

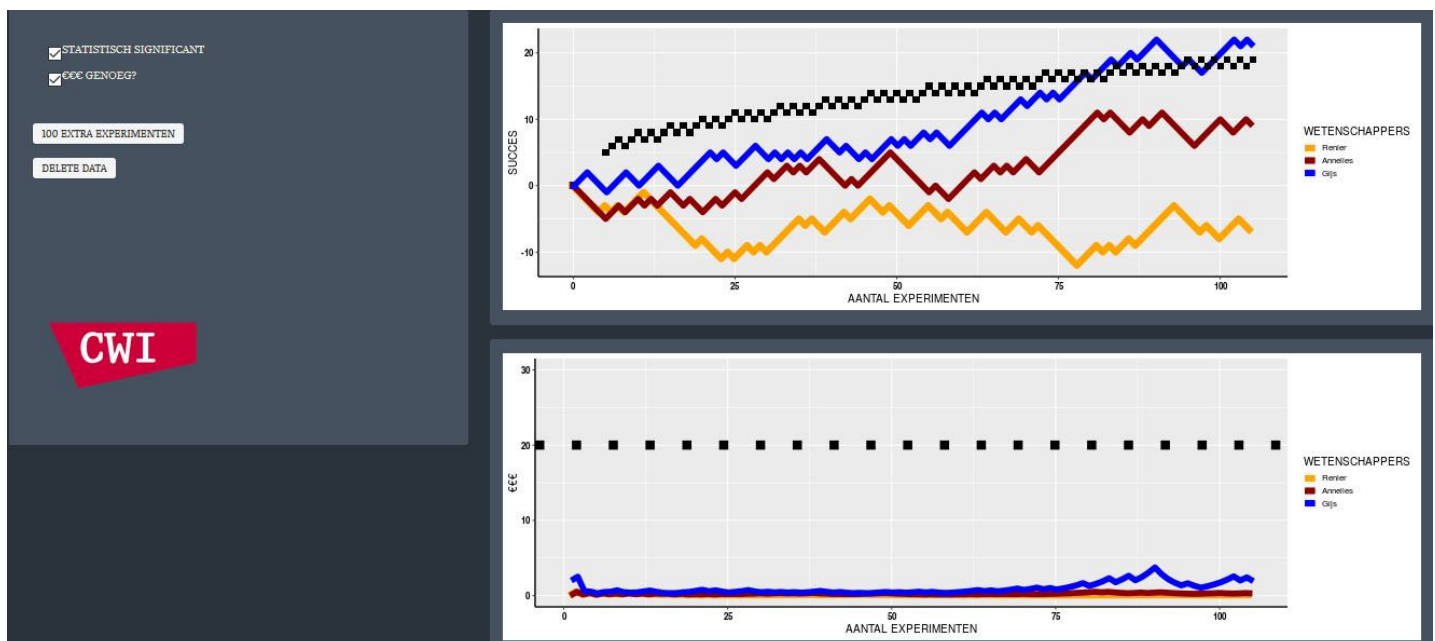
HOME (/) KENNISMAKEN (/KENNISMAKEN)

PROJECTEN (/PROJECTEN) **BLOG (/BLOG)**

CONTACT (/CONTACT) ENGLISH (/ENGLISH)

Inschrijven voor de nieuwsbrief (<http://eepurl.com/cwEqFD>)

WANNEER HEB JE EEN EXPERIMENT VOLDOENDE HERHAALD OM OVERTUIGD TE KUNNEN ZIJN VAN JE RESULTATEN?



Om die vraag draait mijn praatje tijdens de Nacht van de Wetenschap. Probeer het vooral ook zelf uit, met het scorebord dat ik heb gemaakt. Als je in het scorebord aanklikt wat in de tekst hieronder in het **oranje** is aangegeven (zoals het toeval tijdens de Nacht van de Wetenschap op 5 oktober bepaalde) kun je de resultaten van toen reproduceren. Als je nieuwe data wil uitproberen, klik dan gewoon op 'delete data' en genereer nieuwe succesvolle of mislukte experimenten met een muntje of laat de computer gooien met de '20/100 extra experimenten' knop.

NAAR SCOREBORD (<https://significanthelp.shinyapps.io/scorebord/>)

Eveline heeft een promofilmje gemaakt met een voorbeeld van een experiment waarbij het niet lukte om het succesvol te herhalen (zie hier op Youtube (<https://youtu.be/5xx2KlX1OuQ>)). Het blijkt namelijk dat veel wetenschappelijke theorieën waarvan we dachten dat ze al lang in een experiment bewezen waren, als we een nieuw, groter, en beter experiment doen, niet opnieuw te bewijzen zijn. Volgens sommigen komt dat omdat de helft van de theorieën die wetenschappers onderzoeken gewoon fout zijn!*

Dat is dus best wel verspilling. En in het medisch onderzoek hebben ze daar een woord voor: Research Waste**. Verspilling in wetenschappelijk onderzoek. Laat ik een voorbeeld geven. Al sinds de jaren zestig bestaat er een wetenschappelijke theorie die stelt dat een hormoon (corticosteroïde), dat wij zelf ook aanmaken, kan helpen bij kinderen en volwassenen met een hersenvliesontsteking. Sindsdien zijn er 23 experimenten gedaan om die theorie te bewijzen, en hebben we ook vele jaren in deze theorie geloofd. Pas in 2010 werd er op een manier naar al die experimenten bij elkaar gekeken die liet zien dat de theorie niet bewezen was.*** Er was geen bewijs dat het hormoon had geholpen. Veel opvallende experimenten hadden net zo goed toeval kunnen zijn geweest. Als ze daar eerder achter waren gekomen, hadden er minder onnodige experimenten gedaan hoeven worden. Dat had miljoenen euro's kunnen schelen. De hersenvliesontstekingwetenschappers hadden misschien een andere wetenschappelijke theorie kunnen onderzoeken. Misschien hadden we de volwassenen en kinderen met hersenvliesontsteking dan ook beter kunnen helpen.

Bronnen:

* Ioannidis (2005) Why Most Published Research Findings Are False (<https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124>)

** <http://rewardalliance.net/about/> (<http://rewardalliance.net/about/>)

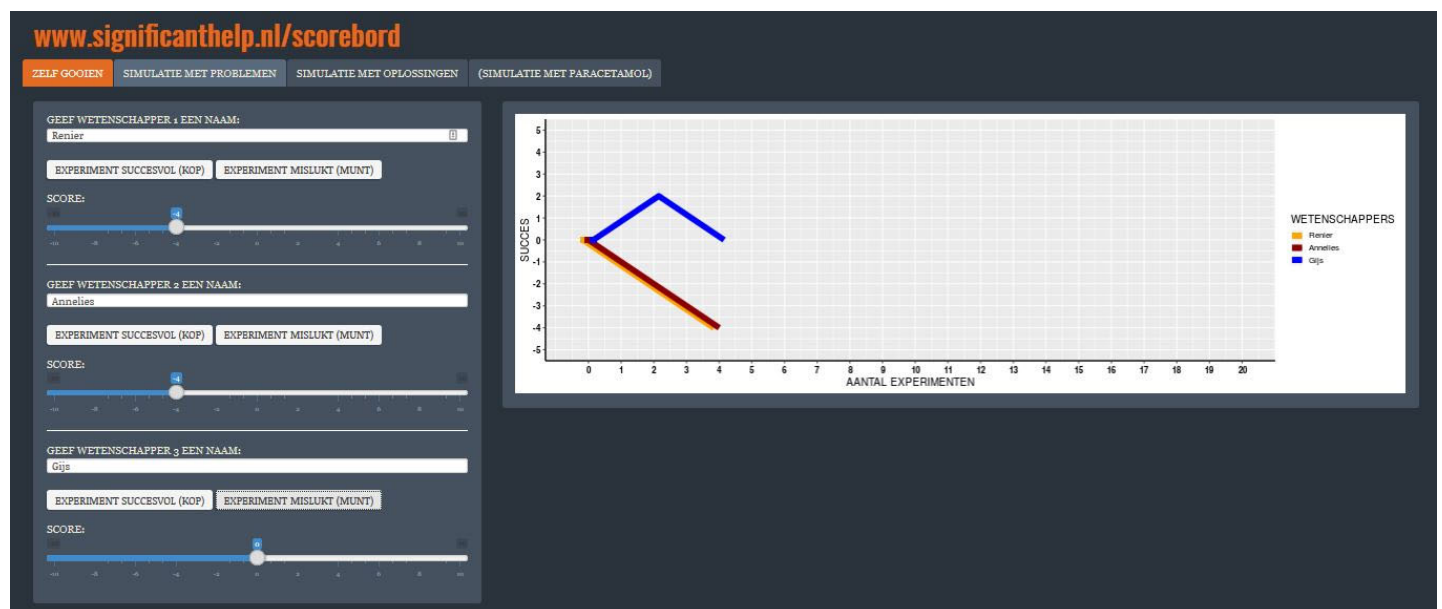
*** Ioannidis (2010) Meta-Research: The Art of Getting it Wrong (<https://onlinelibrary.wiley.com/doi/epdf/10.1002/jrsm.19>)

Praatje

Dit verhaal gaat over toeval. Laat ik met een vraag beginnen: Wie van jullie is wel eens op vakantie een bekende tegengekomen? Handen omhoog als je dit een keer is overkomen. Dat was super-super toevallig toch? Laten we nu eens om ons heen kijken hoeveel handen er omhoog zijn gegaan. Als iets zoveel mensen overkomt, is het dan nog toevallig? Meneer hier vooraan, wie bent u tegengekomen op vakantie? <Mijn automonteur.> Bij mij was het een meisje uit mijn zusjes' voetbalteam. Als we eens stilstaan bij hoeveel mensen we kennen, de kring van mensen om ons heen waar zowel je automonteur als het voetbalteam van mijn zusje bij hoort, dan hebben we best veel bekenden! En dat is precies de reden dat zoveel van ons iemand tegenkomen op vakantie. Ook al is de kans dat we een specifiek iemand tegenkomen, terwijl we dat niet hebben afgesproken, best klein. Omdat we allemaal best vaak op vakantie gaan en best veel mensen kennen, kunnen er heel veel combinaties van mensen en vakanties voorkomen. En als iets met een kleine kans maar genoeg pogingen mag doen om voor te komen, dan gebeurt het toch.

Het toeval is zowel de held als de schurk van dit verhaal; het kan je verrassen, maar ook misleiden. En wij mensen zijn niet zo goed in het toeval herkennen. Iets dat heel erg lastig is voor wetenschappers. Als wetenschappers een wetenschappelijke theorie hebben, moeten ze die namelijk bewijzen met een experiment, bijvoorbeeld met proefpersonen, in een laboratorium of met patiënten in een ziekenhuis. Dat experiment moet slim zijn ontworpen zodat het resultaat bewijs kan leveren dat de wetenschappelijke theorie klopt. Dus moeten wetenschappers beoordelen of het succes van hun experiment er is dankzij hun bijzondere experiment, en niet gewoon vanwege toeval.

Om eens te kijken of wij weten hoe toeval eruit ziet heb ik voor de pauze drie vrijwilligers gevraagd om toeval te maken. **Scorebord verschijnt.** En dit zijn de helden die voor ons muntjes hebben gegooid. Ze kregen een pluspunt voor iedere keer dat ze kop gooiden, en een minpunt voor iedere keer dat ze munt gooiden. Wie zien de scores van Renier in het geel, van Annelies in het rood, en van Gijs in het blauw. Gijs gaat aan de leiding met 2 keer kop. [Renier gooide MUNT, MUNT, MUNT MUNT; Annelies gooide MUNT, MUNT, MUNT, MUNT, Gijs gooide KOP, KOP, MUNT, MUNT]

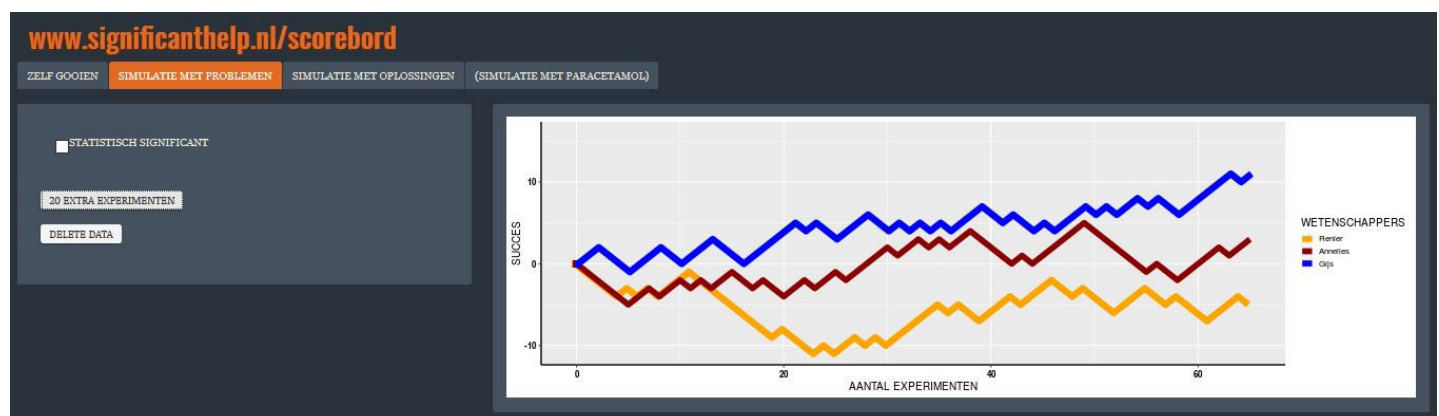


Nu hoop ik dat jullie het met mij eens zijn dat het puur toeval was dat Gijs aan de leiding gaat. Geen van deze drie was namelijk beter in kop gooien, voor iedereen was het puur toeval wat ze gooiden. En dat kan ook het geval zijn voor wetenschappers.

Bijvoorbeeld voor een onderzoek naar hoofdpijn. In een experiment met paracetamol zitten tenminste twee mensen met hoofdpijn: de ene geef je paracetamol, de andere geef je een neppil, en vervolgens ga je kijken wie het eerst van z'n hoofdpijn af is. Alleen net als met het muntjes gooien kan het ook voor een wetenschapper puur toeval zijn of zo'n experiment succesvol is, en laat zien dat je met de paracetamol sneller van je hoofdpijn af bent, of dat het experiment mislukt en de paracetamol juist langzamer was. We hebben immers allemaal wel eens een halfuurtje hoofdpijn gehad, en ook allemaal wel eens een hele dag. Dus misschien was het wel toeval dat de hoofdpijn sneller over was, alsof het succes van het experiment kwam omdat je toevallig kop hebt gegooid.

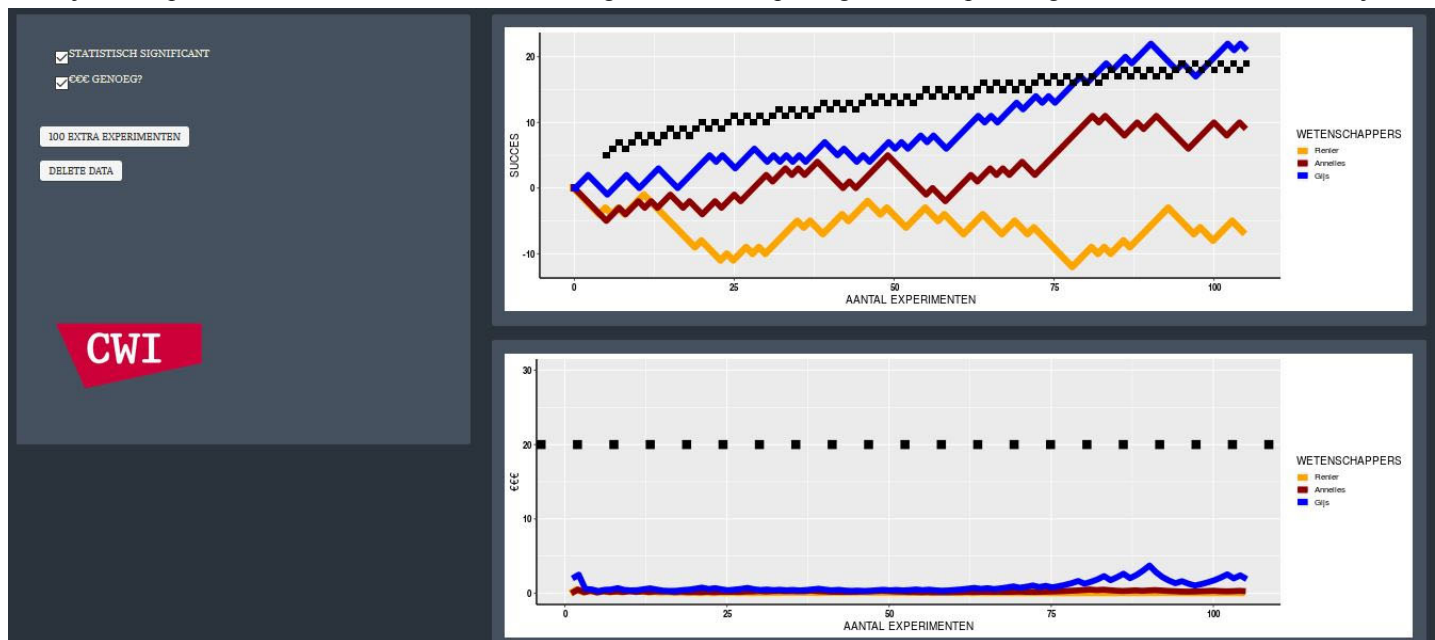
Dus bij deze wil ik je nogmaals voorstellen aan mijn vrijwilligers. Renier in het geel, doet onderzoek naar mentos tegen hoofdpijn. Annelies, in het rood, doet onderzoek naar tiktak tegen hoofdpijn. En Gijs, in het blauw, doet onderzoek naar stoepkrijtjes tegen hoofdpijn, van die dropjes met zo'n witte laag eromheen, ziet er ook wel een beetje uit als een pil. Als geen van die snoepjes helpt tegen hoofdpijn is ieder experiment gewoon een worp met een muntje. Zouden jullie nog een keer willen gooien? Renier gooit kop, een succesvol experiment. **De knop 'EXPERIMENT SUCCESVOL (KOP)' wordt aangeklikt onder Reinier.** Annelies gooit munt, een mislukt experiment. **De knop 'EXPERIMENT MISLUKT (MUNT)' wordt aangeklikt onder Annelies.** Gijs gooit munt, een mislukt experiment. **De knop 'EXPERIMENT MISLUKT (MUNT)' wordt aangeklikt onder Gijs.**

Gijs staat (nog steeds) voor. Je vergeet bijna dat dit nog steeds allemaal toeval is. Maar dat weet je toch wel als je vaker blijft gooien? Je onderzoek herhalen? Laten we nu de computer nog eens heel veel meer muntjes opgooien. **We verplaatsen naar tabblad 'SIMULATIE MET PROBLEMEN'. De knop 'EXTRA EXPERIMENTEN' wordt aangeklikt: 5x.** Vinden jullie dit er nog uitzien als toeval? Het lijkt echt alsof Gijs iets op het spoor is met z'n stoepkrijtjes tegen hoofdpijn! Voor een wetenschapper is dat dus ook erg lastig! Dus gebruiken ze statistiek om hun experimenten te beoordelen. Maar wat ik jullie wil laten zien, is dat de finish die de meeste wetenschappers nu gebruiken, die van statistische significantie, niet zo goed beschermt tegen dit soort toeval. **'STATISTISCH SIGNIFICANT' wordt aangevinkt.** Als je door deze stippellijn gaat, dan denken de meeste wetenschappers dat ze hun wetenschappelijke theorie hebben aangetoond. Bijvoorbeeld dat tiktak helpt tegen hoofdpijn. Maar daarmee houden we onszelf dus voor de gek!



Wetenschappers gaan misschien wel veel te lang door met hun experimenten, omdat ze denken dat ze deze finish nog kunnen halen. Dat zorgt dus voor Research Waste. We kunnen echt miljarden besparen als wetenschappers anders naar hun experimenten gaan kijken, met minder valse hoop. Hoe dan? Daarover gaat mijn onderzoek.

We verplaatsen naar tabblad 'SIMULATIE MET OPLOSSINGEN'. We hebben al voor veel experimenten gevonden hoe het moet. Voor de muntjes-experimenten zie je de grafiek met de nieuwe berekening hieronder. Dit zijn dezelfde experimenten, hier zie je bijvoorbeeld in het begin dat Gijs aan de leiding gaat met 2 keer kop.



De onderste grafiek telt alleen niet gewoon alle successen op, maar remt het na een tijdje af. Net als een slechte pokeraar op beginnersgeluk wat kan winnen, maar het ook weer verliest als je 'm langer laat doorspelen.

Als je op deze manier het succes van je experimenten uitrekent, dan kom je er op tijd achter dat je succes toeval is. Zo zorgen we voor minder Research waste. Ook wetenschappers hebben af en toe wat statistische superpowers nodig.

Dank jullie wel.

Simulatie specs

Voor de simulatie van paracetamolexperimenten heb ik me losjes gebaseerd op bestaand onderzoek (link (<https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011889.pub2/media/CDSR/CD011889/CD011889.pdf>), $RR = 1.3$) en een kans aangenomen van 0.565 dat je in een experiment ziet dat de paracetamolgroep sneller van de hoofdpijn af is dan de placebogroep. Eerlijk gezegd weten we niet echt of dit zo is. Voor een heerlijke statistische discussie over de werkzaamheid van paracetamol bij hoofdpijn, zie deze blogpost van statisticus Stephen Senn: link (<https://errorstatistics.com/2016/08/02/s-senn-painful-dichotomies-guest-post/>)

Sun, 07 October

Method - Mixed Models - Longitudinal Modeling - Generalized Estimating Equations (GEE) - Bayesian Statistics - Discriminant Analysis - Optimal Scaling - Correspondence Analysis - Principal Component Analysis - Monte Carlo Statistics - Bootstrapping - Recommender System - Bayesian networks - K-means - Support Vector Machines - Multiple Testing Correction - High Dimensional Data Analysis - Ridge/Lasso - Splines - Cross-Validation - Random Forests - Bias/Variance Trade-off - Zero-Inflated Poisson Regression - Spatial/Temporal dependence - Finite Population Correction - Power Analysis - Randomized Incomplete Block Design - Survival Analysis - Censoring - Structural Equation Models - Item Response Theory - Imputation - EM Algorithm - Finite Mixture Models - Calibration Problem in Regression - R - SPSS - Weighted Contrasts - Stratified Design - Questionnaire - Permutation Tests - Fixed/Common-Effect Meta-Analysis - Random Effect Meta-Analysis - Sequential Analysis - Bayes Factors - Bayesian T-test - Optional Stopping - Sequential Analysis - Meta-Analysis - Cumulative Meta-Analysis - Test Martingales - Uniformly Most Powerful Tests - Safe Tests

KvK-nummer 67069479

CONTACT

Significant Help

✉ info@significanthelp.nl
(mailto:info@significanthelp.nl)

📄 ([//in02.hostcontrol.com/resources/2e031b77df0692/55365bc8e6/file-object/Significant Help - Algemene Voorwaarden.pdf](https://in02.hostcontrol.com/resources/2e031b77df0692/55365bc8e6/file-object/Significant%20Help%20-%20Algemene%20Voorwaarden.pdf))
Algemene Voorwaarden
([//in02.hostcontrol.com/resources/2e031b77df0692/55365bc8e6/file-object/Significant Help - Algemene Voorwaarden.pdf](https://in02.hostcontrol.com/resources/2e031b77df0692/55365bc8e6/file-object/Significant%20Help%20-%20Algemene%20Voorwaarden.pdf))



(<https://www.linkedin.com/in/judith-ter-schure-a618a840/>)



(<https://www.cwi.nl/research/groups/machine-learning/personview/3056>)