De Spiegel van Berg Domheid en het Dunning-Kruger Effect

Ongeschoold en zich er niet van bewust – verwijzingen naar het Dunning-Kruger-effect zijn populair in de duikgemeenschap. De bevindingen van het oorspronkelijke onderzoek worden echter vaak verkeerd begrepen of verkeerd voorgesteld. Bovendien is het effect zelf misschien niet eens echt. Een (beetje) serieuze diep duik door Tim Blömeke.

Florida Man. De Darwin Awards. Een of andere nitwit bekeek een video over raketten en blies zijn garage op in een poging ontsnappingssnelheid te bereiken. Of probeerde zijn eigen whisky te maken, met hetzelfde resultaat. De verhalen zijn inwisselbaar, maar de moraal is altijd hetzelfde: je weet niet wat je niet weet, en een beetje kennis is gevaarlijk. Ka-boem.

De duikwereld heeft zijn eigen schat aan verhalen van dit type, waarvan er vele gaan over slechte dingen die in grotten gebeuren met duikers die alleen zijn opgeleid om in de oceaan te duiken. Sommige zijn verfraaide, fictieve verslagen, maar er is geen tekort aan waargebeurde. Ongevalsrapporten uit de begindagen van grotverkenning, of non-fictieboeken zoals Robert Kurson's *Shadow Divers*, zorgen voor wat educatieve (en gruwelijke) lectuur.

In 1999 waren er een paar onderzoekers genaamd David Dunning en Justin Kruger die een artikel publiceerden[1] die deze anekdotes van een wetenschappelijke achtergrond voorzag. De onderzoekers voerden testen uit waarin ze deelnemers van verschillende vaardigheidsniveaus vroegen om hun prestaties in een reeks theoretische toetsen te voorspellen. In een notendop concludeerden Dunning en Kruger dat mensen met weinig kundigheid hun prestaties systematisch overschatten, terwijl mensen met veel kundigheid de neiging hebben om die nauwkeuriger te voorspellen en zelfs enigszins te onderschatten.

Het internet pikte hun studie op en verwerkte het tot memes, die zo ver gingen dat ze zelfs namen gaven aan specifieke terreinkenmerken in een fantasievol gevormde curve die beweerde de opkomst, ondergang en wederopstanding van iemands zelfvertrouwen te laten zien in de loop van iemands leercurve: Berg Domheid, de Vallei der Wanhoop, de Helling van Verlichting, het Plateau van Duurzaamheid.

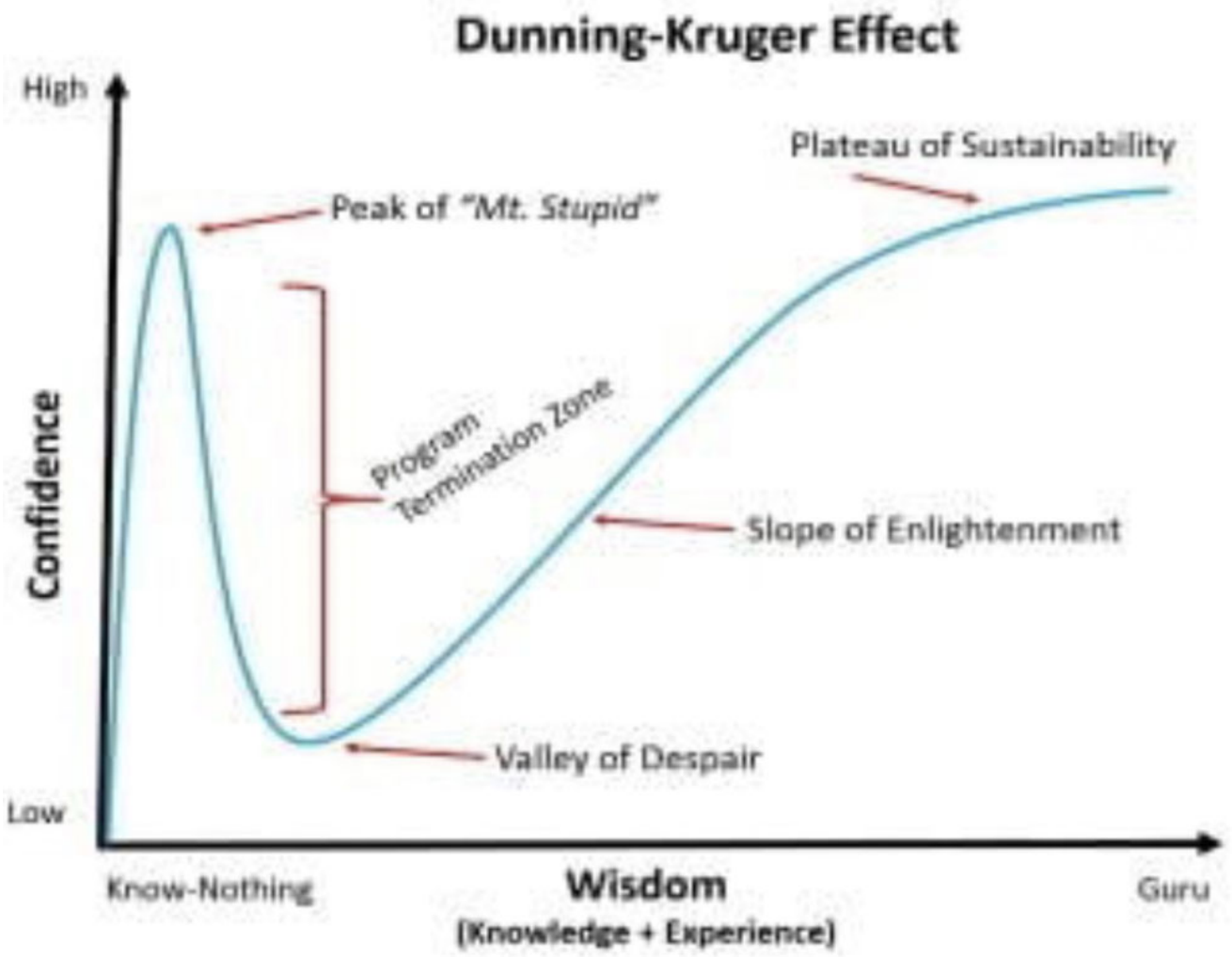

Fig. 1: Je hebt vast wel eens zoiets gezien. Het is geen nauwkeurige weergave van wat Dunning en Kruger zeiden.

De memes waren zo succesvol dat ze zelfs hun weg vonden naar bedrijfsadvies- en managementtrainingsprogramma's. Als gevolg daarvan zijn er nogal wat mensen met dure opleidingen die ze accepteren als een nauwkeurige illustratie van een fenomeen uit de echte wereld. Velen van ons zullen op een of ander moment soortgelijke memes hebben gezien in de context van duiken.

Memes verspreiden zich niet omdat ze waar zijn, maar omdat ze aantrekkelijk zijn, en de bovenstaande doen dat geweldig door ons te herinneren aan de sappige anekdotes die we allemaal zo graag horen. Iedereen heeft een verhaal van "die kerel" (en laten we eerlijk zijn, het zijn meestal mannen die eindigen als hoofdpersoon in verhalen over grote domheid). Een beetje toegepaste scepsis onthult echter al snel waarom we absoluut niet moeten vertrouwen op persoonlijke ervaring als het gaat om het evalueren van empirische claims.

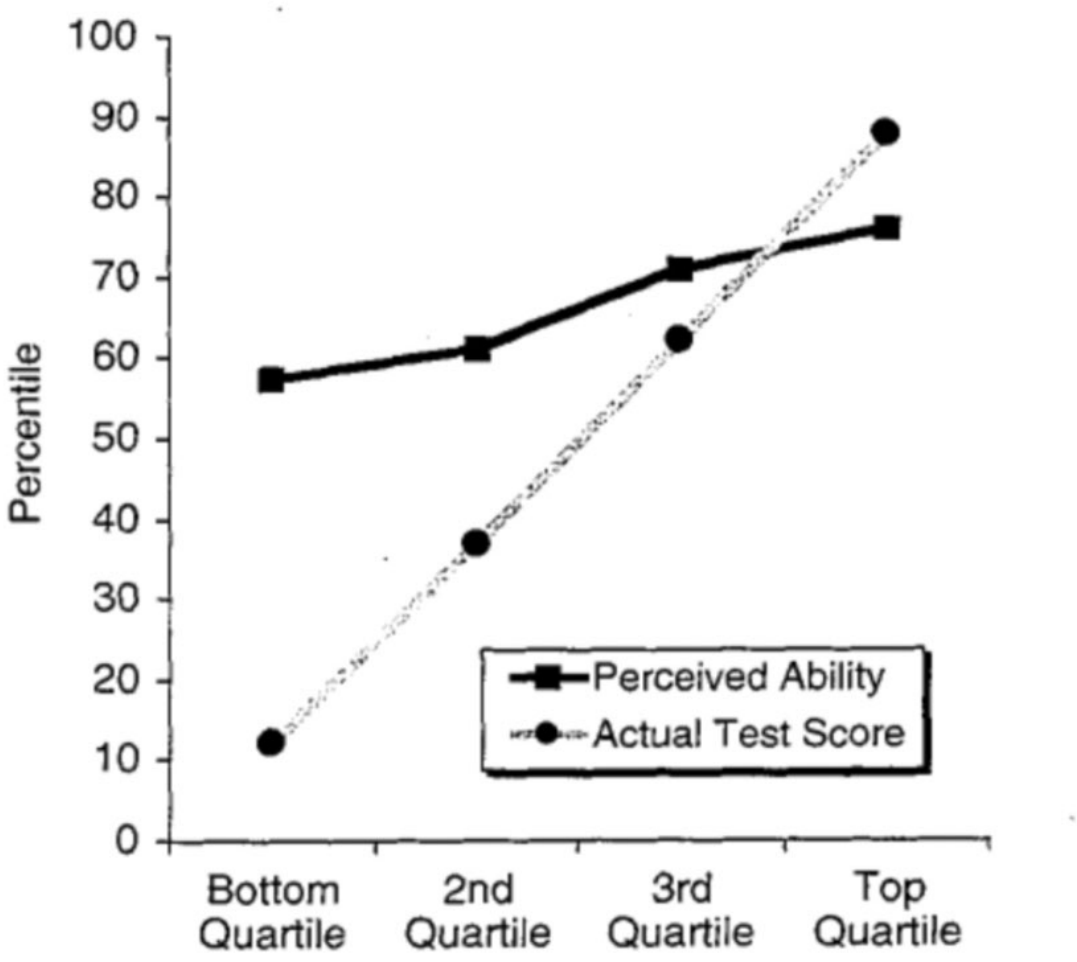

Fig. 2: Deze grafiek toont de oorspronkelijke bevindingen van Dunning en Kruger. Berg Domheid en de Vallei der Wanhoop zijn opvallend afwezig.

Een groot probleem met het vertrouwen op ervaring is dat onze informatie-input verwrongen is. Overmoed kan spectaculair memorabele resultaten opleveren, terwijl gebrek aan vertrouwen zelden aandacht van welke aard dan ook genereert. Iedereen heeft wel eens gehoord van Bob, de open water duiker die met een omgebouwde brandblusser als duikfles een grot inzwom. Hij haalde internationale krantenkoppen toen zijn lichaam werd gevonden. Niemand heeft ooit gehoord van zijn klasgenoot Alice, die haar capaciteiten zodanig onderschatte dat ze (helaas) nooit meer ging duiken. De Bobs van deze wereld worden onderdeel van wat wij onze ervaring noemen, terwijl de Alices snel vergeten worden.

Als iemand die deze gedachten al een tijdje koestert, was ik opgewonden toen ik hoorde dat niet alleen het algemene begrip van Dunning-Kruger, maar ook hun kernclaim onder aanzienlijk vuur is komen te liggen binnen de wetenschappelijke gemeenschap,[2] leidend tot een hoofdartikel in maart 2022 in The Psychologist, met een weerwoord door David Dunning in het volgende nummer. Kritiek concentreert zich rond de stelling dat het effect dat Dunning en Kruger vonden geen kenmerk is van de menselijke psychologie, maar in plaats daarvan een statistisch hulpmiddel dat onbedoeld is gecreëerd door de manier waarop de onderzoekers hun experiment hebben opgezet en hun gegevens hebben geëvalueerd.

Een relatief toegankelijke (en elegante) versie van deze kritiek werd gepubliceerd door de Canadese econoom Blair Fix in een blogpost getiteld "The Dunning-Kruger Effect is Autocorrelation" (april 2022).

"Het Dunning-Kruger-effect komt ook naar boven uit data waaruit dat niet zou moeten. Als je bijvoorbeeld zorgvuldig willekeurige gegevens bewerkt, zodat ze geen Dunning-Kruger-effect bevatten, zul je het effect nog steeds zien. De reden blijkt beschamend simpel te zijn: het Dunning-Kruger-effect heeft niets met de menselijke psychologie te maken. Het is een statistisch hulpmiddel – een verbluffend voorbeeld van autocorrelatie.

[…] De regel met het label 'werkelijke testscore' geeft het gemiddelde percentiel van de testscore van elk kwartiel weer (een mondvol, ik weet het). Alles lijkt in orde, totdat we ons realiseren dat Dunning en Kruger in wezen de testscore (*x*) tegen zichzelf plotten. "

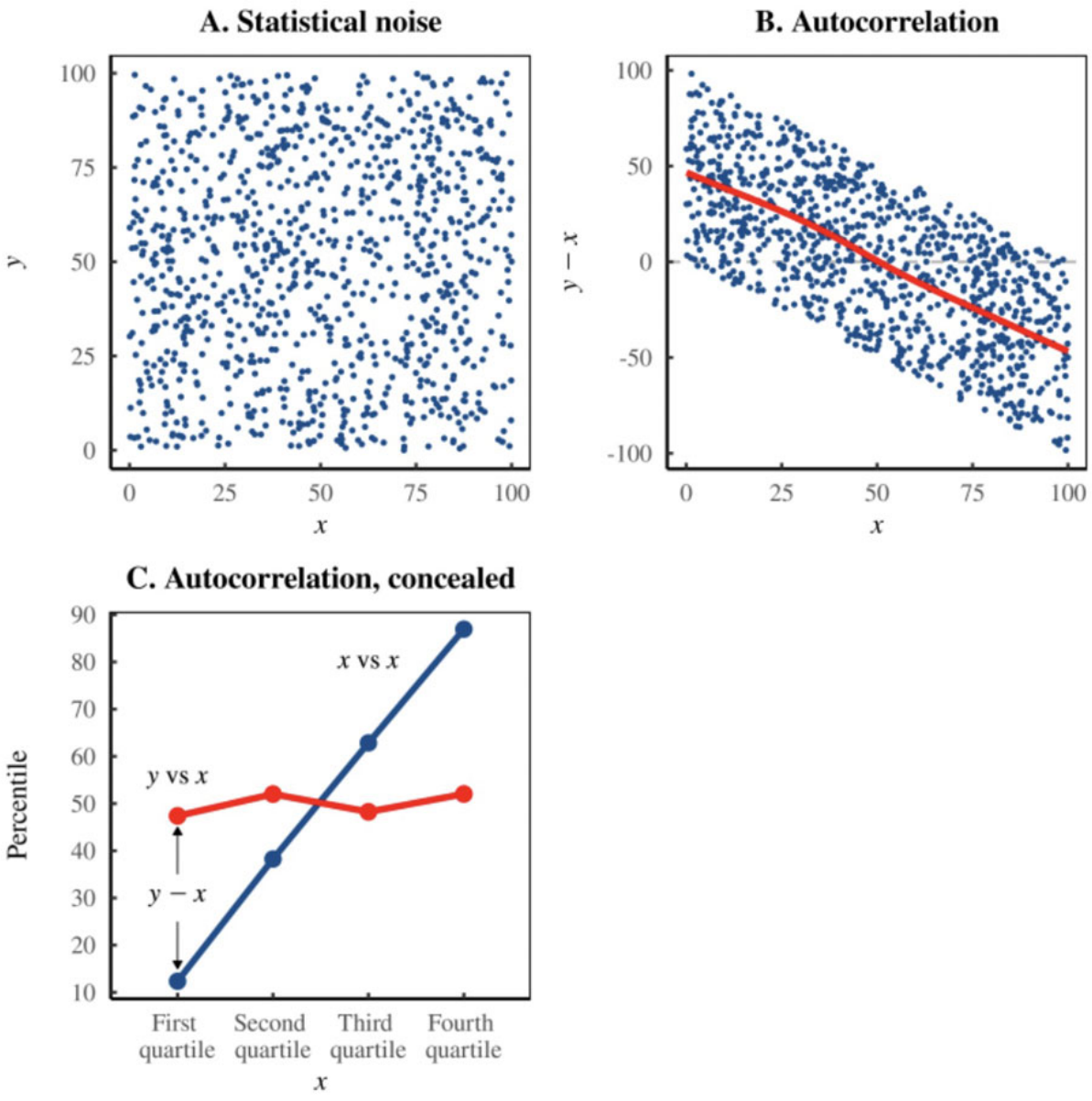

Fig. 3: een wolk van willekeurige gegevens en dezelfde gegevens na het toepassen van aurocorrelatie. Bron: "The Dunning-Kruger Effect is Autocorrelation"

Mijn aanvankelijke opwinding werd prompt gevolgd door een ontnuchterend besef: ik begrijp genoeg wiskunde om het argument overtuigend te vinden, maar niet genoeg om het te verifiëren. Het feit dat de kritiek van Fix en anderen iets bevestigde dat ik *toch al wilde geloven*, hielp ook niet: het zou precies hetzelfde soort cognitieve valstrik kunnen zijn die mensen ertoe brengt de vervormde, meme-ified versie

van Dunning-Kruger's beweringen te accepteren. Wat als ik, door meer te gaan geloven dat het Dunning-Kruger-effect niet echt is, mijn onwetende zelf recht op de top van Berg Domheid zou drijven?

Voor hulp nam ik contact op met een expert, Dr. Stephan Boes, een hoge ambtenaar bij het Bureau voor de Statistiek van Noordrijn-Westfalen in Duitsland[3,] die de kritiek van Fix bevestigde: "De autocorrelatie is er zeker. Ik kan niet precies zeggen hoeveel van het effect het verklaart zonder de gegevens te bekijken, maar het lijkt me heel duidelijk. Er is echter nog een ander probleem verderop in de lijn: deelnemers aan het experiment werd niet echt gevraagd hoe competent ze denken dat ze zijn. Ze werden gevraagd om te voorspellen hoe goed ze het zouden doen *in vergelijking met andere deelnemers*. Daar zijn twee problemen mee: een daarvan is dat om deze voorspelling te doen, zouden deelnemers het vaardigheidsniveau van anderen in de test moeten kennen. Een andere is dat een competitieve ranglijst niet erg geschikt is voor het beschrijven van de verdeling van resultaten voor de meeste taken in de echte wereld, waarbij je meestal een paar mensen hebt die het consequent slecht doen, een paar mensen die consequent uitblinken en een meerderheid met middelmatige resultaten die het misschien beter doen dan hun leeftijdsgenoten in de ene test en slechter in een andere. De manier waarop Dunning en Kruger hun data presenteren, houdt daar helemaal geen rekening mee."

Dit sluit aan bij een andere kritiek op het Dunning-Kruger-effect: in een studie gepubliceerd in 2020, vinden auteurs Gilles E. Gignac en Marcin Zajenkowski dat het *bovengemiddelde syndroom* (ook bekend als *illusoire superioriteit*) een betere verklaring biedt voor de discrepanties tussen voorspelde en werkelijke prestaties ten opzichte van anderen dan die Dunning en Kruger vonden. Illusoire superioriteit beschrijft de observatie dat een meerderheid van de mensen zichzelf slimmer, competenter, betere chauffeurs etc. vindt dan de gemiddelde persoon (wat onmogelijk is; 50% ligt per definitie onder het gemiddelde).

In het licht van deze informatie lijkt het idee om Dunning-Kruger toe te passen in de context van duiken twijfelachtig. Om te beginnen is het beschrijven van het vermogen van duikers in termen van een competitieve ranglijst nutteloos. Het maakt niet uit of je in het bovenste of onderste kwartiel van je Advanced Open Water-klas zat. Waar het om gaat, is dat je vaardigheden voldoende zijn voor de duiken die je maakt – absoluut vermogen, niet relatief. En zelfs als we dat allemaal zouden negeren en Dunning-Kruger zonder meer zouden accepteren, zijn er andere menselijke factoren die een rol spelen: in een PADI-seminar over risicobeheer dat ik ooit bijwoonde, benadrukte de docent dat de meeste duikongevallen zich niet tijdens de opleiding voordoen net uit de opleidingsboot gestapte instructeurs die denken dat ze alles weten. Ongelukken komen vaker voor bij ervaren instructeurs die zelfgenoegzaam worden.

Als we dit allemaal hebben geabsorbeerd, wat moeten we dan doen als onze instructeur of buddy terloops een verwijzing naar Dunning-Kruger of Berg Domheid in de klas of bij een biertje laat vallen? We zouden op kunnen springen en een maniakale tirade kunnen houden over hoe het Dunning-Kruger-effect niet is wat ze denken dat het is, en hoe we in Alert Diver lezen dat het effect misschien niet eens bestaat, en zelfs als het dat wel zou doen, hoe het waarschijnlijk niet van toepassing zou zijn op duiken.

Tenzij je echter vastbesloten bent om de rest van de avond te debatteren over de methodologie van kwantitatieve psychologische studies, regressie naar het gemiddelde en statistische hulpmiddlen gecreëerd door *het* in kaart brengen van *x* versus *(x-y)* wanneer *x* en *y* hetzelfde begrensde waardebereik hebben, zou een beter alternatief zijn om vermeldingen van Dunning-Kruger niet letterlijk maar figuurlijk te nemen: als een culturele code, de korte versie van een waarschuwend verhaal om ons te waarschuwen voor het onderschatten van de moeilijkheid van een taak die we gaan uitproberen. Zelfs als het Dunning-Kruger-effect niet echt is, is overmoed dat zeker wel, in het duiken en elders, en het is meestal gevaarlijker dan het tegenovergestelde. Dat moeten we altijd voor ogen houden.

Moge de hellingen van je leercurve glad en vol plezier zijn.

Voetnoten

[1] Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

[2] Nuhfer, Edward, Christopher Cogan, Steven Fleisher, Eric Gaze, and Karl Wirth. "Random Number Simulations Reveal How Random Noise Affects the Measurements and Graphical Portrayals of Self-Assessed Competency." *Numeracy* 9, Iss. 1 (2016): Article 4. DOI: Random Number Simulations Reveal How Random Noise Affects the Measurements and Graphical Portrayals of Self-Assessed Competency

Gilles E. Gignac, Marcin Zajenkowski, "Het Dunning-Kruger effect is (vooral) en statistisch hulpmiddel: valid benaderingen van het testen van hypothesen met individuele verschildata." *Intelligence*, Nummer 80, 2020, 101449, ISSN 0160-2896, https://doi.org/10.1016/j.intell.2020.101449.

Robert D. McIntosh and Sergio Della Sala, "The persistent irony of the Dunning-Kruger Effect." *The Psychologist,* Journal of the British Psychological Society, vol. 35, March 2020, The persistent irony of the Dunning-Kruger Effect | BPS

David Dunning, "The Dunning-Kruger effect and its discontents." *The Psychologist,* Journal of the British Psychological Society, vol. 35, maart 2020, https://www.bps.org.uk/psychologist/dunning-kruger-effect-and-its-discontents

[3] De gezichtspunten die hier naar voren worden gebracht zijn persoonlijk en geven niet de mening van Dr. Boes' werkgever weer.

Over de auteur

Tim Blömekegeeft les in sport- en techduiken in Taiwan en de Filippijnen. Hij is een fervent grot-, wrak- en CCRduiker en daarnaast schrijvend redacteur en vertaler voor Alert Diver. Hij woont in Taipei, Taiwan. Je kunt hem op Instagram volgen via @timblmk.

Vertaler: Els Knaapen