

<p style="text-align: center;">PROCES OBSERVATIE-INSTRUMENT / BETROUWBAARHEID VAN HET INSTRUMENT</p>

1. OPSTELLEN OBSERVATIE-INSTRUMENT

Stap 1

Voor het opstellen van het observatie-instrument hebben we **in eerste instantie** diverse bronnen geraadpleegd. Op basis van dat **bronnenonderzoek** werd een eerste versie van een observatielijst opgesteld (zie eerste versie observatie-instrument). In een eerste fase hebben we aan de hand van dit observatie-instrument diverse videofragmenten bekeken. Al doende merkten we dat een aantal variabelen uit de observatielijst moeilijk werkbaar waren. Dit bleek bijvoorbeeld het geval bij de variabelen ondernemingszin, taalvaardigheid, muzische vorming en audiovisuele middelen. De gedragsindicatoren van deze variabelen bleken onvoldoende helder omschreven. Bovendien was de variabele muzische vorming nog onvoldoende geïntegreerd uitgewerkt. Door het observeren kwamen we verder tot de vaststelling dat de activiteiten die de studenten uitgevoerd hebben te weinig vrijheid lieten aan de kleuters. De studenten hadden de neiging om de activiteiten te sterk te sturen. Dit alles maakte dat het observeren en scoren zeer tijdsintensief was.

Stap 2

Een volgende stap in het opstellen van een degelijk observatie-instrument was het in **overleg** treden met **experten**, nl. Ferre Laevers en Evelien Buyse. Op basis van deze gesprekken en de eerdere vaststellingen bij het observeren met de eerste versie van het observatie-instrument werd beslist om het observatiekader van Ferre Laevers als basis te nemen en vervolgens die variabelen uit te werken die we zinvol achtten voor ons onderzoek. Op die manier is de tweede versie van het observatie-instrument tot stand gekomen (zie definitieve versie observatie-instrument).

2. WIJZE VAN OBSERVEREN

Om na te gaan of het **observatie-instrument werkbaar** en bovendien ook **betrouwbaar** is, hebben we volgende werkwijze gehanteerd bij het bekijken van diverse videofragmenten.

Fase 1

In een **eerste fase** hebben **8 observatoren** (6 studenten en 2 onderzoekers) 8 onderzoekskinderen gescoord aan de hand het observatie-instrument. Bij het bekijken van de videofragmenten legden we onze focus op 1 onderzoekskind. In eerste instantie gaf elke observator die kleuter een individuele score op de diverse variabelen van het onderzoeksinstrument. Bij de outputvariabelen werd een score voor het totale fragment gegeven. Bij de procesvariabelen welbevinden en betrokkenheid werd een random score (om de 5 minuten) en een score bij kritische / opvallende momenten gegeven. Vervolgens werden de individuele scores aan elkaar meegedeeld en trachtten we tot een compromisscore te komen.

Na enige tijd merkten we dat de individuele scores steeds dichterbij elkaar kwamen te liggen. Het werd dan ook gemakkelijker om tot een compromisscore te komen.

Fase 2

Aangezien we al vrij goed op elkaar afgestemd waren, gingen in een **tweede fase** kleine groepjes van **2 à 3 observatoren** aan de slag. Bij het bekijken van nieuwe fragmenten hanteerden we dezelfde werkwijze als eerder in de grote groep.

Fase 3

In een **derde fase** hebben de 2 onderzoekers alle reeds gescoorde videofragmenten opnieuw bekeken om na te gaan of de compromisscore ook effectief de correcte score en bijgevolg de **expertscore** is.

Vervolgens werd de betrouwbaarheid van het observatie-instrument nagegaan aan de hand van de verkregen gegevens uit fase 1, 2 en 3 (zie punt 3 statistische verwerking).

Fase 4

In een **laatste fase**, nadat de betrouwbaarheid van het instrument werd nagegaan, hebben de 2 onderzoekers de overige onderzoekskinderen geobserveerd en gescoord op de verschillende variabelen (dezelfde werkwijze als hierboven: individuele score, dan tot compromisscore komen). Op die manier hebben we de gegevens van alle betrokken onderzoekskinderen in kaart gebracht.

3. STATISTISCHE VERWERKING

Al deze **scores** werden **geanalyseerd** met het statistisch verwerkingsprogramma SPSS. De betrouwbaarheid van de scores op de variabelen werd nagegaan alsook de betrouwbaarheid van individuele observators. Hieronder vind je eerst de betrouwbaarheid van de scores per variabele en de betrouwbaarheid van de individuele beoordelaars per variabele. Vervolgens vind je een overzicht van al deze gegevens.

3.1 Betrouwbaarheid per variabele (Intra class correlation)

WELBEVINDEN		ICC onderling = .22								
ID	C	1	2	3	4	5	6	7	8	range (#)
FASE 1										
K1_m1	3	3.0	3.0	3.0	4.0	3.0	3.0	2.0	3.0	2 - 4 = 2
K1_m2	3	4.0	3.0	3.0	4.0	3.0	4.0	3.0	3.0	3 - 4 = 1
K2_m1	2	3.0	2.0	1.0	2.0	2.0	3.0		2.0	1 - 3 = 2
K2_m2	2	3.0	2.0	2.0	2.0	3.0	2.0		2.0	2 - 3 = 1
K3_m1	4	4.0	4.0	5.0	4.0	3.0	3.0	5.0	4.0	3 - 5 = 2
K3_m2	4	4.0	4.0	4.0	4.0	3.0	2.0	5.0	4.0	2 - 5 = 3
K4_m1	4	4.0	3.0	4.0	4.0	3.0		4.0	4.0	3 - 4 = 1
K4_m2										
K5_m1	4	4.0	5.0	3.0	3.0	4.0		4.0	5.0	3 - 5 = 2
K5_m2										
FASE 2										
K6_m1	2							2.0	3.0	2 - 3 = 1
K6_m2										
K7_m1	3		2.0			3.0				2 - 3 = 1
K7_m2	3		3.0			3.0				3 - 3 = 0
K8_m1	4	4.0		3.0	3.0					3 - 4 = 1
K8_m2										
	2 - 4	.69	.81	.79	.69	.53	.00	.80	.88	1 - 5

Uit bovenstaande tabel kunnen we afleiden dat de scores op de procesvariabele welbevinden (ICC = .22) een lage betrouwbaarheid vertonen (ICC < .60).

Verder zien we dat de betrouwbaarheid van de beoordelaars 5 en 6 op het vlak van welbevinden niet zo hoog is. De overige beoordelaars zijn betrouwbaar (ICC > .60).

BETROKKENHEID		ICC onderling = .38								
ID	C	1	2	3	4	5	6	7	8	range (#)
FASE 1										
K1_m1	2	3.0	3.0	2.0	3.0	2.0	3.0	2.0	3.0	2 - 3 = 1
K1_m2	3	4.0	4.0	2.0	4.0	3.0	4.0	3.0	3.0	2 - 4 = 2
K2_m1	2	3.0	1.0	1.0	1.0	1.0	3.0		2.0	1 - 3 = 2
K2_m2	3	2.0	2.0	3.0	2.0	3.0	3.0		3.0	2 - 3 = 1
K3_m1	4	4.0	4.0	5.0	3.0	4.0	3.0	5.0	4.0	3 - 5 = 2
K3_m2	4	3.0	4.0	4.0	4.0	3.0	3.0	5.0	4.0	3 - 5 = 2
K4_m1	4	4.0	4.0	4.0	4.0	4.0		5.0	4.0	4 - 5 = 1
K4_m2										
K5_m1	3	3.0	4.0	3.0	3.0	4.0		4.0	4.0	3 - 4 = 1
K5_m2										
FASE 2										
K6_m1	2							2.0	2.0	2 - 2 = 0

K6_m2											
K7_m1	2		2.0			3.0					2 - 3 = 1
K7_m2	3		2.0			4.0					2 - 4 = 2
K8_m1	3	3.0		3.0	4.0						3 - 4 = 1
K8_m2											
	2 - 4	.44	.69	.84	.59	.69	.00	.81	.85		1 - 5

Uit bovenstaande tabel kunnen we afleiden dat de scores op de procesvariabele betrokkenheid (ICC = .38) een lage betrouwbaarheid vertonen (ICC < .60).

Verder zien we dat de betrouwbaarheid van de beoordelaars 1, 4 en 6 bij het meten van de procesvariabele betrokkenheid niet zo hoog is. De overige beoordelaars zijn betrouwbaar (ICC > .60).

ACTIEVE TAALVAARDIGHEID ICC = .91

ID	C	1	2	3	4	5	6	7	8	range (#)
FASE 1										
K1	1	2.0	2.0	2.0	2.0	1.0	1.0	1.0	1.0	1 - 2 = 1
K2	1	1.0	1.0	1.0	1.0	1.0	1.0		1.0	1 - 1 = 0
K3	5	4.0	4.0	4.0	4.0	4.0	3.0	5.0	5.0	3 - 5 = 2
K4	4	4.0	4.0	3.0	4.0	4.0		4.0	4.0	3 - 4 = 1
K5	3	4.0	4.0	3.0	3.0	3.0		3.0	2.0	2 - 4 = 2
FASE 2										
K6	1							1.0	1.0	1 - 1 = 0
K7	3		4.0			2.0				2 - 4 = 2
K8	4	4.0		4.0	4.0					4 - 4 = 0
	1 - 5	.89	.85	.88	.92	.93	.80	1.00	.97	1 - 5

Uit bovenstaande tabel kunnen we afleiden dat de scores op de outputvariabele actieve taalvaardigheid (ICC = .91) een hoge betrouwbaarheid vertonen (ICC > .60).

Alle beoordelaars zijn betrouwbaar bij het scoren van deze outputvariabele (ICC > .60).

SOCIALE COMPETENTIE ICC = .82

ID	C	1	2	3	4	5	6	7	8	range (#)
FASE 1										
K1	1	2.0	2.0	1.0	2.0	1.0	2.0	1.0	1.0	1 - 2 = 1
K2	1	1.0	1.0	1.0	1.0	2.0	1.0		1.0	1 - 2 = 1
K3	3	4.0	4.0	4.0	4.0	4.0	3.0	2.0	3.0	2 - 4 = 2
K4	4	4.0	3.0	5.0	5.0	3.0		5.0	2.0	2 - 5 = 3
K5	3	3.0	4.0	2.0	2.0	3.0		3.0	3.0	2 - 4 = 2
FASE 2										
K6	2							1.0	2.0	1 - 2 = 1
K7	2		2.0			2.0				2 - 2 = 0
K8	4	4.0		4.0	4.0					4 - 4 = 0
	1 - 4	.91	.77	.90	.84	.80	.86	.85	.71	1 - 5

Uit bovenstaande tabel kunnen we afleiden dat de scores op de outputvariabele sociale competentie (ICC = .82) een hoge betrouwbaarheid vertonen (ICC > .60).

Alle beoordelaars zijn betrouwbaar bij het scoren van deze outputvariabele (ICC > .60).

ZELFSTURING **ICC = .48**

ID	C	1	2	3	4	5	6	7	8	range (#)
FASE 1										
K1	2	3.0	4.0	3.0	2.0	2.0	2.0	2.0	3.0	2 - 4 = 2
K2	2	3.0	1.0	2.0	1.0	3.0	3.0		2.0	1 - 3 = 2
K3	4	3.0	5.0	5.0	4.0	3.0	3.0	5.0	3.0	3 - 5 = 2
K4	4	4.0	5.0	5.0	5.0	4.0		5.0	4.0	4 - 5 = 1
K5	4	3.0	3.0	3.0	3.0	2.0		4.0	5.0	2 - 5 = 3
FASE 2										
K6	2							2.0	3.0	2 - 3 = 1
K7	2		3.0			2.0				2 - 3 = 1
K8	4	3.0		3.0	5.0					3 - 5 = 2
2 - 4		.24	.57	.66	.82	.44	.50	.89	.71	1 - 5

Uit bovenstaande tabel kunnen we afleiden dat de scores op de outputvariabele zelfsturing (ICC = .48) een lage betrouwbaarheid vertonen (ICC < .60).

De betrouwbaarheid van de beoordelaars 1, 2, 5 en 6 is bij het meten van de outputvariabele zelfsturing eerder laag. De overige beoordelaars zijn betrouwbaar (ICC > .60).

CREATIVITEIT **ICC = .88**

ID	C	1	2	3	4	5	6	7	8	range (#)
FASE 1										
K1	1	2.0	2.0	2.0	2.0	2.0	1.0	1.0	1.0	1 - 2 = 1
K2	2	2.0	1.0	2.0	2.0	3.0	1.0		2.0	1 - 3 = 2
K3	5	4.0	4.0	4.0	4.0	4.0	3.0	5.0	3.0	3 - 5 = 2
K4	5	5.0	5.0	5.0	5.0	4.0		5.0	5.0	4 - 5 = 1
K5	4	3.0	2.0	2.0	2.0	2.0		4.0	4.0	2 - 4 = 2
FASE 2										
K6	1							1.0	2.0	1 - 2 = 1
K7	2		2.0			2.0				2 - 2 = 0
K8	4	3.0		4.0	4.0					3 - 4 = 1
1 - 5		.84	.78	.77	.77	.65	.74	1.00	.85	1 - 5

Uit bovenstaande tabel kunnen we afleiden dat de scores op de outputvariabele creativiteit (ICC = .88) een hoge betrouwbaarheid vertonen (ICC > .60).

Alle beoordelaars zijn betrouwbaar bij het scoren van deze outputvariabele (ICC > .60).

MUZISCHE **ICC = .84**

ID	C	1	2	3	4	5	6	7	8	range (#)
FASE 1										
K1	1	1.0	2.0	2.0	2.0	1.0	1.0	1.0	1.0	1 - 2 = 1
K2	1	1.0	1.0	1.0	1.0	2.0	1.0		1.0	1 - 2 = 1
K3	4	3.0	4.0	4.0	4.0	4.0	3.0	4.0	2.0	2 - 4 = 2
K4	3	3.0	4.0	3.0	4.0	4.0		2.0	1.0	1 - 4 = 3
K5	4	4.0	3.0	2.0	2.0	3.0		4.0	4.0	2 - 4 = 2
FASE 2										
K6	1							1.0	1.0	1 - 1 = 0
K7	2		3.0			1.0				1 - 3 = 2
K8	2	2.0		3.0	3.0					2 - 3 = 1
1 - 4		.95	.80	.64	.63	.82	.92	.96	.67	1 - 4

Uit bovenstaande tabel kunnen we afleiden dat de scores op de outputvariabele muzische competentie (ICC = .84) een hoge betrouwbaarheid vertonen (ICC > .60).
Alle beoordelaars zijn betrouwbaar bij het scoren van deze outputvariabele (ICC > .60).

3.2 Overzicht Intra class correlation scores met afgesproken 'correcte' score

Schaal	Beoordelaar								
	alle	1	2	3	4	5	6	7	8
WB	.22	.69	.81	.79	.69	.53	.00	.80	.88
BTH	.38	.44	.69	.84	.59	.69	.00	.81	.85
TAAL	.91	.89	.85	.88	.92	.93	.80	1.00	.97
SOCIAAL	.82	.91	.77	.90	.84	.80	.86	.85	.71
ZELFSTURING	.48	.24	.57	.66	.82	.44	.50	.89	.71
CREATIVITEIT	.88	.84	.78	.77	.77	.65	.74	1.00	.85
MUZISCHE	.84	.95	.80	.64	.63	.82	.92	.96	.67

Deze tabel (en ook de voorgaande tabellen) toont ons dat de betrouwbaarheid van de competenties actieve taalvaardigheid, sociale competentie, creativiteit en muzische competentie legitiem is (intra class correlatie > .60). Het bekomen van een consistente score bij het observeren van de procesvariabelen welbevinden (.22) en betrokkenheid (.38) en de competentie zelfsturing (.48) blijkt moeilijker te zijn. De betrouwbaarheid van de outputvariabelen dus hoger dan deze van de procesvariabelen (behalve bij zelfsturing).

Verder kunnen we afleiden dat de individuele observatiescores van beoordelaars 7 en 8 een hoge betrouwbaarheid vertonen (= intra class correlatie met de afgesproken 'correcte' score > .67). Deze vaststelling is gunstig aangezien deze 2 beoordelaars de 2 onderzoekers zijn die verder gegaan zijn met het project.

4 . TOEKOMST

Alle videofragmenten hebben een expertscore per variabele toegewezen gekregen. Het zou zinvol zijn om een handleiding op te stellen zodat kleuteronderwijzers effectief met het observatie-instrument aan de slag zouden kunnen gaan. De videofragmenten en de bijhorende handleiding (met expertscores) vormen dan een basis om de scores bij andere fragmenten / klassituaties aan af te toetsen (cfr. De Leuvense betrokkenheidsschaal – Centrum voor ervaringsgericht onderwijs).