

# Doe dat nog eens?

E. VRIEZE



Hoe ze dit titanenwerk precies voor elkaar gekregen hebben is mij niet helemaal duidelijk, maar de deelnemers aan het Open Science Collaboration (2015) publiceerden een onderzoek in *Science* over de herhaalbaarheid van een serie gepubliceerde artikelen uit 2008, van 3 grote *peer-reviewed* tijdschriften in het domein van de psychologische wetenschappen. Daarvoor hebben meer dan 250 onderzoekers in hun vrije tijd de experimenten van ruim 100 gepubliceerde artikelen herhaald, om te onderzoeken of de resultaten replicerbaar waren. Hoewel 97% van de originele studies significante resultaten claimde, waren slechts 36 van de 100 herhaalde experimenten statistisch significant. Verder waren de effectgroottes van de originele studies gemiddeld tot de helft verminderd na replicatie.

## Replicatie

Dit replicatieproject is een van de eerste systematische onderzoeken waarin de betrouwbaarheid van psychologische experimenten onder de loep genomen wordt. Het resultaat van dit onderzoek suggereert dat gepubliceerde bevindingen in het domein van de psychologische wetenschappen gemiddeld half zo sterk zijn dan ze zich voordoen; slechts één derde is werkelijk statistisch significant. Verder toont het artikel aan dat studies uit de cognitieve psychologie (zoals onderzoek naar aandacht, geheugen, leerparadigma's etc.) beter reproduceerbaar zijn dan die uit de sociale psychologie (hoe mensen elkaar beïnvloeden). Dit laatste zou mogelijk verklaard kunnen worden doordat onderzoek van sociaal psychologen meer afhankelijk is van de context, waardoor de effectgrootte vaak kleiner is.

## Crisis?

Voor vele onderzoekers zijn de bevindingen van het *Science*-artikel een aanwijzing dat de psychologische wetenschappen afstevenen op een onderzoekscrisis en dat een grondige hervorming nodig is. Repliceerbaarheid van data is een kerneigenschap van goed en waardevol wetenschappelijk onderzoek. Al veel langer is er een debat over publicatiebias (de neiging om positieve resultaten van wetenschappelijk onderzoek wél, maar negatieve resultaten niet te publiceren) of 'p-hacking', die wordt aangewakkerd door publicaties van onmogelijke effecten (bijv. precognities), mislukte replicaties van tekstboekexperimenten en enkele prominente fraudezaken wereldwijd. Er lijken gewoontes in ons onderzoeksdomein gekropen, zoals selectieve rapportage, selectieve analyses en slechte specificatie van onderzoekscondities, die ter discussie gesteld moeten worden. Zo is het vrij gangbaar dat een onderzoeker eerst de statistische significantie nakijkt van een kleine steekproef voordat hij of zij beslist meer data te verzamelen. Hoewel misschien onopzettelijk, veroorzaken deze handelingen valse resultaten.

## Andere visie

Andere onderzoekers zijn echter minder ongerust over de onderzoeksresultaten van de Open Science Collaboration (Gilbert e.a. 2016). Zij stellen dat één derde significantie te verwachten was door de statistische methoden van de replicatiestudie. Daarnaast is herhaalbaarheid van één derde van de resultaten misschien nog niet eens zo slecht. Onderzoek van hoog niveau omvat het nemen van risico's en creativiteit, waarbij fout-positieve resultaten nu eenmaal te verwachten zijn.

Ook is het de vraag of de resultaten die niet replicateerbaar bleken in de studie werkelijk onjuist zijn en de succesvol gerepliceerde resultaten de waarheid belichamen. Er zijn immers verschillende redenen waarom de resultaten niet overeen kunnen komen. Zo is er een gerandomiseerde kans: het origineel kan fout zijn, maar dat geldt ook voor de replicatiepoging. Daarnaast kunnen kleine verschillen tussen de uitvoering en de opzet van beide experimenten over het hoofd gezien zijn.

Ten slotte zou je vragen kunnen stellen over hoe vaardig de replicatieteams waren. Misschien zouden deze teams ook slechts één derde significantie repliceren bij het afnemen van sterk gevalideerde, bekende psychologische experimenten, zoals de Stroop-test (waarbij mensen trager een kleur benoemen van een woord, als het woord in een andere kleur gedrukt staat). Daarbij zouden de replicatieteams zelf natuurlijk ook een vooroordeel kunnen hebben waardoor ze (onbewust) de herhaalbaarheid van het originele experiment hebben beïnvloed. Echter, de meeste replicatoren waren gemotiveerd, werkten systematisch en waren in contact met de oorspronkelijke onderzoekers, die toegang hebben gegeven tot materiaal etc. Slechts 3 van de 100 oorspronkelijke onderzoekers weigerden mee te werken. Kortom, het is zo goed als onmogelijk te evalueren of één derde replicateerbaarheid goed, slecht of precies juist is.

## Mooi initiatief

Ik vind dat dit een 3 jaar durend indrukwekkend experiment geweest is dat alleen maar sterk gewaardeerd en aangemoedigd kan worden. Hoe je het ook bekijkt, een groot deel van de onderzochte studies verloren hun sterkte en vaak ook significantie na herhaling van het experiment. Er mogen dan verschillende hypothesen zijn waarom bij herhaling de experimenten geen stand hielden, het is duidelijk dat er nog weinig inzicht is in de precieze oorzaken.

Wetenschappelijk werk zichtbaar maken (transparantie) en onafhankelijke replicatie zijn echter kerningrediënten van alle wetenschappelijk onderzoek. Het Center of Open Science (<https://cos.io>), opgericht door de trekkers van de replicatiestudie, heeft inmiddels een aantal richtlijnen opgesteld voor tijdschriften en onderzoeksteams om transparantie en openheid te vergroten. Verder proberen ze onderzoekers te stimuleren tot een pre-registratie van onderzoeksplannen, met duidelijke hypothesen, systematische onderzoeksstappen en voldoende statistische power, zodat er weinig ruimte is voor p-hacking. Ik ben alvast fan van dit initiatief.

## LITERATUUR

- Gilbert DT, King G, Pettigrew S, Wilson TD. Comment on 'Estimating the reproducibility of psychological science'. *Science* 2016; 351: 6277.
- Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* 2015; 349: 6251.

## AUTEUR

**ELSKE VRIEZE**, UPC K.U. Leuven, campus Leuven, Hereweg 78, B-3000 Leuven, België.  
E-mail: [elske.vrieze@uzleuven.be](mailto:elske.vrieze@uzleuven.be)

Geen strijdige belangen meegedeeld.

## TITLE IN ENGLISH

Do that again please?