

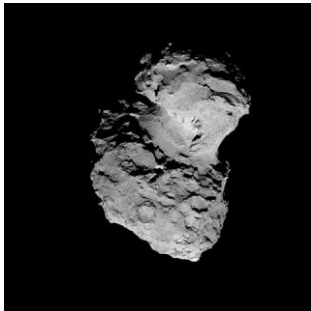


Het meten van kennis en vaardigheid: een lastige klus

Kennis en vaardigheid zijn abstract. Ze zijn lastiger te meten dan bijvoorbeeld de lengte van een tafel. Om toch iets te kunnen zeggen over het kennis- en vaardigheidsniveau van een persoon, gebruiken we een indirecte methode: we stellen een aantal vragen en bepalen het kennis- en vaardigheidsniveau aan de hand van de antwoorden. We nemen daarbij aan dat het aantal goede antwoorden op die vragen iets zegt over het kennis- en vaardigheidsniveau. Voor het gemak spreken we in dit artikel over vaardigheid in plaats van kennis- en vaardigheidsniveau. De vaardigheid geeft dus aan hoe goed je bent in een vak. Experts zijn het er niet over eens of je wel met één cijfer kunt weergeven hoe goed iemand in een vak is. Vaardigheid is namelijk ongrijpbaar en grillig. We gebruiken hier een metafoor om dit nog iets duidelijker te maken.

Een metafoor voor vaardigheid

Vaardigheid zou je kunnen vergelijken met een grillig gevormd stuk steen dat in de ruimte zweeft. Om te weten te komen hoe groot de steen is wordt een foto genomen¹:



Eerst moet bepaald worden hoe je de grootte definieert. Laten we zeggen dat het de grootste rechte afstand is die je kunt meten van de ene kant van de steen naar de andere kant van de steen. Dat zou op deze foto van boven naar beneden zijn. Om de grootte op de foto te 'vertalen' naar een werkelijke grootte heb informatie nodig over de schaal. Is 1 cm op de foto in werkelijkheid 100 meter, 10 kilometer of 1000 kilometer? We noemen dat de schaalfactor. Om in de metafoor te blijven: het nemen van een foto is vergelijkbaar met het afnemen van een examen, de schaalfactor kun je vergelijken met de moeilijkheid van de vragen en het aantal cm op de foto is vergelijkbaar met de score op het examen. De score op een examen (het aantal cm op de foto) zegt niet zo veel over de vaardigheid (de grootte van de steen) als je de moeilijkheid van de vragen (schaalfactor) niet kent. Bijvoorbeeld: als je 5 (van de 10) vragen goed hebt bij heel erg makkelijke vragen is je vaardigheid kleiner dan wanneer je 5 (van de 10) vragen goed hebt bij heel erg moeilijke vragen.

Maar er is nog een ander probleem om de grootte van de steen te meten. Want, wat nu als het brokstuk een kwartslag kantelt? Dan zal een volgende foto een andere lengte opleveren. Om een goed beeld van het brokstuk te krijgen moet je het brokstuk dus eigenlijk van heel veel verschillende kanten fotograferen. Wat nu duidelijk naar voren komt is dat de vaardigheid die door het examen wordt gemeten niet altijd hetzelfde is.

¹ Door ESA/Rosetta/NAV CAM, CC BY-SA IGO 3.0, CC BY-SA 3.0-igo, <https://commons.wikimedia.org/w/index.php?curid=36605324>

Met andere woorden: om een goed beeld van iemands vaardigheid te krijgen moet je moet eigenlijk heel veel examens afnemen. Dan pas kun je iets zeggen over iemands vaardigheid. Maar wat is dat dan, 'iemands vaardigheid'?

Een dik boek

Wat er op een examen gevraagd mag worden staat uitgebreid beschreven in een document, de syllabus. Elk vak op elk niveau heeft zijn eigen syllabus. We beperken ons nu even tot het vak havo biologie. Stel je eens voor dat er een boek zou bestaan waarin *alle* toegestane havo biologie-examenvragen staan. Dat zijn er waarschijnlijk wel meer dan tienduizend. Een heel erg dik boek dus. We gaan er daarbij van uit dat iedereen het erover eens is dat deze vragen passend en compleet zijn om biologie op havoniveau te toetsen. Stel nu dat een leerling al deze vragen in korte tijd kan beantwoorden. Dan geeft de score die hij op al deze vragen heeft gehaald de beste meting van zijn vaardigheid in biologie op havoniveau. We noemen dit zijn werkelijke vaardigheid. Helaas bestaat het dikke boek niet, dus kunnen we iemands werkelijke vaardigheid nooit te weten komen.

Terug naar ons voorbeeld: het examen havo biologie bevat 'slechts' 40-45 vragen. Dat is maar een klein gedeelte van alle duizenden vragen die er in het dikke boek staan. Nu is het examen elk jaar op een soortgelijke manier opgebouwd met zowel makkelijke als moeilijke vragen, een aantal vragen over ecologie, celbiologie, genetica etc. Maar dat neemt niet weg dat het ene examen de leerling mogelijk beter ligt dan het andere examen. Het patroon van de vaardigheid van elke leerling is uniek (en grillig, zie de metafoor). Een vraag die voor heel veel leerlingen moeilijk is, is voor één leerling misschien erg makkelijk omdat er bij hem in de familie iets is gebeurd waardoor deze leerling dat specifieke stukje biologie goed heeft onthouden. Als er (bij toeval) veel van dit soort vragen in het examen terecht komen dan heeft de leerling een voordeel. We zeggen dan dat het examen hem goed ligt. In de rangschikking op score voor dit examen zal deze leerling beter uit de bus komen dan in de rangschikking op de werkelijke vaardigheid. Het kan natuurlijk ook net andersom zijn. Dan zitten er veel vragen in het examen over zaken die de leerling net niet weet. Dan komt hij minder goed uit de bus dan op basis van zijn werkelijke vaardigheid verwacht kon worden. Een examen geeft dus een schatting van de werkelijke vaardigheid van de leerling, maar deze schatting kan een onder- of een overschatting zijn. Doordat het examen veel verschillende onderwerpen bevat wordt de kans vergroot dat een aantal voordelen en nadelen tegen elkaar opwegen. Op deze manier wordt er dus voor gezorgd dat de schatting zo nauwkeurig mogelijk is.

Epiloog

Uit bovenstaande beschrijvingen van de metafoor en het dikke boek komt naar voren dat de score op een toets met de nodige voorzichtigheid moet worden bekeken. Het meten van vaardigheid is (helaas) minder eenduidig dan het meten van de lengte van een tafel. Dit roept wel weer vragen op: wat is de nauwkeurigheid van een examencijfer? En: hoe zorgen we ervoor dat leerlingen niet onterecht zakken voor hun diploma? Hier zal in andere artikelen aandacht aan worden besteed.