Online resource 2 – Tables methodological quality scores

An updated systematic review of childhood physical activity questionnaires Journal: Sports Medicine

Lisan M. Hidding, Mai. J. M. Chinapaw, Mireille N. M. van Poppel, Lidwine B. Mokkink, Teatske M. Altenburg

Corresponding author: Lisan Hidding Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Public and Occupational Health, Amsterdam Public Health research institute, Van der Boechorststraat 7, NL-1081 BT Amsterdam, The Netherlands E-mail: <u>l.hidding@vumc.nl</u>

Questionnaire	Study population ^a	Methodological	Design requirement(s) that determined
		quality ^b	the final methodological quality score ^c
Preschool-age Children's	n = 67	Good	1.Was the percentage of missing items
Physical Activity	Age: 3- to 5-year		given?
Questionnaire (Pre-PAQ)	olds		2.Was there a description of how missing
(proxy) [58]	Sex: 48% girls		items were handled?
			3.Was the sample size included in the
			analysis adequate?
Modified Burdette proxy	n = 107	Fair	4.Were hypotheses regarding
report (proxy) [59]	Age: 3.4±1.2 years		correlations or mean differences
	Sex: pecentage girls		formulated a priori (i.e. before data
	unknown		collection)?
Modified Harro proxy	n = 131	Fair	4.Were hypotheses regarding
report (proxy) [59]	Age: 3.8±1.3		correlations or mean differences
	Sex: pecentage girls		formulated a priori (i.e. before data
	unknown		collection)?
Physical activity	n = 35	Poor	4.Were hypotheses regarding
questionnaire for parents	Age: 4.4±0.7 years		correlations or mean differences
of preschoolers in Mexico	[3-5]		formulated a priori (i.e. before data
[40]	Sex: 51% girls		collection)?
Children's Physical	n = 27	Poor (all comparison	3.Was the sample size included in the
Activity Questionnaire	Age: 4.9±0.7 years	measures)	analysis adequate?
(CPAQ) (proxy) [60] ^d	[4-5]		
	Sex: 38% girls		
Physical activity and	n = 87	Poor	4.Were hypotheses regarding
sedentary behavior proxy	Age: 4-70 months		correlations or mean differences
questionnaire (based on	Sex: 54% girls		formulated a priori (i.e. before data
CHMS) (proxy) [61]			collection)?
Out-of-school Physical	n = 126	Fair	4.Were hypotheses regarding
Activity questionnaire	Age: 11-year-olds		correlations or mean differences
[62]			
		I	

	Sex: 60% girls (in		formulated a priori (i.e. before data
	total sample n=155)		collection)?
Children's Leisure	n=139; [9-12yrs];	Fair	4.Were hypotheses regarding
Activities Study Survey	65% girls		correlations or mean differences
Chinese-version			formulated a priori (i.e. before data
questionnaire (CLASS-C)			collection)?
[50]			
Physical Activity	n = ranging from 73	Fair (all comparison	2.Was there a description of how missing
Questionnaire for Older	(Caltrac) to 97	measures)	items were handled?
Children (PAQ-C) [27]d	(activity rating and		
	Godin 1)		
	Age: 11.3±1.4 years		
	[9-14]		
	Sex: 58%		
Previous Day Physical	n = 37	Fair	3.Was the sample size included in the
Activity Recall (PDPAR)	Age: 10.8±0.1 years		analysis adequate?
[30]	(in total sample		4.Were hypotheses regarding
	n=38)		correlations or mean differences
	Sex: 51% girls		formulated a priori (i.e. before data
			collection)?
Physical Activity	n = 78	Fair	4.Were hypotheses regarding
Questionnaire for older	Age: 11.0 ± 1.2 (total		correlations or mean differences
Children (PAQ-C) (Spanish	sample n=83)		formulated a priori (i.e. before data
version) [52]	Sex: 45 % girls (total		collection)?
	sample n=83)		
Godin Leisure-Time	n = 31	Fair	2.Was there a description of how missing
Exercise Questionnaire	Age: 10.6 ± 0.2		items were handled?
[63]	Sex: 45 % girls		3.Was the sample size included in the
			analysis adequate?
			4.Were hypotheses regarding
			correlations or mean differences
			formulated a priori (i.e. before data
			collection)?

Multimedia Activity Recall	n = 66	Fair	4.Were hypotheses regarding
for Children and	Age: 11.6±0.8 years		correlations or mean differences
Adolescents (MARCA)	Sex: 50% girls		formulated a priori (i.e. before data
[64] ^d			collection)?
Chinese version of the	n = 358	Fair	4.Were hypotheses regarding
Physical Activity	Age: 10.5±1.1 years		correlations or mean differences
Questionnaire for Older	[8-13] (in total		formulated a priori (i.e. before data
Children (PAQ-C) [43]	sample n=742)		collection)?
9	Sex: 46% girls		
Youth Activity Profile	n = 291	Fair	9. Were there any important flaws in the
(YAP) [38]	Age: 9.7±1.0		design or methods of the study?
((n=135), 11.7±0.8		
((n=67), 15.7±1.2		
	(n=89) years		
	Sex: 56% girls		
Food, Health, and Choices	n = 66	Fair	2.Was there a description of how missing
questionnaire (FHC-Q)	Age: <9- to >12-year		items were handled?
[37]	olds		8.For convergent validity: Were the
5	Sex: 50 % girls		measurement properties of the
			comparator instrument(s) adequately
			described?
Self-administered 1	n = 86	Poor	8.For convergent validity: Were the
questionnaire to assess	Age: 10.2 ± 1.1		measurement properties of the
physical activity and	Sex: 54 % girls		comparator instrument(s) adequately
sedentary behaviors [65]			described?
The South American	n = 82	Poor	8.For convergent validity: Were the
Youth/Child	Age: 3- to 10-year		measurement properties of the
Cardiovascular and o	olds		comparator instrument(s) adequately
Environment Study S	Sex: 54 % girls		described?
(SAYCARE) Physical			
Activity (PA)			
questionnaire (proxy)			
[66]			

Canadian Health Measures	n = 878	Poor	4.Were hypotheses regarding
Survey (CHMS) [67]	Age: 8.7 years (95%		correlations or mean differences
	CI 8.5-8.9) [6-11]		formulated a priori (i.e. before data
	Sex: 49% girls		collection)?
Many Rivers Physical	n = 86	Poor	4.Were hypotheses regarding
Activity Recall	Age: 11.1±0.7 years		correlations or mean differences
Questionnaire (MRPARQ)	Sex: 59% girls		formulated a priori (i.e. before data
(modified version of the			collection)?
APARQ) [68]			
Patient Assessment and	n = 18	Poor (all comparison	3.Was the sample size included in the
Council for Exercise	Age: 11.9±2.0 years	measures)	analysis adequate?
(PACE) [69]	Sex: 59% girls		
	(age and sex total		
	sample n=22)		
Self-Administered	n = 90	Poor	8.For convergent validity: Were the
Physical Activity Checklist	Age: 11.4±0.6 (boys)		measurement properties of the
(SAPAC) (Greek version)	11.3±0.6 (girls)		comparator instrument(s) adequately
[49]	years		described?
	Sex: 57% girls		
Assessment of Young	n = 47	Poor (all comparison	4.Were hypotheses regarding
Children's Activity using	Age: 7.7±0.5 years	measures)	correlations or mean differences
Video Technology	Sex: 40% girls		formulated a priori (i.e. before data
(ACTIVITY) [70] ^d			collection)?
Synchronised Nutrition	n = 121	Poor	4.Were hypotheses regarding
and Activity Program	Age: 10.7±2.2 years		correlations or mean differences
(SNAP) [71] ^d	[7-15]		formulated a priori (i.e. before data
	Sex: 60% girls		collection)?
PA questionnaire for	n = 62	Poor (all comparison	4.Were hypotheses regarding
parents and teachers [72] ^d	Age: 7.0±0.7 years	measures)	correlations or mean differences
	[4-8]		formulated a priori (i.e. before data
	Sex: 52% girls		collection)?

Physical Activity	n = 58	Poor	9. Were there any important flaws in the
Questionnaire for older	Age: 7- to 9-year		design or methods of the study?
Children (PAQ-C) [51]	olds		
	Sex: 48 % girls		
The Modified Godin	n = 139	Poor	4.Were hypotheses regarding
Leisure-Time Exercise	Age: 11.1 ± 0.4		correlations or mean differences
Questionnaire [45]	Sex: 52 % girls		formulated a priori (i.e. before data
			collection)?
Parent proxy-report of	n = 167 (validity vs.	Poor (all comparison	4.Were hypotheses regarding
physical activity and	accelerometer), n =	measures)	correlations or mean differences
sedentary activities	125 (validity vs.		formulated a priori (i.e. before data
(proxy) [73]	diary)		collection)?
	Age: 6- to 10-year		8.For convergent validity: Were the
	olds, 13- to 14-year		measurement properties of the
	olds		comparator instrument(s) adequately
	Sex: 51% girls (in		described?
	total sample n=189)		
Diet and lifestyle	N = 446	Poor	4.Were hypotheses regarding
questionnaire [74]	Age: 9.0-11.9 years		correlations or mean differences
	(in total sample		formulated a priori (i.e. before data
	n=563)		collection)?
	Sex: 53% girls (in		8.For convergent validity: Were the
	total sample n=563)		measurement properties of the
			comparator instrument(s) adequately
			described?
ATN-Questionnaire [75]	n = 58	Poor (all comparison	4.Were hypotheses regarding
	Age: 11.4±0.5 years	measures)	correlations or mean differences
	Sex: 54% girls		formulated a priori (i.e. before data
			collection)?
			8.For convergent validity: Were the
			measurement properties of the
			comparator instrument(s) adequately
			described?
		1	

the ENERGY-child	n = 96	Poor	4.Were hypotheses regarding
questionnaire [48]	Age: [11.4±0.6 -		correlations or mean differences
	12.0±0.6 years]		formulated a priori (i.e. before data
	Sex: [31% - 67%		collection)?
	girls]		
A physical activity	n = 4254	Poor	4.Were hypotheses regarding
questionnaire [76]	Age: 11.3 years		correlations or mean differences
	Sex: 51% girls (in		formulated a priori (i.e. before data
	total sample		collection)?
	n=4452)		8.For convergent validity: Were the
			measurement properties of the
			comparator instrument(s) adequately
			described?
Instrument to assess	n = 46	Poor	8.For convergent validity: Were the
children's outdoor active	Age: 9.2 years [7.9-		measurement properties of the
play in various locations	11.7] Sex: 50% girls		comparator instrument(s) adequately
(proxy) [77]			described?
Questions from the	n = 3,940 (organized	Poor	4.Were hypotheses regarding
National Longitudinal	sports question)		correlations or mean differences
Survey of Children and	n = 3,958 (leisure		formulated a priori (i.e. before data
Youth [78]	sports question)		collection)?
	Age: 5th graders		8.For convergent validity: Were the
	Sex: pecentage girls		measurement properties of the
	unknown		comparator instrument(s) adequately
			described?
Physical Activity	n = 132	Poor	8.For convergent validity: Were the
Questionnaire for Older	Age: 10.3±0.6 years		measurement properties of the
Children (PAQ-C) (minor	[9-11]		comparator instrument(s) adequately
modifications) [44]	Sex: 48% girls		described?
A physical activity	n = 224	Good (all	4.Were hypotheses regarding
questionnaire of the	Age 12.2±0.8 years	comparison	correlations or mean differences
Estonian Children	Sex: 0% girls	measures)	formulated a priori (i.e. before data
Personality Behavior and			collection)?

Health Study (ECPBHS)			
[79]			
A physical activity	n = 224	Good (all	4.Were hypotheses regarding
questionnaire of the	Age 12.2±0.8 years	comparison	correlations or mean differences
Estonian Children	Sex: 0% girls	measures)	formulated a priori (i.e. before data
Personality Behavior and			collection)?
Health Study (ECPBHS)			
(proxy) [79]			
3-Day Physical Activity	n = 33	Fair	3.Was the sample size included in the
Record (3DPARecord)	Age: 13.7±0.8 years		analysis adequate?
(Greek version) [33]	Sex: 43% girls (age		4.Were hypotheses regarding
	and sex total sample		correlations or mean differences
	n=40)		formulated a priori (i.e. before data
			collection)?
Seven Day Physical	n = 123	Fair (all comparison	2.Was there a description of how missing
Activity Recall (7 Day-	Age: 14.9 ± 0.9 [13-	measures)	items were handled?
PAR) (Spanish version)	17]		4.Were hypotheses regarding
[80]	Sex: 59 % girls		correlations or mean differences
			formulated a priori (i.e. before data
			collection)?
			8.For convergent validity: Were the
			measurement properties of the
			comparator instrument(s) adequately
			described?
Youth Physical Activity	n = 44	Fair	3.Was the sample size included in the
Questionnaire (YPAQ)	Age: 12.7 [12- 13]		analysis adequate?
[81]	Sex: 61 % girls		4.Were hypotheses regarding
			correlations or mean differences
			formulated a priori (i.e. before data
			collection)?
International Physical	n = 191	Fair	4.Were hypotheses regarding
Activity Questionnaire –	Age: 14.0 ± 0.7		correlations or mean differences
	Sex: 0 % girls		

Short Form (IPAQ-SF)			formulated a priori (i.e. before data
[82]			collection)?
Tartu Physical Activity	n = 191	Fair	4.Were hypotheses regarding
Questionnaire (TPAQ)	Age: 14.0 ± 0.7		correlations or mean differences
[82]	Sex: 0 % girls		formulated a priori (i.e. before data
			collection)?
Physical Activity and	n = 33	Fair	3.Was the sample size included in the
Lifestyle Questionnaire	Age: 13.7±0.8 years		analysis adequate?
(PALQ) (Greek version)	Sex: 43% girls (age		4.Were hypotheses regarding
[33]	and sex total sample		correlations or mean differences
	n=40)		formulated a priori (i.e. before data
			collection)?
Moderate and vigorous	n = 125	Fair	4.Were hypotheses regarding
physical activity items of	Age: 12.2±0.6 years		correlations or mean differences
the Youth Risk Behavior	Sex: 53% girls (age		formulated a priori (i.e. before data
Survey (YRBS) [83]	and sex total sample		collection)?
	n=139)		
3-Day Physical Activity	n = 70	Fair	4.Were hypotheses regarding
Recall (3DPARecall)	Age: 14.0±0.9 years		correlations or mean differences
instrument [20]	[13-16]		formulated a priori (i.e. before data
	Sex: 100% girls		collection)?
International Physical	n = 1021	Fair	4.Were hypotheses regarding
Activity Questionnaire -	Age: 14.3±1.6 years		correlations or mean differences
Short Form (IPAQ - SF)	[12-18]		formulated a priori (i.e. before data
[84]	Sex: 47% girls		collection)?
			8.For convergent validity: Were the
			measurement properties of the
			comparator instrument(s) adequately
			described?
PACE+ questionnaire [85]	n = 235	Fair	4.Were hypotheses regarding
	Age: 14.7±3.1 years		correlations or mean differences
	Sex: 59% girls		

			formulated a priori (i.e. before data
			collection)?
3-Day Physical Activity	n = 155	Fair	4.Were hypotheses regarding
Recall (3DPARecall)	Age: 12.3±0.9 years		correlations or mean differences
(modified for Australian	Sex: 50% girls		formulated a priori (i.e. before data
youth) [86]			collection)?
Single-item activity	n = 96	Fair	4.Were hypotheses regarding
measure [23]	(Accelerometer wear		correlations or mean differences
	time 480 min./d.)		formulated a priori (i.e. before data
	Age: 14.7±0.5 years		collection)?
	Sex: 38% girls (total		
	sample (age and sex		
	total sample n=123)		
	n = 72		
	(Accelerometer >		
	wear time 600		
	min./d.)		
	Age: 14.7±0.5 years		
	Sex: 38% girls (age		
	and sex total sample		
	n=123)		
Oxford Physical Activity	n = 96	Fair	4.Were hypotheses regarding
Questionnaire (OPAQ)	(Accelerometer wear		correlations or mean differences
[23]	time 480 mins/day)		formulated a priori (i.e. before data
	Age: 14.7±0.5 years		collection)?
	Sex: 38% girls (total		
	sample (age and sex		
	total sample n=123)		
	n = 72		
	(Accelerometer >		

	wear time 600		
	min/day)		
	Age: 14.7±0.5 years		
	Sex: 38% girls (age		
	and sex total sample		
	n=123)		
MVPA self-report	n = 203 (5 valid	Fair	4.Were hypotheses regarding
questionnaire [87]	accelerometer days)		correlations or mean differences
	Age: 15.8±0.7 years		formulated a priori (i.e. before data
	Sex: 61% girls		collection)?
	n = 103 (7 valid		
	accelerometer days)		
	Age: 15.8±0.7 (total		
	sample n=203)		
	Sex: 67% girls		
Activity Questionnaire for	n = 42	Fair	3.Was the sample size included in the
Adults and Adolescents	Age: 13.4±1.0 years		analysis adequate?
(AQuAA) [21]	Sex: 50% girls		
Physical Activity	n = ranging from 48	Fair (all comparison	4.Were hypotheses regarding
Questionnaire for	(Caltrac) to 85	measures)	correlations or mean differences
Adolescents (PAQ-A) [88]d	(Activity rating,		formulated a priori (i.e. before data
	Godin 1 and 2)		collection)?
	Age: 16.3±1.5 years		
	Sex: 52% girls		
Modified Physical Activity	n = 88	Fair (all comparison	4.Were hypotheses regarding
Questionnaire for	Age: 14.5±1.7 years	measures)	correlations or mean differences
Adolescents (PAQ-A) [34]	Sex: 42% girls		formulated a priori (i.e. before data
	(age and sex total		collection)?
	sample n=169)		
An adapted version of the	n = 77	Fair	4.Were hypotheses regarding
Assessment of Physical	Age: 13.6 ± 1.1		correlations or mean differences
Activity Levels	Sex: 35 % girls		conclusions of mean uncreateds
Activity Levels	Jea. 33 70 gil 15		

Questionnaire (APALQ)			formulated a priori (i.e. before data
[53]			collection)?
3-Day Physical Activity	n = 219	Fair	4.Were hypotheses regarding
Recall (3DPARecall)	Age: 14.5±1.1 years		correlations or mean differences
instrument (Singaporean	[13-16]		formulated a priori (i.e. before data
version) [42]	Sex: 53% girls (age		collection)?
	and sex total sample		9. Were there any important flaws in the
	n=221)		design or methods of the study?
Web-based physical	n = 342	Fair (all comparison	4.Were hypotheses regarding
Activity Questionnaire for	(pedometer), 391	measures)	correlations or mean differences
Older Children (PAQ-C)	(shuttlerun)		formulated a priori (i.e. before data
[28]	Age: 12.8 years		collection)?
	Sex: 51 % girls		
	(age and sex total		
	sample n=459)		
Physical activity	n = 75	Fair	4.Were hypotheses regarding
questionnaire of the Arab	Age: 16.1±1.1		correlations or mean differences
Teen Lifestyle Study [89]	Sex: 48% girls		formulated a priori (i.e. before data
			collection)?
Previous Day Physical	ACTIVITYGRAM	Poor vs. biotrainer	Poor:
Activity Recall (PDPAR)	n = 147		3.Was the sample size included in the
[31]	Age:12.4±0.4 years	Fair vs.	analysis adequate?
	Sex: 44% girls	questionnaire	
	Biotrainer (first		Fair:
	sample)		4.Were hypotheses regarding
	n = 28 [25-28]		correlations or mean differences
	Age: 12.4±0.5 years		formulated a priori (i.e. before data
	Sex: 50% girls		collection)?
	Biotrainer (second		
	sample) n = 128		
	Age: unknown		
	Sex:36% girls		
	Sex:36% girls		

ACTIVITYGRAM self-	PDPAR	Poor vs. biotrainer	Poor:
report assessment [31]	n = 147		3.Was the sample size included in the
	Age:12.4±0.4 years	Fair vs.	analysis adequate?
	Sex: 44% girls	questionnaire	
			Fair:
	Biotrainer		4.Were hypotheses regarding
	n = 28 [25-28]		correlations or mean differences
	Age: 12.4±0.5 years		formulated a priori (i.e. before data
	Sex: 50% girls		collection)?
MVPA scores of the	n = 76 (vs. acc.)	Fair vs.	4.Were hypotheses regarding
International Physical	Age: 12.7 ± 1.4 (total	accelerometer	correlations or mean differences
Activity Questionnaire	sample n=998)		formulated a priori (i.e. before data
Short form (IPAQ-SF) [90]	Sex: 53 % girls	Poor vs.	collection)?
		questionnaire	
	n = 998 (vs.		
	questionnaire)		
	Age: 12.7 ± 1.4		
	Sex: 50 % girls		
MVPA scores of the Health	n = 76 (vs. acc.)	Fair vs.	4.Were hypotheses regarding
Behavior in School-aged	Age: 12.7 ± 1.4 (total	accelerometer	correlations or mean differences
Children (HBSC) Research	sample n=998)		formulated a priori (i.e. before data
Protocol [90]	Sex: 53 % girls	Poor vs.	collection)?
		questionnaire	
	n = 998 (vs.		
	questionnaire)		
	Age: 12.7 ± 1.4		
	Sex: 50 % girls		
The South American	n = 60	Poor	8.For convergent validity: Were the
Youth/Child	Age: 11- to 18-year		measurement properties of the
Cardiovascular and	olds		comparator instrument(s) adequately
Environment Study	Sex: 56 % girls		described?
(SAYCARE) Physical			

Activity (PA)			
questionnaire [66]			
Pelotas Birth cohort	n = 25	Poor	3.Was the sample size included in the
physical activity	Age: 13.0±0.3 years		analysis adequate?
questionnaire [91]	Sex: 64% girls		
3-Day Physical Activity	n = 20	Poor	3.Was the sample size included in the
Recall (3DPARecall)	Age: 13.3±0.9		analysis adequate?
questionnaire (modified)	Sex: 100% girls		
[92]			
SQUASH [93]	n = 17	Poor	3.Was the sample size included in the
	Age: 17.5±0.6 years		analysis adequate?
	Sex: 53% girls		
International Physical	n = 2018	Poor (all comparison	8.For convergent validity: Were the
Activity Questionnaire for	Age: [12.5-17.5	measures)	measurement properties of the
Adolescents (adapted	years]		comparator instrument(s) adequately
version of the IPAQ) [94]	Sex: 54% girls		described?
Recess Physical Activity	n = 49 (pedometer)	Poor (all comparison	4.Were hypotheses regarding
Recall (RPAR) [95]	Age: 13.3±0.5 years	measures)	correlations or mean differences
	Sex: 65% girls		formulated a priori (i.e. before data
			collection)?
	n = 32 (biotrainer)		
	Age: 12.9±0.8 years		
	Sex: 31% girls		
	n = 32 (actigraph)		
	Age 12.7±0.8 years		
	Sex: 38% girls		
Swedish Adolescent	n = 50	Poor	4.Were hypotheses regarding
Physical Activity	Age: 16.9±0.4 years		correlations or mean differences
Questionnaire (SAPAQ)	Sex: 62% girls		formulated a priori (i.e. before data
[96] ^d			collection)?

Activity Questionnaire for	n = 236	Poor	4.Were hypotheses regarding
Adults and Adolescents	Age: 15.0±1.0 years		correlations or mean differences
(AQuAA) [22]	Sex: 60% girls		formulated a priori (i.e. before data
			collection)?
Computer assisted	n = 2761	Poor	4.Were hypotheses regarding
interview based on	Age: 12- to 19- year		correlations or mean differences
NHANES survey [97]	olds		formulated a priori (i.e. before data
	Sex: 48% girls		collection)?
Previous Day Physical	n = 122	Poor	4.Were hypotheses regarding
Activity Recall (PDPAR-	Age: 13.8 ±1.2 years		correlations or mean differences
24) self-report instrument	Sex: 53% girls		formulated a priori (i.e. before data
[32]			collection)?
Dutch Physical Activity	n = 44	Poor	8.For convergent validity: Were the
Checklist for Adolescents	Age: 14.2±1.8 years		measurement properties of the
(PAQ-A) [35]	Sex: 41% girls		comparator instrument(s) adequately
			described?
Godin-Shephard Survey	n =102	Poor	10.Were design and statistical methods
[98]	Age: 11.2±0.7 years		adequate for the hypotheses to be
	(n=36), 13.6±0.5		tested?
	years (n=36) olds,		
	16.4±0.8 years		
	(n=30) Sex: 51%		
	girls		
Children's Leisure	n = 108	Poor	4.Were hypotheses regarding
Activities Study Survey	Age 12 years		correlations or mean differences
(CLASS) questionnaire	Sex: 58.3% girls		formulated a priori (i.e. before data
(Modified version) [99]			collection)?
			8.For convergent validity: Were the
			measurement properties of the
			comparator instrument(s) adequately
			described?

^a Age presented as mean age ± SD [range]

^b Based on the COSMIN checklist

^c Based on the 'lowest score counts' method: the design requirement that was scored lowest is shown, if multiple design requirements received the lowest score all design requirements receiving this score are shown.
 ^d Study from previous review

Questionnaire	Study population ^a	Methodological	Design requirement(s) that determined
		quality ^b	the final methodological quality score ^c
Preschool-age Children's	n = 103	Good	5.Were the administrations
Physical Activity	Age: 3.8±0.74 years		independent?
Questionnaire (Pre-PAQ)	Sex: 48% girls		7.Were patients stable in the interim
[58]			period on the construct to be measured?
Energy Balance Related	n = 93 preschoolers	Fair	2.Was there a description of how missing
Behaviors (ERBs) self-			items were handled?
administered primary			
caregivers questionniare			
(PCQ), from the ToyBox-			
study (proxy) [46]			
Children's Leisure	n = 58	Fair	2.Was there a description of how missing
Activities Study Survey	Age: 5.3±0.5 years		items were handled?
(CLASS) (proxy) [100] ^d	[5-6]		
	Sex: 37% girls		
Physical activity	n = 21	Poor	3.Was the sample size included in the
questionnaire for parents	Age: 3- to 5-year		analysis adequate?
of preschoolers in Mexico	olds		
[40]	Sex: percentage girls		
	unknown		
Kid Active Q	n = 20	Poor	3.Was the sample size included in the
(webbased)(proxy) [101]	Age: 4.2±1.3 years		analysis adequate?
	[2-6]		
	Sex: 50% girls		
Chinese version of the	n = 92	Good	3.Was the sample size included in the
Physical Activity			analysis adequate?

Table 2 – Test-retest reliability methodological quality scores

Questionnaire for Older	Age: 8- to 13-year		
Children (PAQ-C) [43]	olds		
	Sex: 45% girls		
ATN-questionnaire [41]	n = 87	Good	3.Was the sample size included in the
ATT-questionnaire [41]		dood	-
	Age: 11- to 12-year		analysis adequate?
	olds		
	Sex: percentage girls		
	unknown		
Children's Leisure	n = 214	Good	5.Were the administrations
Activities Study Survey	Age: 10.9±0.9 years		independent?
Chinese-version	[9-12]		
questionnaire (CLASS-C)	Sex: 62% girls		
[50]			
Out-of-school Physical	n = 151	Good	2.Was there a description of how missing
Activity questionnaire	Age: 11-year-olds		items were handled?
[62]	Sex: 60% girls (in		5.Were the administrations
	total sample n=155)		independent?
			7.Were patients stable in the interim
			period on the construct to be measured?
			9.Were the test conditions similar for
			both measurements? E.g. type of
			administration, environment,
			instructions
the ENERGY-child	n = 730	Fair	2.Was there a description of how missing
questionnaire [48]	Age: [11.3±0.5 -		items were handled?
	12.5±0.6 years]		
	Sex: [47% - 58%		
	girls]		
Self-Administered	n = 72	Fair	8.Was the time interval appropriate?
Physical Activity Checklist	Age: 11.5±0.5 years		
	Sex: 49% girls		

(SAPAC) (Greek version)			
[49]			
Physical Activity	n = 84	Fair	2.Was there a description of how missing
Questionnaire for Older	Age: 9- to 14-year		items were handled?
Children (PAQ-C) [29] ^d	olds		
	Sex: 49% girls		
Girls health Enrichment	n = 68	Fair	10.Were there any important flaws in the
Multisite Study Activity	Age: 9.0±0.6 years		design or methods of the study?
Questionnaire (GAQ)	Sex: 100% girls		
[102] ^d			
Food, Health, and Choices	n = 82 (digital vs.	Fair (both groups)	2.Was there a description of how missing
questionnaire (FHC-Q)	paper)		items were handled?
[37]	Age: <9- to >12-year		
	olds		
	Sex: 51 % girls		
	n = 73 (digital vs.		
	digital)		
	Age: <9- to >12-year		
	olds		
	Sex: 45 % girls		
The South American	n = 161	Fair	2.Was there a description of how missing
Youth/Child	Age: 3- to 10-year		items were handled?
Cardiovascular and	olds		11.For continuous scores: Was an
Environment Study	Sex: 50 % girls		intraclass correlation coefficient (ICC)
(SAYCARE) Physical			calculated?
Activity (PA)			
questionnaire (proxy)			
[66]			
Dutch Physical Activity	n = 192	Fair	5.Were the administrations
Checklist for Children	Age: 8.9±1.7 years		independent?
(PAQ-C) [35]	[5-12]		9.Were the test conditions similar for
	Sex: 53% girls		both measurements? E.g. type of

			administration, environment,
			instructions
Instrument to assess	n = 53	Fair	2.Was there a description of how missing
children's outdoor active	Age: 9.5±0.7 years		items were handled?
play in various locations	[8.3-12.3]		
(proxy) [77]	Sex: 42% girls		
Parent proxy-report of	n = 147	2 months time	Fair: 8.Was the time interval
physical activity and	Age: 6- to 10-year	interval: Fair	appropriate?
sedentary activities	olds, 13- to 14-year		
(proxy) [73]	olds	6 months time	Poor: 8.Was the time interval
	Sex: 51% girls (in	interval: Poor	appropriate?
	total sample n=189)		
Physical Activity	n = 83	Poor	8.Was the time interval appropriate?
Questionnaire for older	Age: 11.0 ± 1.2		
Children (PAQ-C) (Spanish	Sex: 45 % girls		
version) [52]			
Godin Leisure-Time	n = 31	Poor	8.Was the time interval appropriate?
Exercise Questionnaire	Age: 10.6 ± 0.2		
[63]	Sex: 45 % girls		
The Modified Godin	n = 139	Poor	8.Was the time interval appropriate?
Leisure-Time Exercise	Age: 11.1 ± 0.4		
Questionnaire	Sex: 52 % girls		
[45]			
Single-item activity	n = 107	Good	2.Was there a description of how missing
measure [23]	Age: 14.7±0.5		items were handled?
	Sex: 38% girls (age		7.Were patients stable in the interim
	and sex total sample		period on the construct to be measured?
	n=123)		
Web-based and paper-	n = 323	Good	2.Was there a description of how missing
based Physical Activity	Age 12.8 years		items were handled?

Questionnaire for Older	Sex: 51% girls		5.Were the administrations
Children (PAQ-C) [28]	(age and sex total		independent?
	sample n=459)		11.For continuous scores: Was an
			intraclass correlation coefficient (ICC)
			calculated?
An adapted version of the	n = 150	Good	5.Were the administrations
Assessment of Physical	Age: 13.6 ± 1.1		independent?
Activity Levels	Sex: 52 % girls		7.Were patients stable in the interim
Questionnaire (APALQ)			period on the construct to be measured?
[53]			9.Were the test conditions similar for
			both measurements? E.g. type of
			administration, environment,
			instructions
International Physical	n = 92	Good	3.Was the sample size included in the
Activity Questionnaire -	Age: 15.9±1.4 years		analysis adequate?
Short Form (IPAQ - SF)	[12-18]		7.Were patients stable in the interim
[84]	Sex: 53% girls		period on the construct to be measured?
			9.Were the test conditions similar for
			both measurements? E.g. type of
			administration, environment,
			instructions
			11.For continuous scores: Was an
			intraclass correlation coefficient (ICC)
			calculated?
Child and Adolescent	n = 77	Good	3.Was the sample size included in the
Physical Activity and	Age: 12±0.8 years		analysis adequate?
Nutrition survey	[11-14]		7.Were patients stable in the interim
(CAPANS-PA) recall	Sex: 51% girls		period on the construct to be measured?
questionnaire [103]			
Activity Questionnaire for	n = 53	Good	2.Was there a description of how missing
Adults and Adolescents	Age: 14.1±1.4 years		items were handled?
(AQuAA) [21]	Sex: 43% girls		3.Was the sample size included in the
			analysis adequate? 7.Were patients
	L		

			stable in the interim period on the
			construct to be measured?
			11.For continuous scores: Was an
			intraclass correlation coefficient (ICC)
			calculated?
Godin-Shephard Survey	n =102	Fair	11.For continuous scores: Was an
[98]	Age: 11.2±0.7 years		intraclass correlation coefficient (ICC)
	(n=36), 13.6±0.5		calculated?
	years (n=36) olds,		
	16.4±0.8 years		
	(n=30) Sex: 51%		
	girls		
VISA-TEEN questionnaire	n = 228	Fair	8.Was the time interval appropriate?
[104]	Age 15.4±1.6 years		
[]	Sex: 46% girls		
	(age and sex total		
	sample n=396)		
Children's Leisure	n = 108	Fair	2.Was there a description of how missing
		rall	items were handled?
Activities Study Survey	Age 12 years		items were nanuleu?
(CLASS) questionnaire	Sex: 58.3% girls		
(Modified version) [99]			
Oxford Physical Activity	n = 104	Fair	2.Was there a description of how missing
Questionnaire (OPAQ)	Age: 14.7±0.5		items were handled?
[23]	Sex: 38% girls (age		
	and sex total sample		
	n=123)		
Quantification de l'activité	n = 121	Fair	7.Were patients stable in the interim
physique en altitude chez	Age: 8- to 16-year		period on the construct to be measured?
les enfants (QAPACE)	olds		
[105] ^d	Sex: 54% girls		

Oxford Physical Activity	n = 87	Fair	2.Was there a description of how missing
Questionnaire (OPAQ)	Age: 13.1±0.9 years		items were handled?
[24] ^d	Sex: 45% girls		
World Health	n = 71	Fair	2.Was there a description of how missing
Organization Health	Age: 14.9±1.6 [13-		items were handled?
Behavior in	18]		
Schoolchildren	Sex: 56% girls		
questionnaire (WHO			
HBSC) [106] ^d			
Selected indicators from	n = 95 (n=44 11-	Fair	2.Was there a description of how missing
the Health Behaviour in	year-olds, n=51 15-		items were handled?
School-aged Children	year-olds)		
(HBSC) questionnaire	Age: [11.7 ± 0.4 to		
(Chinese version) [107]	15.8 ± 0.3 years]		
	Sex: 46% girls		
Selected physical activity	n = 693	Fair	2.Was there a description of how missing
items of the international	Age: 11.1±0.5 years		items were handled?
Health Behavior in School-	and 15.1±0.5yrs		
aged Children (HBSC)	Sex: 49.1% girls		
questionnaire (Czech			
version) [108]			
Measures of in-school and	n = 68	Fair	2.Was there a description of how missing
out-of-school physical	Age: 15.4 years		items were handled?
activity, and travel	Sex: 47% girls		
behaviors of the			
international Healthy			
Environments and active			
living in teenagers – Hong			
Kong (iHealt(H)) study			
[47]			
Physical Activity and	n = 21	Fair	3.Was the sample size included in the
Lifestyle Questionnaire	Age: 13.7±0.8 years		analysis adequate?
			8.Was the time interval appropriate?

[33]and sex total sample re40FairAuser base inputs of how missing readsThe South Americann=177Fair2.Was there a description of how missing removements and of dasYouth/ChildAge: 11- to 18-yearItems were handled?Cardiovascular andoldsItems were handled?Surforment StudySex: 58 % girlsItems were handled?(SAYCARB) PhysicalItemsItercation coefficient (ICC) calculated?Activity (PA)ItemsFair2.Was there a description of how missing questionnaire (66)guestionnaire on(study 2)Fair2.Was there a description of how missing questionnaire onguestionnaire on(study 2)Fair2.Was there a description of how missing questionnaire onguestionnaire on(study 2)Items were handled?guestionnaire on(study 2)ItemsItems were handled?girls unknownSex: percentage of girls unknownItems were handled?Dutch Physical Activityn=94FairSuffer endeministrations(PQ-A)[12-17]Sex: 55% girlsItems were handled?[35]Sex: 55% girlsItems were handled?administration, environment, instructions[37]Age: 14-5±1.1 yearsSuffer endeministration environment, instructions[37]Sex: 55% girlsSex: 55% girlsSuffer endeministration environment, instructions[37]Sex: 55% girlsAge: 14-5±1.1 yearsSuffer endeministration environment, instructions[37]Sex: 55% girls (age an	(PALQ) (Greek version)	Sex: 43% girls (age		
In E-South Americann = 177Fair2.Was there a description of how missing items were handled?Youth/ChildAge: 11- to 18-yearitems were handled?8.Was the time interval appropriate?Cardiovascular andolds11.For continuous scores: Was an intraclass correlation coefficient (ICC)Sex: 58 % girls11.For continuous scores: Was an intraclass correlation coefficient (ICC)Activity (PA)questionnaire [66]Self-administeredn = 61 (study 1), 68Fair2.Was there a description of how missing items were handled?guestionnaire on children's travel to school girls unknownAge: 11- to 14-year girls unknownDutch Physical Activityn = 94Fair5.Were the administrationsDutch Physical Activityn = 94Fair5.Were the cst conditions similar for independent?[35]Sex: 58% girls[36]Sex: 55% girls3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 years instrument (Singaporean instructions[3-Day Physical Activityn = 21 ne221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21 ne221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21 ne2219Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21 ne2219Poor8.Was the time interva	[33]	and sex total sample		
Youth/ChildAge: 11- to 18-yearitems were handled?Cardiovascular andolds8.Was the time interval appropriate?Environment StudySex: 58 % girls1.For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?Activity (PA)questionnaire [66]Self-administeredn = 61 (study 1).60Pair2.Was there a description of how missing items were handled?Guestionnaire on children's travel to schoolAge: 11- to 14-year[39]oldsSex: percentage of girls unknownDutch Physical Activityn = 94Fair5.Were the administrations(FAQ-A)[12-17]-9.Were the test conditions similar for administration[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 years-instrument (Singaporean instrument (Singaporean[13-16]-version) [42]Sex: 53% girls (age and sex total sample-3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Record (3DPARecord)Age: 13.7±0.8 years3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Record (3DPARecord)Age: 13.7±0.8 years3-Day Physical Activity<		n=40)		
Cardioascular and Environment Studyolds8.Was the time interval appropriate?Environment StudySex: 58 % girls11.For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?(SAYCARE) Physicaln = 61 (study 1), 68Fair2.Was there a description of how missing items were handled?guestionnaire [66]n = 61 (study 2)Fair2.Was there a description of how missing items were handled?Gelf-administeredn = 61 (study 2)Fair2.Was there a description of how missing items were handled?Guestionnaire on children's travel to schoolAge: 11- to 14-year girls unknownSex: percentage of girls unknown5.Were the administrationsDutch Physical Activityn = 94Fair5.Were the administrationsChecklist for AdolescentsAge: 13.6±1.4 years (PAQ-A)Juster 13.6±1.4 yearsIndependent?(PAQ-A)[12-17]9.Were the test conditions similar for instructions[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions[3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (30PARecall)Age: 14.5±1.1 years instructionsSex: 53% girls (age and sex total sample n=221)8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Record (3DPARecord)Age: 13.7±0.8 years (Greek version) [33]Sex: 43% girls (age and sex total sample8.Was the time interval appropriate?	The South American	n = 177	Fair	2.Was there a description of how missing
Environment StudySex: 58 % girls11.For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?Activity (PA)questionnaire [66]Self-administeredn = 61 (study 1), 68Fair2.Was there a description of how missing items were handled?questionnaire on(study 2)(39)oldsSex: percentage of girls unknownDutch Physical Activityn = 94Fair5.Were the administrationsChecklist for AdolescentsAge: 13.6±1.4 years (PAQ-A)[35]Sex: 55% girls3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARccall)Age: 14.5±1.1 years andsex total sample n=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Greek version] [42]Sex: 53% girls (age and sex total sample n=221)Poor8.Was the time interval appropriate?Greek version] [33]Sex: 43% girls (age and sex total sample and sex total samplePoor8.Was the time interval appropriate?	Youth/Child	Age: 11- to 18-year		items were handled?
(SAYCARE) Physical Activity (PA)intractase correlation coefficient (ICC) calculated?guestionnaire [66]n = 61 (study 1), 68Fair2.Was there a description of how missing items were handled?guestionnaire on (study 2)(study 2)items were handled?items were handled?children's travel to school (girls unknownAge: 11- to 14-year girls unknownSex: percentage of girls unknownSwere the administrationsDutch Physical Activityn = 94FairS.Were the administrations(fAQ-A)[12-17]Swere the set conditions similar for[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?rescall (3DPARecall)Age: 14.5±1.1 years instrument (Singaporean u adm sex total sample n=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21 Record (3DPARecord)Rec 13.7±0.8 years (Greek version) [33]Poor8.Was the time interval appropriate?	Cardiovascular and	olds		8.Was the time interval appropriate?
Activity (PA) questionnaire [66]Isainacalculated?Self-administeredn = 61 (study 1), 68Fair2.Was there a description of how missing items were handled?questionnaire on(study 2)items were handled?items were handled?children's travel to schoolAge: 11- to 14-yearItems were handled?[39]oldsItems were handled?gars unknownSex: percentage of girls unknownSwere the administrationsDutch Physical Activityn = 94FairSwere the administrations(PAQ-A)[21-7]Jwere the st conditions similar for both measurements? E.g. type of administration, environment, instructions[35]Sex: 55% girlsSwere the administration and diministration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?recall (3DPARecall)Age: 14.5±1.1 yearsInterval administration, environment, instructionsinstrument (Singaporean instrument (Singaporean instructionsI13-16]Age: 14.5±1.1 Poor3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21PoorSwas the time interval appropriate?3-Day Physical ActivityN = 21PoorSwas the time interval appropriate?Greek version [33]Sex: 43% girls (age adi sex total sample iadi sex total sampleInterval appropriate?Greek version [33]Sex: 43% girls (age adi sex total sampleInterval appropriate?	Environment Study	Sex: 58 % girls		11.For continuous scores: Was an
questionnaire [66]n = 61 (study 1), 68Fair2.Was there a description of how missing items were handled?guestionnaire on (study 2) children's travel to school [39]Age: 11 to 14-year olds Sex: percentage of gtrls unknownImage: 11 to 14-year pigrls unknownImage: 11 to 14-year pigrls unknownImage: 11 to 14-year pigrls unknownDutch Physical Activityn = 94Fair5.Were the administrations independent?Dutch Physical Activityn = 94Fair5.Were the administrations(PAQ-A)[12-17]9.Were the test conditions similar for both measurements? E.g. type of administration, environment, instructions[35]Sex: 55% girlsboth measurements? E.g. type of administration3-Day Physical Activityn = 106 Age: 14.5±1.1 years instrument (Singaporean instrument (Singaporean in=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21 Age: 13.7±0.8 years (Greek version) [32]Sex: 43% girls (age and sex total sample and sex total sample and sex total sample8.Was the time interval appropriate?	(SAYCARE) Physical			intraclass correlation coefficient (ICC)
Self-administeredn = 61 (study 1). 68Fair2.Was there a description of how missing items were handled?questionnaire on(study 2)items were handled?items were handled?(alidren's travel to schoolAge: 11 to 14-yearitems were handled?items were handled?[39]oldsSex: percentage of girls unknownSex: percentage of girls unknownitems were handled?Dutch Physical Activityn = 94Fair5.Were the administrations independent?Checklist for AdolescentsAge: 13.6 ± 1.4 yearsindependent?(PAQ-A)[12-17]9.Were the test conditions similar for administration[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5 ± 1.1 years instrument (Singaporean n=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Greek version [33]Sex: 43% girls (age and sex total sampleHoor8.Was the time interval appropriate?Greek version [33]Sex: 43% girls (age and sex total sampleHoorHoorSource ActivityNase total sa	Activity (PA)			calculated?
questionnaire on children's travel to school [39](study 2) Age: 11- to 14-year oldsitems were handled?[39]oldsSex: percentage of girls unknownitems were handled?Dutch Physical Activityn = 94Fair5.Were the administrations independent?Checklist for Adolescents (PAQ-A)Age: 13.6±1.4 years [12-17]S.Were the test conditions similar for both measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall) version [42]Age: 13.7±0.8 years ads ex total sample n=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Greek version) [33]Sex: 53% girls (age and sex total sample and sex total samplePoor8.Was the time interval appropriate?	questionnaire [66]			
 children's travel to school Age: 11- to 14-year olds Sex: percentage of girls unknown Dutch Physical Activity n = 94 Fair S.Were the administrations independent? 9.Were the test conditions similar for both measurements? E.g. type of administration, environment, instructions Sex: 55% girls Sex: 53% girls (age and sex total sample n=221) Sex: 43% girls (age and sex total sample 	Self-administered	n = 61 (study 1), 68	Fair	2.Was there a description of how missing
[39]oldssex: percentage of girls unknownsex: percentage of girls unknownsex: percentage of girls unknownsex: percentage of girls unknownDutch Physical Activityn = 94Fair5.Were the administrationsChecklist for AdolescentsAge: 13.6±1.4 yearsindependent?(PAQ-A)[12-17]9.Were the test conditions similar for both measurements? E.g. type of administration, environment, instructions[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 yearsIntervalPoorstrument (Singaporean ursion) [42]Sex: 53% girls (age and sex total sampleInterval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Greek version [33]Sex: 43% girls (age and sex total sampleInterval appropriate?	questionnaire on	(study 2)		items were handled?
Sex: percentage of girls unknownSex: percentage of girls unknownImage: Sex: percentage of girls unknownFairSex: percentage of girls unknownDutch Physical Activityn = 94Fair5.Were the administrations independent?(PAQ-A)[12-17]9.Were the test conditions similar for both measurements? E.g. type of administration, environment, instructions[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5 ± 1.1 years instrument (Singaporean instrument (Singaporean in=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Greek version [33]Sex: 43% girls (age and sex total sample and sex total sampleFoor8.Was the time interval appropriate?	children's travel to school	Age: 11- to 14-year		
girls unknownFairS.Were the administrationsDutch Physical Activityn = 94FairS.Were the administrationsChecklist for AdolescentsAge: 13.6±1.4 yearsindependent?(PAQ-A)[12-17]9.Were the test conditions similar for[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 yearsNere Sa% girls (age and sex total sampleAge: 13.7±0.8 years3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?3-Day Physical Activityn = 10Poor8.Was the time interval appropriate?Greek version [33]Sex: 43% girls (age and sex total sample and sex total samplePoor8.Was the time interval appropriate?	[39]	olds		
Dutch Physical Activityn = 94Fair5.Were the administrationsChecklist for AdolescentsAge: 13.6±1.4 yearsindependent?(PAQ-A)[12-17]9.Were the test conditions similar for[35]Sex: 55% girlsboth measurements? E.g. type ofadministration, environment, instructionsadministration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 yearsAge: 14.5±1.1 yearsAge: 14.5±1.1 yearsinstrument (Singaporean ursion) [42]Sex: 53% girls (age and sex total sample n=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Greek version) [33]Sex: 43% girls (age and sex total sample and sex total sampleAge: 13.7±0.8 years		Sex: percentage of		
Checklist for AdolescentsAge: 13.6±1.4 yearsindependent?(PAQ-A)[12-17]9.Were the test conditions similar for[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 years instrument (Singaporean[13-16]Poorversion) [42]Sex: 53% girls (age and sex total sample n=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Greek version) [33]Sex: 43% girls (age and sex total sample and sex total samplePoor8.Was the time interval appropriate?		girls unknown		
(PAQ-A)[12-17]9.Were the test conditions similar for both measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 years instrument (Singaporean n=221)Poor8.Was the time interval appropriate?3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?	Dutch Physical Activity	n = 94	Fair	5.Were the administrations
[35]Sex: 55% girlsboth measurements? E.g. type of administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 years8.Was the time interval appropriate?instrument (Singaporean version) [42][13-16]	Checklist for Adolescents	Age: 13.6±1.4 years		independent?
administration, environment, instructions3-Day Physical Activityn = 106Poor8.Was the time interval appropriate?Recall (3DPARecall)Age: 14.5±1.1 yearsIIinstrument (Singaporean[13-16]IIversion) [42]Sex: 53% girls (ageIIand sex total sampleIIIn=221)n=21Poor8.Was the time interval appropriate?Record (3DPARecord)Age: 13.7±0.8 yearsII(Greek version) [33]Sex: 43% girls (ageIIand sex total sampleIIIand sex total sampleIIIBecord (3DPARecord)Age: 13.7±0.8 yearsIIImage: Image:	(PAQ-A)	[12-17]		9.Were the test conditions similar for
Image: Sec: 33% girls (age)Poors.Was the time interval appropriate?8.Call (3DPARecall)Age: 14.5±1.1 yearsFeall (3DPARecall)Age: 14.5±1.1 years1.151I.161Fean (13.16)Fean (13.16)1.161Sex: 53% girls (age)Fean (13.16)Fean (13.16)1.162Sex: 53% girls (age)Fean (13.16)Fean (13.16)1.162Sex: 53% girls (age)Fean (13.16)Fean (13.16)1.172Sex: 13.7±0.8 yearsSex: 43% girls (age)Sex: 43% girls (age)(Greek version) [33]Sex: 43% girls (age)Fean (13.16)1.172Sex: 43% girls (age)Fean (13.16)1.172 <td>[35]</td> <td>Sex: 55% girls</td> <td></td> <td>both measurements? E.g. type of</td>	[35]	Sex: 55% girls		both measurements? E.g. type of
And the second				administration, environment,
Recall (3DPARecall)Age: 14.5±1.1 yearsinstrument (Singaporean[13-16]version) [42]Sex: 53% girls (ageand sex total sampleand sex total samplen=221)n=213-Day Physical Activityn = 21Record (3DPARecord)Age: 13.7±0.8 years(Greek version) [33]Sex: 43% girls (ageand sex total sampleInterval appropriate?				instructions
instrument (Singaporean [13-16] version) [42] Sex: 53% girls (age and sex total sample n=221) 3-Day Physical Activity n = 21 Poor 8.Was the time interval appropriate? Record (3DPARecord) Age: 13.7±0.8 years (Greek version) [33] Sex: 43% girls (age and sex total sample	3-Day Physical Activity	n = 106	Poor	8.Was the time interval appropriate?
version) [42] Sex: 53% girls (age and sex total sample n=221) n=21 Poor 8.Was the time interval appropriate? Record (3DPARecord) Age: 13.7±0.8 years (Greek version) [33] Sex: 43% girls (age and sex total sample	Recall (3DPARecall)	Age: 14.5±1.1 years		
and sex total sample n=221)and sex total sample n=221)and sex total sample3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Record (3DPARecord)Age: 13.7±0.8 yearsImage: 13.7±0.8 yearsImage: 13.7±0.8 years(Greek version) [33]Sex: 43% girls (age) and sex total sampleImage: 13.7±0.8 years	instrument (Singaporean	[13-16]		
n=221)Image: Network<	version) [42]	Sex: 53% girls (age		
3-Day Physical Activityn = 21Poor8.Was the time interval appropriate?Record (3DPARecord)Age: 13.7±0.8 years(Greek version) [33]Sex: 43% girls (age and sex total sample		and sex total sample		
Record (3DPARecord)Age: 13.7±0.8 years(Greek version) [33]Sex: 43% girls (age and sex total sample		n=221)		
(Greek version) [33] Sex: 43% girls (age and sex total sample	3-Day Physical Activity	n = 21	Poor	8.Was the time interval appropriate?
and sex total sample	Record (3DPARecord)	Age: 13.7±0.8 years		
	(Greek version) [33]	Sex: 43% girls (age		
n=40)		and sex total sample		
		n=40)		

Recess Physical Activity	n = 113	Poor	8.Was the time interval appropriate?
Recall (RPAR) [95]	Age 13.1±0.7 years		
	Sex: 48% girls		
Refined 60-min MVPA	n = 138	Poor	8.Was the time interval appropriate?
screening measure [109] ^d	Age: 12.1±0.9		
	Sex: 65% girls		
MVPA scores of the Health	n = 998	Poor	8.Was the time interval appropriate?
Behavior in School-aged	Age: 12.7 ± 1.4		
Children (HBSC) Research	Sex: 50 % girls		
Protocol [90]			
MVPA scores of the	n = 998	Poor	8.Was the time interval appropriate?
International Physical	Age: 12.7 ± 1.4		
Activity Questionnaire	Sex: 50 % girls		
Short form (IPAQ-SF) [90]			
Moderate and vigorous	n = 128	Poor	8.Was the time interval appropriate?
physical activity items of	Age: 12.2±0.6 years		
the Youth Risk Behavior	(in total sample		
Survey (YRBS)	n=139)		
[83]	Sex: 53% girls		

^a Age presented as mean age ± SD [range]

^b Based on the COSMIN checklist

^c Based on the 'lowest score counts' method: the design requirement that was scored lowest is shown, if multiple

design requirements received the lowest score all design requirements receiving this score are shown.

^d Study from previous review

Table 3. Measurement error methodological quality scores
--

Questionnaire	Study population ^a	Methodological	Design requirement(s) that
		quality ^b	determined the final methodological
			quality score ^c
Preschool-age Children's	n = 103	Good	5.Were the administrations
Physical Activity			independent?

Questionnaire (Pre-PAQ)	Age: 3.8±0.74		7.Were patients stable in the interim
[58]	years		period on the construct to be
	Sex: 48% girls		measured?
the ENERGY-child	n = 730	Fair	2.Was there a description of how
questionnaire [48]	Age: [11.3±0.5 -		missing items were handled?
	12.5±0.6 years]		
	Sex: [47% - 58%		
	girls]		
Dutch Physical Activity	n = 192	Fair	5.Were the administrations
Checklist for Children	Age: 8.9±1.7 years		independent?
(PAQ-C) [35]	[5-12]		9.Were the test conditions similar fo
	Sex: 53% girls		both measurements? E.g. type of
			administration, environment,
			instructions
Children's Leisure	n = 109	Fair	2.Was there a description of how
Activities Study Survey	Age: 10.6±0.8		missing items were handled?
(CLASS) [100] ^d	years [10-12] (in		
	total sample		
	n=111)		
	Sex: 63% girls		
ATN-questionnaire	n = 87	Good	3.Was the sample size included in th
(days/week type of	Age: 11- to 12-year		analysis adequate?
transportation) [41]	olds		
	Sex: percentage		
	girls unknown		
3-Day Physical Activity	n = 65	Good	3.Was the sample size included in th
Recall (3DPARecall)	Age: 12.5 ±1.1		analysis adequate?
[19] ^d	years		

	Sex: 64% girls (age		
	and sex in total		
	sample n=320)		
Self-Administered	n = 84	Good	3.Was the sample size included in the
Physical Activity	Age: 12.5 ±1.1		analysis adequate?
Checklist (SAPAC)	years		
(modified) [19] ^d	Sex: 64% girls (age		
	and sex in total		
	sample n=320)		
Measures of in-school	n = 68;	Fair	2.Was there a description of how
and out-of-school	Age: 15.4 years		missing items were handled?
physical activity, and	Sex: 47% girls		
travel behaviors of the			
international Healthy			
Environments and active			
living in teenagers –			
Hong Kong (iHealt(H))			
study [47]			
Dutch Physical Activity	n = 94	Fair	5.Were the administrations
Checklist for Adolescents	Age: 13.6±1.4		independent?
(PAQ-A) [35]	years [12-17]		9.Were the test conditions similar fo
	Sex: 55% girls		both measurements? E.g. type of
			administration, environment,
			instructions

ME: Measurement error; PoA: Percentage of Agreement; PA: physical activity; PE: Physical education; MPA: Moderate Physical Activity; VPA: Vigorous Physical Activity

^a Age presented as mean age ± SD [range]

^b Based on the COSMIN checklist

 $^{\rm c}$ Based on the 'lowest score counts' method: the design requirement that was scored lowest is shown, if multiple

design requirements received the lowest score all design requirements receiving this score are shown.

 $^{\rm d}\,Study$ from previous review