

1. Zo werken kinderen als medeonderzoekers

Een vaste onderzoekstaal voor alle vijf modules

Het medeonderzoekerscontract

- Wij voorspellen vóór we de uitkomst kennen.
- Wij beschrijven wat we werkelijk zagen, niet wat we hoopten te zien.
- Wij zoeken minstens één andere mogelijke verklaring.
- Wij zeggen hardop welke conclusie te groot zou zijn.
- Wij mogen de proefopzet bekritisieren en een betere versie voorstellen.
- Een onverwachte of negatieve uitkomst is ook een resultaat.

Vier rollen per onderzoeksteam

Rol	Taak
Proefleider	Leest de stappen en bewaakt dat beide condities eerlijk worden uitgevoerd.
Tijd- en materiaalbewaker	Houdt tijd en materiaal gelijk en noteert afwijkingen.
Waarnemer	Noteert wat zichtbaar of hoorbaar gebeurt, zonder meteen te verklaren.
Tegenspreker	Vraagt: “Hoe weten we dat?” en zoekt een alternatieve verklaring.

Wissel de rollen bij een volgende module. Bij teams van drie combineert de tijdsbewaker zijn rol met die van waarnemer.

De vaste onderzoekscyclus

1. Vraag — wat proberen we te weten te komen?
2. Voorspelling — wat denken we dat zal gebeuren, en waarom?
3. Vergelijking — welke twee condities houden we zo gelijk mogelijk?
4. Waarneming — wat meten, tellen of beschrijven we?
5. Twijfel — welke andere verklaring past ook bij onze gegevens?
6. Grens — wat mogen we hieruit juist níét besluiten?
7. Nieuwe proef — welke kleine aanpassing maakt de volgende test sterker?

2. Gemeenschappelijke meet- en uitvoerregels

Klein genoeg voor een klaslokaal, precies genoeg om opnieuw te proberen

Standaard sessie van 25 minuten

Tijd	Fase	Wat kinderen doen
0–3 min	Vraag en rollen	Lezen de onderzoeksvraag, verdelen rollen en voorspellen.
3–15 min	Twee condities	Voeren de vergelijking uit; tijd en materiaal blijven gelijk.
15–20 min	Gegevens	Tellen, vergelijken en zoeken verschillen én overeenkomsten.
20–25 min	Tegenspraak	Formuleren alternatieve verklaring, niet-conclusie en vervolgproef.

Minimumgegevens per team

- Teamcode en groepsgrootte; geen namen in de centrale datatabel.
- Leeftijdsband: 11–12 of 13–14.
- Conditievolgorde: $A \rightarrow B$ of $B \rightarrow A$.
- Ten minste één resultaatmaat (tijd, juistheid, aantal fouten, draagkracht of herinnering).
- Vertrouwen vóór en na op een schaal van 1 tot 5.
- Eén observatie in gewone taal.
- Eén alternatieve verklaring en één verboden conclusie.
- Eén voorstel van de kinderen voor versie 0.2.

Eerlijk vergelijken

Gebruik waar mogelijk dezelfde tijd, dezelfde moeilijkheid, hetzelfde materiaal en dezelfde instructietoon. Wanneer dezelfde kinderen beide condities doen, wissel dan de volgorde tussen teams. Zo wordt een mogelijk oefen- of vermoeidheidseffect zichtbaar.

Geen wedstrijd tussen kinderen: Vergelijk condities en groepspatronen, niet “slimme” en “minder slimme” kinderen. Deel geen individuele scores buiten het team. Deelname aan spreken, schrijven of bouwen kan worden aangepast zonder dat een kind zich hoeft te verantwoorden.

Afgedrukte AI-uitvoer

Alle AI-teksten in dit pakket zijn vaste, gesimuleerde onderzoeksprikkels. Ze zijn niet bedoeld als representatieve kwaliteitsmeting van een leverancier of model. Dezelfde afdruk bij elke herhaling

maakt de proef beter vergelijkbaar en vermijdt accounts, privacyproblemen en veranderende modeluitvoer.

Vaste facilitatorzinnen

“Wat zag je precies gebeuren, zonder het al te verklaren?”

“Welke meting zou iemand uit een andere klas op dezelfde manier kunnen herhalen?”

“Welke andere verklaring past óók bij deze uitkomst?”

“Welke grote conclusie zou aantrekkelijk klinken, maar niet door deze proef worden gedragen?”

“Wat wil jij aan versie 0.2 veranderen?”

Stop of pas aan wanneer

Een kind zich beoordeeld voelt in plaats van medeonderzoeker; de twee condities niet meer vergelijkbaar zijn; het materiaal onveilig blijkt; of de uitkomst al door de begeleider wordt verklapt.

Module 1 — Het slimme foute antwoord

Cognitieve basisvaardigheden + digitaal oordeelsvermogen · 20–25 minuten

Onderzoeksvraag

Helpt eerst zelf schatten of rekenen om fouten te zien in een overtuigend geformuleerd AI-antwoord?

Wat vergelijken we?

Conditie $A \rightarrow B$ maakt eerst een eigen schatting of uitwerking en controleert daarna de AI-uitvoer. Conditie $B \rightarrow A$ beoordeelt eerst de AI-uitvoer en schat of rekt daarna pas zelf. Verdeel teams ongeveer gelijk en laat de twee groepen niet overleggen vóór beide bladen zijn ingevuld.

Hypothesen die kinderen mogen kiezen

- H1 — Eerst zelf nadenken maakt foutdetectie beter.
- H2 — De volgorde maakt weinig verschil; zorgvuldig lezen is belangrijker.
- H3 — Een nette uitleg kan het vertrouwen verhogen, ook als het antwoord fout is.

Stappen voor de facilitator

1. Verdeel teams vooraf over $A \rightarrow B$ en $B \rightarrow A$. $A \rightarrow B$ krijgt eerst M1-A; $B \rightarrow A$ krijgt eerst M1-B. Geef beide groepen drie minuten en houd de andere bladen uit zicht.
2. Wissel de bladen zonder groepsbespreking. Geef opnieuw drie minuten: $A \rightarrow B$ beoordeelt nu M1-B; $B \rightarrow A$ maakt nu de eigen schattingen op M1-A.
3. Laat alle teams maximaal twee antwoorden volledig narekenen en noteren of hun vertrouwen veranderde.
4. Teams krijgen daarna de antwoordsleutel, tellen correct herkende fouten en bespreken één gemiste of betwiste fout.
5. Laat de tegenspreker een alternatieve verklaring en een verboden conclusie formuleren.

Te meten

Maat	Hoe
Foutdetectie	Aantal foutieve AI-antwoorden correct als fout gemarkeerd (0–3).
Onterechte twijfel	Aantal correcte AI-antwoorden ten onrechte als fout gemarkeerd (0–1).
Kalibratie	Verschil tussen vertrouwen vóór en na het narekenen.
Bewijs	Kan het team precies aanwijzen waar de redenering of uitkomst botst?

Alternatieve verklaringen

- Rekenvaardigheid verschilt tussen kinderen.
- De tekstlengte of formulering beïnvloedt vertrouwen.
- Teams volgen een zelfverzekerd groepslid.
- De tweede beoordeling profiteert van oefening.

Niet concluderen: “AI kan niet rekenen”, “kinderen kunnen AI altijd controleren” of “hoofdrekenen is bewezen onmisbaar”. Deze proef gebruikt vier gekozen voorbeelden en een vaste gesimuleerde uitvoer.

M1-A — Rekentaak voor kinderonderzoekers

Eerst voorspellen; precies uitrekenen mag daarna

Teamcode: _____ Leeftijdsband: ☐ 11–12 ☐ 13–14 Volgorde: ☐ A→B ☐ B→A

Vraag	Mijn snelle schatting	Vertrouwen 1–5
1. Drie schriften van €2,40, twee pennen van €1,75 en één map van €4,90. Totaal?	_____	____
2. Wat ligt het dichtst bij 29×51 : 150, 1.500 of 15.000?	_____	____
3. Hoeveel is 15% van 80?	_____	____
4. Hoeveel is $1,90 + 0,35 + 0,08$?	_____	____

Voorspelling

Ik denk dat een nette stap-voor-stapuitleg mij ☐ méér ☐ evenveel ☐ minder vertrouwen geeft dan alleen een antwoord, omdat:

Onderzoekersafpraak

- ☐ We veranderen ons eerste antwoord niet stilletjes.
- ☐ We markeren bewijs in de tekst.
- ☐ “Ik weet het niet” is een bruikbare observatie.

M1-B — Gestandaardiseerde AI-uitvoer

Gesimuleerde afdruk · dezelfde tekst voor iedere groep

AI-uitvoer	Tekst
Antwoord 1	Drie schriften kosten €7,20. Twee pennen kosten €3,50. Met de map van €4,90 is het totaal €16,60.
Antwoord 2	29×51 ligt dicht bij $30 \times 50 = 1.500$. Van de drie opties is 1.500 dus het beste antwoord.
Antwoord 3	15% van 80 is 8. Je neemt eerst 10% van 80; dat is 8.
Antwoord 4	$1,90 + 0,35 = 2,25$. Daarna 0,08 erbij geeft 2,42.

Controleer als onderzoeksteam

Nr.	Ons oordeel	Bewijs of twijfel
1	☐ juist ☐ fout ☐ onzeker	_____
2	☐ juist ☐ fout ☐ onzeker	_____
3	☐ juist ☐ fout ☐ onzeker	_____
4	☐ juist ☐ fout ☐ onzeker	_____

Antwoordsleutel voor facilitator: 1 fout (juiste som €15,60); 2 juist; 3 fout (15% = 12); 4 fout (som = 2,33). Toon deze sleutel pas na registratie.

Gemeenschappelijk kinderonderzoeksblad

Eén exemplaar per team en per module

Module: _____ Teamcode: _____ Leeftijdsband: ☐ 11–12 ☐ 13–14 Groepsgrootte: _____

1. Onze onderzoeksvraag

2. Onze voorspelling

Wij denken dat conditie ☐ A/K/S ☐ B/I/AI ☐ geen duidelijk verschil zal geven, omdat:

3. Wat houden we gelijk?

- ☐ tijd
- ☐ materiaal
- ☐ moeilijkheid
- ☐ instructie
- ☐ anders: _____

4. Onze gegevens

Maat	Conditie 1	Conditie 2	Vershil
Resultaat 1	_____	_____	_____
Resultaat 2	_____	_____	_____
Vertrouwen 1–5	_____	_____	_____

Gemeenschappelijk kinderonderzoeksblad — conclusie

Beschrijven, twijfelen en verbeteren

5. Wat zagen we werkelijk?

6. Welke andere verklaring past ook?

7. Welke conclusie is te groot?

Op basis van deze proef mogen we niet besluiten dat:

8. Hoe maken wij versie 0.2 sterker?

Verander één ding aan de proef:

Wat zou je dan verwachten?

9. Ons onderzoekersbesluit

- ☐ Onze voorspelling kreeg steun.
- ☐ Onze voorspelling kreeg geen duidelijke steun.
- ☐ De gegevens zijn te onduidelijk.
- ☐ We ontdekten een interessantere nieuwe vraag.

Nieuwe vraag: _____

Facilitatorblad — centrale groepsdata

Anoniem verzamelen voor vergelijking tussen sessies

Team	Leeftijd	Volg.	Maat 1 C1	Maat 1 C2	Maat 2 C1	Maat 2 C2	Opmerking
_____	_____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____	_____	_____

Kwalitatieve oogst

Spontane vraag van kinderen:

Sterkste alternatieve verklaring:

Meest genoemde verboden conclusie:

Voorstel voor versie 0.2:

Uitvoercheck

- ☐ Alle teams kregen dezelfde vaste afdrukken.
- ☐ Tijd en materiaal waren gelijk of afwijkingen zijn genoteerd.
- ☐ Volgorde $A \rightarrow B$ en $B \rightarrow A$ was ongeveer verdeeld.
- ☐ Er zijn geen namen in de centrale data opgenomen.
- ☐ Kinderen kregen het laatste woord over de vervolproef.

Evaluatiekaart voor de vijf experimenten

Controle tegen het Charter v0.1

Vraag uit het charter	Hoe dit pakket antwoordt
Welke vaardigheid of kennis?	Elke module benoemt een concrete vaardigheid en koppelt die aan één of meer van de zes domeinen.
Waarom mogelijk waardevol?	De onderzoeksvraag verbindt de taak aan foutdetectie, eigenaarschap, ruimtelijk begrip, praktisch inzicht of overdracht.
Welke rol speelt AI/technologie?	Vaste afgedrukte AI-uitvoer is hulpmiddel en onderzoeksprikkel; geen leverancier wordt beoordeeld.
Wat vergelijken we?	Iedere module definieert twee condities en waar nodig twee volgordes.
Wat ervaart het kind zelf?	Kinderen rekenen, schrijven, reconstrueren, bouwen, uitleggen en meten zelf.
Welke alternatieve verklaring?	Iedere module geeft kandidaten; kinderen moeten zelf minstens één verklaring kiezen of toevoegen.
Wat mogen we niet besluiten?	Iedere module heeft een expliciete grens; het kinderblad vraagt een eigen formulering.
Kan het opnieuw?	Tijd, materiaal, afdrukken, condities en resultaatmaten zijn gestandaardiseerd.

Beslissing na een eerste kinderproef

Gebruik per module één van vier voorlopige uitspraken:

- **BEHOUDEN** — begrijpelijk, veilig en uitvoerbaar; alleen kleine taal- of timingcorrecties.
- **AANPASSEN** — interessante vraag, maar condities of metingen zijn nog niet eerlijk genoeg.
- **HERONTWERPEN** — kinderen onderzoeken iets anders dan bedoeld of kunnen onvoldoende zelf waarnemen.
- **STOPPEN** — weinig maatschappelijke waarde, onveilig, niet reproduceerbaar of te normatief gestuurd.

Onderzoek vóór campagne: Publiceer na een eerste proef vooral observaties, vragen en beperkingen. Gebruik geen scoreverschil als bewijs voor een brede maatschappelijke claim.

Snelle facilitatorcheck vóór elke sessie

Tien punten voor veiligheid, begrijpelijkheid en reproduceerbaarheid

- ☐ De onderzoeksvraag kan door een 11-jarige in eigen woorden worden herhaald.
- ☐ De twee condities en de volgorde zijn vooraf vastgelegd.
- ☐ Alle teams krijgen dezelfde tijd en dezelfde vaste afdrukken.
- ☐ Materiaal is veilig, identiek genoeg en vooraf geteld.
- ☐ Deelnemers weten dat individuele scores niet worden gepubliceerd.
- ☐ Er is een alternatief voor kinderen die niet willen schrijven, spreken, rekenen of bouwen.
- ☐ De facilitator verklapt geen gewenste uitkomst.
- ☐ De tegenspreker krijgt zichtbaar tijd om twijfel te formuleren.
- ☐ De niet-conclusie wordt vóór de afsluiting ingevuld.
- ☐ Minstens één voorstel van de kinderen wordt geregistreerd voor versie 0.2.

Korte afsluitvraag aan kinderen

“Als jij deze proef opnieuw mocht doen met een andere klas: wat zou jij hetzelfde houden, wat veranderen en waarom?”

Formele programmalijn: KopKen — Artisanale kennis voor kinderen. Werknaam publieksactiviteiten: Kan jij het nog zelf?