

## **Kritische evaluatie van het IMPACTLAB basisinstrument: Uitgebreide versie met statistische analyses**

### **Behorend bij het IMPACTLAB eindrapport:**

Land-Zandstra, A., Strick, M., Peeters, W., de Jong, P. (2023). Het meten van impact van wetenschapscommunicatie: Eindrapport IMPACTLAB. IMPACTLAB.

IMPACTLAB is een samenwerking tussen de Universiteit Leiden en de Universiteit Utrecht, geïnitieerd en gefinancierd door de Nationale Wetenschaps Agenda (NWA).

Datum: 28 maart 2023

In dit document bespreken we de complete analyses waarop onze conclusies over de kwaliteit van ons basisinstrument, zoals bescheven in ons eindrapport 'Het Meten van Impact van Wetenschapscommunicatie: Eindrapport IMPACTLAB', zijn gebaseerd. Eerst evalueren we aan de hand van statistische toetsen de kwaliteit van het basisinstrument. Daarna beantwoorden we een veelgehoorde vraag onder de gebruikers van onze materialen: wat zijn de belangrijkste vragen van het basisinstrument? Deze laatste analyse is vooral van belang wanneer de basismeting vanwege praktische redenen moet worden ingekort.

### **Inhoudsopgave:**

<i>Kwaliteit en validiteit basisinstrument</i>	<i>P. 2</i>
<i>Resultaten bij kinderen</i>	<i>P. 2</i>
<i>Resultaten bij volwassenen</i>	<i>P. 6</i>
<i>Belangrijkste vragen van het basisinstrument</i>	<i>P.10</i>
<i>Resultaten bij kinderen</i>	<i>P. 10</i>
<i>Resultaten bij volwassenen</i>	<i>P. 12</i>

## Kwaliteit en validiteit basisinstrument

Een goed meetinstrument meet op consistente wijze wat gemeten moet worden. Om de kwaliteit van ons basisinstrument kritisch te evalueren kijken we naar betrouwbaarheid (Hangen de vragen binnen een onderdeel voldoende samen?) en validiteit (Meten de vragen wat ze moeten meten?). Om deze kwaliteiten van ons basisinstrument te beoordelen hebben we metingen van 3 projecten bij elkaar gevoegd. Omdat we vaak aan kinderen net iets andere vragen hebben gesteld dan aan volwassenen, doen we de analyses voor beide doelgroepen apart.

We gebruiken daarvoor de volgende toetsen:

- **Categorische factoranalyse.** Dit is een exploratieve methode waarin we met een open blik naar de data kijken om de onderliggende samenhang tussen verschillende vragen te vinden. Aan de hand van de uitkomsten vormen we ideeën over welke factoren er in de data zitten. Vervolgens kijken we of die overeenkomen met de onderdelen die we met de vragenlijst proberen te meten. Omdat de vragen categorische data opleveren, hebben we gekozen voor de categorische versie van de factoranalyse, de Categorical Principal Components Analysis.
- **Correlatie-analyse.** Dit is een statistische analysetechniek waarmee we kunnen nagaan of de onderdelen in de vragenlijst meten wat ze beogen te meten. We onderzoeken dit door te controleren of de onderdelen van de vragenlijst op een logische manier samenhangen. We berekenen eerst per onderdeel de gemiddelde score om een bij benadering continue variabele te creëren. Daarna berekenen we correlaties tussen de gemiddelden.

## Resultaten bij kinderen

Om de kwaliteit van het basisinstrument binnen de doelgroep kinderen te evalueren, hebben we de datasets van drie metingen samengevoegd:

- Rekenen voor je Leven, november 2021: N = 52 (26 meisjes en 26 jongens)
- Expeditie NEXT, mei 2022: N = 105 (54 meisjes en 51 jongens)
- Leiden Science Family Day, mei 2022: N = 60 (35 meisjes en 25 jongens)

De totale steekproef is N = 217. Het grootste deel (88.9%) zijn kinderen tussen de 6-12 jaar. De gemiddelde leeftijd is M = 9.20 jaar (SD = 2.18).

### 1. Categorische factoranalyse

De categorische factoranalyse onderscheidt vier onderliggende factoren: Wetenschapskapitaal, Plezierbeleving (Emotioneel geheugen), Intensiteit (Emotioneel geheugen) en Effect. We zien dat de emotionele beleving van plezier en intensiteit zijn statistisch van elkaar te onderscheiden zijn: kinderen ervaren plezier en intensiteit als twee verschillende dingen. De factoren komen overeen met de onderdelen die het basisinstrument beoogt te meten. In onderstaande tabel is bovendien te zien dat de vragen in het algemeen goed passen bij de onderdelen waar ze oorspronkelijk bij horen. We kunnen daarom concluderen dat de vragen over het algemeen meten wat ze beoogen te meten.

Wel valt op dat sommige vragen eigenlijk meer samenhangen met een andere factor dan waar ze oorspronkelijk bij horen. De vraag over interessantheid hangt bijvoorbeeld sterker samen met de factor Effect dan met de factor Intensiteit (Emotioneel geheugen), waar hij eigenlijk bij hoort. De vraag over verandering in gedrag laadt hoger op de factor Wetenschapskapitaal dan op de factor Effect, en zo zijn er nog enkele voorbeelden. Hieruit hoeven we niet meteen te concluderen dat deze vragen beter passen bij een ander onderdeel. Inhoudelijk zijn deze vragen bij de juiste onderdelen ondergebracht. De verdere uitkomsten van de validiteitsanalyse tonen bovendien aan dat de vragenlijst in zijn huidige vorm valide en betrouwbaar is (zie verderop).

Dat de vragen per onderdeel niet altijd perfect samenhangen, is te verklaren. Om het basisinstrument zo beknopt mogelijk te houden, hebben we vragen geformuleerd die elk iets anders meten en dus weinig overlappen. Dit komt ten goede aan de praktische bruikbaarheid: je meet zoveel mogelijk aspecten met zo weinig mogelijk vragen. Een nadeel is echter dat de interne consistentie van de vragenlijst hierdoor afneemt. Uit de uitkomsten onder kinderen kunnen we concluderen dat vooral de vragen over interessantheid en verandering van gedrag niet zo duidelijk onder een onderdeel vallen, en dus beter als losse metingen kunnen worden gezien.

	Wetenschaps- kapitaal	Plezierbeleving (Emotioneel geheugen)	Intensiteit (Emotioneel geheugen)	Effect
Wetenschappelijke geletterdheid	.640	.049	-.032	-.151
Houding en waarden	.417	.476	.059	.546

Activiteit	.713	.019	.151	.026
Omgeving	.595	-.052	-.013	.041
Geluk	-.069	.955	.020	.014
Ontspanning	.137	.684	.135	.572
Hoop	.212	.820	.099	.216
Tevredenheid	-.093	.952	.013	-.031
Spanning	.098	.023	.550	.607
Energie	.073	.024	.856	.084
Opwinding	.103	.105	.817	-.105
Interessantheid	.230	.329	.234	.737
Verandering in kennis	-.097	-.005	-.021	.669
Verandering in interesse	.140	-.008	.037	.862
Verandering in houding	-.127	.112	-.117	.590

Verandering in gedrag	.573	.116	.226	.225
-----------------------	------	------	------	------

De vier factoren verklaren ieder voldoende variantie (Eigenwaardes > 1) en hebben een acceptabele betrouwbaarheid (Cronbach's alpha > .60). Ook verklaren de vier factoren gezamenlijk een groot deel van de variantie (65.1%).

	Eigenwaarde	Cronbach's alpha
Wetenschapskapitaal	1.97	.630
Plezierbeleving (Emotioneel geheugen)	3.34	.789
Intensiteit (Emotioneel geheugen)	1.88	.614
Effect	3.22	.791

## 2. Correlatieanalyse

Uit de correlatieanalyse blijkt dat alle onderdelen positief met elkaar samenhangen. Dat betekent dat als deelnemers hoog scoren op een onderdeel, ze meestal ook op de andere onderdelen hoog scoren en vice versa. Vooraf verwachtten we een positief verband tussen emotioneel geheugen en effect. Deze verwachting komt uit: **hoe hoger het beleefde plezier en de beleefde intensiteit, hoe sterker de impact. Verder blijkt dat kinderen met een hoger wetenschapskapitaal gemiddeld meer plezier en intensiteit ervaren, en dat de impact van publieksactiviteiten op hen groter is.**

	Wetenschapskapitaal	Plezierbeleving (Emotioneel geheugen)	Intensiteit (Emotioneel geheugen)
Plezierbeleving (Emotioneel geheugen)	.323**	-	
Intensiteit (Emotioneel geheugen)	.278**	.621**	-

Effect	.415**	.484**	.496**
--------	--------	--------	--------

\*\* = correlaties significant met p-waarde < .01

### Conclusie kwaliteit en validiteit basisinstrument bij doelgroep kinderen

Samengevat zien we dat het basisinstrument bij kinderen een goede meting vormt van wetenschapskapitaal, emotioneel geheugen en impact. Zoals verwacht hangen de vragen binnen een onderdeel in de meeste gevallen goed samen, waardoor we kunnen concluderen dat ze hetzelfde onderdeel meten, al zijn er ook uitzonderingen. Vooral de vragen over interessantheid en verandering van gedrag zijn niet zo makkelijk onder een onderdeel te scharen en kunnen beter als losse metingen worden gezien. Verder blijkt dat plezier en intensiteit weliswaar samenhangen, maar niet hetzelfde zijn. Het is daarom raadzaam om beide onderdelen apart te meten. Al met al concluderen we dat de vragenlijst in zijn huidige vorm valide en betrouwbaar is, en daardoor geschikt voor afname bij kinderen.

### Resultaten bij volwassenen

Om de kwaliteit van het basisinstrument binnen de doelgroep volwassenen te evalueren, hebben we wederom de datasets van drie metingen samengevoegd:

- Stoeplantjesonderzoek, tussen maart en mei 2022: N = 203 (149 vrouwen, 47 mannen, 3 non-binair, 4 'zeg ik liever niet')
- Expeditie NEXT, mei 2022: N = 84 (16 mannen en 68 vrouwen)
- LSFDD, mei 2022: N = 53 (27 mannen, 25 vrouwen, 1 anders)

De totale steekproef is dus N = 340. Leeftijd 17-85 jaar. De gemiddelde leeftijd is M = 51.77 jaar (SD = 15.24).

#### 1. Categorische factoranalyse

De categorische factoranalyse onderscheidt drie onderliggende factoren. De twee aspecten van emotioneel geheugen - plezierbeleving en intensiteit - worden door volwassenen niet van elkaar onderscheiden. Wanneer volwassenen plezier beleven, beleven ze kennelijk ook intensiteit, en vice versa. Zoals te zien is in onderstaande tabel passen de vragen goed bij de onderdelen waar ze oorspronkelijk bij horen. We kunnen daarom concluderen dat de vragen over het algemeen meten wat ze beogen te meten.

Er is een vraag die afwijkt van het verwachte patroon: de vraag over verandering in gedrag hangt eigenlijk meer samen met de factor Emotioneel geheugen dan de factor Effect. Ook dit is te zien in onderstaande tabel. Bij de kinderen laadde deze vraag juist hoger op Wetenschapskapitaal. In verschillende contexten en bij verschillende doelgroepen heeft de uitkomst op deze vraag kennelijk een andere betekenis, wat afbreuk doet aan de validiteit van de vraag. Dit is deels te verklaren door onze keuze om het instrument beknopt te houden en overlap tussen vragen te vermijden, zoals we hiervoor al toelichten bij de resultaten bij kinderen. Desondanks zouden we zeker niet aanraden om de vraag over verandering in gedrag te schrappen uit het basisinstrument. Wel is het raadzaam om deze vraag als een losse meting te gezien, in plaats van een vraag die sterk samenhangt met de andere Effect-vragen. Dit advies is in lijn met de verdere validatie-analyse, waaruit blijkt dat het basisinstrument valide en betrouwbaar is in zijn huidige vorm (zie verderop).

	Wetenschaps- kapitaal	Emotioneel geheugen	Effect
Wetenschappelijke geletterdheid	.845	-.185	.156
Houding en waarden	.812	-.087	.061
Activiteit	.677	.032	-.129
Omgeving	.805	-.094	.140
Geluk	-.013	.859	.083
Ontspanning	-.002	.778	.081
Hoop	-.019	.830	.118

Tevredenheid	.037	.864	.046
Spanning	-.144	.786	.134
Energie	-.114	.850	.093
Opwinding	-.191	.722	-.017
Interessantheid	.025	.730	.214
Verandering in kennis	.087	.247	.813
Verandering in interesse	.088	.379	.724
Verandering in houding	.018	-.079	.839
Verandering in gedrag	-.204	.609	.089

De drie factoren verklaren ieder voldoende variantie (Eigenwaardes > 1) en hebben een acceptabele betrouwbaarheid (Cronbach's alpha > .60). Ook verklaren de vier factoren gezamenlijk een groot deel van de variantie (65.5%).

	Eigenwaarde	Cronbach's alpha
Wetenschapskapitaal	2.61	.689
Emotioneel geheugen	5.81	.891
Effect	2.06	.658



## 2. Correlatieanalyse

Uit de correlatieanalyse blijkt zoals verwacht dat **hoe sterker de emotionele beleving is, hoe hoger de impact**. Interessant is dat er een **zwakke negatieve relatie is tussen wetenschapskapitaal en emotioneel geheugen**. Volwassenen met een relatief hoog wetenschapskapitaal ervaren naar verhouding dus iets minder plezier en intensiteit tijdens een wetenschapscommunicatie-evenement. Wellicht is dit te verklaren door hun vertrouwdheid met wetenschap, waardoor de ervaring minder speciaal is. **Tegelijkertijd beleven volwassenen met een relatief laag wetenschapskapitaal gemiddeld meer plezier en intensiteit**. Misschien is een wetenschapscommunicatie-evenement voor hen iets verrassender. In tegenstelling tot de uitkomsten bij de kinderen, is er bij volwassenen geen direct verband tussen wetenschapskapitaal en impact. **Wetenschapscommunicatie-evenementen hebben in het algemeen dus evenveel impact op volwassenen met meer of minder wetenschapskapitaal**.

	Wetenschapskapitaal	Emotioneel geheugen
Emotioneel geheugen	-.170**	-
Effect	.066	.390**

\*\* = correlaties significant met p-waarde < .01

### Conclusie kwaliteit en validiteit basisinstrument bij doelgroep volwassenen

Samengevat zien we dat het basisinstrument ook bij volwassenen een goede meting vormt van wetenschapskapitaal, emotioneel geheugen en impact. Zoals verwacht hangen de vragen binnen een onderdeel in de meeste gevallen goed samen en meten ze hetzelfde onderdeel. Er was een uitzondering: de vraag over verandering van gedrag is niet makkelijk bij een onderdeel te scharen en kan daarom beter als losse meting worden gezien. Bij volwassenen overlappen de emotionele beleving van plezier en intensiteit en vormen samen een onderdeel. Wanneer volwassenen plezier beleven, beleven ze blijkbaar ook intensiteit, en vice versa. Al met al blijkt de vragenlijst in zijn huidige vorm valide en betrouwbaar, en daardoor geschikt voor afname bij volwassenen.

## Belangrijkste vragen van het basisinstrument

Voor een valide en betrouwbare meting adviseert het IMPACTLAB om alle vragen van het basisinstrument mee te nemen. Toch is er tijdens een publieksactiviteit niet altijd ruimte om het volledige basisinstrument af te nemen. Het is daarom waardevol om te weten welke vragen essentieel zijn, zodat wanneer je de vragenlijst inkort, je in elk geval de belangrijkste vragen meeneemt.

Om de vragenlijst in te korten is ons eerste en belangrijkste advies om eerst zelf na te gaan welke vragen in het basisinstrument voor jou het meest sprekend en relevant zijn. Naar welke aspecten van je publiek ben je nieuwsgierig? Welke emoties wil je teweegbrengen? Welke vorm van impact is relevant, gezien de opzet en het doel van je publieksactiviteit?

Misschien is het toch lastig om deze knopen door te hakken en ben je op zoek naar een onderbouwd advies op basis van ons onderzoek. Hieronder beschrijven we ons advies op basis van statistische analyses. We beschrijven de analyses en adviezen apart voor kinderen en volwassenen.

### Resultaten bij kinderen

In onderstaande tabellen zijn de vragen per onderdeel van meest naar minst representatief gerangschikt op basis van de item-restcorrelatie. We geven per onderdeel aan welke vragen je het beste kunt meenemen als je ervoor kiest om met een ingekorte versie van het basisinstrument te werken.

Bij Wetenschapskapitaal zijn **Activiteit** en **Houding en waarden** de meest representatieve vragen voor kinderen.

Wetenschapskapitaal	
Activiteit	.413
Houding en waarden	.319
Omgeving	.274
Wetenschappelijke geletterdheid	.255

Bij Plezierbeleving (Emotioneel geheugen) zijn **Geluk** en **Tevredenheid** de meest representatieve vragen voor kinderen.

Plezierbeleving (Emotioneel geheugen)	
Geluk	.754
Tevredenheid	.710
Ontspanning	.709
Hoop	.704

Bij Intensiteit (Emotioneel geheugen) zijn **Energie** en **Spanning** de meest representatieve vragen voor kinderen.

Intensiteit (Emotioneel geheugen)	
Energie	.580
Spanning	.501
Interessantheid	.469
Opwinding	.381

Bij Effect zijn **Kennis** en **Interesse** de meest representatieve vragen voor kinderen.

Effect	
Verandering in kennis	.428
Verandering in interesse	.402
Verandering in gedrag	.339
Verandering in houding	.278

## Resultaten bij volwassenen

In onderstaande tabellen zijn de vragen per onderdeel van meest naar minst representatief gerangschikt op basis van de item-restcorrelatie. We geven per onderdeel aan welke vragen je het beste kunt meenemen als je ervoor kiest om met een ingekorte versie van het basisinstrument te werken.

Bij Wetenschapskapitaal zijn **Wetenschappelijke geletterdheid** en **Omgeving** de meest representatieve vragen voor volwassenen.

Wetenschapskapitaal	
Wetenschappelijke geletterdheid	.710
Omgeving	.653
Houding en waarden	.638
Activiteit	.418

Bij Plezierbeleving (Emotioneel geheugen) zijn **Geluk** en **Tevredenheid** de meest representatieve vragen voor volwassenen.

Plezierbeleving (Emotioneel geheugen)	
Geluk	.776
Tevredenheid	.773
Hoop	.726
Ontspanning	.686

Bij Intensiteit (Emotioneel geheugen) zijn **Energie** en **Spanning** de meest representatieve vragen voor volwassenen.

Intensiteit (Emotioneel geheugen)	
-----------------------------------	--

Energie	.803
Spanning	.737
Interessantheid	.622
Opwinding	.514

Bij Effect zijn **Kennis** en **Interesse** de meest representatieve vragen voor volwassenen.

Effect	
Verandering in kennis	.637
Verandering in interesse	.623
Verandering in houding	.368
Verandering in gedrag	.170