	single task	121	7.2	urban	healthy	Turkish
(Garcia-Barrera et al., 2015)	Behavioural Assessment System for Children BASC (Reynolds & Kamphaus, 1992)	EF	Colombia	internal consistency, cross cultural validity, Construct validity (discriminant)	school-age children	parent report	Not stated	848 healthy; 155 clinical	6–11 years	urban	ADHD* and healthy	Spanish
(Sineira et al., 2016)	Child Hoarding Test (Burgess & Shallice, 1997) Brazilian version	EF: inhibition, cognitive flexibility	Brazil	adaptation (development)	school-age children	interview-based task	24 items	139	range: 6–12	urban	healthy	Portuguese
(Sallum et al., 2017)	Self-Ordered Pointing Task SOPT (Petrides & Milner, 1982)	EF: working Memory	Brazil	construct validity (discriminant and convergent), ecological validity	pre-school children	interview-based task	single task	248	range: 3–5	urban	healthy	Portuguese

Table 2 (continued)

Instrument Reference (authors, year)	Name and Version of Instrument	Outcome Measure Assessed	Country/ Study Sites	Type of Study*: Development, validation	Target Population	Mode of Administration	Sub-scales/ number of items of interest	Sample size	Age range and mean	Local Setting	Condition (clinical/ healthy)	Language(s) of Population
(Richard's et al., 2017)	Conjunction Visual Search (CVS) task of tareas de autorregulación Cognitiva TAC battery (Richard's et al., 2017)	EF: inhibition	Argentina	development (but of computer-based task), convergent validity	school-age children	interview-based, computer administered task	1 task: 120 trials	41	range: 6–11, mean: 8.49; standard deviation: 1.47	urban	healthy	Spanish
(Yang et al., 2018)	Dysexecutive Questionnaire DEX (Wilson et al., 1996)	EF	China	structural validity	adolescents	self-report	20 items	1586	mean: 18.9 (5.4)	urban	healthy	Mandarin
(Barreto et al., 2018)	Neuropsychological Assessment of Executive Functions (Spanish acronym: ENFEN) (Porcellano et al., 2009)	EF	Colombia	structural validity, construct validity (convergent and discriminant)	school-age children	interview-based task	Not stated	367 children	range: 6–12; mean 8.9 (1.8)	rural and urban	healthy	Spanish
(Du et al., 2018)	Symptoms and Functional Impairment Rating Scale SFIRS (Du et al., 2018)	EF in ADHD <sup>1</sup> ; working memory, planning, time management, self-monitoring, and emotional control	China	development, structural validity, internal consistency, reliability (test-retest), construct validity	school-age children	parent report	6 sub-scales: hyperactivity-impulsivity, self-control, inattention, self-management, school performance, social interaction	20 development; 412 validation	range: 6–12, 8.7 (1.42)	urban	ADHD <sup>1</sup> and healthy	Mandarin
(Richard's et al., 2018)	Tareas de Autorregulación Cognitiva Battery (TAC)	EF: inhibition, working memory and cognitive flexibility	Argentina	structural validity	children	interview-based task	3 sub-scales used: perceptual inhibition task, 4 WM tasks, 1 cognitive flexibility task	103 children	range: 9–12 years; M=10.84 SD=0.88	urban	healthy	Spanish
(Green et al., 2019)	The Cambridge Neuropsychological Test Automated Battery CANTAB (Huppert et al., 1995)	EF: working memory, inhibition, and attention shifting	Mexico	construct validity (discriminant)	children and adolescents	interview-based task, computer based	6 sub-tests (out of 23): stockings of Cambridge, delayed matching sample, extra-dimensional shift, rapid visual information processing, stop signal task and match to sample visual search task	826 children	range 5–15; mean 9.27 (2.6)	urban	healthy	Spanish
(Ford et al., 2019)	Rapid Assessment of Cognitive and Emotional Regulation RACER (Ford et al., 2019)	EF	Lebanon, Niger	cross cultural validity	children	interview-based, computer administered task	2: inhibition; working memory	2725; 866—Niger, 1859 Lebanon	Lebanon: 9.2 (2.3); Niger: 9.2 (1.4)	urban	healthy	French; Arabic
(Korzeniewski & Ison, 2019)	Executive Function Scale for Children EFS (Korzeniewski & Ison, 2019)	EF	Argentina	content validity, structural validity, internal consistency, reliability (test-retest)	school-age children	parent report	6 sub-scales: attention control, inhibitory control, metacognition, organization, planning and cognitive flexibility	307	7.7 (1.07)	urban	healthy	Spanish



Table 2 (continued)

Instrument Reference (authors, year)	Name and Version of Instrument	Outcome Measure Assessed	Country/ Study Sites	Type of Study*: Development, validation	Target Population	Mode of Administration	Sub-scales/ number of items of interest	Sample size	Age range and mean	Local Setting	Condition (clinical/ healthy)	Language(s) of Population
(Willoughby et al., 2019)	EF Touch (computer-based battery) (Willoughby et al., 2019)	EF: inhibition, cognitive flexibility and working memory	Kenya	construct validity (discriminant)	pre-school children	computer-based, interview administered task	5 tasks/subscales: Go/no-go, variant of Stroop, Spatial conflict arrows, Pick the pictures, Something's the same	193	range: 3–6yrs	urban	healthy	Kiswahili, English
(Gonen et al., 2019)	Preschool Self-Regulation Assessment: PRSA (Smith-Donald et al., 2007)	EF in pre-schoolers: self-regulation	Turkey, USA	cross cultural validity, reliability	pre-school children	interview-based task	3 tasks: The Balance Beam, Tower Task and Pencil Tap tasks	Turkey: 471 children + caregivers; USA: 286 children + caregivers	Turkey: 5.2yrs (2.83 to 5.92); US 4.12 yrs (2.97 to 5.07)	urban	healthy	Turkish, English
(Xu et al., 2020)	computerized version of: Stop Signal task (Xu et al., 2020), Figure Matching task, Spatial Span task; Tower of Hanoi task (Simon, 1975)	EF	China, Hong Kong and USA	cross cultural validity	school-age children and adolescents	self-administered computer-based tasks	4 tasks each measuring: inhibition, cognitive flexibility, working memory, planning	China: 453; Hong Kong 371; UK: 487	HK: 12.2yrs, UK: 11.9yrs; China: 11.93yrs	urban and rural	healthy	Mandarin, Cantonese, English
(Mashadi et al., 2020)	Barkley Deficits in Executive Functioning Scale—Children and Adolescents BDEFS-CA (Barkley, 2012)	EF	Iran	structural validity, internal consistency, reliability (test-retest), construct validity (convergent and discriminant)	children and adolescents	parent report/ self-report	5 subscales: Self-Management to Time, Self-Organization/ Problem Solving, Self-Restraint, Self-Motivation and Self-Regulation of Emotion	2295	range: 6–18; mean: 11.6 (3.34)	rural and urban	mostly healthy, 4.7% clinical	Persian
(Arruda et al., 2020)	EFICA (Executive Function Inventory for Children and Adolescents)- parent and teacher versions (Arruda et al., 2020)	EF	Brazil	structural validity, internal consistency, reliability, construct validity (discriminant)	children and adolescents	parent report/ proxy (teacher) report	Not stated	3284 healthy, 165 ADHD*	range 5–18, mean 8.2 (2.0)	rural and urban	ADHD* and healthy	Portuguese
(Tombokan-Rumakahu & Nitko, 1992)	Indonesian-Vietland Adaptive Behaviour Scales VABS (Sparrow et al., 1984, 2005)	adaptive functioning	Indonesia	adaptation (development), internal consistency, reliability (intra-rater and inter rater), cross cultural validity, Construct validity (discriminant)	children	parent report	4 domains: Communication, Daily Living Skills, Socialization, and Maladaptive Behavior	86; 43 ID*, 43 healthy	range: 6- 18; mean: 11.8 (2.9)	semi-urban	Intellectual disability (ID), healthy	Indonesian
(Goldberg et al., 2009)	VABS- Vietnam version	adaptive functioning	Vietnam	adaptation, structural validity, internal consistency, construct validity (discriminant)	pre-school children	parent report	4 domains: as above for VABS	120 healthy, 31 ID* children	range: 3–6yrs; mean 4.9 (1.1)	urban	intellectual disability and healthy	Vietnamese
(Selvam et al., 2016)	VABS II- Indian Version	adaptive functioning	India	cross cultural validity (discriminant)	pre-school children	parent report	4 domains: as above for VABS	412	2–6 years,	urban	healthy	Kannada
(Amiri et al., 2016)	Iranian Children's Participation Assessment Scale CPAS (Amiri et al., 2016)	Adaptive functioning	Iran	Scale development	Children	parent report/ self-report	8 subscales: ADL, ADL- Play, Leisure, Social Participation, education, Work, Rest/Sleep	40 kids, 21 parents	11.5 (2.8); range 6–18 years	urban	healthy	Persian

Table 2 (continued)

Instrument Reference (authors, year)	Name and Version of Instrument	Outcome Measure Assessed	Country/ Study Sites	Type of Study*: Development, validation	Target Population	Mode of Administration	Sub-scales/ number of items of interest	Sample size	Age range and mean	Local Setting	Condition (clinical/ healthy)	Language(s) of Population
(Animi et al., 2017)	Children's Participation Assessment Scale (CPAS)-parent version	adaptive functioning	Iran	structural validity, reliability (test-retest), construct validity (convergent)	school-age children	parent-report	8 subscales: ADL, Instrumental ADL, Play, Leisure, Social Participation, education, Work, Rest/Sleep	700 parents	range: 6–12; mean 9.45 (1.76)	urban	healthy	Persian
(Munir et al., 1999)	Independent Behaviour Assessment Scale IBAS (Munir et al., 1999)	adaptive functioning	Bangladesh	reliability, internal consistency, construct validity (discriminant)	pre-school and school-age children	parent report, interview-based tasks	4 subscales: Motor skills, Socialisation skills, Communication skills, Daily living skills	1404 healthy	range: 2–9, 222 clinical	rural and urban	healthy and clinical	Bengali
(Holding & Kúsaio-Wékulo, 2009)	Participation in Activities of Daily Living PADL (Holding & Kúsaio-Wékulo, 2009)	adaptive functioning	Kenya	development, internal consistency, construct validity (discriminant)	children	interview-based questionnaire	2 subscales: participation and limitations to participation	study development: 92, validation: 116	range: 6–18; mean: 11.7 (2.9)	rural	assorted chronic conditions (Sickle Cell, DM <sup>2</sup> , etc.); acute encephalopathy; healthy controls	Kiswahili
(Tol et al., 2011)	Child Function Impairment Rating Scale CFIRS (Tol et al., 2011)	adaptive functioning	Indonesia	development, Structural Validity, internal consistency, reliability (test-retest, inter rater), construct validity (convergent and discriminant)	children	FGD, diaries, interview-administered questionnaire	1 unidimensional scale with 4 factors; 11 items,	development: 53, validation: 403 children, 385 parents	mean: 9.9 (1.21)	rural	healthy but exposed to political violence	Indonesian
(Malkawi et al., 2015)	Arabic Preschool Activity Card Sort A-PACS (Beg & La Vesser, 2006)	adaptive functioning, ADL	Jordan	development	school-age children	parent report	7 domains: self-care, community mobility, high physical demand leisure and low physical demand leisure, social inter-action, domestic and education	115 caregivers	4.8	rural and urban	healthy	Arabic
(Animi et al., 2017)	Iranian- Child Participation Questionnaire I-CPQ (Rosenberg et al., 2010)	adaptive functioning	Iran	Structural validity, internal consistency, reliability, construct validity (convergent)	pre-school children	parent report	6 subscales: self-care, home participation, play, leisure, social participation, educational and environment	120	range: 4–6; mean: 5.2	urban	cerebral palsy	Persian

\*Studies as defined according to COSMIN guidelines as any individual validation conducted in any given research project/paper. Thus, some individual papers reported multiple validation “studies” within that single paper

<sup>a</sup>ADHD- Attention Deficit/Hyperactivity Disorder

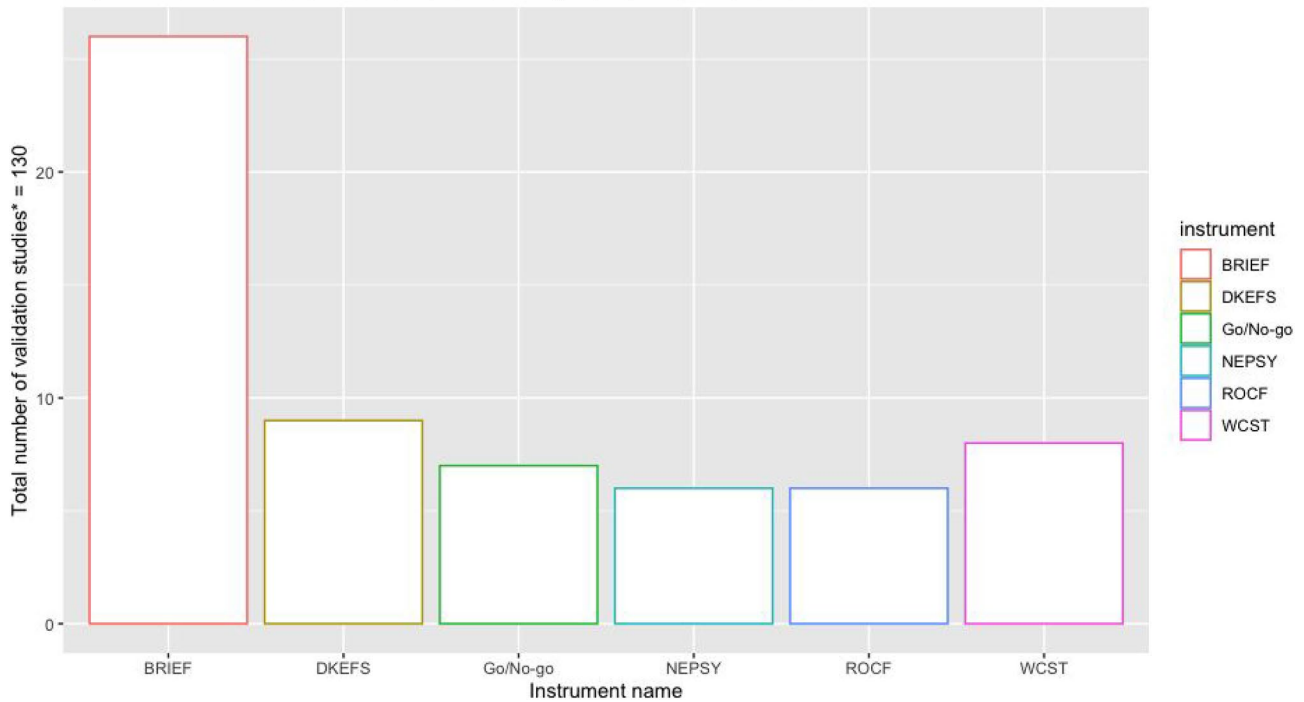
<sup>b</sup>DCD- Development Coordination Disorder

<sup>c</sup>NEPSY- A Developmental Neuropsychological Assessment

<sup>d</sup>ID- Intellectual Disability

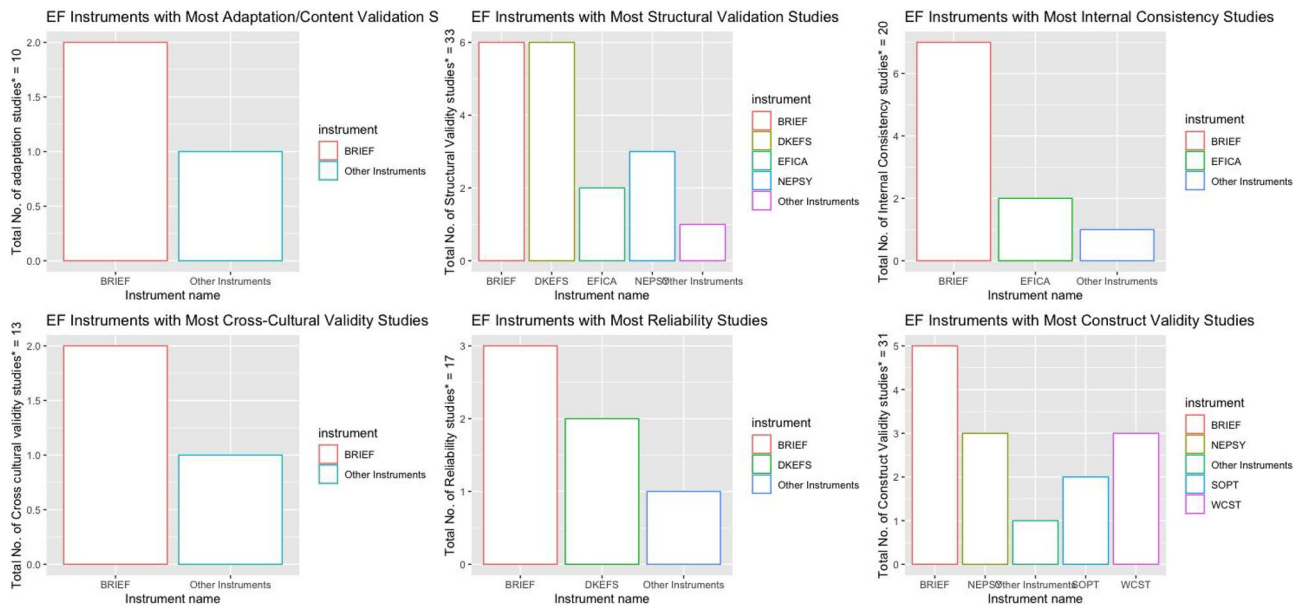
<sup>e</sup>DM- Diabetes Mellitus

Top 5 Executive Function Instruments with Highest Number of Validation Studies



**Fig. 2** Top performing EF instruments in terms of total number of validation studies. \*Studies is defined according to COSMIN guidelines as any individual validation conducted in any given research

project/paper. Thus, some individual papers reported multiple validation “studies” within that single paper

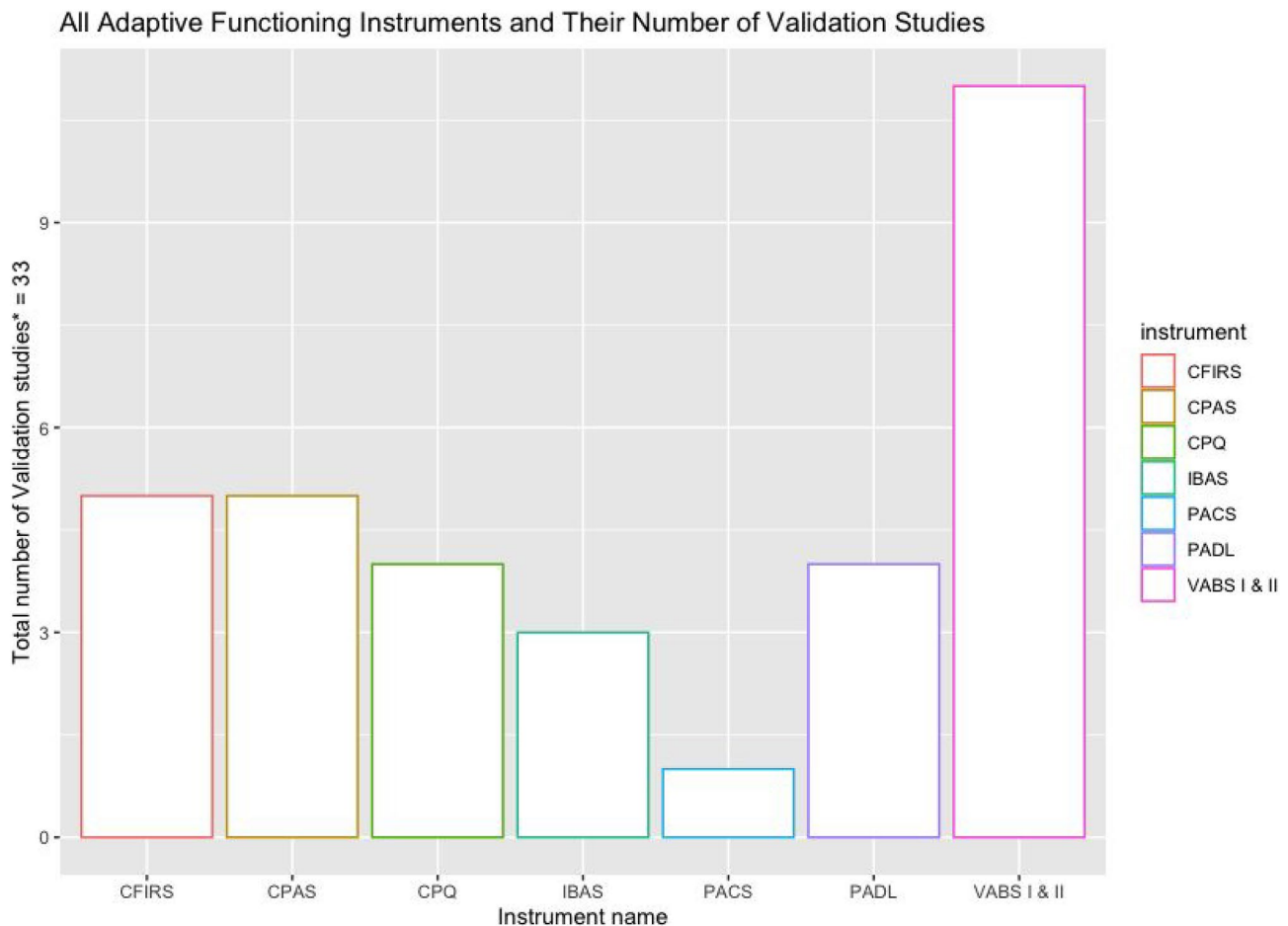


**Fig. 3** Break-down of most validated EF instruments in terms of individual types of Validation studies. \*Studies as defined according to COSMIN guidelines as any individual validation conducted in any

given research project/paper. Thus, some individual papers reported multiple validation “studies” within that single paper

interesting results emerged. Firstly, the BRIEF maintained its status as the most validated instrument regardless of the type of validation- from adaptation studies, through cross-cultural validity studies to construct validity studies. This perhaps speaks to the popularity of the BRIEF even in low-and-middle-income countries. Only the WCST and Go/No-go tests showed a similar consistency of validation with at least 1 validation study in almost all the category types of validation studies although in the graph below this is subsumed under the “other instruments” category (see Table 3 in Appendix III). Apart from BRIEFs and Go/No-go, performance for the other instruments were mostly not consistent across the various types of validation studies. NEPSY for example had all its validation studies in only two categories, including structural validity (3) and construct validity studies (3), while DKEFS featured mostly under structural validity studies (6), reliability studies (2) and Construct validity studies (1) (see Table 3 Appendix III).

Figure 4 is a graph of the most validated adaptive functioning instruments by number of validation studies conducted among children in low-and-middle-income countries. For a full display of results for all validation study types for Adaptive Function instruments used following brain pathology, see Table 4 in Appendix II. A total of 33 individual studies were done for the adaptive function instruments reported in this review. By far, the VABS (in all its various iterations or editions) (Sparrow et al., 1984, 2005) is the most validated instrument for adaptive functioning following brain pathology in children in low-and-middle-income countries with 11 validation studies (33.3%), distantly followed by the CFIRS (Tol et al., 2011) an newly developed tool from Indonesia and the CPAS (Amini et al., 2016) with 5 studies (15.2%) each.



**Fig. 4** Top performing adaptive functioning instruments in terms of total number of validation studies. \*Studies as defined according to COSMIN guidelines as any individual validation conducted in any given research project/paper. Thus, some individual papers reported multiple validation “studies” within that single paper. CFIRS: Child

Function Impairment Rating Scale. CPAS: Children’s Participation Assessment Scale. CPQ: Child Participation Questionnaire. IBAS: Independent Behaviour Assessment Scale. PACS: Preschool Activity Card Sort. PADL: Participation in Activities of Daily Living. VABS: Vineland Adaptive Behaviour Scale

## Discussion

This scoping review was carried out to find out which instruments for assessing executive function and adaptive functioning among children had been validated in low-and-middle-income countries. It also sought to establish which of these instruments stood out in terms of the number and variety of validation studies conducted. This is the first such scoping review in the context of low-and-middle-income countries to the best of the authors' knowledge.

### The Most Validated Instruments

Judging solely by number of validation studies, the BRIEF (Gioia et al., 2000) appears to be the most validated instrument for executive functions among children aged 6 – 18 years in low-and-middle-income countries, while the VABS (Sparrow et al., 1984, 2005) is the most validated for adaptive functioning in a similar context. This finding is corroborated by a similar scoping review of EF instruments used around the world among adolescents where BRIEF was among the top 3 instruments (Nyongesa et al., 2019).

However, the variety of types of validation studies done in non-high-income country was quite limited. None of the instruments found in this present paper had a validation study in all major categories of validation considered (see Table 3 Appendix III). In fact, of the EF tools, apart from the BRIEF (Gioia et al., 2000), WCST (Berg, 1948) and Go/No-go (Luria, 1973) tests, most of the others had validation studies in only 2 to 3 out of the 9 categories of validation studies. Indeed, just 3 types of validation studies- structural validity, construct validity and internal consistency- accounted for a disproportionate 63% of all the validation studies. This showed that validation studies in low-and-middle-income countries was generally skewed towards a few types of studies, highlighting a paucity of studies examining such validation categories as measurement error, responsiveness, content validation and cross-cultural validation. Specific to content validation, unfortunately, relatively few studies focused on adapting or content-validating an existing western-derived instrument (7.9%) or developing a new instrument de novo (3.7%) for the low-and-middle-income country context. In these cases, the implicit but potentially erroneous assumption is that the content of these tools will already be valid in these low-and-middle-income countries (hence making content validation unnecessary), with the focus thus being on other forms of validation (construct, structural etc.). This is noteworthy because one would expect that validating the content of a foreign tool in a new context would be one of the first and most important adaptations to be done before any others are considered (Terwee et al., 2018). Therefore, while many validation studies for EF and

adaptive function instruments have been conducted in low-and-middle-income country contexts, a much wider variety of studies is needed, particularly content validity studies.

Considering the distribution of validation studies in low-and-middle-income countries by world regions, the highest performing regions were Latin America (30.4%), sub-Saharan Africa (21.4%) and the Middle East (17.9%). This is a telling finding because of the implications it has for imposing language limitations in the methodology of such scoping reviews, even if it is for understandable reasons of resource constraints. Apparently, work is progressing in countries like Brazil, Argentina and Iran (the 3 top performing low-and-middle-income countries by number of validation papers). These are non-English speaking countries and are likely to be over-looked by English-based scoping or systematic reviews sadly. Thus, even though in this present study we were also unable to translate and thus include several non-English language papers (see Appendix II for list of such potentially eligible papers that were not included for lack of translation), we did not impose any language restrictions to our search a priori and made the effort to translate as many as we could. Given the performance of these non-English countries or regions, one wonders how skewed the results might have been had such language restrictions been imposed.

Another interesting observation was that Informant-based instruments (either by self, parents or another proxy) were slightly more frequent at 53.7% compared to performance-based measures at 46.3%, which was in contrast to the findings by Nyongesa and colleagues (Nyongesa et al., 2019). This might be because of differences in focus and methodology of the two papers. We focused exclusively on low-and-middle-income countries but also looked broadly at children and adolescents aged 5 – 18 years, while they focused more broadly on all countries (including high-income countries) but looked specifically at adolescents aged 13 – 17 years. In terms of methodology, we looked at 14 databases and imposed no language or date restrictions, while they searched 3 databases and focused on the last 15 years only and on papers published only in English. Evidently when the search is broadened to include high-income countries, performance-based measures dominate while the reverse is true when the search is limited to low-and-middle-income countries. This may possibly be explained by positing that in studies from high-income countries, performance-based measures are most used and largely in experimental, theoretically driven studies that tend to be conducted in controlled environments where performance-based measures work best. In contrast, in studies from low-and-middle-income countries, informant-based questionnaires tend to be preferred, and mostly used in clinical-based studies because they are probably easier to administer and score in a clinical context, and they tend to have better ecological validity (Gioia et al., 2010; Nyongesa et al., 2019) and thus work best in the clinical context.



Further, most validation studies in low-and-middle-income countries took place in urban centres (76.7%). This is somewhat concerning because of the major socioeconomic disparities and standard of living that exist in low-and-middle-income countries between urban and rural populations. The effects of socioeconomic status (SES) on performance in executive function testing has been well documented in South Africa, Australia (Howard et al., 2020) and the United States (Raver et al., 2013). Given that the majority of people in low-and-middle-income countries still live in poverty in rural areas, a lot more validation studies in low-and-middle-income countries involving rural populations are needed to better reflect the realities on the ground in those countries.

### The Cross-Cultural Conundrum

There were relatively few cross-cultural validations (only 8.6% of studies) across board. This is an important observation because this review focused on low-and-middle-income countries in which most of these instruments were not the original countries or cultures of development. Even the BRIEF (Gioia et al., 2000), the instrument with the most validation studies conducted in low-and-middle-income countries had only 2 cross-cultural validation studies done, with only 11 other EF instruments (out of the 40 reviewed) each having a single cross-cultural validation study done. One would have expected more of such cross-cultural studies to perform robust cross-cultural validation of the instruments used. Lev Vygotsky was among the first to warn about the dangers of simply assuming that the concept implied in an item on any measurement instrument would automatically carry forward to another culture and be understood in the same way when that instrument is used in the recipient culture (Vygotsky, 1986 English translation). He warned that if the target audience did not truly hold the same concept as the host audience in the use of a particular set of words (even if they were “translated” words), they were likely to mishandle the item and therefore produce results that were not truly reflective of whatever concept was being assessed (Vygotsky, 1986 English translation). The implication of these is that an assumption could not simply be made by a cross-cultural researcher *ab initio* that any behaviours (or the concepts underlying them) would be developed at all or developed in the same way as within the culture of origin of the researcher (Vygotsky, 1986).

In recent times, this very point has been reinforced by cross-cultural researchers using modern psychometric techniques by demonstrating that assessment of item bias and measurement invariance of any tool used in a cross-cultural context is a crucial step in the comparison of any two groups, for meaningful conclusions to be made (Fischer & Karl, 2019; Gannotti & Handwerker, 2002). The purpose of evaluating invariance is

to confirm whether the responses of different populations on each item differ by more than chance, or put in other words, whether the properties of an instrument are the same in two different groups. Lack of invariance in two groups means there are systematic differences in the way the two groups answer the same questions (for example, the lack of conceptual equivalence that Vygotsky alluded to). The important implication is that if the performance of an instrument is not comparable across two groups (for example, if the factor structure is quite different), one cannot compare the two groups on the construct that is being measured (e.g., compare their mean scores or correlations), and cannot thus draw conclusions that any perceived differences between the two groups are real differences, rather than them simply being an artefact of the fact that the instrument being used behave differently in the two groups. This therefore highlights the utmost importance of cross-cultural validation studies in low-and-middle-income countries for any western-derived assessment tool, and thus belies the unfortunate lack of these studies demonstrated by this scoping review.

### Strengths and Limitations of this Study

This scoping review was conducted following the rigour of the highly recommended PRISMA- ScR guideline (Tricco et al., 2018), ensuring that it was conducted to the highest of methodological standards and can be easily replicated by other researchers. We did not impose any date restrictions and thus went as far back in time as possible for publications available on the databases. Further, the search strategy used was a very thorough one which included listing by name (including name spelling variants) all low-and-middle-income countries in the search strategy, rather than just relying on terms like “developing country”, and so on. We also searched an exceptionally large number of databases (14 in all) and specifically searched the grey literature and in databases tailored towards low-and-middle-income countries. We also endeavoured to include as many non-English papers as we could translate in the study.

However, a significant limitation was that we failed to translate or obtain up to 27 papers that might have been eligible judging solely from their abstracts. However, judging from the eligibility rate after review of full papers of 8.9%, it can be projected that only about 2–3 eligible full papers (i.e., 8.9% of the 27 papers) would probably have been truly eligible of this list but were missed. This would probably not have significantly changed the overall conclusions made in this paper. We have however listed these potentially missed publications in Appendix II for the sake of transparency. We also did not do a robust risk of bias assessment of the methodology and results of the reported validation studies, which would have aided more definitive conclusions to be drawn. But as mentioned, this was outside the scope of this study.

## Conclusion

Quite a number of validation studies have been published on EF and adaptive function assessment tools among children and adolescents in low-and-middle-income countries, however these are woefully inadequate to cover the scope of validations out there. Particularly concerning is the lack of adaptation, content validity and cross-cultural validity studies for western-derived instruments being used in low-and-middle-income countries, as well as studies on development of instruments purposely for low-and-middle-income countries. EF and adaptive function tools that have either been adapted or developed for low-and-middle-income countries are therefore lacking and much needed. The quality of these validation studies though is outside of the scope of this scoping review paper and will be better explored in a subsequent systematic review paper.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11065-021-09529-w>.

**Acknowledgements** The authors would like to acknowledge with gratitude the contribution of Dr Veronica Philips and Mrs Eleanor Barker, Assistant Librarians at the University of Cambridge Medical Library for their advice and immense help in looking over the search strategy and in training the corresponding author in performing systematic reviews.

**Funding** No direct funding for this project was received. However, the Corresponding author is supported by a Commonwealth Scholarship from the Commonwealth Scholarship Secretariat under Award Number: GHCS-2019–169. The University of Cambridge was the sponsor of this study and assumes overall responsibility for the data. Neither the funders nor the sponsor had any role in the planning, design and reporting of this review.

## Declarations

**Ethics Approval** This scoping review was part of the PhD project of the Corresponding author. Ethical approval for the project protocol was obtained from the Psychology Research Ethics Committee of the University of Cambridge with Ethics number: PRE.2020.146.

**Consent to Participate** Not applicable on account of no human or animal participants involved.

**Consent for Publication** Not applicable on account of no primary data analysed.

**Financial and Non-Financial Interests** The authors have no relevant financial or non-financial interests to declare.

**Conflict of Interest** No conflict to report for any of the authors.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated

otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abubakar, A., Ssewanyana, D., de Vries, P. J., & Newton, C. R. (2016). Autism spectrum disorders in sub-Saharan Africa. *The Lancet. Psychiatry*, 3(9), 800–802. [https://doi.org/10.1016/S2215-0366\(16\)30138-9](https://doi.org/10.1016/S2215-0366(16)30138-9).
- Al-Jawahiri, F., & Nielsen, T. R. (2020). Effects of Acculturation on the Cross-Cultural Neuropsychological Test Battery (CNTB) in a Culturally and Linguistically Diverse Population in Denmark. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists*. <https://doi.org/10.1093/arclin/acz083>
- Alloway, T. P. (2007). *Automated Working: Memory Assessment: Manual*. Pearson.
- Amani, M., Gandomani, R. A., & Nesayan, A. (2018). The reliability and validity of behavior rating inventory of executive functions tool teacher's form among Iranian primary school students. *Iranian Rehabilitation Journal*, 16(1), 25–34. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85040818221&partnerID=40&md5=79338f26243ab34a2ef9c96d9fa6364a>
- Amini, M., Hassani Mehraban, A., Haghani, H., Mollazade, E., & Zaree, M. (2017). Factor Structure and Construct Validity of Children Participation Assessment Scale in Activities Outside of School-Parent Version (CPAS-P). *Occupational Therapy in Health Care*, 31(1), 44–60. <https://doi.org/10.1080/07380577.2016.1272733>.
- Amini, M., Mehraban, A. H., Haghani, H., Asgharnezhad, A. A., Mahani, M. K., Hassani Mehraban, A., & Khayat-zadeh Mahani, M. (2016). Development and validation of Iranian children's participation assessment scale. *Medical Journal of the Islamic Republic of Iran*, 30(1), 333. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem2&NEWS=N&AN=27390703>
- Arruda, M. A., Arruda, R., & Anunciação, L. (2020). Psychometric properties and clinical utility of the executive function inventory for children and adolescents: A large multistage populational study including children with adhd. *Applied Neuropsychology: Child*. <https://doi.org/10.1080/21622965.2020.1726353>
- Bailey, C. E. (2007). Cognitive Accuracy and Intelligent Executive Function in the Brain and in Business. *Annals of the New York Academy of Sciences*, 1118(1), 122–141. <https://doi.org/10.1196/annals.1412.011>.
- Bakar, E. E., Taner, Y. I., Soysal, A. S., Karakas, S., & Turgay, A. (2011). Behavioral rating inventory and laboratory tests measure different aspects of executive functioning in boys: A validity study [Davranış derecelendirme envanteri ve laboratuvar testleri erkek çocuklarda yönetici fonksiyonların farklı yönlerini ölçmektedir. *Klinik Psikofarmakoloji Bulteni*, 21(4), 302–316. <https://doi.org/10.5455/BCP.20111004014003>
- Bakare, M. O., Munir, K. M., & Bello-Mojeed, M. A. (2014). Public health and research funding for childhood neurodevelopmental disorders in Sub-Saharan Africa: a time to balance priorities. *Healthcare in Low-Resource Settings*, 2(1). <https://doi.org/10.4081/hls.2014.1559>
- Barkley, R. A. (2012). *Barkley deficits in executive functioning scale—children and adolescents (BDEFS-CA)*. Guilford Press.
- Barreto, L. C. R., Pulido, J. H. P., Torres, J. D. C., & Estupinan, G. P. F. (2018). Psychometric properties and standardization of the

- ENFEN test in rural and urban areas of Tunja city (Colombia). *Diversitas-Perspectivas En Psicología*, 14(2), 339–350. <https://doi.org/10.15332/s1794-9998.2018.0002.10>
- Bathelt, J., Holmes, J., Astle, D. E., Holmes, J., Gathercole, S., Astle, D., et al. (2018). Data-Driven Subtyping of Executive Function-Related Behavioral Problems in Children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 57(4), 252–262.e4. <https://doi.org/10.1016/j.jaac.2018.01.014>
- Bennett, P. C., Ong, B., & Ponsford, J. (2005). Measuring executive dysfunction in an acute rehabilitation setting: Using the dys-executive questionnaire (DEX). *Journal of the International Neuropsychological Society*, 11(4), 376–385. <https://doi.org/10.1017/S1355617705050423>
- Berg, C., & LaVesser, P. (2006). The preschool activity card sort. *OTJR: Occupation, Participation and Health*, 26(4), 143–151.
- Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, 39(1), 15–22. <https://doi.org/10.1080/00221309.1948.9918159>
- Bitta, M., Kariuki, S. M., Abubakar, A., & Newton, C. R. J. C. (2017). Burden of neurodevelopmental disorders in low and middle-income countries: A systematic review and meta-analysis. *Wellcome Open Research*, 2, 121. <https://doi.org/10.12688/wellcomeopenres.13540.3>
- Blackburn, H. L., & Benton, A. L. (1957). Revised administration and scoring of the Digit Span Test. *Journal of Consulting Psychology*, 21(2), 139–143. <https://doi.org/10.1037/H0047235>
- Burgess, P. W., & Shallice, T. (1997). *The hayling and brixton tests*.
- Burkey, M. D., Murray, S. M., Bangirana, P., Familiar, I., Opoka, R. O., Nakasujja, N., Boivin, M., & Bass, J. K. (2015). Executive function and attention-deficit/hyperactivity disorder in Ugandan children with perinatal HIV exposure. *Global Mental Health* (Cambridge, England), 2, e4. <https://doi.org/10.1017/gmh.2015.2>
- Chernoff, M. C., Laughton, B., Ratswana, M., Familiar, I., Fairlie, L., Vhembo, T., Kamthunzi, P., Kabugho, E., Joyce, C., Zimmer, B., Ariansen, J. L., Jean-Philippe, P., & Boivin, M. J. (2018). Validity of Neuropsychological Testing in Young African Children Affected by HIV. *Journal of Pediatric Infectious Diseases*, 13(3), 185–201. <https://doi.org/10.1055/s-0038-1637020>
- Cochrane Library. (2012). *LMIC Filters | Cochrane Effective Practice and Organisation of Care*. <https://epoc.cochrane.org/lmic-filters>
- Collins, A., & Koehlin, E. (2012). Reasoning, Learning, and Creativity: Frontal Lobe Function and Human Decision-Making. *PLoS Biology*, 10(3).
- de Bustamante Carim, D., Miranda, M. C., & Bueno, O. F. A. (2012). Tradução e Adaptação para o Português do Behavior Rating Inventory of Executive Function—BRIEF = Translation and adaptation into Portuguese of the Behavior Rating Inventory of Executive Function—BRIEF. *Psicologia: Reflexão e Crítica*, 25(4), 653–661. <https://doi.org/10.1590/S0102-79722012000400004>
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System*.
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Du, Y., Li, M., Jiang, W., Li, Y., & Coghill, D. R. (2018). Developing the Symptoms and Functional Impairment Rating Scale: A Multi-Dimensional ADHD Scale. *Psychiatry Investigation*, 15(1), 13–23. <https://doi.org/10.4306/pi.2018.15.1.13>
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School Readiness and Later Achievement. *Developmental Psychology*, 43(6), 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.1428>
- Ellefsen, M. R., Ng, F.F.-Y., Wang, Q., & Hughes, C. (2017). Efficiency of Executive Function: A Two-Generation Cross-Cultural Comparison of Samples From Hong Kong and the United Kingdom. *Psychological Science*, 28(5), 555–566. <https://doi.org/10.1177/0956797616687812>
- Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Frontiers in Psychology*, 10(JULY), 1–18. <https://doi.org/10.3389/fpsyg.2019.01507>
- Ford, C. B., Kim, H. Y., Brown, L., Aber, J. L., & Sheridan, M. A. (2019). A cognitive assessment tool designed for data collection in the field in low- and middle-income countries. *Research in Comparative and International Education*, 14(1), 141–157. <https://doi.org/10.1177/1745499919829217>
- Gannotti, M. E., & Handwerker, W. P. (2002). Puerto Rican understandings of child disability: methods for the cultural validation of standardized measures of child health. *Social Science & Medicine* (1982), 55(12), 2093–2105. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med4&NEWS=N&AN=12409123>
- García-Barrera, M. A., Karr, J. E., Duran, V., Drenfeld, E., & Pineda, D. A. (2015). Cross-cultural validation of a behavioral screener for executive functions: Guidelines for clinical use among Colombian children with and without ADHD. *Psychological Assessment*, 27(4), 1349–1363. <https://doi.org/10.1037/pas0000117>
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior rating inventory of executive function: BRIEF*.
- Gioia, G. A., Kenworthy, L., & Isquith, P. K. (2010). Executive function in the real world: BRIEF lessons from Mark Ylvisaker. *The Journal of Head Trauma Rehabilitation*, 25(6), 433–439. <https://doi.org/10.1097/HTR.0b013e3181fbc272>
- Goldberg, M. R., Dill, C. A., Shin, J. Y., & Nguyen, V. N. (2009). Reliability and validity of the Vietnamese Vineland Adaptive Behavior Scales with preschool-age children. *Research in Developmental Disabilities*, 30(3), 592–602. <https://doi.org/10.1016/j.ridd.2008.09.001>
- Gonen, M., Guler-Yildiz, T., Ulker-Erdem, A., Garcia, A., Raikes, H., Acar, I. H., et al. (2019). Examining the Association Between Executive Functions and Developmental Domains of Low-Income Children in the United States and Turkey. *Psychological Reports*, 122(1), 155–179. <https://doi.org/10.1177/0033294118756334>
- Green, R., Till, C., Al-Hakeem, H., Cribbie, R., Tellez-Rojo, M. M., Osorio, E., et al. (2019). Assessment of neuropsychological performance in Mexico City youth using the Cambridge Neuropsychological Test Automated Battery (CANTAB). *Journal of Clinical and Experimental Neuropsychology*, 41(3), 246–256. <https://doi.org/10.1080/13803395.2018.1529229>
- Hall, P. A., & Marteau, T. M. (2014). Executive function in the context of chronic disease prevention: Theory, research and practice. *Preventive Medicine*, 68, 44–50. <https://doi.org/10.1016/j.ypmed.2014.07.008>
- Holding, P., Anum, A., van de Vijver, F. J. R., Vokhiwa, M., Bugase, N., Hossen, T., Makasi, C., Baiden, F., Kimute, O., Bangre, O., Hasan, R., Nanga, K., Sefenu, R. P. S., A-Hayat, N., Khan, N., Oduro, A., Rashid, R., Samad, R., Singlovic, J., ... Gomes, M. (2018). Can we measure cognitive constructs consistently within and across cultures? Evidence from a test battery in Bangladesh, Ghana, and Tanzania. *Applied Neuropsychology: Child*, 7(1), 1–13. <https://doi.org/10.1080/21622965.2016.1206823>
- Holding, P., & Kitsao-Wekulo, P. (2009). Is assessing participation in daily activities a suitable approach for measuring the impact of disease on child development in African children? *Journal of Child and Adolescent Mental Health*, 21(2), 127–138. <https://doi.org/10.2989/JCAMH.2009.21.2.4.1012>
- Howard, S. J., Cook, C. J., Everts, L., Melhuish, E., Scerif, G., Norris, S., Twine, R., Kahn, K., & Draper, C. E. (2020). Challenging socioeconomic status: A cross-cultural comparison of early executive function. *Developmental Science*, 23(1), e12854. <https://doi.org/10.1111/desc.12854>
- Huppert, F. A., Brayne, C., Gill, C., Paykel, E. S., & Beardsall, L. (1995). CAMCOG—a concise neuropsychological test to assist dementia diagnosis: Socio-demographic determinants in an elderly population sample. *British Journal of Clinical Psychology*, 34(4), 529–541.
- Injoque-Ricle, I., Calero, A. D., Alloway, T. P., & Burin, D. I. (2011). Assessing working memory in Spanish-speaking children:



- Automated Working Memory Assessment battery adaptation. *Learning and Individual Differences*, 21(1), 78–84. <https://doi.org/10.1016/j.lindif.2010.09.012>.
- Karbach, J., & Unger, K. (2014). Executive control training from middle childhood to adolescence. In *Frontiers in Psychology* (Vol. 5, Issue MAY, p. 390). Frontiers Research Foundation. <https://doi.org/10.3389/fpsyg.2014.00390>
- Kashala, E., Tylleskar, T., Elgen, I., & Kayembe, K. (2005). Attention deficit and hyperactivity disorder among school children in Kinshasa, Democratic Republic of Congo. *African Health*. <https://www.ajol.info/index.php/ahs/article/view/7014>
- Kieling, C., Baker-Henningham, H., Belfer, M., Conti, G., Ertem, I., Omigbodun, O. O., et al. (2011). Child and adolescent mental health worldwide: Evidence for action. *The Lancet*, 378(9801), 1515–1525. [https://doi.org/10.1016/S0140-6736\(11\)60827-1](https://doi.org/10.1016/S0140-6736(11)60827-1).
- Korkman, M. (1998). NEPSY. A developmental neuropsychological assessment. *Test Materials and Manual*.
- Korzeniowski, C., & Ison, M. (2019). Executive functioning scale for schoolchildren: An analysis of psychometric properties [Escala de Funcionamiento Ejecutivo para Escolares: Análisis de las Propiedades Psicométricas]. *Psicología Educativa*, 25(2), 147–157. <https://doi.org/10.5093/psed2019a4>.
- Law, M. (2002). Participation in the occupations of everyday life. *American Journal of Occupational Therapy*, 56(6), 640–649.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21(1), 59–80. <https://doi.org/10.1348/026151003321164627>.
- Lezak, M. D. (1983). *Neuropsychology Assessment* (2nd ed.). Oxford University Press. [https://books.google.co.uk/books?hl=en&lr=&id=FroDVkVKA2EC&oi=fnd&pg=PA3&dq=Lezak,+M.+D.+\(1983\).+&ots=q61fZPsp7T&sig=OdtXkNBRaoXOAsZLbi9fU-mEIPk&redir\\_esc=y#v=onepage&q=Lezak%2C+M.+D.+\(1983\).&f=false](https://books.google.co.uk/books?hl=en&lr=&id=FroDVkVKA2EC&oi=fnd&pg=PA3&dq=Lezak,+M.+D.+(1983).+&ots=q61fZPsp7T&sig=OdtXkNBRaoXOAsZLbi9fU-mEIPk&redir_esc=y#v=onepage&q=Lezak%2C+M.+D.+(1983).&f=false)
- Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment*. Oxford University Press.
- Lunt, L., Bramham, J., Morris, R. G., Bullock, P. R., Selway, R. P., Xenitidis, K., & David, A. S. (2012). Prefrontal cortex dysfunction and ‘Jumping to Conclusions’: Bias or deficit? *Journal of Neuropsychology*, 6(1), 65–78. <https://doi.org/10.1111/j.1748-6653.2011.02005.x>.
- Luria, A. R. (1973). *Psychophysiology of the frontal lobes*. Academic Press.
- Malek, A., Hekmati, I., Amiri, S., Pirezadeh, J., & Gholizadeh, H. (2013). The standardization of Victoria Stroop color-word test among Iranian bilingual adolescents. *Archives of Iranian Medicine*, 16(7), 380–384. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84879381324&partnerID=40&md5=30ac419347b132927ea9a155b6bf3468>
- Malkawi, S. H., Hamed, R. T., Abu-Dahab, S. M. N., AlHeresh, R. A., & Holm, M. B. (2015). Development of the Arabic Version of the Preschool Activity Card Sort (A-PACS). *Child: Care, Health and Development*, 41(4), 559–568. <https://doi.org/10.1111/cch.12209>
- Mashhadi, A., Maleki, Z. H., Hasani, J., & Rasoolzadeh Tabatabaei, S. K. (2020). Psychometric properties of Persian version of the Barkley deficits in executive functioning scale—children and adolescents. *Applied Neuropsychology: Child*. <https://doi.org/10.1080/21622965.2020.1726352>
- Matson, J. L., Rivet, T. T., Fodstad, J. C., Dempsey, T., & Boisjoli, J. A. (2009). Examination of adaptive behavior differences in adults with autism spectrum disorders and intellectual disability. *Research in Developmental Disabilities*, 30(6), 1317–1325. <https://doi.org/10.1016/j.ridd.2009.05.008>.
- Merikangas, K. R. K., Nakamura, E. F. E., & Kessler, R. C. (2009). Epidemiology of mental disorders in children and adolescents. *Dialogues in Clinical Neuroscience*, 11(1), 7–20. <http://www.ncbi.nlm.nih.gov/pubmed/19432384>
- Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions. *Current Directions in Psychological Science*, 21(1), 8–14. <https://doi.org/10.1177/09637214111429458>.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>.
- Mokkink, L. B., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2018). *COSMIN manual for systematic reviews of PROMs COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) user manual*. [www.cosmin.nl](http://www.cosmin.nl)
- Morgan, P. L., Li, H., Farkas, G., Cook, M., Pun, W. H., & Hillemeier, M. M. (2017). Executive functioning deficits increase kindergarten children’s risk for reading and mathematics difficulties in first grade. *Contemporary Educational Psychology*, 50, 23–32. <https://doi.org/10.1016/j.cedpsych.2016.01.004>.
- Munir, S. Z., Zaman, S., & McConachie, H. (1999). Development of an Independent Behaviour Assessment Scale for Bangladesh. *Journal of Applied Research in Intellectual Disabilities*, 12(3), 241–252. <https://doi.org/10.1111/j.1468-3148.1999.tb00080.x>.
- Nampijja, M., Apule, B., Lule, S., Akurut, H., Muhangi, L., Elliott, A. M., Alcock, K. J., M., N., B., A., S., L., H., A., L., M., A.M., E., Nampijja, M., Apule, B., Lule, S., Akurut, H., Muhangi, L., Elliott, A. M., & Alcock, K. J. (2010). Adaptation of western measures of cognition for assessing 5-year-old semi-urban Ugandan children. *British Journal of Educational Psychology*, 80(1), 15–30. <https://doi.org/10.1348/000709909X460600>
- Nyongesa, M. K., Ssewanyana, D., Mutua, A. M., Chongwo, E., Scerif, G., Newton, C. R., & Abubakar, A. (2019). Assessing Executive Function in Adolescence: A Scoping Review of Existing Measures and Their Psychometric Robustness. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00311>
- Ostrosky-Solis, F., Ramírez, M., Lozano, A., Picasso, H., & Vélez, A. (2004). Culture or education? Neuropsychological test performance of a Maya indigenous population. *International Journal of Psychology*, 39(1), 36–46. <https://doi.org/10.1080/00207590344000277>.
- Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, 20(3), 249–262. [https://doi.org/10.1016/0028-3932\(82\)90100-2](https://doi.org/10.1016/0028-3932(82)90100-2).
- Pineda, D. A., Puerta, I. C., Aguirre, D. C., Garcia-Barrera, M. A., & Kamphaus, R. W. (2007). The role of neuropsychologic tests in the diagnosis of attention deficit hyperactivity disorder. *Pediatric Neurology*, 36(6), 373–381. <https://doi.org/10.1016/j.pediatrneuol.2007.02.002>.
- Pluck, G., Amraoui, D., & Fornell-Villalobos, I. (2019). Brief communication: Reliability of the D-KEFS Tower Test in samples of children and adolescents in Ecuador. *Applied Neuropsychology: Child*, 1–7. <https://doi.org/10.1080/21622965.2019.1629922>
- Portellano, J. A., Martínez, R., & Zumárraga, L. (2009). *ENFEN: Evaluación Neuropsicológica de las funciones ejecutivas en niños*. TEA Ediciones.
- Poulsen, A. A., & Ziviani, J. M. (2006). *Children’s participation beyond the school grounds*.
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147–1157. <https://doi.org/10.1007/s1136-018-1798-3>.

- Raver, C. C., Blair, C., & Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: new perspectives on models of differential susceptibility. *Developmental Psychology*, 49(2), 292–304. <https://doi.org/10.1037/a0028343>
- Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique. (Les problems.). [The psychological examination in cases of traumatic encephalopathy. Problems.]. *Archives De Psychologie*, 28, 215–285.
- Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System for Children: Manual*. American Guidance Service, Inc. <https://www.sciencedirect.com/science/article/pii/S002244059490037X?via%3Dihub>
- Richard's, M. M., Introzzi, I., Zamora, E., & Vernucci, S. (2017). Analysis of internal and external validity criteria for a computerized visual search task: A pilot study. *Applied Neuropsychology: Child*, 6(2), 110–119. <https://doi.org/10.1080/21622965.2015.1083433>.
- Richard's, M. M., Vernucci, S., Stelzer, F., Introzzi, I., & Guàrdia-Olmos, J. (2018). Exploratory data analysis of executive functions in children: A new assessment battery. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, 1–8 <https://doi.org/10.1007/s12144-018-9860-4>
- Rincon Diaz, M. A., & Rey Anaconda, C. A. (2017). Adaptation and evaluation of the psychometric properties of brief-p in colombian preschoolers. *Revista Ces Psicología*, 10(1), 48–62. <https://doi.org/10.21615/cesp.10.1.4>
- Rosenberg, L., Jarus, T., & Bart, O. (2010). Development and initial validation of the Children Participation Questionnaire (CPQ). <https://doi.org/10.3109/09638281003611086>, 32(20), 1633–1644. <https://doi.org/10.3109/09638281003611086>
- Rosete, H. S. (2018). Normative data on a neuropsychological screening instrument for school-aged adolescents in Chiang Mai, Thailand [ProQuest Information & Learning]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering* (Vol. 79, Issues 1-B(E)). <https://ezp.lib.cam.ac.uk/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2017-54456-296&site=ehost-live&scope=site>
- Rosetti, M. F., Ulloa, R. E., Reyes-Zamorano, E., Palacios-Cruz, L., de la Peña, F., & Hudson, R. (2018). A novel experimental paradigm to evaluate children and adolescents diagnosed with attention-deficit/hyperactivity disorder: Comparison with two standard neuropsychological methods. *Journal of Clinical and Experimental Neuropsychology*, 40(6), 576–585. <https://doi.org/10.1080/13803395.2017.1393501>.
- Rosetti, M. F., Ulloa, R. E., Vargas-Vargas, I. L., Reyes-Zamorano, E., Palacios-Cruz, L., de la Peña, F., Larralde, H., & Hudson, R. (2016). Evaluation of children with ADHD on the Ball-Search Field Task. *Scientific Reports* 6 <https://doi.org/10.1038/srep19664>
- Ruffieux, N., Njamnshi, A. K., Mayer, E., Sztajzel, R., Eta, S. C., Doh, R. F., et al. (2010). Neuropsychology in cameroon: First normative data for cognitive tests among school-aged children. *Child Neuropsychology*, 16(1), 1–19. <https://doi.org/10.1080/09297040902802932>.
- Sallum, I., da Mata, F. G., Cheib, N. F., Mathias, C. W., Miranda, D. M., & Malloy-Diniz, L. F. (2017). Development of a version of the self-ordered pointing task: A working memory task for Brazilian preschoolers. *Clinical Neuropsychologist*, 31(2), 459–470. <https://doi.org/10.1080/13854046.2016.1275818>.
- Sartori, R. F., Valentini, N. C., Nobre, G. C., & Fonseca R. P. (2020). Motor and verbal inhibitory control: Development and validity of the go/No-Go app test for children with development coordination disorder. *Applied Neuropsychology-Child*, 1–10 <https://doi.org/10.1080/21622965.2020.1726178>
- Selvam, S., Thomas, T., Shetty, P., Thennarasu, K., Raman, V., Khanna, D., Mehra, R., Kurpad, A. V., & Srinivasan, K. (2018). Development of norms for executive functions in typically-developing Indian urban preschool children and its association with nutritional status. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, 24(2), 226–246. <https://doi.org/10.1080/09297049.2016.1254761>
- Selvam, S., Thomas, T., Shetty, P., Zhu, J., Raman, V., Khanna, D., et al. (2016). Norms for developmental milestones using VABS-II and association with anthropometric measures among apparently healthy urban Indian preschool children. *Psychological Assessment*, 28(12), 1634–1645. <https://doi.org/10.1037/pas0000295>.
- Senturk, N., Yeniceri, N., Alp, I. E., & Altan-Atalay, A. (2014). An Exploratory Study on the Junior Brixton Spatial Rule Attainment Test in 6-to 8-Year-Olds. *Journal of Psychoeducational Assessment*, 32(2), 123–132. <https://doi.org/10.1177/0734282913490917>.
- Shallice, T., Marzocchi, G. M., Coser, S., Del Savio, M., Meuter, R. F., & Rumiati, R. I. (2002). Executive function profile of children with attention deficit hyperactivity disorder. *Developmental Neuropsychology*, 21(1), 43–71.
- Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: Elaboration and explanation. *BMJ (online)*, 349(January), 1–25. <https://doi.org/10.1136/bmj.g7647>.
- Simblett, S. K., Badham, R., Greening, K., Adlam, A., Ring, H., & Bateman, A. (2012). Validating independent ratings of executive functioning following acquired brain injury using Rasch analysis. *Neuropsychological Rehabilitation*, 22(6), 874–889. <https://doi.org/10.1080/09602011.2012.703956>.
- Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7(2), 268–288.
- Siqueira, L. S., Gonçalves, H. A., Hübner, L. C., & Fonseca, R. P. (2016). Development of the Brazilian version of the child Hayling test [Desenvolvimento da versão brasileira do Teste Hayling Infantil]. *Trends in Psychiatry and Psychotherapy*, 38(3), 164–174. <https://doi.org/10.1590/2237-6089-2016-0019>.
- Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, 22(2), 173–187. <https://doi.org/10.1016/j.ecresq.2007.01.002>.
- Sobeh, J., & Spijkers, W. (2012). Development of attention functions in 5-to 11-year-old Arab children as measured by the German Test Battery of Attention Performance (KITAP): A pilot study from Syria. *Child Neuropsychology*, 18(2), 144–167. <https://doi.org/10.1080/09297049.2011.594426>.
- Sparrow, S. S., Balla, D. A., Cicchetti, D. V., & Harrison, P. L. (1984). *Vineland adaptive behavior scales*.
- Sparrow, S. S., Cicchetti, D. V., Balla, D. A., & Doll, E. A. (2005). *Vineland adaptive behavior scales: Survey forms manual*. American Guidance Service.
- St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology*, 59(4), 745–759. <https://doi.org/10.1080/17470210500162854>.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643.
- Stuss, D. T. (1983). Frontal lobe lesions and behavior localization. In A. Kertesz (Ed.), *Localization in neuropsychology* (pp. 429–454). Academic Press.
- Stuss, D. T. (2011). Functions of the frontal lobes: Relation to executive functions. *Journal of the International Neuropsychological Society*, 17(5), 759–765. <https://doi.org/10.1017/S135561771000695>.
- Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research*



- Psychologische Forschung*, 63(3–4), 289–298. <https://doi.org/10.1007/s004269900007>.
- Taşcu, A., Iliescu, D., & Avram, E. (2012). Adapting and implementing the Wisconsin Card Sorting Test in Romania. A look at the clinical data. *Procedia - Social and Behavioral Sciences*, 33, 1022–1026. <https://doi.org/10.1016/j.sbspro.2012.01.277>.
- Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. In *Quality of Life Research* (Vol. 27, Issue 5, pp. 1159–1170). Springer International Publishing. <https://doi.org/10.1007/s11136-018-1829-0>
- Thorell, L. B., Lazarević, N., Milovanović, I., Bugarski Ignjatović, V., Lazarević, N., Milovanovic, I., & Ignjatovic, V. B. (2020). Psychometric properties of the Teenage Executive Functioning Inventory (TEXI): A freely available questionnaire for assessing deficits in working memory and inhibition among adolescents. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, 26(6), 1–8. <https://doi.org/10.1080/09297049.2020.1726885>.
- Thorell, L. B., & Nyberg, L. (2008). The Childhood Executive Functioning Inventory (CHEXI): A new rating instrument for parents and teachers. *Developmental Neuropsychology*, 33(4), 536–552. <https://doi.org/10.1080/87565640802101516>.
- Tol, W. A., Komprou, I. H., Jordans, M. J. D., & Susanty, D. (2011). Developing a function impairment measure for children affected by political violence: A mixed methods approach in Indonesia. *International Journal for Quality in Health Care*, 23(4), 375–383. <https://doi.org/10.1093/intqhc/mzr032>
- Tombakan-Runtukahu, J., & Nitko, A. J. (1992). Translation, cultural adjustment, and validation of a measure of adaptive behavior. *Research in Developmental Disabilities*, 13(5), 481–501. [https://doi.org/10.1016/0891-4222\(92\)90004-P](https://doi.org/10.1016/0891-4222(92)90004-P).
- Trevisan, B. T., Dias, N. M., Berberian, A. A., & Seabra, A. G. (2017). Childhood executive functioning inventory: Adaptation and psychometric properties of the Brazilian version [Childhood executive functioning inventory: Adaptação e propriedades psicométricas da versão Brasileira]. *Psico-USF*, 22(1), 63–74. <https://doi.org/10.1590/1413-82712017220106>.
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>.
- Tseng, W. L., & Gau, S. S. F. (2013). Executive function as a mediator in the link between attention-deficit/hyperactivity disorder and social problems. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 54(9), 996–1004. <https://doi.org/10.1111/jcpp.12072>.
- Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, ed.). Cambridge, ma: mit Press.
- Willoughby, M. T. T., Piper, B., Kwayumba, D., & McCune, M. (2019). Measuring executive function skills in young children in Kenya. *Child Neuropsychology*, 25(4), 425–444. <https://doi.org/10.1080/09297049.2018.1486395>.
- Wilson, B. A., Alderman, N., Burgess, P. W., Emslie, H., & Evans, J. J. (1996). *BADS: Behavioural assessment of the dysexecutive syndrome*. Pearson.
- Wong, A., Rodriguez, M., Quevedo, L., Fernandez de Cossio, L., Borges, A., Reyes, A., Corral, R., Blanco, F., & Alvarez, M. (2012). Questionnaire of executive function for dancers: an ecological approach. *Assessment*, 19(3), 383–387. <https://doi.org/10.1177/1073191110397483>
- World Bank Group. (2019). World Bank Country and Lending Groups. World Bank Group. <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
- World Health Organization. (2001). *International classification of functioning, disability and health: ICF*. World Health Organization.
- Xu, C., Ellefson, M. R., Ng, F. F.-Y., Wang, Q., & Hughes, C. (2020). An East-West contrast in executive function: Measurement invariance of computerized tasks in school-aged children and adolescents. *Journal of Experimental Child Psychology*, 199, 104929. <https://doi.org/10.1016/j.jecp.2020.104929>
- Yang, Z.-Y., Sun, S.-F., Lui, S. S. Y., Shi, H.-S., Xie, D.-J., & Xie, W.-L., et al. (2018). An attempt at revisiting the factor structure of the Dysexecutive Questionnaire in the Chinese setting. *Psych Journal*, 7(1), 25–30. <https://doi.org/10.1002/pchj.198>.
- Zarrabi, M., Shahrivar, Z., Tehrani Doost, M., Khademi, M., & Zargari Nejad, G. (2015). Concurrent Validity of the Behavior Rating Inventory of Executive Function in Children With Attention Deficit Hyperactivity Disorder. *Iranian Journal of Psychiatry and Behavioral Sciences*, 9(1), e213–e213. <https://doi.org/10.17795/ijpbs213>
- Zimmermann, P., Fimm, B., & Gondan, M. (2002). *Testbatterie zur aufmerksamkeitsprüfung für kinder:(KiTAP);[das schloß der geister]*. Psytest.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.