Appendix

Scoring instruments

Appendix table 1 – Paediatric Basic Life Support assessment instrument

The greyed-out *italic* item is not applicable in the video-recorded scenario tests

				Likert scale		
PBLS	Item description	1	2	3	4	5
	Recognise	No attempt to recognise	-	Recognises unresponsiveness	-	Efficiently recognises
Responsiveness	unresponsiveness	unresponsiveness	1	but inefficiently		unresponsiveness
Call for help	Recognise need for help and alert surroundings both	No attempt to call for help	-	Calls for help but not done efficiently	-	Efficiently calls for help
	by loud verbal call out and using telephone					
Open airway	Establish open airways including mouth inspection,	No attempt to open airway	-	Establishes open airway but inappropriately	-	Appropriately establishes open airway
Орен ан жау	appropriate head and jaw positioning.					

						Efficiently
						assesses
	Assess breathing			Assesses		breathing and
	and recognise	No attempt to	-	 breathing but 		recognises
Check breathing	respiratory arrest	assess breathing		inefficiently		respiratory arrest
	or abnormal					or abnormal
	breathing					breathing
		No attempt to		Delivers some		Consistently
	Provide high	provide rescue	_	effective rescue		provides effective
Rescue breaths	quality initial	breaths		breaths		rescue breaths
	rescue breaths	breaths		breatiis	ı	rescue preatifs
	Provide high	No attempt to		Provides some		Consistently
	quality	provide	-	high quality	-	provides high
	compressions.	compressions				quality
Compressions	Adequate rate,					compressions
	compression					
	depth and correct					
	hand placement					
	Provide high	No attempt to		Provides some		Consistently
	quality	provide ventilations	-	effective	-	provides effective
	ventilations in	provide ventilations		ventilations		ventilations
Ventilations	general during					
	CPR with					
	adequate chest					
	rise.					
Time factor	Act effectively	No optimal use of	-	Optimal time use	-	Consistently
Time factor	Act checuvery	140 Optimal use of		Spanial anne ase	•	Соловсения

	with minimised	time		inconsistently	optimal time use	
	hands-off time, no					
	delays in					
	treatment and fast					
	call for help					
	Call for Automatic	No attempt to call		Calls for AED but	1	Adequate use of
Use of AED	External	for AED	-	used inadequately	-	AED
	defibrillator and					
	appropriate use					

Appendix table 2 Foreign Body Airway Obstruction Management assessment instrument

The greyed-out *italic* items are not applicable in the video-recorded scenario tests.

FBAOM	Item			Likert scale		
scoring	description	1	2	3	4	5
Identify different stages of foreign body airway obstruction	Distinguish effective and ineffective cough	No attempt to distinguish cough	-	Inconsistently distinguishes effective and ineffective coughs Recognises	-	Consistently distinguishes effective and ineffective coughs
Identify consciousness	unresponsivenes s	recognise responsiveness	-	responsiveness but inefficiently	-	recognises responsiveness
Call for help	Recognise need for help and alert surroundings	No attempt to call for help	-	Calls for help but not done efficiently	-	Efficiently calls for help
Back blows	Provide high quality back blows with adequate force and correct placement of impact	No attempt to provide back blows	-	Provides some high quality back blows	-	Consistently provides high quality back blows
Chest thrusts	Provide high	No attempt to	-	Provides some high	-	Consistently

/abdominal	quality chest	provide thrusts		quality thrusts		provides high
thrusts	thrust or					quality thrusts
according to	abdominal					
age	thrust according					
uge	to age					
Identify loss of	Recognise	No reaction to				Reacts
	changes in	loss of	-	Reacts appropriately	-	appropriately and
consciousness	condition and	consciousness		but inefficiently		efficiently
and change to	act					
CPR	appropriately					
	Assess breathing					Efficiently
	and recognise					assesses breathing
	respiratory	No attempt to		Assesses breathing		and recognises
	arrest or	assess breathing		but inefficiently		respiratory arrest
Assessment of	abnormal					or abnormal
breathing	breathing					breathing
	requiring					Dreathing
	ventilator					
	support	No otto ment to				Consistently
	Provide high	No attempt to		Provides some		Consistently
	quality	provide		effective ventilations		provides effective
	ventilations if	ventilations				ventilations
Ventilation	patient stops					
	breathing with					
	adequate chest					
	rise					

Flowchart for collecting validity evidence

Appendix figure 1: Flowchart for collecting validity evidence

The flowchart describes the five sources of validity evidence and the study design to collect evidence in the five categories.

Step 1 - Content evidence

Content evidence examines the relation between the content of the assessment and the skill level (construct) it is intended to measure and evidence sources include expert reviews.

International experts identified the essential assessment items in a recent study and the research group developed descriptive anchors for 5-point scoring scales for each item.

Step 2 - Response process

The response process examines the actions of test takers and raters and to the extent it fits the underlying intended construct and evidence sources include how to minimise bias.

Pilot tests of the scoring instruments was performed.
Videorecorded scenario tests were scored with the
assessment instrument by two blinded raters after participating
in a rater training course.

Step 3 - Internal structure

The internal structure evidence examines how well the internal components of the test match the intended construct and sources of evidence involves the reliability of the measurements and internal consistency of the items.

Generalizability theory was used to analyse variance in assessment scores. A d-study was performed to determine the number of raters and tests needed for reliable judgments. Internal consistency was assessed by Cronbach's alpha

Step 4 - Relation to other variables

Relation to other variables examines the correlation between the scores and other measurements were a correlation would be expected and evidence sources includes external measures such as different levels of training or other measurements of performance such as overall performance scores.

Assessment scores for participants with three different expected skill levels were analysed.

Correlation of the overall performance score and the assessment scores were analysed.

Step 5 - Consequences

Consequences evidence examines the intended and unintended consequences of assessment results and sources includes consequences of passing and possible unintended effects of testing.

The contrasting groups method was used to determine a pass/fail level and teh consequences for the groups of participants were analysed.

Individual item scores

The figures (appendix figure 2 and 3) portrait the mean individual item scores by each group. The mean individual item scores for PBLS increased for all items from untrained to trained laypersons to lifeguards.

ANOVA and post hoc analysis of individual items for both PBLS and FBAOM items across the three groups is visible in appendix table 3. The ANOVA was significant for all PBLS and FBAOM items except the FBAOM item "Call for help".

In post hoc analysis for PBLS all items but "Call for help" differed from untrained to trained laypersons. There were no statistically significant differences between trained laypersons and lifeguards. "Call for help" was significantly higher in the lifeguard group compared with the untrained layperson group (t(21)=-3.27 p=0.008)

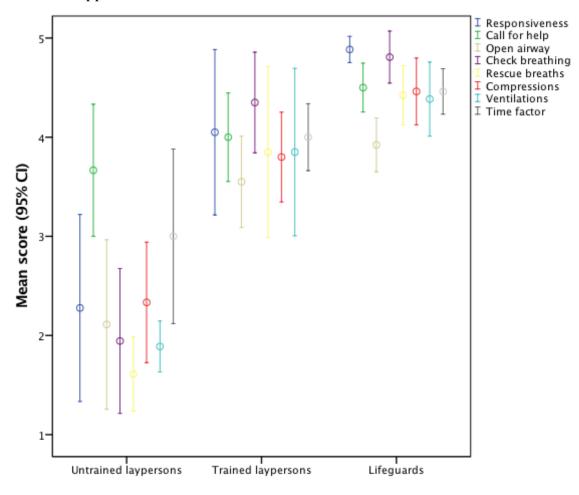
For the FBAOM test "Call for help" was non-significant in an one-way ANOVA across the three groups (F(2,28) = 2.27, p=0.12). The post hoc analysis found significant differences between untrained and trained laypersons' scores for "backblows" (t(18)=-5.40 p<0.001), trained laypersons and lifeguards (t(21)=3.76 p=0.003) and not between intermediates and lifeguards (t(21)=-2.07 p=0.14).

The FBAOM items "Chest thrusts "and "change to CPR" was significantly different between all three individual groups (p<0.001).

The internal consistency analysed by Cronbach's alpha standardised items were 0.94 for the PBLS eight item scale and 0.64 for the FBAOM four item scale. The ANOVA above identified the call for help score as non-significantly different and an analysis of the internal consistency when this item was left out resulted in a Cronbach's alpha of 0.93.

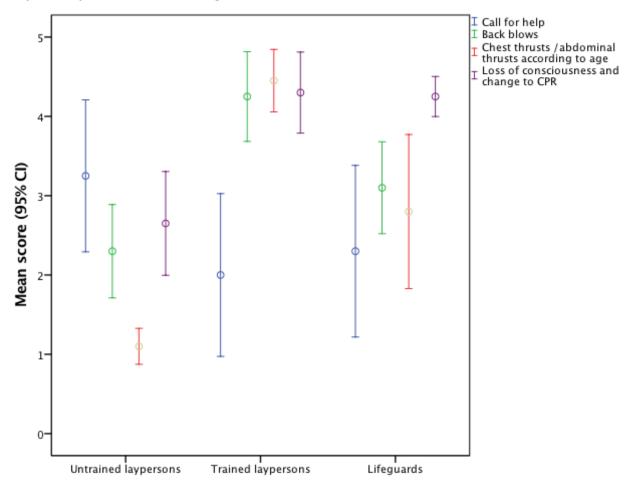
Appendix Figure 2 – Paediatric Basic Life Support scores for individual assessment items by group

The figure illustrates the mean scores with 95% confidence intervals for the eight Paediatric Basic Life Support items.



Appendix figure 3 – Foreign Body Airway Obstruction scores for individual assessment items by group

The figure illustrates the mean scores with 95% confidence intervals for the four Foreign Body Airway Obstruction Management items.



Appendix table 3

The table shows the results of individual items ANOVA and the post hoc Bonferroni corrected analysis for both Paediatric Basic Life Support and Foreign Body Airway Obstruction Management items.

Item	Untrain	Trained	Lifeguar	ANOVA	Post hoc analysis			
	ed	layperso	ds,		Untrained	Untraine	Trained	
	layperso	ns, mean	mean		vs. trained	d	laypers	
	ns, mean	(95%CI)	(95%CI)		layperson	layperso	ons vs.	
	(95%CI)				s	ns vs.	lifeguar	
						lifeguar	ds	
						ds		
PBLS	<u> </u>	<u> </u>	<u> </u>	<u> </u>	<u> </u>	<u> </u>	<u> </u>	
				F(2,30)				
Responsive	2.55	4.05	4.88	=14.77		t(21)=-	t(21)=-	
ness	(1.51 -	(3.21 -	(4.75 -	p<0.00	t(18)=-2.56	5.79	2.54	
	3.58)	4.88)	5.02)	1	p=0.01	p<0.001	p=0.19	
				F(2,30)				
Call for help	3.65			=5.47		t(21)=-	t(21)=-	
Can for help	(3.06 -	4 (3.55 -	4.5 (4.25	p=0.00	t(18)=-1.08	3.27	2.32	
	4.23)	4.44)	- 4.75)	9	p=0.65	p=0.008	p=0.19	
				F(2,29)				
Open	2.11	3.55	3.92	=16.63		t(20)=-	t(21)=-	
airway	(1.25 -	(3.08 -	(3.65 -	p<0.00	t(17)=-3.5	5.33	1.64	
	2.96)	4.01)	4.19)	1	p<0.001	p<0.001	p=0.73	

				F(2,30)			
Check	1.85	4.35	4.81	=54.89		t(21)=-	t(21)=-
breathing	(1.17 -	(3.84 -	(4.54 -	p<0.00	t(18)=-6.7	10.04	1.91
	2.52)	4.85)	5.07)	1	p<0.001	p<0.001	p=0.40
				F(2,30)			
Rescue	1.65	3.85	4.42	=38.60		t(21)=-	t(21)=-
breaths	(1.31 -	(2.98 -	(4.12 -	p<0.00	t(18)=-5.38	13.58	1.56
	1.98)	4.71)	4.72)	1	p<0.001	p<0.001	p=0.27
				F(2,30)			
Compressio	2.45		4.46	=25.66		t(21)=-	t(21)=-
ns	(1.85 -	3.8 (3.34	(4.12 -	p<0.00	t(18)=-4.09	6.95	2.66
	3.04)	- 4.25)	4.8)	1	p<0.001	p<0.001	p=0.08
				F(2,30)			
Ventilations		3.85	4.38	=30.80		t(21)=-	t(21)=-
ventulations	1.9 (1.67	(3.00 -	(4.01 -	p<0.00	t(18)=-5.05	11.57	1.41
	- 2.12)	4.69)	4.76)	1	p<0.001	p<0.001	p=0.33
				F(2,30)			
Time factor	3.05		4.46	=11.85		t(21)=-	t(21)=-
Time factor	(2.26 -	4 (3.66 -	(4.23 -	p<0.00	t(18)=-2.53	4.36	2.6
	3.83)	4.33)	4.69)	1	p=0.01	p<0.001	p=0.37
FBAOM	<u>I</u>	ı	ı	ı	<u> </u>	<u>I</u>	1
Call fan hale	3.25	2 (0.97 -	2.18	F(2,28)	t(18)=2.02	t(19)=1.7	t(19)=-
Call for help	(2.29 -	3.02)	(1.18 -	=2.27	p=0.18	2 p=0.29	0.29

	4.20)		3.18)	p=0.12			p>0.999
				F(2,30)			
Back blows		4.25		=15.30		t(21)=-	t(21)=3.
Dack Diows	2.3 (1.71	(3.68 -	3 (2.52 -	p<0.00	t(18)=-5.4	2.07	76
	- 2.88)	4.81)	3.48)	1	p<0.001	p=0.14	p=0.003
Chest							
thrusts							
/abdominal				F(2,30)			
thrusts		4.45	2.92	=40.11	t(18)=-	t(21)=-	t(21)=3.
according	1.1 (0.87	(4.05 -	(2.19 -	p<0.00	16.69	4.64	71
to age	- 1.32)	4.84)	3.65)	1	p<0.001	p<0.001	p<0.001
Loss of							
consciousne				F(2,29)			
ss and		4.45	2.92	=15.80		t(20)=-	t(20)=0.
change to	1.1 (0.87	(4.05 -	(2.19 -	p<0.00	t(18)=-4.5	4.63	83
CPR	- 1.32)	4.84)	3.65)	1	p<0.001	p<0.001	p<0.001

Results of generalizability analysis

Appendix table 4

The table shows the variance components of the different facets and interactions; and relative contribution to the overall variance for Paediatric Basic Life Support and Foreign Body Airway Obstruction. In addition, an interpretation of the results from the different components is provided.

Source		Paediat	tric Basic Life	Support	Foreign B	ody airway	obstruction
of	Descriptio	Variance	Relative	Interpretati	Variance	Relative	Interpretati
varianc	n	component	contributio	on	component	contributi	on
e		s	n (%)		s	on (%)	
Particip	Systematic	57.874	48.3	Most of the	148.407	61.4	Most of the
ants	variance			variance is			variance is
	between			related to			related to
	participants			intended			intended
				differences			differences
				between			between
				participants			participants
Tests	Systematic	0	0	Cases were	2.124	0.9	Cases were
	variance			equally			almost
	among tests			difficult			equally
							difficult
Raters	Systematic	8.362	7	A minor	23.529	9.7	A minor
	variance			fraction of			fraction of
	among			the variance			the variance
	raters			was due to			was due to
				differences			differences
				between			between
				raters			raters
Interact	Trend for	24.154	20.2	The tests	20.098	8.30	The tests

betwee n tests and particip antsassess participants and differentlytheir assessment of participantstheir assessment of participantsInteract ion particip attersTrend for assess participants6.3125.3There was only minor bias between participants differently00There was no bias between raters and due to effective blindingInteract ion betwee n raters and on casesTrend for rater to betwee n raters differently00The raters found the cases equally challenging00The raters found the cases equally challengingInteract testsResidual ion variability23.11619.3Anticipated unexplained error47.65119.7Anticipated unexplained errorInteract raters included in and the test design23.11619.3Anticipated error47.65119.7Anticipated error	ion	tests to			varied in			varied in
n tests and and and participants and differently ants participants assessment of participants assessment of participants ants assessment of participants ants Interact ion participants and participants and participants Trend for assess and bias between bias between bias between assess and differently participants and ants participants and participants due to effective blinding blinding blinding Interact ion rater to betwee assess Interact start Trend for assess and participants assess and participants assess and participants and on cases tests 0 0 The raters and participants due to effective blinding blinding blinding Interact tests Trend for assess and participants assess and participants and on cases tests 0 0 The raters and participants assess as assess and participants and and the test and participants and the test asses as assess as asses and participants assess and participants assess asses are assess and participants assess and participants assess and participants assess and participants are assess are asse								
and differently particip ants Interact Trend for assess participants participants Interact Trend for assess participants and differently participants Interact Trend for assess participants and differently participants Interact Trend for assess participants and differently participants and differently participants Interact Trend for assess participants and assess participants and due to assess participants and between participants and assess participants and and the test participants and the test participants and the test participants and and the test design participants and assess participants and and the test design participants and of the participants and and the test design participants and of the participants and of the participants and and participants and participants and and participants and and participants and and participants and p								
particip ants Interact Trend for raters to betwee and differently participants Interact Trend for o assess and participants and differently particip ants Interact Trend for o only minor bias between raters and participants due to effective blinding Interact Trend for rater to betwee assess cases equally on cases tests Interact Residual variability betwee not cases on the test of the								
ants Interact Irend for raters to betwee assess participants and differently particip ants Interact Interact Irend for raters to between the same of the state of		uniterentity						
Interact Trend for raters to betwee assess participants and differently particip ants Interact Trend for raters to bias between raters and participants due to effective blinding Interact Trend for rater to betwee assess fiests Interact Residual ion variability betwee and facets notes; not raters included in and the test particip disassess Interact Trend for to for tests, not raters included in and the test particip disasses is assess only ass	1				participants			participants
ion raters to between assess bias between n raters and participants and differently participants due to effective blinding Interact Trend for rater to betwee assess cases equally n raters and differently and on cases tests Interact Residual 23.116 19.3 Anticipated ion variability betwee and facets not raters included in and the test participants due to effective blinding Interact residual cases on the participants due to effective blinding on lass between raters and participants due to effective blinding Interact Trend for to to the found the cases equally cases equally cases equally cases equally cases equally challenging and to the test design		Tuesd Con	C 212	г o	The	0	0	Th
betwee n raters and participants and participants and differentlybias between raters and participants due to effective blindingraters and due to effective blindingInteract ion rater to betwee n raters differently and on cases0 <br< th=""><th></th><th></th><th>6.312</th><th>5.3</th><th></th><th>U</th><th>U</th><th></th></br<>			6.312	5.3		U	U	
n raters and participants and participants differentlyraters and due to effective blindingparticipants due to effective blindingInteract ion raters and ion cases testsTrend for assess differently on cases00The raters found the cases equally challenging unexplained error00The raters found the cases equally challenging unexplained errorInteract tests, not raters included in and the test particip23.11619.3Anticipated error47.65119.7Anticipated unexplained error					-			
and differently due to effective blinding Interact Trend for rater to betwee assess differently on cases Interact Residual variability variability and the test particip design Interact included in and the test particip Interact included in and the test particip with and series and series included in and the test particip Interact due to effective blinding due to effective blinding Interact effective blinding Interact found the found the cases equally cases equally cases equally challenging challenging with a series of the case of the								
particip ants due to effective blinding Interact Trend for 0 0 The raters found the assess cases equally challenging on cases Interact Residual variability variability betwee and facets not raters included in and the test particip design effective blinding due to effective blinding blinding 0 0 The raters of 0 0 The raters found the cases equally cases equally challenging challenging challenging 47.651 19.7 Anticipated unexplained error error								
ants effective blinding Interact Trend for 0 0 The raters found the betwee assess cases equally challenging and on cases tests Interact Residual 23.116 19.3 Anticipated ion variability betwee and facets not raters included in and the test particip design	and	differently			-			
Interact Trend for rater to found the betwee assess cases equally on cases tests Interact Residual variability betwee and facets not raters included in and the test particip design Interact Trend for 0 0 The raters of 0 0 The raters found the found the cases equally cases equally challenging challenging challenging challenging and 47.651 19.7 Anticipated unexplained error error	particip				due to			effective
Interact rater to rater to betwee assess differently on cases tests Interact Residual variability betwee and facets raters included in and the test particip design Trend for rater to found the found the cases equally challenging challenging challenging and to make the test of the	ants				effective			blinding
ionrater tofound thefound thebetweeassesscases equallycases equallyn ratersdifferentlychallengingchallengingandon casescasestestslineractResidual23.11619.3Anticipated47.65119.7Anticipatedionvariabilityunexplainedunexplainedbetweeand facetserrorerrorn tests,noterrorerrorratersincluded inandthe testparticipdesigndesign					blinding			
betweeassesscases equallyn ratersdifferentlychallengingandon casestests23.11619.3Anticipatedionvariabilityunexplainedbetweeand facetserrornoterrorratersincluded inandthe testparticipdesign	Interact	Trend for	0	0	The raters	0	0	The raters
n raters differently and on cases tests Interact Residual variability unexplained error n tests, not raters included in and the test particip design challenging challengi	ion	rater to			found the			found the
and testson casesdesidual ton23.11619.3Anticipated unexplained error47.65119.7Anticipated unexplained errorbetwee n tests, raters and participon t the test designerrorerror	betwee	assess			cases equally			cases equally
testsInteractResidual23.11619.3Anticipated47.65119.7Anticipatedionvariabilityunexplainedunexplainedbetweeand facetserrorerrorn tests,noterrorratersincluded inerrorandthe testerrorparticipdesign	n raters	differently			challenging			challenging
Interact Residual variability variability betwee and facets not raters included in and the test particip design 23.116 19.3 Anticipated 47.651 19.7 Anticipated unexplained error error error	and	on cases						
ion variability unexplained unexplained betwee and facets error error ntests, not raters included in and the test particip design	tests							
betwee and facets error error n tests, not included in error error and the test particip design error error	Interact	Residual	23.116	19.3	Anticipated	47.651	19.7	Anticipated
n tests, not raters included in and the test particip design	ion	variability			unexplained			unexplained
raters included in and the test design	betwee	and facets			error			error
and the test design	n tests,	not						
particip design	raters	included in						
	and	the test						
lauta	particip	design						
ants	ants							