

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.



BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmjjournals.org/>).

If you have any questions on BMJ Open's open peer review process please email
info.bmjopen@bmj.com

BMJ Open

Psychometric properties of early childhood development assessment tools in low-and-middle-income countries: a systematic review

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2024-096365
Article Type:	Original research
Date Submitted by the Author:	09-Nov-2024
Complete List of Authors:	Bliznashka, Lilia; International Food Policy Research Institute, Nutrition, Diets, and Health Unit; University of Edinburgh, Global Academy of Agriculture and Food Systems Hentschel, Elizabeth; Harvard T H Chan School of Public Health, Department of Global Health and Population Ali, Nazia Binte; Harvard T H Chan School of Public Health, Department of Global Health and Population Hunt, Xanthe; Stellenbosch University, Department of Global Health; Africa Health Research Institute (AHRI) Neville, Sarah ; Brown University School of Public Health, International Health Institute Olney, D.; International Food Policy Research Institute, Nutrition, Diets, and Health Unit Pitchik, Helen; University of California Berkeley, School of Public Health Roy, Aditi; Centre for Chronic Disease Control Seiden, Jonathan; Harvard Graduate School of Education Solís-Cordero, Katherine ; University of Costa Rica, School of Nursing Thapa, Aradhana; Emory University Hubert Department of Global Health Jeong, Joshua; Emory University Hubert Department of Global Health
Keywords:	PAEDIATRICS, Child, Community child health < PAEDIATRICS, Patient Reported Outcome Measures, Psychometrics

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Psychometric properties of early childhood development assessment tools in low-and-middle-income countries: a systematic review

Lilia Bliznashka, PhD^{1,2,†}, Elizabeth Hentschel, PhD³, Nazia Binte Ali, MPH³, Xanthe Hunt, PhD^{4,5}, Sarah Elizabeth Neville, PhD⁶, Deanna Olney, PhD¹, Helen O. Pitchik, PhD⁷, Aditi Roy, PhD⁸, Jonathan Seiden, EdM⁹, Katherine Solís Cordero, PhD¹⁰, Aradhana Thapa, MPH¹¹, Joshua Jeong, ScD¹¹

¹ International Food Policy Research Institute, Washington, DC, 20001

² Global Academy of Agriculture and Food Systems, University of Edinburgh, UK

³ Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA

⁴ Institute for Life Course Health Research, Department of Global Health, Stellenbosch University, South Africa

⁵ Africa Health Research Institute (AHRI), Somkele, South Africa

⁶ Brown University School of Public Health, International Health Institute, Providence, RI, USA

⁷ Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, California, USA

⁸ Centre for Chronic Disease Control, New Delhi, India

⁹ Harvard Graduate School of Education, Cambridge, MA, USA

¹⁰ School of Nursing, University of Costa Rica, San José, Costa Rica

¹⁰ Hubert Department of Global Health, Emory University, Atlanta, GA, USA

1
2
3 † After the first, second, and last authors, all authors are listed in alphabetical order by last name.
4
5
6

7 **Corresponding author:**
8
9

10 Lilia Bliznashka
11
12 1.lbliznashka@cgiar.org
13 1201 I St NW, Washington, DC 20001, US
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1 **Abstract**

2 **Background:** Valid and reliable measurement of early childhood development (ECD) is critical
3 for monitoring and evaluating ECD-related policies and programmes.

4 **Objective:** To systematically appraised available evidence on the psychometric properties of tools
5 used to measure ECD in low- and middle-income countries (LMICs).

6 **Design:** We searched MEDLINE, Embase, PubMed, PsycInfo, SciELO, and BVS. We included
7 studies that examined the reliability, validity, and measurement invariance of tools assessing ECD
8 in children 0-6 years of age living in LMICs. Each study was independently screened by two
9 researchers and data extracted by one randomly assigned researcher.

10 **Results:** A total of 153 articles covering 114 tools met inclusion criteria. Most reported
11 psychometric properties were internal consistency reliability ($n=114$, 65%), concurrent validity
12 ($n=78$, 45%), test-retest reliability ($n=72$, 41%), convergent validity ($n=72$, 41%), and structural
13 validity ($n=69$, 39%). Measurement invariance was least commonly reported ($n=17$, 10%). Most
14 articles came from Brazil, China, India, and South Africa. Most psychometric evidence was from
15 urban ($n=89$, 51%) or urban-rural ($n=39$, 22%) contexts. Study samples focused on children aged
16 6-17.9 or 48-59.9 months. The most assessed developmental domain was language ($n=105$, 60%),
17 followed by motor ($n=98$, 56%) and cognitive ($n=77$, 44%).

18 **Conclusions:** Psychometric evidence is fragmented, limited, and heterogeneous. More rigorous
19 psychometric analyses, especially on measurement invariance, are needed to establish the quality
20 and accuracy of ECD tools for use in LMICs.

21 **PROSPERO registration:** CRD42022372305

23 Key messages

24 What is already known on this topic

- A critical aspect to monitoring and evaluating interventions and policies to improve early childhood development (ECD) in children <6 years of age in low- and middle-income countries (LMICs) is the availability of reliable and valid assessment tools.
 - Existing literature largely focuses on individual ECD tools, ECD domains, or psychometric properties.

30 What this study adds

- This systematic review provides a comprehensive systematic appraisal of available psychometric evidence of ECD tools used in LMICs.
 - The analysis is unique in that it disaggregates articles by those developing or creating new ECD tools and those using or adapting existing tools. It further provides the first disaggregation by urban and rural setting.
 - The analysis additionally looks at psychometric evidence for specific child age groups to provide a more in-depth understanding of adequate tools to measure ECD in the early years.

38 How this study might affect research, practice or policy

- The findings highlight the need for rigorous research and psychometric analyses in diverse child populations to ensure improved population level assessment and monitoring and evaluation of ECD policies and programs.

43 INTRODUCTION

44 The Sustainable Development Goals (SDGs) recognise the importance of early childhood
45 development (ECD),¹ which sets the foundation for children's later learning and economic
46 outcomes.²⁻⁴ Rigorous ECD measurement is critical for accurate comparability of children's skills
47 across populations and time, evaluating the effectiveness of ECD interventions, and tracking ECD-
48 related policies.

49 In low-and-middle-income countries (LMICs), >140 tools (60% of which originated in
50 high-income countries (HICs)) have been used to assess ECD in children 0-8 years old.⁵ Directly
51 applying ECD tools from HICs in LMICs can be problematic without psychometric evidence for
52 new cultures and contexts. Few studies to-date have synthesised the evidence on the psychometric
53 properties of ECD tools.

54 Evidence pertaining to a tool's reliability, validity, and measurement invariance in a given
55 context is critical for selecting an ECD outcome or indicator. Other factors include the purpose of
56 measurement, the population and age-range of interest, and the developmental domain(s) of
57 interest.⁶ Evidence of reliability and validity ensures consistent and accurate ECD measurement,⁷
58 whereas evidence of measurement invariance guarantees assessment of the same construct across
59 countries, sub-groups, and time.^{7,8}

60 Prior reviews have provided guidance for selecting ECD tools for use in LMICs^{5,9-11} and
61 underscored that evidence on tool reliability and validity is fundamental.^{9,10} However, prior work
62 has often focused on individual ECD domains,¹²⁻¹⁴ which has limited use for population level
63 assessment or tracking of SDG-related policies, or a subset of psychometric properties,¹⁰ which is
64 not evidence of reliability and validity as a whole. Further, prior reviews do not disaggregate use
65 of ECD tools by HICs vs LMICs or urban vs rural settings,¹²⁻¹⁴ despite known and persistent

1
2
3 66 disparities in ECD by country income level and urban/rural residence^{15,16} and the fact that young
4 67 child populations are increasingly diverse due to migration and urbanization.¹³ These are important
5 68 distinctions given that psychometric properties can vary by population characteristics.
6
7
8

9
10 69 We reviewed available evidence on 11 psychometric properties of tools used to assess ECD
11 70 in children 0-6 years old living in LMICs. Our findings can assist stakeholders in selection of ECD
12 71 tools with adequate psychometric evidence for the intended use, and inform what research is
13 72 needed to improve how we track ECD-related SDGs, programmes, and policies in LMICs.
14
15
16
17
18
19
20
21
22
23

24 73 **METHODS**

25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

74 74 We identified articles in any language through MEDLINE, Embase, PubMed, PsycInfo,
75 SciELO, and BVS (**Supplemental Methods, Supplemental Table 1**). Full-text, peer-reviewed
76 articles were included if: (1) the studies were conducted in a LMIC, (2) included children 0-6 years
77 old, (3) included at least one ECD domain (cognitive, language, motor, social-emotional,
78 attention/executive function, personal-social, and pre-academic/academic) defined in
79 **Supplemental Table 2**, (4) developed a new tool or adapted an existing one, and provided primary
80 evidence of at least one of 11 psychometric properties in terms of reliability, validity, or
81 measurement invariance (**Table 1**), and (5) were published between January 1, 2007 and March 9,
82 2023. We considered all ECD tools regardless of their intended or actual use (diagnosis, screening,
83 or surveillance). We included both articles that adapted original tools (“adaptation articles”
84 hereafter) and articles that developed new tools (“development articles” hereafter). Cited
85 references were reviewed for potential inclusion. We excluded articles where a tool was used to
86 assess an outcome measure (e.g., trials reporting impacts on ECD outcomes), but the article did
87 not include measurement objectives.¹⁷ Among articles with an explicit measurement focus, we
88 not include measurement objectives.¹⁷ Among articles with an explicit measurement focus, we

1
2
3 89 excluded those reporting only on convergent validity (e.g., correlations with socio-demographic
4 variables) which alone provides limited psychometric evidence.¹⁷
5
6

7 91 Search results were imported into Covidence, where duplicates were automatically
8 removed. Two reviewers independently screened titles and abstracts for inclusion and reviewed
9 full texts of retained articles. Disagreements in screening or full text review were resolved through
10 discussion with a third reviewer. Each article was extracted by one reviewer using a standard data
11 extraction sheet (**Supplemental Methods**). For quality assurance, 20% of articles were randomly
12 selected for independent extraction by a second reviewer. Any discrepancies regarding data
13 extraction were resolved by a third reviewer. Authors of included articles were not contacted when
14 information was missing or unclear.
15
16

17 99 Seven reviewers assessed risk of bias using a checklist created by the authors
18 (**Supplemental Methods**), which contained three categories of bias due to: (1)
19 training/administration (not assessed for tools relying on self-assessment), (2) selective reporting
20 (only assessed for studies reporting on convergent or predictive validity), and (3) missing data.
21
22 100 Each article was rated separately on each category using the following ratings: low risk, some
23 concerns, high risk, and very high risk. We also assessed indirectness of populations¹⁸ by assessing
24 whether the sample was limited to a specific setting, the tool covered the entire age range it was
25 intended for, subgroups were generalisable, and results were generalisable. Where insufficient
26 information was provided in the article, we rated the study as “unable to assess”. Disagreements
27 for articles assessed by two reviewers (20%) were resolved through discussion with a third
28 reviewer.
29
30

31 110 Data analysis was conducted at the article-tool level because some articles reported on
32 multiple tools. We created binary indicators for whether evidence on each one of the 11
33
34

1
2
3 112 psychometric properties was reported. In longitudinal studies, we considered the psychometric
4 properties reported at any time point. We then summarised the evidence by country, location, age
5
6 113 group at assessment, and type of article (development vs adaptation). Results were synthesised
7
8 114 narratively.
9
10 115
11
12 116
13
14 117 **Results**

15
16 118 Our search strategy identified 6,430 records from six databases and two other sources
17
18 119 (**Figure 1**). After removing duplicates, 5,388 records were excluded during title and abstract
19 screening. After full text review of the remaining 250 records, 98 records were excluded. We
20 included 153 articles, covering 114 tools. Four articles reported on multiple tools, resulting in 175
21
22 121 article-tool combinations (referred to as articles for brevity).
23
24 122
25
26 123
27
28 123 Most articles were adaptation article (n=140, 80%) (**Figure 2, Supplemental Table 3**).
29
30 124 The most often evaluated psychometric properties were internal consistency reliability (n=114,
31
32 125 65%), concurrent validity (n=78, 45%), test-retest reliability (n=72, 41%) convergent validity
33
34 126 (n=72, 41%), and structural validity (n=69, 39%). Measurement invariance was the least frequently
35
36 127 reported psychometric property (n=17, 10%) and was primarily evaluated over countries and child
37
38 128 sex. The number of articles increased over time (**Supplemental Figure 1**). Between 2007-2010,
39
40 129 most included articles were development articles; since 2011, most included articles have been
41
42 130 adaptation articles.
43
44
45 131
46
47 131 Psychometric evidence came from 53 countries (**Figure 3**); one-third of articles came from
48
49 132 Brazil (n=22, 12%), China (n=18, 10%), India (n=13, 7%), and South Africa (n=10, 5%). For 40%
50
51 133 of countries represented, there was only one article reporting psychometric evidence
52
53 134 (**Supplemental Table 4**). Psychometric evidence for each ECD tool included was generally
54
55
56
57
58
59
60

1
2
3 135 limited to 1-2 countries with a few notable exceptions: the Global Scales for Early Development
4 136 (GSED, 32 countries), the Caregiver Reported Early Development Instrument (CREDI, 18
5 137 countries), the International Development and Early Learning Assessment (IDELA, 17 countries),
6 138 the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III, 12 countries), and
7 139 the Ages and Stages Questionnaire-3 (ASQ-3, 11 countries) (**Supplemental Table 5**).
8
9
10
11
12
13
14

140 Most psychometric evidence came from urban (n=89, 51%) or urban-rural (n=39, 22%)
141 settings. In South Asia and Sub-Saharan Africa, a similar number of articles originated from urban,
142 rural, and urban-rural settings (**Supplemental Table 6**).
143
144

145 For development articles, a single article reported psychometric properties for all tools
146 except for the CREDI covered in two articles (**Supplemental Table 7**). For adaptation articles,
147 ASQ-3 and BSID-III were most often studied (13 and 11 articles, respectively). For 75% of tools
148 only a single article provided psychometric evidence (**Supplemental Table 8**).
149
150

151 Although included tools targeted 0-71.9-month-old children, most studies focused on
152 children 12–17.9 month (n=83, 47%), 6-11.9 months (n=78, 45%), or 54-59.9 months (n=74, 42%)
153 old. Articles largely did not cover the full age groups defined here, or the full age range the ECD
154 tool can be used for. Some articles included age groups as narrow as one month (**Supplemental**
155 **Figures 2-3**).
156
157

158 Most articles reported on language (n=105, 60%), motor (n=98, 56%), and cognitive (n=77,
159 44%) development (**Supplemental Table 9**). Academic/pre-academic was the least studied
160 domain (n=8, 5%). Within domains, internal consistency reliability was most frequently reported
161 (**Figure 5**).
162
163

164 Bias due to training/administration was assessed for 53 articles (30%): 24% had low risk
165 of bias (n=42) and 6% some concerns (n=11) (**Supplemental Tables 10-11**). Most articles (n=80,
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
7510
7511
7512
7513
7514
7515
7516
7517
7518
7519
7520
7521
7522
7523
7524
7525
7526
7527
7528
7529
7530
7531
7532
7533
7534
7535
7536
7537
7538
7539
7540
7541
7542
7543
7544
7545
7546
7547
7548
7549
7550
7551
7552
7553
7554
7555
7556
7557
7558
7559
7560
7561
7562
7563
7564
7565
7566
7567
7568
7569
7570
7571
7572
7573
7574
7575
7576
7577
7578
7579
7580
7581
7582
7583
7584
7585
7586
7587
7588
7589
7590
7591
7592
7593
7594
7595
7596
7597
7598
7599
75100
75101
75102
75103
75104
75105
75106
75107
75108
75109
75110
75111
75112
75113
75114
75115
75116
75117
75118
75119
75120
75121
75122
75123
75124
75125
75126
75127
75128
75129
75130
75131
75132
75133
75134
75135
75136
75137
75138
75139
75140
75141
75142
75143
75144
75145
75146
75147
75148
75149
75150
75151
75152
75153
75154
75155
75156
75157
75158
75159
75160
75161
75162
75163
75164
75165
75166
75167
75168
75169
75170
75171
75172
75173
75174
75175
75176
75177
75178
75179
75180
75181
75182
75183
75184
75185
75186
75187
75188
75189
75190
75191
75192
75193
75194
75195
75196
75197
75198
75199
75200
75201
75202
75203
75204
75205
75206
75207
75208
75209
75210
75211
75212
75213
75214
75215
75216
75217
75218
75219
75220
75221
75222
75223
75224
75225
75226
75227
75228
75229
75230
75231
75232
75233
75234
75235
75236
75237
75238
75239
75240
75241
75242
75243
75244
75245
75246
75247
75248
75249
75250
75251
75252
75253
75254
75255
75256
75257
75258
75259
75260
75261
75262
75263
75264
75265
75266
75267
75268
75269
75270
75271
75272
75273
75274
75275
75276
75277
75278
75279
75280
75281
75282
75283
75284
75285
75286
75287
75288
75289
75290
75291
75292
75293
75294
75295
75296
75297
75298
75299
752100
752101
752102
752103
752104
752105
752106
752107
752108
752109
752110
752111
752112
752113
752114
752115
752116
752117
752118
752119
752120
752121
752122
752123
752124
752125
752126
752127
752128
752129
752130
752131
752132
752133
752134
752135
752136
752137
752138
752139
752140
752141
752142
752143
752144
752145
752146
752147
752148
752149
752150
752151
752152
752153
752154
752155
752156
752157
752158
752159
752160
752161
752162
752163
752164
752165
752166
752167
752168
752169
752170
752171
752172
752173
752174
752175
752176
752177
752178
752179
752180
752181
752182
752183
752184
752185
752186
752187
752188
752189
752190
752191
752192
752193
752194
752195
752196
752197
752198
752199
752200
752201
752202
752203
752204
752205
752206
752207
752208
752209
752210
752211
752212
752213
752214
752215
752216
752217
752218
752219
752220
752221
752222
752223
752224
752225
752226
752227
752228
752229
752230
752231
752232
752233
752234
752235
752236
752237
752238
752239
752240
752241
752242
752243
752244
752245
752246
752247
752248
752249
752250
752251
752252
752253
752254
752255
752256
752257
752258
752259
752260
752261
752262
752263
752264
752265
752266
752267
752268
752269
752270
752271
752272
752273
752274
752275
752276
752277
752278
752279
752280
752281
752282
752283
752284
752285
752286
752287
752288
752289
752290
752291
752292
752293
752294
752295
752296
752297
752298
752299
752300
752301
752302
752303
752304
752305
752306
752307
752308
752309
752310
752311
752312
752313
752314
752315
752316
752317
752318
752319
752320
752321
752322
752323
752324
752325
752326
752327
752328
752329
752330
752331
752332
752333
752334
752335
752336
752337
752338
752339
752340
752341
752342
752343
752344
752345
752346
752347
752348
752349
752350
752351
752352
752353
752354
752355
752356
752357
752358
752359
752360
752361
752362
752363
752364
752365
752366
752367
752368
752369
752370
752371
752372
752373
752374
752375
752376
752377
752378
752379
752380
752381
752382
752383
752384
752385
752386
752387
752388
752389
752390
752391
752392
752393
752394
752395
752396
752397
752398
752399
752400
752401
752402
752403
752404
752405
752406
752407
752408
752409
752410
752411
752412
752413
752414
752415
752416
752417
752418
752419
752420
752421
752422
752423
752424
752425
752426
752427
752428
752429
752430
752431
752432
752433
752434
752435
752436
752437
752438
752439
752440
752441
752442
752443
752444
752445
752446
752447
752448
752449
752450
752451
752452
752453
752454
752455
752456
752457
752458
752459
752460
752461
752462
752463
752464
752465
752466
752467
752468
752469
752470
752471
752472
752473
752474
752475
752476
752477
752478
752479
752480
752481
752482
752483
752484
752485
752486
752487
752488
752489
752490
752491
752492
752493
752494
752495
752496
752497
752498
752499
752500
752501
752502
752503
752504
752505
752506
752507
752508
752509
752510
752511
752512
752513
752514
752515
752516
752517
752518
752519
752520
752521
752522
752523
752524
752525
752526
752527
752528
752529
752530
752531
752532
752533
752534
752535
752536
752537
752538
752539
752540
752541
752542
752543
752544
752545
752546
752547
752548
752549
752550
752551
752552
752553
752554
752555
752556
752557
752558
752559
752560
752561
752562
752563
752564
752565
752566
752567
752568
752569
752570
752571
752572
752573
752574
752575
752576
752577
752578
752579
752580
752581
752582
752583
752584
752585
752586
752587
752588
752589
752590
752591
752592
752593
752594
752595
752596
752597
752598
752599
752600
752601
752602
752603
752604
752605
752606
752607
752608
752609
752610
752611
752612
752613
752614
752615
752616
752617
752618
752619
752620
752621
752622
752623
752624
752625
752626
752627
752628
752629
752630
752631
752632
752633
752634
752635
752636
752637
752638
752639
752640
752641
752642
752643
752644
752645
752646
752647
752648
752649
752650
752651
752652
752653
752654
752655
752656
752657
752658
752659
752660
752661
752662
752663
752664
752665
752666
752667
752668
752669
752670
752671
752672
752673
752674
752675
752676
752677
752678
752679
752680
752681
752682
752683
752684
752685
752686
752687
752688
752689
752690
752691
752692
752693
752694
752695
752696
752697
752698
752699
752700
752701
752702
752703
752704
752705
752706
752707
752708
752709
752710
752711
752712
752713
752714
752715
752716
752717
752718
752719
752720
752721
752722
752723
752724
752725
752726
752727
752728
752729
752730
752731
752732
752733
752734
752735
752736
752737
752738
752739
7527340
7527341
7527342
7527343
7527344
7527345
7527346
7527347
7527348
7527349
7527350
7527351
7527352
7527353
7527354
7527355
7527356
7527357
7527358
7527359
7527360<br

1
2
3 158 82%) had low risk of bias due to selective reporting. Risk of bias due to missing data was assessed
4
5 159 for 130 articles (74%): 28% had some concerns (n=49) and 4% high risk (n=4). For indirectness,
6
7 160 18 articles (10%) were rated as generalisable and 81 (46%) as probably generalisable.
8
9
10 161
11
12 162 **DISCUSSION**
13
14
15 163 Based on 153 articles, available evidence on 11 psychometric properties of 114 ECD tools
16
17 164 for children 0-6 years old is fragmented, limited, and heterogeneous. The most frequently provided
18
19 165 evidence was on internal consistency reliability, test-retest reliability and/or concurrent,
20
21 convergent, and structural validity. Psychometric evidence on measurement invariance was the
22
23 166 least commonly available. Although evidence came from 53 LMICs, four countries were most
24
25 167 represented. Most evidence came from urban or urban-rural settings. The most studied tools were
26
27 168 ASQ-3 and BSID-III.
28
29
30
31 170 Our findings support ECD measurement trends:¹⁹ much of the work on psychometric
32
33 171 properties is recent, with ECD tools being developed and adapted concurrently. Psychometric
34
35 172 efforts remain limited to a few ECD tools,^{20,21} individual ECD domains,¹²⁻¹⁴ and few psychometric
36
37 173 properties.^{10,12} This fragmentation is evidenced here by included articles focusing on individual
38
39 174 countries, limited age ranges, and single developmental domains. The resulting heterogeneous
40
41 psychometric evidence can hinder comparability and large-scale monitoring of ECD policies and
42
43 175 programmes within and across LMICs, which is crucial for identifying and implementing effective
44
45 176 approaches to support ECD.
46
47
48
49 178 Despite efforts to consolidate ECD measurement through tools like CREDI, GSED, and
50
51 179 IDELA, such tools do not fully meet research, programmatic, and policy needs as evidenced by
52
53 180 the increase of adaptation and development articles since 2015. This is not surprising given that
54
55
56
57
58
59
60

1
2
3 181 no ECD tool is suitable for all populations.²² Our disaggregation by development and adaptation
4 articles permitted better understanding of the psychometric evidence available and highlighted
5 evidence gaps for both existing and newly developed tools.
6
7
8
9

10 184 This review highlights four important limitations of existing psychometric evidence for
11 ECD tools in LMICs. First, although most tools are designed for a wide age range, the
12 psychometric evidence behind most tools pertained to narrower age ranges and in some cases as
13 narrow as one month. This may have limited applicability to diverse age ranges (given that there
14 is a natural variability in child development in the early years¹³) in these specific contexts.
15
16

17 189 Relatedly, most psychometric evidence pertained to urban contexts. Given existing urban-rural
18 disparities in ECD^{15,16} and increasingly diverse young child populations in urban settings,¹³ ECD
19 tools whose psychometric properties were examined only in urban settings might be inadequate
20 for rural settings. Those developing or adapting ECD tools should consider establishing
21 psychometric properties across the full intended child age range and across both urban and rural
22 settings. Although this implies longer, more expensive, and logistically difficult studies, it would
23 ensure the tool can be used broadly and in more diverse populations.
24
25

26 196 A second limitation is the very little evidence on measurement invariance. While several
27 studies reported on samples drawn from multiple countries, few conducted statistical analysis to
28 test for equivalence across countries, thus providing no evidence of measurement invariance.⁸
29
30 199 Cross-country invariance, which guarantees assessment of the same construct across countries, is
31 key for tracking global SDG goals. Further, there was a scarcity of articles on measurement
32 invariance over time, a property which guarantees observed changes can be attributed to
33 interventions/policies rather than fluctuations in the measure's stability over time.^{8,23} Psychometric
34 evidence of measurement invariance even for the most widely used tools to assess intervention
35
36

1
2
3 204 impact like BSID²⁴ is limited in LMICs. This may be because establishing measurement invariance
4 over time is particularly challenging due to the natural variability in child development in early
5 life.¹³ Measurement non-invariance can lead to incorrect conclusions about
6 intervention/programme effectiveness.⁸ More work in this domain is needed particularly for tools
7 widely used for policy making and programme evaluations. Measurement work should be
8 considered relative to competing and more urgent ECD priorities in LMICs.
9
10
11
12
13
14
15
16

17 210 Third, consistent with existing literature, we found limited psychometric evidence on tools
18 measuring socio-emotional and personal-social development.^{13,23} This is surprising given that
19 these two domains are among the most culturally specific,¹³ implying they require more
20 comprehensive and rigorous adaptation. Psychometric evidence on tools to assess
21 attention/executive function and academic/preacademic development was most limited. Without
22 additional work to establish a psychometric base, this poses major challenges for those seeking to
23 monitor these domains in early life.
24
25
26
27
28
29
30
31
32

33 217 Lastly, our assessment of study quality demonstrated a common lack of reporting and
34 transparency in training of assessors and data management. Risk of bias could not be assessed for
35 many included studies or was considered high. Better reporting standards and guidelines for
36 psychometric studies can help strengthen the field and ensure more critical evaluation of the
37 evidence is possible.
38
39
40
41
42
43

44 222 Nevertheless, some ECD tools had multiple forms of psychometric properties assessed.
45 Using IDELA as an example: one article examined three types of reliability and one type of
46 validity, drawing on pilots from 12 countries,¹⁹ followed by another article examining
47 measurement invariance across countries.²⁵ Subsequent articles have examined additional
48 psychometric properties, albeit in individual countries.^{26,27} Although such examples of
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 227 consecutively examining psychometric properties should be the norm, they are often financially
4 and logically infeasible and the timing does not always align with programmatic and policy
5 agendas. When prior evidence is unavailable and psychometric studies not possible, statistical
6 analysis should be conducted to verify psychometric properties and at a minimum should include
7 reporting internal consistency reliability and structural validity (where relevant). This can help
8 build evidence across multiple settings and populations and confirm the usefulness of ECD tools
9 with diverse populations.
10
11
12
13
14
15
16
17
18

19 234 Several strengths of our review should be noted. We used extensive search filters, made no
20 restrictions on domains, included both single- and multi-domain tools, and had a rigorously trained
21 team with diverse backgrounds in ECD measurement in LMICs. Nevertheless, we acknowledge
22 some limitations. We excluded studies published prior to 2007, the search start year when the
23 inaugural *Lancet* Series on Child Development in Developing Countries was published which
24 fundamentally changed the breadth and scale of ECD research in LMICs. Therefore, the search
25 period captured the most important years for the evolution of psychometric research on ECD tools
26 in LMICs. Further, although we applied no language restrictions, we were limited to the languages
27 spoken by the review team. We were unable to conduct full text review and extraction for articles
28 in Chinese, Farsi, and Turkish. Consequently, results from China, Iran, and Turkey may be
29 underrepresented.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

CONCLUSION

247 Psychometric evidence is fragmented, limited, and heterogeneous. More research is
248 warranted to establish the applicability of existing tools in diverse populations, including urban
249 and rural settings, and on establishing measurement invariance over countries and time.
250
251
252
253
254
255
256
257
258
259

1
2
3 250 Nevertheless, the results by country and ECD tool presented here can serve stakeholders in
4 selecting tools with at least some available psychometric evidence. To improve monitoring,
5 evaluation, and accountability for ECD globally, psychometric evidence should be a key
6 consideration when selecting ECD tools. Improved reporting for psychometric studies can help
7 ensure transparency, replication, and adequate ability to assess the quality of evidence.
8
9 254
10
11
12
13
14
15
16
17 255
18
19 256 **Acknowledgements**
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

257 **Contributors:** LB conceptualised the review, designed the methodology, screened articles,
258 extracted and analysed data, drafted and revised the manuscript. EH extracted data and analysed
259 data, drafted and revised the manuscript. NBA and AR extracted data, and revised the manuscript.
260 XH and KSC screened articles, extracted data, and revised the manuscript. SEN and JS extracted
261 data and revised the manuscript. DO revised the manuscript. HOP screened articles and revised
262 the manuscript. AT curated the data and revised the manuscript. JJ conceptualised the review,
263 designed the methodology, screened articles, and revised the manuscript. All authors read and
264 approved the final manuscript. LB is the guarantor.

265 **Funding:** This work was supported by the CGIAR Initiative on Resilient Cities. We would like to
266 thank all funders who supported this research through their contributions to the CGIAR Trust
267 Fund: <https://www.cgiar.org/funders>. SEN's time was supported by the National Institute of
268 Mental Health (grant number T32 MH078788).

269 **Competing interests:** JS has contributed and continues to contribute to the development of the
270 CREDI, GSED, IDELA and AIM-ECD. The remaining authors declare no conflicts of interest.

271 **Patient consent for publication:** Not applicable.

272 **Ethical considerations:** Not applicable.

- 1
2
3 273 **Provenance and peer review:** Not commissioned; externally peer reviewed.
4
5 274 **Data availability statement:** All data are included in the manuscript and appendix.
6
7 275
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

References

1. United Nations. Transforming our world: the 2030 Agenda for Sustainable Development. Published online 2015. <https://sdgs.un.org/sites/default/files/publications/21252030Agenda%20for%20Sustainable%20Development%20web.pdf>
2. Walker SP, Chang SM, Vera-Hernandez M, Grantham-McGregor S. Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behavior. *Pediatrics*. 2011;127(5):849-857. doi:10.1542/peds.2010-2231
3. Gertler P, Heckman J, Pinto R, et al. Labor market returns to an early childhood stimulation intervention in Jamaica. *Science* (80-). 2014;344(6187):998-1001. doi:10.1126/science.1251178
4. Campbell F, Conti G, Heckman JJ, et al. Early Childhood Investments Substantially Boost Adult Health. *Science* (80-). 2014;343(6178):1478-1485. doi:10.1126/science.1248429
5. Fernald LCH, Prado EL, Kariger PK, Raikes A. A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries. Published online 2017.
6. Pushparatnam A, Seiden JM, Luna Bazaldua DA. Guiding Questions for Choosing the Right Tools to Measure Early Childhood Outcomes: Why, What, Who, and How. Published online 2022. <https://openknowledge.worldbank.org/handle/10986/37030>
7. American Psychological Association. APA Dictionary of Psychology. Accessed February 26, 2023. <https://dictionary.apa.org/>
8. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Dev Rev*. 2016;41:71-90. doi:10.1016/j.dr.2016.06.004
9. Semrud-Clikeman M, Romero RAA, Prado EL, Shapiro EG, Bangirana P, John CC. Selecting measures for the neurodevelopmental assessment of children in low- and middle-income countries. *Child Neuropsychol*. Published online September 9, 2016;1-42. doi:10.1080/09297049.2016.1216536
10. Munoz-Chereau B, Ang L, Dockrell J, Outhwaite L, Heffernan C. Measuring early child development across low and middle-income countries: A systematic review. *J Early Child Res*. 2021;19(4):443-470. doi:10.1177/1476718X211020031
11. Boggs D, Milner KM, Chandna J, et al. Rating early child development outcome measurement tools for routine health programme use. *Arch Dis Child*. 2019;104(Suppl 1):S22-S33. doi:10.1136/archdischild-2018-315431
12. Griffiths A, Toovey R, Morgan PE, Spittle AJ. Psychometric properties of gross motor assessment tools for children: a systematic review. *BMJ Open*. 2018;8(10):e021734. doi:10.1136/bmjopen-2018-021734
13. Halle TG, Darling-Churchill KE. Review of measures of social and emotional development. *J Appl Dev Psychol*. 2016;45:8-18. doi:10.1016/j.appdev.2016.02.003
14. Humphrey N, Kalambouka A, Wigelsworth M, Lendrum A, Deighton J, Wolpert M. Measures of Social and Emotional Skills for Children and Young People. *Educ Psychol Meas*. 2011;71(4):617-637. doi:10.1177/0013164410382896
15. Lu C, Cuartas J, Fink G, et al. Inequalities in early childhood care and development in low/middle-income countries: 2010–2018. *BMJ Glob Heal*. 2020;5(2):e002314. doi:10.1136/bmjjgh-2020-002314
16. McCoy DC, Peet ED, Ezzati M, et al. Early Childhood Developmental Status in Low- and Middle-Income Countries: National, Regional, and Global Prevalence Estimates Using Predictive Modeling. *PLoS Med*. 2016;13(6):1-18. doi:10.1371/journal.pmed.1002034

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
17. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs). Published online 2018:1-78. https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf
 18. GRADE Working Group. *Handbook for Grading the Quality of Evidence and the Strength of Recommendations Using the GRADE Approach. Updated October 2013.* (Schünemann H, Brozek J, Guyatt G, Oxman A, eds.); 2013. <https://gdt.gradepro.org/app/handbook/handbook.html>
 19. Pisani L, Borisova I, Dowd AJ. Developing and validating the International Development and Early Learning Assessment (IDELA). *Int J Educ Res.* 2018;91(July):1-15. doi:10.1016/j.ijer.2018.06.007
 20. Kersten P, Czuba K, McPherson K, et al. A systematic review of evidence for the psychometric properties of the Strengths and Difficulties Questionnaire. *Int J Behav Dev.* 2016;40(1):64-75. doi:10.1177/0165025415570647
 21. Velikonja T, Edbrooke-Childs J, Calderon A, Sleed M, Brown A, Deighton J. The psychometric properties of the Ages & Stages Questionnaires for ages 2-2.5: a systematic review. *Child Care Health Dev.* 2017;43(1):1-17. doi:10.1111/cch.12397
 22. Sabanathan S, Wills B, Gladstone M. Child development assessment tools in low-income and middle-income countries: how can we use them more appropriately? *Arch Dis Child.* 2015;100(5):482-488. doi:10.1136/archdischild-2014-308114
 23. Gridley N, Blower S, Dunn A, Bywater T, Bryant M. Psychometric Properties of Child (0–5 Years) Outcome Measures as used in Randomized Controlled Trials of Parent Programs: A Systematic Review. *Clin Child Fam Psychol Rev.* 2019;22(3):388-405. doi:10.1007/s10567-019-00277-1
 24. Jeong J, Franchett EE, Ramos de Oliveira C V., Rehmani K, Yousafzai AK. Parenting interventions to promote early child development in the first three years of life: A global systematic review and meta-analysis. *PLOS Med.* 2021;18(5):e1003602. doi:10.1371/journal.pmed.1003602
 25. Halpin PF, Wolf S, Yoshikawa H, et al. Measuring early learning and development across cultures: Invariance of the IDELA across five countries. *Dev Psychol.* 2019;55(1):23-37. doi:10.1037/dev0000626
 26. Wolf S, Halpin P, Yoshikawa H, Dowd AJ, Pisani L, Borisova I. Measuring school readiness globally: Assessing the construct validity and measurement invariance of the International Development and Early Learning Assessment (IDELA) in Ethiopia. *Early Child Res Q.* 2017;41(June):21-36. doi:10.1016/j.ecresq.2017.05.001
 27. Pisani L, Seiden J, Wolf S. Longitudinal evidence on the predictive validity of the International Development and Early Learning Assessment (IDELA). *Educ Assessment, Eval Account.* 2022;34(2):173-194. doi:10.1007/s11092-021-09374-8

136/bmjopen-2024-096365 on 1 May 2025. Downloaded from https://bmjopen.bmjjournals.org/ by guest user. Related to article DOI: 10.1136/bmjopen-2024-096365. © Author(s) (or their employer(s)) 2024. All rights reserved. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits other parties to copy and redistribute the article without prior permission or charge, for the non-commercial research and educational purposes, provided that:

Table 1 Definitions of the psychometric properties used in the selection criteria for included articles

Domain	Measurement property	Example test statistic	Definition
Reliability			The consistency of a test or measurement, i.e., how consistently a measure produces similar results with repeated measures over a short period of time or across assessors at the same time point. This can also be thought of as the correlation between observed scores across replications.
	Test-retest reliability	correlation coefficient	Correlation between scores from the same test from assessments conducted over a short time interval.
	Inter-rater reliability	kappa, Bland-Altman test	The extent to which independent assessors produce similar ratings in judging the same abilities or characteristics in the same target person at the same time.
	Internal consistency reliability	Cronbach's alpha, alpha	Degree of interrelatedness among items on the same tool, i.e., how well the items work together to provide information on an underlying construct.
Validity			The degree to which the tool measures what it is supposed to measure, i.e., the degree to which the tool reflects the underlying construct.
	Content/face validity		The degree to which the content of the tool is adequate for the construct being measured, i.e., assessing the extent to which a tool appears to reflect the underlying construct.
	Concurrent/criterion validity	correlation coefficient; regression estimate	The degree to which scores on one measurement tool are related to scores obtained at about the same point in time from another tool considered the gold standard.
	Convergent validity	correlation coefficient; regression estimate	Evidence that scores on a test or measurement are associated with theoretically related measures or variables.

Domain	Measurement property	Example test statistic	Definition
	Predictive	correlation coefficient; regression estimate	Evidence that a score correlates with a variable that can only be assessed at some point after the test has been administered or the measurement made, e.g., evidence that scores now are correlated with scores at a future point.
	Structural validity (dimensionality)	Exploratory factor analysis: number of factors, eigen values Confirmatory factor analysis: model fit statistics such as Comparative Fit Index (CF), Root mean square error of approximation (RMSEA)	The degree to which the scores of assessment are an adequate reflection of the dimensionality of the construct to be measured.
Invariance			The property when a scale or construct provides the same results across different samples, populations, settings, or characteristics.
	Measurement invariance over countries	Likelihood ratio chi squared statistic and p-value from freeing parameters across groups	The degree to which an assessment of construct provides the same results across separate samples in different countries.
	Measurement invariance over time	Likelihood ratio chi squared statistic and p-value from freeing parameters across groups	The degree to which an assessment of construct provides the same results across time.
	Measurement invariance over other groups	Likelihood ratio chi squared statistic and p-value from freeing parameters across groups	The degree to which an assessment of construct provides the same results across different groups.

Note: All definitions based on the APA Dictionary of Psychology.⁷

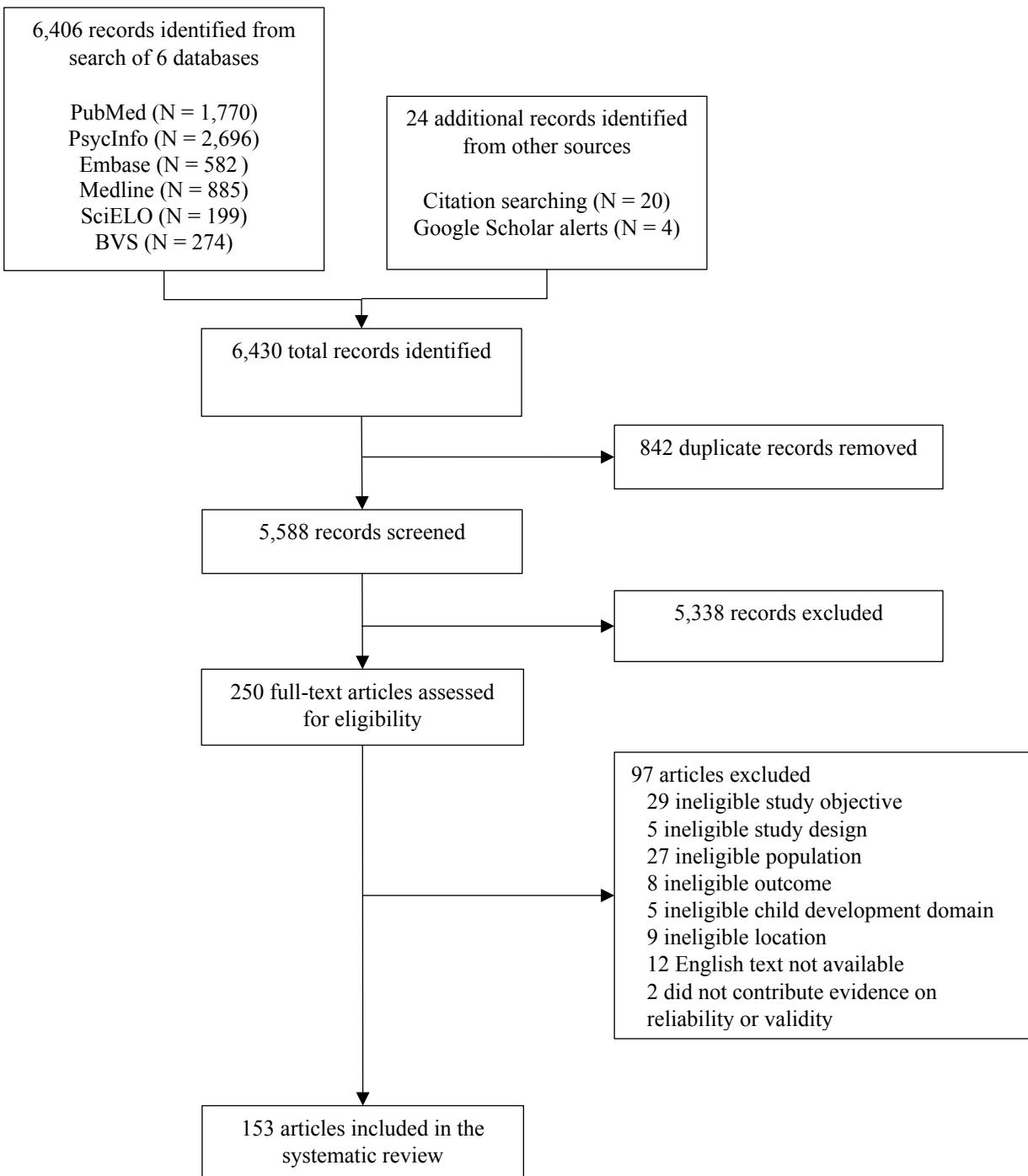
Figure Captions

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram of search results and included articles

Figure 2 Number of included article-tools and type of psychometric evidence provided

Figure 3 Countries where studies providing evidence on at least one psychometric property were conducted

Figure 4 Number of articles providing psychometric evidence by article type and early childhood development domain



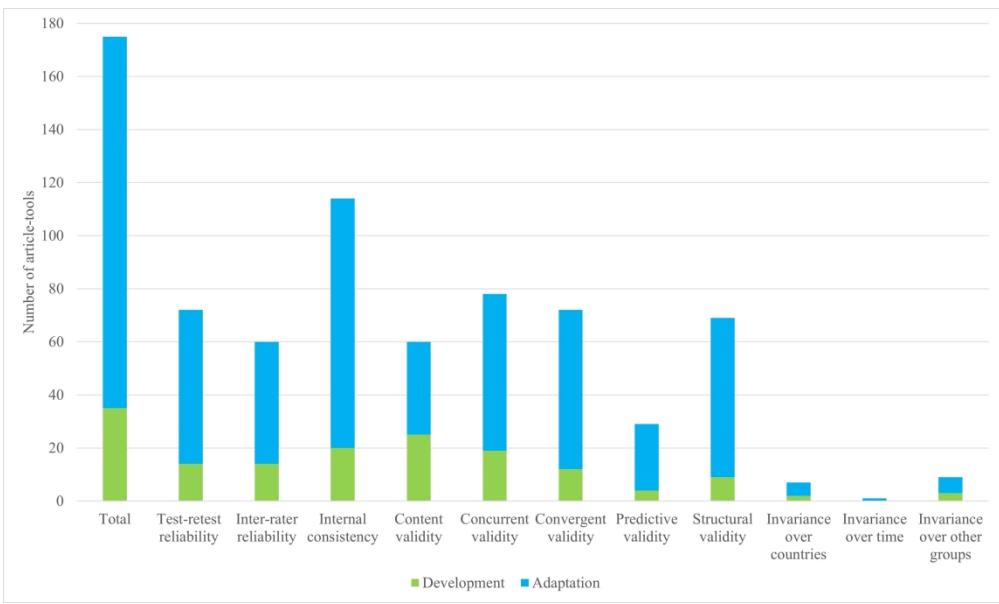


Figure 2 Number of included article-tools and type of psychometric evidence provided

276x164mm (300 x 300 DPI)

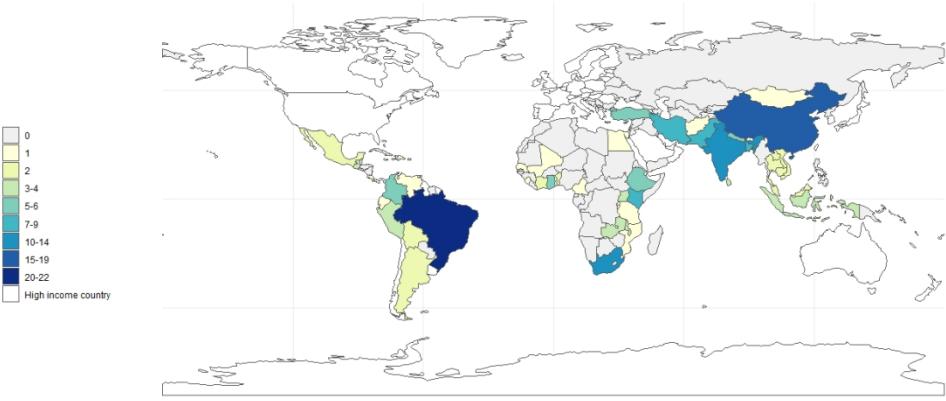


Figure 3 Countries where studies providing evidence on at least one psychometric property were conducted

345x211mm (96 x 96 DPI)

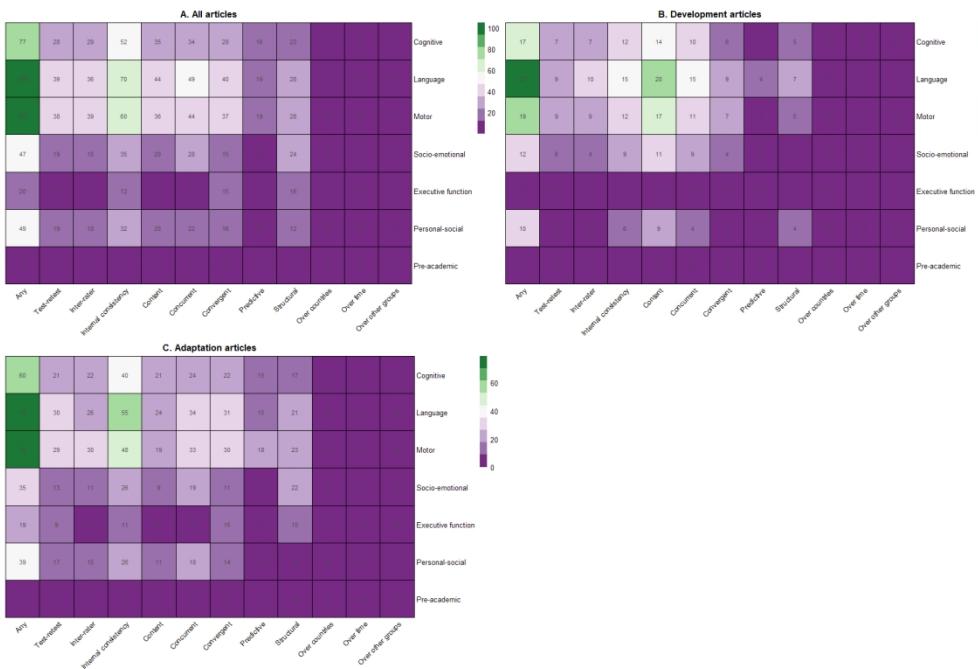


Figure 4 Number of articles providing psychometric evidence by article type and early childhood development domain

476x317mm (96 x 96 DPI)

Supplemental appendix

Table of Contents

Supplemental Methods	2
Search strategy and selection criteria	2
Data analysis	2
Supplemental Table 1 Search strategy as executed in PubMed	4
Supplemental Table 2 Definitions of the seven early childhood domains of interest	7
Supplemental Table 3 Number of included articles by type of psychometric evidence provided	8
Supplemental Table 4 List of countries and assessment tools used in included articles.....	9
Supplemental Table 5 List of child development assessment tools and countries where they were applied.....	30
Supplemental Table 6 Number of articles providing psychometric evidence by world region and urban/rural location.....	37
Supplemental Table 7 Number of articles providing psychometric evidence by child development assessment tools (development articles)	38
Supplemental Table 8 Number of articles providing psychometric evidence by child development assessment tools (adaptation articles)	41
Supplemental Table 9 Number of articles providing psychometric evidence by article type and early childhood development domain.....	47
Supplemental Table 10 Number of articles by category of risk of bias	49
Supplemental Table 11 Risk of bias assessment by included article	50
Supplemental Figure 1 Number of articles providing evidence on psychometrics by type of article	55
Supplemental Figure 2 Number of articles and tools providing psychometric evidence by child age group	56
Supplemental Figure 3 Number of articles providing evidence on psychometrics by child age at assessment.....	57
References of all included articles	58

Supplemental Methods

Search strategy and selection criteria

The search strategy used medical subject headings (MeSH), keywords, and free text words along four key elements: population, construct, measurement properties, and location. Search terms were combined using Boolean operators. Truncation wildcards were used to include variations of the search terms. The search strategy used a combination of searches through titles, abstracts, and keywords. The COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) measurement properties filter was used for the third search element.¹ The search strategy was piloted in PubMed and then adapted for each database.

Articles were included if at least 50% of the study sample was within the range 0-6 years and the average age in the sample was <6 years or unspecified. If information on the child age range was not available or clear, the article was included. Multi-country articles were included if ≥50% of the countries were LMICs. If an article reported that measurement properties were reported elsewhere (e.g., referencing another study and thus providing secondary evidence), the article was ineligible. We included prospective, retrospective, cross-sectional, and longitudinal quantitative study designs.

We excluded articles studying children with developmental disabilities and disorders (e.g., autism spectrum disorder, cerebral palsy), and children with physical disabilities that impair performance on ECD measures (e.g., deafness, blindness). We excluded the following study designs: animal studies, simulation studies, case studies, opinions, letters, preprints, protocols, conference abstracts, ecological studies, dissertations/theses, reviews or systematic reviews, and meta-analyses.

1. Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009; **18**: 1115–23.

Data analysis

One reviewer developed the data extraction sheet, and two reviewers piloted it. Revisions were made through discussion between the two reviewers. Seven reviewers were trained on using the data extraction sheet through pilot extractions of included articles. Further revisions were made to the data extraction sheet based on discussions of the piloting process. Data extraction included information on publication details, study meta-data, characteristics of the ECD tool, type of administration, ECD domains assessed, and psychometric properties with respect to reliability, validity, and measurement invariance.

1
2
3 One reviewer created a risk of bias checklist by adapting items from the COSMIN¹ and Cochrane's
4 ROBINS-E² risk of bias tools. The checklist was then refined via discussions with two other
5 reviewers and after piloting by four reviewers. Seven reviewers were trained and conducted the
6 risk of bias assessment at the article level. Disagreements for articles assessed by two reviewers
7 (20%) were resolved through discussion with a third reviewer. The ROBINS-E risk of bias ratings
8 were used: low risk (little-to-no risk of bias in this domain), some concerns (some risk of bias in
9 this domain, but not clear it is an important risk of bias), high risk (important problems in this
10 domain), and very high risk (the study is very problematic in this domain).
11
12

13 For all tools, child age at assessment was converted to months based on the information provided
14 in the article. For longitudinal studies, age at first assessment was used. For multi-site/multi-
15 country studies, the full age range at assessment was used across the sites/countries. Because the
16 objective the paper was to take stock of the available evidence rather than to report the adequacy
17 and relevance of specific psychometric properties, we did not summarize evidence on the
18 psychometric properties themselves.
19
20

- 21
- 22 1. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews
23 of Patient-Reported Outcome Measures (PROMs). 2018; : 1–78.
 - 24 2. ROBINS-E Development Group. Risk Of Bias In Non-randomized Studies - of Exposure
25 (ROBINS-E). Launch version, 1 June 2022. 2022.
<https://www.riskofbias.info/welcome/robins-e-tool>.
- 26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Supplemental Table 1 Search strategy as executed in PubMed

(“Child”[MeSH] OR “Child, Preschool”[MeSH] OR “child*”[tiab] OR “Infant”[MeSH] OR “infant*”[tiab] OR “paediatric”[tiab] OR “pediatric”[tiab] OR “toddler*”[tiab])

AND

("Child Development"[MeSH] OR "child development*"[tiab] OR "early child development"[tiab] OR "early childhood development"[tiab] OR "ECD"[tiab] OR "Child Behavior"[MeSH] OR "child behavior"[tiab] OR "child behaviour"[tiab] OR "Cognition"[MeSH] OR "Executive Function"[MeSH] OR "executive function*"[tiab] OR "Emotional Intelligence"[MeSH] OR "attachment"[tiab] OR "attention"[tiab] OR "cognition"[tiab] OR "cognitive development"[tiab] OR "cognitive skills"[tiab] OR "neurocognitive"[tiab] OR "communication"[tiab] OR "language"[tiab] OR "language skills"[tiab] OR "language development"[tiab] OR "receptive language"[tiab] OR "expressive language"[tiab] OR "speech development"[tiab] OR "communication skills"[tiab] OR "Motor Skills"[MeSH] OR "motor skills"[tiab] OR "fine motor"[tiab] OR "gross motor"[tiab] OR "motor"[tiab] OR "prosocial"[tiab] OR "socialisation"[tiab] OR "socialization"[tiab] OR "socioemotion*"[tiab] OR "socio-emotion*"[tiab] OR "neurodevelopment"[tiab] OR "neurobehavior*"[tiab] OR "neurobehaviour*"[tiab] OR "neurobehavioral*"[tiab] OR "neurobehavioural*"[tiab] OR "Intelligence"[MeSH] OR "intelligence"[tiab] OR "IQ"[tiab] OR "psychomotor"[tiab] OR "sensorimotor"[tiab] OR "attention"[tiab] OR "inhibition"[tiab] OR "impulse control"[tiab] OR "working memory"[tiab] OR "problem solving"[tiab] OR "self-regulation"[tiab] OR "prosocial"[tiab] OR "psychosocial development"[tiab] OR "internali*"[tiab] OR "externali*"[tiab] OR "empathy"[tiab] OR "neuropsychology*"[tiab] OR "school readiness"[tiab] or "numeracy"[tiab] or "literacy"[tiab] OR "pre-academic"[tiab] OR "early learning"[tiab] or "personal-social"[tiab] OR "personal social"[tiab])

AND

((“developing countr*”[tiab] OR “under developed countr*”[tiab] OR lmic*[tiab] OR “less developed”[tiab] OR “low income”[tiab] OR “lower income”[tiab] OR “low- and middle-income”[tiab] OR “low middle income”[tiab] OR “resource poor”[tiab] OR “resource constrained”[tiab] OR “low resource”[tiab] OR “limited resource*”[tiab] OR “resource limited”[tiab]) AND (Afghan*[tiab] OR Africa[tiab] OR African[tiab] OR Algeria*[tiab] OR “American Samoa*”[tiab] OR Angola*[tiab] OR Argentin*[tiab] OR Bangladesh*[tiab] OR Barbad*[tiab] OR Belorussian[tiab] OR Beliz*[tiab] OR Benin*[tiab] OR Bhutan*[tiab] OR Bolivia*[tiab] OR Botswan*[tiab] OR Brazil*[tiab] OR “Burkina Faso”[tiab] OR Burkinabe[tiab] OR Burund*[tiab] OR Cambodia*[tiab] OR Cameroon*[tiab] OR “Cape Verde”[tiab] OR “Cape Verdean”[tiab] OR “Central African Republic”[tiab] OR Chad*[tiab] OR Chile*[tiab] OR China[tiab] OR Chinese[tiab] OR Colombia*[tiab] OR Comoros[tiab] OR Comorian[tiab] OR Congo[tiab] OR Congolese[tiab] OR Costa Rica*[tiab] OR “Côte d’Ivoire”[tiab] OR “Ivory Coast”[tiab] OR Ivorian[tiab] OR Croatia*[tiab] OR Croat[tiab] OR

1
2
3 Cuba*[tiab] OR Djibouti*[tiab] OR Dominica*[tiab] OR "Dominican Republic"[tiab] OR
4 Ecuador*[tiab] OR Egypt*[tiab] OR "El Salvador"[tiab] OR Salvadorian[tiab] OR "Equatorial
5 Guinea"[tiab] OR Guinean[tiab] OR Eritrea*[tiab] OR Ethiopia*[tiab] OR Fiji*[tiab] OR
6 Gabon*[tiab] OR Gambia*[tiab] OR Gaza[tiab] OR Gazan[tiab] OR Ghana[tiab] OR
7 Ghanaian[tiab] OR Grenad*[tiab] OR Guatemala*[tiab] OR Guinea[tiab] OR Guyan*[tiab] OR
8 Haiti*[tiab] OR Hondura*[tiab] OR Hungar*[tiab] OR India[tiab] OR Indian[tiab] OR
9 Indonesia*[tiab] OR Iran*[tiab] OR Iraq*[tiab] OR Jamaica*[tiab] OR Jordan*[tiab] OR
10 Kenya[tiab] OR Kenyan[tiab] OR Kiribati[tiab] OR Korea*[tiab] OR Kyrgy*[tiab] OR
11 Laos[tiab] OR Laotian*[tiab] OR Lebanon[tiab] OR Lebanese[tiab] OR Lesotho[tiab] OR
12 Liberia*[tiab] OR Libya*[tiab] OR Macedonia*[tiab] OR Madagasca*[tiab] OR Malawi*[tiab]
13 OR Malaysia*[tiab] OR Maldives[tiab] OR Maldivian[tiab] OR Mali[tiab] OR Malian*[tiab]
14 OR "Marshall Islands"[tiab] OR Mauritania*[tiab] OR Mauritius[tiab] OR Mauritian[tiab] OR
15 Mayotte[tiab] OR Mexic*[tiab] OR Micronesia*[tiab] OR Moldov*[tiab] OR Mongolia*[tiab]
16 OR Morocco*[tiab] OR Mozambique[tiab] OR Mozambican[tiab] OR Myanmar[tiab] OR
17 Namibia*[tiab] OR Nepal*[tiab] OR Nevis[tiab] OR Nicaragua*[tiab] OR Niger*[tiab] OR
18 "Northern Mariana Islands"[tiab] OR Oman*[tiab] OR Pakistan*[tiab] OR Palau*[tiab] OR
19 Panama*[tiab] OR "Papua New Guinea"[tiab] OR Paraguay*[tiab] OR Peru*[tiab] OR
20 Philippine*[tiab] OR Filipino*[tiab] OR Poland[tiab] OR Polish[tiab] OR Rwanda*[tiab] OR
21 Samoa*[tiab] OR Sao Tome*[tiab] OR Principe[tiab] OR Senegal*[tiab] OR Seychell*[tiab]
22 OR Sierra Leon*[tiab] OR Solomon Island*[tiab] OR Somali*[tiab] OR South Africa*[tiab] OR
23 Sri Lanka*[tiab] OR "Saint Kitts"[tiab] OR "St Kitts"[tiab] OR "Saint Lucia"[tiab] OR "St
24 Lucia"[tiab] OR "Saint Vincent" [tiab] OR "St Vincent"[tiab] OR Sudan*[tiab] OR
25 Suriname*[tiab] OR Swaziland Or Swazi[tiab] OR Syria*[tiab] OR Tajik*[tiab] OR
26 Tanzania*[tiab] OR Thailand[tiab] OR Thai[tiab] OR "Timor Leste"[tiab] OR Togo*[tiab] OR
27 Tonga*[tiab] OR Trinidad[tiab] OR Trinidadian[tiab] OR Tobago[tiab] OR Tobagonian[tiab]
28 OR Tunisia*[tiab] OR Turkey[tiab] OR Turkish[tiab] OR Uganda*[tiab] OR Uruguay*[tiab] OR
29 Vanuat*[tiab] OR Venezuela*[tiab] OR Vietnam*[tiab] OR "West Bank"[tiab] OR
30 Yemen*[tiab] OR Zambia*[tiab] OR Zimbabwe*))

31 AND
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(instrumentation[sh] OR methods[sh] OR "Validation Studies"[pt] OR "Comparative
Study"[pt] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR
clinometr*[tw] OR "outcome assessment (health care)"[MeSH] OR "outcome assessment"[tiab]
OR "outcome measure*[tw] OR "observer variation"[MeSH] OR "observer variation"[tiab]
OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR
reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR
valid*[tiab] OR "coefficient of variation"[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR
homogeneous[tiab] OR "internal consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR
alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR
reduction*[tiab])) OR agreement[tw] OR precision[tw] OR imprecision[tw] OR "precise
values"[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND
(test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[tw] OR result[tw] OR results[tw] OR test[tw] OR tests[tw])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known group"[tiab] OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR "factor structure"[tiab] OR "factor structures"[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR "item discriminant"[tiab] OR "interscale correlation*[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab])) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab])) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab] OR "measurement invariance"[tiab] OR "invariance"[tiab] OR assessment[tiab] OR adapt*[tiab])

NOT

("editorial"[Publication Type] OR "comment"[Publication Type] OR "case reports"[Publication Type] OR "congress"[Publication Type] OR "consensus development conference"[Publication Type] OR "interview"[Publication Type] OR "lecture"[Publication Type] OR "legal case"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "Observational Study, Veterinary"[Publication Type] OR "Patient Education Handout"[Publication Type] OR "practice guideline"[Publication Type] OR "preprint"[Publication Type])

1
2
3
4
5
Supplemental Table 2 Definitions of the seven early childhood domains of interest

6 Domain	7 Definition
8 Cognitive	9 Cognitive development includes abilities such as general intellectual ability, problem solving, reasoning, visual-spatial ability, memory, learning, etc.
10 Language	11 Language development includes the ability to understand and express verbal communication, and vocabulary development and acquisition.
12 Motor	13 Motor development includes gross motor development (the ability to control and coordinate gross movements of the legs and arms (e.g., jumping, throwing)) and fine motor development (the ability to control and coordinate fine movements of the fingers and toes).
14 Socio-emotional/temperament	15 Socio-emotional development is the ability to regulate emotional responses and social interactions. It includes behaviour problems, social competency, emotional competency, and self-regulation. 16 Temperament includes biological influences on the experience and expression of emotion: extraversion (positive affect, activity level, impulsivity, risk-taking), negative affectivity (fear, anger, sadness, discomfort) and effortful control (attention shifting and focusing, perceptual sensitivity, inhibitory and activational control). 17 Socio-emotional development and temperament are often overlapping constructs, particularly in the early years.
18 Attention/Executive function	19 Executive function, including attention, entails international control over behaviour and cognition. It covers inhibitory control, cognitive flexibility, planning, and working memory.
20 Personal-social/adaptive	21 Personal-social and adaptive behaviour comprises the ability to perform daily-life and self-help skills, such as feeding, dressing, toilet training, recognising others, interacting with others, and adjusting to new situations.
22 Academic/pre-academic	23 Preacademic skills are skills needed to learn reading and math, such as counting and letters.

50
51 Note: Definitions based on (Boggs et al., 2019; Fernald et al., 2017).
52
53
54
55
56
57
58
59
60

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES). Copyright © 2024 uses related to text and data mining, AI training, and similar technologies.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
Supplemental Table 3 Number of included articles by type of psychometric evidence provided

Type of article	Total number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance		
		Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Invariance over countries	Invariance over time	Invariance over other groups
All types of articles	175	72	60	114	60	78	72	29	69	7	1	9
Development articles	35	14	14	20	25	19	12	4	9	2	0	3
Adaptation articles	140	58	46	94	35	59	60	25	60	5	1	6

Supplemental Table 4 List of countries and assessment tools used in included articles

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance		References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over time	Over other groups	
Afghanistan	International Development and Early Learning Assessment (IDELA)	1	0	0	0	0	0	0	0	1	0	0	Halpin 2019
Argentina	Bayley Infant Neurodevelopmental Screener (BINS)	1	1	1	0	0	0	0	0	0	0	0	McCarthy 2012
Argentina	Escala de Habilidades Sociales (Scale of Social Abilities)	1	0	0	1	0	0	0	0	1	0	0	Lacunza 2009
Bangladesh	Ages & Stages Questionnaire Inventory (ASQ:I)	1	0	1	1	0	1	1	0	0	0	0	Pitchik 2023
Bangladesh	Bayley Scales of Infant Development, Second Edition (BSID-II)	1	0	0	0	1	0	0	0	0	0	0	Parveen 2014
Bangladesh	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	1	0	0	0	1	0	0	Pendergast 2018
Bangladesh	Dimensions of Mastery Questionnaire (DMQ-18)	1	1	1	1	1	1	0	0	1	0	0	Shaoli 2019
Bangladesh	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018

136/bmjopen-2024096365
For peer review only
on May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). At training and similar technologies.

136/bmjopen-2024-096365 | For peer review only | Downloaded from <http://bmjopen.bmjjournals.org> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - Affiliation et similar technologies

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance across countries, including for uses related to text and data mining	Over time	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural				
Bangladesh	Rapid Neurodevelopmental Assessment (RNDA)	2	0	2	0	1	2	0	1	0		0	0	Khan 2010; Khan 2013
Bangladesh	The language inventory	1	1	0	0	1	1	1	1	0		0	0	Hamadani 2010
Bangladesh	WHO Gross Motor Milestones	1	0	1	0	0	1	1	1	0		0	0	Hamadani 2013
Benin	Mullen Scales of Early Learning (MSEL)	2	0	1	0	0	1	1	1	0		0	0	Koura 2013; Boivin 2021
Bhutan	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0		0	0	Pisani 2018
Bolivia	Bayley Infant Neurodevelopmental Screener (BINS)	1	1	1	0	0	0	0	0	0		0	0	McCarthy 2012
Bolivia	International Development and Early Learning Assessment (IDELA)	1	0	0	0	0	0	0	0	1		0	0	Halpin 2019
Brazil	Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	1	0	0	0	0	0	0	0	1		0	0	Anunciação 2019
Brazil	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	0	1	0	0	0	0	1		0	1	Filgueiras 2013
Brazil	Ages & Stages Questionnaires: Social Emotional (ASQ:SE-BR)	1	0	0	0	0	0	0	0	1	0	1	0	Chen 2017

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance selected by copyright, including for uses related to text and data mining, AI training, and similar technologies.	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Brazil	Alberta Infant Motor Scale (AIMS)	4	3	3	1	1	3	0	1	0	0	0	Saccani 2016; Siegle 2018; de Albuquerque 2018; Valentini 2012
Brazil	Battelle Developmental Inventory Second Edition, BDI-2	1	0	0	0	1	0	0	0	0	0	0	de Albuquerque 2022
Brazil	Bayley Infant Neurodevelopment Screener (BINS)	2	1	1	1	1	1	0	1	1	0	0	Guedes 2011; McCarthy 2012
Brazil	Bayley Scales of Infant Development, Third Edition (BSID-III)	2	1	0	1	1	1	0	0	2	0	0	Pendergast 2018; Madaschi 2016
Brazil	Caregiver Reported Early Development Instruments (CREDI), short-form	1	1	0	1	1	1	1	0	0	0	0	Altafim 2018
Brazil	Denver Developmental Screening Test (Denver-II)	1	0	1	1	1	0	1	0	1	0	0	Lopez Boo 2020
Brazil	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1	0	1	Sincovich 2019
Brazil	Motor Development Scale	1	0	0	1	0	0	0	0	1	0	0	Okuda 2020
Brazil	Oxford Neurodevelopment Assessment (OX-NDA)	1	1	1	1	0	1	0	0	0	0	0	Fernandes 2022

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance selected by copyright, including for uses related to text and data mining	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Brazil	Preschool and Kindergarten Behavior Scale (PKBS)	1	0	0	1	0	0	0	0	1	0	0	Dias 2011
Brazil	Questionário para Avaliação do Desenvolvimento Infantil (QAD-PIPAS)	1	0	0	0	1	0	0	0	0	0	0	Venancio 2020
Brazil	Screening for Identification of Oral Language Difficulties by Preschool Teachers (SIOLD)	1	0	0	0	0	0	0	0	1	0	0	Puglisi 2020
Brazil	Self-ordered Pointing Task (SOPT)	1	0	0	0	0	1	1	0	0	0	0	Sallum 2017
Brazil	Test of Gross Motor Development (TGMD-2)	1	0	1	1	0	0	1	0	1	0	0	Bandeira 2020
Cambodia	Angkor Hospital for Children Developmental Milestone Assessment Tool (AHC DMAT)	1	0	0	0	1	0	0	0	0	0	0	Ngoun 2012
Cambodia	Developmental Milestone Chart (DMC)	1	0	0	0	1	0	0	0	0	0	0	Scherzer 2009
Cameroon	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	0	0	0	1	0	1	0	Lohaus 2014
China	Ages & Stages Questionnaire Inventory (ASQ:I)	1	1	0	1	1	1	0	0	0	0	0	Xie 2017

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance corrected by copyright, including for uses related to text and data mining, AI training, and similar technologies.	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
China	Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	1	1	1	1	1	0	0	0	1	0	0	Bian 2017
China	Ages & Stages Questionnaire: Social-Emotional, First Edition (ASQ:SE-1)	1	1	1	1	0	1	0	0	0	0	0	Xie 2019 (Journal of Child & Dev Studies)
China	Ages & Stages Questionnaire-3 (ASQ-3)	2	1	1	1	0	2	1	0	1	0	0	Yue 2019; Bian 2012
China	Caregiver Reported Early Development Instruments (CREDI), long-form	1	0	0	1	0	1	1	0	0	0	0	Li 2020
China	Child Behavior Checklist/1.5-5 (CBCL/1.5-5) and the Caregiver-Teacher Report Form (C-TRF)	1	0	0	0	0	0	0	0	1	0	0	Liu 2011
China	Child Social Preference Scale	1	0	0	1	0	0	0	1	1	0	0	Li 2016
China	Chinese Preschool Readiness Scale (CPRS)	1	0	0	1	1	1	0	0	1	0	0	Xie 2019 (Early Educ & Dev)
China	Chinese version of Urban Infant-Toddler Social and Emotional Assessment	1	1	1	1	1	1	0	0	1	0	0	Jianduan 2009
China	Chinese Vocabulary Checklist	1	0	0	0	0	0	1	0	0	0	0	Hao 2008

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance			References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over time	Over other groups	First author & year	
China	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1	0	1	Sincovich 2019	
China	Emotional and Behavioral Screener (EBS)	1	0	1	1	0	1	0	0	1	0	0	Zhang 2018	
China	Movement Assessment Battery for Children- second edition (MBAC-2)	1	1	1	1	1	1	0	0	1	0	0	Hua 2013	
China	Preschool Learning Behavior Scale (PLBS)	1	0	0	1	0	1	0	0	1	0	0	Wu 2019 (JPA)	
China	Social Skills domain of System-Rating Scales (SSIS-RS)	1	0	0	1	0	0	1	0	1	0	0	Wu 2019 (PS)	
China	The Chinese Early Language and Communication Questionnaire	1	0	0	1	0	1	0	0	0	0	0	Zhang 2021	
China	The Chinese Intelligence Scale for Young Children	1	0	0	0	0	0	0	0	1	0	1	Guo 2009	
Colombia	Ages & Stages Questionnaire-3 (ASQ-3)	2	2	0	2	0	1	1	1	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020	
Colombia	Battelle Developmental Inventory Second Edition, BDI-2	2	2	0	2	0	1	1	1	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020	
Colombia	Bayley Infant Neurodevelopmental Screener (BINS)		1	1	0	0	0	0	0	0	0	0	McCarthy 2012	

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance selected by copyright, including for uses related to text and data mining, AI training and similar technologies.	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Colombia	Bayley Scales of Infant Development, Third Edition (BSID-III)	2	2	0	2	0	1	1	1	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	Denver Developmental Screening Test (Denver-II)	2	2	0	2	0	1	1	1	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	MacArthur-Bates Communicative Development Inventories CDI	1	0	0	1	1	0	1	0	0	0	0	Lara Díaz 2011
Colombia	MacArthur-Bates Communicative Development Inventories Short-Forms I vocabulary checklist	2	2	0	2	0	1	1	1	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	MacArthur-Bates Communicative Development Inventories Short-Forms II vocabulary checklist	2	2	0	2	0	1	1	1	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	Preschool Language Scale-3 (PLS- 3)	1	0	0	1	0	0	0	0	1	0	0	Romero 2013
Colombia	WHO Gross Motor Milestones	2	2	0	2	0	1	1	1	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020
Costa Rica	EDIN II (Evaluación del Desarrollo Integral de la niña y el niño)	1	0	0	0	1	0	1	0	0	0	0	Cordero 2019

136/bmjopen-2024-096365 | For peer review only | Downloaded from <http://bmjopen.bmjjournals.org> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - www.abes.fr. Copyright © 2024 BMJ Publishing Group Ltd. All rights reserved.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance			References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over time	Over other groups	First author & year	
Côte d'Ivoire	Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	1	0	0	1	0	0	1	0	1	0	0	Barthel 2021	
Côte d'Ivoire	Kilifi Developmental Inventory	1	0	1	1	0	0	0	1	1	0	0	Barthel 2022	
Dominican Republic	Malawi Developmental Assessment Tool (MDAT), renamed Dominican Child Development Screening Tool	1	0	1	1	0	0	1	0	0	0	0	Sánchez-Vincitore 2019	
Ecuador	Bayley Infant Neurodevelopmental Screener (BINS)	1	1	1	0	0	0	0	0	0	0	0	McCarthy 2012	
Egypt	International Development and Early Learning Assessment (IDE LA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018	
Ethiopia	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	1	1	0	0	1	1	1	0	0	Hanlon 2016	
Ethiopia	Denver Developmental Screening Test (Denver-II)	1	1	1	0	1	0	0	0	0	0	0	Abessa 2016	
Ethiopia	International Development and Early Learning Assessment (IDE LA)	3	1	1	1	1	2	2	0	2	1	0	1	Halpin 2019; Pisani 2018; Wolf 2017

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance selected by copyright, including for uses related to text and data mining, AI training, and similar technologies	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Georgia	Ages & Stages Questionnaire-3 (ASQ-3)	1	1	0	1	1	0	1	0	0	0	0	Zirakashvili 2018
Ghana	Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	1	0	0	1	0	0	1	0	1	0	0	Barthel 2021
Ghana	International Development and Early Learning Assessment (IDELA)	1	0	0	1	0	0	0	1	1	0	0	Pisani 2022
Ghana	Kilifi Developmental Inventory	1	0	1	1	0	0	0	1	1	0	0	Barthel 2022
Ghana	Rapid Assessment of Cognitive and Emotional Regulation (RACER)	1	1	1	0	0	1	1	0	0	0	0	Yuan 2022
Ghana	Vocabulary Checklists	1	0	0	0	0	1	0	1	0	0	0	Prado 2018
Grenada	INTER-GROWTH-21st Project Neurodevelopment Assessment (INTER-NDA)	1	1	1	1	0	0	0	0	0	0	0	Waechter 2022
Guatemala	Ages & Stages Questionnaire-3 (ASQ-3)	1	1	0	1	0	1	0	1	0	0	0	Colbert 2021
Guatemala	Mullen Scales of Early Learning (MSEL)	1	1	1	1	0	0	1	1	0	0	0	Colbert 2020

136/bmjopen-2024-096365 | published online 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). For peer review only. © Author(s) 2024. This article is an open access publication

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance	Over time	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural				
Guatemala	Rapid Neurodevelopmental Assessment (RNDA)	1	0	1	0	1	1	1	0	0	0	0	0	Thompson 2015
India	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	1	1	1	0	0	0	0	0	0	0	Kvestad 2013
India	Bayley Scales of Infant Development, Second Edition (BSID-II)	1	0	0	0	0	1	0	0	0	0	0	0	Biasini 2015
India	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	1	0	0	0	1	0	0	0	Pendergast 2018
India	Caregiver Reported Early Development Instruments (CREDI), long-form	1	0	0	1	0	1	1	0	0	0	0	0	Alderman 2021
India	DEvelopmental assessment on an E-Platform (DEEP)	2	0	0	0	1	1	0	0	0	0	0	0	Bhavnani 2019; Mukherjee 2020
India	Developmental Assessment Tool for Anganwadis (DATA)	1	0	0	1	1	0	0	0	1	0	0	0	Nair 2009
India	Developmental Assessment Tool for Anganwadis (DATA-II)	1	0	0	1	1	0	0	0	1	0	0	0	Nair 2012
India	International Development and Early Learning Assessment (IDEA)	1	1	1	1	1	1	1	0	0	0	0	0	Pisani 2018

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance selected by copyright, including for uses related to text and data mining, AI training and similar technologies.	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
India	Parents' Evaluation of Developmental Status (PEDS)	1	0	0	0	0	1	0	0	0	0	0	Mukherjee 2022
India	Screening Test of Early Language Development-Test version (STELD-T)	1	0	0	1	0	1	0	0	0	0	0	Pathak 2023
India	Bracken School Readiness Assessment, Third Edition (BSRA-3)	1	0	0	1	1	0	0	0	0	0	0	Shah 2013
India	Vineland Adaptive Behaviour Scale (VABS II)	1	1	1	0	0	0	1	0	0	0	0	Kumar 2016
Indonesia	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	1	1	1	1	0	0	0	0	0	Yunilda 2022
Indonesia	Early Development Instrument	1	0	1	1	1	1	0	1	0	0	0	Brinkman 2017
Indonesia	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018
Iran	Ages & Stages Questionnaire (ASQ)	1	0	1	1	1	0	0	0	1	0	0	Vameghi 2013
Iran	Bayley Scales of Infant and Toddler Development (BSID)	1	1	1	1	1	0	1	0	1	0	0	Azari 2017
Iran	Childhood Nonverbal Communication Scale (CNCS)	1	0	0	1	1	1	1	0	1	0	0	Oryadi-Zanjani 2020

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org> by [Agence Bibliographique de l'Enseignement Supérieur \(ABES\)](#). For personal use only. Not for commercial use. For peer review only.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Iran	Children's Behavior Questionnaire (CBQ)	1	1	0	1	1	1	1	0	1	0	Golmohammadi 2022
Iran	Elicited Imitation Test (EIT)	1	0	1	0	0	1	0	0	0	0	Mehrani 2018
Iran	Farsi Narrative Norms Instrument	1	1	1	0	1	0	0	0	0	0	Soleymani 2016
Iran	Infant Neurological International Battery	1	0	1	0	0	1	0	0	0	0	Soleimani 2007
Jamaica	Intelligibility in Context Scale (ICS) / ICS-Jamaican Creole	1	1	1	1	0	1	1	0	1	0	Washington 2017
Kenya	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	1	0	1	1	0	1	0	1	0	McHenry 2021
Kenya	Child behaviour checklist (CBCL)	1	1	0	1	0	0	1	0	1	0	Kariuki 2016
Kenya	Communicative Development Inventory (CDI)	1	1	0	1	0	1	0	0	0	0	Alcock 2015
Kenya	Developmental Milestones Checklist (DMC)	1	1	0	1	1	1	1	0	1	0	Abubakar 2010
Kenya	Executive Function Touch (EF Touch)	2	0	1	1	1	1	2	0	1	0	Willoughby 2019 (Child Neuropsych); Willoughby 2019 (Dev Science)
Kenya	Kilifi Developmental Inventory	1	1	1	1	1	1	1	0	0	0	Abubakar 2008

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance corrected by copyright, including for uses relating to text and data mining	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Kenya	Receptive and expressive vocabulary test battery (English, Luo and Swahili)	1	0	1	1	0	1	1	1	0	0	0	Knauer 2019
Kiribati	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1	0	1	Sincovich 2019
Laos	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1	0	1	Sincovich 2019
Laos	MELQO Measure of Development and Learning Outcomes	1	0	0	0	0	0	1	0	1	0	1	Gomez 2022
Malawi	International Development and Early Learning Assessment (IDE LA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018
Malawi	Malawi Developmental Assessment Tool (MDAT)	1	1	1	0	1	0	1	0	0	0	0	Gladstone 2010
Malawi	Vocabulary Checklists	1	0	0	0	0	1	0	1	0	0	0	Prado 2018
Malaysia	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	1	0	0	0	0	1	0	0	Zakaria 2012
Malaysia	Malay Preschool Language Assessment Tool (MPLAT)	1	1	1	1	1	0	0	0	0	0	0	Razak 2010

136/bmjopen-2024-096365
For peer review only
Downloaded from <http://bmjopen.bmjjournals.org> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Ateliers d'apprentissage et de formation à distance et de la recherche et de la technologie.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance	References	
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Mali	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018
Mexico	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	1	1	0	0	0	0	1	0	0	Ortiz-Leon 2018
Mexico	Spanish-language MacArthur–Bates Communicative Development Inventories short forms (S-CDI SFI and SFII)	1	1	0	0	0	0	1	0	0	0	0	Jackson-Maldonado 2013
Mongolia	Mongolian Rapid Baby Scale (MORBAS)	1	0	0	0	1	1	0	0	0	0	0	Dagvadorj 2015
Mozambique,	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018
Nepal	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	1	1	0	0	1	0	0	0	0	Shrestha 2019
Nepal	Bayley Scales of Infant Development, Third Edition (BSID-III)	3	0	2	1	1	0	0	0	1	1	0	Pendergast 2018; Ranjitkar 2018; Manandhar 2016
Nepal	The Test of Infant Motor Performance (TIMP)	1	0	1	1	0	1	0	1	0	0	0	Kvestad 2023

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance corrected by copyright, including for uses related to text and data mining, AI training, and similar technologies.	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Pakistan	Bayley Scales of Infant Development, Second Edition (BSID-II)	1	0	0	0	0	1	0	0	0	0	0	Biasini 2015
Pakistan	Bayley Scales of Infant Development, Third Edition (BSID-III)	2	0	0	0	1	1	0	1	1	0	0	Pendergast 2018; Rasheed 2023
Pakistan	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018
Pakistan	Executive function battery	1	0	0	1	1	0	1	1	1	0	0	Obradovic 2019
Pakistan	Executive Function Touch (EF Touch)	1	0	0	0	0	0	1	0	1	0	0	Toor 2022
Pakistan	MacArthur Communicative Development Inventory Level II	1	1	1	1	1	1	0	0	0	0	0	Rasheed 2017
Pakistan	Strengths and Difficulties Questionnaire (SDQ)	1	1	1	1	1	1	1	0	1	0	0	Finch 2018
Pakistan	Wechsler Preschool and Primary Scale of Intelligence, third edition (WPPSI-III)	1	0	1	1	1	0	1	0	1	0	0	Rasheed 2018
Peru	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	1	0	0	0	1	1	0	Pendergast 2018

136/bmjopen-2024-096365 | For peer review only | 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - A training and similar technologies.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Peru	Child Behavior Scale (CBS)	1	0	0	1	0	0	0	0	1	0	0
Peru	Escala de Conductas de Aprendizaje Preescolar (ECAP), a Spanish translation of the Preschool Learning Behaviors Scale (PLBS)	1	0	0	0	0	0	0	0	1	0	0
Rwanda	International Development and Early Learning Assessment (IDE LA)	1	1	1	1	1	1	1	0	0	0	0
Samoa	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1	0	1
Senegal	Milestones Checklist; Vocabulary Inventory	1	0	0	1	1	1	1	0	0	0	0
Sierra Leone	Saccadic reaction time (SRTs)	1	1	0	0	0	0	1	0	0	0	0
South Africa	Ages & Stages Questionnaire (ASQ)	1	0	0	0	1	0	0	0	0	0	0
South Africa	Ages & Stages Questionnaire-3 (ASQ-3)	2	0	0	2	0	0	1	0	1	0	0
South Africa	Bayley Infant Neurodevelopmental Screener (BINS)	1	0	0	1	0	0	1	0	1	0	0

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance selected by copyright, including for uses related to text and data mining, AI training, and similar technologies	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
South Africa	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	1	0	0	0	1	0	0	Pendergast 2018
South Africa	Early Learning Outcomes Measure (ELOM)	1	0	1	1	0	0	0	0	1	0	1	Snelling 2019
South Africa	Little Developmental Coordination Disorder Questionnaire (Little DCDQ)	1	0	0	1	0	1	0	0	0	0	0	Venter 2015
South Africa	Mullen Scales of Early Learning (MSEL)	1	0	0	0	1	0	0	0	0	0	0	Bornman 2010
South Africa	Parents' Evaluation of Developmental status (PEDS)	1	0	0	0	0	1	0	0	0	0	0	Abdoola 2021
South Africa	The infant gross motor screening test (IGMST)	1	1	1	0	1	1	0	0	0	0	0	Hilburn 2011
South Africa	The Road to Health Booklet (RTHB)	1	0	0	0	0	1	0	0	0	0	0	van der Linde 2015
Sri Lanka	Child Behaviour Assessment Instrument (CBAI)	1	1	0	1	1	1	0	0	0	0	0	Samarakkody 2010
Sri Lanka	Denver Developmental Screening Test for Sri Lankan Children (DDST-SL)	1	0	1	0	0	0	0	0	0	0	0	Wijedasa 2012

136/bmjopen-2024-096365 | Accepted by copyright, including for use related to teaching, AI training, and similar technologies.
 For peer review only - 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES).

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance			References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over time	Over other groups	First author & year	
Sri Lanka	Ragama Early Assessment for Children (REACH)	1	0	1	1	1	1	0	0	0	0	0	0	Caldera 2023
Tanzania	Caregiver Reported Early Development Instruments (CREDI)	1	1	1	1	1	1	1	0	0	0	0	0	McCoy 2017
Thailand	Parents' Evaluation of Developmental Status (PEDS)	1	0	0	0	0	0	1	0	0	0	0	0	Chunsawan 2016
Thailand	Thai Speech and Language Test (TSLT)	1	0	1	1	1	0	0	0	0	0	0	0	Prathanee 2008
The Gambia	Mullen Scales of Early Learning (MSEL)	1	0	0	0	0	0	1	1	0	0	0	0	Milosavljevic 2019
Tonga	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1	0	0	1	Sincovich 2019
Turkey	Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	1	1	1	1	0	1	0	0	0	0	0	0	Kucuker 2011
Turkey	Bayley Infant Neurodevelopmental Screener (BINS)	1	0	0	0	0	0	0	1	0	0	0	0	Soysal 2014
Turkey	Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	1	1	1	1	0	1	0	0	0	0	0	0	Karabekiroglu 2009
Turkey	Emotional Regulation Checklist	1	1	0	1	1	1	0	0	1	0	0	0	Danişman 2016

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance corrected by copyright, including for uses related to AI training, and similar technologies.	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Turkey	Kaufman Survey of Early Academic and Language Skills (K-SEALS)	1	1	1	1	1	0	0	0	1	0	0	Uyanik 2014
Turkey	Strengths and Difficulties Questionnaire (SDQ)	1	0	0	1	0	1	1	0	0	0	0	Dursun 2020
Tuvalu	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1	0	1	Sincovich 2019
Uganda	Bayley Scales of Infant Development, bead threading task	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	British Ability Scales – 3rd ed., block design task	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	International Development and Early Learning Assessment (IDELA)	1	0	0	0	0	0	0	0	1	0	0	Halpin 2019
Uganda	Kilifi Developmental Inventory, coin box task	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	Kilifi Picture Vocabulary Test	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	Modified Fagan Test for Infant Intelligence (FTII)	1	0	0	0	0	0	0	1	0	0	0	Familiar-Lopez 2022
Uganda	Movement Assessment Battery for Children, balancing on one leg	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance selected by copyright, including for uses related to text and data mining, AI training, and similar technologies.	First author & year	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Uganda	NEPSY verbal fluency task	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	NEPSY, knock tap game	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	NEPSY, sentence repetition task	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	Sky Search in the Tests of Everyday Attention for Children, Sky Search Task	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	Tap Once Tap Twice Task	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Uganda	Tobii Professional Studio eye tracking programming	1	0	0	0	0	1	0	0	0	0	0	Boivin 2017
Uganda	Wisconsin Card Sort Test	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Venezuela	Bayley Infant Neurodevelopmental Screener (BINS)	1	1	1	0	0	0	0	0	0	0	0	McCarthy 2012
Viet Nam	Vietnamese version of the Intelligibility in Context Scale (ICS-VN).	1	0	0	1	0	1	1	0	0	0	0	Phạm 2017
Vietnam	International Development and Early Learning Assessment (IDEA)	1	0	0	0	0	0	0	0	1	1	0	Halpin 2019

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Open access DOI: 10.4236/ojs.2024.096363	Copyright, including for uses related to text and data mining	Number of articles providing evidence on measurement invariance	References	
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural					
Zambia	Ages & Stages Questionnaire-3 (ASQ-3)	2	0	0	2	0	0	1	0	1			0	0	van Heerden 2017; Hsiao 2016
Zambia	Bayley Scales of Infant Development, Second Edition (BSID-II)	1	0	0	0	0	1	0	0	0			0	0	Biasini 2015
Zambia	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0			0	0	Pisani 2018

For peer review only - <http://bmjopen.bmj.com/site/about/guidelines.xhtml>

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from http://bmjopen.bmjjournals.org/ by Enseignement Supérieur (ABES) - AI training, and similar technologies. Protected by copyright. including for unauthorised distribution to text mining, data mining, AI training, and similar technologies.

Supplemental Table 5 List of child development assessment tools and countries where they were applied

Name assessment tool	Number of countries where the tool was applied	List of countries where the tool was applied	Number of articles where tool was applied	Age range at assessment (in months)	Type of administration	References (first author & year)
Ages & Stages Questionnaire (ASQ)	2	Iran; South Africa	1	4-60	Indirect/reported	Vameghi 2013; Bornman 2010
Ages & Stages Questionnaire Inventory (ASQ:I)	2	Bangladesh; China	1	1-27	Indirect/reported	Pitchik 2023; Xie 2017
Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	3	Brazil; China; Turkey	3	4-72	Indirect/reported	Anunciação 2019; Bian 2017; Kucuker 2011
Ages & Stages Questionnaire: Social-Emotional, First Edition (ASQ:SE-1)	1	China	1	15-63	Indirect/reported	Xie 2019 (Journal of Child & Dev Studies)
Ages & Stages Questionnaire-3 (ASQ-3)	11	Brazil; China; Colombia; Georgia; Guatemala; India; Indonesia; Mexico; Nepal; South Africa; Zambia	13	1-66	Indirect/reported	Filgueiras 2013; Yue 2019; Rubio-Codina 2016; Rubio-Codina 2020; Zirakashvili 2018; Colbert 2021; Kvistad 2013; Yunilda 2022; Ortiz-Leon 2018; Shrestha 2019; van Heerden 2017; Hsiao 2016; Bian 2012
Ages & Stages Questionnaires: Social Emotional (ASQ:SE-BR)	1	Brazil	1	21-66	Indirect/reported	Chen 2017
Alberta Infant Motor Scale (AIMS)	1	Brazil	4	0-19	Direct assessment	Saccani 2016; Siegle 2018; de Albuquerque 2018; Valentini 2012
Angkor Hospital for Children Developmental Milestone Assessment Tool (AHC DMAT)	1	Cambodia		1-72	Direct assessment indirect/reported	Ngoun 2012
Battelle Developmental Inventory Second Edition, BDI-2	2	Brazil; Colombia	3	0-95	Direct assessment indirect/reported	de Albuquerque 2022; Rubio-Codina 2016; Rubio-Codina 2020
Bayley Infant Neurodevelopmental Screener (BINS)	8	Argentina, Brazil, Ecuador, Bolivia, Colombia, Venezuela; Turkey; South Africa	4	3-24	Direct assessment	McCarthy 2012; Soysal 2014; Rodriguez 2020; Guedes 2011
Bayley Scales of Infant and Toddler Development (BSID)	1	Iran	1	1-42	Not reported	Azari 2017
Bayley Scales of Infant Development, bead threading task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010

1	Bayley Scales of Infant Development, Second Edition (BSID-II)	4	Bangladesh, India, Pakistan, Zambia	2	1-42	Direct assessment	Parveen 2014
2	Bayley Scales of Infant Development, Third Edition (BSID-III)	12	Bangladesh, Brazil, India, Nepal, Pakistan, Peru, South Africa; Colombia; Ethiopia; Cameroon; Kenya; Malaysia	11	6-42	Direct assessment	Pendergast 2018; Madaschi 2016; Rubio-Codina 2016; Rubio-Codina 2020; Hanlon 2016; Lohaus 2014; McHenry 2021; Zakaria 2012; Ranjitkar 2018; Manandhar 2016; Rasheed 2023
3	Bracken School Readiness Assessment, Third Edition (BSRA-3)	1	India	1	36-60	Direct assessment	Shah 2013
4	Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	3	Côte d'Ivoire; Ghana; Turkey	2	23-32	Indirect/reported	Barthel 2021; Karabekiroglu 2009
5	British Ability Scales – 3rd ed., block design task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
6	Caregiver Reported Early Development Instruments (CREDI)	18	17 countries not listed, Tanzania	2	0-35	Indirect/reported	Waldman 2021; McCoy 2017
7	Caregiver Reported Early Development Instruments (CREDI), long-form	2	China; India	2	5-36	Indirect/reported	Li 2020; Alderman 2021
8	Caregiver Reported Early Development Instruments (CREDI), short-form	18	17 countries not listed; Brazil	2	0-35	Indirect/reported	McCoy 2018; Altafim 2018
9	Child Behavior Checklist/1.5-5 (CBCL/1.5-5) and the Caregiver-Teacher Report Form (C-TRF)	1	China	1	50-71	Indirect/reported	Liu 2011
10	Child Behavior Scale (CBS)	1	Peru	1	24-72	Indirect/reported	Meyer 2011
11	Child Behaviour Assessment Instrument (CBAI)	1	Sri Lanka	1	48-72	Indirect/reported	Samarakkody 2010
12	Child behaviour checklist (CBCL)	1	Kenya	1	12-72	Indirect/reported	Kariuki 2016
13	Child Social Preference Scale	1	China	1	36-60	Indirect/reported	Li 2016
14	Childhood Nonverbal Communication Scale (CNCS)	1	Iran	1	3-18	Indirect/reported	Oryadi-Zanjani 2020
15	Children's Behavior Questionnaire (CBQ)	1	Iran	1	48-72	Indirect/reported	Golmohammadi 2022
16	Chinese Preschool Readiness Scale (CPRS)	1	China	1	31-62	Indirect/reported	Xie 2019 (Early Educ & Dev)
17	Chinese version of Urban Infant-Toddler Social and Emotional Assessment	1	China	1	12-36	Indirect/reported	Jianduan 2009
18	Chinese Vocabulary Checklist	1	China	1	12-30	Indirect/reported	Hao 2008
19	Communicative Development Inventory (CDI)	1	Kenya	1	6-36	Indirect/reported	Alcock 2015
20	Denver Developmental Screening Test for Sri Lankan Children (DDST-SL)	1	Sri Lanka	1	0-80	Direct assessment	Wijedasa 2012

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmj.com/> on June 13, 2025 at Agence Biographique de l'Etat. Copyright © The Authors. Licensee BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in other forms, provided the original author(s) and publisher are credited.

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). Attribution and similar technologies. © The Author(s) 2024. Licensee BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original author(s) and source are given.

1	Denver Developmental Screening Test (Denver-II)	3	Brazil; Colombia; Ethiopia	4	0-70.6	Direct assessment, indirect/report	Lopez Boo 2020; Rubio-Codina 2016; Rubio-Codina 2020; Abessa 2016
2	DEvelopmental assessment on an E-Platform (DEEP)	1	India	2	34-40	Direct assessment	Bhavnani 2019; Mukherjee 2020
3	Developmental Assessment Tool for Anganwadis (DATA)	1	India	1	19-36	Direct assessment	Nair 2009
4	Developmental Assessment Tool for Anganwadis (DATA-II)	1	India	1	36-48	Direct assessment	Nair 2012
5	Developmental Milestone Chart (DMC)	1	Cambodia	1	1-96	Direct assessment, indirect/report	Scherzer 2009
6	Developmental Milestones Checklist (DMC)	1	Kenya	1	2.53–12.06	Indirect/reported	Abubakar 2010
7	Dimensions of Mastery Questionnaire (DMQ-18)	1	Bangladesh	1	36-72	Indirect/reported	Shaoli 2019
8	Early Development Instrument	1	Indonesia	1	48-96	Indirect/reported	Brinkman 2017
9	Early Human Capability Index (eHCI)	7	Brazil, Kiribati, China, Laos, Samoa, Tonga, Tuvalu	1	24-72	Indirect/reported	Sincovich 2019
10	Early Learning Outcomes Measure (ELOM)	1	South Africa	1	50-69	Direct assessment	Snelling 2019
11	EDIN II (Evaluación del Desarrollo Integral de la niña y el niño)	1	Costa Rica	1	0-72	Direct assessment	Cordero, 2019
12	Elicited Imitation Test (EIT)	1	Iran	1	36-83	Direct assessment	Mehrani 2018
13	Emotional and Behavioral Screener (EBS)	1	China	1	48-84	Indirect/reported	Zhang 2018
14	Emotional Regulation Checklist	1	Turkey	1	48-60	Indirect/reported	Danişman 2016
15	Escala de Conductas de Aprendizaje Preescolar (ECAP), a Spanish translation of the Preschool Learning Behaviors Scale (PLBS)	1	Peru	1	24-72	Indirect/reported	Hahn 2009
16	Escala de Habilidades Sociales (Scale of Social Abilities)	1	Argentina	1	36-60	Indirect/reported	Lacunza 2009
17	Executive function battery	1	Pakistan	1		Direct assessment	Obradovic 2019
18	Executive Function Touch (EF Touch)	2	Kenya	3	36-72	Self-administration	Willoughby 2019 (Child Neuropsych); Willoughby 2019 (Dev Science); Toor 2022
19	Farsi Narrative Norms Instrument	1	Iran	1	60-72	Direct assessment	Soleymani 2016
20	Global Scales for Early Development	32	Not listed	1		Direct assessment, indirect/report	McCray 2023

1	Infant Neurological International Battery	1	Iran	1	4-18	Direct assessment	Soleimani 2007
2	Intelligibility in Context Scale (ICS) / ICS-Jamaican Creole	1	Jamaica	1	39-75	Indirect/report	Washington 2017
3	INTER- GROWTH-21st Project Neurodevelopment Assessment (INTER-NDA)	1	Grenada	1	22-30	Direct assessment indirect/report	Waechter 2022
4	International Development and Early Learning Assessment (IDELA)	17	Afghanistan, Ethiopia, Uganda, Vietnam, Bolivia, Bangladesh, Bhutan, Egypt, India, Indonesia, Mali, Malawi, Mozambique, Pakistan, Rwanda, Zambia, Ghana	4	36-96	Direct assessment	Halpin 2019; Pisani 2018; Wolf 2017; Pisani 2022
5	Kaufman Survey of Early Academic and Language Skills (K-SEALS)	1	Turkey		61-72	Direct assessment	Uyanik 2014
6	Kilifi Developmental Inventory	2	Ghana, Côte d'Ivoire, Kenya	2	6-35	Direct assessment	Barthel 2022; Abubakar 2008
7	Kilifi Developmental Inventory, coin box task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
8	Kilifi Picture Vocabulary Test	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
9	Little Developmental Coordination Disorder Questionnaire (Little DCDQ)	1	South Africa	1	36-49	Indirect/reported	Venter 2015
10	MacArthur Communicative Development Inventory Level II	1	Pakistan	1		Indirect/reported	Rasheed 2017
11	MacArthur-Bates Communicative Development Inventories CDI	1	Colombia	1	8-30	Indirect/reported	Lara Díaz 2011
12	MacArthur-Bates Communicative Development Inventories Short-Forms I vocabulary checklist	1	Colombia	2	6-42	Indirect/reported	Rubio-Codina 2016; Rubio-Codina 2020
13	MacArthur-Bates Communicative Development Inventories Short-Forms II vocabulary checklist	1	Colombia	2	6-42	Indirect/reported	Rubio-Codina 2016; Rubio-Codina 2020
14	Malawi Developmental Assessment Tool (MDAT)	1	Malawi	1	0-84	Direct assessment, indirect/reported	Gladstone 2010
15	Malawi Developmental Assessment Tool (MDAT), renamed Dominican Child Development Screening Tool	1	Dominican Republic	1	2-24	Direct assessment, indirect/reported	Sánchez-Vincitore 2019

136/bmjopen-2024-096365 on 31 May 2025. Downloaded from http://bmjopen.bmjjournals.org/ by Enseignement Supérieur (ABES). All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in whole or in part, without the prior permission of the copyright holder.

1	Malay Preschool Language Assessment Tool (MPLAT)	1	Malaysia	1	48-83	Direct assessment	Razak 2010
2	MELQO Measure of Development and Learning Outcomes	1	Laos	1	48-49	Direct assessment	Gomez 2022
3	Milestones Checklist; Vocabulary Inventory	1	Senegal	1	4-30	Indirect/report	Weber 2018
4	Modified Fagan Test for Infant Intelligence (FTII)	1	Uganda	1	6-12	Direct assessment	Familiar-Lopez 2022
5	Mongolian Rapid Baby Scale (MORBAS)	1	Mongolia	1	0-42	Indirect/report	Dagvadorj 2015
6	Motor Development Scale	1	Brazil	1	48-72	Not reported	Okuda 2020
7	Movement Assessment Battery for Children-second edition (MBAC-2)	1	China	1	36-72	Direct assessment	Hua 2013
8	Movement Assessment Battery for Children, balancing on one leg	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
9	Mullen Scales of Early Learning (MSEL)	4	Guatemala; South Africa; The Gambia	5	0-68	Direct assessment	Koura 2013; Boivin 2021; Colbert 2020; Bornman 2010; Milosavljevic 2019
10	NEPSY verbal fluency task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
11	NEPSY, knock tap game	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
12	NEPSY, sentence repetition task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
13	Oxford Neurodevelopment Assessment (OX-NDA)	1	Brazil	1		Direct assessment/indirect/report	Fernandes 2022
14	Parents' Evaluation of Developmental Status (PEDS)	3	India; South Africa; Thailand	3	0-30	Indirect/report	Mukherjee 2022; Abdoola 2021; Chunswan 2016
15	Preschool and Kindergarten Behavior Scale (PKBS)	1	Brazil	1	36-72	Indirect/report	Dias 2011
16	Preschool Language Scale-3 (PLS- 3)	1	Colombia	1	48-84	Direct assessment	Romero 2013
17	Preschool Learning Behavior Scale (PLBS)	1	China	1		Indirect/report	Wu 2019 (JPA)
18	Questionário para Avaliação do Desenvolvimento Infantil (QAD-PIPAS)	1	Brazil	1	0-59	Indirect/report	Venancio 2020
19	Ragama Early Assessment for Children (REACH)	1	Sri Lanka	1	24-60	Indirect/report	Caldera 2023
20	Rapid Assessment of Cognitive and Emotional Regulation (RACER)	1	Ghana	1	48-72	Self-administration	Yuan 2022
21	Rapid Neurodevelopmental Assessment (RNDA)	2	Bangladesh; Guatemala	3	0-60	Direct assessment	Khan 2010; Khan 2013; Thompson 2015
22	Receptive and expressive vocabulary test battery (English, Luo and Swahili)	1	Kenya	1	24-83	Direct assessment	Knauer 2019

136/bmjopen-2024-096365 on 11 May 2024. Downloaded from http://bmjopen.bmjjournals.org/ on 13 June 2025 at Agence Biographique de l'Enseignement Supérieur (ABES). For peer review only. Copyright © The Author(s). Licensee BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original author(s) and source are cited.

Saccadic reaction time (SRTs)	1	Sierra Leone	1	7-11.9	Self-administration	Leppänen 2022
Screening for Identification of Oral Language Difficulties by Preschool Teachers (SIOLD)	1	Brazil	1	51-65	Indirect/report	Puglisi 2020
Screening Test of Early Language Development-Test version (STELD-T)	1	India	1	6-21	Indirect/report	Pathak 2023
Self-ordered Pointing Task (SOPT)	1	Brazil	1	42-60	Self-administration	Sallum 2017
Sky Search in the Tests of Everyday Attention for Children, Sky Search Task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
Social Skills domain of System-Rating Scales (SSIS-RS)	1	China	1	36-60	Indirect/report	Wu 2019 (PS)
Spanish-language MacArthur–Bates Communicative Development Inventories short forms (S-CDI SFI and SFII)	1	Mexico	1	8-30	Indirect/report	Jackson-Maldonado 2013
Strengths and Difficulties Questionnaire (SDQ)	2	Pakistan; Turkey	2	20-51	Indirect/report	Finch 2018; Dursun 2020
Tap Once Tap Twice Task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
Test of Gross Motor Development (TGMD-2)	1	Brazil	1	36-60	Direct assessment	Bandeira 2020
Thai Speech and Language Test (TSLT)	1	Thailand	1	3-24	Direct assessment	Prathanee 2008
The Chinese Early Language and Communication Questionnaire	1	China	1	2-48	Indirect/report	Zhang 2021
The Chinese Intelligence Scale for Young Children	1	China	1	36-84	Direct assessment	Guo 2009
The infant gross motor screening test (IGMST)	1	South Africa	1	6-18	Direct assessment	Hilburn 2011
The language inventory	1	Bangladesh	1	12-18	Indirect/report	Hamadani 2010
The Road to Health Booklet (RTHB)	1	South Africa	1	6-12	Indirect/report	van der Linde 2015
The Test of Infant Motor Performance (TIMP)	1	Nepal	1	2-3	Direct assessment	Kvestad 2023
Tobii Professional Studio eye tracking programming	1	Uganda	1	44-65	Direct assessment	Boivin 2017
Vietnamese version of the Intelligibility in Context Scale (ICS-VN).	1	Viet Nam	1	24-71	Indirect/report	Pham 2017
Vineland Adaptive Behaviour Scale (VABS II)	1	India	1	36-108	Indirect/report	Kumar 2016
Vocabulary Checklists	2	Malawi, Ghana	1	12-17; 48-72	Indirect/report	Prado 2018

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmj.com/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - www.abes.fr. Copyright © 2024 The Author(s). Published by BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the [Creative Commons Attribution-NonCommercial License](https://creativecommons.org/licenses/by-nd/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and publisher are credited and that the original publication in this journal is cited, in accordance with the rules of the license. Any further distribution of this version of the article may be subject to the terms and conditions imposed by the copyright holder or licensor. Those who redistribute the article are responsible for ensuring that it is used in accordance with the terms of the License.

Wechsler Preschool and Primary Scale of Intelligence, thid edition (WPPSI-III)	1	Pakistan	1		Direct assessment	Rasheed 2018
WHO Gross Motor Milestones	2	Bangladesh; Colombia	3	3-42	Direct assessment, indirect/report	Hamadani 2013; Rubio-Codina 2016; Rubio-Codina 2020
Wisconsin Card Sort Test	1	Uganda	1	55-67	Direct assessment	Nampijja 2010

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

Supplemental Table 6 Number of articles providing psychometric evidence by world region and urban/rural location

Region	Urban	Rural	Urban-rural
Caribbean	2	0	0
Central America	1	3	0
East Asia	10	1	8
Oceania	0	0	1
North Africa and the Middle East	0	0	1
North America	1	0	0
South America	18	0	7
South Asia	11	13	9
Southeast Asia	5	1	3
Sub-Saharan Africa	10	15	13
West Asia	8	0	2

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES). Copyright © 2024 BMJ Publishing Group Ltd. All rights reserved. No reuse allowed without permission.

Supplemental Table 7 Number of articles providing psychometric evidence by child development assessment tools (development articles)

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance			References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over countries	Over time	Over other groups	
Angkor Hospital for Children Developmental Milestone Assessment Tool (AHC DMAT)	1	1	0	0	0	1	0	0	0	0	0	0	0	Ngoun 2012
Caregiver Reported Early Development Instruments (CREDI)	2	18	2	1	2	2	2	2	0	1	0	0	0	McCoy 2017; Waldman 2021
Caregiver Reported Early Development Instruments (CREDI), short-form	1	17	1	0	1	1	1	1	0	0	0	0	1	McCoy 2018
Child Behaviour Assessment Instrument (CBAI)	1	1	1	0	1	1	1	0	0	0	0	0	0	Samarakkody 2010
Childhood Nonverbal Communication Scale (CNCS)	1	1	0	0	1	1	1	1	0	1	0	0	0	Oryadi-Zanjani 2020
Chinese Preschool Readiness Scale (CPRS)	1	1	0	0	1	1	1	0	0	1	0	0	0	Xie 2019 (Early Educ & Dev)
Chinese Vocabulary Checklist	1	1	0	0	0	0	0	1	0	0	0	0	0	Hao 2008
DEvelopmental assessment on an E-Platform (DEEP)	1	1	0	0	0	1	0	0	0	0	0	0	0	Bhavnani 2019
Developmental Assessment Tool for Anganwadis (DATA)	1	1	0	0	1	1	0	0	0	1	0	0	0	Nair 2009
Developmental Assessment Tool for Anganwadis (DATA-II)	1	1	0	0	1	1	0	0	0	1	0	0	0	Nair 2012
Developmental Milestone Chart (DMC)	1	1	0	0	0	1	0	0	0	0	0	0	0	Scherzer 2009
Developmental Milestones Checklist (DMC)	1	1	1	0	1	1	1	1	0	1	0	0	0	Abubakar 2010
Early Learning Outcomes Measure (ELOM)	1	1	0	1	1	0	0	0	0	1	0	0	1	Snelling 2019
Elicited Imitation Test (EIT)	1	1	0	1	0	0	1	0	0	0	0	0	1	Mehrani 2018
Escala de Habilidades Sociales (Scale of Social Abilities)	1	1	0	0	1	0	0	0	0	1	0	0	0	Lacunza 2009

136/bmjopen-2024-003650 © 2024 BMJ Publishing Group. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution, and reproduction in other forms, provided the original author(s) and publisher are credited and a link is made to the journal's website.

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Farsi Narrative Norms Instrument	1	1	1	1	0	1	0	0	0	0	0	Soleymani 2016
Global Scales for Early Development	1	32	0	0	1	1	0	0	0	0	0	McCray 2023
International Development and Early Learning Assessment (IDELA)	1	12	1	1	1	1	1	1	0	0	0	Pisani 2018
Kilifi Developmental Inventory	1	1	1	1	1	1	1	1	0	0	0	Abubakar 2008
Malawi Developmental Assessment Tool (MDAT)	1	1	1	1	0	1	0	1	0	0	0	Gladstone 2010
Malay Preschool Language Assessment Tool (MPLAT)	1	1	1	1	1	1	0	0	0	0	0	Razak 2010
Mongolian Rapid Baby Scale (MORBAS)	1	1	0	0	0	1	1	0	0	0	0	Dagvadorj 2015
Oxford Neurodevelopment Assessment (OX-NDA)	1	1	1	1	1	0	1	0	0	0	0	Fernandes 2022
Questionário para Avaliação do Desenvolvimento Infantil (QAD-PIPAS)	1	1	0	0	0	1	0	0	0	0	0	Venancio 2020
Ragama Early Assessment for Children (REACH)	1	1	0	1	1	1	1	0	0	0	0	Caldera 2023
Rapid Neurodevelopmental Assessment (RNDA)	1	1	0	1	0	1	1	0	1	0	0	Khan 2010
Receptive and expressive vocabulary test battery (English, Luo and Swahili)	1	1	0	1	1	0	1	1	1	0	0	Knauer 2019
Saccadic reaction time (SRTs)	1	1	1	0	0	0	0	1	0	0	0	Leppänen 2022
Screening for Identification of Oral Language Difficulties by Preschool Teachers (SIOLD)	1	1	0	0	0	0	0	0	0	1	0	Puglisi 2020
Screening Test of Early Language Development-Test version (STELD-T)	1	1	0	0	1	0	1	0	0	0	0	Pathak 2023
Thai Speech and Language Test (TSLT)	1	1	0	1	1	1	0	0	0	0	0	Prathanee 2008
The infant gross motor screening test (IGMST)	1	1	1	1	0	1	1	0	0	0	0	Hilburn 2011
The language inventory	1	1	1	0	0	1	1	1	1	0	0	Hamadani 2010

136/bmjopen2024-00365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). Enseignement Supérieur (ABES) is the national library of France. It collects, preserves, makes available and promotes the production of knowledge in the fields of higher education, research and training. It also supports the development of digital technologies for learning and teaching.

Name of assessment tool	Number of articles providing psycho-metric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance			References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over other countries	Over time	Over other groups	
Vocabulary Checklists	1	1	0	0	0	0	1	0	1	0	0	0	0	Prado 2018

136/bmjopen2024-00365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES). Copyright © 2024 BMJ Publishing Group Ltd. All rights reserved. No reuse allowed without permission.

Supplemental Table 8 Number of articles providing psychometric evidence by child development assessment tools (adaptation articles)

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement equivalence including for uses related to text and data mining, AI training, and similar technologies.	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Ages & Stages Questionnaire (ASQ)	2	2	0	1	1	2	0	0	0	1	0	Vameghi 2013; Bornman 2010
Ages & Stages Questionnaire Inventory (ASQ:I)	2	2	1	1	2	1	2	1	0	0	0	Xie 2017; Pitchik 2023
Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	3	3	2	2	2	1	1	0	0	2	0	Kucuker 2011; Anunciação 2019; Bian 2017
Ages & Stages Questionnaire: Social-Emotional, First Edition (ASQ:SE-1)	1	1	1	1	1	0	1	0	0	0	0	Xie 2019 (Journal of Child & Dev Studies)
Ages & Stages Questionnaire-3 (ASQ-3)	13	11	5	5	12	3	5	5	2	4	0	Zirakashvili 2018; Colbert 2021; Rubio-Codina 2016; Rubio-Codina 2020; Shrestha 2019; Hsiao 2016; Filgueiras 2013; van Heerden 2017; Ortiz-Leon 2018; Kvistad 2013; Yue 2019; Yunilda 2022; Bian 2012
Ages & Stages Questionnaires: Social Emotional (ASQ:SE-BR)	1	1	0	0	0	0	0	0	0	1	0	Chen 2017
Alberta Infant Motor Scale (AIMS)	4	1	3	3	2	1	3	1	1	0	0	Saccani 2016; Siegle 2018; de Albuquerque 2018; Valentini 2012
Battelle Developmental Inventory Second Edition, BDI-2	3	2	2	0	2	1	1	1	1	0	0	Rubio-Codina 2016; Rubio-Codina 2020; de Albuquerque 2022
Bayley Infant Neurodevelopment Screener (BINS)	4	8	1	1	2	1	1	1	2	2	0	Guedes 2011; McCarthy 2012; Soysal 2014; Rodriguez 2020
Bayley Scales of Infant and Toddler Development (BSID)	1	1	1	1	1	1	0	1	0	1	0	Azari 2017

136/bmjopen-2024-096365 On 31 May 2025. Downloaded from <http://bmjopen.bmjjournals.com/> at Agence Bibliographique de l'Enseignement Supérieur (ABES).
Copyright © The Author(s) 2024. Published by BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (CC BY-NC 4.0), which permits non-commercial reuse, distribution, and reproduction in its entirety, without prior permission or licensing from the author(s) or publisher. For commercial reuse of this article, please go to the BMJ Publishing Group Ltd website (www.bmjjournals.com) and follow the BMJ Publishing Group Ltd Terms and Conditions.
For peer review only - <http://bmjopen.bmjjournals.com/site/about/guidelines.xhtml>

136/bmjopen.2024.06655 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - At training and similar technologies.

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement equivalence	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Bayley Scales of Infant Development, bead threding task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Bayley Scales of Infant Development, Second Edition (BSID-II)	2	4	0	0	0	1	1	0	0	0	0	Biasini 2015; Parveen 2014
Bayley Scales of Infant Development, Third Edition (BSID-III)	11	12	4	3	7	2	3	3	4	5	2	Rubio-Codina 2016; Rubio-Codina 2020; Hanlon 2016; Ranjitkar 2018; Madaschi 2016; Pendergast 2018; Manandhar 2016; Zakaria 2012; Lohaus 2014; McHenry 2021; Rasheed 2023
Bracken School Readiness Assessment, Third Edition (BSRA-3)	1	1	0	0	1	1	0	0	0	0	0	Shah 2013
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	2	3	1	1	2	0	1	1	0	1	0	Karabekiroglu 2009; Barthel 2021
British Ability Scales – 3rd ed., block design task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Caregiver Reported Early Development Instruments (CREDI), long-form	2	2	0	0	2	0	2	2	0	0	0	Alderman 2021; Li 2020
Caregiver Reported Early Development Instruments (CREDI), short-form	1	1	1	0	1	1	1	1	0	0	0	Altafim 2018
Child Behavior Checklist/1.5-5 (CBCL/1.5-5) and the Caregiver-Teacher Report Form (C-TRF)	1	1	0	0	0	0	0	0	0	1	0	Liu 2011
Child Behavior Scale (CBS)	1	1	0	0	1	0	0	0	0	1	0	Meyer 2011
Child behaviour checklist (CBCL)	1	1	1	0	1	0	0	1	0	1	0	Kariuki 2016
Child Social Preference Scale	1	1	0	0	1	0	0	0	1	1	0	Li 2016
Children's Behavior Questionnaire (CBQ)	1	1	1	0	1	1	1	1	0	1	0	Golmohammadi 2022
Chinese version of Urban Infant-Toddler Social and Emotional Assessment	1	1	1	1	1	1	1	0	0	1	0	Jianduan 2009

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement equivalence			References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over countries	Over time	Over other groups	
Communicative Development Inventory (CDI)	1	1	1	0	1	0	1	0	0	0	0	0	0	Alcock 2015
Denver Developmental Screening Test for Sri Lankan Children (DDST-SL)	1	1	0	1	0	0	0	0	0	0	0	0	0	Wijedasa 2012
Denver Developmental Screening Test (Denver-II)	4	3	3	2	3	2	1	2	1	1	0	0	0	Abessa 2016; Rubio-Codina 2016; Rubio-Codina 2020; Lopez Boo 2020
DEvelopmental assessment on an E-Platform (DEEP)	1	1	0	0	0	0	1	0	0	0	0	0	0	Mukherjee 2020
Dimensions of Mastery Questionnaire (DMQ-18)	1	1	1	1	1	1	1	0	0	1	0	0	0	Shaoli 2019
Early Development Instrument	1	1	0	1	1	1	1	0	1	0	0	0	0	Brinkman 2017
Early Human Capability Index (eHCI)	1	7	0	0	1	0	0	0	0	1	0	1	1	Sincovich 2019
EDIN II (Evaluación del Desarrollo Integral de la niña y el niño)	1	1	0	0	0	1	0	1	0	0	0	0	0	Cordero, 2019
Emotional and Behavioral Screener (EBS)	1	1	0	1	1	0	1	0	0	1	0	0	0	Zhang 2018
Emotional Regulation Checklist	1	1	1	0	1	1	1	0	0	1	0	0	0	Danişman 2016
Escala de Conductas de Aprendizaje Preescolar (ECAP), a Spanish translation of the Preschool Learning Behaviors Scale (PLBS)	1	1	0	0	0	0	0	0	0	1	0	0	0	Hahn 2009
Executive function battery	1	1	0	0	1	1	0	1	1	1	0	0	0	Obradovic 2019
Executive Function Touch (EF Touch)	3	2	0	1	1	1	1	3	0	2	0	0	0	Willoughby 2019 (Child Neuropsych); Willoughby 2019 (Dev Science); Toor 2022
Infant Neurological International Battery	1	1	0	1	0	0	1	0	0	0	0	0	0	Soleimani 2007
Intelligibility in Context Scale (ICS) / ICS-Jamaican Creole	1	1	1	1	1	0	1	1	0	1	0	0	0	Washington 2017

136/bmjopen.2024.026655 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). For peer review only.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement equivalence			References
				Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over countries	Over time	Over other groups	
INTER- GROWTH-21st Project Neurodevelopment Assessment (INTER-NDA)	1	1	1	1	1	1	0	0	0	0	0	0	0	0	Waechter 2022
International Development and Early Learning Assessment (IDELA)	3	17	1	1	1	1	0	1	1	1	2	1	1	1	Halpin 2019; Pisani 2022; Wolf 2017
Kaufman Survey of Early Academic and Language Skills (K-SEALS)	1	1	1	1	1	1	1	0	0	0	1	0	0	0	Uyanik 2014
Kilifi Developmental Inventory	1	2	0	1	1	1	0	0	0	1	1	0	0	0	Barthel 2022
Kilifi Developmental Inventory, coin box task	1	1	1	0	1	1	0	0	1	0	1	0	0	0	Nampijja 2010
Kilifi Picture Vocabulary Test	1	1	1	0	1	1	0	0	1	0	1	0	0	0	Nampijja 2010
Little Developmental Coordination Disorder Questionnaire (Little DCDQ)	1	1	0	0	1	1	0	1	0	0	0	0	0	0	Venter 2015
MacArthur Communicative Development Inventory Level II	1	1	1	1	1	1	1	1	0	0	0	0	0	0	Rasheed 2017
MacArthur-Bates Communicative Development Inventories CDI	1	1	0	0	1	1	0	1	0	0	0	0	0	0	Lara Diaz 2011
MacArthur-Bates Communicative Development Inventories Short-Forms I vocabulary checklist	2	1	2	0	2	0	0	1	1	1	0	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020
MacArthur-Bates Communicative Development Inventories Short-Forms II vocabulary checklist	2	1	2	0	2	0	0	1	1	1	0	0	0	0	Rubio-Codina 2016; Rubio-Codina 2020
Malawi Developmental Assessment Tool (MDAT), renamed Dominican Child Development Screening Tool	1	1	0	1	1	1	0	0	1	0	0	0	0	0	Sánchez-Vincitore 2019
MELQO Measure of Development and Learning Outcomes	1	1	0	0	0	0	0	0	1	0	1	1	0	1	Gomez 2022
Milestones Checklist; Vocabulary Inventory	1	1	0	0	1	1	1	1	1	0	0	0	0	0	Weber 2018
Modified Fagan Test for Infant Intelligence (FTII)	1	1	0	0	0	0	0	0	0	1	0	0	0	0	Familiar-Lopez 2022

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement equivalence	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Motor Development Scale	1	1	0	0	1	0	0	0	0	1	0	Okuda 2020
Movement Assessment Battery for Children- second edition (MBAC-2)	1	1	1	1	1	1	1	0	0	1	0	Hua 2013
Movement Assessment Battery for Children, balancing on one leg	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Mullen Scales of Early Learning (MSEL)	5	4	1	2	1	1	1	3	3	0	0	Colbert 2020; Koura 2013; Milosavljevic 2019; Bornman 2010; Boivin 2021
NEPSY verbal fluency task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
NEPSY, knock tap game	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
NEPSY, sentence repetition task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Parents' Evaluation of Developmental status (PEDS)	3	3	0	0	0	0	3	0	0	0	0	Abdoola 2021; Chunswan 2016; Mukherjee 2022
Preschool and Kindergarten Behavior Scale (PKBS)	1	1	0	0	1	0	0	0	0	1	0	Dias 2011
Preschool Language Scale-3 (PLS- 3)	1	1	0	0	1	0	0	0	0	1	0	Romero 2013
Preschool Learning Behavior Scale (PLBS)	1	1	0	0	1	0	1	0	0	1	0	Wu 2019 (JPA)
Rapid Assessment of Cognitive and Emotional Regulation (RACER)	1	1	1	1	0	0	1	1	0	0	0	Yuan 2022
Rapid Neurodevelopmental Assessment (RNDA)	2	2	0	2	0	1	2	1	0	0	0	Thompson 2015; Khan 2013
Self-ordered Pointing Task (SOPT)	1		0	0	0	0	1	1	0	0	0	Sallum 2017
Sky Search in the Tests of Everyday Attention for Children, Sky Search Task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Social Skills domain of System-Rating Scales (SSIS-RS)	1	1	0	0	1	0	0	1	0	1	0	Wu 2019 (PS)
Spanish-language MacArthur–Bates Communicative Development Inventories short forms (S-CDI SFI and SFII)	1	1	1	0	0	0	0	1	0	0	0	Jackson-Maldonado 2013

136/bmjopen.2024.66565 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> at Agence Bibliographique de l'Enseignement Supérieur (ABES). © 2024 The Author(s). Published by BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (CC BY-NC 4.0), which permits use, distribution and reproduction in other formats, provided the original author(s) and publisher are credited and a link is made to the journal's content page or full-text PDF in the BMJ Open service. See: <http://creativecommons.org/licenses/by-nc/4.0/>

136/bmjopen.2024.026655 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). Attributions required. © 2024 BMJ Publishing Group Ltd. All rights reserved.

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement equivalence		References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Over countries	Over other groups	
Strengths and Difficulties Questionnaire (SDQ)	2	2	1	1	2	1	2	2	0	1	0	0	Finch 2018; Dursun 2020
Tap Once Tap Twice Task	1	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010
Test of Gross Motor Development (TGMD-2)	1	1	0	1	1	0	0	1	0	1	0	0	Bandeira 2020
The Chinese Early Language and Communication Questionnaire	1	1	0	0	1	0	1	0	0	0	0	0	Zhang 2021
The Chinese Intelligence Scale for Young Children	1	1	0	0	0	0	0	0	0	1	0	1	Guo 2009
The Road to Health Booklet (RTHB)	1	1	0	0	0	0	1	0	0	0	0	0	van der Linde 2015
The Test of Infant Motor Performance (TIMP)	1	1	0	1	1	0	1	0	1	0	0	0	Kvestad 2023
Tobii Professional Studio eye tracking programming	1	1	0	0	0	0	1	0	0	0	0	0	Boivin 2017
Vietnamese version of the Intelligibility in Context Scale (ICS-VN).	1	1	0	0	1	0	1	1	0	0	0	0	Pham 2017
Vineland Adaptive Behaviour Scale (VABS II)	1	1	1	1	0	0	0	1	0	0	0	0	Kumar 2016
Wechsler Preschool and Primary Scale of Intelligence, third edition (WPPSI-III)	1	1	0	1	1	1	0	1	0	1	0	0	Rasheed 2018
WHO Gross Motor Milestones	3	2	2	1	2	0	2	2	2	0	0	0	Hamadani 2013; Rubio-Codina 2016; Rubio-Codina 2020
Wisconsin Card Sort Test	1	1	1	0	1	0	0	1	0	1	0	0	Nampijja 2010

Supplemental Table 9 Number of articles providing psychometric evidence by article type and early childhood development domain

Child development domain	Total number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance		
		Test-retest	Inter-rater	Internal consistency	Content/face	Concurrent /Criterion	Convergent	Predictive	Structural	Over countries	Over time	Over other groups
All articles (N=175 article-tool dyads)												
Cognitive	77	28	29	52	35	34	28	16	22	0	3	
Language	105	39	36	70	44	49	40	19	28	0	6	
Motor	98	38	39	60	36	44	37	19	28	0	4	
Socioemotional/temperament	47	19	15	35	20	28	15	3	24	1	4	
Attention/executive function	20	9	3	12	2	5	15	1	16	0	3	
Personal-social/adaptive	49	19	18	32	20	22	16	6	12	0	1	
Academic/preacademic	8	1	3	4	1	2	2	1	7	0	4	
Development articles (N=35 article-tool dyads)												
Cognitive	17	7	7	12	14	10	6	1	5	0	2	
Language	27	9	10	15	20	15	9	4	7	0	2	
Motor	19	9	9	12	17	11	7	1	5	0	2	
Socioemotional/temperament	12	6	4	9	11	9	4	0	2	0	1	
Attention/executive function	1	0	1	1	0	0	0	0	1	0	1	
Personal-social/adaptive	10	2	3	6	9	4	2	1	4	0	0	
Academic/preacademic	1	0	1	1	0	0	0	0	1	0	1	
Adaptation articles (N=140 article-tool dyads)												
Cognitive	60	21	22	40	21	24	22	15	17	1	0	1
Language	78	30	26	55	24	34	31	15	21	2	0	4
Motor	79	29	30	48	19	33	30	18	23	3	0	2
Socioemotional/temperament	35	13	11	26	9	19	11	3	22	2	1	3

136/bmjopen-2024-096365 © 2024 BMJ Publishing Group Ltd. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>). See the terms and conditions (<http://creativecommons.org/licenses/by-nc/4.0/legalcode>) at <http://bmjopen.bmjjournals.com>.

Attention/ executive function	19	9	2	11	2	5	15	1	15	0	2
Personal-social/ adaptive	39	17	15	26	11	18	14	5	8	0	1
Academic/ preacademic	7	1	1	3	1	2	2	1	6	0	3

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmj.com/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - including for uses related to text and data mining, AI training, and similar technologies.

For peer review only

1
2
3
4
5
6
Supplemental Table 10 Number of articles by category of risk of bias
7
8
9

	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data
Low risk	42	80	74
Some concerns	11	8	49
High risk	0	1	7
Unable to assess	122	8	44
Not applicable	0	78	1

	Indirectness (are results generalisable?)
Yes	18
Probably yes	81
Probably no	49
No	24
No information	3

Supplemental Table 11 Risk of bias assessment by included article

First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Abdoola 2021	unable to assess	N/A	unable to assess	probably no
Abessa 2016	low risk	N/A	unable to assess	probably yes
Abubakar 2008	unable to assess	Low risk	low risk	probably yes
Abubakar 2010	unable to assess	Some concerns	High risk	No
Alcock 2015	unable to assess	N/A	low risk	Probably yes
Alderman 2021	Low risk	unable to assess	Low risk	Yes
Altafim 2018	low risk	unable to assess	some concerns	probably yes
Anunciação 2019	unable to assess	n/a	some concerns	Yes
Azari 2017	unable to assess	high risk	unable to assess	probably no
Bandeira 2020	unable to assess	low risk	Low risk	probably yes
Barthel 2021	unable to assess	low risk	low risk	probably yes
Barthel 2022	Some concerns	Low risk	Low risk	Probably no
Bhavnani 2019	unable to assess	N/A	low risk	Yes
Bian 2012	unable to assess	N/A	Low risk	no
Bian 2017	unable to assess	n/a	unable to assess	Yes
Biasini 2015	unable to assess	N/A	Some concerns	Probably yes
Boivin 2017	unable to assess	Low risk	Low risk	No
Boivin 2021	unable to assess	low risk	low risk	probably yes
Bornman 2010	unable to assess	N/A	Low risk	no
Bornman 2010	unable to assess	N/A	low risk	no
Brinkman 2017	some concerns	Low risk	unable to assess	probably yes
Caldera 2023	unable to assess	N/A	unable to assess	Yes
Chen 2017	unable to assess	N/A	some concerns	yes
Chunsuwan 2016	unable to assess	N/A	unable to assess	probably no
Colbert 2020	low risk	low risk	unable to assess	probably yes
Colbert 2021	unable to assess	low risk	Low risk	No information
Cordero, 2019	unable to assess	N/A	unable to assess	Yes
Dagvadorj 2015	unable to assess	N/A	Low risk	Probably no
Danişman 2016	unable to assess	N/A	high risk	no
de Albuquerque 2018	unable to assess	N/A	Low risk	probably yes
de Albuquerque 2022	unable to assess	N/A	Low risk	probably yes
Dias 2011	unable to assess	N/A	low risk	Probably yes
Dursun 2020	unable to assess	low risk	Some concerns	probably no
Familiar-Lopez 2022	unable to assess	Some concerns	Some concerns	No
Fernandes 2022	unable to assess	low risk	low risk	probably yes
Filgueiras 2013	unable to assess	N/A	unable to assess	probably yes
Finch 2018	low risk	Low risk	low risk	probably no
Gladstone 2010	unable to assess	low risk	some concerns	Probably yes

First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Golmohammadi 2022	unable to assess	Low risk	low risk	probably no
Gomez 2022	unable to assess	Low risk	Low risk	Probably yes
Guedes 2011	unable to assess	low risk	unable to assess	probably no
Guo 2009	unable to assess	N/A	unable to assess	Probably yes
Hahn 2009	unable to assess	N/A	high risk	probably yes
Halpin 2019	unable to assess	N/A	unable to assess	yes
Hamadani 2010	low risk	low risk	high risk	probably yes
Hamadani 2013	unable to assess	low risk	low risk	Probably yes
Hanlon 2016	Some concerns	Some concerns	Low risk	Probably no
Hao 2008	unable to assess	low risk	unable to assess	no
Hilburn 2011	low risk	N/A	unable to assess	Probably yes
Hsiao 2016	some concerns	low risk	low risk	probably yes
Hua 2013	unable to assess	N/A	Some concerns	probably yes
Jackson-Maldonado 2013	unable to assess	low risk	Low risk	Yes
Jianduan 2009	unable to assess	N/A	unable to assess	probably yes
Karabekiroglu 2009	unable to assess	N/A	low risk	no
Kariuki 2016	Some concerns	Low risk	Some concerns	Probably yes
Khan 2010	unable to assess	Some concerns	low risk	Probably no
Khan 2013	unable to assess	N/A	unable to assess	probably yes
Knauer 2019	unable to assess	low risk	some concerns	probably yes
Koura 2013	low risk	unable to assess	unable to assess	probably no
Kucuker 2011	unable to assess	N/A	some concerns	Probably yes
Kumar 2016	unable to assess	some concerns	low risk	probably no
Kvestad 2013	Low risk	N/A	some concerns	Probably yes
Kvestad 2023	low risk	low risk	some concerns	probably no
Lacunza 2009	unable to assess	N/A	unable to assess	probably no
Lara Díaz 2011	unable to assess	low risk	Low risk	Probably yes
Leppänen 2022	unable to assess	low risk	high risk	probably no
Li 2016	unable to assess	Low risk	Some concerns	No
Li 2020	unable to assess	unable to assess	low risk	probably yes
Liu 2011	unable to assess	N/A	Some concerns	probably yes
Lohaus 2014	unable to assess	Low risk	Some concerns	No
Lopez Boo 2020	Some concerns	low risk	some concerns	probably yes
Madashchi 2016	unable to assess	N/A	Some concerns	probably yes
Manandhar 2016	low risk	N/A	unable to assess	probably no
McCarthy 2012	low risk	N/A	unable to assess	probably yes
McCoy 2017	Low risk	low risk	low risk	yes
McCoy 2018	unable to assess	low risk	low risk	yes
McCray 2023	unable to assess	N/A	Low risk	Probably yes
McHenry 2021	low risk	low risk	some concerns	Probably no
Mehrani 2018	unable to assess	N/A	Low risk	Probably no
Meyer 2011	unable to assess	N/A	unable to assess	Probably no

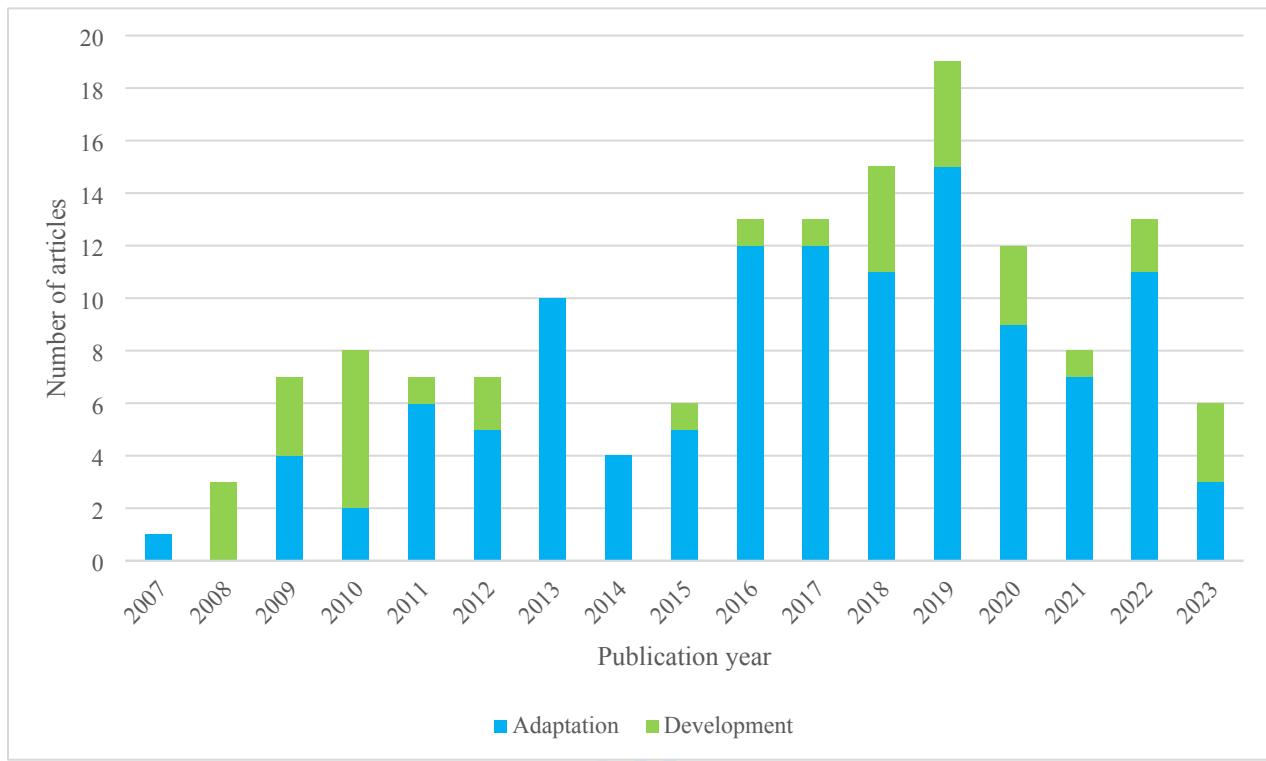
First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Milosavljevic 2019	low risk	low risk	some concerns	no
Mukherjee 2020	unable to assess	N/A	low risk	No
Mukherjee 2022	unable to assess	N/A	low risk	probably yes
Nair 2009	unable to assess	N/A	unable to assess	probably yes
Nair 2012	unable to assess	N/A	unable to assess	probably yes
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Ngoun 2012	unable to assess	N/A	unable to assess	yes
Obradovic 2019	unable to assess	Low risk	low risk	probably yes
Okuda 2020	unable to assess	N/A	unable to assess	probably yes
Ortiz-Leon 2018	unable to assess	N/A	some concerns	probably no
Oryadi-Zanjani 2020	some concerns	N/A	low risk	probably yes
Parveen 2014	unable to assess	N/A	N/A	no
Pathak 2023	unable to assess	N/A	low risk	probably no
Pendergast 2018	low risk	N/A	Some concerns	yes
Phạm 2017	unable to assess	low risk	low risk	probably no
Pisani 2018	unable to assess	low risk	unable to assess	yes
Pisani 2022	unable to assess	low risk	some concerns	probably yes
Pitchik 2023	low risk	low risk	low risk	probably yes
Prado 2018	Low risk	Low risk	low risk	No
Prathanee 2008	low risk	N/A	low risk	probably yes
Puglisi 2020	Some concerns	N/A	low risk	Probably yes
Ranjitkar 2018	low risk	low risk	low risk	probably yes
Rasheed 2017	some concerns	N/A	unable to assess	probably no
Rasheed 2018	low risk	Low risk	unable to assess	probably yes
Rasheed 2023	unable to assess	Low risk	unable to assess	no information
Razak 2010	unable to assess	N/A	low risk	probably no
Rodriguez 2020	low risk	unable to assess	unable to assess	probably yes
Romero 2013	unable to assess	N/A	unable to assess	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes

First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	Some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	Some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	low risk	Low risk	some concerns	Probably yes
Saccani 2016	low risk	N/A	unable to assess	yes
Sallum 2017	unable to assess	low risk	some concerns	probably yes
Samarakkody 2010	unable to assess	N/A	unable to assess	yes
Sánchez-Vincitore 2019	low risk	Low risk	low risk	probably no
Scherzer 2009	unable to assess	N/A	unable to assess	probably yes
Shah 2013	unable to assess	n/a	unable to assess	no
Shaoli 2019	unable to assess	n/a	low risk	probably yes
Shrestha 2019	some concerns	low risk	some concerns	probably no
Siegle 2018	unable to assess	N/A	some concerns	probably no
Sincovich 2019	unable to assess	N/A	some concerns	probably no
Snelling 2019	unable to assess	N/A	some concerns	probably yes
Soleimani 2007	low risk	N/A	Some concerns	no information
Soleymani 2016	unable to assess	N/A	Low risk	No
Soysal 2014	unable to assess	some concerns	some concerns	no
Thompson 2015	low risk	unable to assess	unable to assess	probably no
Toor 2022	unable to assess	Low risk	some concerns	No
Uyanik 2014	unable to assess	N/A	Some concerns	Probably no
Valentini 2012	unable to assess	Low risk	unable to assess	Probably no
Vameghi 2013	unable to assess	N/A	high risk	probably no
van der Linde 2015	unable to assess	low risk	low risk	probably no
van Heerden 2017	low risk	N/A	unable to assess	no
Venancio 2020	unable to assess	N/A	Some concerns	Probably yes
Venter 2015	unable to assess	N/A	high risk	probably no
Waechter 2022	low risk	N/A	unable to assess	probably yes
Waldman 2021	unable to assess	Some concerns	Low risk	Probably yes
Washington 2017	unable to assess	Low risk	Some concerns	No
Weber 2018	unable to assess	low risk	some concerns	probably yes
Wijedasa 2012	low risk	N/A	low risk	yes
Willoughby 2019 (Child Neuropsych)	low risk	low risk	low risk	probably yes

First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Willoughby 2019 (Dev Science)	Low risk	Low risk	Some concerns	Probably yes
Wolf 2017	Some concerns	low risk	unable to assess	no
Wu 2019 (JPA)	unable to assess	n/a	unable to assess	Probably yes
Wu 2019 (PS)	unable to assess	Low risk	low risk	probably yes
Xie 2017	low risk	N/A	some concerns	no
Xie 2019 (Early Educ & Dev)	unable to assess	N/A	Some concerns	Probably no
Xie 2019 (Journal of Child & Dev Studies)	unable to assess	N/A	some concerns	probably yes
Yuan 2022	low risk	low risk	low risk	probably yes
Yue 2019	low risk	unable to assess	low risk	yes
Yunilda 2022	unable to assess	unable to assess	low risk	probably no
Zakaria 2012	unable to assess	N/A	unable to assess	no
Zhang 2018	unable to assess	N/A	some concerns	Probably no
Zhang 2021	unable to assess	n/a	unable to assess	Probably yes
Zirakashvili 2018	unable to assess	Some concerns	Some concerns	Probably yes

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

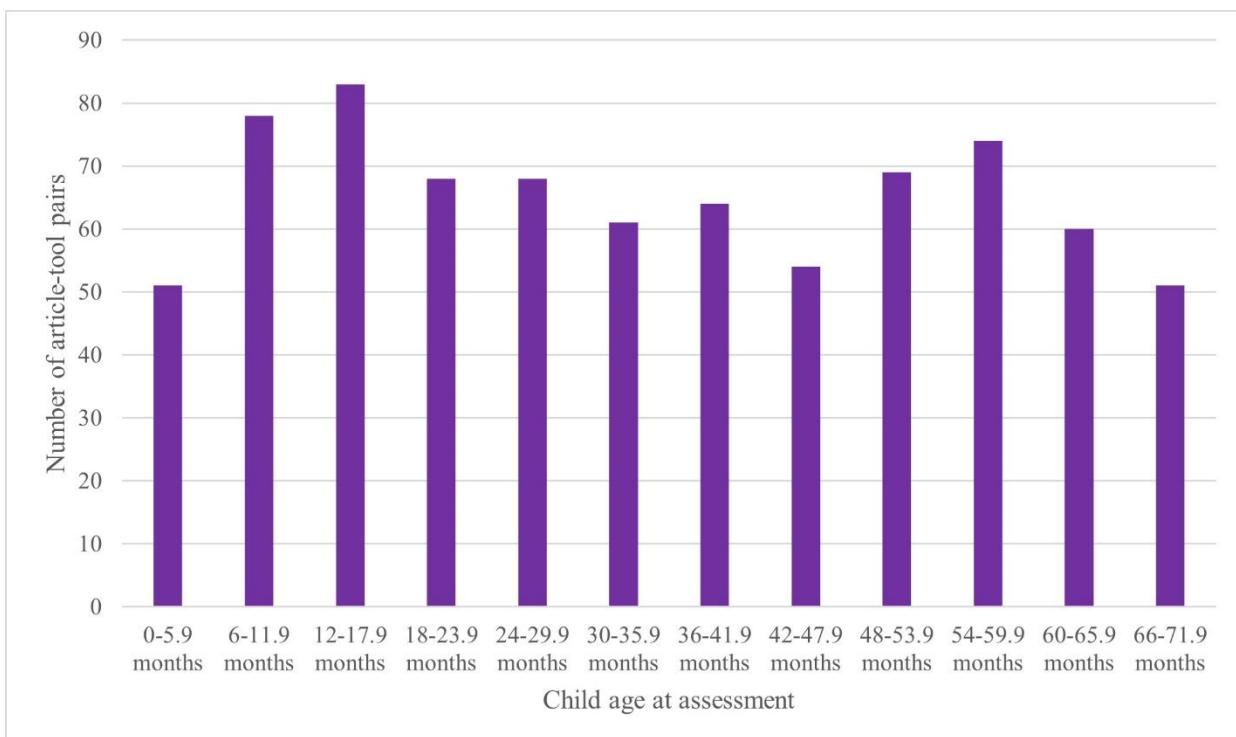
Supplemental Figure 1 Number of articles providing evidence on psychometrics by type of article



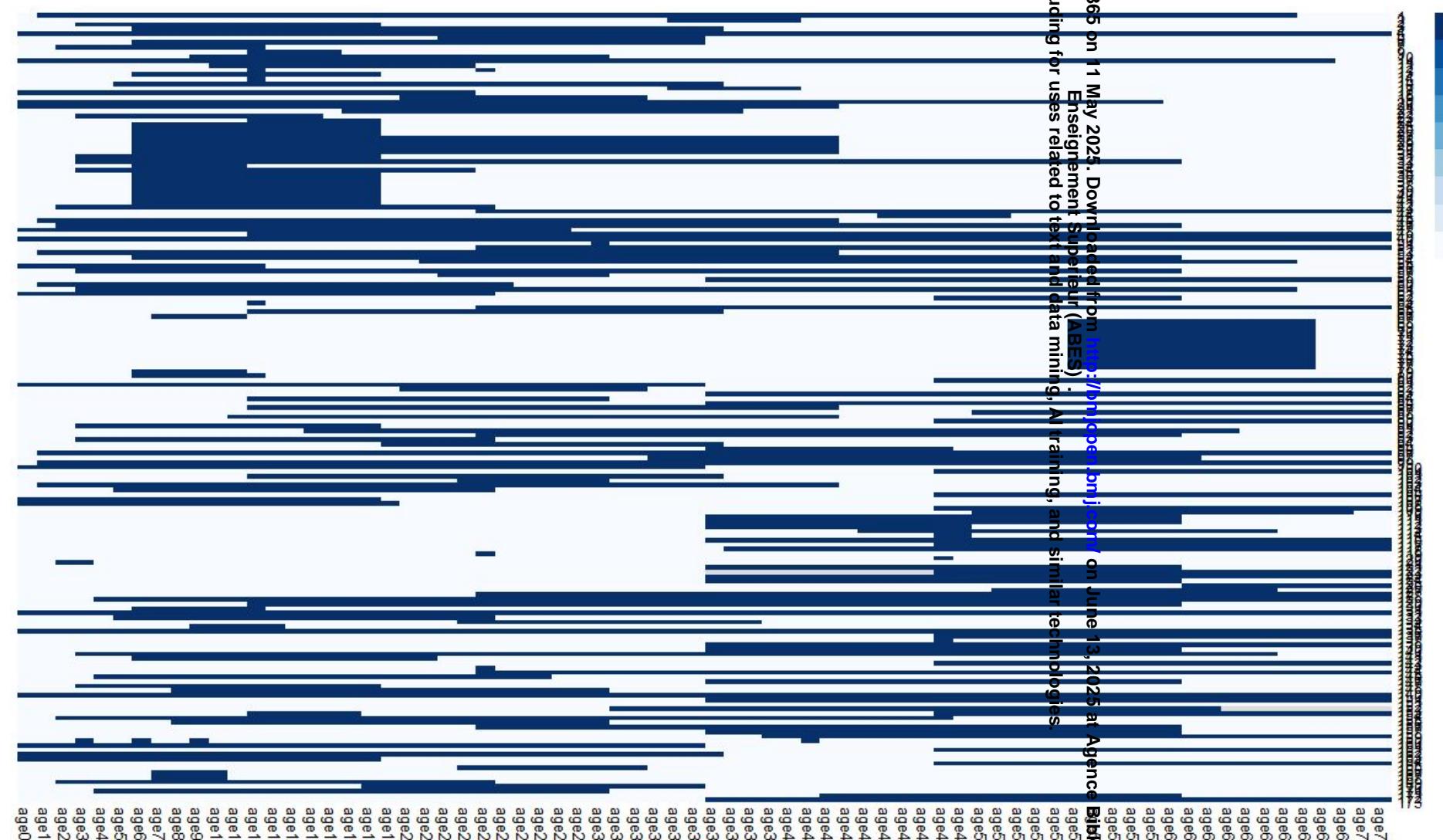
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). Excluded by copyright. Excluding for uses related to text and data mining, AI training, and similar technologies.

Supplemental Figure 2 Number of articles and tools providing psychometric evidence by child age group



Supplemental Figure 3 Number of articles providing evidence on psychometrics by child age at assessment



Note: Y-axis lists each article-tool combination included in the systematic review ordered by tool in alphabetical order.

References of all included articles

- Abdoola, S., Swanepoel, D. W., Van Der Linde, J., & Glascoe, F. P. (2021). Detecting developmental delays in infants from a low-income South African community: comparing the BSID-III and PEDS tools. *Early Child Development and Care*, 191(4), 545–554. <https://doi.org/10.1080/03004430.2019.1628027>
- Abessa, T. G., Worku, B. N., Kibebew, M. W., Valy, J., Lemmens, J., Thijs, H., Yimer, W. K., Kolsteren, P., & Granitzer, M. (2016). Adaptation and standardization of a Western tool for assessing child development in non-Western low-income context. *BMC Public Health*, 16(1), 1–13. <https://doi.org/10.1186/s12889-016-3288-2>
- Abubakar, A., Holding, P., Van De Vijver, F., Bomu, G., & Van Baar, A. (2010). Developmental monitoring using caregiver reports in a resource-limited setting: The case of Kilifi, Kenya. *Acta Paediatrica, International Journal of Paediatrics*, 99(2), 291–297. <https://doi.org/10.1111/j.1651-2227.2009.01561.x>
- Abubakar, A., Holding, P., van Baar, A., Newton, C. R. J. C., & van de Vijver, F. J. R. (2008). Monitoring psychomotor development in a resource-limited setting: An evaluation of the Kilifi Developmental Inventory. *Annals of Tropical Paediatrics*, 28(3), 217–226. <https://doi.org/10.1179/146532808X335679>
- Albuquerque, P. L. de, Guerra, M. Q. de F., Lima, M. de C., & Eickmann, S. H. (2018). Concurrent validity of the Alberta Infant Motor Scale to detect delayed gross motor development in preterm infants: A comparative study with the Bayley III. *Developmental Neurorehabilitation*, 21(6), 408–414. <https://doi.org/10.1080/17518423.2017.1323974>
- Alcock, K. J., Rimba, K., Holding, P., Kitsao-Wekulo, P., Abubakar, A., & Newton, C. R. J. C. (2015). Developmental inventories using illiterate parents as informants: Communicative Development Inventory (CDI) adaptation for two Kenyan languages. *Journal of Child Language*, 42(4), 763–785. <https://doi.org/10.1017/S0305000914000403>
- Alderman, H., Friedman, J., Ganga, P., Kak, M., & Rubio-Codina, M. (2021). Assessing the performance of the Caregiver Reported Early Development Instruments (CREDI) in rural India. *Annals of the New York Academy of Sciences*, 1492(1), 58–72. <https://doi.org/10.1111/nyas.14543>
- Altafim, E. R. P., McCoy, D. C., Brentani, A., Escobar, A. M. de U., Grisi, S. J. F. E., & Fink, G. (2020). Measuring early childhood development in Brazil: validation of the Caregiver Reported Early Development Instruments (CREDI). *Jornal de Pediatria*, 96(1), 66–75. <https://doi.org/10.1016/j.jped.2018.07.008>
- Anunciação, L., Squires, J., Clifford, J., & Landeira-Fernandez, J. (2019). Confirmatory analysis and normative tables for the Brazilian Ages and Stages Questionnaires: Social–Emotional. *Child: Care, Health and Development*, 45(3), 387–393. <https://doi.org/10.1111/cch.12649>

- Azari, N., Soleimani, F., Vameghi, R., Sajedi, F., Shahshahani, S., Karimi, H., Kraskian, A., Shahrokhi, A., Teymour, R., & Gharib, M. (2017). A psychometric study of the Bayley scales of infant and toddler development in Persian language children. *Iranian Journal of Child Neurology*, 11(1), 50–56.
- Bandeira, P. F. R., Duncan, M., Pessoa, M. L., Soares, I., Da Silva, L., Mota, J., & Martins, C. (2020). Tgmd-2 short version: Evidence of validity and associations with sex, age, and bmi in preschool children. *Journal of Motor Learning and Development*, 8(3), 528–543. <https://doi.org/10.1123/JMLD.2019-0040>
- Barthel, D., Kriston, L., Guo, N., Yao, E. D., Armel, K. E. J., Fordjour, D., Mohammed, Y., Kotchi, C. E. B., Eberhardt, K. A., Hinz, R., Ehrhardt, S., & Bindt, C. (2022). Psychometric properties of the Brief Infant-Toddler Social and Emotional Assessment (BITSEA) in two-year-old Ivorian and Ghanaian children. *Applied Developmental Science*, 26(3), 567–577. <https://doi.org/10.1080/10888691.2021.1931216>
- Barthel, D., Kriston, L., Yao, E. D., Fordjour, D., Armel, K. E. J., Eberhardt, K. A., Hinz, R., Ehrhardt, S., & Bindt, C. (2022). Psychomotor development in two-year-old Ivorian and Ghanaian children – Psychometric properties of the Kilifi Developmental Inventory. *British Journal of Developmental Psychology*, 40(4), 471–486. <https://doi.org/10.1111/bjdp.12428>
- Bhavnani, S., Mukherjee, D., Dasgupta, J., Verma, D., Parameshwaran, D., Divan, G., Sharma, K. K., Thiagarajan, T., & Patel, V. (2019). Development, feasibility and acceptability of a gamified cognitive DEvelopmental assessment on an E-Platform (DEEP) in rural Indian pre-schoolers—a pilot study. *Global Health Action*, 12(1). <https://doi.org/10.1080/16549716.2018.1548005>
- Bian, X., Xie, H., Squires, J., & Chen, C. Y. (2017). Adapting a Parent-Completed, Socioemotional Questionnaire in China: the Ages & Stages Questionnaires: Social-Emotional. *Infant Mental Health Journal*, 38(2), 258–266. <https://doi.org/10.1002/imhj.21626>
- Bian, X., Yao, G., Squires, J., Hoselton, R., Chen, C. I., Murphy, K., Wei, M., & Fang, B. (2012). Translation and use of parent-completed developmental screening test in Shanghai. *Journal of Early Childhood Research*, 10(2), 162–175. <https://doi.org/10.1177/1476718X11430071>
- Biasini, F. J., De Jong, D., Ryan, S., Thorsten, V., Bann, C., Bellad, R., Mahantshetti, N. S., Dhaded, S. M., Pasha, O., Chomba, E., Goudar, S. S., Carlo, W. A., & McClure, E. (2015). Development of a 12 month screener based on items from the Bayley II Scales of Infant Development for use in Low Middle Income countries. *Early Human Development*, 91(4), 253–258. <https://doi.org/10.1016/j.earlhumdev.2015.02.001>
- Boivin, M. J., Weiss, J., Chhaya, R., Seffren, V., Awadu, J., Sikorskii, A., & Giordani, B. (2017). The feasibility of automated eye tracking with the early childhood vigilance test of attention

- 1
2
3 in younger HIV-exposed Ugandan children. *Neuropsychology*, 31(5), 525–534.
4 https://doi.org/10.1037/neu0000382
5
6 Boivin, M. J., Zoumenou, R., Sikorskii, A., Fievet, N., Alao, J., Davidson, L., Cot, M.,
7 Massougbedji, A., & Bodeau-Livinec, F. (2021). Neurodevelopmental assessment at one
8 year of age predicts neuropsychological performance at six years in a cohort of West
9 African Children. *Child Neuropsychology*, 27(4), 548–571.
10 https://doi.org/10.1080/09297049.2021.1876012
11
12 Bornman, J., Sevcik, R. A., Romski, M. A., & Pae, H. K. (2010). Successfully translating
13 language and culture when adapting assessment measures. *Journal of Policy and Practice in*
14 *Intellectual Disabilities*, 7(2), 111–118. https://doi.org/10.1111/j.1741-1130.2010.00254.x
15
16 Brinkman, S. A., Kinnell, A., Maika, A., Hasan, A., Jung, H., & Pradhan, M. (2017). Validity
17 and Reliability of the Early Development Instrument in Indonesia. *Child Indicators*
18 *Research*, 10(2), 331–352. https://doi.org/10.1007/s12187-016-9372-4
19
20 Caldera, A. ., Wickremasinghe, A. R., Muttiah, N., Godamunne, P. K. S., Jayasena, B. .,
21 Chathurika, L. K. E., Perera, K. M. N., Mendis, M., Tilakarathne, D., Peiris, M. K. R. .,
22 Wijesinghe, T., Senarathna, N. ., Saubhagya, W. D. L., Chandraratne, M., & Sumanasena,
23 S. . (2023). REACH for the preschoolers; a developmental assessment tool for 2–5 year old
24 children in Sri Lanka. *BMC Pediatrics*, 23(1), 80. https://doi.org/10.1186/s12887-023-
25 03895-5
26
27 Chen, C. Y., Xie, H., Filgueiras, A., Squires, J., Anunciação, L., & Landeira-Fernandez, J.
28 (2017). Examining the Psychometric Properties of the Brazilian Ages & Stages
29 Questionnaires-Social-Emotional: Use in Public Child Daycare Centers in Brazil. *Journal of*
30 *Child and Family Studies*, 26(9), 2412–2425. https://doi.org/10.1007/s10826-017-0770-0
31
32 Chunsuwan, I., Hansakunachai, T., & Pornsamrit, S. (2016). Parent Evaluation of Developmental
33 Status (PEDS) in screening: The Thai experience. *Pediatrics International*, 58(12), 1277–
34 1283. https://doi.org/10.1111/ped.13055
35
36 Colbert, A. M., Connery, A. K., Lamb, M. M., Bauer, D., Olson, D., Paniagua-Avila, A.,
37 Martínez, M. A., Arroyave, P., Hernández, S., Mirella Calvimontes, D., Bolaños, G. A., El
38 Sahly, H. M., Muñoz, F. M., & Asturias, E. J. (2021). Caregiver rating of early childhood
39 development: Reliability and validity of the ASQ-3 in rural Guatemala. *Early Human*
40 *Development*, 161(November 2020). https://doi.org/10.1016/j.earlhumdev.2021.105453
41
42 Colbert, A. M., Lamb, M. M., Asturias, E. J., Muñoz, F. M., Bauer, D., Arroyave, P., Hernández,
43 S., Martínez, M. A., Paniagua-Avila, A., Olson, D., Calvimontes, D. M., Bolaños, G. A., El
44 Sahly, H. M., & Connery, A. K. (2020). Reliability and Validity of an Adapted and
45 Translated Version of the Mullen Scales of Early Learning (AT-MSEL) in Rural
46 Guatemala. *Child: Care, Health and Development*, 46(3), 327–335.
47 https://doi.org/10.1111/cch.12748
48
49
50
51
52
53
54
55
56
57
58
59
60

- Dagvadorj, A., Takehara, K., Bavuusuren, B., Morisaki, N., Gochoo, S., & Mori, R. (2015). The quick and easy Mongolian Rapid Baby Scale shows good concurrent validity and sensitivity. *Acta Paediatrica, International Journal of Paediatrics*, 104(3), e94–e99. <https://doi.org/10.1111/apa.12884>
- Danişman, Ş., Dereli Iman, E., Demircan, Z. A., & Yaya, D. (2016). Una aproximación al enfoque Reggio Emilia: Claves para entender por qué motiva a los estudiantes. *Electronic Journal of Research in Educational Psychology*, 14(3), 534–556.
- de Albuquerque, K. A., da Cunha, A. C. B., & Berkovits, M. D. (2022). Cross-cultural adaptation of the Battelle Development Inventory, 2nd Edition for Brazil. *Estudos de Psicologia*, 39, 1–13. <https://doi.org/10.1590/1982-0275202239e190170>
- Dias, T. P., Freitas, L. C., Del Prette, Z. A. P., & Del Prette, A. (2011). Validação da escala de comportamentos sociais de pré-escolares para o Brasil. *Psicologia Em Estudo*, 16(3), 447–457. <https://doi.org/10.1590/S1413-73722011000300012>
- Dursun, O. B., Öğütlü, H., & Esin, İ. S. (2020). Psychometric properties of Turkish version of the strength and difficulties questionnaire for age 2-4. *Noropsikiyatri Arsivi*, 57(1), 44–49. <https://doi.org/10.29399/npa.23674>
- Familiar-Lopez, I., Sikorskii, A., Chhaya, R., Holmes, A., Arima, E. G., Caesar, O. J., Nakasujja, N., & Boivin, M. J. (2023). Predictive validation of Ugandan infant eye-tracking test for memory of human faces. *Child Neuropsychology*, 29(3), 486–502. <https://doi.org/10.1080/09297049.2022.2099537>
- Fernandes, M., Bassani, D., Albernaz, E., Bertoldi, A. D., Silveira, M. F., Matijsevich, A., Anselmi, L., Cruz, S., Halal, C. S., Tovo-Rodrigues, L., Cruz, G. I. N., Metgud, D., & Santos, I. S. (2022). Construction and validation of the Oxford Neurodevelopment Assessment (OX-NDA) in 1-year-old Brazilian children. *BMC Pediatrics*, 22(1), 1–12. <https://doi.org/10.1186/s12887-022-03794-1>
- Ferreira-Vasques, A. T., Santos, C. F., & Lamônica, D. A. C. (2019). Transcultural adaptation process of the Griffiths-III Mental Development Scale. *Child: Care, Health and Development*, 45(3), 403–408. <https://doi.org/10.1111/cch.12664>
- Filgueiras, A., Pires, P., Maissonette, S., & Landeira-Fernandez, J. (2013). Psychometric properties of the Brazilian-adapted version of the Ages and Stages Questionnaire in public child daycare centers. *Early Human Development*, 89(8), 561–576. <https://doi.org/10.1016/j.earlhhumdev.2013.02.005>
- Finch, J. E., Yousafzai, A. K., Rasheed, M., & Obradović, J. (2018). Measuring and understanding social-emotional behaviors in preschoolers from rural Pakistan. *PLoS ONE*, 13(11), 1–22. <https://doi.org/10.1371/journal.pone.0207807>

- 1
2
3 Flórez Romero, R., Castro Martínez, J., & Arias Velandia, N. (2013). Propiedades psicométricas
4 de la escala de lenguaje para preescolares (PLS-3) colombianos TT - Psychometric
5 Properties of the Pre-school Language Scale (PLS-3) with Colombian Children TT -
6 Propriedades psicométricas da Escala de linguagem para pré-escola. *Pensamiento*
7 *Psicológico*, 11(1), 131–140.
8
9
- 10 Gladstone, M., Lancaster, G. a., Umar, E., Nyirenda, M., Kayira, E., van den Broek, N. R., &
11 Smyth, R. L. (2010). The Malawi Developmental Assessment Tool (MDAT): The Creation,
12 Validation, and Reliability of a Tool to Assess Child Development in Rural African
13 Settings. *PLoS Medicine*, 7(5), e1000273. <https://doi.org/10.1371/journal.pmed.1000273>
- 14
15 Golmohammadi, G., Zarifian, T., Qhanbari, S., Bakhshi, E., & Sakhai, F. (2022). Cross-cultural
16 adaptation and validation of the Persian version of Children's Behavior Questionnaire in
17 Iranian children. *Current Psychology*, 41(6), 4041–4053. <https://doi.org/10.1007/s12144-020-00918-7>
- 18
19
20
21 Gomez, J. A., Brown, J. L., & Spier, E. (2022). Factor structure, measurement invariance, and
22 construct validity of MELQO measure of child development and early learning in Laos.
23 *International Journal of School and Educational Psychology*, 10(2), 181–207.
24 <https://doi.org/10.1080/21683603.2020.1791770>
- 25
26
27 Guedes, D. Z., Primi, R., & Kopelman, B. I. (2011). BINS validation - Bayley
28 neurodevelopmental screener in Brazilian preterm children under risk conditions. *Infant*
29 *Behavior and Development*, 34(1), 126–135. <https://doi.org/10.1016/j.infbeh.2010.11.001>
- 30
31
32 Guo, B., Aveyard, P., & Dai, X. (2009). The Chinese Intelligence Scale for Young Children.
33 *Educational and Psychological Measurement*, 69(3), 459–474.
34 <https://doi.org/10.1177/0013164409332209>
- 35
36
37 Hahn, K. R., Schaefer, B. A., Merino, C., & Worrell, F. C. (2009). The factor structure of
38 preschool learning behaviors scale scores in peruvian children. *Canadian Journal of School*
39 *Psychology*, 24(4), 318–331. <https://doi.org/10.1177/0829573509349015>
- 40
41 Halpin, P. F., Wolf, S., Yoshikawa, H., Rojas, N., Kabay, S., Pisani, L., & Dowd, A. J. (2019).
42 Measuring early learning and development across cultures: Invariance of the IDELA across
43 five countries. *Developmental Psychology*, 55(1), 23–37.
44 <https://doi.org/10.1037/dev0000626>
- 45
46
47 Hamadani, J. D., Baker-Henningham, H., Tofail, F., Mehrin, F., Huda, S. N., & Grantham-
48 McGregor, S. M. (2010). Validity and Reliability of Mothers' Reports of Language
49 Development in 1-Year-Old Children in a Large-Scale Survey in Bangladesh. *Food and*
50 *Nutrition Bulletin*, 31(2_suppl2), S198–S206.
51 <https://doi.org/10.1177/15648265100312S212>
- 52
53
54 Hamadani, J. D., Tofail, F., Cole, T., & Grantham-McGregor, S. (2013). The relation between
55 age of attainment of motor milestones and future cognitive and motor development in
- 56
57
58
59
60

- 1
2
3 Bangladeshi children. *Maternal and Child Nutrition*, 9(SUPPL. 1), 89–104.
4 https://doi.org/10.1111/mcn.12020
5
6 Hanlon, C., Medhin, G., Worku, B., Tomlinson, M., Alem, A., Dewey, M., & Prince, M. (2016).
7 Adapting the Bayley Scales of infant and toddler development in Ethiopia: evaluation of
8 reliability and validity. *Child: Care, Health and Development*, 42(5), 699–708.
9 https://doi.org/10.1111/cch.12371
10
11
12 Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese.
13 *Behavior Research Methods*, 40(3), 728–733. https://doi.org/10.3758/BRM.40.3.728
14
15 Hilburn, N., Potterton, J., Stewart, A., & BeckerC, P. (2011). The development of a screening
16 tool to evaluate gross motor function in HIV-infected infants. *AIDS Care - Psychological
17 and Socio-Medical Aspects of AIDS/HIV*, 23(12), 1619–1625.
18 https://doi.org/10.1080/09540121.2011.579942
19
20
21 Hsiao, C., Richter, L., Makusha, T., Matafwali, B., van Heerden, A., & Mabaso, M. (2017). Use
22 of the ages and stages questionnaire adapted for South Africa and Zambia. *Child: Care,
23 Health and Development*, 43(1), 59–66. https://doi.org/10.1111/cch.12413
24
25
26 Hua, J., Gu, G., Meng, W., & Wu, Z. (2013). Age band 1 of the Movement Assessment Battery
27 for Children-Second Edition: Exploring its usefulness in mainland China. *Research in
28 Developmental Disabilities*, 34(2), 801–808. https://doi.org/10.1016/j.ridd.2012.10.012
29
30 Jackson-Maldonado, D., Marchman, V. A., & Fernald, L. C. H. (2013). Short-form versions of
31 the Spanish MacArthur-Bates Communicative Development Inventories. *Applied
32 Psycholinguistics*, 34(4), 837–868. https://doi.org/10.1017/S0142716412000045
33
34 Jianduan, Z., Huishan, W., Shuhua, S., Xiaonan, H., Guoyan, L., Guangli, L., & Junxin, S.
35 (2009). Reliability and validity of standardized Chinese version of Urban Infant-Toddler
36 Social and Emotional Assessment. *Early Human Development*, 85(5), 331–336.
37 https://doi.org/10.1016/j.earlhundev.2008.12.012
38
39
40 Karabekiroglu, K., Rodopman-Arman, A., Ay, P., Ozkesen, M., Akbas, S., Tasdemir, G. N.,
41 Boke, O., & Peksen, Y. (2009). The reliability and validity of the Turkish version of the
42 brief infant-toddler social emotional assessment (BITSEA). *Infant Behavior and
43 Development*, 32(3), 291–297. https://doi.org/10.1016/j.infbeh.2009.03.003
44
45
46 Kariuki, S. M., Abubakar, A., Murray, E., Stein, A., & Newton, C. R. J. C. (2016). Evaluation of
47 psychometric properties and factorial structure of the pre-school child behaviour checklist at
48 the Kenyan Coast. *Child and Adolescent Psychiatry and Mental Health*, 10(1), 1–10.
49 https://doi.org/10.1186/s13034-015-0089-9
50
51
52 Khan, N. Z., Muslima, H., Shilpi, A. B., Begum, D., Parveen, M., Akter, N., Ferdous, S., Nahar,
53 K., McConachie, H., & Darmstadt, G. L. (2013). Validation of rapid neurodevelopmental
54
55
56
57
58
59
60

1
2
3 assessment for 2- to 5-year-old children in bangladesh. *Pediatrics*, 131(2).
4 https://doi.org/10.1542/peds.2011-2421
5

6 Khan, N. Z., Muslina, H., Begum, D., Shilpi, A. B., Akhter, S., Bilkis, K., Begum, N., Parveen,
7 M., Ferdous, S., Morshed, R., Batra, M., & Darmstadt, G. L. (2010). Validation of rapid
8 neurodevelopmental assessment instrument for under-two-year-old children in Bangladesh.
9 *Pediatrics*, 125(4). https://doi.org/10.1542/peds.2008-3471
10
11

12 Knauer, H. A., Kariger, P., Jakiela, P., Ozier, O., & Fernald, L. C. H. (2019). Multilingual
13 assessment of early child development: Analyses from repeated observations of children in
14 Kenya. *Developmental Science*, 22(5), 1–23. https://doi.org/10.1111/desc.12875
15
16

17 Koura, K. G., Boivin, M. J., Davidson, L. L., Ouédraogo, S., Zoumenou, R., Alao, M. J., Garcia,
18 A., Massougbedji, A., Cot, M., & Bodeau-Livinec, F. (2013). Usefulness of child
19 development assessments for low-resource settings in Francophone Africa. *Journal of
20 Developmental and Behavioral Pediatrics*, 34(7), 486–493.
21
22 https://doi.org/10.1097/DBP.0b013e31829d211c
23

24 Kucuker, S., Kapci, E. G., & Uslu, R. I. (2011). Evaluation of the turkish version of the “ages
25 and stages questionnaires: Social-emotional” in identifying children with social-emotional
26 problems. *Infants and Young Children*, 24(2), 207–220.
27
28 https://doi.org/10.1097/IYC.0b013e31820eae26
29

30 Kumar, R., Shankar, K., Kush, V., Kumar, C., Bhave, A., & Agarwal, V. (2016). Adaptation:
31 Vineland Adaptive Behavior Scale for 3-9 year-old Indian children. *International Journal
32 on Disability and Human Development*, 15(1), 49–55. https://doi.org/10.1515/ijdhd-2014-
33 0026
34

35 Kvestad, I., Silpakar, J. S., Hysing, M., Ranjitkar, S., Strand, T. A., Swinger, C., Shrestha, M.,
36 Chandyo, R. K., & Ulak, M. (2023). The reliability and predictive ability of the Test of
37 Infant Motor Performance (TIMP) in a community-based study in Bhaktapur, Nepal. *Infant
38 Behavior and Development*, 70(January), 101809.
39
40 https://doi.org/10.1016/j.infbeh.2023.101809
41

42 Kvestad, I., Taneja, S., Kumar, T., Bhandari, N., Strand, T. A., & Hysing, M. (2013). The
43 assessment of developmental status using the Ages and Stages questionnaire-3 in nutritional
44 research in north Indian young children. *Nutrition Journal*, 12(1), 1–11.
45
46 https://doi.org/10.1186/1475-2891-12-50
47

48 Lacunza, A. B., Castro Solano, A., & Contini, N. (2009). Habilidades sociales preescolares: una
49 escala para niños de contextos de pobreza. *Revista de Psicología*, 27(1), 3–28.
50
51 https://doi.org/10.18800/psico.200901.001
52

53 Lara Díaz, M. F. L., Fonseca, Á. M. G., Bohórquez, D. M. G., Guechá, C. M., & Sellabona, E. S.
54 (2011). Normativización del inventario del desarrollo comunicativo MacArthur-bates al
55 Español, Colombia. *Revista Latinoamericana de Psicología*, 43(2), 241–254.
56
57

- Leppänen, J. M., Butcher, J. W., Godbout, C., Stephenson, K., Taylor Hendrixson, D., Griswold, S., Rogers, B. L., Webb, P., Koroma, A. S., & Manary, M. J. (2022). Assessing infant cognition in field settings using eye-tracking: A pilot cohort trial in Sierra Leone. *BMJ Open*, 12(2). <https://doi.org/10.1136/bmjopen-2021-049783>
- Li, Y., Zhu, J. J., Coplan, R. J., Gao, Z. Q., Xu, P., Li, L., & Zhang, H. (2016). Assessment and Implications of Social Withdrawal Subtypes in Young Chinese Children: The Chinese Version of the Child Social Preference Scale. *Journal of Genetic Psychology*, 177(3), 97–101. <https://doi.org/10.1080/00221325.2016.1174100>
- Li, Y., Tang, L., Bai, Y., Zhao, S., & Shi, Y. (2020). Reliability and validity of the Caregiver Reported Early Development Instruments (CREDI) in impoverished regions of China. *BMC Pediatrics*, 20(1). <https://doi.org/10.1186/s12887-020-02367-4>
- Liu, J., Cheng, H., & Leung, P. W. L. (2011). The application of the preschool child behavior checklist and the Caregiver-Teacher Report Form to Mainland Chinese children: Syndrome structure, gender differences, country effects, and inter-informant agreement. *Journal of Abnormal Child Psychology*, 39(2), 251–264. <https://doi.org/10.1007/s10802-010-9452-8>
- Lohaus, A., Lamm, B., Keller, H., Teubert, M., Fassbender, I., Glüer, M., Borchert, S., Vöhringer, I., Teiser, J., Freitag, C., Suhrke, J., Knopf, M., & Schwarzer, G. (2014). Gross and Fine Motor Differences Between Cameroonian and German Children Aged 3 to 40 Months: Results of a Cross-Cultural Longitudinal Study. *Journal of Cross-Cultural Psychology*, 45(8), 1328–1341. <https://doi.org/10.1177/0022022114537703>
- Lopez Boo, F., Cubides Mateus, M., & Llonch Sabatés, A. (2020). Initial psychometric properties of the Denver II in a sample from Northeast Brazil. *Infant Behavior and Development*, 58(August 2019), 101391. <https://doi.org/10.1016/j.infbeh.2019.101391>
- Madraschi, V., Mecca, T. P., Macedo, E. C., & Paula, C. S. (2016). Bayley-III scales of infant and toddler development: Transcultural adaptation and psychometric properties. *Paideia*, 26(64), 189–197. <https://doi.org/10.1590/1982-43272664201606>
- Manandhar, S. R., Dulal, S., Manandhar, D. S., Saville, N., & Prost, A. (2016). Acceptability and Reliability of the Bayley Scales of Infant Development III Cognitive and Motor Scales among Children in Makwanpur. *Journal of Nepal Health Research Council*, 14(32), 47–50.
- McCarthy, A. M., Wehby, G. L., Barron, S., Aylward, G. P., Castilla, E. E., Javois, L. C., Goco, N., & Murray, J. C. (2012). Application of neurodevelopmental screening to a sample of South American infants: The Bayley Infant Neurodevelopmental Screener (BINS). *Infant Behavior and Development*, 35(2), 280–294. <https://doi.org/10.1016/j.infbeh.2011.12.003>
- McCoy, D. C., Sudfeld, C. R., Bellinger, D. C., Muhihi, A., Ashery, G., Weary, T. E., Fawzi, W., & Fink, G. (2017). Development and validation of an early childhood development scale for use in low-resourced settings. *Population Health Metrics*, 15(1), 1–18. <https://doi.org/10.1186/s12963-017-0122-8>

- McCoy, D. C., Waldman, M., & Fink, G. (2018). Measuring early childhood development at a global scale: Evidence from the Caregiver-Reported Early Development Instruments. *Early Childhood Research Quarterly*, 45, 58–68. <https://doi.org/10.1016/j.ecresq.2018.05.002>
- McCrory, G., McCoy, D., Kariger, P., Janus, M., Black, M. M., Chang, S. M., Tofail, F., Eekhout, I., Waldman, M., Van Buuren, S., Khanam, R., Sazawal, S., Nizar, A., Schönbeck, Y., Zongo, A., Brentani, A., Zhang, Y., Dua, T., Cavallera, V., ... Gladstone, M. (2023). The creation of the Global Scales for Early Development (GSED) for children aged 0-3 years: Combining subject matter expert judgements with big data. *BMJ Global Health*, 8(1), 1–11. <https://doi.org/10.1136/bmjjh-2022-009827>
- McHenry, M. S., Oyungu, E., Yang, Z., Hines, A. C., Ombitsa, A. R., Vreeman, R. C., Abubakar, A., & Monahan, P. O. (2021). Cultural adaptation of the Bayley Scales of Infant and Toddler Development, 3rd Edition for use in Kenyan children aged 18–36 months: A psychometric study. *Research in Developmental Disabilities*, 110(December 2020), 103837. <https://doi.org/10.1016/j.ridd.2020.103837>
- Mehrani, M. B. (2018). An elicited imitation test for measuring preschoolers' language development. *Psychological Reports*, 121(4), 767–786. <https://doi.org/10.1177/0033294117732976>
- Meyer, E. L., Schaefer, B. A., Soto, C. M., Simmons, C. S., Anguiano, R., Brett, J., Holman, A., Martin, J. F., Hata, H. K., Roberts, K. J., Mello, Z. R., & Worrell, F. C. (2011). Factor structure of child behavior scale scores in peruvian preschoolers. *Psychology in the Schools*, 48(10), 931–942. <https://doi.org/10.1002/pits.20596>
- Milosavljevic, B., Vellekoop, P., Maris, H., Halliday, D., Drammeh, S., Sanyang, L., Darboe, M. K., Elwell, C., Moore, S. E., & Lloyd-Fox, S. (2019). Adaptation of the Mullen Scales of Early Learning for use among infants aged 5- to 24-months in rural Gambia. *Developmental Science*, 22(5), 1–17. <https://doi.org/10.1111/desc.12808>
- Mukherjee, D., Bhavnani, S., Swaminathan, A., Verma, D., Parameshwaran, D., Divan, G., Dasgupta, J., Sharma, K., Thiagarajan, T. C., & Patel, V. (2020). Proof of Concept of a Gamified DEvelopmental Assessment on an E-Platform (DEEP) Tool to Measure Cognitive Development in Rural Indian Preschool Children. *Frontiers in Psychology*, 11(June), 1–12. <https://doi.org/10.3389/fpsyg.2020.01202>
- Mukherjee, S. B., Verma, S., Sharma, S., & Aneja, S. (2022). Diagnostic Accuracy of Parents' Evaluation of Developmental Status (PEDS), PEDS Developmental Milestones, and PEDS Combined in Indian Children Aged Less than 2 Years. *Indian Journal of Pediatrics*, 89(5), 459–465. <https://doi.org/10.1007/s12098-020-03651-y>
- Nair, M. K. C., Russell, P. S., Rekha, R. S., Lakshmi, M. A., Latha, S., & Rajee, K. (2009). Validation of developmental assessment tool for Anganwadis (DATA). *Indian Pediatrics*, 46 Suppl(2), s27-36.

- Nair, M. K. C., & Russell, P. S. (2013). Development and normative validation of Developmental Assessment Tool for Anganwadis for 3- to 4-year-old children (DATA-II). *Journal of Clinical Epidemiology*, 66(1), 23–29.
<https://doi.org/10.1016/j.jclinepi.2012.06.011>
- 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>
- Ngoun, C., Stoey, L. S., van't Ende, K., & Kumar, V. (2012). Creating a Cambodia-specific developmental milestone screening tool - A pilot study. *Early Human Development*, 88(6), 379–385. <https://doi.org/10.1016/j.earlhumdev.2011.09.014>
- Obradović, J., Finch, J. E., Portilla, X. A., Rasheed, M. A., Tirado-Strayer, N., & Yousafzai, A. K. (2019). Early executive functioning in a global context: Developmental continuity and family protective factors. *Developmental Science*, 22(5). <https://doi.org/10.1111/desc.12795>
- Okuda, P. M. M., Félix, E., Cogo-Moreira, H., Liu, T., Surkan, P. J., Martins, S. S., & Caetano, S. C. (2020). Construct validity of the motor development scale (MDS). *Trends in Psychiatry and Psychotherapy*, 42(1), 82–85. <https://doi.org/10.1590/2237-6089-2018-0114>
- Ortiz-León, S., Granados-Rojas, A., Cavazos-Olivio, J., Benito-Avendaño, L. De, Obregón-García, V. H., Duran-Avendaño, X. M., Ramírez-Pérez, R. V., Cárdenas-Medina, J. A., Vargas-Terréz, B. E., Luna-Guevara, G., Rodríguez-Machain, A. C., & Jaimes-Medrano, A. L. (2018). Internal and inter-rater reliability of the ASQ-3 in Mexican preschoolers. *Salud Mental*, 41(2), 65–72. <https://doi.org/10.17711/SM.0185-3325.2018.011>
- Oryadi-Zanjani, M. M. (2020). Development of the Childhood Nonverbal Communication Scale. *Journal of Autism and Developmental Disorders*, 50(4), 1238–1248.
<https://doi.org/10.1007/s10803-019-04356-8>
- Parveen, M., Rahman, S. T., Islam, S., Zaman, S. S., Hamadani, J., & Khan, N. Z. (2014). Adaptation of items of bayley scales of infant development-II (BSID-II) suitable for Bangladeshi infants. *Dhaka University Journal of Biological Sciences*, 23(2), 187–195.
<https://doi.org/10.3329/dujbs.v23i2.20099>
- Pathak, S., & Sovani-Kelkar, P. (2023). Development of a screening tool to identify babies at risk of language delay in India: A preliminary study. *Language Acquisition*, 30(1), 29–49.
<https://doi.org/10.1080/10489223.2022.2078212>
- Pendergast, L. L., Schaefer, B. A., Murray-Kolb, L. E., Svensen, E., Shrestha, R., Rasheed, M. A., Scharf, R. J., Kosek, M. N., Vasquez, A. O., Maphula, A., Costa, H., Rasmussen, Z. A., Yousafzai, A., Tofail, F., Seidman, J. C., Acosta, A. M., de Burga, R. R., Chavez, C. B., Flores, J. T., ... Strand, T. (2018). Assessing development across cultures: Invariance of the

- 1
2
3 Bayley-III scales across seven international MAL-ED sites. *School Psychology Quarterly*,
4 33(4), 604–614. <https://doi.org/10.1037/spq0000264>
- 5
6
7 Phạm, B., McLeod, S., & Harrison, L. J. (2017). Validation and norming of the Intelligibility in
8 Context Scale in Northern Viet Nam. *Clinical Linguistics and Phonetics*, 31(7–9), 665–681.
9 <https://doi.org/10.1080/02699206.2017.1306110>
- 10
11 Pisani, L., Borisova, I., & Dowd, A. J. (2018). Developing and validating the International
12 Development and Early Learning Assessment (IDELA). *International Journal of
13 Educational Research*, 91(July), 1–15. <https://doi.org/10.1016/j.ijer.2018.06.007>
- 14
15 Pisani, L., Seiden, J., & Wolf, S. (2022). Longitudinal evidence on the predictive validity of the
16 International Development and Early Learning Assessment (IDELA). *Educational
17 Assessment, Evaluation and Accountability*, 34(2), 173–194.
18 <https://doi.org/10.1007/s11092-021-09374-8>
- 19
20 Pitchik, H. O., Tofail, F., Akter, F., Shoab, A. K. M., Sultana, J., Huda, T. M. N., Rahman, M.,
21 Winch, P. J., Luby, S. P., & Fernald, L. C. H. (2023). Concurrent validity of the Ages and
22 Stages Questionnaire Inventory and the Bayley Scales of Infant and Toddler Development
23 in rural Bangladesh. *BMC Pediatrics*, 23(1), 1–11. <https://doi.org/10.1186/s12887-022-03800-6>
- 24
25 Prado, E. L., Phuka, J., Ocansey, E., Maleta, K., Ashorn, P., Ashorn, U., Adu-Afarwuah, S.,
26 Oaks, B. M., Lartey, A., & Dewey, K. G. (2018). A method to develop vocabulary
27 checklists in new languages and their validity to assess early language development.
28 *Journal of Health, Population, and Nutrition*, 37(1), 13. <https://doi.org/10.1186/s41043-018-0145-1>
- 29
30 Prathanee, B., Pongjanyakul, A., & Chano, J. (2008). Thai speech and language test for children
31 between 1 and 2 years of age. *International Journal of Language and Communication
32 Disorders*, 43(1), 125–140. <https://doi.org/10.1080/13682820601181208>
- 33
34 Puglisi, M. L., Blasi, H. F., & Snowling, M. J. (2020). Screening for the identification of oral
35 language difficulties in brazilian preschoolers: A validation study. *Language, Speech, and
36 Hearing Services in Schools*, 51(3), 852–865. https://doi.org/10.1044/2020_LSHSS-19-00083
- 37
38 Ranjitkar, S., Kvestad, I., Strand, T. A., Ulak, M., Shrestha, M., Chandyo, R. K., Shrestha, L., &
39 Hysing, M. (2018). Acceptability and reliability of the Bayley Scales of infant and Toddler
40 Development-III Among Children in Bhaktapur, Nepal. *Frontiers in Psychology*, 9(JUL),
41 1–10. <https://doi.org/10.3389/fpsyg.2018.01265>
- 42
43 Rasheed, M. A., Tabussum, G., Habib, Z. H., & Yousafzai, A. K. (2017). Adaptation of
44 MacArthur Communicative Development Inventory in rural Pakistan – useful tool for early
45 childhood studies. *Child: Care, Health and Development*, 43(3), 427–434.
46 <https://doi.org/10.1111/cch.12450>
- 47
48
49
50
51
52
53
54
55
56
57
58
59
60

- Rasheed, M. A., Kvestad, I., Shaheen, F., Memon, U., & Strand, T. A. (2023). The predictive validity of Bayley Scales of Infant and Toddler Development-III at 2 years for later general abilities: Findings from a rural, disadvantaged cohort in Pakistan. *PLOS Global Public Health*, 3(1), e0001485. <https://doi.org/10.1371/journal.pgph.0001485>
- Rasheed, M. A., Pham, S., Memon, U., Siyal, S., Obradović, J., & Yousafzai, A. K. (2018). Adaptation of the Wechsler Preschool and Primary Scale of Intelligence-III and lessons learned for evaluating intelligence in low-income settings. *International Journal of School and Educational Psychology*, 6(3), 197–207. <https://doi.org/10.1080/21683603.2017.1302851>
- Razak, R. A., Madison, C. L., Siow, Y. K., & Aziz, M. A. A. (2010). Preliminary Content Validity and Reliability of a Newly Developed Malay Preschool Language Assessment Tool. *Asia Pacific Journal of Speech, Language and Hearing*, 13(4), 217–234. <https://doi.org/10.1179/136132810805334985>
- Rodriguez, V. J., Zegarac, M., La Barrie, D. L., Parrish, M. S., Matseke, G., Peltzer, K., & Jones, D. L. (2020). Validation of the Bayley Infant Neurodevelopmental Screener among HIV-Exposed Infants in Rural South Africa. *Journal of Acquired Immune Deficiency Syndromes*, 85(4), 507–516. <https://doi.org/10.1097/QAI.0000000000002479>
- Rubio-Codina, M., Araujo, M. C., Attanasio, O., Muñoz, P., & Grantham-McGregor, S. (2016). Concurrent validity and feasibility of short tests currently used to measure early childhood development in large scale studies. *PLoS ONE*, 11(8), 1–17. <https://doi.org/10.1371/journal.pone.0160962>
- Rubio-Codina, M., & Grantham-McGregor, S. (2020). Predictive validity in middle childhood of short tests of early childhood development used in large scale studies compared to the Bayley-III, the Family Care Indicators, height-for-age, and stunting: A longitudinal study in Bogota, Colombia. *PLoS ONE*, 15(4), 1–20. <https://doi.org/10.1371/journal.pone.0231317>
- Saccani, R., Valentini, N. C., & Pereira, K. R. G. (2016). New Brazilian developmental curves and reference values for the Alberta infant motor scale. *Infant Behavior and Development*, 45, 38–46. <https://doi.org/10.1016/j.infbeh.2016.09.002>
- 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>
- Samarakkody, D. C., Fernando, D. N., Perera, H., McClure, R. J., & De Silva, H. (2010). The Child Behaviour Assessment Instrument: Development and validation of a measure to screen for externalising child behavioural problems in community setting. *International Journal of Mental Health Systems*, 4, 1–8. <https://doi.org/10.1186/1752-4458-4-13>

- Sánchez-Vincitore, L. V., Schaettle, P., & Castro, A. (2019). Validation of the Malawi Developmental Assessment Tool for children in the Dominican Republic: Preliminary results. *PLoS ONE*, 14(8), 1–11. <https://doi.org/10.1371/journal.pone.0221162>
- Scherzer, A. L. (2009). Experience in Cambodia with the use of a culturally relevant developmental milestone chart for children in low- and middle-income countries. *Journal of Policy and Practice in Intellectual Disabilities*, 6(4), 287–292. <https://doi.org/10.1111/j.1741-1130.2009.00234.x>
- Shah, M. B., Schaefer, B. A., & Clark, T. P. (2013). Reliability of Bracken School Readiness Assessment, Third Edition Scores With Young Children in Mumbai, India. *International Journal of School and Educational Psychology*, 1(1), 58–66. <https://doi.org/10.1080/21683603.2013.782595>
- Shaoli, S. S., Islam, S., Haque, S., & Islam, A. (2019). Validating the Bangla version of the Dimensions of Mastery Questionnaire (DMQ-18) for preschoolers. *Asian Journal of Psychiatry*, 44(May 2019), 143–149. <https://doi.org/10.1016/j.ajp.2019.07.044>
- Shrestha, M., Strand, T. A., Ulak, M., Chandyo, R. K., Ranjitkar, S., Hysing, M., Shrestha, L., & Kvestad, I. (2019). The feasibility of the Ages and Stages Questionnaire for the assessment of child development in a community setting in Nepal. *Child: Care, Health and Development*, 45(3), 394–402. <https://doi.org/10.1111/cch.12654>
- Siegle, C. B. H., & dos Santos Cardoso de Sá, C. (2018). Concurrent validity between instruments of assessment of motor development in infants exposed to HIV. *Infant Behavior and Development*, 50(February), 198–206. <https://doi.org/10.1016/j.infbeh.2018.01.005>
- Sincovich, A., Gregory, T., Zanon, C., Santos, D. D., Lynch, J., & Brinkman, S. A. (2019). Measuring early childhood development in multiple contexts: The internal factor structure and reliability of the early Human Capability Index in seven low and middle income countries. *BMC Pediatrics*, 19(1), 1–14. <https://doi.org/10.1186/s12887-019-1852-5>
- Snelling, M., Dawes, A., Biersteker, L., Girdwood, E., & Tredoux, C. (2019). The development of a South African Early Learning Outcomes Measure: A South African instrument for measuring early learning program outcomes. *Child: Care, Health and Development*, 45(2), 257–270. <https://doi.org/10.1111/cch.12641>
- Soleimani, F., & Dadkhah, A. (2007). Validity and reliability of infant Neurological International Battery for detection of gross motor developmental delay in Iran. *Child: Care, Health and Development*, 33(3), 262–265. <https://doi.org/10.1111/j.1365-2214.2006.00704.x>
- Soleymani, Z., Nematzadeh, S., Tehrani, L. G., Rahgozar, M., & Schneider, P. (2016). Language sample analysis: development of a valid language assessment tool and determining the reliability of outcome measures for Farsi-speaking children. *European Journal of Developmental Psychology*, 13(2), 275–291. <https://doi.org/10.1080/17405629.2016.1152175>

Solís Cordero K, Quezada Ugalde AM, González Serrano S, Zúñiga Baldi C. Costa Rican test: Evaluation of the children's integral development from 0 months to 6 years of age (EDIN II). *Enfermería Actual Costa Rica* 2019; 37.
DOI:<http://dx.doi.org/10.15517/revenf.v0ino.37.34959>.

Soysal, A. S., Gucuyener, K., Ergenekon, E., Turan, Ö., Koc, E., Turkylmaz, C., Önal, E., & Atalay, Y. (2014). The prediction of later neurodevelopmental status of preterm infants at ages 7 to 10 years using the bayley infant neurodevelopmental screener. *Journal of Child Neurology*, 29(10), 1349–1355. <https://doi.org/10.1177/0883073813520495>

Thompson, L., Peñaloza, R. A., Stormfields, K., Kooistra, R., Valencia-Moscoso, G., Muslima, H., & Khan, N. Z. (2015). Validation and adaptation of rapid neurodevelopmental assessment instrument for infants in Guatemala. *Child: Care, Health and Development*, 41(6), 1131–1139. <https://doi.org/10.1111/cch.12279>

Toor, H. K., & Hanif, R. (2022). Executive Function Touch battery: Translation and preliminary measure validation for Pakistani preschoolers. *PLoS ONE*, 17(9 September), 1–14. <https://doi.org/10.1371/journal.pone.0274431>

Uyanik, Ö., & Kandir, A. (2014). Adaptation of the Kaufman Survey of Early Academic and Language Skills to Turkish Children Aged 61 to 72 Months. *Kuram ve Uygulamada Eğitim Bilimleri*, 14(2), 682–692. <https://doi.org/10.12738/estp.2014.2.1682>

Valentini, N. C., & Saccani, R. (2012). Brazilian Validation of the Alberta Infant Motor Scale. *Physical Therapy*, 92(3), 440–447. <https://doi.org/10.2522/ptj.20110036>

Vameghi, R., Sajedi, F., Kraskian Mojembari, A., Habiollahi, A., Lornezhad, H. R., & Delavar, B. (2013). Cross-cultural adaptation, validation and standardization of Ages and Stages Questionnaire (ASQ) in Iranian children. *Iranian Journal of Public Health*, 42(5), 522–528.

van Der Linde, J., Swanepoel, D. W., Glascoe, F. P., Louw, E. M., & Vinck, B. (2015). Developmental screening in South Africa: Comparing the national developmental checklist to a standardized tool. *African Health Sciences*, 15(1), 188–196. <https://doi.org/10.4314/ahs.v15i1.25>

van Heerden, A., Hsiao, C., Matafwali, B., Louw, J., & Richter, L. (2017). Support for the feasibility of the ages and stages questionnaire as a developmental screening tool: A cross-sectional study of South African and Zambian children aged 2-60 months. *BMC Pediatrics*, 17(1), 1–9. <https://doi.org/10.1186/s12887-017-0802-3>

Venancio, S. I., Bortoli, M. C., Frias, P. G., Giugliani, E. R. J., Alves, C. R. L., & Santos, M. O. (2020). Development and validation of an instrument for monitoring child development indicators. *Jornal de Pediatria (Versão Em Português)*, 96(6), 778–789. <https://doi.org/10.1016/j.jpedp.2019.10.008>

- Venter, A., Pienaar, A. E., & Coetzee, D. (2015). Suitability of the 'Little DCDQ' for the identification of DCD in a selected group of 3–5-year-old South African children.' *Early Child Development and Care*, 185(8), 1359–1371.
<https://doi.org/10.1080/03004430.2014.1000887>
- Waechter, R., Evans, R., Hanna, S., Murray, T., Mobley, C., Holmes, S., Isaac, R., Wolfe, R., Andrew, E., Landon, B., & Fernandes, M. (2022). Adaptation of the INTERGROWTH-21st neurodevelopment assessment (INTER-NDA) to the context of the English-speaking Caribbean. *BMC Pediatrics*, 22(1), 1–10. <https://doi.org/10.1186/s12887-021-03039-7>
- Waldman, M., McCoy, D. C., Seiden, J., Cuartas, J., Fink, G., Altafim, E., Brentani, A., Castellanos, A., Chen, A., Chomat, A. M., Fawzi, W., de Pin˜eres, C. G., Hamadani, J., Henao, N., Jervis, P., Kane, C., Measelle, J., Medrano, P., Pisani, L., ... Yousafzai, A. K. (2021). Validation of motor, cognitive, language, and socio-emotional subscales using the Caregiver Reported Early Development Instruments: An application of multidimensional item factor analysis. *International Journal of Behavioral Development*, 45(4), 368–377.
<https://doi.org/10.1177/01650254211005560>
- Washington, K. N., McDonald, M. M., McLeod, S., Crowe, K., & Devonish, H. (2017). Validation of the intelligibility in context scale for Jamaican Creole-speaking preschoolers. *American Journal of Speech-Language Pathology*, 26(3), 750–761.
https://doi.org/10.1044/2016_AJSLP-15-0103
- Weber, A. M., Marchman, V. A., Diop, Y., & Fernald, A. (2018). Validity of caregiver-report measures of language skill for Wolof-learning infants and toddlers living in rural African villages. *Journal of Child Language*, 45(4), 939–958.
<https://doi.org/10.1017/S0305000917000605>
- Wijedasa, D. (2012). Developmental screening in context: Adaptation and standardization of the Denver Developmental Screening Test-II (DDST-II) for Sri Lankan children. *Child: Care, Health and Development*, 38(6), 889–899. <https://doi.org/10.1111/j.1365-2214.2011.01332.x>
- Willoughby, M. 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>
- Willoughby, M. T., Piper, B., Oyanga, A., & Merseth King, K. (2019). Measuring executive function skills in young children in Kenya: Associations with school readiness. *Developmental Science*, 22(5). <https://doi.org/10.1111/desc.12818>
- Wolf, S., Halpin, P., Yoshikawa, H., Dowd, A. J., Pisani, L., & Borisova, I. (2017). Measuring school readiness globally: Assessing the construct validity and measurement invariance of the International Development and Early Learning Assessment (IDEA) in Ethiopia. *Early Childhood Research Quarterly*, 41(June), 21–36.
<https://doi.org/10.1016/j.ecresq.2017.05.001>

- 1
2
3 Wu, Z., Hu, B. Y., & Fan, X. (2019). Cross-Cultural Validity of Preschool Learning Behavior
4 Scale in Chinese Cultural Context. *Journal of Psychoeducational Assessment*, 37(1), 125–
5 130. <https://doi.org/10.1177/0734282916651538>
- 6
7 Wu, Z., Mak, M. C. K., Hu, B. Y., He, J., & Fan, X. (2019). A validation of the Social Skills
8 domain of the Social Skills Improvement System-Rating Scales with Chinese preschoolers.
9 *Psychology in the Schools*, 56(1), 126–147. <https://doi.org/10.1002/pits.22193>
- 10
11 Xie, H., Bian, X., Chen, C. Y., Squires, J., & Lu, P. (2019). Examining the Convergent Evidence
12 of a Parent-Completed, Social-Emotional Screening Tool in China. *Journal of Child and
13 Family Studies*, 28(6), 1471–1480. <https://doi.org/10.1007/s10826-019-01362-3>
- 14
15 Xie, H., Clifford, J., Squires, J., Chen, C. Y., Bian, X., & Yu, Q. (2017). Adapting and validating
16 a developmental assessment for Chinese infants and toddlers: The ages & stages
17 questionnaires: Inventory. *Infant Behavior and Development*, 49(June), 281–295.
18 <https://doi.org/10.1016/j.infbeh.2017.09.004>
- 19
20 Xie, S., & Li, H. (2019). Development and Validation of the Chinese Preschool Readiness Scale.
21 *Early Education and Development*, 30(4), 522–539.
22 <https://doi.org/10.1080/10409289.2019.1596462>
- 23
24 Yuan, H., Ocansey, M., Adu-Afarwuah, S., Sheridan, M., Hamoudi, A., Okronipa, H.,
25 Kumordzie, S. M., Oaks, B. M., & Prado, E. (2022). Evaluation of a tablet-based
26 assessment tool for measuring cognition among children 4–6 years of age in Ghana. *Brain
and Behavior*, 12(10), 1–12. <https://doi.org/10.1002/brb3.2749>
- 27
28 Yue, A., Jiang, Q., Wang, B., Abbey, C., Medina, A., Shi, Y., & Rozelle, S. (2019). Concurrent
29 validity of the Ages and Stages Questionnaire and the Bayley Scales of Infant Development
30 III in China. *PLoS ONE*, 14(9). <https://doi.org/10.1371/journal.pone.0221675>
- 31
32 Yunilda, E., Gunardi, H., Medise, B. E., & Oswari, H. (2023). The Indonesian version of Ages
33 and Stages Questionnaire III accuracy compared to Bayley Scales of Infant Development
34 III. *Infant and Child Development*, 32(1), 1–12. <https://doi.org/10.1002/icd.2387>
- 35
36 Zakaria, S., Chua, B. S., Sombuling, A., Ahmad, S., & Hashmi, S. I. (2012). Reliability and
37 Validity for Malay Version of Bayley Scales of Infant and Toddler Development-Third
38 Edition (Bayley-III): Preliminary Study. *International Proceedings of Economics
39 Development and Research*, 40(June), 116–119.
- 40
41 Zhang, L., Li, Z., & He, K. (2018). Validation of the Chinese Revision of Emotional and
42 Behavioral Screener among Preschoolers (4–7 years old). *Journal of Child and Family
43 Studies*, 27(3), 771–779. <https://doi.org/10.1007/s10826-017-0940-0>
- 44
45 Zhang, Y., Xu, X., Jiang, Y., Sun, W., Wang, Y., Song, Y., Dong, S., Zhu, Q., Jiang, F., &
46 Sheng, L. (2021). Early language and communication development in Chinese children:
47
- 48
49
50
51
52
53
54
55
56
57
58
59
60

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

1
2
3 Adaption and validation of a parent report instrument. *International Journal of Speech-*
4 *Language Pathology*, 23(4), 372–381. <https://doi.org/10.1080/17549507.2020.1817558>
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Zirakashvili, M., Gabunia, M., Tatishvili, N., Ediberidze, T., Lomidze, G., Chachava, T., & Hix-Small, H. (2018). Cultural Adaptation and Psychometric Validation of the Ages and Stages Questionnaires for use in Georgia. *Journal of Child and Family Studies*, 27(3), 739–749. <https://doi.org/10.1007/s10826-017-0917-z>

For peer review only

BMJ Open

Psychometric properties of early childhood development assessment tools in low-and-middle-income countries: a systematic review

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2024-096365.R1
Article Type:	Original research
Date Submitted by the Author:	27-Mar-2025
Complete List of Authors:	Bliznashka, Lilia; International Food Policy Research Institute, Nutrition, Diets, and Health Unit; University of Edinburgh, Global Academy of Agriculture and Food Systems Hentschel, Elizabeth; Harvard T H Chan School of Public Health, Department of Global Health and Population Ali, Nazia Binte; Harvard T H Chan School of Public Health, Department of Global Health and Population Hunt, Xanthe; Stellenbosch University, Department of Global Health; Africa Health Research Institute (AHRI) Neville, Sarah ; Brown University School of Public Health, International Health Institute Olney, D.; International Food Policy Research Institute, Nutrition, Diets, and Health Unit Pitchik, Helen; University of California Berkeley, School of Public Health Roy, Aditi; Centre for Chronic Disease Control Seiden, Jonathan; Harvard Graduate School of Education Solís-Cordero, Katherine ; University of Costa Rica, School of Nursing Thapa, Aradhana; Emory University Hubert Department of Global Health Jeong, Joshua; Emory University Hubert Department of Global Health
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Global health, Paediatrics
Keywords:	PAEDIATRICS, Child, Community child health < PAEDIATRICS, Patient Reported Outcome Measures, Psychometrics

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Psychometric properties of early childhood development assessment tools in low-and-middle-income countries: a systematic review

Lilia Bliznashka, PhD^{1,2,†}, Elizabeth Hentschel, PhD³, Nazia Binte Ali, MPH³, Xanthe Hunt, PhD^{4,5}, Sarah Elizabeth Neville, PhD⁶, Deanna Olney, PhD¹, Helen O. Pitchik, PhD⁷, Aditi Roy, PhD⁸, Jonathan Seiden, EdM⁹, Katherine Solís Cordero, PhD¹⁰, Aradhana Thapa, MPH¹¹, Joshua Jeong, ScD¹¹

¹ International Food Policy Research Institute, Washington, DC, 20001

² Global Academy of Agriculture and Food Systems, University of Edinburgh, UK

³ Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA

⁴ Institute for Life Course Health Research, Department of Global Health, Stellenbosch University, South Africa

⁵ Africa Health Research Institute (AHRI), Somkele, South Africa

⁶ Brown University School of Public Health, International Health Institute, Providence, RI, USA

⁷ Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, California, USA

⁸ Centre for Chronic Disease Control, New Delhi, India

⁹ Harvard Graduate School of Education, Cambridge, MA, USA

¹⁰ School of Nursing, University of Costa Rica, San José, Costa Rica

¹⁰ Hubert Department of Global Health, Emory University, Atlanta, GA, USA

1
2
3 † After the first, second, and last authors, all authors are listed in alphabetical order by last name.
4
5
6

7 **Corresponding author:**
8
9

10 Lilia Bliznashka
11
12

13 l.bliznashka@cgiar.org
14
15

16 1201 I St NW, Washington, DC 20001, US
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1 **Abstract**

2 **Objective:** Valid and reliable measurement of early childhood development (ECD) is critical for
3 monitoring and evaluating ECD-related policies and programmes. Although ECD tools developed
4 in high-income countries may be applicable to low- and middle-income countries (LMICs),
5 directly applying them in LMICs can be problematic without psychometric evidence for new
6 cultures and contexts. Our objective was to systematically appraised available evidence on the
7 psychometric properties of tools used to measure ECD in LMIC.

8 **Design:** A systematic review following the Preferred Reporting Items for Systematic Reviews and
9 Meta-Analyses guidelines.

10 **Data sources:** MEDLINE, Embase, PubMed, PsycInfo, SciELO, and BVS were searched from
11 inception to February 2025.

12 **Eligibility criteria:** We included studies that examined the reliability, validity, and measurement
13 invariance of tools assessing ECD in children 0-6 years of age living in LMICs.

14 **Data extraction and synthesis:** Each study was independently screened by two researchers and
15 data extracted by one randomly assigned researcher. Risk of bias was assessed using a checklist
16 developed by the study team assessing bias due to training/administration, selective reporting, and
17 missing data. Results were synthesized narratively by country, location, age group at assessment,
18 and developmental domain.

19 **Results:** A total of 160 articles covering 117 tools met inclusion criteria. Most reported
20 psychometric properties were internal consistency reliability ($n=117$, 64%), concurrent validity
21 ($n=81$, 45%), convergent validity ($n=74$, 41%), test-retest reliability ($n=73$, 40%), and structural
22 validity ($n=72$, 40%). Measurement invariance was least commonly reported ($n=16$, 9%). Most
23 articles came from Brazil, China, India, and South Africa. Most psychometric evidence was from

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

1
2
3 24 urban (n=92, 51%) or urban-rural (n=41, 23%) contexts. Study samples focused on children aged
4
5 25 6-17.9 or 48-59.9 months. The most assessed developmental domains were language (n=111,
6
7 26 61%), motor (n=104, 57%) and cognitive (n=82, 45%). Bias due to missing data was most
8
9 27 common.

10
11
12 28 **Conclusions:** Psychometric evidence is fragmented, limited, and heterogeneous. More rigorous
13
14 psychometric analyses, especially on measurement invariance, are needed to establish the quality
15
16 and accuracy of ECD tools for use in LMICs.

17
18 31
19
20 32

1
2
3 **33 Strengths and limitations of this study**

- 4
5 • This systematic review used extensive search filters, made no restrictions on
6 developmental domains, and included both single- and multi-domain tools.
7
8 • The review team was rigorously trained and brought diverse backgrounds in ECD
9 measurement in LMICs.
10
11 • We could not conduct full text review and extraction for articles in Chinese, Farsi, and
12 Turkish because these languages were not spoken by the review team.
13
14 • We did not use a validated tool to assess risk of bias, which may limit the comparability of
15 our findings.

43 INTRODUCTION

44 The Sustainable Development Goals (SDGs) recognise the importance of early childhood
45 development (ECD),¹ which sets the foundation for children's later learning and economic
46 outcomes.^{2–4} Rigorous ECD measurement is critical for accurate comparability of children's skills
47 across populations and time, evaluating the effectiveness of ECD interventions, and tracking ECD-
48 related policies.

49 In low-and-middle-income countries (LMICs), >140 tools (60% of which originated in
50 high-income countries (HICs)) have been used to assess ECD in children 0–8 years old.⁵ Although
51 ECD tools developed in HICs may be applicable to LMICs, directly applying ECD tools from
52 HICs in LMICs can be problematic without psychometric evidence for new cultures and contexts.
53 Few studies to-date have synthesised the evidence on the psychometric properties of ECD tools.

54 Evidence pertaining to a tool's reliability, validity, and measurement invariance in a given
55 context is critical for selecting an ECD outcome or indicator. Other factors include the purpose of
56 measurement, the population and age-range of interest, the developmental domain(s) of interest,
57 and administration time and cost.⁶ Evidence of reliability and validity ensures consistent and
58 accurate ECD measurement,⁷ whereas evidence of measurement invariance guarantees assessment
59 of the same construct across countries and sub-groups.^{7,8}

60 Prior reviews have provided guidance for selecting ECD tools for use in LMICs^{5,9–11} and
61 underscored that evidence on tool reliability and validity is fundamental.^{9,10} However, prior work
62 has often focused on individual ECD domains,^{12–14} which has limited use for population level
63 assessment or tracking of SDG-related policies, or a subset of psychometric properties,¹⁰ which is
64 not evidence of reliability and validity as a whole. Further, prior reviews do not disaggregate use
65 of ECD tools by HICs vs LMICs or urban vs rural settings,^{12–14} despite known and persistent

1
2
3 66 disparities in ECD by country income level and urban/rural residence^{15,16} and the fact that young
4 67 child populations are increasingly diverse due to migration and urbanization.¹³ These are important
5 68 distinctions given that psychometric properties can vary by population characteristics.
6
7

8 69 We reviewed available evidence on 10 psychometric properties of tools used to assess ECD
9 70 in children 0-6 years old living in LMICs. We summarised the current landscape by ECD tool,
10 71 developmental domain, country, and age group, focusing on 10 types of psychometric evidence.
11 72 We sought to deepen our understanding and consistently summarise whether psychometric
12 73 evidence exists for tools used to measure ECD in LMICs. Our findings can assist stakeholders in
13 74 selection of ECD tools for their intended use and context, and inform what research is needed to
14 75 improve how we track ECD-related SDGs, programmes, and policies in LMICs.
15
16 76
17
18 77 **METHODS**
19
20 78 **Search strategy and selection criteria**
21
22 79 We identified articles in any language through MEDLINE, Embase, PubMed, PsycInfo,
23 80 SciELO, and BVS. The search strategy used medical subject headings (MeSH), keywords, and
24 81 free text words along four key elements: population, construct, measurement properties, and
25 82 location (**Supplemental Tables 1**). Search terms were combined using Boolean operators.
26 83 Truncation wildcards were used to include variations of the search terms. The search strategy used
27 84 a combination of searches through titles, abstracts, and keywords. The COnsensus-based Standards
28 85 for the selection of health status Measurement INstruments (COSMIN) measurement properties
29 86 filter was used for the third search element.¹⁷ The search strategy was piloted in PubMed and then
30 87 adapted for the remaining databases.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Full-text, peer-reviewed articles were included if: (1) the study was conducted in a LMIC, (2) included children 0-6 years old, (3) included at least one ECD domain (cognitive, language, motor, social-emotional, attention/executive function, personal-social, and pre-academic/academic) defined in **Supplemental Table 2**, (4) developed a new tool or adapted an existing one, and provided primary evidence of at least one of 10 psychometric properties in terms of reliability, validity, or measurement invariance (**Table 1**), and (5) were published between January 1, 2007 and March 9, 2023. An updated search was conducted on February 28, 2025 and seven additional studies were identified for inclusion: five through the database searches and two from a Google Scholar alert. The 10 psychometric properties were selected based on prior systematic reviews on psychometrics properties of ECD tools,^{9,10,12,18,19} classical test theory,²⁰ and reviews of measurement in cross-cultural psychology.^{21,22} We considered all ECD tools regardless of their intended or actual use (diagnosis, screening, or surveillance). We included both articles that adapted original tools (“adaptation articles” hereafter) and articles that developed new tools (“development articles” hereafter). Articles were included if at least 50% of the study sample was within the range 0-6 years and the average age in the sample was <6 years or unspecified. If information on the child age range was not available or clear, the article was included. Multi-country articles were included if ≥50% of the countries were LMICs. If an article reported that measurement properties were reported elsewhere (e.g., referencing another study and thus providing secondary evidence), the article was ineligible. We included prospective, retrospective, cross-sectional, and longitudinal quantitative study designs. Cited references were reviewed for potential inclusion.

We excluded articles where a tool was used to assess an outcome measure (e.g., trials reporting impacts on ECD outcomes), but the article did not include measurement objectives.²³

1
2
3 111 Among articles with an explicit measurement focus, we excluded those reporting only on
4 convergent validity (e.g., correlations with socio-demographic variables) which alone provides
5 limited psychometric evidence.²³ We excluded articles studying children with developmental
6 disabilities and disorders (e.g., autism spectrum disorder, cerebral palsy), and children with
7 physical disabilities that impair performance on ECD measures (e.g., deafness, blindness). Lastly,
8 we excluded the following study designs: animal studies, simulation studies, case studies,
9 opinions, letters, preprints, protocols, conference abstracts, ecological studies, dissertations/theses,
10 reviews or systematic reviews, and meta-analyses.
11
12
13
14
15
16
17
18
19
20
21
22

119 **Data Extraction**

23
24 120 Search results were imported into Covidence, where duplicates were automatically
25 removed. Two reviewers independently screened titles and abstracts for inclusion and reviewed
26 full texts of retained articles. Disagreements in screening or full text review were resolved through
27 discussion with a third reviewer.
28
29
30
31
32

33 124 One reviewer developed the data extraction sheet, and two reviewers piloted it. Revisions
34 were made through discussion between the two reviewers. Seven reviewers were trained on using
35 the data extraction sheet through pilot extractions of included articles. Further revisions were made
36 to the data extraction sheet based on discussions of the piloting process. Data extraction included
37 information on publication details, study meta-data, characteristics of the ECD tool, type of
38 administration, ECD domains assessed, and psychometric properties with respect to reliability,
39 validity, and measurement invariance. After piloting, each article was extracted by one reviewer.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54 131 For quality assurance, 20% of articles were randomly selected for independent extraction by a
55 second reviewer. Any discrepancies regarding data extraction were resolved by a third reviewer.
56
57
58
59
60 133 Authors of included articles were not contacted when information was missing or unclear.

134 Risk of Bias Assessment

135 Since we aimed to assess the quality of the underlying studies, rather than the quality of
136 the psychometric properties, we could not use a validated risk of bias tool and instead developed
137 a new one. One reviewer created a risk of bias checklist by adapting items from the COSMIN²³
138 and Cochrane's ROBINS-E²⁴ risk of bias tools. The checklist was then refined via discussions with
139 two other reviewers and a psychometrician, and after piloting by four reviewers. The final checklist
140 contained three categories of bias due to: (1) training/administration (not assessed for tools relying
141 on self-assessment), (2) selective reporting (only assessed for studies reporting on convergent or
142 predictive validity), and (3) missing data. Each article was rated separately on each category using
143 the ROBINS-E risk of bias ratings: low risk, some concerns, high risk, and very high risk. We also
144 assessed indirectness of populations²⁵ by assessing whether the sample was limited to a specific
145 setting, the tool covered the entire age range it was intended for, subgroups were generalisable,
146 and results were generalisable. Where insufficient information was provided in the article, we rated
147 the study as "unable to assess". Seven reviewers were trained and conducted the risk of bias
148 assessment at the article level. Disagreements for articles assessed by two reviewers (20%) were
149 resolved through discussion with a third reviewer.

150 Data Synthesis

151 Data analysis was conducted at the article-tool level because some articles reported on
152 multiple tools. We created binary indicators for whether evidence on each one of the 10
153 psychometric properties was reported. In longitudinal studies, we considered the psychometric
154 properties reported at any time point. We then summarised the evidence by country, location, age
155 group at assessment, type of article (development vs adaptation), developmental domain, and ECD
156 tool. For all tools, child age at assessment was converted to months based on the information

1
2
3 157 provided in the article. For longitudinal studies, age at first assessment was used. For multi-
4 site/multi-country studies, the full age range at assessment was used across the sites/countries.
5
6 159 Because the objective the paper was to take stock of the available evidence rather than to report
7
8 160 the adequacy and relevance of specific psychometric properties, we did not summarize evidence
9
10 161 on the psychometric properties themselves. Results were synthesised narratively.
11
12
13
14

15 162 This review followed the Preferred Reporting Items for Systematic Reviews and Meta-
16 Analyses (PRISMA) guidelines.²⁶ It was preregistered on PROSPERO: CRD42022372305.
17
18

19 164 Patient and Public Involvement

20
21

22 165 Patients and/or the public were not involved in the design, conduct, reporting or
23
24 166 dissemination plans of this research.
25
26

27 167 Results

28
29

30 168 Study characteristics

31
32

33 169 Our search strategy identified 6,430 records from six databases and two other sources
34
35 170 (**Figure 1**). After removing duplicates, 5,388 records were excluded during title and abstract
36 screening. After full text review of the remaining 250 records, 98 records were excluded. After
37 updating the search in February 2025, we included seven more articles (five identified from the
38
39 172 six databases and two identified from other sources). We included 160 articles, covering 117 tools.
40
41 173 Four articles reported on multiple tools, resulting in 182 article-tool combinations (referred to as
42
43 174 articles for brevity).
44
45

46 176 Psychometric evidence available

47
48

49 177 Most articles were adaptation article (n=145, 80%) (**Figure 2, Supplemental Table 3**).
50
51 178 The most often evaluated psychometric properties were internal consistency reliability (n=117,
52
53 179 64%), concurrent validity (n=81, 45%), convergent validity (n=74, 41%), test-retest reliability
54
55
56
57
58
59

1
2
3 180 (n=73, 40%), and structural validity (n=72, 40%). Measurement invariance was the least frequently
4 reported psychometric property (n=16, 9%) and was primarily evaluated over countries and child
5 sex. The number of articles increased over time (**Supplemental Figure 1**). Between 2007-2010,
6
7 181 most included articles were development articles; since 2011, most included articles have been
8 adaptation articles.
9
10 182
11 183
12 184
13

14 185 *Psychometric evidence available by country*

15
16
17 186 Psychometric evidence came from 55 countries (**Figure 3**); one-third of articles came from
18 Brazil (n=22, 12%), China (n=18, 10%), India (n=14, 8%), and South Africa (n=10, 5%). For 40%
19 of countries represented, there was only one article reporting psychometric evidence
20
21 188
22 189 (**Supplemental Table 4**). Psychometric evidence for each ECD tool included was generally
23 limited to 1-2 countries with a few notable exceptions: the Global Scales for Early Development
24 (GSED, 32 countries), the Caregiver Reported Early Development Instrument (CREDI, 18
25 countries), the International Development and Early Learning Assessment (IDELA, 17 countries),
26
27 190 countries), the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III, 12 countries), and
28
29 191 countries), the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III, 12 countries), and
30
31 192 the Ages and Stages Questionnaire-3 (ASQ-3, 11 countries) (**Supplemental Table 5**).
32
33
34 193
35
36 194
37
38 195 *Psychometric evidence available by location*

39
40 196 Most psychometric evidence came from urban (n=92, 51%) or urban-rural (n=41, 23%)
41 settings. In South Asia and Sub-Saharan Africa, a similar number of articles originated from urban,
42
43 197 rural, and urban-rural settings (**Supplemental Table 6**).
44
45 198

46 199 *Psychometric evidence available by ECD tool*

47
48 200 For development articles, a single article reported psychometric properties for all tools
49 except for the CREDI covered in two articles (**Supplemental Table 7**). For adaptation articles,
50
51 201 ASQ-3 and BSID-III were most often studied (14 and 12 articles, respectively). For 75% of tools
52
53 202
54
55
56
57
58
59
60

1
2
3 203 only a single article provided psychometric evidence (**Supplemental Table 8**). Other frequently
4 studied tools included the Mullen Scales of Early Learning (n=5 articles), the Alberta Infant Motor
5 Scale (n=4 articles), the Bayley Infant Neurodevelopmental Screener (n=4 articles), Denver
6 Developmental Screening Test (Denver-II, n=4 articles), and International Development and Early
7 Learning Assessment (IDEA, n=4 articles) (**Supplemental Table 5**).
8
9
10
11
12
13
14
15 208 *Psychometric evidence available by age group*
16
17
18 209 Although included tools targeted 0-71.9-month-old children, most studies focused on
19 children 12–17.9 month (n=86, 47%), 6-11.9 months (n=82, 45%), or 54-59.9 months (n=76, 42%)
20 old. Articles largely did not cover the full age groups defined here, or the full age range the ECD
21 tool can be used for. Some articles included age groups as narrow as one month (**Supplemental
22 Figures 2-3**).
23
24
25
26
27
28 214 *Psychometric evidence available by developmental domain*
29
30
31 215 Most articles reported on language (n=111, 61%), motor (n=104, 57%), and cognitive (n=82,
32 45%) development (**Supplemental Table 9**). Academic/pre-academic was the least studied
33 domain (n=10, 5%). Within domains, internal consistency reliability was most frequently reported
34
35
36
37
38 218 (**Figure 4**).
39
40
41 219 *Risk of bias assessment*
42
43
44 220 Bias due to training/administration was assessed for 59 articles (32%). Of these, 80% had
45 low risk of bias (n=47) and 20% some concerns (n=12) (**Supplemental Tables 10-11**). Risk of
46
47
48
49 222 bias due to selective reporting was assessed for 95 articles (52%). Of these, most articles (n=86,
50 91%) had low risk of bias. Risk of bias due to missing data was assessed for 137 articles (75%).
51
52
53
54 224 Of these, 48% had some concerns (n=52) and 5% high risk (n=7). For indirectness, 20 articles
55 (11%) were rated as generalisable and 82 (45%) as probably generalisable.
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

226

227 DISCUSSION

228 Based on 160 articles, available evidence on 10 psychometric properties of 117 ECD tools
229 for children 0-6 years old is fragmented, limited, and heterogeneous. The most frequently provided
230 evidence was on internal consistency reliability, test-retest reliability and/or concurrent,
231 convergent, and structural validity. Psychometric evidence on measurement invariance was the
232 least commonly available. Although evidence came from 55 LMICs, four countries were most
233 represented. Most evidence came from urban or urban-rural settings. The most studied tools were
234 ASQ-3 and BSID-III.

235 Our findings support ECD measurement trends found in other work:²⁷ much of the work
236 on psychometric properties is recent, with ECD tools being developed and adapted concurrently.
237 Psychometric efforts remain limited to a few ECD tools,^{28,29} individual ECD domains,^{12–14} and
238 few psychometric properties.^{10,12} This fragmentation is evidenced by included articles focusing on
239 individual countries, limited age ranges, and single developmental domains. In addition, included
240 studies focused on individual reliability or validity properties (e.g., internal consistency reliability
241 and concurrent reliability, respectively), thus providing a limited picture of reliability and validity
242 as a whole. The resulting heterogeneous psychometric evidence can hinder comparability and
243 large-scale monitoring of ECD policies and programmes within and across LMICs, which is
244 crucial for identifying and implementing effective approaches to support ECD.

245 Despite efforts to consolidate ECD measurement through tools like CREDI, GSED, and
246 IDELA, such tools do not fully meet research, programmatic, and policy needs as evidenced by
247 the increase of adaptation and development articles since 2015. This is not surprising given that
248 no ECD tool is suitable for all populations.¹⁸ Our disaggregation by development and adaptation

1
2
3 249 articles permitted better understanding of the psychometric evidence available and highlighted
4 250 evidence gaps for both existing and newly developed tools.
5
6

7 251 This review highlights four important limitations of existing psychometric evidence for
8 252 ECD tools in LMICs. First, although most tools are designed for a wide age range, the
9 253 psychometric evidence behind most tools pertained to narrower age ranges and in some cases as
10 254 narrow as one month. This may have limited applicability to diverse age ranges (given that there
11 255 is a natural variability in child development in the early years¹³) in these specific contexts.
12 256 Relatedly, most psychometric evidence pertained to urban contexts. Given existing urban-rural
13 257 disparities in ECD^{15,16} and increasingly diverse young child populations in urban settings,¹³ ECD
14 258 tools whose psychometric properties were examined only in urban settings might be inadequate
15 259 for rural settings. Those developing or adapting ECD tools should consider establishing
16 260 psychometric properties across the full intended child age range and across both urban and rural
17 261 settings. Although this implies longer, more expensive, and logistically difficult studies, it would
18 262 ensure the tool can be used broadly and in more diverse populations.
19
20

21 263 A second limitation is the very little evidence on measurement invariance. While several
22 264 studies reported on samples drawn from multiple countries, few conducted statistical analysis to
23 265 test for equivalence across countries, thus providing no evidence of measurement invariance.⁸
24 266 Cross-country invariance, which guarantees assessment of the same construct across countries, is
25 267 key for tracking global SDG goals. More work in this domain is needed particularly for tools
26 268 widely used for policy making and programme evaluations. Measurement work should be
27 269 considered relative to competing and more urgent ECD priorities in LMICs.
28
29

30 270 Third, consistent with existing literature, we found limited psychometric evidence on tools
31 271 measuring socio-emotional and personal-social development.^{13,19} This is surprising given that
32
33

1
2
3 272 these two domains are among the most culturally specific,¹³ implying they require more
4 comprehensive and rigorous adaptation. In addition, psychometric evidence on tools to assess
5 attention/executive function and academic/preacademic development was very limited. Without
6 additional work to establish a psychometric base, this poses major challenges for those seeking to
7 monitor these domains in early life.
8
9
10
11
12
13

14
15 277 Lastly, in contrast to prior reviews which assessed the quality of psychometric properties
16 using COSMIN guidelines,^{11,12} we assessed the risk of bias and quality of the underlying studies
17 themselves. We observed a common lack of reporting and transparency in training of assessors
18 and data management. Risk of bias could not be assessed for many included studies or was
19 considered high. Better reporting standards and guidelines for psychometric studies can help
20 strengthen the field and ensure more critical evaluation of the evidence is possible.
21
22
23
24
25
26
27

28
29 283 Nevertheless, some ECD tools had multiple forms of psychometric properties assessed.
30 Using IDELA as an example: one article examined three types of reliability and one type of
31 validity, drawing on pilots from 12 countries,²⁷ followed by another article examining
32 measurement invariance across countries.³⁰ Subsequent articles have examined additional
33 psychometric properties, albeit in individual countries.^{31,32} Although such examples of
34 consecutively examining psychometric properties should be the norm, they are often financially
35 and logically infeasible and the timing does not always align with programmatic and policy
36 agendas. Likewise, ample psychometric evidence was available for ECD tools that have been
37 implemented for longer duration (e.g., ASQ and BSID), or have had more available funding (e.g.,
38 GSED). As a result, the quantity of psychometric evidence available should not be the criterion
39 used to determine the psychometric quality or usability of an ECD tool. When prior psychometric
40 evidence is unavailable and psychometric studies not possible, statistical analysis should be
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 295 conducted to verify psychometric properties and at a minimum should include reporting internal
4 consistency reliability and structural validity (where relevant). This can help build evidence across
5 multiple settings and populations and confirm the usefulness of ECD tools with diverse
6 populations. In addition, since our review focused on whether psychometric evidence exists, our
7 findings on the availability of psychometric evidence do not inform our understanding of the
8 underlying quality, strength, or rigor of the psychometric evidence. An important next step in this
9 line of work is to fully unpack the utility of existing psychometric evidence. The results of
10 psychometric analyses along with other characteristics of an ECD tool (e.g., domain assessed, age
11 range, and administration time and cost among others) should be used to determine the most
12 relevant ECD tool for the given context.
13
14

15 305 Several strengths of our review should be noted. We used extensive search filters, made no
16 restrictions on domains, included both single- and multi-domain tools, and had a rigorously trained
17 team with diverse backgrounds in ECD measurement in LMICs. Nevertheless, we acknowledge
18 some limitations. We excluded studies published prior to 2007, the search start year when the
19 inaugural *Lancet* Series on Child Development in Developing Countries was published which
20 fundamentally changed the breadth and scale of ECD research in LMICs. Therefore, the search
21 period captured the most important years for the evolution of psychometric research on ECD tools
22 in LMICs. Further, we did not use a validated risk of bias tool, which may limit the comparability
23 of our findings. Lastly, although we applied no language restrictions, we were limited to the
24 languages spoken by the review team. We were unable to conduct full text review and extraction
25 for articles in Chinese, Farsi, and Turkish. Consequently, results from China, Iran, and Turkey
26 may be underrepresented.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

318 CONCLUSION

319 Psychometric evidence on ECD tools used in LMICs is fragmented, limited, and
320 heterogeneous. More research is warranted to establish the applicability of existing tools in diverse
321 populations, including urban and rural settings, and on establishing measurement invariance over
322 countries. Nevertheless, the results by country and ECD tool presented here can serve stakeholders
323 by providing a database of available psychometric evidence for ECD tools in LMICs. To improve
324 monitoring, evaluation, and accountability for ECD globally, psychometric evidence should be a
325 key consideration when selecting ECD tools together with other important consideration including
326 the purpose of measurement, available resources for training and administration, and the
327 population and developmental domain of interest. As psychometric properties can vary by
328 geography, population, and age, among other characteristics, greater psychometric validation can
329 help facilitate ECD tool selection across diverse contexts in LMICs. Improved reporting for
330 psychometric studies can help ensure transparency, replication, and adequate ability to assess the
331 quality of evidence.

332

333 Acknowledgements

334 **Contributors:** LB conceptualised the review, designed the methodology, screened articles,
335 extracted and analysed data, drafted and revised the manuscript. EH extracted data and analysed
336 data, drafted and revised the manuscript. NBA and AR extracted data, and revised the manuscript.
337 XH and KSC screened articles, extracted data, and revised the manuscript. SEN and JS extracted
338 data and revised the manuscript. DO revised the manuscript. HOP screened articles and revised
339 the manuscript. AT curated the data and revised the manuscript. JJ conceptualised the review,

1
2
3 340 designed the methodology, screened articles, and revised the manuscript. All authors read and
4
5 341 approved the final manuscript. LB is the guarantor.
6
7
8

9 342 **Funding:** This work was supported by the CGIAR Initiative on Resilient Cities. We would like to
10
11 343 thank all funders who supported this research through their contributions to the CGIAR Trust
12
13 344 Fund: <https://www.cgiar.org/funders>. SEN's time was supported by the National Institute of
14
15 345 Mental Health (grant number T32 MH078788).
16
17

18 346 **Competing interests:** JS has contributed and continues to contribute to the development of the
19
20 347 CREDI, GSED, IDELA and AIM-ECD. The remaining authors declare no conflicts of interest.
21
22

23 348 **Patient consent for publication:** Not applicable.
24
25

26 349 **Ethical considerations:** Not applicable.
27
28

29 350 **Provenance and peer review:** Not commissioned; externally peer reviewed.
30
31

32 351 **Data availability statement:** All data are included in the manuscript and appendix.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

1. United Nations. Transforming our world: the 2030 Agenda for Sustainable Development. Published online 2015. <https://sdgs.un.org/sites/default/files/publications/21252030Agenda%20for%20Sustainable%20Development%20web.pdf>
2. Walker SP, Chang SM, Vera-Hernandez M, Grantham-McGregor S. Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behavior. *Pediatrics*. 2011;127(5):849-857. doi:10.1542/peds.2010-2231
3. Gertler P, Heckman J, Pinto R, et al. Labor market returns to an early childhood stimulation intervention in Jamaica. *Science* (80-). 2014;344(6187):998-1001. doi:10.1126/science.1251178
4. Campbell F, Conti G, Heckman JJ, et al. Early Childhood Investments Substantially Boost Adult Health. *Science* (80-). 2014;343(6178):1478-1485. doi:10.1126/science.1248429
5. Fernald LCH, Prado EL, Kariger PK, Raikes A. A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries. Published online 2017.
6. Pushparatnam A, Seiden JM, Luna Bazaldua DA. Guiding Questions for Choosing the Right Tools to Measure Early Childhood Outcomes: Why, What, Who, and How. Published online 2022. <https://openknowledge.worldbank.org/handle/10986/37030>
7. American Psychological Association. APA Dictionary of Psychology. Accessed February 26, 2023. <https://dictionary.apa.org/>
8. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Dev Rev*. 2016;41:71-90. doi:10.1016/j.dr.2016.06.004
9. Semrud-Clikeman M, Romero RAA, Prado EL, Shapiro EG, Bangirana P, John CC. Selecting measures for the neurodevelopmental assessment of children in low- and middle-income countries. *Child Neuropsychol*. Published online September 9, 2016;1-42. doi:10.1080/09297049.2016.1216536
10. Munoz-Chereau B, Ang L, Dockrell J, Outhwaite L, Heffernan C. Measuring early child development across low and middle-income countries: A systematic review. *J Early Child Res*. 2021;19(4):443-470. doi:10.1177/1476718X211020031
11. Boggs D, Milner KM, Chandna J, et al. Rating early child development outcome measurement tools for routine health programme use. *Arch Dis Child*. 2019;104(Suppl 1):S22-S33. doi:10.1136/archdischild-2018-315431
12. Griffiths A, Toovey R, Morgan PE, Spittle AJ. Psychometric properties of gross motor assessment tools for children: a systematic review. *BMJ Open*. 2018;8(10):e021734. doi:10.1136/bmjopen-2018-021734
13. Halle TG, Darling-Churchill KE. Review of measures of social and emotional development. *J Appl Dev Psychol*. 2016;45:8-18. doi:10.1016/j.appdev.2016.02.003
14. Humphrey N, Kalambouka A, Wigelsworth M, Lendrum A, Deighton J, Wolpert M. Measures of Social and Emotional Skills for Children and Young People. *Educ Psychol Meas*. 2011;71(4):617-637. doi:10.1177/0013164410382896
15. Lu C, Cuartas J, Fink G, et al. Inequalities in early childhood care and development in low/middle-income countries: 2010–2018. *BMJ Glob Heal*. 2020;5(2):e002314. doi:10.1136/bmjjgh-2020-002314
16. McCoy DC, Peet ED, Ezzati M, et al. Early Childhood Developmental Status in Low- and Middle-Income Countries: National, Regional, and Global Prevalence Estimates Using Predictive Modeling. *PLoS Med*. 2016;13(6):1-18. doi:10.1371/journal.pmed.1002034

17. Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res*. 2009;18(8):1115-1123. doi:10.1007/s11136-009-9528-5

18. Sabanathan S, Wills B, Gladstone M. Child development assessment tools in low-income and middle-income countries: how can we use them more appropriately? *Arch Dis Child*. 2015;100(5):482-488. doi:10.1136/archdischild-2014-308114

19. Gridley N, Blower S, Dunn A, Bywater T, Bryant M. Psychometric Properties of Child (0–5 Years) Outcome Measures as used in Randomized Controlled Trials of Parent Programs: A Systematic Review. *Clin Child Fam Psychol Rev*. 2019;22(3):388-405. doi:10.1007/s10567-019-00277-1

20. Crocker LM, Algina J. *Introduction to Classical and Modern Test Theory*. First Edit. Holt, Rinehart, and Winston; 1986.

21. Hui CH, Triandis HC. Measurement in Cross-Cultural Psychology. *J Cross Cult Psychol*. 1985;16(2):131-152. doi:10.1177/0022002185016002001

22. Peña ED. Lost in Translation: Methodological Considerations in Cross-Cultural Research. *Child Dev*. 2007;78(4):1255-1264. doi:10.1111/j.1467-8624.2007.01064.x

23. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs). Published online 2018:1-78. https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf

24. ROBINS-E Development Group. Risk Of Bias In Non-randomized Studies - of Exposure (ROBINS-E). Launch version, 1 June 2022. Published online 2022. <https://www.riskofbias.info/welcome/robins-e-tool>

25. GRADE Working Group. *Handbook for Grading the Quality of Evidence and the Strength of Recommendations Using the GRADE Approach. Updated October 2013*. (Schünemann H, Brożek J, Guyatt G, Oxman A, eds.); 2013. <https://gdt.gradepro.org/app/handbook/handbook.html>

26. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097

27. Pisani L, Borisova I, Dowd AJ. Developing and validating the International Development and Early Learning Assessment (IDELA). *Int J Educ Res*. 2018;91(July):1-15. doi:10.1016/j.ijer.2018.06.007

28. Kersten P, Czuba K, McPherson K, et al. A systematic review of evidence for the psychometric properties of the Strengths and Difficulties Questionnaire. *Int J Behav Dev*. 2016;40(1):64-75. doi:10.1177/0165025415570647

29. Velikonja T, Edbrooke-Childs J, Calderon A, Sleed M, Brown A, Deighton J. The psychometric properties of the Ages & Stages Questionnaires for ages 2-2.5: a systematic review. *Child Care Health Dev*. 2017;43(1):1-17. doi:10.1111/cch.12397

30. Halpin PF, Wolf S, Yoshikawa H, et al. Measuring early learning and development across cultures: Invariance of the IDELA across five countries. *Dev Psychol*. 2019;55(1):23-37. doi:10.1037/dev0000626

31. Wolf S, Halpin P, Yoshikawa H, Dowd AJ, Pisani L, Borisova I. Measuring school readiness globally: Assessing the construct validity and measurement invariance of the International Development and Early Learning Assessment (IDELA) in Ethiopia. *Early Child Res Q*. 2017;41(June):21-36. doi:10.1016/j.ecresq.2017.05.001

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

- 1
2
3 32. Pisani L, Seiden J, Wolf S. Longitudinal evidence on the predictive validity of the
4 International Development and Early Learning Assessment (IDELA). *Educ Assessment,
5 Eval Account*. 2022;34(2):173-194. doi:10.1007/s11092-021-09374-8
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
- For peer review only

Table 1 Definitions of the psychometric properties used in the selection criteria for included articles

Domain	Measurement property	Example test statistic	Definition
Reliability			The consistency of a test or measurement, i.e., how consistently a measure produces similar results with repeated measures over a short period of time or across assessors at the same time point. This can also be thought of as the correlation between observed scores across replications.
	Test-retest reliability	correlation coefficient	Correlation between scores from the same test from assessments conducted over a short time interval.
	Inter-rater reliability	kappa, Bland-Altman test	The extent to which independent assessors produce similar ratings in judging the same abilities or characteristics in the same target person at the same time.
	Internal consistency reliability	Cronbach's alpha, alpha	Degree of interrelatedness among items on the same tool, i.e., how well the items work together to provide information on an underlying construct.
Validity			The degree to which the tool measures what it is supposed to measure, i.e., the degree to which the tool reflects the underlying construct.
	Content/face validity		The degree to which the content of the tool is adequate for the construct being measured, i.e., assessing the extent to which a tool appears to reflect the underlying construct.
	Concurrent/criterion validity	correlation coefficient; regression estimate	The degree to which scores on one measurement tool are related to scores obtained at about the same point in time from another tool considered the gold standard.
	Convergent validity	correlation coefficient; regression estimate	Evidence that scores on a test or measurement are associated with theoretically related measures or variables.

136/bmjopen-2024-096365 on 11 MAY 2025. Downloaded from <http://bmjopen.bmj.com/> on June 13, 2025 at Agence Bibliographique de l'enseignement Supérieur (ABES). For personal research only. Not for commercial use. For more information, see terms and conditions (https://www.bmjjournals.org/page/info/about/policies/terms.html) and similar technologies.

Domain	Measurement property	Example test statistic	Definition
	Predictive	correlation coefficient; regression estimate	Evidence that a score correlates with a variable that can only be assessed at some point after the test has been administered or the measurement made, e.g., evidence that scores now are correlated with scores at a future point.
	Structural validity (dimensionality)	Exploratory factor analysis: number of factors, eigen values Confirmatory factor analysis: model fit statistics such as Comparative Fit Index (CF), Root mean square error of approximation (RMSEA)	The degree to which the scores of assessment are an adequate reflection of the dimensionality of the construct to be measured.
Invariance			The property when a scale or construct provides the same results across different samples, populations, settings, or characteristics.
	Measurement invariance over countries	Likelihood ratio chi squared statistic and p-value from freeing parameters across groups	The degree to which an assessment of construct provides the same results across separate samples in different countries.
	Measurement invariance over other groups	Likelihood ratio chi squared statistic and p-value from freeing parameters across groups	The degree to which an assessment of construct provides the same results across different groups.

Note: All definitions based on the APA Dictionary of Psychology.⁷

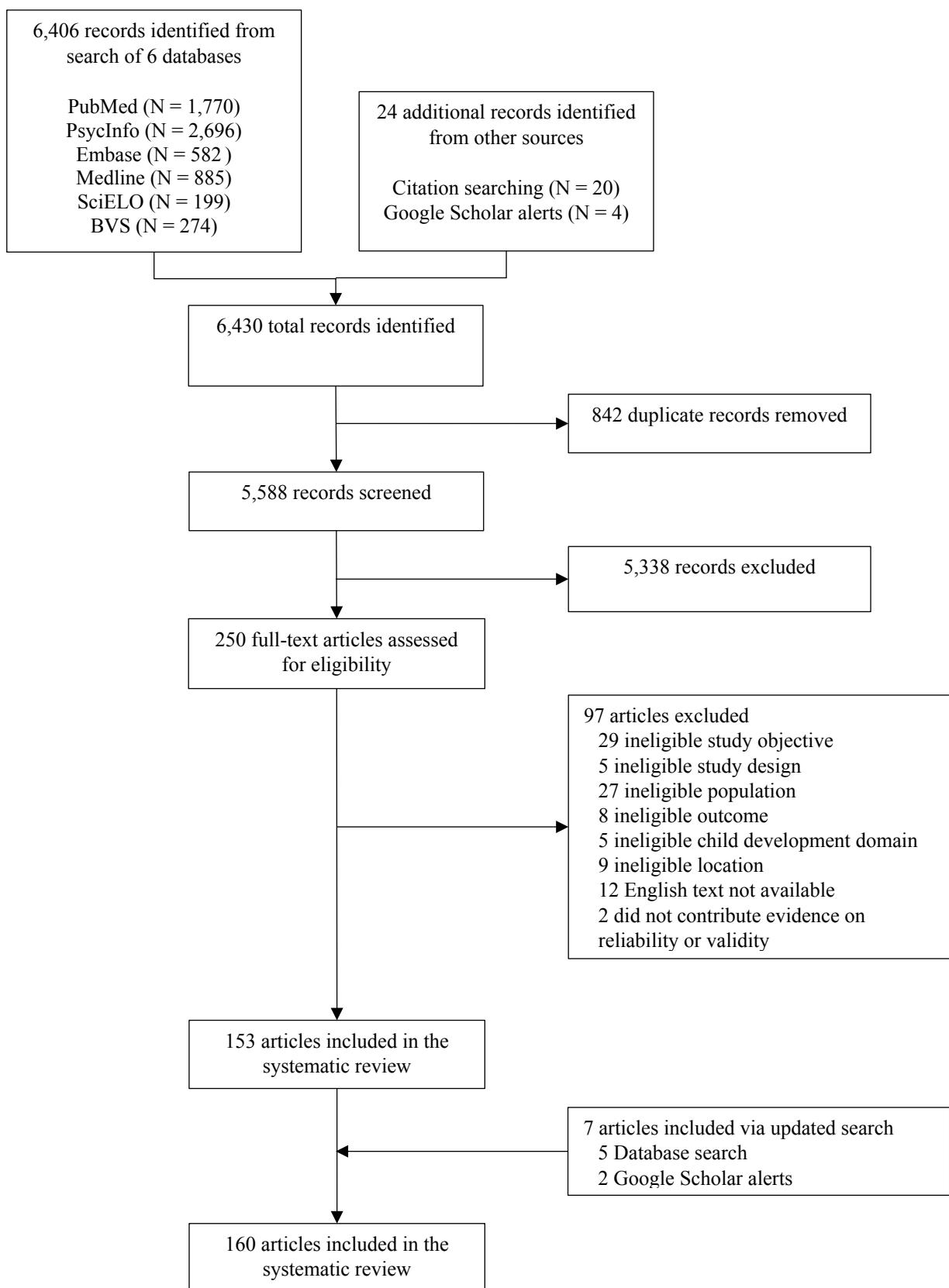
Figure Captions

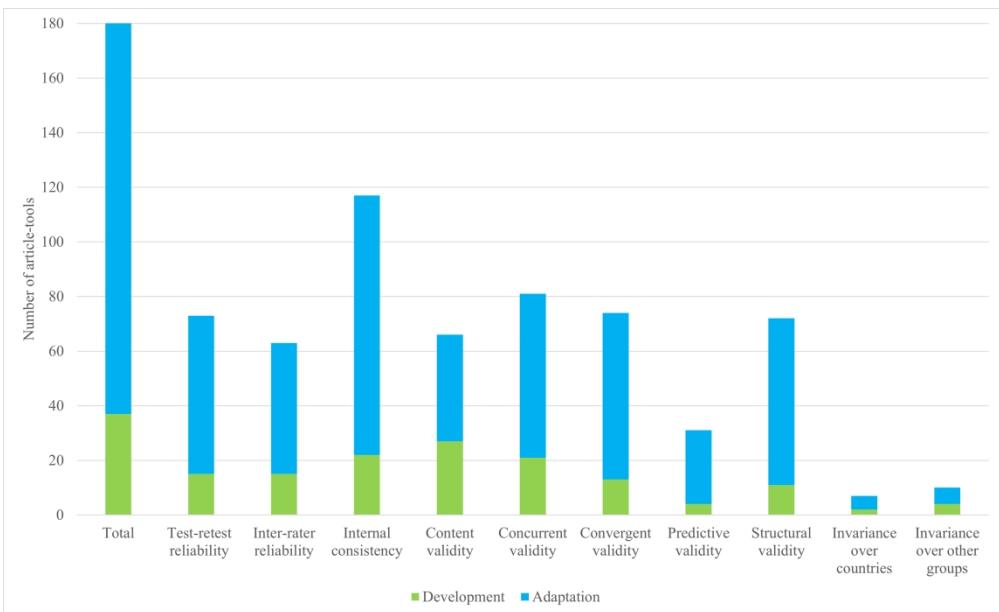
Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram of search results and included articles

Figure 2 Number of included article-tools and type of psychometric evidence provided

Figure 3 Countries where studies providing evidence on at least one psychometric property were conducted

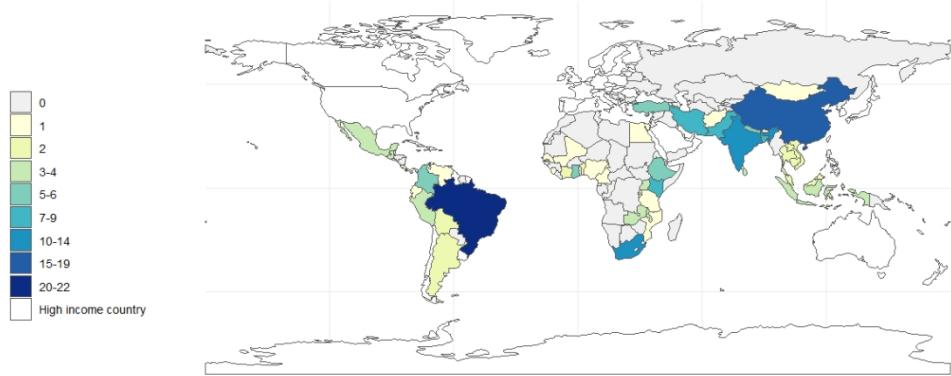
Figure 4 Number of articles providing psychometric evidence by article type and early childhood development domain





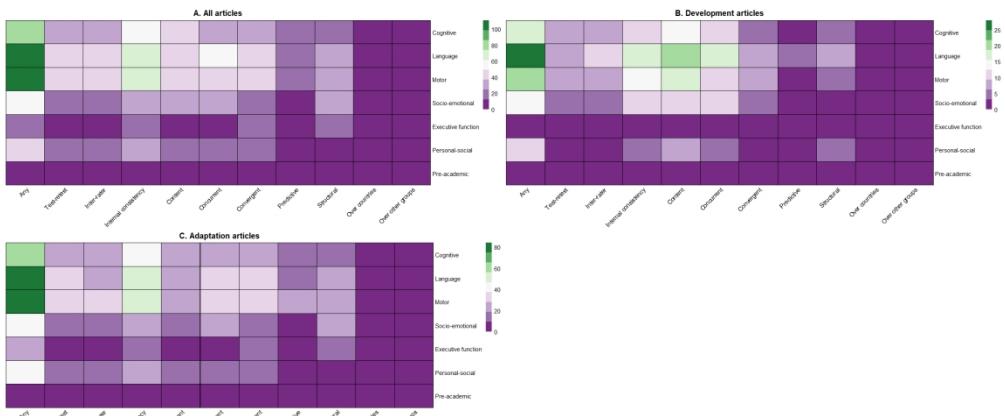
Number of included article-tools and type of psychometric evidence provided

276x168mm (300 x 300 DPI)



21 Countries where studies providing evidence on at least one psychometric property were conducted
22

23 265x116mm (115 x 115 DPI)
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Number of articles providing psychometric evidence by article type and early childhood development domain

593x264mm (108 x 108 DPI)

Supplemental appendix

Table of Contents

Supplemental Table 1 Search strategy as executed in PubMed	2
Supplemental Table 2 Definitions of the seven early childhood domains of interest	15
Supplemental Table 3 Number of included articles by type of psychometric evidence provided	16
Supplemental Table 4 List of countries and assessment tools used in included articles.....	17
Supplemental Table 5 List of child development assessment tools and countries where they were applied.....	38
Supplemental Table 6 Number of articles providing psychometric evidence by world region and urban/rural location.....	45
Supplemental Table 7 Number of articles providing psychometric evidence by child development assessment tools (development articles).....	46
Supplemental Table 8 Number of articles providing psychometric evidence by child development assessment tools (adaptation articles).....	49
Supplemental Table 9 Number of articles providing psychometric evidence by article type and early childhood development domain.....	56
Supplemental Table 10 Number of articles by category of risk of bias	57
Supplemental Table 11 Risk of bias assessment by included article	58
Supplemental Figure 1 Number of articles providing evidence on psychometrics by type of article	63
Supplemental Figure 2 Number of articles and tools providing psychometric evidence by child age group	64
Supplemental Figure 3 Number of articles providing evidence on psychometrics by child age at assessment.....	64
References of all included articles	66

1
2
3
4
5
6 **Supplemental Table 1** Search strategy as executed in each database

PubMed
(“Child”[MeSH] OR “Child, Preschool”[MeSH] OR “child*”[tiab] OR “Infant”[MeSH] OR “infant*”[tiab] OR “paediatric”[tiab] OR “pediatric”[tiab] OR “toddler*”[tiab])
AND
(“Child Development”[MeSH] OR “child development*”[tiab] OR “early child development”[tiab] OR “early childhood development”[tiab] OR “ECD”[tiab] OR “Child Behavior”[MeSH] OR “child behavior”[tiab] OR “child behaviour”[tiab] OR “Cognition”[MeSH] OR “Executive Function”[MeSH] OR “executive function*”[tiab] OR “Emotional Intelligence”[MeSH] OR “attachment”[tiab] OR “attention”[tiab] OR “cognition”[tiab] OR “cognitive development”[tiab] OR “cognitive skills”[tiab] OR “neurocognitive”[tiab] OR “communication”[tiab] OR “language”[tiab] OR “language skills”[tiab] OR “language development”[tiab] OR “receptive language”[tiab] OR “expressive language”[tiab] OR “speech development”[tiab] OR “communication skills”[tiab] OR “Motor Skills”[MeSH] OR “motor skills”[tiab] OR “fine motor”[tiab] OR “gross motor”[tiab] OR “motor”[tiab] OR “prosocial”[tiab] OR “socialisation”[tiab] OR “socialization”[tiab] OR “socioemotion*”[tiab] OR “socio-emotion*”[tiab] OR “neurodevelopment”[tiab] OR “neurobehavior”[tiab] OR “neurobehaviour”[tiab] OR “neurobehavioral”[tiab] OR “neurobehavioural”[tiab] OR “Intelligence”[MeSH] OR “intelligence”[tiab] OR “IQ”[tiab] OR “psychomotor”[tiab] OR “sensorimotor”[tiab] OR “attention”[tiab] OR “inhibition”[tiab] OR “impulse control”[tiab] OR “working memory”[tiab] OR “problem solving”[tiab] OR “self-regulation”[tiab] OR “prosocial”[tiab] OR “psychosocial development”[tiab] OR “internali*”[tiab] OR “externali*”[tiab] OR “empathy”[tiab] OR “neuropsychology*”[tiab] OR “school readiness”[tiab] or “numeracy”[tiab] or “literacy”[tiab] OR “pre-academic”[tiab] OR “early learning”[tiab] or “personal-social”[tiab] OR “personal social”[tiab])
AND
((“developing countr*”[tiab] OR “under developed countr*”[tiab] OR “Imic*”[tiab] OR “less developed”[tiab] OR “low income”[tiab] OR “lower income”[tiab] OR “low- and middle-income”[tiab] OR “low middle income”[tiab] OR “resource poor”[tiab] OR “resource constrained”[tiab] OR “low resource”[tiab] OR “limited resource*”[tiab] OR “resource limited”[tiab]) AND (Afghan*[tiab] OR Africa[tiab] OR African[tiab] OR Algeria*[tiab] OR “American Samoa*”[tiab] OR Angola*[tiab] OR Argentin*[tiab] OR Bangladesh*[tiab] OR Barbad*[tiab] OR Belorussian[tiab] OR Beliz*[tiab] OR Benin*[tiab] OR Bhutan*[tiab] OR Bolivia*[tiab] OR Botswan*[tiab] OR Brazil*[tiab] OR “Burkina Faso”[tiab] OR Burkinabe[tiab] OR Burund*[tiab] OR Cambodia*[tiab] OR Cameroon*[tiab] OR “Cape Verde”[tiab] OR “Cape Verdean”[tiab] OR “Central African Republic”[tiab] OR Chad*[tiab] OR Chile*[tiab] OR China[tiab] OR Chinese[tiab] OR Colombia*[tiab] OR Comoros[tiab] OR Comorian[tiab] OR Congo[tiab] OR Congolese[tiab] OR Costa Rica*[tiab] OR “Côte

Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

1
2
3 d'Ivoire"[tiab] OR "Ivory Coast"[tiab] OR Ivorian[tiab] OR Croatia*[tiab] OR Croat[tiab] OR
4 Cuba*[tiab] OR Djibouti*[tiab] OR Dominica*[tiab] OR "Dominican Republic"[tiab] OR
5 Ecuador*[tiab] OR Egypt*[tiab] OR "El Salvador"[tiab] OR Salvadorian[tiab] OR "Equatorial
6 Guinea"[tiab] OR Guinean[tiab] OR Eritrea*[tiab] OR Ethiopia*[tiab] OR Fiji*[tiab] OR
7 Gabon*[tiab] OR Gambia*[tiab] OR Gaza[tiab] OR Gazan[tiab] OR Ghana[tiab] OR
8 Ghanaian[tiab] OR Grenad*[tiab] OR Guatemala*[tiab] OR Guinea[tiab] OR Guyan*[tiab] OR
9 Haiti*[tiab] OR Hondu[tiab] OR Hungar*[tiab] OR India[tiab] OR Indian[tiab] OR
10 Indonesia*[tiab] OR Iran*[tiab] OR Iraq*[tiab] OR Jamaica*[tiab] OR Jordan*[tiab] OR
11 Kenya[tiab] OR Kenyan[tiab] OR Kiribati[tiab] OR Korea*[tiab] OR Kyrgy*[tiab] OR
12 Laos[tiab] OR Laotian*[tiab] OR Lebanon[tiab] OR Lebanese[tiab] OR Lesotho[tiab] OR
13 Liberia*[tiab] OR Libya*[tiab] OR Macedonia*[tiab] OR Madagasca*[tiab] OR Malawi*[tiab]
14 OR Malaysia*[tiab] OR Maldives[tiab] OR Maldivian[tiab] OR Mali[tiab] OR Malian*[tiab]
15 OR "Marshall Islands"[tiab] OR Mauritania*[tiab] OR Mauritius[tiab] OR Mauritian[tiab] OR
16 Mayotte[tiab] OR Mexic*[tiab] OR Micronesia*[tiab] OR Moldov*[tiab] OR Mongolia*[tiab]
17 OR Morocco*[tiab] OR Mozambique[tiab] OR Mozambican[tiab] OR Myanmar[tiab] OR
18 Namibia*[tiab] OR Nepal*[tiab] OR Nevis[tiab] OR Nicaragua*[tiab] OR Niger*[tiab] OR
19 "Northern Mariana Islands"[tiab] OR Oman*[tiab] OR Pakistan*[tiab] OR Palau*[tiab] OR
20 Panama*[tiab] OR "Papua New Guinea"[tiab] OR Paraguay*[tiab] OR Peru*[tiab] OR
21 Philippine*[tiab] OR Filipino*[tiab] OR Poland[tiab] OR Polish[tiab] OR Rwanda*[tiab] OR
22 Samoa*[tiab] OR Sao Tome*[tiab] OR Principe[tiab] OR Senegal*[tiab] OR Seychell*[tiab]
23 OR Sierra Leon*[tiab] OR Solomon Island*[tiab] OR Somali*[tiab] OR South Africa*[tiab] OR
24 Sri Lanka*[tiab] OR "Saint Kitts"[tiab] OR "St Kitts"[tiab] OR "Saint Lucia"[tiab] OR "St
25 Lucia"[tiab] OR "Saint Vincent" [tiab] OR "St Vincent"[tiab] OR Sudan*[tiab] OR
26 Suriname*[tiab] OR Swaziland Or Swazi[tiab] OR Syria*[tiab] OR Tajik*[tiab] OR
27 Tanzania*[tiab] OR Thailand[tiab] OR Thai[tiab] OR "Timor Leste"[tiab] OR Togo*[tiab] OR
28 Tonga*[tiab] OR Trinidad[tiab] OR Trinidadian[tiab] OR Tobago[tiab] OR Tobagonian[tiab]
29 OR Tunisia*[tiab] OR Turky[tiab] OR Turkish[tiab] OR Uganda*[tiab] OR Uruguay*[tiab] OR
30 Vanuat*[tiab] OR Venezuela*[tiab] OR Vietnam*[tiab] OR "West Bank"[tiab] OR
31 Yemen*[tiab] OR Zambia*[tiab] OR Zimbabwe*))

32 AND

33 (instrumentation[sh] OR methods[sh] OR "Validation Studies"[pt] OR "Comparative
34 Study"[pt] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR
35 clinometr*[tw] OR "outcome assessment (health care)"[MeSH] OR "outcome assessment"[tiab]
36 OR "outcome measure*"[tw] OR "observer variation"[MeSH] OR "observer variation"[tiab]
37 OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR
38 reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR
39 valid*[tiab] OR "coefficient of variation"[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR
40 homogeneous[tiab] OR "internal consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR
41 alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR
42 reduction*[tiab]))) OR agreement[tw] OR precision[tw] OR imprecision[tw] OR "precise
43 values"[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND

(test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw])) AND (measure[tw] OR measures[tw] OR findings[tw] OR result[tw] OR results[tw] OR test[tw] OR tests[tw])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known group"[tiab] OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR "factor structure"[tiab] OR "factor structures"[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR "item discriminant"[tiab] OR "interscale correlation*"[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab])) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR (small*[tiab] AND (real[tiab] OR detectable[tiab])) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab] OR "measurement invariance"[tiab] OR "invariance"[tiab] OR assessment[tiab] OR adapt*[tiab])

NOT

("editorial"[Publication Type] OR "comment"[Publication Type] OR "case reports"[Publication Type] OR "congress"[Publication Type] OR "consensus development conference"[Publication Type] OR "interview"[Publication Type] OR "lecture"[Publication Type] OR "legal case"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "Observational Study, Veterinary"[Publication Type] OR "Patient Education Handout"[Publication Type] OR "practice guideline"[Publication Type] OR "preprint"[Publication Type])

MEDLINE

(Child OR Child, Preschool OR child OR Infant OR infant OR paediatric OR pediatric OR toddler)

AND

(Child Development OR child development OR early child development OR early childhood development OR ECD OR Child Behavior OR child behavior OR child behaviour OR Cognition OR Executive Function OR executive function OR Emotional Intelligence OR attachment OR attention OR cognition OR cognitive development OR cognitive skills OR neurocognitive OR communication OR language OR language skills OR language development OR receptive language OR expressive language OR speech development OR communication skills OR Motor Skills OR motor skills OR fine motor OR gross motor OR prosocial OR socialisation OR socialization OR socioemotion OR socio-emotion OR neurodevelopment OR neurobehavior OR neurobehaviour OR neurobehavioral OR neurobehavioural OR Intelligence OR intelligence OR IQ OR psychomotor OR sensorimotor OR attention OR inhibition OR impulse control OR working memory OR problem solving OR self-regulation OR prosocial OR psychosocial development OR internali OR externali OR empathy)

AND

((developing countr OR under developed countr OR lmic OR less developed OR low income OR lower income OR low- and middle-income OR low middle income OR resource poor OR resource constrained OR low resource OR limited resource OR resource limited) AND (Afghan OR Africa OR African OR Algeria OR American Samoa OR Angola OR Argentin OR Bangladesh OR Barbad OR Belorussian OR Beliz OR Benin OR Bhutan OR Bolivia OR Botswan OR Brazil OR Burkina Faso OR Burkinabe OR Burund OR Cambodia OR Cameroon OR Cape Verde OR Cape Verdean OR Central African Republic OR Chad OR Chile OR China OR Chinese OR Colombia OR Comoros OR Comorian OR Congo OR Congolese OR Costa Rica OR Cote d Ivoire OR Ivory Coast OR Ivorian OR Croatia OR Croat OR Cuba OR Djibouti OR Dominica OR Dominican Republic OR Ecuador OR Egypt OR El Salvador OR Salvadorian OR Equatorial Guinea OR Guinean OR Eritrea OR Ethiopia OR Fiji OR Gabon OR Gambia OR Gaza OR Gazan OR Ghana OR Ghanaian OR Grenad OR Guatemala OR Guinea OR Guyan OR Haiti OR Hondura OR Hungar OR India OR Indian OR Indonesia OR Iran OR Iraq OR Jamaica OR Jordan OR Kenya OR Kenyan OR Kiribati OR Korea OR Kyrgy OR Laos OR Laotian OR Lebanon OR Lebanese OR Lesotho OR Liberia OR Libya OR Macedonia OR Madagasca OR Malawi OR Malaysia OR Maldives OR Maldivian OR Mali OR Malian OR Marshall Islands OR Mauritania OR Mauritius OR Mauritian OR Mayotte OR Mexic OR Micronesia OR Moldov OR Mongolia OR Morocc OR Mozambique OR Mozambican OR Myanmar OR Namibia OR Nepal OR Nevis OR Nicaragua OR Niger OR Northern Mariana Islands OR Oman OR Pakistan OR Palau OR Panama OR Papua New Guinea OR Paraguay OR Peru OR Philippine OR Filipino OR Poland OR Polish OR Rwanda OR Samoa OR Sao Tome OR Principe OR Senegal OR Seychell OR Sierra Leon OR Solomon Island OR Somali OR

1
2
3 South Africa OR Sri Lanka OR Saint Kitts OR St Kitts OR Saint Lucia OR St Lucia OR Saint
4 Vincent OR St Vincent OR Sudan OR Suriname OR Swaziland OR Swazi OR Syria OR Tajik
5 OR Tanzania OR Thailand OR Thai OR Timor Leste OR Togo OR Tonga OR Trinidad OR
6 Trinidadian OR Tobago OR Tobagonian OR Tunisia OR Turky OR Turkish OR Uganda OR
7 Uruguay OR Vanuat OR Venezuela OR Vietnam OR West Bank OR Yemen OR Zambia OR
8 Zimbabwe))
9
10
11 AND
12
13
14 (instrumentation OR methods OR Validation Studies OR Comparative Study OR psychometrics
15 OR psychometr OR clinimetr OR clinometr OR outcome assessment OR outcome measure OR
16 observer variation OR observer variation OR Health Status Indicators OR reproducibility of
17 results OR reproducib OR discriminant analysis OR reliab OR unreliab OR valid OR coefficient
18 of variation OR coefficient OR homogeneity OR homogeneous OR internal consistency OR
19 (cronbach AND (alpha OR alphas)) OR (item AND (correlation OR selection OR reduction))
20 OR agreement OR precision OR imprecision OR precise values OR test-retest OR (test AND
21 retest) OR (reliab AND (test OR retest)) OR stability OR interrater OR inter-rater OR intrarater
22 OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver
23 OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician
24 OR inratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer
25 OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR
26 interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant
27 OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappas OR repeatab
28 OR ((replicab OR repeated) AND (measure OR measures OR findings OR result OR results OR
29 test OR tests)) OR generaliza OR generalisa OR concordance OR (intraclass AND correlation)
30 OR discriminative OR known group OR factor analysis OR factor analyses OR factor structure
31 OR factor structures OR dimension OR subscale OR (multitrait AND scaling AND (analysis
32 OR analyses)) OR item discriminant OR interscale correlation OR error OR errors OR
33 individual variability OR interval variability OR rate variability OR (variability AND (analysis
34 OR values)) OR (uncertainty AND (measurement OR measuring)) OR standard error of
35 measurement OR sensitiv OR responsive OR (limit AND detection) OR minimal detectable
36 concentration OR interpretab OR ((minimal OR minimally OR clinical OR clinically) AND
37 (important OR significant OR detectable) AND (change OR difference)) OR (small AND (real
38 OR detectable) AND (change OR difference)) OR meaningful change OR ceiling effect OR
39 floor effect OR Item response model OR IRT OR Rasch OR Differential item functioning OR
40 DIF OR computer adaptive testing OR item bank OR cross-cultural equivalence)
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Embase

(Child OR Child, Preschool OR child OR Infant OR infant OR paediatric OR pediatric OR toddler)

AND

(Child Development OR child development OR early child development OR early childhood development OR ECD OR Child Behavior OR child behavior OR child behaviour OR Cognition OR Executive Function OR executive function OR Emotional Intelligence OR attachment OR attention OR cognition OR cognitive development OR cognitive skills OR neurocognitive OR communication OR language OR language skills OR language development OR receptive language OR expressive language OR speech development OR communication skills OR Motor Skills OR motor skills OR fine motor OR gross motor OR prosocial OR socialisation OR socialization OR socioemotion OR socio-emotion OR neurodevelopment OR neurobehavior OR neurobehaviour OR neurobehavioral OR neurobehavioural OR Intelligence OR intelligence OR IQ OR psychomotor OR sensorimotor OR attention OR inhibition OR impulse control OR working memory OR problem solving OR self-regulation OR prosocial OR psychosocial development OR internali OR externali OR empathy)

AND

((developing countr OR under developed countr OR lmic OR less developed OR low income OR lower income OR low- and middle-income OR low middle income OR resource poor OR resource constrained OR low resource OR limited resource OR resource limited) AND (Afghan OR Africa OR African OR Algeria OR American Samoa OR Angola OR Argentin OR Bangladesh OR Barbad OR Belorussian OR Beliz OR Benin OR Bhutan OR Bolivia OR Botswan OR Brazil OR Burkina Faso OR Burkinabe OR Burund OR Cambodia OR Cameroon OR Cape Verde OR Cape Verdean OR Central African Republic OR Chad OR Chile OR China OR Chinese OR Colombia OR Comoros OR Comorian OR Congo OR Congolese OR Costa Rica OR Cote d Ivoire OR Ivory Coast OR Ivorian OR Croatia OR Croat OR Cuba OR Djibouti OR Dominica OR Dominican Republic OR Ecuador OR Egypt OR El Salvador OR Salvadorian OR Equatorial Guinea OR Guinean OR Eritrea OR Ethiopia OR Fiji OR Gabon OR Gambia OR Gaza OR Gazan OR Ghana OR Ghanaian OR Grenad OR Guatemala OR Guinea OR Guyan OR Haiti OR Hondura OR Hungar OR India OR Indian OR Indonesia OR Iran OR Iraq OR Jamaica OR Jordan OR Kenya OR Kenyan OR Kiribati OR Korea OR Kyrgy OR Laos OR Laotian OR Lebanon OR Lebanese OR Lesotho OR Liberia OR Libya OR Macedonia OR Madagasca OR Malawi OR Malaysia OR Maldives OR Maldivian OR Mali OR Malian OR Marshall Islands OR Mauritania OR Mauritius OR Mauritian OR Mayotte OR Mexic OR Micronesia OR Moldov OR Mongolia OR Morocc OR Mozambique OR Mozambican OR Myanmar OR Namibia OR Nepal OR Nevis OR Nicaragua OR Niger OR Northern Mariana Islands OR Oman OR Pakistan OR Palau OR Panama OR Papua New Guinea OR Paraguay OR Peru OR Philippine OR Filipino OR Poland OR Polish OR Rwanda OR Samoa OR Sao Tome OR Principe OR Senegal OR Seychell OR Sierra Leon OR Solomon Island OR Somali OR

1
2
3 South Africa OR Sri Lanka OR Saint Kitts OR St Kitts OR Saint Lucia OR St Lucia OR Saint
4 Vincent OR St Vincent OR Sudan OR Suriname OR Swaziland OR Swazi OR Syria OR Tajik
5 OR Tanzania OR Thailand OR Thai OR Timor Leste OR Togo OR Tonga OR Trinidad OR
6 Trinidadian OR Tobago OR Tobagonian OR Tunisia OR Turky OR Turkish OR Uganda OR
7 Uruguay OR Vanuat OR Venezuela OR Vietnam OR West Bank OR Yemen OR Zambia OR
8 Zimbabwe))
9
10
11 AND
12
13
14 (instrumentation OR methods OR Validation Studies OR Comparative Study OR psychometrics
15 OR psychometr OR clinimetr OR clinometr OR outcome assessment OR outcome measure OR
16 observer variation OR observer variation OR Health Status Indicators OR reproducibility of
17 results OR reproducib OR discriminant analysis OR reliab OR unreliab OR valid OR coefficient
18 of variation OR coefficient OR homogeneity OR homogeneous OR internal consistency OR
19 (cronbach AND (alpha OR alphas)) OR (item AND (correlation OR selection OR reduction))
20 OR agreement OR precision OR imprecision OR precise values OR test-retest OR (test AND
21 retest) OR (reliab AND (test OR retest)) OR stability OR interrater OR inter-rater OR intrarater
22 OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver
23 OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician
24 OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer
25 OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR
26 interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant
27 OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappas OR repeatab
28 OR ((replicab OR repeated) AND (measure OR measures OR findings OR result OR results OR
29 test OR tests)) OR generaliza OR generalisa OR concordance OR (intraclass AND correlation)
30 OR discriminative OR known group OR factor analysis OR factor analyses OR factor structure
31 OR factor structures OR dimension OR subscale OR (multitrait AND scaling AND (analysis
32 OR analyses)) OR item discriminant OR interscale correlation OR error OR errors OR
33 individual variability OR interval variability OR rate variability OR (variability AND (analysis
34 OR values)) OR (uncertainty AND (measurement OR measuring)) OR standard error of
35 measurement OR sensitiv OR responsive OR (limit AND detection) OR minimal detectable
36 concentration OR interpretab OR ((minimal OR minimally OR clinical OR clinically) AND
37 (important OR significant OR detectable) AND (change OR difference)) OR (small AND (real
38 OR detectable) AND (change OR difference)) OR meaningful change OR ceiling effect OR
39 floor effect OR Item response model OR IRT OR Rasch OR Differential item functioning OR
40 DIF OR computer adaptive testing OR item bank OR cross-cultural equivalence)
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

PsycInfo

(Child OR Child, Preschool OR child OR Infant OR infant OR paediatric OR pediatric OR toddler)

AND

(Child Development OR child development OR early child development OR early childhood development OR ECD OR Child Behavior OR child behavior OR child behaviour OR Cognition OR Executive Function OR executive function OR Emotional Intelligence OR attachment OR attention OR cognition OR cognitive development OR cognitive skills OR neurocognitive OR communication OR language OR language skills OR language development OR receptive language OR expressive language OR speech development OR communication skills OR Motor Skills OR motor skills OR fine motor OR gross motor OR prosocial OR socialisation OR socialization OR socioemotion OR socio-emotion OR neurodevelopment OR neurobehavior OR neurobehaviour OR neurobehavioral OR neurobehavioural OR Intelligence OR intelligence OR IQ OR psychomotor OR sensorimotor OR attention OR inhibition OR impulse control OR working memory OR problem solving OR self-regulation OR prosocial OR psychosocial development OR internali OR externali OR empathy)

AND

((developing countr OR under developed countr OR lmic OR less developed OR low income OR lower income OR low- and middle-income OR low middle income OR resource poor OR resource constrained OR low resource OR limited resource OR resource limited) AND (Afghan OR Africa OR African OR Algeria OR American Samoa OR Angola OR Argentin OR Bangladesh OR Barbad OR Belorussian OR Beliz OR Benin OR Bhutan OR Bolivia OR Botswan OR Brazil OR Burkina Faso OR Burkinabe OR Burund OR Cambodia OR Cameroon OR Cape Verde OR Cape Verdean OR Central African Republic OR Chad OR Chile OR China OR Chinese OR Colombia OR Comoros OR Comorian OR Congo OR Congolese OR Costa Rica OR Cote d Ivoire OR Ivory Coast OR Ivorian OR Croatia OR Croat OR Cuba OR Djibouti OR Dominica OR Dominican Republic OR Ecuador OR Egypt OR El Salvador OR Salvadorian OR Equatorial Guinea OR Guinean OR Eritrea OR Ethiopia OR Fiji OR Gabon OR Gambia OR Gaza OR Gazan OR Ghana OR Ghanaian OR Grenad OR Guatemala OR Guinea OR Guyan OR Haiti OR Hondura OR Hungar OR India OR Indian OR Indonesia OR Iran OR Iraq OR Jamaica OR Jordan OR Kenya OR Kenyan OR Kiribati OR Korea OR Kyrgy OR Laos OR Laotian OR Lebanon OR Lebanese OR Lesotho OR Liberia OR Libya OR Macedonia OR Madagasca OR Malawi OR Malaysia OR Maldives OR Maldivian OR Mali OR Malian OR Marshall Islands OR Mauritania OR Mauritius OR Mauritian OR Mayotte OR Mexic OR Micronesia OR Moldov OR Mongolia OR Morocc OR Mozambique OR Mozambican OR Myanmar OR Namibia OR Nepal OR Nevis OR Nicaragua OR Niger OR Northern Mariana Islands OR Oman OR Pakistan OR Palau OR Panama OR Papua New Guinea OR Paraguay OR Peru OR Philippine OR Filipino OR Poland OR Polish OR Rwanda OR Samoa OR Sao Tome OR Principe OR Senegal OR Seychell OR Sierra Leon OR Solomon Island OR Somali OR

1
2
3 South Africa OR Sri Lanka OR Saint Kitts OR St Kitts OR Saint Lucia OR St Lucia OR Saint
4 Vincent OR St Vincent OR Sudan OR Suriname OR Swaziland OR Swazi OR Syria OR Tajik
5 OR Tanzania OR Thailand OR Thai OR Timor Leste OR Togo OR Tonga OR Trinidad OR
6 Trinidadian OR Tobago OR Tobagonian OR Tunisia OR Turky OR Turkish OR Uganda OR
7 Uruguay OR Vanuat OR Venezuela OR Vietnam OR West Bank OR Yemen OR Zambia OR
8 Zimbabwe))
9
10
11 AND
12
13
14 (instrumentation OR methods OR Validation Studies OR Comparative Study OR psychometrics
15 OR psychometr OR clinimetr OR clinometr OR outcome assessment OR outcome measure OR
16 observer variation OR observer variation OR Health Status Indicators OR reproducibility of
17 results OR reproducib OR discriminant analysis OR reliab OR unreliab OR valid OR coefficient
18 of variation OR coefficient OR homogeneity OR homogeneous OR internal consistency OR
19 (cronbach AND (alpha OR alphas)) OR (item AND (correlation OR selection OR reduction))
20 OR agreement OR precision OR imprecision OR precise values OR test-retest OR (test AND
21 retest) OR (reliab AND (test OR retest)) OR stability OR interrater OR inter-rater OR intrarater
22 OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver
23 OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician
24 OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer
25 OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR
26 interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant
27 OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappas OR repeatab
28 OR ((replicab OR repeated) AND (measure OR measures OR findings OR result OR results OR
29 test OR tests)) OR generaliza OR generalisa OR concordance OR (intraclass AND correlation)
30 OR discriminative OR known group OR factor analysis OR factor analyses OR factor structure
31 OR factor structures OR dimension OR subscale OR (multitrait AND scaling AND (analysis
32 OR analyses)) OR item discriminant OR interscale correlation OR error OR errors OR
33 individual variability OR interval variability OR rate variability OR (variability AND (analysis
34 OR values)) OR (uncertainty AND (measurement OR measuring)) OR standard error of
35 measurement OR sensitiv OR responsive OR (limit AND detection) OR minimal detectable
36 concentration OR interpretab OR ((minimal OR minimally OR clinical OR clinically) AND
37 (important OR significant OR detectable) AND (change OR difference)) OR (small AND (real
38 OR detectable) AND (change OR difference)) OR meaningful change OR ceiling effect OR
39 floor effect OR Item response model OR IRT OR Rasch OR Differential item functioning OR
40 DIF OR computer adaptive testing OR item bank OR cross-cultural equivalence)
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

SciELO

(Child OR Child, Preschool OR child OR Infant OR infant OR paediatric OR pediatric OR toddler)

AND

(Child Development OR child development OR early child development OR early childhood development OR ECD OR Child Behavior OR child behavior OR child behaviour OR Cognition OR Executive Function OR executive function OR Emotional Intelligence OR attachment OR attention OR cognition OR cognitive development OR cognitive skills OR neurocognitive OR communication OR language OR language skills OR language development OR receptive language OR expressive language OR speech development OR communication skills OR Motor Skills OR motor skills OR fine motor OR gross motor OR prosocial OR socialisation OR socialization OR socioemotion OR socio-emotion OR neurodevelopment OR neurobehavior OR neurobehaviour OR neurobehavioral OR neurobehavioural OR Intelligence OR intelligence OR IQ OR psychomotor OR sensorimotor OR attention OR inhibition OR impulse control OR working memory OR problem solving OR self-regulation OR prosocial OR psychosocial development OR internali OR externali OR empathy)

AND

(instrumentation OR methods OR Validation Studies OR Comparative Study OR psychometrics OR psychometr OR clinimetr OR clinometr OR outcome assessment OR outcome measure OR observer variation OR observer variation OR Health Status Indicators OR reproducibility of results OR reproducib OR discriminant analysis OR reliab OR unreliab OR valid OR coefficient of variation OR coefficient OR homogeneity OR homogeneous OR internal consistency OR (cronbach AND (alpha OR alphas)) OR (item AND (correlation OR selection OR reduction)) OR agreement OR precision OR imprecision OR precise values OR test-retest OR (test AND retest) OR (reliab AND (test OR retest)) OR stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappas OR repeatab OR ((replicab OR repeated) AND (measure OR measures OR findings OR result OR results OR test OR tests)) OR generaliza OR generalisa OR concordance OR (intraclass AND correlation) OR discriminative OR known group OR factor analysis OR factor analyses OR "factor structure" OR "factor structures" OR dimension OR subscale OR (multitrait AND scaling AND (analysis OR analyses)) OR "item discriminant" OR "interscale correlation" OR error OR errors OR "individual variability" OR "interval variability" OR "rate variability" OR (variability AND (analysis OR values)) OR (uncertainty AND (measurement OR measuring)) OR "standard error of measurement" OR "sensitive" OR "responsive" OR (limit AND detection) OR "minimal

1
2
3
4
5
6
7
8
9
10 detectable concentration" OR interpretab OR ((minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference)) OR (small AND (real OR detectable) AND (change OR difference)) OR "meaningful change" OR ceiling effect OR floor effect OR "Item response model" OR IRT OR Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR item bank OR "cross-cultural equivalence")
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

BVS

(Child OR Child, Preschool OR child OR Infant OR infant OR paediatric OR pediatric OR toddler)

AND

(Child Development OR child development OR early child development OR early childhood development OR ECD OR Child Behavior OR child behavior OR child behaviour OR Cognition OR Executive Function OR executive function OR Emotional Intelligence OR attachment OR attention OR cognition OR cognitive development OR cognitive skills OR neurocognitive OR communication OR language OR language skills OR language development OR receptive language OR expressive language OR speech development OR communication skills OR Motor Skills OR motor skills OR fine motor OR gross motor OR prosocial OR socialisation OR socialization OR socioemotion OR socio-emotion OR neurodevelopment OR neurobehavior OR neurobehaviour OR neurobehavioral OR neurobehavioural OR Intelligence OR intelligence OR IQ OR psychomotor OR sensorimotor OR attention OR inhibition OR impulse control OR working memory OR problem solving OR self-regulation OR prosocial OR psychosocial development OR internali OR externali OR empathy)

AND

(instrumentation OR methods OR Validation Studies OR Comparative Study OR psychometrics OR psychometr OR clinimetr OR clinometr OR outcome assessment OR outcome measure OR observer variation OR observer variation OR Health Status Indicators OR reproducibility of results OR reproducib OR discriminant analysis OR reliab OR unreliab OR valid OR coefficient of variation OR coefficient OR homogeneity OR homogeneous OR internal consistency OR (cronbach AND (alpha OR alphas)) OR (item AND (correlation OR selection OR reduction)) OR agreement OR precision OR imprecision OR precise values OR test-retest OR (test AND retest) OR (reliab AND (test OR retest)) OR stability OR interrater OR inter-rater OR intrarater OR intra-rater OR intertester OR inter-tester OR intratester OR intra-tester OR interobserver OR inter-observer OR intraobserver OR intra-observer OR intertechnician OR inter-technician OR intratechnician OR intra-technician OR interexaminer OR inter-examiner OR intraexaminer OR intra-examiner OR interassay OR inter-assay OR intraassay OR intra-assay OR interindividual OR inter-individual OR intraindividual OR intra-individual OR interparticipant OR inter-participant OR intraparticipant OR intra-participant OR kappa OR kappas OR repeatab OR ((replicab OR repeated) AND (measure OR measures OR findings OR result OR results OR test OR tests)) OR generaliza OR generalisa OR concordance OR (intraclass AND correlation) OR discriminative OR known group OR factor analysis OR factor analyses OR "factor structure" OR "factor structures" OR dimension OR subscale OR (multitrait AND scaling AND (analysis OR analyses)) OR "item discriminant" OR "interscale correlation" OR error OR errors OR "individual variability" OR "interval variability" OR "rate variability" OR (variability AND (analysis OR values)) OR (uncertainty AND (measurement OR measuring)) OR "standard error of measurement" OR "sensitive" OR "responsive" OR (limit AND detection) OR "minimal

1
2
3
4
5
6
7
8
9
10 detectable concentration" OR interpretab OR ((minimal OR minimally OR clinical OR clinically) AND (important OR significant OR detectable) AND (change OR difference)) OR (small AND (real OR detectable) AND (change OR difference)) OR "meaningful change" OR ceiling effect OR floor effect OR "Item response model" OR IRT OR Rasch OR "Differential item functioning" OR DIF OR "computer adaptive testing" OR item bank OR "cross-cultural equivalence")

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Supplemental Table 2 Definitions of the seven early childhood domains of interest

Domain	Definition
Cognitive	Cognitive development includes abilities such as general intellectual ability, problem solving, reasoning, visual-spatial ability, memory, learning, etc.
Language	Language development includes the ability to understand and express verbal communication, and vocabulary development and acquisition.
Motor	Motor development includes gross motor development (the ability to control and coordinate gross movements of the legs and arms (e.g., jumping, throwing)) and fine motor development (the ability to control and coordinate fine movements of the fingers and toes).
Socio-emotional/temperament	<p>Socio-emotional development is the ability to regulate emotional responses and social interactions. It includes behaviour problems, social competency, emotional competency, and self-regulation.</p> <p>Temperament includes biological influences on the experience and expression of emotion: extraversion (positive affect, activity level, impulsivity, risk-taking), negative affectivity (fear, anger, sadness, discomfort) and effortful control (attention shifting and focusing, perceptual sensitivity, inhibitory and activational control).</p> <p>Socio-emotional development and temperament are often overlapping constructs, particularly in the early years.</p>
Attention/Executive function	Executive function, including attention, entails international control over behaviour and cognition. It covers inhibitory control, cognitive flexibility, planning, and working memory.
Personal-social/adaptive	Personal-social and adaptive behaviour comprises the ability to perform daily-life and self-help skills, such as feeding, dressing, toilet training, recognising others, interacting with others, and adjusting to new situations.
Academic/pre-academic	Preacademic skills are skills needed to learn reading and math, such as counting and letters.

Note: Definitions based on (Boggs et al., 2019; Fernald et al., 2017).

Supplemental Table 3 Number of included articles by type of psychometric evidence provided

Type of article	Total number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance	
		Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural	Invariance over countries	Invariance over other groups
All types of articles	182	73	63	117	66	81	74	31	72	7	10
Development articles	37	15	15	22	27	21	13	4	11	2	4
Adaptation articles	145	58	48	95	39	60	61	27	61	5	6

ected by copyright, including for research purposes related to text and data mining, AI training, and similar technologies.

136/bmjopen-2024-063655 © 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - At training and similar technologies.

Supplemental Table 4 List of countries and assessment tools used in included articles

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance on May 2025. Downloaded from http://bmjopen.bmjjournals.org/ on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - At training and similar technologies.	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Afghanistan	International Development and Early Learning Assessment (IDELA)	1	0	0	0	0	0	0	0	1	0	Halpin 2019
Argentina	Bayley Infant Neurodevelopmental Screener (BINS)	1	1	1	0	0	0	0	0	0	0	McCarthy 2012
Argentina	Escala de Habilidades Sociales (Scale of Social Abilities)	1	0	0	1	0	0	0	0	1	0	Lacunza 2009
Bangladesh	Ages & Stages Questionnaire Inventory (ASQ:I)	1	0	1	1	0	1	1	0	0	0	Pitchik 2023
Bangladesh	Bayley Scales of Infant Development, Second Edition (BSID-II)	1	0	0	0	1	0	0	0	0	0	Parveen 2014
Bangladesh	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	1	0	0	0	1	0	Pendergast 2018
Bangladesh	Dimensions of Mastery Questionnaire (DMQ-18)	1	1	1	1	1	1	0	0	1	0	Shaoli 2019
Bangladesh	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	Pisani 2018

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Bangladesh	Rapid Neurodevelopmental Assessment (RNDA)	2	0	2	0	1	2	0	1	0		0	Khan 2010; Khan 2013
Bangladesh	The language inventory	1	1	0	0	1	1	1	1	0		0	Hamadani 2010
Bangladesh	WHO Gross Motor Milestones	1	0	1	0	0	1	1	1	0		0	Hamadani 2013
Benin	Mullen Scales of Early Learning (MSEL)	2	0	1	0	0	1	1	1	0		0	Koura 2013; Boivin 2021
Bhutan	International Development and Early Learning Assessment (IDE LA)	1	1	1	1	1	1	1	0	0		0	Pisani 2018
Bolivia	Bayley Infant Neurodevelopmental Screener (BINS)	1	1	1	0	0	0	0	0	0		0	McCarthy 2012
Bolivia	International Development and Early Learning Assessment (IDE LA)	1	0	0	0	0	0	0	0	1		0	Halpin 2019
Brazil	Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	1	0	0	0	0	0	0	0	1		0	Anunciação 2019
Brazil	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	0	1	0	0	0	0	1		1	Filgueiras 2013
Brazil	Ages & Stages Questionnaires: Social Emotional (ASQ:SE-BR)	1	0	0	0	0	0	0	0	1	0	0	Chen 2017

136/bmjopen2024-096365 | 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). For peer review only. © Author(s) (or their employer(s)) 2024. All rights reserved. No reuse allowed without permission.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Brazil	Alberta Infant Motor Scale (AIMS)	4	3	3	1	1	3	0	1	0	0	0	Saccani 2016; Siegle 2018; de Albuquerque 2018; Valentini 2012
Brazil	Battelle Developmental Inventory Second Edition, BDI-2	1	0	0	0	1	0	0	0	0	0	0	de Albuquerque 2022
Brazil	Bayley Infant Neurodevelopment Screener (BINS)	2	1	1	1	1	1	0	1	1	0	0	Guedes 2011; McCarthy 2012
Brazil	Bayley Scales of Infant Development, Third Edition (BSID-III)	2	1	0	1	1	1	0	0	2	0	0	Pendergast 2018; Madaschi 2016
Brazil	Caregiver Reported Early Development Instruments (CREDI), short-form	1	1	0	1	1	1	1	0	0	0	0	Altafim 2018
Brazil	Denver Developmental Screening Test (Denver-II)	1	0	1	1	1	0	1	0	1	0	0	Lopez Boo 2020
Brazil	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1	0	1	Sincovich 2019
Brazil	Motor Development Scale	1	0	0	1	0	0	0	0	1	0	0	Okuda 2020
Brazil	Oxford Neurodevelopment Assessment (OX-NDA)	1	1	1	1	0	1	0	0	0	0	0	Fernandes 2022

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Brazil	Preschool and Kindergarten Behavior Scale (PKBS)	1	0	0	1	0	0	0	0	1		0	Dias 2011
Brazil	Questionário para Avaliação do Desenvolvimento Infantil (QAD-PIPAS)	1	0	0	0	1	0	0	0	0		0	Venancio 2020
Brazil	Screening for Identification of Oral Language Difficulties by Preschool Teachers (SIOLD)	1	0	0	0	0	0	0	0	1		0	Puglisi 2020
Brazil	Self-ordered Pointing Task (SOPT)	1	0	0	0	0	1	1	0	0		0	Sallum 2017
Brazil	Test of Gross Motor Development (TGMD-2)	1	0	1	1	0	0	1	0	1		0	Bandeira 2020
Cambodia	Angkor Hospital for Children Developmental Milestone Assessment Tool (AHC DMAT)	1	0	0	0	1	0	0	0	0		0	Ngoun 2012
Cambodia	Developmental Milestone Chart (DMC)	1	0	0	0	1	0	0	0	0		0	Scherzer 2009
Cameroon	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	0	0	0	1	0	1	0	Lohaus 2014
China	Ages & Stages Questionnaire Inventory (ASQ:I)	1	1	0	1	1	1	0	0	0	0	0	Xie 2017

136/bmjopen2024-096355 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). For peer review only. © Author(s) 2024. Licensee BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution Non Commercial No Derivatives License (CC BY-NC-ND). For more information, please see the [http://creativecommons.org/licenses/by-nc-nd/4.0/](#).

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
China	Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	1	1	1	1	1	0	0	0	1		0	Bian 2017
China	Ages & Stages Questionnaire: Social-Emotional, First Edition (ASQ:SE-1)	1	1	1	1	0	1	0	0	0		0	Xie 2019 (Journal of Child & Dev Studies)
China	Ages & Stages Questionnaire-3 (ASQ-3)	2	1	1	1	0	2	1	0	1		0	Yue 2019; Bian 2012
China	Caregiver Reported Early Development Instruments (CREDI), long-form	1	0	0	1	0	1	1	0	0		0	Li 2020
China	Child Behavior Checklist/1.5-5 (CBCL/1.5-5) and the Caregiver-Teacher Report Form (C-TRF)	1	0	0	0	0	0	0	0	1		0	Liu 2011
China	Child Social Preference Scale	1	0	0	1	0	0	0	1	1		1	Li 2016
China	Chinese Preschool Readiness Scale (CPRS)	1	0	0	1	1	1	0	0	1		0	Xie 2019 (Early Educ & Dev)
China	Chinese version of Urban Infant-Toddler Social and Emotional Assessment	1	1	1	1	1	1	0	0	1	0	0	Jianduan 2009
China	Chinese Vocabulary Checklist	1	0	0	0	0	0	1	0	0	0	0	Hao 2008

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
China	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1		1	Sincovich 2019
China	Emotional and Behavioral Screener (EBS)	1	0	1	1	0	1	0	0	1		0	Zhang 2018
China	Movement Assessment Battery for Children- second edition (MBAC-2)	1	1	1	1	1	1	0	0	1		0	Hua 2013
China	Preschool Learning Behavior Scale (PLBS)	1	0	0	1	0	1	0	0	1		0	Wu 2019 (JPA)
China	Social Skills domain of System-Rating Scales (SSIS-RS)	1	0	0	1	0	0	1	0	1		0	Wu 2019 (PS)
China	The Chinese Early Language and Communication Questionnaire	1	0	0	1	0	1	0	0	0		0	Zhang 2021
China	The Chinese Intelligence Scale for Young Children	1	0	0	0	0	0	0	0	1		1	Guo 2009
Colombia	Ages & Stages Questionnaire-3 (ASQ-3)	2	2	0	2	0	1	1	1	0		0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	Battelle Developmental Inventory Second Edition, BDI-2	2	2	0	2	0	1	1	1	0		0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	Bayley Infant Neurodevelopmental Screener (BINS)		1	1	0	0	0	0	0	0		0	McCarthy 2012

136/bmjopen2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - www.abes.fr/. Copyright © Peer Review Only and similar technologies.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Colombia	Bayley Scales of Infant Development, Third Edition (BSID-III)	2	2	0	2	0	1	1	1	0		0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	Denver Developmental Screening Test (Denver-II)	2	2	0	2	0	1	1	1	0		0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	MacArthur-Bates Communicative Development Inventories CDI	1	0	0	1	1	0	1	0	0		0	Lara Díaz 2011
Colombia	MacArthur-Bates Communicative Development Inventories Short-Forms I vocabulary checklist	2	2	0	2	0	1	1	1	0		0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	MacArthur-Bates Communicative Development Inventories Short-Forms II vocabulary checklist	2	2	0	2	0	1	1	1	0		0	Rubio-Codina 2016; Rubio-Codina 2020
Colombia	Preschool Language Scale-3 (PLS- 3)	1	0	0	1	0	0	0	0	1		0	Romero 2013
Colombia	WHO Gross Motor Milestones	2	2	0	2	0	1	1	1	0		0	Rubio-Codina 2016; Rubio-Codina 2020
Costa Rica	EDIN II (Evaluación del Desarrollo Integral de la niña y el niño)	1	0	0	0	1	0	1	0	0		0	Cordero 2019

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Côte d'Ivoire	Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	1	0	0	1	0	0	1	0	1		0	Barthel 2021
Côte d'Ivoire	Kilifi Developmental Inventory	1	0	1	1	0	0	0	1	1		0	Barthel 2022
Dominican Republic	Malawi Developmental Assessment Tool (MDAT), renamed Dominican Child Development Screening Tool	1	0	1	1	0	0	1	0	0		0	Sánchez-Vincitore 2019
Ecuador	Bayley Infant Neurodevelopmental Screener (BINS)	1	1	1	0	0	0	0	0	0		0	McCarthy 2012
Egypt	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0		0	Pisani 2018
Ethiopia	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	1	1	0	0	1	1	1		0	Hanlon 2016
Ethiopia	Denver Developmental Screening Test (Denver-II)	1	1	1	0	1	0	0	0	0		0	Abessa 2016
Ethiopia	Global Scales for Early Development	1	0	1	0	1	0	0	0	0		0	Yibeltal 2025

136/bmjopen2024-096355 | 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - www.abes.fr/. Copyright © 2024 Author(s). All rights reserved. No reuse allowed without permission.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Ethiopia	International Development and Early Learning Assessment (IDELA)	3	1	1	1	1	2	2	0	2		1	Halpin 2019; Pisani 2018; Wolf 2017
Georgia	Ages & Stages Questionnaire-3 (ASQ-3)	1	1	0	1	1	0	1	0	0		0	Zirakashvili 2018
Ghana	Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	1	0	0	1	0	0	1	0	1		0	Barthel 2021
Ghana	International Development and Early Learning Assessment (IDELA)	1	0	0	1	0	0	0	1	1		0	Pisani 2022
Ghana	Kilifi Developmental Inventory	1	0	1	1	0	0	0	1	1		0	Barthel 2022
Ghana	Rapid Assessment of Cognitive and Emotional Regulation (RACER)	1	1	1	0	0	1	1	0	0		0	Yuan 2022
Ghana	Vocabulary Checklists	1	0	0	0	0	1	0	1	0		0	Prado 2018
Grenada	Grenada Learning and Memory Scale	1	1	1	1	1	1	1	0	1		0	Blackmon 2024
Grenada	INTER- GROWTH-21st Project Neurodevelopment Assessment (INTER-NDA)	1	1	1	1	0	0	0	0	0		0	Waechter 2022
Guatemala	Ages & Stages Questionnaire-3 (ASQ-3)	1	1	0	1	0	1	0	1	0	0	0	Colbert 2021

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Guatemala	Mullen Scales of Early Learning (MSEL)	1	1	1	1	0	0	1	1	0		0	Colbert 2020
Guatemala	Rapid Neurodevelopmental Assessment (RNDA)	1	0	1	0	1	1	1	0	0		0	Thompson 2015
India	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	1	1	1	0	0	0	0		0	Kvestad 2013
India	Bayley Scales of Infant Development, Second Edition (BSID-II)	1	0	0	0	0	1	0	0	0		0	Biasini 2015
India	Bayley Scales of Infant Development, Third Edition (BSID-III)	2	0	0	0	1	0	0	1	1		0	Pendergast 2018; Koshy 2024
India	Caregiver Reported Early Development Instruments (CREDI), long-form	1	0	0	1	0	1	1	0	0		0	Alderman 2021
India	DEvelopmental assessment on an E-Platform (DEEP)	2	0	0	0	1	1	0	0	0		0	Bhavnani 2019; Mukherjee 2020
India	Developmental Assessment Tool for Anganwadis (DATA)	1	0	0	1	1	0	0	0	1		0	Nair 2009
India	Developmental Assessment Tool for Anganwadis (DATA-II)	1	0	0	1	1	0	0	0	1		0	Nair 2012
India	International Development and Early Learning Assessment (IDEA)	1	1	1	1	1	1	1	0	0		0	Pisani 2018

136/bmjopen2024-096355
Number of articles providing evidence on measurement invariance
over countries
11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES).
For peer review only - AI training, AI and similar technologies.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
India	Parents' Evaluation of Developmental Status (PEDS)	1	0	0	0	0	1	0	0	0	0	0	Mukherjee 2022
India	Screening Test of Early Language Development-Test version (STELD-T)	1	0	0	1	0	1	0	0	0	0	0	Pathak 2023
India	Bracken School Readiness Assessment, Third Edition (BSRA-3)	1	0	0	1	1	0	0	0	0	0	0	Shah 2013
India	Vineland Adaptive Behaviour Scale (VABS II)	1	1	1	0	0	0	1	0	0	0	0	Kumar 2016
Indonesia	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	1	1	1	1	0	0	0	0	0	Yunilda 2022
Indonesia	Early Development Instrument	1	0	1	1	1	1	0	1	0	0	0	Brinkman 2017
Indonesia	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018
Iran	Ages & Stages Questionnaire (ASQ)	1	0	1	1	1	0	0	0	1	0	0	Vameghi 2013
Iran	Bayley Scales of Infant and Toddler Development (BSID)	1	1	1	1	1	0	1	0	1	0	0	Azari 2017
Iran	Childhood Nonverbal Communication Scale (CNCS)	1	0	0	1	1	1	1	0	1	0	0	Oryadi-Zanjani 2020
Iran	Children's Behavior Questionnaire (CBQ)	1	1	0	1	1	1	1	0	1	0	0	Golmohammadi 2022

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Iran	Elicited Imitation Test (EIT)	1	0	1	0	0	1	0	0	0		1	Mehrani 2018
Iran	Farsi Narrative Norms Instrument	1	1	1	0	1	0	0	0	0		0	Soleymani 2016
Iran	Infant Neurological International Battery	1	0	1	0	0	1	0	0	0		0	Soleimani 2007
Jamaica	Intelligibility in Context Scale (ICS) / ICS-Jamaican Creole	1	1	1	1	0	1	1	0	1		0	Washington 2017
Kenya	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	1	0	1	1	0	1	0	1		0	McHenry 2021
Kenya	Child behaviour checklist (CBCL)	1	1	0	1	0	0	1	0	1		0	Kariuki 2016
Kenya	Communicative Development Inventory (CDI)	1	1	0	1	0	1	0	0	0		0	Alcock 2015
Kenya	Developmental Milestones Checklist (DMC)	1	1	0	1	1	1	1	0	1		0	Abubakar 2010
Kenya	Executive Function Touch (EF Touch)	2	0	1	1	1	1	2	0	1		0	Willoughby 2019 (Child Neuropsych); Willoughby 2019 (Dev Science)
Kenya	Global Scales of Early Development	1	0	0	0	1	1	1	0	0	0	0	Jeong 2025
Kenya	Kilifi Developmental Inventory	1	1	1	1	1	1	1	0	0	0	0	Abubakar 2008

136/bmjopen2024-096365
 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). At training, and similar technologies.
 For peer review only

136/bmjopen2024-096365 | 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). For peer review only. © 2024 Author(s). Licensee BMJ Publishing Group Ltd.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Kenya	Receptive and expressive vocabulary test battery (English, Luo and Swahili)	1	0	1	1	0	1	1	1	0		0	Knauer 2019
Kiribati	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1		1	Sincovich 2019
Laos	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1		1	Sincovich 2019
Laos	MELQO Measure of Development and Learning Outcomes	1	0	0	0	0	0	1	0	1		1	Gomez 2022
Malawi	International Development and Early Learning Assessment (IDE LA)	1	1	1	1	1	1	1	0	0		0	Pisani 2018
Malawi	Malawi Developmental Assessment Tool (MDAT)	1	1	1	0	1	0	1	0	0		0	Gladstone 2010
Malawi	Vocabulary Checklists	1	0	0	0	0	1	0	1	0		0	Prado 2018
Malaysia	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	1	0	0	0	0	1		0	Zakaria 2012
Malaysia	Malay Preschool Language Assessment Tool (MPLAT)	1	1	1	1	1	0	0	0	0		0	Razak 2010

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Mali	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018
Mexico	Ages & Stages Questionnaire-3 (ASQ-3)	1	0	1	1	0	0	0	0	1	0	0	Ortiz-Leon 2018
Mexico	Early Childhood Development Index 2030	1	0	0	1	1	1	0	0	1	0	1	Halpin 2024
Mexico	Spanish-language MacArthur–Bates Communicative Development Inventories short forms (S-CDI SFI and SFII)	1	1	0	0	0	0	1	0	0	0	0	Jackson-Maldonado 2013
Mongolia	Mongolian Rapid Baby Scale (MORBAS)	1	0	0	0	1	1	0	0	0	0	0	Dagvadorj 2015
Mozambique,	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0	0	0	Pisani 2018
Nepal	Ages & Stages Questionnaire-3 (ASQ-3)	2	0	2	1	1	0	1	1	0	0	0	Shrestha 2019; Shrestha 2024
Nepal	Bayley Scales of Infant Development, Third Edition (BSID-III)	3	0	2	1	1	0	0	0	1	0	0	Pendergast 2018; Ranjitkar 2018; Manandhar 2016

136/bmjopen2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). For peer review only. AI training and similar technologies.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Nigeria	Early Childhood Development Index (ECDI)	1	0	0	1	1	0	0	0	1		0	Enelamah 2024
Nepal	The Test of Infant Motor Performance (TIMP)	1	0	1	1	0	1	0	1	0		0	Kvestad 2023
Pakistan	Bayley Scales of Infant Development, Second Edition (BSID-II)	1	0	0	0	0	1	0	0	0		0	Biasini 2015
Pakistan	Bayley Scales of Infant Development, Third Edition (BSID-III)	2	0	0	0	1	1	0	1	1		0	Pendergast 2018; Rasheed 2023
Pakistan	International Development and Early Learning Assessment (IDE LA)	1	1	1	1	1	1	1	0	0		0	Pisani 2018
Pakistan	Executive function battery	1	0	0	1	1	0	1	1	1		0	Obradovic 2019
Pakistan	Executive Function Touch (EF Touch)	1	0	0	0	0	0	1	0	1		0	Toor 2022
Pakistan	MacArthur Communicative Development Inventory Level II	1	1	1	1	1	1	0	0	0		0	Rasheed 2017
Pakistan	Strengths and Difficulties Questionnaire (SDQ)	1	1	1	1	1	1	1	0	1		0	Finch 2018
Pakistan	Wechsler Preschool and Primary Scale of Intelligence, third edition (WPPSI-III)	1	0	1	1	1	0	1	0	1	0	0	Rasheed 2018

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Palestine	Early Childhood Development Index 2030	1	0	0	1	1	1	0	0	1		1	Halpin 2024
Peru	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	1	0	0	0	1		0	Pendergast 2018
Peru	Child Behavior Scale (CBS)	1	0	0	1	0	0	0	0	1		0	Meyer 2011
Peru	Escala de Conductas de Aprendizaje Preescolar (ECAP), a Spanish translation of the Preschool Learning Behaviors Scale (PLBS)	1	0	0	0	0	0	0	0	1		0	Hahn 2009
Rwanda	International Development and Early Learning Assessment (IDELA)	1	1	1	1	1	1	1	0	0		0	Pisani 2018
Samoa	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1		1	Sincovich 2019
Senegal	Milestones Checklist; Vocabulary Inventory	1	0	0	1	1	1	1	0	0		0	Weber 2018
Sierra Leone	Saccadic reaction time (SRTs)	1	1	0	0	0	0	1	0	0		0	Leppänen 2022
South Africa	Ages & Stages Questionnaire (ASQ)	1	0	0	0	1	0	0	0	0		0	Bornman 2010
South Africa	Ages & Stages Questionnaire-3 (ASQ-3)	2	0	0	2	0	0	1	0	1		0	van Heerden 2017; Hsiao 2016

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
South Africa	Bayley Infant Neurodevelopmental Screener (BINS)	1	0	0	1	0	0	1	0	1	0	Rodriguez 2020
South Africa	Bayley Scales of Infant Development, Third Edition (BSID-III)	1	0	0	0	1	0	0	0	1	0	Pendergast 2018
South Africa	Early Learning Outcomes Measure (ELOM)	1	0	1	1	0	0	0	0	1	1	Snelling 2019
South Africa	Little Developmental Coordination Disorder Questionnaire (Little DCDQ)	1	0	0	1	0	1	0	0	0	0	Venter 2015
South Africa	Mullen Scales of Early Learning (MSEL)	1	0	0	0	1	0	0	0	0	0	Bornman 2010
South Africa	Parents' Evaluation of Developmental status (PEDS)	1	0	0	0	0	1	0	0	0	0	Abdoola 2021
South Africa	The infant gross motor screening test (IGMST)	1	1	1	0	1	1	0	0	0	0	Hilburn 2011
South Africa	The Road to Health Booklet (RTHB)	1	0	0	0	0	1	0	0	0	0	van der Linde 2015
Sri Lanka	Child Behaviour Assessment Instrument (CBAI)	1	1	0	1	1	1	0	0	0	0	Samarakkody 2010

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Sri Lanka	Denver Developmental Screening Test for Sri Lankan Children (DDST-SL)	1	0	1	0	0	0	0	0	0		0	Wijedasa 2012
Sri Lanka	Ragama Early Assessment for Children (REACH)	1	0	1	1	1	1	0	0	0		0	Caldera 2023
Tanzania	Caregiver Reported Early Development Instruments (CREDI)	1	1	1	1	1	1	1	0	0		0	McCoy 2017
Thailand	Parents' Evaluation of Developmental Status (PEDS)	1	0	0	0	0	1	0	0	0		0	Chunsuwan 2016
Thailand	Thai Speech and Language Test (TSLT)	1	0	1	1	1	0	0	0	0		0	Prathanee 2008
The Gambia	Mullen Scales of Early Learning (MSEL)	1	0	0	0	0	0	1	1	0		0	Milosavljevic 2019
Tonga	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1		1	Sincovich 2019
Turkey	Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	1	1	1	1	0	1	0	0	0		0	Kucuker 2011
Turkey	Bayley Infant Neurodevelopmental Screener (BINS)	1	0	0	0	0	0	0	1	0		0	Soysal 2014
Turkey	Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	1	1	1	1	0	1	0	0	0		0	Karabekiroglu 2009

136/bmjopen
2024/09/6365
on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) -
Related to text and data mining, AI training, and similar technologies.

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries	Over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Turkey	Emotional Regulation Checklist	1	1	0	1	1	1	0	0	1		0	Danişman 2016
Turkey	Kaufman Survey of Early Academic and Language Skills (K-SEALS)	1	1	1	1	1	0	0	0	1		0	Uyanık 2014
Turkey	Strengths and Difficulties Questionnaire (SDQ)	1	0	0	1	0	1	1	0	0		0	Dursun 2020
Tuvalu	Early Human Capability Index (eHCI)	1	0	0	1	0	0	0	0	1		1	Sincovich 2019
Uganda	Bayley Scales of Infant Development, bead threading task	1	1	0	1	0	0	1	0	1		0	Nampijja 2010
Uganda	British Ability Scales – 3rd ed., block design task	1	1	0	1	0	0	1	0	1		0	Nampijja 2010
Uganda	International Development and Early Learning Assessment (IDELA)	1	0	0	0	0	0	0	0	1		0	Halpin 2019
Uganda	Kilifi Developmental Inventory, coin box task	1	1	0	1	0	0	1	0	1		0	Nampijja 2010
Uganda	Kilifi Picture Vocabulary Test	1	1	0	1	0	0	1	0	1		0	Nampijja 2010
Uganda	Modified Fagan Test for Infant Intelligence (FTII)	1	0	0	0	0	0	0	1	0		0	Familiar-Lopez 2022

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Uganda	Movement Assessment Battery for Children, balancing on one leg	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Uganda	NEPSY verbal fluency task	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Uganda	NEPSY, knock tap game	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Uganda	NEPSY, sentence repetition task	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Uganda	Sky Search in the Tests of Everyday Attention for Children, Sky Search Task	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Uganda	Tap Once Tap Twice Task	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Uganda	Tobii Professional Studio eye tracking programming	1	0	0	0	0	1	0	0	0	0	Boivin 2017
Uganda	Wisconsin Card Sort Test	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Venezuela	Bayley Infant Neurodevelopmental Screener (BINS)	1	1	1	0	0	0	0	0	0	0	McCarthy 2012
Viet Nam	Vietnamese version of the Intelligibility in Context Scale (ICS-VN).	1	0	0	1	0	1	1	0	0	0	Pham 2017

Country	Name of assessment tool	Number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over countries and over groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Vietnam	International Development and Early Learning Assessment (IDE LA)	1	0	0	0	0	0	0	0	1	0	Halpin 2019
Zambia	Ages & Stages Questionnaire-3 (ASQ-3)	2	0	0	2	0	0	1	0	1	0	van Heerden 2017; Hsiao 2016
Zambia	Bayley Scales of Infant Development, Second Edition (BSID-II)	1	0	0	0	0	1	0	0	0	0	Biasini 2015
Zambia	International Development and Early Learning Assessment (IDE LA)	1	1	1	1	1	1	1	0	0	0	Pisani 2018

Supplemental Table 5 List of child development assessment tools and countries where they were applied

Name assessment tool	Number of countries where the tool was applied	List of countries where the tool was applied	Number of articles where tool was applied	Age range at assessment (in months)	Type of administration	References (first author & year)
Ages & Stages Questionnaire (ASQ)	2	Iran; South Africa	1	4-60	Indirect/reported	Vameghi 2013; Bornman 2010
Ages & Stages Questionnaire Inventory (ASQ:I)	2	Bangladesh; China	1	1-27	Indirect/reported	Pitchik 2023; Xie 2017
Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	3	Brazil; China; Turkey	3	4-72	Indirect/reported	Anunciação 2019; Bian 2017; Kucuker 2011
Ages & Stages Questionnaire: Social-Emotional, First Edition (ASQ:SE-1)	1	China	1	15-63	Indirect/reported	Xie 2019 (Journal of Child & Dev Studies)
Ages & Stages Questionnaire-3 (ASQ-3)	11	Brazil; China; Colombia; Georgia; Guatemala; India; Indonesia; Mexico; Nepal; South Africa; Zambia	14	1-66	Indirect/reported	Filgueiras 2013; Yue 2019; Rubio-Codina 2016; Rubio-Codina 2020; Zirakashvili 2018; Colbert 2021; Kvistad 2013; Yunilda 2022; Ortiz-Leon 2018; Shrestha 2019; van Heerden 2017; Hsiao 2016; Bian 2012; Shrestha 2024
Ages & Stages Questionnaires: Social Emotional (ASQ:SE-BR)	1	Brazil	1	21-66	Indirect/reported	Chen 2017
Alberta Infant Motor Scale (AIMS)	1	Brazil	4	0-19	Direct assessment	Saccani 2016; Siegle 2018; de Albuquerque 2018; Valentini 2012
Angkor Hospital for Children Developmental Milestone Assessment Tool (AHC DMAT)	1	Cambodia	1	1-72	Direct assessment	Ngoun 2012
Battelle Developmental Inventory Second Edition, BDI-2	2	Brazil; Colombia	3	0-95	Direct assessment indirect/reported	de Albuquerque 2022; Rubio-Codina 2016; Rubio-Codina 2020
Bayley Infant Neurodevelopmental Screener (BINS)	8	Argentina, Brazil, Ecuador, Bolivia, Colombia, Venezuela; Turkey; South Africa	4	3-24	Direct assessment	McCarthy 2012; Soysal 2014; Rodriguez 2020; Guedes 2011
Bayley Scales of Infant and Toddler Development (BSID)	1	Iran	1	1-42	Not reported	Azari 2017

1 2 3 4	Name assessment tool	Number of countries where the tool was applied	List of countries where the tool was applied	Number of articles where tool was applied	Age range at assessment (in months)	Type of administration	136/bmjopen-2024-096365 on 11 May 2025. Downloaded from http://bmjopen.bmjjournals.org/	References (first author & year)
5 6 7 8	Bayley Scales of Infant Development, bead threading task	1	Uganda	1	55-67	Direct assessment	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Nampijja 2010
9 10 11 12 13 14 15	Bayley Scales of Infant Development, Second Edition (BSID-II)	4	Bangladesh, India, Pakistan, Zambia	2	1-42	Direct assessment	Bibliographique de l'	Parveen 2014
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Bayley Scales of Infant Development, Third Edition (BSID-III)	12	Bangladesh, Brazil, India, Nepal, Pakistan, Peru, South Africa; Colombia; Ethiopia; Cameroon; Kenya; Malaysia	12	6-42	Direct assessment	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Pendergast 2018; Madaschi 2016; Rubio-Codina 2016; Rubio-Codina 2020; Hanlon 2016; Lohaus 2014; McHenry 2021; Zakaria 2012; Ranjitkar 2018; Manandhar 2016; Rasheed 2023; Koshy 2024
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Bracken School Readiness Assessment, Third Edition (BSRA-3)	1	India	1	36-60	Direct assessment	Bibliographique de l'	Shah 2013
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	3	Côte d'Ivoire; Ghana; Turkey	2	23-32	Indirect/reporter	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Barthel 2021; Karabekiroglu 2009
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	British Ability Scales – 3rd ed., block design task	1	Uganda	1	55-67	Direct assessment	Bibliographique de l'	Nampijja 2010
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Caregiver Reported Early Development Instruments (CREDI)	18	17 countries not listed, Tanzania	2	0-35	Indirect/reporter	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Waldman 2021; McCoy 2017
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Caregiver Reported Early Development Instruments (CREDI), long-form	2	China; India	2	5-36	Indirect/reporter	Bibliographique de l'	Li 2020; Alderman 2021
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Caregiver Reported Early Development Instruments (CREDI), short-form	18	17 countries not listed; Brazil	2	0-35	Indirect/reporter	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	McCoy 2018; Altafim 2018
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Child Behavior Checklist/1.5-5 (CBCL/1.5-5) and the Caregiver-Teacher Report Form (C-TRF)	1	China	1	50-71	Indirect/reporter	Bibliographique de l'	Liu 2011
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Child Behavior Scale (CBS)	1	Peru	1	24-72	Indirect/reporter	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Meyer 2011
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Child Behaviour Assessment Instrument (CBAI)	1	Sri Lanka	1	48-72	Indirect/reporter	Bibliographique de l'	Samarakkody 2010
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Child behaviour checklist (CBCL)	1	Kenya	1	12-72	Indirect/reporter	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Kariuki 2016
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Child Social Preference Scale	1	China	1	36-60	Indirect/reporter	Bibliographique de l'	Li 2016
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Childhood Nonverbal Communication Scale (CNCS)	1	Iran	1	3-18	Indirect/reporter	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Oryadi-Zanjani 2020
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Children's Behavior Questionnaire (CBQ)	1	Iran	1	48-72	Indirect/reporter	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Golmohammadi 2022
24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47	Chinese Preschool Readiness Scale (CPRS)	1	China	1	31-62	Indirect/reporter	Enseignement Supérieur (ABES). Autodétermination, Apprentissage et Techniques de	Xie 2019 (Early Educ & Dev)

Name assessment tool	Number of countries where the tool was applied	List of countries where the tool was applied	Number of articles where tool was applied	Age range at assessment (in months)	Type of administration	References (first author & year)
Chinese version of Urban Infant-Toddler Social and Emotional Assessment	1	China	1	12-36	Indirect/reported	Jianduan 2009
Chinese Vocabulary Checklist	1	China	1	12-30	Indirect/reported	Hao 2008
Communicative Development Inventory (CDI)	1	Kenya	1	6-36	Indirect/reported	Alcock 2015
Denver Developmental Screening Test for Sri Lankan Children (DDST-SL)	1	Sri Lanka	1	0-80	Direct assessment	Wijedasa 2012
Denver Developmental Screening Test (Denver-II)	3	Brazil; Colombia; Ethiopia	4	0-70.6	Direct assessment; indirect/reported	Lopez Boo 2020; Rubio-Codina 2016; Rubio-Codina 2020; Abessa 2016
DEvelopmental assessment on an E-Platform (DEEP)	1	India	2	34-40	Direct assessment	Bhavnani 2019; Mukherjee 2020
Developmental Assessment Tool for Anganwadis (DATA)	1	India	1	19-36	Direct assessment	Nair 2009
Developmental Assessment Tool for Anganwadis (DATA-II)	1	India	1	36-48	Direct assessment	Nair 2012
Developmental Milestone Chart (DMC)	1	Cambodia	1	1-96	Direct assessment; indirect/reported	Scherzer 2009
Developmental Milestones Checklist (DMC)	1	Kenya	1	2.53-12.06	Indirect/reported	Abubakar 2010
Dimensions of Mastery Questionnaire (DMQ-18)	1	Bangladesh	1	36-72	Indirect/reported	Shaoli 2019
Early Childhood Development Index (ECDI)	1	Nigeria	1	36-48	Indirect/reported	Enelamah 2024
Early Childhood Development Index 2030 (ECDI2030)	2	Mexico, Palestine	1	24-59	Indirect/reported	Halpin 2024
Early Development Instrument	1	Indonesia	1	48-96	Indirect/reported	Brinkman 2017
Early Human Capability Index (eHCI)	7	Brazil, Kiribati, China, Laos, Samoa, Tonga, Tuvalu	1	24-72	Indirect/reported	Sincovich 2019
Early Learning Outcomes Measure (ELOM)	1	South Africa	1	50-69	Direct assessment	Snelling 2019
EDIN II (Evaluación del Desarrollo Integral de la niña y el niño)	1	Costa Rica	1	0-72	Direct assessment	Cordero, 2019
Elicited Imitation Test (EIT)	1	Iran	1	36-83	Direct assessment	Mehrani 2018
Emotional and Behavioral Screener (EBS)	1	China	1	48-84	Indirect/reported	Zhang 2018
Emotional Regulation Checklist	1	Turkey	1	48-60	Indirect/reported	Danişman 2016

1 2 3 4	Name assessment tool	Number of countries where the tool was applied	List of countries where the tool was applied	Number of articles where tool was applied	Age range at assessment (in months)	Type of administration	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	References (first author & year)
5 6 7 8 9	Escala de Conductas de Aprendizaje Preescolar (ECAP), a Spanish translation of the Preschool Learning Behaviors Scale (PLBS)	1	Peru	1	24-72	Indirect/reporting	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Hahn 2009
10 11 12 13 14	Escala de Habilidades Sociales (Scale of Social Abilities)	1	Argentina	1	36-60	Indirect/reporting	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Lacunza 2009
15 16 17 18 19	Executive function battery	1	Pakistan	1		Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Obradovic 2019
20 21 22 23 24	Executive Function Touch (EF Touch)	2	Kenya	3	36-72	Self-administration	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Willoughby 2019 (Child Neuropsych); Willoughby 2019 (Dev Science); Toor 2022
25 26 27 28 29	Farsi Narrative Norms Instrument	1	Iran	1	60-72	Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Soleymani 2016
30 31 32 33 34	Global Scales for Early Development	32	Not listed Ethiopia, Kenya	3	0-36	Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	McCrory 2023; Jeong 2025; Yibeltal 2025
35 36 37 38 39	Grenada Learning and Memory Scale	1	Grenada	1	36-72	Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Blackmon 2024
40 41 42 43 44	Infant Neurological International Battery	1	Iran	1	4-18	Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Soleimani 2007
45 46 47	Intelligibility in Context Scale (ICS) / ICS-Jamaican Creole	1	Jamaica	1	39-75	Indirect/reporting	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Washington 2017
48 49 50 51 52	INTER- GROWTH-21st Project Neurodevelopment Assessment (INTER-NDA)	1	Grenada	1	22-30	Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Waechter 2022
53 54 55 56 57	International Development and Early Learning Assessment (IDEA)	17	Afghanistan, Ethiopia, Uganda, Vietnam, Bolivia, Bangladesh, Bhutan, Egypt, India, Indonesia, Mali, Malawi, Mozambique, Pakistan, Rwanda, Zambia, Ghana	4	36-96	Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Halpin 2019; Pisani 2018; Wolf 2017; Pisani 2022
58 59 60 61 62	Kaufman Survey of Early Academic and Language Skills (K-SEALS)	1	Turkey		61-72	Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Uyanik 2014
63 64 65 66 67	Kilifi Developmental Inventory	2	Ghana, Côte d'Ivoire, Kenya	2	6-35	Direct assessment	136/bmjopen-2024-096365 on 13 May 2025. Downloaded from http://bmjopen.bmjjournals.com/	Enseignement Supérieur (ABES) · AI training, and similar technologies.	Barthel 2022; Abubakar 2008

Name assessment tool	Number of countries where the tool was applied	List of countries where the tool was applied	Number of articles where tool was applied	Age range at assessment (in months)	Type of administration, including for what purpose, including for research purposes reported to text and data mining, training and similar technologies	References (first author & year)
Kilifi Developmental Inventory, coin box task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
Kilifi Picture Vocabulary Test	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
Little Developmental Coordination Disorder Questionnaire (Little DCDQ)	1	South Africa	1	36-49	Indirect/reported	Venter 2015
MacArthur Communicative Development Inventory Level II	1	Pakistan	1		Indirect/reported	Rasheed 2017
MacArthur-Bates Communicative Development Inventories CDI	1	Colombia	1	8-30	Indirect/reported	Lara Diaz 2011
MacArthur-Bates Communicative Development Inventories Short-Forms I vocabulary checklist	1	Colombia	2	6-42	Indirect/reported	Rubio-Codina 2016; Rubio-Codina 2020
MacArthur-Bates Communicative Development Inventories Short-Forms II vocabulary checklist	1	Colombia	2	6-42	Indirect/reported	Rubio-Codina 2016; Rubio-Codina 2020
Malawi Developmental Assessment Tool (MDAT)	1	Malawi	1	0-84	Direct assessment; indirect/reported	Gladstone 2010
Malawi Developmental Assessment Tool (MDAT), renamed Dominican Child Development Screening Tool	1	Dominican Republic	1	2-24	Direct assessment; indirect/reported	Sánchez-Vincitore 2019
Malay Preschool Language Assessment Tool (MPLAT)	1	Malaysia	1	48-83	Direct assessment	Razak 2010
MELQO Measure of Development and Learning Outcomes	1	Laos	1	48-49	Direct assessment	Gomez 2022
Milestones Checklist; Vocabulary Inventory	1	Senegal	1	4-30	Indirect/reported	Weber 2018
Modified Fagan Test for Infant Intelligence (FTII)	1	Uganda	1	6-12	Direct assessment	Familiar-Lopez 2022
Mongolian Rapid Baby Scale (MORBAS)	1	Mongolia	1	0-42	Indirect/reported	Dagvadorj 2015
Motor Development Scale	1	Brazil	1	48-72	Not reported	Okuda 2020
Movement Assessment Battery for Children-second edition (MBAC-2)	1	China	1	36-72	Direct assessment	Hua 2013
Movement Assessment Battery for Children, balancing on one leg	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
Mullen Scales of Early Learning (MSEL)	4	Guatemala; South Africa; The Gambia	5	0-68	Direct assessment	Koura 2013; Boivin 2021; Colbert 2020; Bornman 2010; Milosavljevic 2019

Name assessment tool	Number of countries where the tool was applied	List of countries where the tool was applied	Number of articles where tool was applied	Age range at assessment (in months)	Type of administration	References (first author & year)
NEPSY verbal fluency task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
NEPSY, knock tap game	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
NEPSY, sentence repetition task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
Oxford Neurodevelopment Assessment (OX-NDA)	1	Brazil	1		Direct assessment/indirect/reported	Fernandes 2022
Parents' Evaluation of Developmental Status (PEDS)	3	India; South Africa; Thailand	3	0-30	Indirect/reported	Mukherjee 2022; Abdoola 2021; Chunswan 2016
Preschool and Kindergarten Behavior Scale (PKBS)	1	Brazil	1	36-72	Indirect/reported	Dias 2011
Preschool Language Scale-3 (PLS- 3)	1	Colombia	1	48-84	Direct assessment	Romero 2013
Preschool Learning Behavior Scale (PLBS)	1	China	1		Indirect/reported	Wu 2019 (JPA)
Questionário para Avaliação do Desenvolvimento Infantil (QAD-PIPAS)	1	Brazil	1	0-59	Indirect/reported	Venancio 2020
Ragama Early Assessment for Children (REACH)	1	Sri Lanka	1	24-60	Indirect/reported	Caldera 2023
Rapid Assessment of Cognitive and Emotional Regulation (RACER)	1	Ghana	1	48-72	Self-administration	Yuan 2022
Rapid Neurodevelopmental Assessment (RNDA)	2	Bangladesh; Guatemala	3	0-60	Direct assessment	Khan 2010; Khan 2013; Thompson 2015
Receptive and expressive vocabulary test battery (English, Luo and Swahili)	1	Kenya	1	24-83	Direct assessment	Knauer 2019
Saccadic reaction time (SRTs)	1	Sierra Leone	1	7-11.9	Self-administration	Leppänen 2022
Screening for Identification of Oral Language Difficulties by Preschool Teachers (SIOLD)	1	Brazil	1	51-65	Indirect/reported	Puglisi 2020
Screening Test of Early Language Development-Test version (STELD-T)	1	India	1	6-21	Indirect/reported	Pathak 2023
Self-ordered Pointing Task (SOPT)	1	Brazil	1	42-60	Self-administration	Sallum 2017
Sky Search in the Tests of Everyday Attention for Children, Sky Search Task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010

Name assessment tool	Number of countries where the tool was applied	List of countries where the tool was applied	Number of articles where tool was applied	Age range at assessment (in months)	Type of administration	References (first author & year)
Social Skills domain of System-Rating Scales (SSIS-RS)	1	China	1	36-60	Indirect/reported	Wu 2019 (PS)
Spanish-language MacArthur–Bates Communicative Development Inventories short forms (S-CDI SFI and SFII)	1	Mexico	1	8-30	Indirect/reported	Jackson-Maldonado 2013
Strengths and Difficulties Questionnaire (SDQ)	2	Pakistan; Turkey	2	20-51	Indirect/reported	Finch 2018; Dursun 2020
Tap Once Tap Twice Task	1	Uganda	1	55-67	Direct assessment	Nampijja 2010
Test of Gross Motor Development (TGMD-2)	1	Brazil	1	36-60	Direct assessment	Bandeira 2020
Thai Speech and Language Test (TSLT)	1	Thailand	1	3-24	Direct assessment	Prathanee 2008
The Chinese Early Language and Communication Questionnaire	1	China	1	2-48	Indirect/reported	Zhang 2021
The Chinese Intelligence Scale for Young Children	1	China	1	36-84	Direct assessment	Guo 2009
The infant gross motor screening test (IGMST)	1	South Africa	1	6-18	Direct assessment	Hilburn 2011
The language inventory	1	Bangladesh	1	12-18	Indirect/reported	Hamadani 2010
The Road to Health Booklet (RTHB)	1	South Africa	1	6-12	Indirect/reported	van der Linde 2015
The Test of Infant Motor Performance (TIMP)	1	Nepal	1	2-3	Direct assessment	Kvestad 2023
Tobii Professional Studio eye tracking programming	1	Uganda	1	44-65	Direct assessment	Boivin 2017
Vietnamese version of the Intelligibility in Context Scale (ICS-VN).	1	Viet Nam	1	24-71	Indirect/reported	Pham 2017
Vineland Adaptive Behaviour Scale (VABS II)	1	India	1	36-108	Indirect/reported	Kumar 2016
Vocabulary Checklists	2	Malawi, Ghana	1	12-17; 48-72	Indirect/reported	Prado 2018
Wechsler Preschool and Primary Scale of Intelligence, thid edition (WPPSI-III)	1	Pakistan	1		Direct assessment	Rasheed 2018
WHO Gross Motor Milestones	2	Bangladesh; Colombia	3	3-42	Direct assessment, indirect/report	Hamadani 2013; Rubio-Codina 2016; Rubio-Codina 2020
Wisconsin Card Sort Test	1	Uganda	1	55-67	Direct assessment	Nampijja 2010

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmj.com/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES) - selected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
Supplemental Table 6 Number of articles providing psychometric evidence by world region and urban/rural location

Region	Urban	Rural	Urban-rural
Caribbean	2	0	0
Central America	1	3	0
East Asia	10	1	8
Oceania	0	0	1
North Africa and the Middle East	0	0	2
North America	1	0	1
South America	18	0	7
South Asia	13	13	9
Southeast Asia	5	1	3
Sub-Saharan Africa	11	16	14
West Asia	8	0	2

Supplemental Table 7 Number of articles providing psychometric evidence by child development assessment tools (development articles)

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence of measurement invariance	References	
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural			
Angkor Hospital for Children Developmental Milestone Assessment Tool (AHC DMAT)	1	1	0	0	0	1	0	0	0	0	0	Ngoun 2012	
Caregiver Reported Early Development Instruments (CREDI)	2	18	2	1	2	2	2	2	0	1	0	McCoy 2017; Waldman 2021	
Caregiver Reported Early Development Instruments (CREDI), short-form	1	17	1	0	1	1	1	1	0	0	1	McCoy 2018	
Child Behaviour Assessment Instrument (CBAI)	1	1	1	0	1	1	1	0	0	0	0	Samarakkody 2010	
Childhood Nonverbal Communication Scale (CNCS)	1	1	0	0	1	1	1	1	0	1	0	Oryadi-Zanjani 2020	
Chinese Preschool Readiness Scale (CPRS)	1	1	0	0	1	1	1	0	0	1	0	Xie 2019 (Early Educ & Dev)	
Chinese Vocabulary Checklist	1	1	0	0	0	0	0	1	0	0	0	Hao 2008	
DEvelopmental assessment on an E-Platform (DEEP)	1	1	0	0	0	1	0	0	0	0	0	Bhavnani 2019	
Developmental Assessment Tool for Anganwadis (DATA)	1	1	0	0	1	1	0	0	0	1	0	Nair 2009	
Developmental Assessment Tool for Anganwadis (DATA-II)	1	1	0	0	1	1	0	0	0	1	0	Nair 2012	
Developmental Milestone Chart (DMC)	1	1	0	0	0	1	0	0	0	0	0	Scherzer 2009	
Developmental Milestones Checklist (DMC)	1	1	1	0	1	1	1	1	0	1	0	Abubakar 2010	
Early Childhood Development Index 2030 (ECDI203)	1	2	0	0	1	1	1	0	0	1	0	1	Halpin 2024
Early Learning Outcomes Measure (ELOM)	1	1	0	1	1	0	0	0	0	1	0	1	Snelling 2019
Elicited Imitation Test (EIT)	1	1	0	1	0	0	1	0	0	0	0	1	Mehrani 2018

136/bmjopen-2024-096365 © 2024 BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (CC BY-NC 4.0), which permits others to distribute, remix, adapt, and build upon this work, and license their derivative works under the same license, without further permission or charge, provided the original author(s) and the copyright owner(s) are properly cited.

136/bmjopen.2024.096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> at Agence Bibliographique de l'Enseignement Supérieur (ABES). For peer review only - AI training and similar technologies.

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence of measurement invariance	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Escala de Habilidades Sociales (Scale of Social Abilities)	1	1	0	0	1	0	0	0	0	1	0	Lacunza 2009
Farsi Narrative Norms Instrument	1	1	1	1	0	1	0	0	0	0	0	Soleymani 2016
Global Scales for Early Development	1	32	0	0	1	1	0	0	0	0	0	McCray 2023
Grenada Learning and Memory Scalea	1	1	1	1	1	1	1	1	0	1	0	Blackmon 2024
International Development and Early Learning Assessment (IDEA)	1	12	1	1	1	1	1	1	0	0	0	Pisani 2018
Kilifi Developmental Inventory	1	1	1	1	1	1	1	1	0	0	0	Abubakar 2008
Malawi Developmental Assessment Tool (MDAT)	1	1	1	1	0	1	0	1	0	0	0	Gladstone 2010
Malay Preschool Language Assessment Tool (MPLAT)	1	1	1	1	1	1	0	0	0	0	0	Razak 2010
Mongolian Rapid Baby Scale (MORBAS)	1	1	0	0	0	1	1	0	0	0	0	Dagvadorj 2015
Oxford Neurodevelopment Assessment (OX-NDA)	1	1	1	1	1	0	1	0	0	0	0	Fernandes 2022
Questionário para Avaliação do Desenvolvimento Infantil (QAD-PIPAS)	1	1	0	0	0	1	0	0	0	0	0	Venancio 2020
Ragama Early Assessment for Children (REACH)	1	1	0	1	1	1	1	0	0	0	0	Caldera 2023
Rapid Neurodevelopmental Assessment (RNDA)	1	1	0	1	0	1	1	0	1	0	0	Khan 2010
Receptive and expressive vocabulary test battery (English, Luo and Swahili)	1	1	0	1	1	0	1	1	1	0	0	Knauer 2019
Saccadic reaction time (SRTs)	1	1	1	0	0	0	0	1	0	0	0	Leppänen 2022
Screening for Identification of Oral Language Difficulties by Preschool Teachers (SIOLD)	1	1	0	0	0	0	0	0	0	1	0	Puglisi 2020
Screening Test of Early Language Development-Test version (STELD-T)	1	1	0	0	1	0	1	0	0	0	0	Pathak 2023

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence of measurement invariance Copyright, including for uses related to text and data mining, AI training, and similar technologies.	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Thai Speech and Language Test (TSLT)	1	1	0	1	1	1	0	0	0	0	0	Prathanee 2008
The infant gross motor screening test (IGMST)	1	1	1	1	0	1	1	0	0	0	0	Hilburn 2011
The language inventory	1	1	1	0	0	1	1	1	1	0	0	Hamadani 2010
Vocabulary Checklists	1	1	0	0	0	0	1	0	1	0	0	Prado 2018

136/bmjopen-2024-19636 DOI: <https://doi.org/10.1136/bmjopen-2024-19636> © May 2025. Downloaded from <http://bmjopen.bmjjournals.com/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). Evidently for uses related to text and data mining, AI training, and similar technologies.

Supplemental Table 8 Number of articles providing psychometric evidence by child development assessment tools (adaptation articles)

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance (including for uses related to text and data mining, AI training, and similar technologies.)	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Ages & Stages Questionnaire (ASQ)	2	2	0	1	1	2	0	0	0	1	0	Vameghi 2013; Bornman 2010
Ages & Stages Questionnaire Inventory (ASQ:I)	2	2	1	1	2	1	2	1	0	0	0	Xie 2017; Pitchik 2023
Ages & Stages Questionnaire: Social-Emotional (ASQ-SE)	3	3	2	2	2	1	1	0	0	2	0	Kucuker 2011; Anunciação 2019; Bian 2017
Ages & Stages Questionnaire: Social-Emotional, First Edition (ASQ:SE-1)	1	1	1	1	1	0	1	0	0	0	0	Xie 2019 (Journal of Child & Dev Studies)
Ages & Stages Questionnaire-3 (ASQ-3)	14	11	5	6	12	4	5	5	3	4	0	Zirakashvili 2018; Colbert 2021; Rubio-Codina 2016; Rubio-Codina 2020; Shrestha 2019; Hsiao 2016; Filgueiras 2013; van Heerden 2017; Ortiz-Leon 2018; Kvistad 2013; Yue 2019; Yunilda 2022; Bian 2012; Shrestha 2024
Ages & Stages Questionnaires: Social Emotional (ASQ:SE-BR)	1	1	0	0	0	0	0	0	1	0	0	Chen 2017
Alberta Infant Motor Scale (AIMS)	4	1	3	3	2	1	3	1	1	0	0	Saccani 2016; Siegle 2018; de Albuquerque 2018; Valentini 2012
Battelle Developmental Inventory Second Edition, BDI-2	3	2	2	0	2	1	1	1	1	0	0	Rubio-Codina 2016; Rubio-Codina 2020; de Albuquerque 2022
Bayley Infant Neurodevelopment Screener (BINS)	4	8	1	1	2	1	1	1	2	2	0	Guedes 2011; McCarthy 2012; Soysal 2014; Rodriguez 2020
Bayley Scales of Infant and Toddler Development (BSID)	1	1	1	1	1	1	0	1	0	1	0	Azari 2017

Name of assessment tool	Number of articles providing psycho-metric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Bayley Scales of Infant Development, bead threeding task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Bayley Scales of Infant Development, Second Edition (BSID-II)	2	4	0	0	0	1	1	0	0	0	0	Biasini 2015; Parveen 2014
Bayley Scales of Infant Development, Third Edition (BSID-III)	12	12	4	3	7	2	3	3	5	5	2	Rubio-Codina 2016; Rubio-Codina 2020; Hanlon 2016; Ranjitkar 2018; Madaschi 2016; Pendegast 2018; Manandhar 2016; Zakaria 2012; Lohaus 2014; McHenry 2021; Rasheed 2023; Koshy 2024
Bracken School Readiness Assessment, Third Edition (BSRA-3)	1	1	0	0	1	1	0	0	0	0	0	Shah 2013
Brief Infant-Toddler Social and Emotional Assessment (BITSEA)	2	3	1	1	2	0	1	1	0	1	0	Karabekiroglu 2009; Barthel 2021
British Ability Scales – 3rd ed., block design task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Caregiver Reported Early Development Instruments (CREDI), long-form	2	2	0	0	2	0	2	2	0	0	0	Alderman 2021; Li 2020
Caregiver Reported Early Development Instruments (CREDI), short-form	1	1	1	0	1	1	1	1	0	0	0	Altafim 2018
Child Behavior Checklist/1.5-5 (CBCL/1.5-5) and the Caregiver-Teacher Report Form (C-TRF)	1	1	0	0	0	0	0	0	0	1	1	Liu 2011
Child Behavior Scale (CBS)	1	1	0	0	1	0	0	0	0	1	0	Meyer 2011
Child behaviour checklist (CBCL)	1	1	1	0	1	0	0	1	0	1	0	Kariuki 2016
Child Social Preference Scale	1	1	0	0	1	0	0	0	1	1	0	Li 2016
Children's Behavior Questionnaire (CBQ)	1	1	1	0	1	1	1	1	0	1	0	Golmohammadi 2022

136/bmjopen.2024.598365
on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> at Agence Bibliographique de l'Etat (Superior Institute of Statistics - INSTAT) on 13 June 2025. Copyright © 2024 BMJ Publishing Group Ltd. All rights reserved.
For peer review only.

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Chinese version of Urban Infant-Toddler Social and Emotional Assessment	1	1	1	1	1	1	1	0	0	1	0	Jianduan 2009
Communicative Development Inventory (CDI)	1	1	1	0	1	0	1	0	0	0	0	Alcock 2015
Denver Developmental Screening Test for Sri Lankan Children (DDST-SL)	1	1	0	1	0	0	0	0	0	0	0	Wijedasa 2012
Denver Developmental Screening Test (Denver-II)	4	3	3	2	3	2	1	2	1	1	0	Abessa 2016; Rubio-Codina 2016; Rubio-Codina 2020; Lopez Boo 2020
DEvelopmental assessment on an E-Platform (DEEP)	1	1	0	0	0	0	1	0	0	0	0	Mukherjee 2020
Dimensions of Mastery Questionnaire (DMQ-18)	1	1	1	1	1	1	1	0	0	1	0	Shaoli 2019
Early Childhood Development Index (ECDI)	1	1	0	0	1	1	0	0	0	1	0	Enelamah 2024
Early Development Instrument	1	1	0	1	1	1	1	0	1	0	0	Brinkman 2017
Early Human Capability Index (eHCI)	1	7	0	0	1	0	0	0	0	1	0	Sincovich 2019
EDIN II (Evaluación del Desarrollo Integral de la niña y el niño)	1	1	0	0	0	1	0	1	0	0	0	Cordero, 2019
Emotional and Behavioral Screener (EBS)	1	1	0	1	1	0	1	0	0	1	0	Zhang 2018
Emotional Regulation Checklist	1	1	1	0	1	1	1	0	0	1	0	Danişman 2016
Escala de Conductas de Aprendizaje Preescolar (ECAP), a Spanish translation of the Preschool Learning Behaviors Scale (PLBS)	1	1	0	0	0	0	0	0	0	1	0	Hahn 2009
Executive function battery	1	1	0	0	1	1	0	1	1	1	0	Obradovic 2019
Executive Function Touch (EF Touch)	3	2	0	1	1	1	1	3	0	2	0	Willoughby 2019 (Child Neuropsych); Willoughby 2019 (Dev Science); Toor 2022

Name of assessment tool	Number of articles providing psycho-metric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Global Scales of Early Development	2	2	0	1	0	2	1	1	0	0	0	Yibeltal 2025; Jeong 2025
Infant Neurological International Battery	1	1	0	1	0	0	1	0	0	0	0	Soleimani 2007
Intelligibility in Context Scale (ICS) / ICS-Jamaican Creole	1	1	1	1	1	0	1	1	0	1	0	Washington 2017
INTER- GROWTH-21st Project Neurodevelopment Assessment (INTER-NDA)	1	1	1	1	1	0	0	0	0	0	0	Waechter 2022
International Development and Early Learning Assessment (IDE LA)	3	17	1	1	1	0	1	1	1	2	1	Halpin 2019; Pisani 2022; Wolf 2017
Kaufman Survey of Early Academic and Language Skills (K-SEALS)	1	1	1	1	1	1	0	0	0	1	0	Uyanik 2014
Kilifi Developmental Inventory	1	2	0	1	1	0	0	0	1	1	0	Barthel 2022
Kilifi Developmental Inventory, coin box task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Kilifi Picture Vocabulary Test	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Little Developmental Coordination Disorder Questionnaire (Little DCDQ)	1	1	0	0	1	0	1	0	0	0	0	Venter 2015
MacArthur Communicative Development Inventory Level II	1	1	1	1	1	1	1	0	0	0	0	Rasheed 2017
MacArthur-Bates Communicative Development Inventories CDI	1	1	0	0	1	1	0	1	0	0	0	Lara Díaz 2011
MacArthur-Bates Communicative Development Inventories Short-Forms I vocabulary checklist	2	1	2	0	2	0	1	1	1	0	0	Rubio-Codina 2016; Rubio-Codina 2020
MacArthur-Bates Communicative Development Inventories Short-Forms II vocabulary checklist	2	1	2	0	2	0	1	1	1	0	0	Rubio-Codina 2016; Rubio-Codina 2020

136/bmjopen.2024098365
on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> at Agence Bibliographique de l'Enseignement Supérieur (ABES) on 13 June 2025. Copyright © The Author(s). Licensee BMJ Publishing Group Limited. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in other forms, provided the original author(s) and publisher are credited.

Name of assessment tool	Number of articles providing psychometric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Malawi Developmental Assessment Tool (MDAT), renamed Dominican Child Development Screening Tool	1	1	0	1	1	0	0	1	0	0	0	Sánchez-Vincitore 2019
MELQO Measure of Development and Learning Outcomes	1	1	0	0	0	0	0	1	0	1	1	Gomez 2022
Milestones Checklist; Vocabulary Inventory	1	1	0	0	1	1	1	1	0	0	0	Weber 2018
Modified Fagan Test for Infant Intelligence (FTII)	1	1	0	0	0	0	0	0	1	0	0	Familiar-Lopez 2022
Motor Development Scale	1	1	0	0	1	0	0	0	0	1	0	Okuda 2020
Movement Assessment Battery for Children- second edition (MBAC-2)	1	1	1	1	1	1	1	0	0	1	0	Hua 2013
Movement Assessment Battery for Children, balancing on one leg	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Mullen Scales of Early Learning (MSEL)	5	4	1	2	1	1	1	3	3	0	0	Colbert 2020; Koura 2013; Milosavljevic 2019; Bornman 2010; Boivin 2021
NEPSY verbal fluency task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
NEPSY, knock tap game	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
NEPSY, sentence repetition task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Parents' Evaluation of Developmental status (PEDS)	3	3	0	0	0	0	3	0	0	0	0	Abdoola 2021; Chunswan 2016; Mukherjee 2022
Preschool and Kindergarten Behavior Scale (PKBS)	1	1	0	0	1	0	0	0	0	1	0	Dias 2011
Preschool Language Scale-3 (PLS- 3)	1	1	0	0	1	0	0	0	0	1	0	Romero 2013
Preschool Learning Behavior Scale (PLBS)	1	1	0	0	1	0	1	0	0	1	0	Wu 2019 (JPA)
Rapid Assessment of Cognitive and Emotional Regulation (RACER)	1	1	1	1	0	0	1	1	0	0	0	Yuan 2022

Name of assessment tool	Number of articles providing psycho-metric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
Rapid Neurodevelopmental Assessment (RNDA)	2	2	0	2	0	1	2	1	0	0	0	Thompson 2015; Khan 2013
Self-ordered Pointing Task (SOPT)	1		0	0	0	0	1	1	0	0	0	Sallum 2017
Sky Search in the Tests of Everyday Attention for Children, Sky Search Task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Social Skills domain of System-Rating Scales (SSIS-RS)	1	1	0	0	1	0	0	1	0	1	0	Wu 2019 (PS)
Spanish-language MacArthur–Bates Communicative Development Inventories short forms (S-CDI SFI and SFII)	1	1	1	0	0	0	0	1	0	0	0	Jackson-Maldonado 2013
Strengths and Difficulties Questionnaire (SDQ)	2	2	1	1	2	1	2	2	0	1	0	Finch 2018; Dursun 2020
Tap Once Tap Twice Task	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010
Test of Gross Motor Development (TGMD-2)	1	1	0	1	1	0	0	1	0	1	0	Bandeira 2020
The Chinese Early Language and Communication Questionnaire	1	1	0	0	1	0	1	0	0	0	0	Zhang 2021
The Chinese Intelligence Scale for Young Children	1	1	0	0	0	0	0	0	0	1	0	Guo 2009
The Road to Health Booklet (RTHB)	1	1	0	0	0	0	1	0	0	0	0	van der Linde 2015
The Test of Infant Motor Performance (TIMP)	1	1	0	1	1	0	1	0	1	0	0	Kvestad 2023
Tobii Professional Studio eye tracking programming	1	1	0	0	0	0	1	0	0	0	0	Boivin 2017
Vietnamese version of the Intelligibility in Context Scale (ICS-VN)	1	1	0	0	1	0	1	1	0	0	0	Pham 2017
Vineland Adaptive Behaviour Scale (VABS II)	1	1	1	1	0	0	0	1	0	0	0	Kumar 2016
Wechsler Preschool and Primary Scale of Intelligence, thid edition (WPPSI-III)	1	1	0	1	1	1	0	1	0	1	0	Rasheed 2018

136/bmjopen.2024098365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES) - Enseignement Supérieur (ABES). Copyright © 2024 The Author(s). Published by BMJ Publishing Group Ltd on behalf of BMJ Publishing Group Ltd. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits other people and organisations to copy and redistribute the article without prior permission or permission from the author(s) and without additional permission or fees. A copy of this license is available in the Additional Information section in the online version of this article.

Name of assessment tool	Number of articles providing psycho-metric evidence	Number of countries where the tool was applied	Number of articles providing evidence on reliability			Number of articles providing evidence on validity					Number of articles providing evidence on measurement invariance over other groups	References
			Test-retest	Inter-rater	Internal consistency	Content	Concurrent	Convergent	Predictive	Structural		
WHO Gross Motor Milestones	3	2	2	1	2	0	2	2	2	0	0	Hamadani 2013; Rubio-Codina 2016; Rubio-Codina 2020
Wisconsin Card Sort Test	1	1	1	0	1	0	0	1	0	1	0	Nampijja 2010

Supplemental Table 9 Number of articles providing psychometric evidence by article type and early childhood development domain

Child development domain	Total number of articles	Number of articles providing evidence on reliability			Number of articles providing evidence on validity				Number of articles providing evidence on measurement invariance	
		Test-retest	Inter-rater	Internal consistency	Content/face	Concurrent /Criterion	Convergent	Predictive	Structural	Over countries
All articles (N=182 article-tool dyads)										
Cognitive	82	28	31	53	39	35	29	18		3
Language	111	39	38	72	49	51	41	21		4
Motor	104	38	41	62	41	46	38	21		5
Socioemotional/ temperament	52	19	17	37	25	30	16	4		4
Attention/ executive function	22	10	4	14	4	7	16	1		1
Personal-social/ adaptive	49	19	18	32	20	22	16	6		0
Academic/ preacademic	10	2	4	6	3	4	3	1		2
Development articles (N=37 article-tool dyads)										
Cognitive	17	7	7	12	14	10	6	1		2
Language	28	9	10	16	21	16	9	4		2
Motor	20	9	9	13	18	12	7	1		2
Socioemotional/ temperament	13	6	4	10	12	10	4	0		2
Attention/ executive function	3	1	2	3	2	2	1	0		0
Personal-social/ adaptive	10	2	3	6	9	4	2	1		0
Academic/ preacademic	3	1	2	1	2	2	1	0		0
Adaptation articles (N=145 article-tool dyads)										
Cognitive	65	21	24	41	25	25	23	17		1
Language	83	30	28	56	28	35	32	17		2
Motor	84	29	32	49	23	34	31	20		3
Socioemotional/ temperament	39	13	13	27	13	20	12	4		2
Attention/ executive function	19	9	2	11	2	5	15	1		1
Personal-social/ adaptive	39	17	15	26	11	18	14	5		0
Academic/ preacademic	7	1	1	3	1	2	2	1		2

136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> at Agence Bibliographique de l'Enseignement Supérieur (ABES) on 13 June 2025 at 14:44:48. For peer review only - http://bmjopen.bmjjournals.org/site/about/guidelines.xhtml

1
2
3 **Supplemental Table 10** Number of articles by category of risk of bias
4
5
6

	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data
Low risk	47	86	78
Some concerns	12	8	52
High risk	0	1	7
Unable to assess	123	9	44
Not applicable	0	78	1

	Indirectness (are results generalisable?)
Yes	20
Probably yes	82
Probably no	53
No	24
No information	3

Supplemental Table 11 Risk of bias assessment by included article

First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Abdoola 2021	unable to assess	N/A	unable to assess	probably no
Abessa 2016	low risk	N/A	unable to assess	probably yes
Abubakar 2008	unable to assess	Low risk	low risk	probably yes
Abubakar 2010	unable to assess	Some concerns	High risk	No
Alcock 2015	unable to assess	N/A	low risk	Probably yes
Alderman 2021	Low risk	unable to assess	Low risk	Yes
Altafim 2018	low risk	unable to assess	some concerns	probably yes
Anunciação 2019	unable to assess	n/a	some concerns	Yes
Azari 2017	unable to assess	high risk	unable to assess	probably no
Bandeira 2020	unable to assess	low risk	Low risk	probably yes
Barthel 2021	unable to assess	low risk	low risk	probably yes
Barthel 2022	Some concerns	Low risk	Low risk	Probably no
Bhavnani 2019	unable to assess	N/A	low risk	Yes
Bian 2012	unable to assess	N/A	Low risk	no
Bian 2017	unable to assess	n/a	unable to assess	Yes
Biasini 2015	unable to assess	N/A	Some concerns	Probably yes
Blackmon 2024	Some concerns	Low risk	Some concerns	Probably no
Boivin 2017	unable to assess	Low risk	Low risk	No
Boivin 2021	unable to assess	low risk	low risk	probably yes
Bornman 2010	unable to assess	N/A	Low risk	no
Bornman 2010	unable to assess	N/A	low risk	no
Brinkman 2017	some concerns	Low risk	unable to assess	probably yes
Caldera 2023	unable to assess	N/A	unable to assess	Yes
Chen 2017	unable to assess	N/A	some concerns	yes
Chunswan 2016	unable to assess	N/A	unable to assess	probably no
Colbert 2020	low risk	low risk	unable to assess	probably yes
Colbert 2021	unable to assess	low risk	Low risk	No information
Cordero, 2019	unable to assess	N/A	unable to assess	Yes
Dagvadorj 2015	unable to assess	N/A	Low risk	Probably no
Danişman 2016	unable to assess	N/A	high risk	no
de Albuquerque 2018	unable to assess	N/A	Low risk	probably yes
de Albuquerque 2022	unable to assess	N/A	Low risk	probably yes
Dias 2011	unable to assess	N/A	low risk	Probably yes
Dursun 2020	unable to assess	low risk	Some concerns	probably no
Enelamah 2024	Unable to assess	N/A	Low risk	Yes
Familiar-Lopez 2022	unable to assess	Some concerns	Some concerns	No
Fernandes 2022	unable to assess	low risk	low risk	probably yes
Filgueiras 2013	unable to assess	N/A	unable to assess	probably yes

First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Finch 2018	low risk	Low risk	low risk	probably no
Gladstone 2010	unable to assess	low risk	some concerns	Probably yes
Golmohammadi 2022	unable to assess	Low risk	low risk	probably no
Gomez 2022	unable to assess	Low risk	Low risk	Probably yes
Guedes 2011	unable to assess	low risk	unable to assess	probably no
Guo 2009	unable to assess	N/A	unable to assess	Probably yes
Hahn 2009	unable to assess	N/A	high risk	probably yes
Halpin 2019	unable to assess	N/A	unable to assess	yes
Halpin 2024	Low risk	Low risk	Some concerns	Yes
Hamadani 2010	low risk	low risk	high risk	probably yes
Hamadani 2013	unable to assess	low risk	low risk	Probably yes
Hanlon 2016	Some concerns	Some concerns	Low risk	Probably no
Hao 2008	unable to assess	low risk	unable to assess	no
Hilburn 2011	low risk	N/A	unable to assess	Probably yes
Hsiao 2016	some concerns	low risk	low risk	probably yes
Hua 2013	unable to assess	N/A	Some concerns	probably yes
Jackson-Maldonado 2013	unable to assess	low risk	Low risk	Yes
Jeong 2025	Low risk	Low risk	Low risk	Probably yes
Jianduan 2009	unable to assess	N/A	unable to assess	probably yes
Karabekiroglu 2009	unable to assess	N/A	low risk	no
Kariuki 2016	Some concerns	Low risk	Some concerns	Probably yes
Khan 2010	unable to assess	Some concerns	low risk	Probably no
Khan 2013	unable to assess	N/A	unable to assess	probably yes
Knauer 2019	unable to assess	low risk	some concerns	probably yes
Koshy 2024	Low risk	Low risk	Low risk	Probably no
Koura 2013	low risk	unable to assess	unable to assess	probably no
Kucuker 2011	unable to assess	N/A	some concerns	Probably yes
Kumar 2016	unable to assess	some concerns	low risk	probably no
Kvestad 2013	Low risk	N/A	some concerns	Probably yes
Kvestad 2023	low risk	low risk	some concerns	probably no
Lacunza 2009	unable to assess	N/A	unable to assess	probably no
Lara Díaz 2011	unable to assess	low risk	Low risk	Probably yes
Leppänen 2022	unable to assess	low risk	high risk	probably no
Li 2016	unable to assess	Low risk	Some concerns	No
Li 2020	unable to assess	unable to assess	low risk	probably yes
Liu 2011	unable to assess	N/A	Some concerns	probably yes
Lohaus 2014	unable to assess	Low risk	Some concerns	No
Lopez Boo 2020	Some concerns	low risk	some concerns	probably yes
Madashchi 2016	unable to assess	N/A	Some concerns	probably yes
Manandhar 2016	low risk	N/A	unable to assess	probably no
McCarthy 2012	low risk	N/A	unable to assess	probably yes
McCoy 2017	Low risk	low risk	low risk	yes

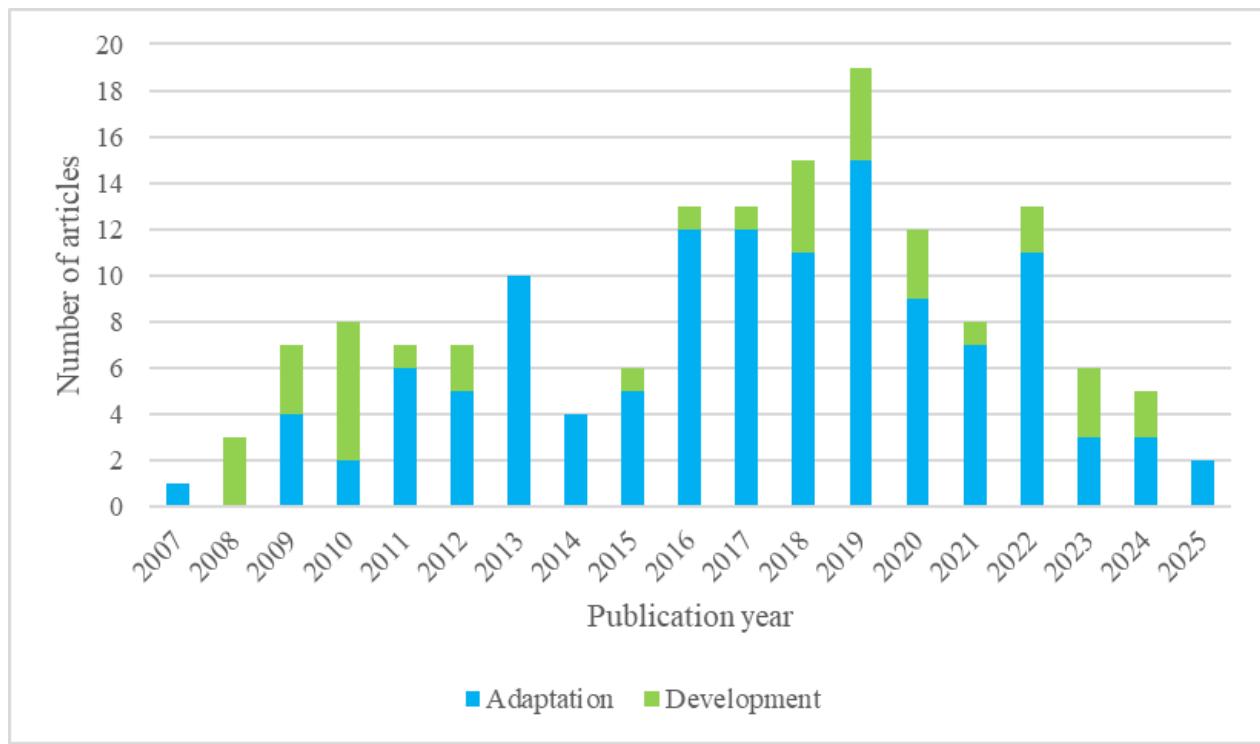
First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
McCoy 2018	unable to assess	low risk	low risk	yes
McCray 2023	unable to assess	N/A	Low risk	Probably yes
McHenry 2021	low risk	low risk	some concerns	Probably no
Mehrani 2018	unable to assess	N/A	Low risk	Probably no
Meyer 2011	unable to assess	N/A	unable to assess	Probably no
Milosavljevic 2019	low risk	low risk	some concerns	no
Mukherjee 2020	unable to assess	N/A	low risk	No
Mukherjee 2022	unable to assess	N/A	low risk	probably yes
Nair 2009	unable to assess	N/A	unable to assess	probably yes
Nair 2012	unable to assess	N/A	unable to assess	probably yes
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Nampijja 2010	unable to assess	low risk	low risk	probably no
Ngoun 2012	unable to assess	N/A	unable to assess	yes
Obradovic 2019	unable to assess	Low risk	low risk	probably yes
Okuda 2020	unable to assess	N/A	unable to assess	probably yes
Ortiz-Leon 2018	unable to assess	N/A	some concerns	probably no
Oryadi-Zanjani 2020	some concerns	N/A	low risk	probably yes
Parveen 2014	unable to assess	N/A	N/A	no
Pathak 2023	unable to assess	N/A	low risk	probably no
Pendergast 2018	low risk	N/A	Some concerns	yes
Pham 2017	unable to assess	low risk	low risk	probably no
Pisani 2018	unable to assess	low risk	unable to assess	yes
Pisani 2022	unable to assess	low risk	some concerns	probably yes
Pitchik 2023	low risk	low risk	low risk	probably yes
Prado 2018	Low risk	Low risk	low risk	No
Prathanee 2008	low risk	N/A	low risk	probably yes
Puglisi 2020	Some concerns	N/A	low risk	Probably yes
Ranjitkar 2018	low risk	low risk	low risk	probably yes
Rasheed 2017	some concerns	N/A	unable to assess	probably no
Rasheed 2018	low risk	Low risk	unable to assess	probably yes
Rasheed 2023	unable to assess	Low risk	unable to assess	no information
Razak 2010	unable to assess	N/A	low risk	probably no
Rodriguez 2020	low risk	unable to assess	unable to assess	probably yes

First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Romero 2013	unable to assess	N/A	unable to assess	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	low risk	Probably yes
Rubio-Codina 2016	low risk	Low risk	Some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	Some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	unable to assess	Low risk	some concerns	Probably yes
Rubio-Codina 2020	low risk	Low risk	some concerns	Probably yes
Saccani 2016	low risk	N/A	unable to assess	yes
Sallum 2017	unable to assess	low risk	some concerns	probably yes
Samarakkody 2010	unable to assess	N/A	unable to assess	yes
Sánchez-Vincitore 2019	low risk	Low risk	low risk	probably no
Scherzer 2009	unable to assess	N/A	unable to assess	probably yes
Shah 2013	unable to assess	n/a	unable to assess	no
Shaoli 2019	unable to assess	n/a	low risk	probably yes
Shrestha 2019	some concerns	low risk	some concerns	probably no
Shrestha 2024	Low risk	Low risk	Some concerns	Probably no
Siegle 2018	unable to assess	N/A	some concerns	probably no
Sincovich 2019	unable to assess	N/A	some concerns	probably no
Snelling 2019	unable to assess	N/A	some concerns	probably yes
Soleimani 2007	low risk	N/A	Some concerns	no information
Soleymani 2016	unable to assess	N/A	Low risk	No
Soysal 2014	unable to assess	some concerns	some concerns	no
Thompson 2015	low risk	unable to assess	unable to assess	probably no
Toor 2022	unable to assess	Low risk	some concerns	No
Uyanik 2014	unable to assess	N/A	Some concerns	Probably no
Valentini 2012	unable to assess	Low risk	unable to assess	Probably no
Vameghi 2013	unable to assess	N/A	high risk	probably no
van der Linde 2015	unable to assess	low risk	low risk	probably no
van Heerden 2017	low risk	N/A	unable to assess	no
Venancio 2020	unable to assess	N/A	Some concerns	Probably yes
Venter 2015	unable to assess	N/A	high risk	probably no
Waechter 2022	low risk	N/A	unable to assess	probably yes
Waldman 2021	unable to assess	Some concerns	Low risk	Probably yes

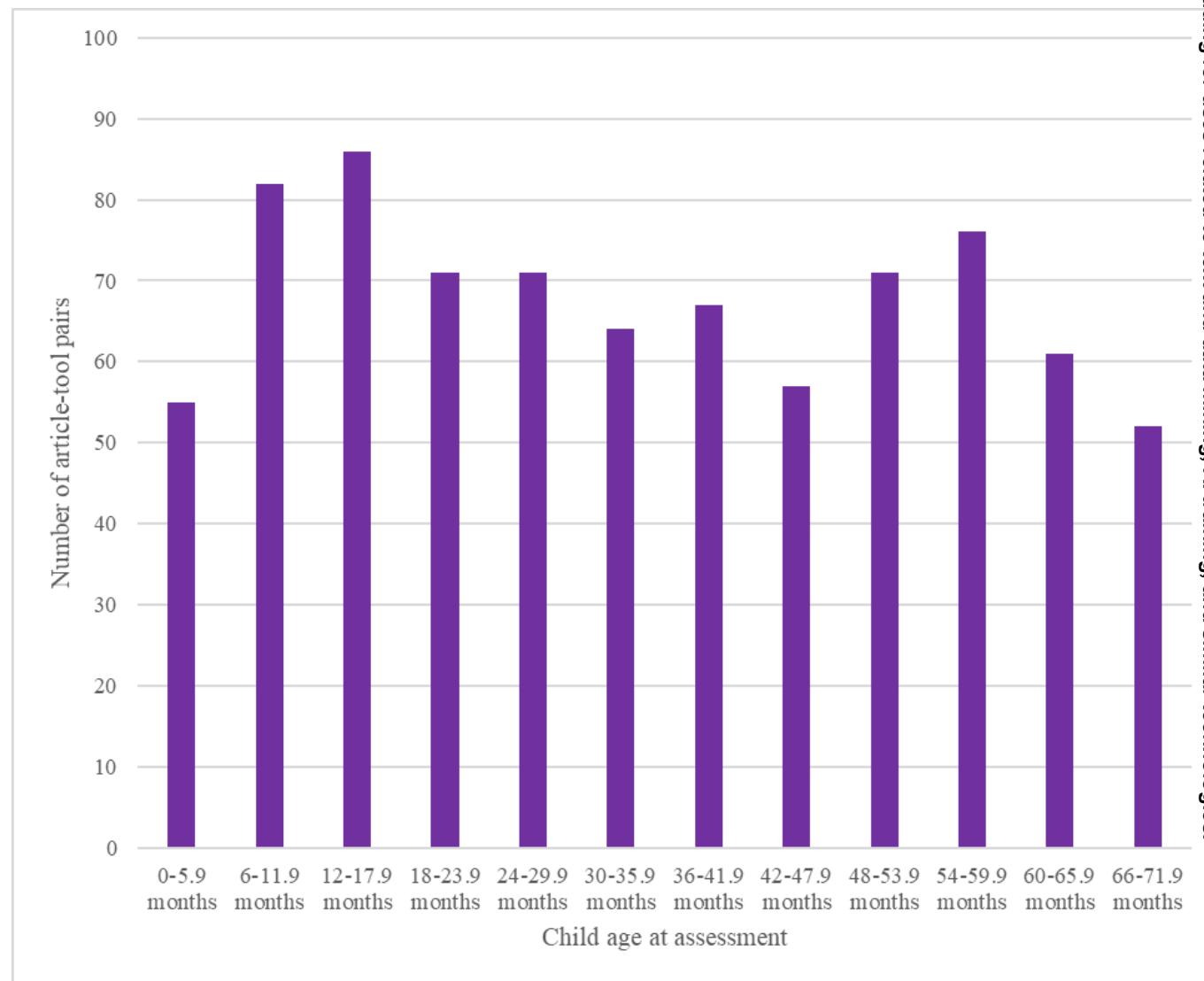
First author & year	Risk of bias due to training or administration	Risk of bias due to selective reporting	Risk of bias due to missing data	Are results generalizable?
Washington 2017	unable to assess	Low risk	Some concerns	No
Weber 2018	unable to assess	low risk	some concerns	probably yes
Wijedasa 2012	low risk	N/A	low risk	yes
Willoughby 2019 (Child Neuropsych)	low risk	low risk	low risk	probably yes
Willoughby 2019 (Dev Science)	Low risk	Low risk	Some concerns	Probably yes
Wolf 2017	Some concerns	low risk	unable to assess	no
Wu 2019 (JPA)	unable to assess	n/a	unable to assess	Probably yes
Wu 2019 (PS)	unable to assess	Low risk	low risk	probably yes
Xie 2017	low risk	N/A	some concerns	no
Xie 2019 (Early Educ & Dev)	unable to assess	N/A	Some concerns	Probably no
Xie 2019 (Journal of Child & Dev Studies)	unable to assess	N/A	some concerns	probably yes
Yibeltal 2025	Low risk	Low risk	Low risk	Probably no
Yuan 2022	low risk	low risk	low risk	probably yes
Yue 2019	low risk	unable to assess	low risk	yes
Yunilda 2022	unable to assess	unable to assess	low risk	probably no
Zakaria 2012	unable to assess	N/A	unable to assess	no
Zhang 2018	unable to assess	N/A	some concerns	Probably no
Zhang 2021	unable to assess	n/a	unable to assess	Probably yes
Zirakashvili 2018	unable to assess	Some concerns	Some concerns	Probably yes

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

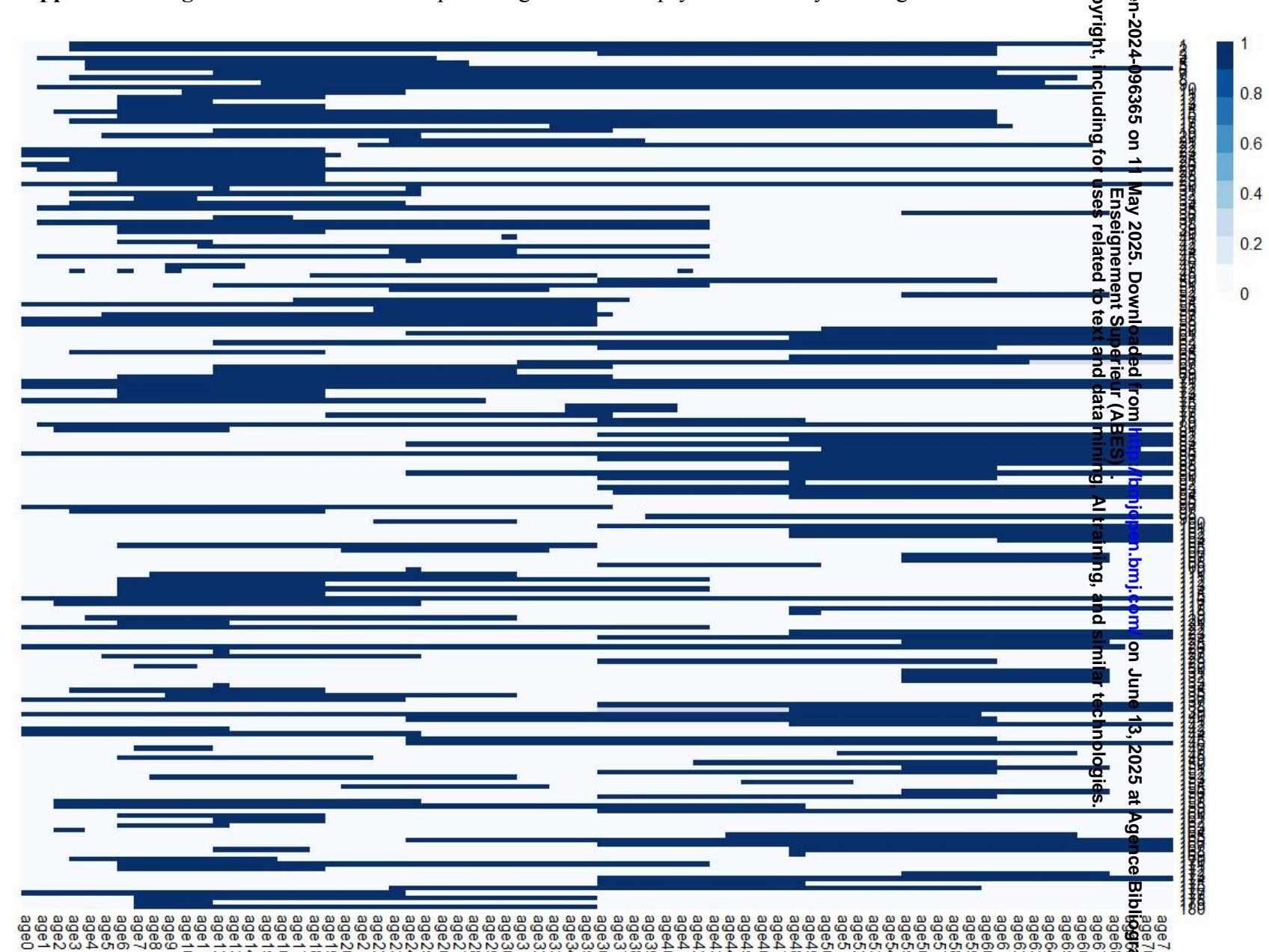
Supplemental Figure 1 Number of articles providing evidence on psychometrics by type of article



Supplemental Figure 2 Number of articles and tools providing psychometric evidence by child age group



136/bmjopen-2024-096365 on 11 May 2025. Downloaded from <http://bmjopen.bmjjournals.org/> on June 13, 2025 at Agence Bibliographique de l'Enseignement Supérieur (ABES). Excluded by copyright. Excluding for uses related to text and data mining, AI training, and similar technologies.

Supplemental Figure 3 Number of articles providing evidence on psychometrics by child age at assessment

Note: Y-axis lists each article-tool combination included in the systematic review ordered by tool in alphabetical order.

References of all included articles

- Abdoola, S., Swanepoel, D. W., Van Der Linde, J., & Glascoe, F. P. (2021). Detecting developmental delays in infants from a low-income South African community: comparing the BSID-III and PEDS tools. *Early Child Development and Care*, 191(4), 545–554. <https://doi.org/10.1080/03004430.2019.1628027>
- Abessa, T. G., Worku, B. N., Kibebew, M. W., Valy, J., Lemmens, J., Thijs, H., Yimer, W. K., Kolsteren, P., & Granitzer, M. (2016). Adaptation and standardization of a Western tool for assessing child development in non-Western low-income context. *BMC Public Health*, 16(1), 1–13. <https://doi.org/10.1186/s12889-016-3288-2>
- Abubakar, A., Holding, P., Van De Vijver, F., Bomu, G., & Van Baar, A. (2010). Developmental monitoring using caregiver reports in a resource-limited setting: The case of Kilifi, Kenya. *Acta Paediatrica, International Journal of Paediatrics*, 99(2), 291–297. <https://doi.org/10.1111/j.1651-2227.2009.01561.x>
- Abubakar, A., Holding, P., van Baar, A., Newton, C. R. J. C., & van de Vijver, F. J. R. (2008). Monitoring psychomotor development in a resource-limited setting: An evaluation of the Kilifi Developmental Inventory. *Annals of Tropical Paediatrics*, 28(3), 217–226. <https://doi.org/10.1179/146532808X335679>
- Albuquerque, P. L. de, Guerra, M. Q. de F., Lima, M. de C., & Eickmann, S. H. (2018). Concurrent validity of the Alberta Infant Motor Scale to detect delayed gross motor development in preterm infants: A comparative study with the Bayley III. *Developmental Neurorehabilitation*, 21(6), 408–414. <https://doi.org/10.1080/17518423.2017.1323974>
- Alcock, K. J., Rimba, K., Holding, P., Kitsao-Wekulo, P., Abubakar, A., & Newton, C. R. J. C. (2015). Developmental inventories using illiterate parents as informants: Communicative Development Inventory (CDI) adaptation for two Kenyan languages. *Journal of Child Language*, 42(4), 763–785. <https://doi.org/10.1017/S0305000914000403>
- Alderman, H., Friedman, J., Ganga, P., Kak, M., & Rubio-Codina, M. (2021). Assessing the performance of the Caregiver Reported Early Development Instruments (CREDI) in rural India. *Annals of the New York Academy of Sciences*, 1492(1), 58–72. <https://doi.org/10.1111/nyas.14543>
- Altafim, E. R. P., McCoy, D. C., Brentani, A., Escobar, A. M. de U., Grisi, S. J. F. E., & Fink, G. (2020). Measuring early childhood development in Brazil: validation of the Caregiver Reported Early Development Instruments (CREDI). *Jornal de Pediatria*, 96(1), 66–75. <https://doi.org/10.1016/j.jped.2018.07.008>
- Anunciação, L., Squires, J., Clifford, J., & Landeira-Fernandez, J. (2019). Confirmatory analysis and normative tables for the Brazilian Ages and Stages Questionnaires: Social–Emotional. *Child: Care, Health and Development*, 45(3), 387–393. <https://doi.org/10.1111/cch.12649>

- Azari, N., Soleimani, F., Vameghi, R., Sajedi, F., Shahshahani, S., Karimi, H., Kraskian, A., Shahrokhi, A., Teymour, R., & Gharib, M. (2017). A psychometric study of the Bayley scales of infant and toddler development in Persian language children. *Iranian Journal of Child Neurology*, 11(1), 50–56.
- Bandeira, P. F. R., Duncan, M., Pessoa, M. L., Soares, I., Da Silva, L., Mota, J., & Martins, C. (2020). Tgmd-2 short version: Evidence of validity and associations with sex, age, and bmi in preschool children. *Journal of Motor Learning and Development*, 8(3), 528–543. <https://doi.org/10.1123/JMLD.2019-0040>
- Barthel, D., Kriston, L., Guo, N., Yao, E. D., Armel, K. E. J., Fordjour, D., Mohammed, Y., Kotchi, C. E. B., Eberhardt, K. A., Hinz, R., Ehrhardt, S., & Bindt, C. (2022). Psychometric properties of the Brief Infant-Toddler Social and Emotional Assessment (BITSEA) in two-year-old Ivorian and Ghanaian children. *Applied Developmental Science*, 26(3), 567–577. <https://doi.org/10.1080/10888691.2021.1931216>
- Barthel, D., Kriston, L., Yao, E. D., Fordjour, D., Armel, K. E. J., Eberhardt, K. A., Hinz, R., Ehrhardt, S., & Bindt, C. (2022). Psychomotor development in two-year-old Ivorian and Ghanaian children – Psychometric properties of the Kilifi Developmental Inventory. *British Journal of Developmental Psychology*, 40(4), 471–486. <https://doi.org/10.1111/bjdp.12428>
- Bhavnani, S., Mukherjee, D., Dasgupta, J., Verma, D., Parameshwaran, D., Divan, G., Sharma, K. K., Thiagarajan, T., & Patel, V. (2019). Development, feasibility and acceptability of a gamified cognitive DEvelopmental assessment on an E-Platform (DEEP) in rural Indian pre-schoolers—a pilot study. *Global Health Action*, 12(1). <https://doi.org/10.1080/16549716.2018.1548005>
- Bian, X., Xie, H., Squires, J., & Chen, C. Y. (2017). Adapting a Parent-Completed, Socioemotional Questionnaire in China: the Ages & Stages Questionnaires: Social-Emotional. *Infant Mental Health Journal*, 38(2), 258–266. <https://doi.org/10.1002/imhj.21626>
- Bian, X., Yao, G., Squires, J., Hoselton, R., Chen, C. I., Murphy, K., Wei, M., & Fang, B. (2012). Translation and use of parent-completed developmental screening test in Shanghai. *Journal of Early Childhood Research*, 10(2), 162–175. <https://doi.org/10.1177/1476718X11430071>
- Biasini, F. J., De Jong, D., Ryan, S., Thorsten, V., Bann, C., Bellad, R., Mahantshetti, N. S., Dhaded, S. M., Pasha, O., Chomba, E., Goudar, S. S., Carlo, W. A., & McClure, E. (2015). Development of a 12 month screener based on items from the Bayley II Scales of Infant Development for use in Low Middle Income countries. *Early Human Development*, 91(4), 253–258. <https://doi.org/10.1016/j.earlhumdev.2015.02.001>
- Blackmon, K., Evans, R., Mohammed, L., Burgen, K. S., Ingraham, E., Punch, B., Isaac, R., Murray, T., Noel, J., Belmar-Roberts, C., Waechter, R., & Landon, B. (2024). The Grenada Learning and Memory Scale: Psychometric features and normative data in Caribbean

1
2
3 preschool children. *Journal of the International Neuropsychological Society*, 2024, 856–
4 866. <https://doi.org/10.1017/S1355617724000481>

5
6 Boivin, M. J., Weiss, J., Chhaya, R., Seffren, V., Awadu, J., Sikorskii, A., & Giordani, B. (2017).
7 The feasibility of automated eye tracking with the early childhood vigilance test of attention
8 in younger HIV-exposed Ugandan children. *Neuropsychology*, 31(5), 525–534.
9 <https://doi.org/10.1037/neu0000382>

10
11 Boivin, M. J., Zoumenou, R., Sikorskii, A., Fievet, N., Alao, J., Davidson, L., Cot, M.,
12 Massougbedji, A., & Bodeau-Livinec, F. (2021). Neurodevelopmental assessment at one
13 year of age predicts neuropsychological performance at six years in a cohort of West
14 African Children. *Child Neuropsychology*, 27(4), 548–571.
15 <https://doi.org/10.1080/09297049.2021.1876012>

16
17 Bornman, J., Sevcik, R. A., Romski, M. A., & Pae, H. K. (2010). Successfully translating
18 language and culture when adapting assessment measures. *Journal of Policy and Practice in
19 Intellectual Disabilities*, 7(2), 111–118. <https://doi.org/10.1111/j.1741-1130.2010.00254.x>

20
21 Brinkman, S. A., Kinnell, A., Maika, A., Hasan, A., Jung, H., & Pradhan, M. (2017). Validity
22 and Reliability of the Early Development Instrument in Indonesia. *Child Indicators
23 Research*, 10(2), 331–352. <https://doi.org/10.1007/s12187-016-9372-4>

24
25 Caldera, A., Wickremasinghe, A. R., Muttiah, N., Godamunne, P. K. S., Jayasena, B.,
26 Chathurika, L. K. E., Perera, K. M. N., Mendis, M., Tilakarathne, D., Peiris, M. K. R.,
27 Wijesinghe, T., Senarathna, N., Saubhagya, W. D. L., Chandraratne, M., & Sumanasena,
28 S. (2023). REACH for the preschoolers; a developmental assessment tool for 2–5 year old
29 children in Sri Lanka. *BMC Pediatrics*, 23(1), 80. <https://doi.org/10.1186/s12887-023-03895-5>

30
31 Chen, C. Y., Xie, H., Filgueiras, A., Squires, J., Anunciação, L., & Landeira-Fernandez, J.
32 (2017). Examining the Psychometric Properties of the Brazilian Ages & Stages
33 Questionnaires-Social-Emotional: Use in Public Child Daycare Centers in Brazil. *Journal of
34 Child and Family Studies*, 26(9), 2412–2425. <https://doi.org/10.1007/s10826-017-0770-0>

35
36 Chunsuwan, I., Hansakunachai, T., & Pornsamrit, S. (2016). Parent Evaluation of Developmental
37 Status (PEDS) in screening: The Thai experience. *Pediatrics International*, 58(12), 1277–
38 1283. <https://doi.org/10.1111/ped.13055>

39
40 Colbert, A. M., Connery, A. K., Lamb, M. M., Bauer, D., Olson, D., Paniagua-Avila, A.,
41 Martínez, M. A., Arroyave, P., Hernández, S., Mirella Calvimontes, D., Bolaños, G. A., El
42 Sahly, H. M., Muñoz, F. M., & Asturias, E. J. (2021). Caregiver rating of early childhood
43 development: Reliability and validity of the ASQ-3 in rural Guatemala. *Early Human
44 Development*, 161(November 2020). <https://doi.org/10.1016/j.earlhumdev.2021.105453>

45
46 Colbert, A. M., Lamb, M. M., Asturias, E. J., Muñoz, F. M., Bauer, D., Arroyave, P., Hernández,
47 S., Martínez, M. A., Paniagua-Avila, A., Olson, D., Calvimontes, D. M., Bolaños, G. A., El

- Sahly, H. M., & Connery, A. K. (2020). Reliability and Validity of an Adapted and Translated Version of the Mullen Scales of Early Learning (AT-MSEL) in Rural Guatemala. *Child: Care, Health and Development*, 46(3), 327–335.
<https://doi.org/10.1111/cch.12748>
- Dagvadorj, A., Takehara, K., Bavuusuren, B., Morisaki, N., Gochoo, S., & Mori, R. (2015). The quick and easy Mongolian Rapid Baby Scale shows good concurrent validity and sensitivity. *Acta Paediatrica, International Journal of Paediatrics*, 104(3), e94–e99.
<https://doi.org/10.1111/apa.12884>
- Danişman, Ş., Dereli Iman, E., Demircan, Z. A., & Yaya, D. (2016). Una aproximación al enfoque Reggio Emilia: Claves para entender por qué motiva a los estudiantes. *Electronic Journal of Research in Educational Psychology*, 14(3), 534–556.
- de Albuquerque, K. A., da Cunha, A. C. B., & Berkovits, M. D. (2022). Cross-cultural adaptation of the Battelle Development Inventory, 2nd Edition for Brazil. *Estudos de Psicologia*, 39, 1–13. <https://doi.org/10.1590/1982-0275202239e190170>
- Dias, T. P., Freitas, L. C., Del Prette, Z. A. P., & Del Prette, A. (2011). Validação da escala de comportamentos sociais de pré-escolares para o Brasil. *Psicologia Em Estudo*, 16(3), 447–457. <https://doi.org/10.1590/S1413-73722011000300012>
- Dursun, O. B., Öğütlü, H., & Esin, İ. S. (2020). Psychometric properties of Turkish version of the strength and difficulties questionnaire for age 2-4. *Noropsikiyatri Arsivi*, 57(1), 44–49.
<https://doi.org/10.29399/npa.23674>
- Enelamah, N. V., Rao, S., Lombe, M., Yu, M., Newransky, C., Villodas, M. L., Foell, A., & Neibitt, V. (2024). Examining the Utility of the Early Childhood Development Index (ECDI) among Children in the Nigeria Context. *Children*, 11(3).
<https://doi.org/10.3390/children11030361>
- Familiar-Lopez, I., Sikorskii, A., Chhaya, R., Holmes, A., Arima, E. G., Caesar, O. J., Nakasujja, N., & Boivin, M. J. (2023). Predictive validation of Ugandan infant eye-tracking test for memory of human faces. *Child Neuropsychology*, 29(3), 486–502.
<https://doi.org/10.1080/09297049.2022.2099537>
- Fernandes, M., Bassani, D., Albernaz, E., Bertoldi, A. D., Silveira, M. F., Matijsevich, A., Anselmi, L., Cruz, S., Halal, C. S., Tovo-Rodrigues, L., Cruz, G. I. N., Metgud, D., & Santos, I. S. (2022). Construction and validation of the Oxford Neurodevelopment Assessment (OX-NDA) in 1-year-old Brazilian children. *BMC Pediatrics*, 22(1), 1–12.
<https://doi.org/10.1186/s12887-022-03794-1>
- Ferreira-Vasques, A. T., Santos, C. F., & Lamônica, D. A. C. (2019). Transcultural adaptation process of the Griffiths-III Mental Development Scale. *Child: Care, Health and Development*, 45(3), 403–408. <https://doi.org/10.1111/cch.12664>

- 1
2
3 Filgueiras, A., Pires, P., Maissonette, S., & Landeira-Fernandez, J. (2013). Psychometric
4 properties of the Brazilian-adapted version of the Ages and Stages Questionnaire in public
5 child daycare centers. *Early Human Development*, 89(8), 561–576.
6 https://doi.org/10.1016/j.earlhhumdev.2013.02.005
7
8 Finch, J. E., Yousafzai, A. K., Rasheed, M., & Obradović, J. (2018). Measuring and
9 understanding social-emotional behaviors in preschoolers from rural Pakistan. *PLoS ONE*,
10 13(11), 1–22. https://doi.org/10.1371/journal.pone.0207807
11
12 Flórez Romero, R., Castro Martínez, J., & Arias Velandia, N. (2013). Propiedades psicométricas
13 de la escala de lenguaje para preescolares (PLS-3) colombianos TT - Psychometric
14 Properties of the Pre-school Language Scale (PLS-3) with Colombian Children TT -
15 Propriedades psicométricas da Escala de linguagem para pré-escola. *Pensamiento
16 Psicológico*, 11(1), 131–140.
17
18 Gladstone, M., Lancaster, G. a., Umar, E., Nyirenda, M., Kayira, E., van den Broek, N. R., &
19 Smyth, R. L. (2010). The Malawi Developmental Assessment Tool (MDAT): The Creation,
20 Validation, and Reliability of a Tool to Assess Child Development in Rural African
21 Settings. *PLoS Medicine*, 7(5), e1000273. https://doi.org/10.1371/journal.pmed.1000273
22
23 Golmohammadi, G., Zarifian, T., Qhanbari, S., Bakhshi, E., & Sakhai, F. (2022). Cross-cultural
24 adaptation and validation of the Persian version of Children's Behavior Questionnaire in
25 Iranian children. *Current Psychology*, 41(6), 4041–4053. https://doi.org/10.1007/s12144-
26 020-00918-7
27
28 Gomez, J. A., Brown, J. L., & Spier, E. (2022). Factor structure, measurement invariance, and
29 construct validity of MELQO measure of child development and early learning in Laos.
30 *International Journal of School and Educational Psychology*, 10(2), 181–207.
31 https://doi.org/10.1080/21683603.2020.1791770
32
33 Guedes, D. Z., Primi, R., & Kopelman, B. I. (2011). BINS validation - Bayley
34 neurodevelopmental screener in Brazilian preterm children under risk conditions. *Infant
35 Behavior and Development*, 34(1), 126–135. https://doi.org/10.1016/j.infbeh.2010.11.001
36
37 Guo, B., Aveyard, P., & Dai, X. (2009). The Chinese Intelligence Scale for Young Children.
38 *Educational and Psychological Measurement*, 69(3), 459–474.
39 https://doi.org/10.1177/0013164409332209
40
41 Hahn, K. R., Schaefer, B. A., Merino, C., & Worrell, F. C. (2009). The factor structure of
42 preschool learning behaviors scale scores in peruvian children. *Canadian Journal of School
43 Psychology*, 24(4), 318–331. https://doi.org/10.1177/0829573509349015
44
45 Halpin, P. F., de Castro, E. F., Petrowski, N., & Cappa, C. (2024). Monitoring early childhood
46 development at the population level: The ECDI2030. *Early Childhood Research Quarterly*,
47 67(November 2023), 1–12. https://doi.org/10.1016/j.ecresq.2023.11.004
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Halpin, P. F., Wolf, S., Yoshikawa, H., Rojas, N., Kabay, S., Pisani, L., & Dowd, A. J. (2019).
4 Measuring early learning and development across cultures: Invariance of the IDELA across
5 five countries. *Developmental Psychology*, 55(1), 23–37.
6 https://doi.org/10.1037/dev0000626
7
8 Hamadani, J. D., Baker-Henningham, H., Tofail, F., Mehrin, F., Huda, S. N., & Grantham-
9 McGregor, S. M. (2010). Validity and Reliability of Mothers' Reports of Language
10 Development in 1-Year-Old Children in a Large-Scale Survey in Bangladesh. *Food and*
11 *Nutrition Bulletin*, 31(2_suppl2), S198–S206.
12 https://doi.org/10.1177/15648265100312S212
13
14 Hamadani, J. D., Tofail, F., Cole, T., & Grantham-McGregor, S. (2013). The relation between
15 age of attainment of motor milestones and future cognitive and motor development in
16 Bangladeshi children. *Maternal and Child Nutrition*, 9(SUPPL. 1), 89–104.
17 https://doi.org/10.1111/mcn.12020
18
19 Hanlon, C., Medhin, G., Worku, B., Tomlinson, M., Alem, A., Dewey, M., & Prince, M. (2016).
20 Adapting the Bayley Scales of infant and toddler development in Ethiopia: evaluation of
21 reliability and validity. *Child: Care, Health and Development*, 42(5), 699–708.
22 https://doi.org/10.1111/cch.12371
23
24 Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese.
25 *Behavior Research Methods*, 40(3), 728–733. https://doi.org/10.3758/BRM.40.3.728
26
27 Hilburn, N., Potterton, J., Stewart, A., & BeckerC, P. (2011). The development of a screening
28 tool to evaluate gross motor function in HIV-infected infants. *AIDS Care - Psychological*
29 *and Socio-Medical Aspects of AIDS/HIV*, 23(12), 1619–1625.
30 https://doi.org/10.1080/09540121.2011.579942
31
32 Hsiao, C., Richter, L., Makusha, T., Matafwali, B., van Heerden, A., & Mabaso, M. (2017). Use
33 of the ages and stages questionnaire adapted for South Africa and Zambia. *Child: Care,*
34 *Health and Development*, 43(1), 59–66. https://doi.org/10.1111/cch.12413
35
36 Hua, J., Gu, G., Meng, W., & Wu, Z. (2013). Age band 1 of the Movement Assessment Battery
37 for Children-Second Edition: Exploring its usefulness in mainland China. *Research in*
38 *Developmental Disabilities*, 34(2), 801–808. https://doi.org/10.1016/j.ridd.2012.10.012
39
40 Jackson-Maldonado, D., Marchman, V. A., & Fernald, L. C. H. (2013). Short-form versions of
41 the Spanish MacArthur-Bates Communicative Development Inventories. *Applied*
42 *Psycholinguistics*, 34(4), 837–868. https://doi.org/10.1017/S0142716412000045
43
44 Jeong, J., McCann, J. K., Onyango, S., & Ochieng, M. (2025). Validation of the Global Scales of
45 Early Development (GSED) tool in rural Western Kenya. *BMC Public Health*, 25(1).
46 https://doi.org/10.1186/s12889-025-21801-9
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- Jianduan, Z., Huishan, W., Shuhua, S., Xiaonan, H., Guoyan, L., Guangli, L., & Junxin, S. (2009). Reliability and validity of standardized Chinese version of Urban Infant-Toddler Social and Emotional Assessment. *Early Human Development*, 85(5), 331–336. <https://doi.org/10.1016/j.earlhumdev.2008.12.012>
- Karabekiroglu, K., Rodopman-Arman, A., Ay, P., Ozkesen, M., Akbas, S., Tasdemir, G. N., Boke, O., & Peksen, Y. (2009). The reliability and validity of the Turkish version of the brief infant-toddler social emotional assessment (BITSEA). *Infant Behavior and Development*, 32(3), 291–297. <https://doi.org/10.1016/j.infbeh.2009.03.003>
- Kariuki, S. M., Abubakar, A., Murray, E., Stein, A., & Newton, C. R. J. C. (2016). Evaluation of psychometric properties and factorial structure of the pre-school child behaviour checklist at the Kenyan Coast. *Child and Adolescent Psychiatry and Mental Health*, 10(1), 1–10. <https://doi.org/10.1186/s13034-015-0089-9>
- Khan, N. Z., Muslima, H., Shilpi, A. B., Begum, D., Parveen, M., Akter, N., Ferdous, S., Nahar, K., McConachie, H., & Darmstadt, G. L. (2013). Validation of rapid neurodevelopmental assessment for 2- to 5-year-old children in bangladesh. *Pediatrics*, 131(2). <https://doi.org/10.1542/peds.2011-2421>
- Khan, N. Z., Muslima, H., Begum, D., Shilpi, A. B., Akhter, S., Bilkis, K., Begum, N., Parveen, M., Ferdous, S., Morshed, R., Batra, M., & Darmstadt, G. L. (2010). Validation of rapid neurodevelopmental assessment instrument for under-two-year-old children in Bangladesh. *Pediatrics*, 125(4). <https://doi.org/10.1542/peds.2008-3471>
- Knauer, H. A., Kariger, P., Jakielo, P., Ozier, O., & Fernald, L. C. H. (2019). Multilingual assessment of early child development: Analyses from repeated observations of children in Kenya. *Developmental Science*, 22(5), 1–23. <https://doi.org/10.1111/desc.12875>
- Koshy, B., Srinivasan, M., Scharf, R., Strand, T. A., Mohan, V. R., Beulah, R., John, S., Mulyiyil, J., & Kang, G. (2024). Stability and predictability of Bayley Scales of Infant and Toddler Development: evidence from a south Indian birth cohort prospective study. BMJ Open, 14(11), e082624. <https://doi.org/10.1136/bmjopen-2023-082624>
- Koura, K. G., Boivin, M. J., Davidson, L. L., Ouédraogo, S., Zoumenou, R., Alao, M. J., Garcia, A., Massougbedji, A., Cot, M., & Bodeau-Livinec, F. (2013). Usefulness of child development assessments for low-resource settings in Francophone Africa. *Journal of Developmental and Behavioral Pediatrics*, 34(7), 486–493. <https://doi.org/10.1097/DBP.0b013e31829d211c>
- Kucuker, S., Kapci, E. G., & Uslu, R. I. (2011). Evaluation of the turkish version of the “ages and stages questionnaires: Social-emotional” in identifying children with social-emotional problems. *Infants and Young Children*, 24(2), 207–220. <https://doi.org/10.1097/IYC.0b013e31820eae26>

- Kumar, R., Shankar, K., Kush, V., Kumar, C., Bhave, A., & Agarwal, V. (2016). Adaptation: Vineland Adaptive Behavior Scale for 3-9 year-old Indian children. *International Journal on Disability and Human Development*, 15(1), 49–55. <https://doi.org/10.1515/ijdhd-2014-0026>
- Kvestad, I., Silpakar, J. S., Hysing, M., Ranjitkar, S., Strand, T. A., Schwinger, C., Shrestha, M., Chandyo, R. K., & Ulak, M. (2023). The reliability and predictive ability of the Test of Infant Motor Performance (TIMP) in a community-based study in Bhaktapur, Nepal. *Infant Behavior and Development*, 70(January), 101809. <https://doi.org/10.1016/j.infbeh.2023.101809>
- Kvestad, I., Taneja, S., Kumar, T., Bhandari, N., Strand, T. A., & Hysing, M. (2013). The assessment of developmental status using the Ages and Stages questionnaire-3 in nutritional research in north Indian young children. *Nutrition Journal*, 12(1), 1–11. <https://doi.org/10.1186/1475-2891-12-50>
- Lacunza, A. B., Castro Solano, A., & Contini, N. (2009). Habilidades sociales preescolares: una escala para niños de contextos de pobreza. *Revista de Psicología*, 27(1), 3–28. <https://doi.org/10.18800/psico.200901.001>
- Lara Díaz, M. F. L., Fonseca, Á. M. G., Bohórquez, D. M. G., Guechá, C. M., & Sellabona, E. S. (2011). Normativización del inventario del desarrollo comunicativo MacArthur-bates al Español, Colombia. *Revista Latinoamericana de Psicología*, 43(2), 241–254.
- Leppänen, J. M., Butcher, J. W., Godbout, C., Stephenson, K., Taylor Hendrixson, D., Griswold, S., Rogers, B. L., Webb, P., Koroma, A. S., & Manary, M. J. (2022). Assessing infant cognition in field settings using eye-tracking: A pilot cohort trial in Sierra Leone. *BMJ Open*, 12(2). <https://doi.org/10.1136/bmjopen-2021-049783>
- Li, Y., Zhu, J. J., Coplan, R. J., Gao, Z. Q., Xu, P., Li, L., & Zhang, H. (2016). Assessment and Implications of Social Withdrawal Subtypes in Young Chinese Children: The Chinese Version of the Child Social Preference Scale. *Journal of Genetic Psychology*, 177(3), 97–101. <https://doi.org/10.1080/00221325.2016.1174100>
- Li, Y., Tang, L., Bai, Y., Zhao, S., & Shi, Y. (2020). Reliability and validity of the Caregiver Reported Early Development Instruments (CREDI) in impoverished regions of China. *BMC Pediatrics*, 20(1). <https://doi.org/10.1186/s12887-020-02367-4>
- Liu, J., Cheng, H., & Leung, P. W. L. (2011). The application of the preschool child behavior checklist and the Caregiver-Teacher Report Form to Mainland Chinese children: Syndrome structure, gender differences, country effects, and inter-informant agreement. *Journal of Abnormal Child Psychology*, 39(2), 251–264. <https://doi.org/10.1007/s10802-010-9452-8>
- Lohaus, A., Lamm, B., Keller, H., Teubert, M., Fassbender, I., Glüer, M., Borchert, S., Vöhringer, I., Teiser, J., Freitag, C., Suhrke, J., Knopf, M., & Schwarzer, G. (2014). Gross and Fine Motor Differences Between Cameroonian and German Children Aged 3 to 40

- 1
2
3 Months: Results of a Cross-Cultural Longitudinal Study. *Journal of Cross-Cultural*
4 *Psychology*, 45(8), 1328–1341. <https://doi.org/10.1177/0022022114537703>
- 5
6 Lopez Boo, F., Cubides Mateus, M., & Llonch Sabatés, A. (2020). Initial psychometric
7 properties of the Denver II in a sample from Northeast Brazil. *Infant Behavior and*
8 *Development*, 58(August 2019), 101391. <https://doi.org/10.1016/j.infbeh.2019.101391>
- 9
10 Madaschi, V., Mecca, T. P., Macedo, E. C., & Paula, C. S. (2016). Bayley-III scales of infant and
11 toddler development: Transcultural adaptation and psychometric properties. *Paideia*,
12 26(64), 189–197. <https://doi.org/10.1590/1982-43272664201606>
- 13
14 Manandhar, S. R., Dulal, S., Manandhar, D. S., Saville, N., & Prost, A. (2016). Acceptability and
15 Reliability of the Bayley Scales of Infant Development III Cognitive and Motor Scales
16 among Children in Makwanpur. *Journal of Nepal Health Research Council*, 14(32), 47–50.
- 17
18 McCarthy, A. M., Wehby, G. L., Barron, S., Aylward, G. P., Castilla, E. E., Javois, L. C., Goco,
19 N., & Murray, J. C. (2012). Application of neurodevelopmental screening to a sample of
20 South American infants: The Bayley Infant Neurodevelopmental Screener (BINS). *Infant*
21 *Behavior and Development*, 35(2), 280–294. <https://doi.org/10.1016/j.infbeh.2011.12.003>
- 22
23 McCoy, D. C., Sudfeld, C. R., Bellinger, D. C., Muhihi, A., Ashery, G., Weary, T. E., Fawzi, W.,
24 & Fink, G. (2017). Development and validation of an early childhood development scale for
25 use in low-resourced settings. *Population Health Metrics*, 15(1), 1–18.
<https://doi.org/10.1186/s12963-017-0122-8>
- 26
27 McCoy, D. C., Waldman, M., & Fink, G. (2018). Measuring early childhood development at a
28 global scale: Evidence from the Caregiver-Reported Early Development Instruments. *Early*
29 *Childhood Research Quarterly*, 45, 58–68. <https://doi.org/10.1016/j.ecresq.2018.05.002>
- 30
31 McCray, G., McCoy, D., Kariger, P., Janus, M., Black, M. M., Chang, S. M., Tofail, F., Eekhout,
32 I., Waldman, M., Van Buuren, S., Khanam, R., Sazawal, S., Nizar, A., Schönbeck, Y.,
33 Zongo, A., Brentani, A., Zhang, Y., Dua, T., Cavallera, V., ... Gladstone, M. (2023). The
34 creation of the Global Scales for Early Development (GSED) for children aged 0–3 years:
35 Combining subject matter expert judgements with big data. *BMJ Global Health*, 8(1), 1–11.
<https://doi.org/10.1136/bmjjgh-2022-009827>
- 36
37 McHenry, M. S., Oyungu, E., Yang, Z., Hines, A. C., Ombitsa, A. R., Vreeman, R. C.,
38 Abubakar, A., & Monahan, P. O. (2021). Cultural adaptation of the Bayley Scales of Infant
39 and Toddler Development, 3rd Edition for use in Kenyan children aged 18–36 months: A
40 psychometric study. *Research in Developmental Disabilities*, 110(December 2020), 103837.
<https://doi.org/10.1016/j.ridd.2020.103837>
- 41
42 Mehrani, M. B. (2018). An elicited imitation test for measuring preschoolers' language
43 development. *Psychological Reports*, 121(4), 767–786.
<https://doi.org/10.1177/0033294117732976>
- 44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- Meyer, E. L., Schaefer, B. A., Soto, C. M., Simmons, C. S., Anguiano, R., Brett, J., Holman, A., Martin, J. F., Hata, H. K., Roberts, K. J., Mello, Z. R., & Worrell, F. C. (2011). Factor structure of child behavior scale scores in peruvian preschoolers. *Psychology in the Schools*, 48(10), 931–942. <https://doi.org/10.1002/pits.20596>
- Milosavljevic, B., Vellekoop, P., Maris, H., Halliday, D., Drammeh, S., Sanyang, L., Darboe, M. K., Elwell, C., Moore, S. E., & Lloyd-Fox, S. (2019). Adaptation of the Mullen Scales of Early Learning for use among infants aged 5- to 24-months in rural Gambia. *Developmental Science*, 22(5), 1–17. <https://doi.org/10.1111/desc.12808>
- Mukherjee, D., Bhavnani, S., Swaminathan, A., Verma, D., Parameshwaran, D., Divan, G., Dasgupta, J., Sharma, K., Thiagarajan, T. C., & Patel, V. (2020). Proof of Concept of a Gamified DEvelopmental Assessment on an E-Platform (DEEP) Tool to Measure Cognitive Development in Rural Indian Preschool Children. *Frontiers in Psychology*, 11(June), 1–12. <https://doi.org/10.3389/fpsyg.2020.01202>
- Mukherjee, S. B., Verma, S., Sharma, S., & Aneja, S. (2022). Diagnostic Accuracy of Parents' Evaluation of Developmental Status (PEDS), PEDS Developmental Milestones, and PEDS Combined in Indian Children Aged Less than 2 Years. *Indian Journal of Pediatrics*, 89(5), 459–465. <https://doi.org/10.1007/s12098-020-03651-y>
- Nair, M. K. C., Russell, P. S., Rekha, R. S., Lakshmi, M. A., Latha, S., & Rajee, K. (2009). Validation of developmental assessment tool for Anganwadis (DATA). *Indian Pediatrics*, 46 Suppl(2), s27-36.
- Nair, M. K. C., & Russell, P. S. (2013). Development and normative validation of Developmental Assessment Tool for Anganwadis for 3- to 4-year-old children (DATA-II). *Journal of Clinical Epidemiology*, 66(1), 23–29. <https://doi.org/10.1016/j.jclinepi.2012.06.011>
- 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>
- Ngoun, C., Stoey, L. S., van't Ende, K., & Kumar, V. (2012). Creating a Cambodia-specific developmental milestone screening tool - A pilot study. *Early Human Development*, 88(6), 379–385. <https://doi.org/10.1016/j.earlhumdev.2011.09.014>
- Obradović, J., Finch, J. E., Portilla, X. A., Rasheed, M. A., Tirado-Strayer, N., & Yousafzai, A. K. (2019). Early executive functioning in a global context: Developmental continuity and family protective factors. *Developmental Science*, 22(5). <https://doi.org/10.1111/desc.12795>
- Okuda, P. M. M., Félix, E., Cogo-Moreira, H., Liu, T., Surkan, P. J., Martins, S. S., & Caetano, S. C. (2020). Construct validity of the motor development scale (MDS). *Trends in Psychiatry and Psychotherapy*, 42(1), 82–85. <https://doi.org/10.1590/2237-6089-2018-0114>

- Ortiz-León, S., Granados-Rojas, A., Cavazos-Olivo, J., Benito-Avendaño, L. De, Obregón-García, V. H., Duran-Avendaño, X. M., Ramírez-Pérez, R. V., Cárdenas-Medina, J. A., Vargas-Terréz, B. E., Luna-Guevara, G., Rodríguez-Machain, A. C., & Jaimes-Medrano, A. L. (2018). Internal and inter-rater reliability of the ASQ-3 in Mexican preschoolers. *Salud Mental*, 41(2), 65–72. <https://doi.org/10.17711/SALUD.0185-3325.2018.011>
- Oryadi-Zanjani, M. M. (2020). Development of the Childhood Nonverbal Communication Scale. *Journal of Autism and Developmental Disorders*, 50(4), 1238–1248. <https://doi.org/10.1007/s10803-019-04356-8>
- Parveen, M., Rahman, S. T., Islam, S., Zaman, S. S., Hamadani, J., & Khan, N. Z. (2014). Adaptation of items of bayley scales of infant development-II (BSID-II) suitable for Bangladeshi infants. *Dhaka University Journal of Biological Sciences*, 23(2), 187–195. <https://doi.org/10.3329/dujbs.v23i2.20099>
- Pathak, S., & Sovani-Kelkar, P. (2023). Development of a screening tool to identify babies at risk of language delay in India: A preliminary study. *Language Acquisition*, 30(1), 29–49. <https://doi.org/10.1080/10489223.2022.2078212>
- Pendergast, L. L., Schaefer, B. A., Murray-Kolb, L. E., Svensen, E., Shrestha, R., Rasheed, M. A., Scharf, R. J., Kosek, M. N., Vasquez, A. O., Maphula, A., Costa, H., Rasmussen, Z. A., Yousafzai, A., Tofail, F., Seidman, J. C., Acosta, A. M., de Burga, R. R., Chavez, C. B., Flores, J. T., ... Strand, T. (2018). Assessing development across cultures: Invariance of the Bayley-III scales across seven international MAL-ED sites. *School Psychology Quarterly*, 33(4), 604–614. <https://doi.org/10.1037/spq0000264>
- Phạm, B., McLeod, S., & Harrison, L. J. (2017). Validation and norming of the Intelligibility in Context Scale in Northern Viet Nam. *Clinical Linguistics and Phonetics*, 31(7–9), 665–681. <https://doi.org/10.1080/02699206.2017.1306110>
- Pisani, L., Borisova, I., & Dowd, A. J. (2018). Developing and validating the International Development and Early Learning Assessment (IDELA). *International Journal of Educational Research*, 91(July), 1–15. <https://doi.org/10.1016/j.ijer.2018.06.007>
- Pisani, L., Seiden, J., & Wolf, S. (2022). Longitudinal evidence on the predictive validity of the International Development and Early Learning Assessment (IDELA). *Educational Assessment, Evaluation and Accountability*, 34(2), 173–194. <https://doi.org/10.1007/s11092-021-09374-8>
- Pitchik, H. O., Tofail, F., Akter, F., Shoab, A. K. M., Sultana, J., Huda, T. M. N., Rahman, M., Winch, P. J., Luby, S. P., & Fernald, L. C. H. (2023). Concurrent validity of the Ages and Stages Questionnaire Inventory and the Bayley Scales of Infant and Toddler Development in rural Bangladesh. *BMC Pediatrics*, 23(1), 1–11. <https://doi.org/10.1186/s12887-022-03800-6>

- Prado, E. L., Phuka, J., Ocansey, E., Maleta, K., Ashorn, P., Ashorn, U., Adu-Afarwuah, S., Oaks, B. M., Lartey, A., & Dewey, K. G. (2018). A method to develop vocabulary checklists in new languages and their validity to assess early language development. *Journal of Health, Population, and Nutrition*, 37(1), 13. <https://doi.org/10.1186/s41043-018-0145-1>
- Prathanee, B., Pongjanyakul, A., & Chano, J. (2008). Thai speech and language test for children between 1 and 2 years of age. *International Journal of Language and Communication Disorders*, 43(1), 125–140. <https://doi.org/10.1080/13682820601181208>
- Puglisi, M. L., Blasi, H. F., & Snowling, M. J. (2020). Screening for the identification of oral language difficulties in brazilian preschoolers: A validation study. *Language, Speech, and Hearing Services in Schools*, 51(3), 852–865. https://doi.org/10.1044/2020_LSHSS-19-00083
- Ranjitkar, S., Kvestad, I., Strand, T. A., Ulak, M., Shrestha, M., Chandyo, R. K., Shrestha, L., & Hysing, M. (2018). Acceptability and reliability of the Bayley Scales of infant and Toddler Development-III Among Children in Bhaktapur, Nepal. *Frontiers in Psychology*, 9(JUL), 1–10. <https://doi.org/10.3389/fpsyg.2018.01265>
- Rasheed, M. A., Tabussum, G., Habib, Z. H., & Yousafzai, A. K. (2017). Adaptation of MacArthur Communicative Development Inventory in rural Pakistan – useful tool for early childhood studies. *Child: Care, Health and Development*, 43(3), 427–434. <https://doi.org/10.1111/cch.12450>
- Rasheed, M. A., Kvestad, I., Shaheen, F., Memon, U., & Strand, T. A. (2023). The predictive validity of Bayley Scales of Infant and Toddler Development-III at 2 years for later general abilities: Findings from a rural, disadvantaged cohort in Pakistan. *PLOS Global Public Health*, 3(1), e0001485. <https://doi.org/10.1371/journal.pgph.0001485>
- Rasheed, M. A., Pham, S., Memon, U., Siyal, S., Obradović, J., & Yousafzai, A. K. (2018). Adaptation of the Wechsler Preschool and Primary Scale of Intelligence-III and lessons learned for evaluating intelligence in low-income settings. *International Journal of School and Educational Psychology*, 6(3), 197–207. <https://doi.org/10.1080/21683603.2017.1302851>
- Razak, R. A., Madison, C. L., Siow, Y. K., & Aziz, M. A. A. (2010). Preliminary Content Validity and Reliability of a Newly Developed Malay Preschool Language Assessment Tool. *Asia Pacific Journal of Speech, Language and Hearing*, 13(4), 217–234. <https://doi.org/10.1179/136132810805334985>
- Rodriguez, V. J., Zegarac, M., La Barrie, D. L., Parrish, M. S., Matseke, G., Peltzer, K., & Jones, D. L. (2020). Validation of the Bayley Infant Neurodevelopmental Screener among HIV-Exposed Infants in Rural South Africa. *Journal of Acquired Immune Deficiency Syndromes*, 85(4), 507–516. <https://doi.org/10.1097/QAI.0000000000002479>

- Rubio-Codina, M., Araujo, M. C., Attanasio, O., Muñoz, P., & Grantham-McGregor, S. (2016). Concurrent validity and feasibility of short tests currently used to measure early childhood development in large scale studies. *PLoS ONE*, 11(8), 1–17. <https://doi.org/10.1371/journal.pone.0160962>
- Rubio-Codina, M., & Grantham-McGregor, S. (2020). Predictive validity in middle childhood of short tests of early childhood development used in large scale studies compared to the Bayley-III, the Family Care Indicators, height-for-age, and stunting: A longitudinal study in Bogota, Colombia. *PLoS ONE*, 15(4), 1–20. <https://doi.org/10.1371/journal.pone.0231317>
- Saccani, R., Valentini, N. C., & Pereira, K. R. G. (2016). New Brazilian developmental curves and reference values for the Alberta infant motor scale. *Infant Behavior and Development*, 45, 38–46. <https://doi.org/10.1016/j.infbeh.2016.09.002>
- 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>
- Samarakkody, D. C., Fernando, D. N., Perera, H., McClure, R. J., & De Silva, H. (2010). The Child Behaviour Assessment Instrument: Development and validation of a measure to screen for externalising child behavioural problems in community setting. *International Journal of Mental Health Systems*, 4, 1–8. <https://doi.org/10.1186/1752-4458-4-13>
- Sánchez-Vincitore, L. V., Schaettle, P., & Castro, A. (2019). Validation of the Malawi Developmental Assessment Tool for children in the Dominican Republic: Preliminary results. *PLoS ONE*, 14(8), 1–11. <https://doi.org/10.1371/journal.pone.0221162>
- Scherzer, A. L. (2009). Experience in Cambodia with the use of a culturally relevant developmental milestone chart for children in low- and middle-income countries. *Journal of Policy and Practice in Intellectual Disabilities*, 6(4), 287–292. <https://doi.org/10.1111/j.1741-1130.2009.00234.x>
- Shah, M. B., Schaefer, B. A., & Clark, T. P. (2013). Reliability of Bracken School Readiness Assessment, Third Edition Scores With Young Children in Mumbai, India. *International Journal of School and Educational Psychology*, 1(1), 58–66. <https://doi.org/10.1080/21683603.2013.782595>
- Shaoli, S. S., Islam, S., Haque, S., & Islam, A. (2019). Validating the Bangla version of the Dimensions of Mastery Questionnaire (DMQ-18) for preschoolers. *Asian Journal of Psychiatry*, 44(May 2019), 143–149. <https://doi.org/10.1016/j.ajp.2019.07.044>
- Shrestha, M., Kvestad, I., Hysing, M., Ranjitkar, S., Ulak, M., Chandyo, R. K., & Strand, T. A. (2024). The relationship between the ages and stages questionnaire, 3rd edition scores in early childhood and future cognitive abilities in young Nepalese children. *BMC Pediatrics*, 24(1), 642. <https://doi.org/10.1186/s12887-024-05112-3>

- Shrestha, M., Strand, T. A., Ulak, M., Chandyo, R. K., Ranjitkar, S., Hysing, M., Shrestha, L., & Kvestad, I. (2019). The feasibility of the Ages and Stages Questionnaire for the assessment of child development in a community setting in Nepal. *Child: Care, Health and Development*, 45(3), 394–402. <https://doi.org/10.1111/cch.12654>
- Siegle, C. B. H., & dos Santos Cardoso de Sá, C. (2018). Concurrent validity between instruments of assessment of motor development in infants exposed to HIV. *Infant Behavior and Development*, 50(February), 198–206. <https://doi.org/10.1016/j.infbeh.2018.01.005>
- Sincovich, A., Gregory, T., Zanon, C., Santos, D. D., Lynch, J., & Brinkman, S. A. (2019). Measuring early childhood development in multiple contexts: The internal factor structure and reliability of the early Human Capability Index in seven low and middle income countries. *BMC Pediatrics*, 19(1), 1–14. <https://doi.org/10.1186/s12887-019-1852-5>
- Snelling, M., Dawes, A., Biersteker, L., Girdwood, E., & Tredoux, C. (2019). The development of a South African Early Learning Outcomes Measure: A South African instrument for measuring early learning program outcomes. *Child: Care, Health and Development*, 45(2), 257–270. <https://doi.org/10.1111/cch.12641>
- Soleimani, F., & Dadkhah, A. (2007). Validity and reliability of infant Neurological International Battery for detection of gross motor developmental delay in Iran. *Child: Care, Health and Development*, 33(3), 262–265. <https://doi.org/10.1111/j.1365-2214.2006.00704.x>
- Soleymani, Z., Nematzadeh, S., Tehrani, L. G., Rahgozar, M., & Schneider, P. (2016). Language sample analysis: development of a valid language assessment tool and determining the reliability of outcome measures for Farsi-speaking children. *European Journal of Developmental Psychology*, 13(2), 275–291. <https://doi.org/10.1080/17405629.2016.1152175>
- Solís Cordero K, Quezada Ugalde AM, González Serrano S, Zúñiga Baldi C. Costa Rican test: Evaluation of the children's integral development from 0 months to 6 years of age (EDIN II). Enfermería Actual Costa Rica 2019; 37. DOI:<http://dx.doi.org/10.15517/revenf.v0ino.37.34959>.
- Soysal, A. S., Gucuyener, K., Ergenekon, E., Turan, Ö., Koc, E., Turkyilmaz, C., Önal, E., & Atalay, Y. (2014). The prediction of later neurodevelopmental status of preterm infants at ages 7 to 10 years using the bayley infant neurodevelopmental screener. *Journal of Child Neurology*, 29(10), 1349–1355. <https://doi.org/10.1177/0883073813520495>
- Thompson, L., Peñaloza, R. A., Stormfields, K., Kooistra, R., Valencia-Moscoso, G., Muslima, H., & Khan, N. Z. (2015). Validation and adaptation of rapid neurodevelopmental assessment instrument for infants in Guatemala. *Child: Care, Health and Development*, 41(6), 1131–1139. <https://doi.org/10.1111/cch.12279>

- 1
2
3 Toor, H. K., & Hanif, R. (2022). Executive Function Touch battery: Translation and preliminary
4 measure validation for Pakistani preschoolers. *PLoS ONE*, 17(9 September), 1–14.
5 https://doi.org/10.1371/journal.pone.0274431
6
7 Uyanik, Ö., & Kandir, A. (2014). Adaptation of the Kaufman Survey of Early Academic and
8 Language Skills to Turkish Children Aged 61 to 72 Months. *Kuram ve Uygulamada Eğitim
9 Bilimleri*, 14(2), 682–692. https://doi.org/10.12738/estp.2014.2.1682
10
11 Valentini, N. C., & Saccani, R. (2012). Brazilian Validation of the Alberta Infant Motor Scale.
12 *Physical Therapy*, 92(3), 440–447. https://doi.org/10.2522/ptj.20110036
13
14 Vameghi, R., Sajedi, F., Kraskian Mojembari, A., Habiollahi, A., Lornezhad, H. R., & Delavar,
15 B. (2013). Cross-cultural adaptation, validation and standardization of Ages and Stages
16 Questionnaire (ASQ) in Iranian children. *Iranian Journal of Public Health*, 42(5), 522–528.
17
18 van Der Linde, J., Swanepoel, D. W., Glascoe, F. P., Louw, E. M., & Vinck, B. (2015).
19 Developmental screening in South Africa: Comparing the national developmental checklist
20 to a standardized tool. *African Health Sciences*, 15(1), 188–196.
21 https://doi.org/10.4314/ahs.v15i1.25
22
23 van Heerden, A., Hsiao, C., Matafwali, B., Louw, J., & Richter, L. (2017). Support for the
24 feasibility of the ages and stages questionnaire as a developmental screening tool: A cross-
25 sectional study of South African and Zambian children aged 2–60 months. *BMC Pediatrics*,
26 17(1), 1–9. https://doi.org/10.1186/s12887-017-0802-3
27
28 Venancio, S. I., Bortoli, M. C., Frias, P. G., Giugliani, E. R. J., Alves, C. R. L., & Santos, M. O.
29 (2020). Development and validation of an instrument for monitoring child development
30 indicators. *Jornal de Pediatria (Versão Em Português)*, 96(6), 778–789.
31 https://doi.org/10.1016/j.jpedp.2019.10.008
32
33 Venter, A., Pienaar, A. E., & Coetzee, D. (2015). Suitability of the ‘Little DCDQ’ for the
34 identification of DCD in a selected group of 3–5-year-old South African children.’ *Early
35 Child Development and Care*, 185(8), 1359–1371.
36 https://doi.org/10.1080/03004430.2014.1000887
37
38 Waechter, R., Evans, R., Hanna, S., Murray, T., Mobley, C., Holmes, S., Isaac, R., Wolfe, R.,
39 Andrew, E., Landon, B., & Fernandes, M. (2022). Adaptation of the INTERGROWTH-21st
40 neurodevelopment assessment (INTER-NDA) to the context of the English-speaking
41 Caribbean. *BMC Pediatrics*, 22(1), 1–10. https://doi.org/10.1186/s12887-021-03039-7
42
43 Waldman, M., McCoy, D. C., Seiden, J., Cuartas, J., Fink, G., Altafim, E., Brentani, A.,
44 Castellanos, A., Chen, A., Chomat, A. M., Fawzi, W., de Pin˜eres, C. G., Hamadani, J.,
45 Henao, N., Jervis, P., Kane, C., Measelle, J., Medrano, P., Pisani, L., ... Yousafzai, A. K.
46 (2021). Validation of motor, cognitive, language, and socio-emotional subscales using the
47 Caregiver Reported Early Development Instruments: An application of multidimensional
48

- item factor analysis. *International Journal of Behavioral Development*, 45(4), 368–377. <https://doi.org/10.1177/01650254211005560>
- Washington, K. N., McDonald, M. M., McLeod, S., Crowe, K., & Devonish, H. (2017). Validation of the intelligibility in context scale for Jamaican Creole-speaking preschoolers. *American Journal of Speech-Language Pathology*, 26(3), 750–761. https://doi.org/10.1044/2016_AJSLP-15-0103
- Weber, A. M., Marchman, V. A., Diop, Y., & Fernald, A. (2018). Validity of caregiver-report measures of language skill for Wolof-learning infants and toddlers living in rural African villages. *Journal of Child Language*, 45(4), 939–958. <https://doi.org/10.1017/S0305000917000605>
- Wijedasa, D. (2012). Developmental screening in context: Adaptation and standardization of the Denver Developmental Screening Test-II (DDST-II) for Sri Lankan children. *Child: Care, Health and Development*, 38(6), 889–899. <https://doi.org/10.1111/j.1365-2214.2011.01332.x>
- Willoughby, M. 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>
- Willoughby, M. T., Piper, B., Oyanga, A., & Merseth King, K. (2019). Measuring executive function skills in young children in Kenya: Associations with school readiness. *Developmental Science*, 22(5). <https://doi.org/10.1111/desc.12818>
- Wolf, S., Halpin, P., Yoshikawa, H., Dowd, A. J., Pisani, L., & Borisova, I. (2017). Measuring school readiness globally: Assessing the construct validity and measurement invariance of the International Development and Early Learning Assessment (IDEA) in Ethiopia. *Early Childhood Research Quarterly*, 41(June), 21–36. <https://doi.org/10.1016/j.ecresq.2017.05.001>
- Wu, Z., Hu, B. Y., & Fan, X. (2019). Cross-Cultural Validity of Preschool Learning Behavior Scale in Chinese Cultural Context. *Journal of Psychoeducational Assessment*, 37(1), 125–130. <https://doi.org/10.1177/0734282916651538>
- Wu, Z., Mak, M. C. K., Hu, B. Y., He, J., & Fan, X. (2019). A validation of the Social Skills domain of the Social Skills Improvement System-Rating Scales with Chinese preschoolers. *Psychology in the Schools*, 56(1), 126–147. <https://doi.org/10.1002/pits.22193>
- Xie, H., Bian, X., Chen, C. Y., Squires, J., & Lu, P. (2019). Examining the Convergent Evidence of a Parent-Completed, Social-Emotional Screening Tool in China. *Journal of Child and Family Studies*, 28(6), 1471–1480. <https://doi.org/10.1007/s10826-019-01362-3>
- Xie, H., Clifford, J., Squires, J., Chen, C. Y., Bian, X., & Yu, Q. (2017). Adapting and validating a developmental assessment for Chinese infants and toddlers: The ages & stages

- questionnaires: Inventory. *Infant Behavior and Development*, 49(June), 281–295.
<https://doi.org/10.1016/j.infbeh.2017.09.004>
- Xie, S., & Li, H. (2019). Development and Validation of the Chinese Preschool Readiness Scale. *Early Education and Development*, 30(4), 522–539.
<https://doi.org/10.1080/10409289.2019.1596462>
- Yibeltal, K., Teklehaymanot, A., Workneh, F., Fasil, N., Jensen, S. K. G., Chin, T. I., North, K., Haymanot, B., Worku, A., Lee, A. C., & Berhane, Y. (2025). Adaptation and Implementation of the Global Scales for Early Development (GSED) Tool in Ethiopia. *Children*, 12(3), 299. <https://doi.org/10.3390/children12030299>
- Yuan, H., Ocansey, M., Adu-Afarwuah, S., Sheridan, M., Hamoudi, A., Okronipa, H., Kumordzie, S. M., Oaks, B. M., & Prado, E. (2022). Evaluation of a tablet-based assessment tool for measuring cognition among children 4–6 years of age in Ghana. *Brain and Behavior*, 12(10), 1–12. <https://doi.org/10.1002/brb3.2749>
- Yue, A., Jiang, Q., Wang, B., Abbey, C., Medina, A., Shi, Y., & Rozelle, S. (2019). Concurrent validity of the Ages and Stages Questionnaire and the Bayley Scales of Infant Development III in China. *PLoS ONE*, 14(9). <https://doi.org/10.1371/journal.pone.0221675>
- Yunilda, E., Gunardi, H., Medise, B. E., & Oswari, H. (2023). The Indonesian version of Ages and Stages Questionnaire III accuracy compared to Bayley Scales of Infant Development III. *Infant and Child Development*, 32(1), 1–12. <https://doi.org/10.1002/icd.2387>
- Zakaria, S., Chua, B. S., Sombuling, A., Ahmad, S., & Hashmi, S. I. (2012). Reliability and Validity for Malay Version of Bayley Scales of Infant and Toddler Development-Third Edition (Bayley-III): Preliminary Study. *International Proceedings of Economics Development and Research*, 40(June), 116–119.
- Zhang, L., Li, Z., & He, K. (2018). Validation of the Chinese Revision of Emotional and Behavioral Screener among Preschoolers (4–7 years old). *Journal of Child and Family Studies*, 27(3), 771–779. <https://doi.org/10.1007/s10826-017-0940-0>
- Zhang, Y., Xu, X., Jiang, Y., Sun, W., Wang, Y., Song, Y., Dong, S., Zhu, Q., Jiang, F., & Sheng, L. (2021). Early language and communication development in Chinese children: Adaption and validation of a parent report instrument. *International Journal of Speech-Language Pathology*, 23(4), 372–381. <https://doi.org/10.1080/17549507.2020.1817558>
- Zirakashvili, M., Gabunia, M., Tatishvili, N., Ediberidze, T., Lomidze, G., Chachava, T., & Hix-Small, H. (2018). Cultural Adaptation and Psychometric Validation of the Ages and Stages Questionnaires for use in Georgia. *Journal of Child and Family Studies*, 27(3), 739–749. <https://doi.org/10.1007/s10826-017-0917-z>