

Phygital

Physical + Digital

**Vijf activiteiten, drie groepen,
24 kinderen gemeten
voor en na.**

**"De kinderen die normaal afhaken,
doen bij Phygital wel mee."**

Dit verslag laat zien wat er in juli 2026 daadwerkelijk is gedaan, wat de metingen opleveren, waar de opzet tegen grenzen aanliep en wat er nodig is om de volgende ronde wel harde uitspraken te kunnen doen.

PILOTVERSLAG

Gamen als haak naar bewegen

Haalbaarheidspilot: uitvoering,
metingen en lessen

Gemeente Amsterdam Zuidoost
Afdeling Jeugd & Sport

PERIODE

8 t/m 22 juli 2026

LOCATIE

De Rozemarn

ACTIVITEITEN

5 dagdelen

GEMETEN KINDEREN

24

In het kort

Wat deze pilot heeft opgeleverd, en wat hij bewust niet bewijst

Phygital verzorgde in juli 2026 vijf activiteiten tijdens de zomerschool van PPOZO op basisschool De Rozemarn. Kinderen spelen een game, en wat ze op het scherm doen vertalen we direct naar een fysieke opdracht in de zaal. De game is de haak, het bewegen is het doel. Bij 24 kinderen is het reactievermogen voor en na gemeten met een eigen ontwikkelde meetapp: 188 losse metingen, 3.225 geregistreerde prikkels.

Lees dit verslag als een haalbaarheidspilot, niet als een effectmeting

Deze pilot moest laten zien of de aanpak uitvoerbaar is, of de kinderen meedoen en of er betrouwbaar te meten valt binnen een zomerschool. Op die drie vragen geeft dit verslag antwoord.

De vraag of het programma het reactievermogen van kinderen verbetert, kan met vijf activiteiten, drie groepen en tien kinderen in een traject niet worden beantwoord. Wij doen dat dan ook niet. Waar in dit verslag cijfers staan die de goede kant op wijzen, staat er steeds bij hoe hard ze zijn en wat er nodig is om er wel een uitspraak op te baseren.

5ACTIVITEITEN
UITGEVOERD**3**GROEPEN
BEREIKT**24**KINDEREN
GEMETEN**188**METINGEN
AFGENOMEN**3.225**PRIKKELS
GEREGISTREERD

DE VIJF CONCLUSIES

1. De aanpak werkt in de praktijk. Bij alle drie de groepen werd de vertaalslag van scherm naar zaal begrepen en meteen opgepakt. Fortnite bleek de sterkste haak: die groep kinderen wilde zonder uitzondering meedoen.
2. Alleen de groep die een echt traject volgde, laat vooruitgang zien op snelheid en nauwkeurigheid samen. De groep van De Rozemarn deed drie lessen over dertien dagen en verbeterde 10,1% op de prestatiescore, waarbij het foutpercentage daalde van 14,8% naar 9,6%. Die uitkomst blijft overeind als het kind met de grootste vooruitgang buiten beschouwing blijft.
3. De groep die binnen een uur voor- en nagemeten werd, werd wel sneller maar maakte meer fouten (10,2% naar 11,9%). Dat is harder tikken, geen leren. Meten binnen dezelfde les kost bovendien ruim veertig van de zestig lesminuten en levert daarvoor niets bruikbaars op.
4. De opzet raakte versnipperd. De bedoeling was een vaste groep vier lessen lang te volgen. Dat lukte niet, waardoor de meetkracht over drie groepen verdeeld raakte in plaats van geconcentreerd.
5. De aantallen zijn te klein voor bewijs, en dat claimen we ook niet. Met tien kinderen was alleen een effect van meer dan het dubbele van wat wij meten aantoonbaar geweest. Wat deze pilot wel oplevert: een uitvoerbare aanpak, kinderen die meedoen, en een werkend meetsysteem waarmee een serieus vervolg de vraag wel kan beantwoorden.

Waar dit verslag eerlijk over is

Met tien kinderen in de trajectgroep en drie lessen is er geen statistisch bewijs te leveren dat gamen het reactievermogen van kinderen verbetert. Het verschil dat we meten is klein en kan ook door toeval, dagvorm of gewenning aan de meespellen zijn ontstaan. Er was geen controlegroep.

Wat wel opvalt: de trajectgroep is de enige groep waar snelheid en nauwkeurigheid samen vooruitgingen, en die uitkomst hangt niet aan een enkel kind. Daar staat tegenover dat een van de vergelijkingsgroepen in een enkel dagdeel een even sterke verbetering liet zien. Beide staan in dit verslag.

Wat Phygital doet

Van scherm naar zaal, in dezelfde beweging

Phygital koppelt gamen aan bewegen. Niet als beloning achteraf ("eerst sporten, dan mag je gamen"), maar als een doorlopende lijn binnen dezelfde les. Twee kinderen spelen op het scherm, de rest van de groep voert in de zaal uit wat er in de game gebeurt. Dat noemen we de vertaalslag.

Zo ziet een vertaalslag eruit

Bij Fortnite gaat het om bouwen, dekking zoeken en snel reageren op wat er om je heen gebeurt. In de zaal wordt dat: op het signaal van de begeleider zo snel mogelijk achter een pion duiken, in vier tellen een "muur" van matten opbouwen, of bij een kleurcommando naar de juiste hoek sprinten. Wie op het scherm te laat reageert, ligt eruit. In de zaal geldt hetzelfde. De kinderen zien meteen dat het om dezelfde vaardigheid gaat: kijken, kiezen, uitvoeren, en dat allemaal onder tijdsdruk.

Bij FIFA vertaalden we de acties op het veld naar de zaal: een sprint na balverlies, kappen en draaien bij een richtingswissel, een strafschooproutine met een keeper die pas op het laatste moment een kant kiest. De spelvormen komen uit het draaiboek dat bij deze pilot hoort.

Waarom dit werkt bij deze doelgroep

- De drempel is laag. Een kind dat niet op voetbal zit en zich ongemakkelijk voelt bij gymles, kent Fortnite wel en heeft daar wel status.
- De opdracht is meteen duidelijk. Je hoeft geen sportterm uit te leggen: "doe wat er op het scherm gebeurt" is genoeg.
- Er is altijd een rol. Wie niet aan de beurt is op het scherm, doet in de zaal mee, en dat is niet de minder leuke rol.
- Het meten is zelf een spel. De kinderen willen hun eigen tijd verbeteren, dus de meting kost geen motivatie maar levert het juist op.

DE OPDRACHT VAN DE GEMEENTE

Gemeente Amsterdam Zuidoost, afdeling Jeugd & Sport, gaf opdracht voor vijf zomeractiviteiten tijdens de zomerschool van PPOZO op basisschool De Rozemarn, voor kinderen uit de bovenbouw, tegen een vast bedrag van EUR 1.000 inclusief alles (offerte 2026-PHY-001). Onderdeel van de opdracht was een 0-meting en een eindmeting van het reactievermogen, zodat er meer overblijft dan een indruk.

De vijf activiteiten

Wat er per dag is gedaan, en met wie

Dag	Datum	Groep	Wat we hebben gedaan	Meting
Dag 1	wo 8 juli	De Rozemarn (11)	Kennismaking, uitleg van de aanpak, 0-meting reactievermogen en reactiespellen in de zaal.	0-meting
Dag 2	vr 10 juli	De Rozemarn	FIFA op het scherm, met in de zaal de bewegingen die daar bij horen: sprint na balverlies, kappen en draaien, strafschoproutine.	geen
Dag 3	wo 15 juli	Wereldwijs (13)	Fortnite met de vertaalslag naar de zaal. Voor- en nameting binnen dezelfde les van een uur.	voor + na
Dag 4	di 21 juli	De Rozemarn (10)	Fortnite met de vertaalslag, eindmeting en uitreiking van de diploma's aan de kinderen die het hele traject hebben gevolgd.	eindmeting
Dag 5	wo 22 juli	Derde groep	Fortnite met de vertaalslag. Bewust geen voor- en nameting, omdat op dag 3 bleek dat daar binnen een uur geen ruimte voor is.	geen

Waarom de groep van De Rozemarn de kern van dit verslag is

De Rozemarn is de enige groep die we op dag 1, dag 2 en dag 4 hebben gehad. Deze kinderen kregen drie lessen over een periode van dertien dagen, met de 0-meting op de eerste dag en de eindmeting op de laatste. Daar zit tijd tussen waarin iets kan veranderen. Op deze groep ligt daarom de nadruk, en deze kinderen kregen op dag 4 een diploma omdat zij een volledig traject hebben afgerond.

De metingen bij Wereldwijs op dag 3 zijn om een andere reden waardevol: ze laten precies zien waarom een voor- en nameting binnen een uur niet werkt. Dat is een resultaat op zichzelf.

GELEVERD TEGENOVER AFGESPROKEN

Afspraak uit offerte 2026-PHY-001	Geleverd	Status
Vijf activiteiten van 45 tot 60 minuten in juli 2026	Vijf activiteiten uitgevoerd op 8, 10, 15, 21 en 22 juli	Volledig
Locatie De Rozemarn, zomerschool PPOZO	Alle vijf de activiteiten op De Rozemarn	Volledig
0-meting en nameting van het reactievermogen	188 metingen bij 24 kinderen, waarvan 10 kinderen met een volledig traject	Volledig, opzet afgeweken
Deelnamecijfers per les	Per meetdag vastgelegd; op de twee dagen zonder meting alleen als aanwezigheid geteld	Deels
Kinderen van 9 tot en met 12 jaar	De zomerschool deelde groepen in met kinderen van 8 en 9 jaar	Afgeweken

Afspraak uit offerte 2026-PHY-001	Geleverd	Status
Vaste teams gedurende de hele pilot	Niet mogelijk doordat de groepssamenstelling per dag wisselde	Niet gehaald
Certificaten voor de deelnemers	Uitgereikt aan de trajectgroep op dag 4	Volledig
Sportmateriaal, schermen, flyers, eindrapportage binnen EUR 1.000	Alles geleverd, geen meerkosten in rekening gebracht	Volledig

Twee afspraken zijn niet gehaald, en beide om dezelfde reden: de groepsindeling lag bij de zomerschool en niet bij ons. De kinderen die we kregen waren jonger dan de afgesproken 9 tot 12 jaar, en vaste teams over vijf lessen waren onmogelijk toen bleek dat de groepen per dag wisselden. Dat heeft de uitvoering niet in de weg gestaan, maar wel de opbouw die het programma bedoeld had.

Bereik

Over de vijf dagdelen zijn drie groepen bereikt. Van 24 kinderen zijn metingen vastgelegd (13 jongens, 11 meisjes, 8 en 9 jaar). De derde groep op dag 5 deed wel mee, maar is bewust niet gemeten en de groepsgrootte is niet vastgelegd. Wij rekenen daarom met 24 kinderen: het werkelijke bereik ligt hoger, maar wij nemen geen getal op dat wij niet kunnen onderbouwen.

Hoe we gemeten hebben

Vier spellen van dertig seconden, op een tablet, voor en na

Voor deze pilot hebben wij zelf een meetsysteem gebouwd: de Phygital Reactiemeter. Elk kind speelt vier korte spellen van maximaal dertig seconden. De app registreert per prikkel de reactietijd in milliseconden en telt fouten apart, zodat een kind dat lukraak tikt niet per ongeluk een goede score haalt. De volledige meting duurt ongeveer drie minuten per kind.

Waarom we het zelf hebben gebouwd

Dit systeem hoorde niet bij de opdracht. De afspraak ging uit van een bestaande reactietimer-app, of anders een fluit en een stopwatch. Wij wilden zelf weten of je het op deze manier kon testen, en zijn het daarom gaan bouwen. Het is buiten de opdracht om ontwikkeld, in eigen tijd, en er is niets voor in rekening gebracht.

Dat pakte beter uit dan de afgesproken werkwijze. Een stopwatchmeting bij kinderen van acht is niet nauwkeurig genoeg om iets uit af te leiden, en bestaande testapparatuur kost al snel meer dan de hele opdracht waard was. Nu is er meer gemeten dan was afgesproken: niet alleen reactietijd maar ook fouten, niet alleen een eindscore maar elke losse poging, en niet op papier maar direct digitaal en controleerbaar. Dat de cijfers in dit verslag tot op de individuele reactie na te rekenen zijn, komt daar rechtstreeks uit voort.

Dat wij zowel de meetlat als de analyse leveren, is een gerechtvaardigd bezwaar

Het instrument is gebouwd door de partij die ook de lessen geeft en die een vervolgopdracht wil. Dat is een reden om kritisch te zijn, en wij vinden dat terecht.

Daarom is de opzet zo gemaakt dat u ons niet hoeft te geloven: de ruwe meetgegevens zijn als bestand op te vragen, elke losse reactie is bewaard, de rekenregels staan in bijlage B en dit verslag wordt automatisch uit die gegevens gegenereerd. Iedereen met het databestand komt op dezelfde getallen uit. Bij een vervolg kan de gemeente de analyse desgewenst door een derde laten doen; wij leveren de gegevens dan aan zonder voorbewerking.

Spel	Wat het meet	Hoe het werkt	Wat een fout is
Tik op groen	Eenvoudige reactietijd	Het scherm wordt groen op een willekeurig moment. Tik zo snel mogelijk. Vijf keer.	Tikken voordat het groen is
Groen of rood	Impulsrem (go/no-go)	Tik wel bij groen, tik niet bij rood. Rood is ook aan de vorm te herkennen, zodat kleurenblindheid geen rol speelt.	Toch op rood tikken
Links of rechts	Keuzereactietijd	Er verschijnt een pijl. Tik op de kant waar de pijl heen wijst.	De verkeerde kant kiezen
Vang de stip	Volgehouden tempo	Dertig seconden lang verschijnen stippen op willekeurige plekken. Vang er zo veel mogelijk.	Naast de stip tikken

Wat we hebben gedaan om de meting eerlijk te houden

- Elk kind mag eerst oefenen. Die oefenronde wordt niet meegeteld, zodat de eerste meting niet vooral onwennigheid meet.

- De reactietijd wordt pas gestempeld nadat het scherm daadwerkelijk is bijgewerkt, niet op het moment dat de app het bevel geeft. Anders telt de tekentijd van het apparaat mee in de score van het kind.
- Fouten vervuilen het gemiddelde niet: ze worden apart geteld. Snelheid en nauwkeurigheid zijn zo los van elkaar te lezen, en dat blijkt in dit verslag precies het verschil te maken.
- De 0-meting en de eindmeting gebruiken exact dezelfde spellen, dezelfde duur en hetzelfde apparaat.
- De app werkt volledig offline. In een gymzaal zonder wifi gaat er zo niets verloren.

Wat deze getallen niet zijn

De metingen zijn gedaan op een aanraakscherm, niet met laboratoriumapparatuur. Een tablet voegt vertraging toe tussen de aanraking en de registratie. De absolute getallen in dit verslag zijn daarom niet een op een te vergelijken met waarden uit wetenschappelijk onderzoek.

Overigens is er ook in dat onderzoek geen normtabel waaraan we ons zouden kunnen meten. Gepubliceerde gemiddelden voor eenvoudige reactietijd bij kinderen van deze leeftijd lopen uiteen van ongeveer 350 tot ver boven 700 milliseconden, en dat verschil komt vrijwel volledig door de opzet van de test, niet door de kinderen. De waarden die wij meten (347 tot 480 milliseconden) vallen midden in dat bereik, wat aangeeft dat de meting plausibele waarden oplevert.

Voor dit doel is dat afdoende: we vergelijken elk kind met zichzelf, op hetzelfde apparaat en met hetzelfde spel. Het verschil tussen de 0-meting en de eindmeting is de meetwaarde, niet het getal op zich.

Hoe betrouwbaar is zo een meting eigenlijk

Een reactiemeting bij kinderen is geen weegschaal. Onderzoek naar vergelijkbare taken bij basisschoolkinderen, ook in een gewoon klaslokaal met de bijbehorende drukte, laat zien dat de gemiddelde reactietijd redelijk stabiel is tussen twee meetmomenten, maar dat foutmaten en uitschieters dat duidelijk minder zijn. Praktisch betekent dat twee dingen voor dit verslag.

- Een verschil bij een enkel kind zegt vrijwel niets. Een kind dat slecht heeft geslapen of net ruzie had, meet die dag trager, en dat verschil kan groter zijn dan het effect van een heel programma.
- Een verschil op groepsniveau, in dezelfde richting, over meerdere spellen, zegt wel iets. Daarom staan in dit verslag steeds de groepscijfers voorop en zijn de individuele uitkomsten alleen in de bijlage opgenomen.

Dit is ook de reden dat wij bij een vervolg een tussenmeting en een nameting na zes weken voorstellen: hoe meer meetmomenten, hoe beter een echte verandering te onderscheiden is van een toevallige dagvorm.

Persoonsgegevens

In dit verslag staan geen namen van kinderen. Voor dit verslag zijn zij omgenummerd naar R1 tot en met R11 en W1 tot en met W13.

Onderwerp	Stand van zaken
Vastgelegde gegevens	Per kind alleen een voornaam, leeftijd, geslacht en de meetuitslagen. Geen achternaam, adres, geboortedatum, foto of contactgegevens.
Opslag en toegang	Beveiligde server in Europa, toegang met een gedeelde toegangscode, versleutelde verbinding, dagelijkse back-up met bewaartermijn van dertig dagen.
Toestemming	De deelname liep via de zomerschool, die de ouders informeerde. Er is geen aparte schriftelijke toestemming per kind vastgelegd voor de meting.

Onderwerp	Stand van zaken
Verwerkersovereenkomst	Niet gesloten met school of gemeente.
Bewaartermijn	Nog niet formeel vastgelegd. Voorstel: gegevens vernietigen zodra de eindrapportage is aanvaard, tenzij de gemeente ze voor vervolgonderzoek wil bewaren.
Beeldmateriaal	Foto's zijn alleen gebruikt na mondelinge toestemming ter plaatse. Voor publicatie is schriftelijke toestemming van de ouders nodig.

Wij benoemen de open punten hier zelf, omdat ze bij een vervolg als eerste geregeld moeten zijn: schriftelijke toestemming per kind, een verwerkersovereenkomst met de opdrachtgever, een vastgelegde bewaartermijn en, als er structureel gegevens van minderjarigen worden verzameld, een toets of een gegevensbeschermingseffectbeoordeling nodig is. Wij stellen voor dat de gemeente daarin de kaders geeft en wij ze uitvoeren.

De Rozemarn: de trajectgroep

11 kinderen, drie lessen, dertien dagen tussen de 0-meting en de eindmeting

De trajectgroep van De Rozemarn bestond uit 11 kinderen van 8 en 9 jaar (7 jongens, 4 meisjes). Van 10 kinderen is zowel een volledige 0-meting als een volledige eindmeting beschikbaar; een kind was op dag 4 afwezig en telt daarom niet mee. De 0-meting was op 8 juli, de eindmeting op 21 juli: dertien dagen ertussen, met daartussen drie lessen.

10 van 11 KINDEREN TWEE KEER GEMETEN	10,1% BETERE PRESTATIESCORE	7 van 10 KINDEREN VOORUIT	14,8% > 9,6% FOUTPERCENTAGE
--	---------------------------------------	-------------------------------------	--

WAAROM SNELHEID ALLEEN NIET GENOEG IS

Reactiesnelheid is een verleidelijk cijfer, want het daalt vanzelf als een kind roekelozer gaat tikken. Wie sneller reageert maar vaker de verkeerde kant kiest, is niet beter geworden. Daarom staat in dit verslag de prestatiescore voorop: de reactietijd gedeeld door het aandeel goede reacties. Een kind dat sneller wordt en even nauwkeurig blijft, verbetert op die score. Een kind dat sneller wordt door meer fouten te maken, niet. Het is de gebruikelijke maat wanneer snelheid en zorgvuldigheid tegen elkaar in kunnen bewegen, en dat is precies wat hier gebeurt.

Zo hebben we de ruwe meetgegevens behandeld

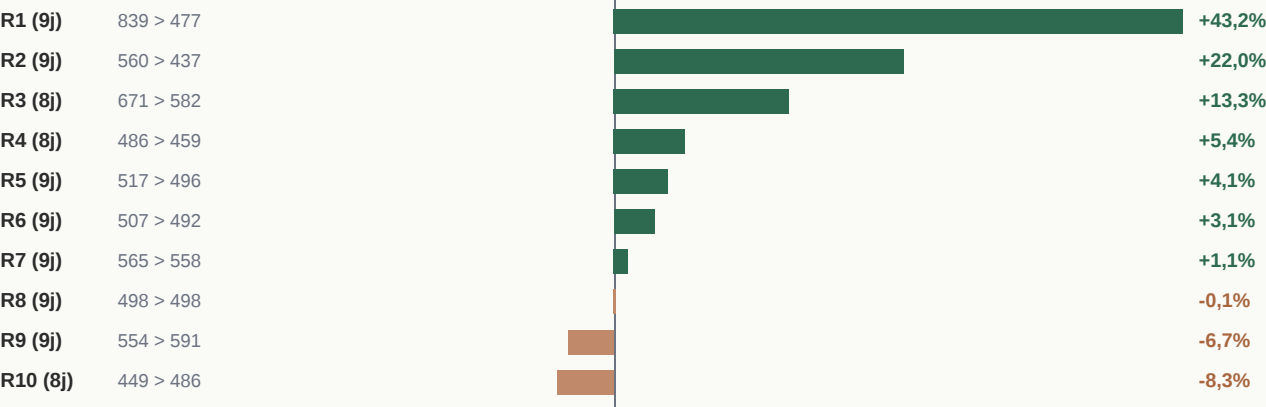
Voordat er iets berekend is, zijn op alle groepen dezelfde regels toegepast. Losse reacties onder 150 milliseconden tellen niet mee, want dat is anticiperen en geen reageren. Reacties boven 1.500 milliseconden tellen ook niet mee, want dat is afleiding. Een spel telt alleen mee als er daarna nog minstens vier geldige pogingen over zijn, en een kind telt alleen mee als er minstens tien minuten tussen de twee metingen zit.

In totaal vielen 19 van de 5.313 losse reacties af, 0,4% van alle metingen. Deze regels zijn vooraf vastgelegd en op elke groep gelijk toegepast, zodat er achteraf niets naar een gewenste uitkomst is toegerekend.

DE PRESTATIESCORE PER KIND

Verandering in prestatiescore per kind

snelheid gecorrigeerd voor fouten | 7 van de 10 kinderen ging vooruit



rechts van de lijn = vooruitgegaan, links = achteruitgegaan

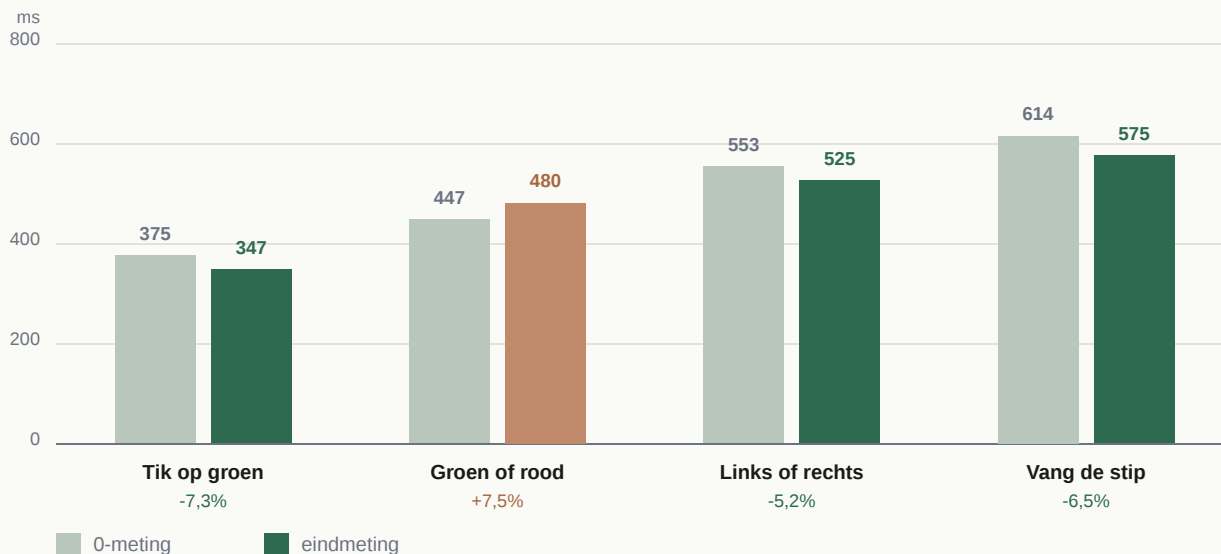
De groep ging van 565 naar 508, een verbetering van 10,1%. 7 van de 10 kinderen ging vooruit. De twee kinderen met de grootste winst verbeterden 43,2% en 22,0%; de twee kinderen die achteruitgingen verloren 8,3% en 6,7%. Die laatste twee noemen we er expliciet bij, omdat een verslag dat alleen de uitschieters naar boven vermeldt geen verslag is maar een folder.

Belangrijker dan het gemiddelde is of de uitkomst overeind blijft zonder het beste kind. Dat is hier zo: laat je het kind met de grootste vooruitgang weg, dan blijft er een verbetering van 4,3% over met een vrijwel gelijke effectmaat (0,43 tegenover 0,48). De uitkomst hangt dus niet aan een enkel kind.

SNELHEID EN NAUWKEURIGHEID PER SPEL

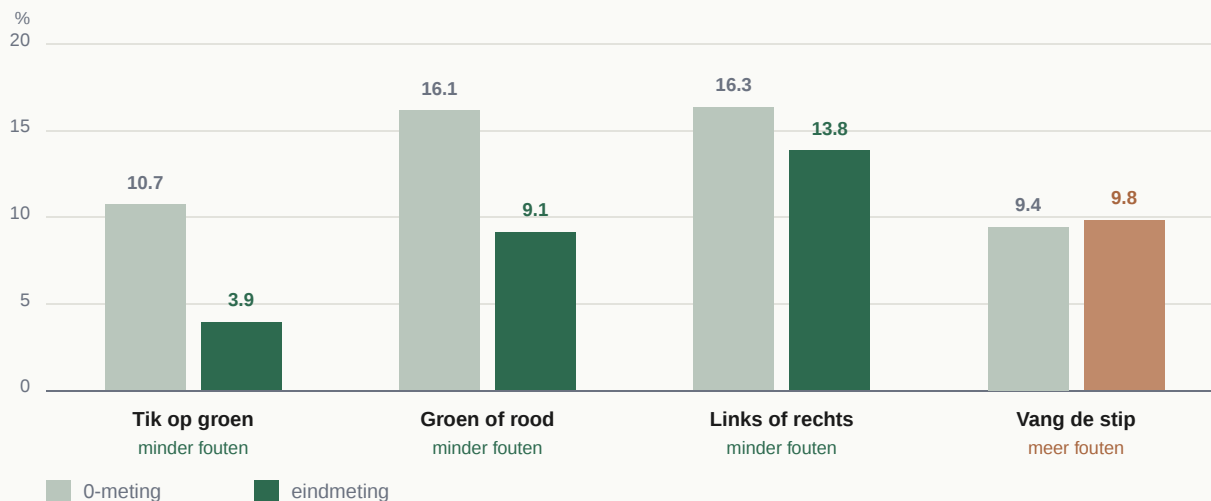
Gemiddelde reactietijd per spel, 0-meting tegenover eindmeting

lager is beter | 10 kinderen



Foutpercentage per spel, 0-meting tegenover eindmeting

lager is beter | aandeel foute reacties van alle reacties



Op 2 van de vier spellen ging de groep er zowel in snelheid als in nauwkeurigheid op vooruit: Tik op groen en Links of rechts. Bij Tik op groen, het spel dat de zuiverste reactietijd meet, daalde het foutpercentage van 10,7% naar 3,9%. Bij de andere twee spellen ging het niet in beide richtingen goed, en dat hoort in dit verslag te staan.

Groen of rood: trager, maar met minder fouten

Bij Groen of rood ging de gemiddelde tijd omhoog, van 447 naar 480 milliseconden. Dat is het enige spel waarbij je soms juist niet moet tikken. Tegelijk daalde het foutpercentage van 16,1% naar 9,1%. Twee lezingen zijn mogelijk: de kinderen zijn beheerster geworden en kijken beter voordat ze reageren, of ze zijn simpelweg voorzichtiger gaan spelen. Met deze gegevens is niet uit te maken welke van de twee het is. Op de prestatiescore, die beide kanten meeweegt, komt dit spel per saldo licht positief uit.

Vang de stip: meer stippen, maar ook meer missers

Bij Vang de stip ving de groep in dezelfde dertig seconden 498 stippen tegen 467 bij de 0-meting. Het aantal missers steeg echter mee, van 48 naar 54. Dit spel belooft doorzetten, en dat is precies waarom het geen zuivere reactiemaat is: sneller tikken levert bijna vanzelf meer treffers en meer missers op. Het staat in dit verslag als beeld van het tempo dat de kinderen volhouden, niet als bewijs.

Wat de toets zegt, en wat de toets niet kan zeggen

Op de prestatiescore wonnen de kinderen gemiddeld 57 punten. Getoetst geeft dat $t(9) = 1,52$ met een effectmaat van 0,48. Dat is niet significant, en het betrouwbaarheidsinterval loopt van -28 tot 142: ook geen enkele verbetering past nog binnen de meetruimte.

Belangrijker dan het woord significant is wat deze pilot überhaupt kon aantonen. Met 10 kinderen was alleen een effect vanaf ongeveer 0,99 betrouwbaar aantoonbaar geweest, oftewel ruim tweemaal zo groot als wat we hebben gemeten. Een niet-significante uitkomst was bij deze groepsgrootte dus het meest waarschijnlijke resultaat, ook als het programma precies zou doen wat wij hopen.

Dat is geen excuus achteraf: het is de reden waarom in het vervolgvoorstel achterin staat hoeveel kinderen en hoeveel lessen er nodig zijn om de vraag wel te beantwoorden.

Wereldwijs: de les die de meetopzet blootlegde

13 kinderen, voor- en nameting binnen dezelfde les van een uur

Op dag 3 werkten we met groep Wereldwijs: 13 kinderen van 8 en 9 jaar. Omdat deze groep maar één keer beschikbaar was, hebben we de voormeting en de nameting binnen dezelfde les gedaan. Alle 13 kinderen hebben beide metingen volledig afgerond, en 12 daarvan voldoen aan de meetregels. De tijd tussen de twee metingen was per kind mediaan 31 minuten, met een maximum van 42 minuten.

De Rozemarn drie lessen, dertien dagen tussen de metingen PRESTATIESCORE 565 > 508 VERANDERING 10,1% beter FOUTPERCENTAGE 14,8% > 9,6% Sneller en nauwkeuriger tegelijk, en dat blijft zo als je het beste kind weglaat.	Wereldwijs een les, mediaan 31 minuten tussen de metingen PRESTATIESCORE 526 > 514 VERANDERING 2,3% beter FOUTPERCENTAGE 10,2% > 11,9% Sneller, maar met meer fouten. Onder de streep blijft er vrijwel niets over.
--	---

Op de kale reactiesnelheid ziet Wereldwijs er niet slechter uit dan De Rozemarn: 9 van de 12 kinderen werd sneller. Wie alleen daarnaar kijkt, zou concluderen dat één les van een uur net zo goed werkt als drie lessen over twee weken. Dat is precies de conclusie waar de nauwkeurigheid voor behoedt.

Bij Wereldwijs steeg het foutpercentage op de reactiespellen namelijk van 10,2% naar 11,9%. Bij Vang de stip liep het aantal missers op van 21 naar 64. De kinderen kenden het spel inmiddels, wisten dat het om snelheid ging en gingen harder tikken. Op de prestatiescore, die snelheid en fouten samen weegt, blijft er dan 2,3% over tegenover 10,1% bij De Rozemarn. En laat je bij Wereldwijs het beste kind weg, dan slaat de uitkomst om naar een lichte achteruitgang.

Een kind dat we uit de vergelijking hebben gehaald

Bij Wereldwijs is een kind buiten de analyse gelaten. De voormeting van dit kind werd pas afgenomen nadat de rest van de klas beide metingen al achter de rug had, en de nameting volgde ruim binnen een minuut daarna. Dat zijn geen twee meetmomenten maar één, en het kind is daarom volgens de vooraf vastgestelde regel uitgesloten.

Wij noemen dit niet omdat het de uitkomst gunstiger maakt, maar omdat het precies laat zien wat er misgaat als je in een uur zowel les wilt geven als twee keer wilt meten.

WAT TWEE GROEPEN BUITEN DEZE PILOT LATEN ZIEN

De Reactiemeter is in dezelfde periode gebruikt bij twee groepen buiten deze opdracht, ook met beide metingen binnen één dagdeel. Die groepen zijn geen controlegroep, want zij kregen een ander programma. Ze zijn wel leerzaam, en niet alleen in ons voordeel.

Groep	Opzet	Prestatiescore	Verandering	Fout%
De Rozemarn	3 lessen, 13 dagen ertussen	565 > 508	10,1%	14,8% > 9,6%
Wereldwijs	1 les, 31 min ertussen	526 > 514	2,3%	10,2% > 11,9%
Vergelijkingsgroep 1	1 dagdeel, ander programma	545 > 553	-1,5%	10,5% > 12,3%
Vergelijkingsgroep 2	1 dagdeel, ander programma	518 > 485	6,4%	3,4% > 2,1%

De Rozemarn boekt de grootste verbetering op de prestatiescore en is samen met vergelijkingsgroep 2 de enige groep waar het foutpercentage daalt. Bij Wereldwijs en vergelijkingsgroep 1 gaat de nauwkeurigheid juist achteruit.

Het cijfer dat tegen ons pleit, en waarom het er toch staat

Vergelijkingsgroep 2 verbeterde in één dagdeel 6,4% op de prestatiescore, met een effectmaat van 0,99: de sterkste en meest consistente uitkomst in de hele meetreeks, sterker dan die van onze eigen trajectgroep over dertien dagen.

Dat is een ongemakkelijk gegeven en het hoort in dit verslag. Het betekent dat een enkele sessie een verbetering kan opleveren die net zo groot of groter is dan wat wij na drie lessen meten. Daarmee kan de vooruitgang bij De Rozemarn niet zonder meer aan het programma worden toegeschreven: gewenning aan de meetspellen kan een deel of zelfs het geheel van het verschil verklaren.

Wat er wel staat, is dat De Rozemarn de enige groep is waarbij snelheid en nauwkeurigheid samen de goede kant opgingen en dat dit overeind blijft zonder het beste kind. Dat is een aanwijzing die een vervolg rechtvaardigt, en niets meer dan dat. Alleen een opzet met een controlegroep kan deze vraag beslechten, en die staat in het voorstel achterin.

Waar we tegenaan liepen

Drie praktische problemen, en wat we ervan hebben geleerd

1. De opzet raakte versnipperd over drie groepen

De bedoeling was helder: één vaste groep, minimaal vier lessen, zodat we dezelfde kinderen konden volgen en steeds een nieuwe spelvorm konden opbouwen. Dat was met de leerkrachten van De Rozemarn ook zo afgesproken. Bij de uitvoering bleek de zomerschool daar anders in te staan: het werd niet eerlijk gevonden tegenover de andere groepen om één groep vaker mee te laten doen. Die afweging is goed te volgen. Voor het meten was het wel ongelukkig, want daardoor kwamen we uit op één groep van 11 kinderen die we drie keer zagen, en twee groepen die we eenmalig zagen.

De les is niet dat de zomerschool iets fout deed, maar dat de afspraak over de groepsindeling vooraf op het juiste niveau vastgelegd had moeten worden: niet met de leerkrachten, maar met de organisatie die over het programma gaat. Dat is een afspraak die in de opdrachtverlening thuishoort.

2. Een uur is te kort voor een voor- en nameting

Op dag 3 hebben we voor- en nagemeten binnen dezelfde les. De meetgegevens laten zien hoe duur dat was: het voormetblok liep van 08:09 tot 08:30 en het nameetblok van 08:43 tot 09:04. Bij elkaar ruim veertig minuten actief meten binnen een venster van vijftig minuten, met twee schermen en 13 kinderen. Wat overblijft is te weinig om iets te doen dat het meten waard is.

Even veelzeggend is dat de metingen door elkaar heen gingen lopen: bij een kind werd de voormeting pas afgenomen toen de rest van de klas de nameting al had gedaan, met nog geen minuut tussen zijn eigen twee metingen. Dat kind is uit de analyse gelaten.

Dit hebben we tijdens de pilot zelf gecorrigeerd: op dag 5 hebben we bewust helemaal niet gemeten en de volle tijd aan de activiteit besteed. Dat leverde een betere les op dan dag 3.

3. Twee schermen is de bottleneck

Met twee schermen spelen er twee kinderen tegelijk en wacht de rest. In de zaal is dat opgelost met de vertaalslag, want daar doet iedereen mee. Bij het meten is het niet opgelost: meten is per kind. Meer meetapparaten, of meten in een aparte ruimte terwijl de rest doorgaat met de les, lost dit direct op. De Reactiemeter draait op elke telefoon of tablet en werkt offline, dus dat is een kwestie van organisatie, niet van techniek.

4. We hebben te weinig vastgelegd naast de meting

In ons eigen draaiboek staan zeven soorten gegevens die per les vastgelegd zouden worden: presentie, actieve deelname op een vast moment, een duimmeting, observaties van een paar focuskinderen, letterlijke uitspraken van kinderen, foto's en de reactiemeting. In de praktijk is vooral die laatste consequent vastgelegd. Daardoor kunnen we het verhaal over betrokkenheid en gedrag wel vertellen, maar niet onderbouwen. Voor een programma dat juist op meedoen mikt, is dat een gemis dat wij onszelf aanrekenen.

Dit is direct op te lossen: het gaat om een formulier van een halve minuut per kind per les, en de begeleider die de les geeft kan dat niet tegelijk doen. Bij een vervolg hoort er iemand bij te zijn die alleen observeert en registreert.

Wat we hierdoor niet hebben kunnen aantonen

Er is geen groep die vier of meer lessen heeft gevolgd en er is geen controlegroep die wel gemeten is maar niet heeft meegedaan. Daarmee blijft onbeantwoord of het verschil bij De Rozemarn door het programma komt of door gewenning aan de meetspellen. De vergelijkingsgroepen laten zien dat die tweede verklaring serieus genomen moet worden.

Er is bovendien geen enkel beweeggegeven verzameld: geen actieve minuten, geen intensiteit, geen deelnamegraad per les. Dat is een terechte kritiek op een programma dat uit een beweegdoelstelling wordt betaald. Reactievermogen is een indirecte maat, gekozen omdat hij objectief en snel af te nemen is, niet omdat hij het beleidsdoel dekt. Bij een vervolg horen daar beweegminuten en deelname naast te staan.

Wat wel werkte

Wat we in drie groepen consequent hebben gezien

Fortnite is de sterkste haak

Bij alle drie de groepen waar we Fortnite hebben gebruikt, was de betrokkenheid direct hoog. Kinderen die bij de kennismaking afwachting waren, deden bij Fortnite meteen mee. Dat is geen meetresultaat maar een consistente observatie over drie verschillende groepen, en het is voor de praktijk het belangrijkste gegeven: zonder betrokkenheid komt er van bewegen niets terecht.

De vertaalslag wordt begrepen, ook door de jongste kinderen

De stap van "dit gebeurt op het scherm" naar "dit doen wij nu in de zaal" hoefde bij geen enkele groep meer dan een keer uitgelegd te worden. De kinderen vulden hem zelfs zelf aan met eigen voorstellen. Dat is belangrijker dan het klinkt: het betekent dat de aanpak niet afhankelijk is van een specifieke begeleider en dus overdraagbaar is.

De game keert de rangorde in de zaal om

Dit is volgens ons de kern van waarom het werkt, en het is geen meetresultaat maar een observatie. Een kind dat bij de gymles nooit als eerste wordt gekozen, kan in Fortnite of FIFA juist degene zijn die het uitlegt. Door de game als vertrekpunt te nemen, verandert wie er in de zaal vooropgaat. Datzelfde kind neemt die positie mee naar de fysieke opdracht, en doet daar dus ook mee.

Game-taal is gedeelde taal

In deze klassen zit altijd een deel van de kinderen dat het Nederlands nog aan het opbouwen is. Woorden als respawn, skill move en level kennen zij allemaal, ongeacht wat er thuis wordt gesproken. En de opdrachten in de zaal zijn voordoen en nadoen, niet uitleggen en begrijpen. Dat is een bewuste keuze, geen toeval: het verlaagt de drempel om mee te doen voor precies de kinderen die anders afwachten.

Bekendheid verklaart waarschijnlijk meer dan training

Het verschil tussen de trajectgroep en de eenmalige groep is niet alleen een cijfer. De trajectgroep kende op dag 4 de begeleiders, de zaal, de app en de opdracht al. Wereldwijd kreeg alles in een uur en ging vooral harder tikken om te winnen. Wij denken dat vertrouwdheid hier een sterkere verklaring is dan een trainingseffect, en dat is voor een vervolg minstens zo bruikbaar: het pleit voor doorlopende groepen en vaste begeleiders, niet voor losse activiteiten.

Metten werkt als spel, niet als test

De kinderen wilden hun eigen tijd verbeteren en vroegen uit zichzelf om nog een poging. Dat de eindmeting bij De Rozemarn door tien van de elf kinderen volledig is afgerond, terwijl het om vier spellen met volle aandacht gaat, komt daar rechtstreeks uit voort.

Het diploma sloot het traject af

De kinderen van De Rozemarn kregen op dag 4 een diploma, omdat zij als enigen het volledige traject hadden gevolgd. Dat is bewust gedaan: het maakt zichtbaar dat er een begin en een eind aan zit en dat er iets is afgerond, wat bij losse activiteiten ontbreekt. Voor kinderen die weinig gelegenheid hebben om iets af te maken, is dat geen bijzaak.



De trajectgroep van De Rozemarn na de uitreiking op dag 4, 21 juli 2026. Dit zijn de kinderen die we op dag 1, dag 2 en dag 4 hebben begeleid en die zowel de 0-meting als de eindmeting hebben afgerond.

Een punt dat we niet overslaan: Fortnite en de leeftijd

Fortnite heeft een PEGI 12-classificatie en de kinderen in deze pilot waren 8 en 9 jaar. Dat verdient een eerlijke toelichting, geen stilzwijgen.

In de praktijk speelt vrijwel elk kind in deze groep het spel al thuis; wij introduceren het niet. Wij gebruiken het in een begeleide setting, op een gedeeld scherm met de begeleider erbij, in de bouwmodus en de spelvormen zonder gevechtselementen, zonder online contact met onbekenden, zonder stemchat en zonder aankopen. De game duurt bovendien maar een deel van de les: het grootste deel van de tijd bewegen de kinderen in de zaal.

Voor een vervolg leggen we dit vooraf schriftelijk voor aan school en ouders, met de mogelijkheid om een alternatieve titel te kiezen. Voor scholen die liever geen shooter-titel in huis halen, is FIFA of Rocket League een volwaardig alternatief, al is de haak bij Fortnite aantoonbaar sterker.

Waarom dit past in Zuidoost

De aansluiting bij bestaand beleid en bestaande cijfers

Deze pilot staat niet op zichzelf. De opgave waar hij op inspeelt, is in Zuidoost scherper dan elders in de stad, en het programma sluit aan bij kaders die er al liggen.

DE OPGAVE IN CIJFERS

Cijfer	Wat het zegt	Bron
33%	van de tienjarigen in Zuidoost had overgewicht of obesitas, tegenover 22 procent in Amsterdam als geheel	O&S Amsterdam 2024, op basis van GGD-onderzoek 2019
75%	van de scholieren uit Zuidoost sport wekelijks in de vrije tijd, tegenover 85 procent stadsbreed	GGD Corona Gezondheidsmonitor Jeugd 2021
62%	van de inwoners van Zuidoost sport regelmatig: het laagste aandeel van alle stadsdelen	Sportmonitor Amsterdam 2022
15%	van de Amsterdamse jeugd haalt de beweegnorm van een uur per dag	GGD Gezondheidsmonitor Jeugd 2023

Het gaat hier om de kinderen die in die cijfers aan de verkeerde kant staan: kinderen die niet op een sportclub zitten en bij een gewone gymles afhaken. Precies die groep is met een klassiek sportaanbod het moeilijkst te bereiken, en precies bij die groep werkt een haak die zij zelf al kennen.

AANSLUITING BIJ BESTAANDE KADERS

- Amsterdamse Aanpak Gezond Gewicht: Zuidoost is daarin een expliciet benoemd aandachtsgebied, samen met Noord en Nieuw-West.
- Visie sport- en beweegbeleid 2023-2026: de opgave om sportdeelname juist in de kwetsbare wijken te verhogen.
- Masterplan Zuidoost 2021-2040: langjarige inzet op jeugd, met sport en voorzieningen op scholen als concrete interventie.
- De Rijke Schooldag: het programma is ontworpen als kant-en-klaar blok dat binnen dat rooster past, en is in deze pilot ook binnen een bestaand schoolprogramma uitgevoerd.
- Sportakkoord II en het Gezond en Actief Leven Akkoord: het kader waaruit gemeenten dit type activiteiten tot en met 2026 bekostigen.

WAT DIT TOEVOEGT AAN WAT ER AL IS

In Zuidoost is al veel sportaanbod. De vraag die er dan toe doet, is niet of er nog een activiteit bij moet, maar wie er met dit aanbod bereikt wordt die anders niet bereikt zou zijn.

- Het aanbod richt zich op de kinderen die niet op een club zitten. Wie al voetbalt, heeft dit niet nodig.
- Het draait binnen bestaande structuren: zomerschool, naschools aanbod, De Rijke Schooldag. Er is geen nieuwe locatie, vereniging of vervoer voor nodig.
- De begeleiding komt uit de wijk en spreekt de taal van de kinderen, letterlijk en in de onderwerpen die zij kennen.
- Er komt een meting mee. Dat is bij dit type activiteit ongebruikelijk, en het maakt verantwoording over meerdere jaren mogelijk in plaats van alleen deelnamecijfers.

Een punt van timing dat wij zelf aankaarten

De middelen uit het Gezond en Actief Leven Akkoord en Sportakkoord II lopen tot en met 2026. Hoe de bekostiging daarna wordt geregeld, was bij het schrijven van dit verslag nog niet uitgekristalliseerd. Voor een vervolg is het daarom verstandig nu al te kijken welke route na 2026 openstaat, bijvoorbeeld via het beleidskader sociale basis of de gebiedsgerichte opgaven van het stadsdeel.

Wij noemen dit omdat een goed programma dat op een financieringsgat stukloopt, alsnog niets oplevert.

Conclusie

Wat deze pilot heeft aangetoond, en wat de volgende stap zou moeten zijn

WAT DEZE PILOT HEEFT AANGETOOND

- De aanpak is uitvoerbaar binnen een reguliere zomerschool, met vijf activiteiten binnen de afgesproken tijd en het afgesproken budget.
- De kinderen doen mee. Bij drie verschillende groepen was de betrokkenheid direct hoog en werd de vertaalslag van scherm naar zaal zonder moeite begrepen.
- Er ligt een werkend meetsysteem dat wij zelf hebben gebouwd en dat blijft bestaan. 188 metingen bij 24 kinderen zijn zonder uitval vastgelegd, ook zonder internet in de zaal. Het hoorde niet bij de opdracht en is er niet voor betaald: wij hebben het gebouwd omdat we zelf wilden weten of het zo kon. Het is nu direct inzetbaar op andere locaties, en dat is een blijvende opbrengst die los staat van wat de lessen hebben opgeleverd.
- De enige groep die een echt traject volgde, is ook de enige groep die sneller en nauwkeuriger werd. Dat is een aanwijzing, geen bewijs, maar het is wel de aanwijzing die je zou verwachten als de aanpak werkt.

WAT DEZE PILOT NIET HEEFT AANGETOOND

- Niet dat het programma het reactievermogen van kinderen verbetert. Daarvoor zijn de aantallen te klein, was er geen controlegroep en waren er te weinig lessen.
- Niet dat er blijvend effect is. Er is na de eindmeting niet opnieuw gemeten.
- Niet dat de winst zich vertaalt naar meer bewegen buiten de les. Dat is niet gemeten en zou in een vervolg wel gevraagd moeten worden.

VOORSTEL VOOR EEN VERVOLG

De pilot heeft precies opgeleverd wat een pilot moet opleveren: een werkende aanpak, een werkend meetsysteem en een scherp beeld van wat er in de opzet niet klopte. Op basis daarvan stellen wij het volgende vervolg voor.

Onderdeel	In deze pilot	Voorstel vervolg
Looptijd	5 activiteiten in 3 weken zomerschool	12 aaneengesloten weken binnen het schooljaar, niet in een vakantie
Aantal lessen	3 lessen voor de trajectgroep	24 sessies, tweemaal per week op een vaste dag en tijd
Groepen	3 groepen, waarvan 1 met een traject	Per school 1 vaste klas plus 1 vergelijkbare klas die alleen gemeten wordt
Locaties	1 school	2 basisscholen in Zuidoost, zodat overdraagbaarheid zichtbaar wordt
Meetmomenten	0-meting en eindmeting	Oefensessie, 0-meting een week vooraf, eindmeting 48 uur na de laatste les, nameting na 4 tot 6 weken
Meettijd per les	Tot 40 minuten op meetdagen	Maximaal 10 minuten, met 6 tot 8 apparaten of een meetmoment buiten de les

Onderdeel	In deze pilot	Voorstel vervolg
Wat we meten	Alleen reactievermogen	Beweegminuten als hoofdmaat, daarnaast aanwezigheid per kind per les, uitval en reactievermogen
Vastleggen in de zaal	Niet systematisch	Vaste observatielijst per les door iemand die niet lesgeeft

De voorwaarde die wij vooraf willen vastleggen

Dat de groepen twaalf weken lang dezelfde blijven, wordt voor de start schriftelijk bevestigd door de school en door de organisatie die over het naschoolse programma gaat. Dat is precies wat in deze pilot misging, en het is met een afspraak van dertig minuten te voorkomen.

Het bezwaar dat het oneerlijk is als een groep vaker meedoet, lossen we op door de vergelijkingsklas het programma na afloop alsnog te geven. Dan krijgt iedereen het, alleen niet tegelijk.

Met deze opzet ligt het aantal kinderen ruim boven wat nodig is om een verschil van de omvang die wij nu zien aan te tonen, is er voor het eerst een vergelijkingsgroep, en telt de dosis mee: vierentwintig sessies is het niveau waarop bij dit type programma effect verwacht mag worden. Drie lessen is dat niet.

RISICO'S EN HOE WE ZE AFDEKKEN

Risico	Hoe we het afdekken
De groepsindeling wijzigt alsnog	Schriftelijke bevestiging vooraf als opschortende voorwaarde; zonder vaste groep starten we niet
Het programma hangt op een of twee personen	Draaiboek is uitgeschreven per les; in het vervolg leiden we twee begeleiders uit de wijk op die de lessen zelfstandig kunnen geven
Techniek valt uit in de zaal	De meetapp werkt volledig offline en synchroniseert later; er is altijd een reserveapparaat
Kinderen haken na verloop van tijd af	Aanwezigheid per kind per les registreren en bij twee keer afwezig contact opnemen; wisselende spelvormen uit het draaiboek
Discussie over schermtijd en gameleeftijd	Vooraf schriftelijk voorleggen aan school en ouders, met een alternatieve titel als keuze
Financiering valt weg na 2026	Nu al de route na 2026 bepalen, samen met het stadsdeel

Wij denken dat het reëel is om de volgende stap niet als onderzoek te zien maar als doorontwikkeling met een fatsoenlijke meting eronder: het programma draaien zoals het bedoeld is, en tegelijk de vraag beantwoorden die deze pilot open moest laten. De aanpak, het draaiboek en het meetsysteem liggen er. Wat ontbrak was tijd met dezelfde kinderen.

Bijlage A: alle meetwaarden

Per kind, per meting. Alle waarden komen rechtstreeks uit het meetsysteem.

De Rozemarn, de trajectgroep

Reactietijd is het gemiddelde van de drie reactiespellen. De prestatiescore is die reactietijd gedeeld door het aandeel goede reacties: hoe lager, hoe beter. Fouten zijn geteld over de drie reactiespellen, stippen zijn het aantal gevangen stippen in dertig seconden, waarbij meer beter is.

Kind	Lft	Reactietijd (ms)	Prestatiescore	Verschil	Stippen	Fouten
R1	9	518 > 463	839 > 477	+43,2%	47 > 47	7 > 1
R2	9	444 > 396	560 > 437	+22,0%	47 > 55	5 > 2
R3	8	422 > 485	671 > 582	+13,3%	46 > 46	7 > 3
R4	8	421 > 459	486 > 459	+5,4%	51 > 56	2 > 0
R5	9	455 > 434	517 > 496	+4,1%	45 > 49	2 > 2
R6	9	507 > 467	507 > 492	+3,1%	47 > 48	0 > 1
R7	9	533 > 509	565 > 558	+1,1%	42 > 45	1 > 2
R8	9	422 > 450	498 > 498	-0,1%	42 > 51	3 > 2
R9	9	413 > 407	554 > 591	-6,7%	48 > 51	7 > 6
R10	8	449 > 439	449 > 486	-8,3%	52 > 50	0 > 2
Groep		458 > 451	565 > 508	+10,1%	467 > 498	34 > 21

Niet meegeteld: N. (geen eindmeting afgerond).

Groep Wereldwijs

Voor- en nameting binnen dezelfde les, mediaan 31 minuten ertussen.

Kind	Lft	Score voor	Score na	Verschil	Tijd ertussen	Fouten
W1	9	752	509	+32,2%	25 min	11 > 1
W2	8	503	430	+14,5%	29 min	1 > 0
W3	8	538	470	+12,6%	42 min	3 > 2
W4	8	518	458	+11,7%	42 min	2 > 3
W5	9	695	623	+10,4%	24 min	4 > 5
W6	8	523	480	+8,3%	34 min	2 > 1
W7	9	483	462	+4,4%	23 min	2 > 3
W8	9	490	525	-7,2%	14 min	3 > 2

Kind	Lft	Score voor	Score na	Verschil	Tijd ertussen	Fouten
W9	9	476	555	-16,5%	29 min	0 > 3
W10	9	456	538	-18,0%	35 min	0 > 6
W11	9	439	521	-18,6%	32 min	0 > 4
W12	9	442	599	-35,6%	39 min	0 > 2
Groep		526	514	+2,3%	31 min	28 > 32

Niet meegeteld: D. (minder dan 10 minuten tussen beide metingen).

Groepsgemiddelden per spel

Spel	Groep	0-meting	Eind	Verschil	Fout% voor	Fout% na
Tik op groen	De Rozemarn	375 ms	347 ms	-7,3%	10,7%	3,9%
	Wereldwijs	378 ms	362 ms	-4,2%	11,8%	6,3%
Groen of rood	De Rozemarn	447 ms	480 ms	+7,5%	16,1%	9,1%
	Wereldwijs	459 ms	452 ms	-1,4%	9,8%	12,7%
Links of rechts	De Rozemarn	553 ms	525 ms	-5,2%	16,3%	13,8%
	Wereldwijs	548 ms	532 ms	-2,8%	9,5%	14,6%
Vang de stip	De Rozemarn	614 ms	575 ms	-6,5%	9,4%	9,8%
	Wereldwijs	581 ms	572 ms	-1,6%	3,4%	9,7%

Bijlage B: verantwoording

Hoe de cijfers in dit verslag tot stand zijn gekomen

Herkomst van de gegevens

Alle getallen in dit verslag komen rechtstreeks uit de database van de Phygital Reactiemeter, het meetsysteem dat tijdens de pilot is gebruikt. De cijfers zijn niet met de hand overgenomen: dit verslag wordt gegenereerd uit een export van de meetgegevens, zodat er geen overtypfouten in kunnen sluipen. De ruwe meetgegevens zijn op verzoek beschikbaar als CSV-bestand.

Rekenwijze

- Meetregels, vooraf vastgesteld en op elke groep gelijk toegepast: losse reacties onder 150 ms (anticipatie) en boven 1.500 ms (afleiding) tellen niet mee; een spel telt alleen mee met minstens 4 geldige pogingen; een kind telt alleen mee met minstens 10 minuten tussen beide metingen. Hierdoor viel 0,4% van alle losse reacties af (19 van 5.313).
- Reactietijd: het gemiddelde van de geldige reacties per spel, in milliseconden, afgerond op hele getallen.
- Foutpercentage: foute reacties gedeeld door alle reacties (goede plus foute) binnen dezelfde spellen.
- Prestatiescore: de reactietijd gedeeld door het aandeel goede reacties, gemiddeld over de drie reactiespellen. Vang de stip telt hierin niet mee omdat dat een telspel is met een andere schaal. Deze maat is gekozen omdat snelheid en nauwkeurigheid in deze gegevens tegen elkaar in bewegen; alleen naar snelheid kijken zou een kind belonen dat harder en slordiger tikt.
- Meetellende kinderen: bij De Rozemarn 10 van de 11, bij Wereldwijs 12 van de 13. De uitgesloten kinderen en de reden staan in bijlage A.
- Toetsing: gepaarde t-toets op de prestatiescore, met de effectmaat Cohen dz, een 95-procents betrouwbaarheidsinterval en een berekening van het kleinste effect dat bij deze groepsgrootte aantoonbaar was.
- Gevoeligheidscontrole: elke uitkomst is opnieuw berekend zonder het kind met de grootste vooruitgang. Waar dat het beeld verandert, staat dat in de tekst.

Wat er bewust niet is gedaan

- Er zijn geen metingen weggelaten om het resultaat gunstiger te maken. Ook de spellen waarop de groep achteruitgang staan in dit verslag.
- Er is niet geselecteerd op kinderen die vooruitgingen. Alle kinderen met een volledige meting staan in bijlage A.
- Er is geen effect toegeschreven aan het programma dat op basis van deze opzet niet toe te schrijven is.

Over het meetsysteem

De Reactiemeter is door Phygital zelf ontwikkeld voor deze pilot en blijft beschikbaar. Het is een webapp die als app op een telefoon of tablet te installeren is, volledig offline werkt en synchroniseert zodra er weer verbinding is. Groepen zijn te scheiden en per groep is een PDF- of CSV-export te maken. Het systeem staat op een beveiligde server met toegangscode en dagelijkse back-up. Het is buiten de opdracht om gebouwd en niet in rekening gebracht: voor een vervolg is er geen ontwikkeling nodig en geen licentie, en het kan zonder extra kosten op meer locaties tegelijk draaien.