

# De ene deskundige is de andere niet

Hoe de rechter empirisch gefundeerd bewijs kan waarderen

Raimond Giard & Harald Merckelbach<sup>1</sup>

Om de feiten te kunnen ophelderen, zijn rechters regelmatig aangewezen op deskundigen. Maar hoe goed zijn die deskundigen en wat als zij elkaar tegenspreken? Bij het beantwoorden van deze vragen worden rechters gehinderd door de kennisparadox. Die komt erop neer dat rechters in voorkomende gevallen zich moeten verlaten op experts, maar deze experts niet de maat kunnen nemen juist vanwege hun gebrek aan specialistische kennis. De kennisparadox wordt vaak op een nogal fatalistische manier gepresenteerd: zo is het en daar zullen rechters mee moeten leren leven. Auteurs hebben daar een andere opvatting over. Zij denken dat rechters – als zij snappen hoe kennisverwerving plaatsvindt en hoe de hiërarchie van solide kennis er uitziet – de deskundige op een gerichte manier kunnen bevragen zodat zij tot een beargumenteerde weging kunnen komen van het deskundigenbericht.

## 1. Het waarderen van deskundigenbewijs

Bij iedere civiel-, bestuurs- of strafrechtelijke procedure zal de rechter het ingebrachte bewijsmateriaal dienen te waarderen en deze waardering ligt vervolgens ten grondslag aan zijn beslissing.<sup>2</sup> In zijn algemeenheid heeft de wetgever gekozen voor een vrij bewijsstelsel.<sup>3</sup> De rechter zal al het ingebrachte bewijsmateriaal naar eigen inzicht moeten waarderen. Achter de vrije bewijsleer gaat de gedachte schuil dat het ondoenlijk is om voor een breed scala aan problemen een alomvattend systeem van specifieke regels voor de bewijswaardering en -beslissing op te stellen.

In veel gevallen zal de rechter weinig moeite hebben met het beoordelen van bewijsmateriaal, omdat het probleem betrekkelijk eenvoudig is en zowel formeel- als materieelrechtelijk duidelijk ingekaderd. Zodra echter een bewijsmiddel bestaat uit een deskundigenbericht of uit wetenschappelijke publicaties waarop partijen zich beroepen, rijst de vraag hoe betrouwbaar de aangereikte kennis is. Dan gaat het om de epistemische dimensie van de bewijswaardering.<sup>4</sup>

Enkele voorbeelden van kwesties waarin de rechter niet om de waardering van deskundigenbewijs heen kan. Een schilder overlijdt aan blaaskanker: kwam dat door de beroepsmatige blootstelling aan kankerverwekkende stoffen, zoals een medisch expert beweert?<sup>5</sup> Een horecagelegenheid brandt af: was dat door nalatigheid van het personeel, zoals een branddeskundige betoogt?<sup>6</sup> Huizen

lopen schade op door aardbevingen: kwam dat volledig door gasexploitatie, zoals een bouwexpert meent, en welke schade werd daarmee aangericht?<sup>7</sup> Een gemeentebestuur wordt verweten in een anonieme brief ernstige beschuldigingen te hebben geuit over een aantal collega's. Levert schriftkundig onderzoek voldoende betrouwbaar bewijs voor het vaststellen van het auteurschap?<sup>8</sup>

Iedere procedure kent een debat tussen procespartijen, die tot verschillende conclusies komen en zich daarbij nogal eens beroepen op door deskundigen aangedragen kennis. Het bewijsmateriaal van experts is steeds *speciale* kennis, zoveel is duidelijk. Maar deze kennis kan variëren van anekdotische observaties tot aan langs

### Auteurs

1. Prof. dr. R.W.M. Giard is arts, jurist en methodoloog. Hij is emeritus hoogleraar methodologie en aansprakelijkheid, Erasmus School of Law, EUR. Prof. dr. H.L.G.J. Merckelbach is hoogleraar rechtspsychologie, Universiteit van Maastricht.

### Noten

2. Omwille van de leesbaarheid schrijven we 'zijn', maar we bedoelen 'haar of zijn'.  
3. Zie voor het civiele recht art. 152 lid 2 Rv

en voor het strafrecht met name art.339 en 340 Sv.

4. Giard 2016a.

5. HR 7 juni 2013,

ECLI:NL:HR:2013:BZ1721.

6. HR 5 december 2003,

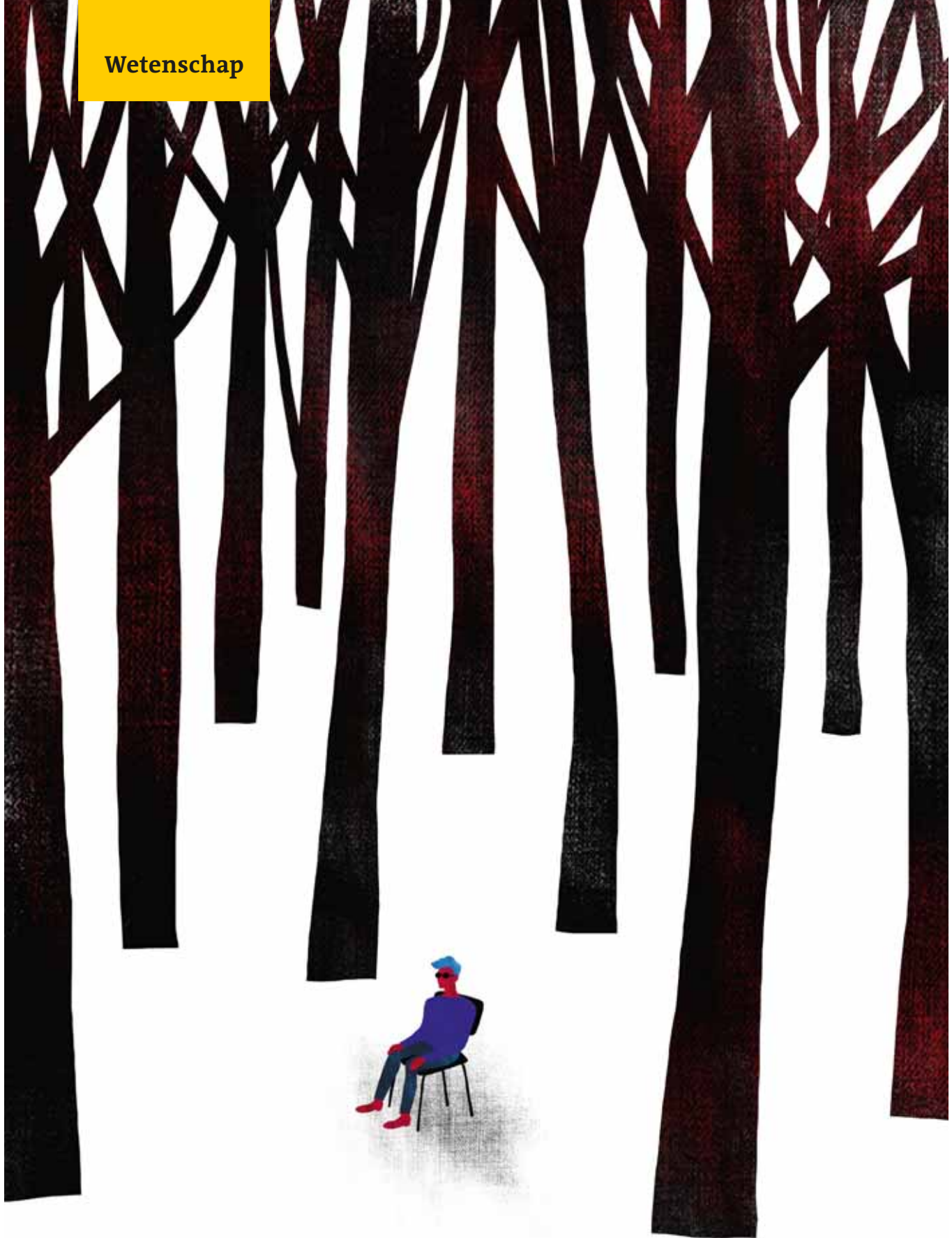
ECLI:NL:HR:2003:AN8478.

7. Rb. Noord-Nederland, 1 maart 2017,

ECLI:NL:RBNNE:2017:715.

8. Rb. Arnhem 26 mei 2011,

ECLI:NL:RBARN:2011:BQ6202.



De bomen en het bos © Fanatic Studio / Alamy

streng-experimentele weg getoetste inzichten. Hoe kunnen rechters bedachtzaam om gaan met het deskundigenbewijs?<sup>9</sup> In wat volgt schetsen we een richting waarin het antwoord gevonden kan worden.

## 2. Eerst snappen dan oordelen

Een rechter kan pas goed oordelen over een kwestie wanneer hij begrijpt hoe het zo heeft kunnen gebeuren.<sup>10</sup>

Dat begrijpen berust op het ophelderen van het causale netwerk dat tot de gebeurtenis heeft geleid: dat is 'verklaren'.<sup>11</sup> Verklaringen staan dus in iedere procedure centraal, zijn causaal van aard en in essentie steeds antwoorden op waarom- en waardoor-vragen: waarom kreeg deze schilder blaaskanker? Waarom brandde deze horecagelegenheid af? Waardoor werd deze aardbeving veroorzaakt en waarom liepen de huizen daardoor schade op?

Waarom denken we te weten dat ambtenaar X deze brief heeft geschreven?

Rechtvaardigen hebben begrijpelijkerwijze de neiging zich op de rechtsvraag te concentreren. Maar als aan een solide rechterlijk oordeel een juiste verklaring ten grondslag moet liggen, ontkomen rechters er niet aan zich een opinie te vormen over de kwaliteit van de door de deskundigen naar voren geschoven verklaringen. De weg waarlangs naar de best mogelijke verklaring wordt gezocht, is die van abductief redeneren: er bestaan verschillende concurrerende duidingen en daaruit zal een beredeneerde keuze moeten worden gemaakt.<sup>12</sup> Er zijn allerlei criteria geformuleerd voor het gewicht dat aan een verklaring kan worden toegekend.<sup>13</sup> Voor het civiele recht noemt De Bock bijvoorbeeld de volgende ijkpunten voor het waarderen van deskundigenbewijzen: relevantie, consistentie, kwaliteit en coherentie.<sup>14</sup> Daarnaast benadrukt zij dat de reikwijdte van de bewijsmiddelen tegen de achtergrond van de door de wederpartij aangedragen alternatieve hypothesen moet worden bezien.<sup>15</sup>

Zulke criteria zijn te globaal. Zo kunnen eenvoudige en heldere formuleringen een deugd zijn, maar op zichzelf vormen ze nog geen betrouwbaar waarheidscriterium: stereotypen zijn bijvoorbeeld simpel, maar daarom nog niet waar.<sup>16</sup> Coherentie maakt een argumentatie misschien wel geloofwaardig, maar daardoor niet per sé juist: de coherentiebias.<sup>17</sup> De reputatie van een deskundige garandeert evenmin juiste uitkomsten vanwege het gevaar van autoriteitsbias.<sup>18</sup> Wat betreft de reikwijdte van de bewijsmiddelen is er het risico van de illusie van een explanatoire diepgang.<sup>19</sup> Dit betekent dat een ogenschijnlijke plausibele verklaring te oppervlakkig blijft en de kern van het probleem niet wezenlijk verheldert.

Al deze criteria vereisen steeds nadere precisering en zijn daarom onwerkbaar. Naar onze mening is de rechter meer gebaat bij een methodologische invalshoek die te ontleen is aan de empirische wetenschappen.<sup>20</sup> Centraal daarin staat niet *wie* de deskundige is en zelfs niet *zozeer wat* hij zegt, maar eerder *hoe* de kennis die hij ter verklaring naar voren schuift is verkregen.

### 3. Hoe kennis wordt verworven

Deskundigen worden wel omschreven als de planken over het moeras van de rechterlijke onwetendheid.<sup>21</sup> Maar de stevigheid van die planken varieert al naar gelang het type kennis dat de deskundige inbrengt. In de eerste plaats kan de deskundige verwijzen naar *casuïstiek*, zoals een gevalsbeschrijving van een schilder met blaaskanker.<sup>22</sup> Casuïstiek levert het bestaansbewijs voor iets – ‘het is ooit voorgekomen dat (een schilder blaaskanker kreeg)’ – maar zegt niets over causaliteit. Daarom levert casuïstiek wankelende kennis op. Die kennis wordt al wat robuuster als de deskundige verwijst naar een hele reeks van soortgelijke gevallen die in de vakliteratuur is gepubliceerd. Zo’n reeks

## Deskundigen worden wel omschreven als de planken over het moeras van de rechterlijke onwetendheid

adstrueert dat ‘het vaker voorkomt dat (schilders blaaskanker krijgen)’. Maar causale conclusies zijn hieruit nog steeds niet te trekken. Hooguit valt te vermoeden dat – om bij het voorbeeld te blijven – blootstelling aan toxische verfstoffen blaaskanker veroorzaakt.

Nog steviger is de plank van specialistische kennis wanneer de deskundige verwijst naar een verzameling van soortgelijke casus die gerekruteerd werden uit een welomschreven populatie en binnen een duidelijk afgebakende tijdsperiode (‘van alle beroepsschilders in Zuid-Nederland tussen 2000 en 2015 kreeg x% blaaskanker’). Dat is *observatieel* onderzoek. Zo kan bijvoorbeeld worden vastgesteld hoe vaak blaaskanker onder schilders voorkomt. Dat type onderzoek is echter louter descriptief.

Als het vermoeden bestaat dat blootstelling aan schadelijke stoffen tijdens het schilderen een oorzakelijke factor is, kan een systematische vergelijking inzicht bieden. Gedurende een bepaalde periode wordt geobserveerd hoe vaak blaaskanker voorkomt bij schilders en bij een controlegroep van mensen die niet als schilder werken. Blijkt de kans inderdaad groter bij schilders? Dat is *analytisch* oftewel *quasi-experimenteel* onderzoek en het levert kennis op die dichter in de buurt komt van een causale verklaring, maar zeker niet perfect is. Want als geconstateerd wordt dat bij schilders blaaskanker vaker optreedt, is dat voorshands geen bewijs van een causale relatie tussen blootstelling aan schadelijke stoffen en blaaskanker; er is alleen een correlatie tussen deze twee variabelen en die sluit op zichzelf niet uit dat er nog een derde onderliggende oorzaak is. Bij wijze van voorbeeld: misschien roken schilders wel vaker en meer dan niet-schilders en is roken verantwoordelijk voor de hogere frequentie van blaaskanker onder schilders.<sup>23</sup>

Voor het ophelderen van causale relaties is *experimenteel* onderzoek een superieure benadering. Een voorbeeld is onderzoek naar de vraag of juist de combinatie van alcohol en een antidepressivum aanzet tot agressief gedrag. Nee, dat blijkt niet het geval als je in een experimentele opzet sommigen proefpersonen drank geeft, anderen een antidepressivum, weer anderen een combinatie van de twee en nog weer anderen helemaal niets en ze dan provokeert. De onderzoekers beoordelen de eindtoestand, in dit geval de agressie van de proefpersoon. Maar ze mogen geen weet hebben van de behandeling die de proefpersoon ten deel

9. Zie ook de dissertatie van Hoving (2017)

daarover voor het strafrecht: hij vraagt zich af hoe de omgang met het deskundigenbewijs in het strafproces kan worden verbeterd.

10. Keil 2006.

11. Strevens 2013.

12. Lipton 2004; Pardo & Allen 2008.

13. Zemla et al. 2017.

14. De Bock 2011 § 6.6 (p. 259-275) en § 7.9 (p. 322-325).

15. De Bock 2011, p. 218.

16. We doelen hier op wat in jargon de *representativeness bias* of *representative-*

*ness heuristic* heet (ook: *pigeonholing*):

zie Gigerenzer & Gaissmaier 2011.

17. Amaya 2013.

18. Schröder 2016.

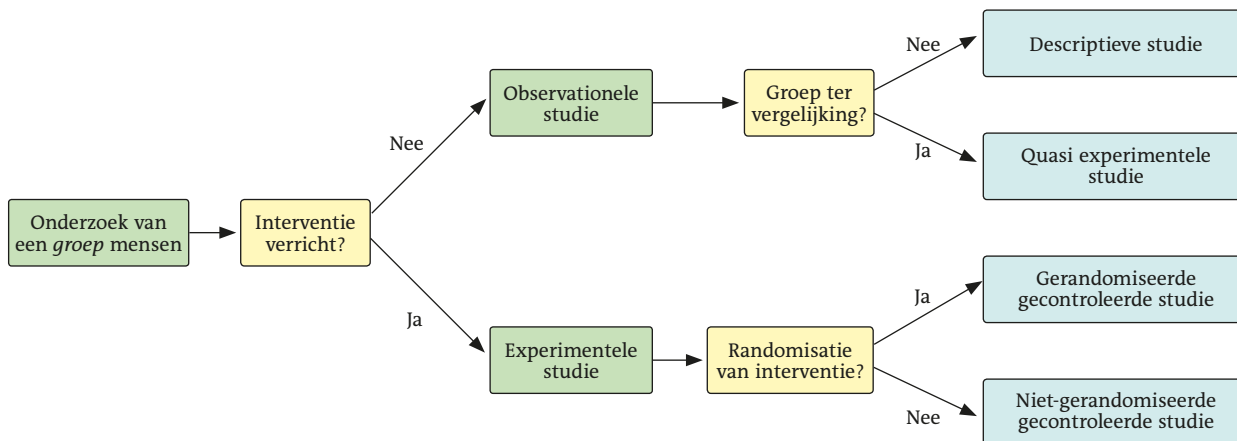
19. In de literatuur aangeduid als *de illusion of explanatory depth* (IOED). Zie daarover Keil 2006 en Fischer & Keil 2013.

20. Giard 2016a, hfdst. 3 (p. 43 e.v.).

21. Ettekoven 2016.

22. We volgen hier verder de taxonomie van Grimes & Schulz 2002.

23. Maccoun 1998.



Figuur 1: Verschillen in onderzoeksmethodologie (bewerkt naar Grimes & Schulz, 2002).

viel: ze zijn 'geblindeerd'. Proefpersonen die alcohol en een antidepressivum hebben gehad reageren daar dus niet met buitensporige agressie op; eerder het tegendeel bleek het geval. Dat is harde kennis over (de afwezigheid van) een causale samenhang die strafrechtelijke relevantie kan hebben als ze door een deskundige naar voren wordt gebracht.<sup>24</sup> Experimenteel onderzoek biedt dus een stevige plank van specialistische kennis. Of bijvoorbeeld *mediation* een betere oplossing is voor geschilbeslechting dan een procedure in rechte, laat zich alleen maar goed met dergelijk experimenteel onderzoek beantwoorden.<sup>25</sup>

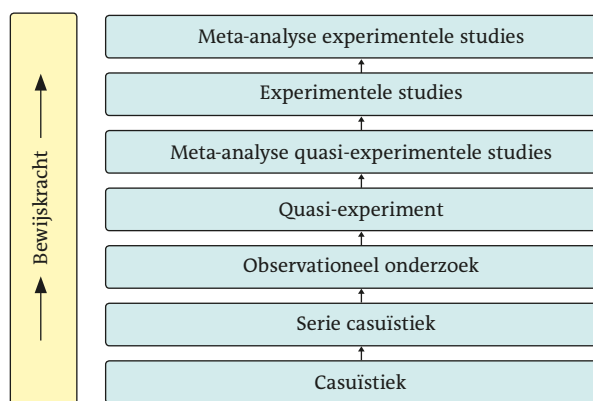
Om redenen die met ethiek te maken hebben, zijn niet alle kwesties experimenteel benaderbaar. De kwestie van toxische blootstelling en blaaskanker is niet op te helderen door een groep van proefpersonen wel en een andere niet gedurende lange tijd bloot te stellen aan verf. Wetenschappers zijn dan aangewezen op het eerdergenoemde quasi-experiment, dat wil zeggen de systematische vergelijking tussen twee bestaande groepen – denk aan schilders en metselaars – die over lange tijd met elkaar worden vergeleken voor wat betreft blaaskanker.

Of het nu om quasi-experimenteel of experimenteel onderzoek gaat, als randvoorwaarde geldt wel dat de naast elkaar gestelde groepen ook daadwerkelijk enkel en uitsluitend verschillen voor wat betreft de factor die wordt onderzocht (schilderen, alcohol/antidepressiva, mediatioon). Dat heet de *ceteris paribus-doctrine*.<sup>26</sup> Door het lot te laten toewijzen – in jargon: randomisatie – of iemand wel of niet zal worden blootgesteld aan een interventie is daarvoor de beste garantie. Wanneer randomisatie om ethische redenen niet mogelijk is, kan een wetenschapper statistische technieken aanwenden – in jargon covariatie-analyse – om de te vergelijken groepen zoveel als mogelijk gelijk te trekken. Zo worden overige factoren die een oorzakelijke (bij)rol kunnen spelen – bijvoorbeeld roken in het geval van schilders – door het lot of met een statistische techniek gelijk over de verschillende onderzoeksgroepen verdeeld en kan het optreden van een verschijnsel dan aan één oorzakelijke factor worden toegeschreven, mits toeval als verklaring voldoende valt uit te sluiten.<sup>27</sup> De meest degelijke plank van specialistische kennis is als de deskundige kan verwijzen naar onderzoek waarin de resultaten van verschillende wetenschappelijke studies over hetzelfde onderwerp zijn samengevoegd. Dan hebben

we het over een synthese door middel van een zogenaamde meta-analyse.<sup>28</sup> Daarbij wordt in de wetenschappelijke literatuur uitputtend gezocht naar publicaties over een bepaald onderwerp. Aan de hand van strenge selectiecriteria, met name over de bij het onderzoek gehanteerde methodologie, worden de beste studies geselecteerd. Na deze schifting worden de uitkomsten van de geselecteerde studies samengevoegd en geanalyseerd. Door de grotere aantallen onderzochte personen krijgen de resultaten meer zeggingskracht (in jargon: *power*). Zo laten meta-analyses zien dat er inderdaad een stevige samenhang bestaat tussen blootstelling aan giftige stoffen en blaaskanker, zelfs als de invloed van roken wordt verdisconteerd.<sup>29</sup> Figuur 1 geeft een samenvatting van hoe specialistische kennis kan worden vergaard.

#### 4. Hiërarchie in bewijskracht

Ons punt is dat er een hiërarchie van specialistische kennis bestaat. Zo zijn de resultaten van experimenteel onderzoek robuuster dan die van een casuïstische mededeling. Voorts: de uitkomsten van welk type onderzoek dan ook zijn steeds omgeven met een onzekerheidsmarge. Ze zijn provisorisch en kunnen om die reden nooit aanspraak maken op de definitieve waarheid. Wetenschappelijk gezien bestaat er niet zoiets als waterdicht bewijs. De uitkomsten van onderzoek kunnen daarom het best in bewijskracht worden uitgedrukt. Figuur 2 geeft daarvan een hiërarchisch overzicht.



Figuur 2: Hiërarchische indeling van typen empirisch onderzoek

Met dit schema valt het onderzoek waarop een partij of deskundige zich beroept beter te plaatsen. Wanneer een expert *ex cathedra* een mening naar voren brengt zonder die deugdelijk wetenschappelijk te verankeren, valt dat buiten dit schema en moet het als onbruikbaar worden beoordeeld. In alle andere gevallen kan de specialistische kennis waarop de deskundige zich beroept langs de meetlat van figuur 2 worden gelegd.

## De uitkomsten van onderzoek kunnen het best in bewijskracht worden uitgedrukt

Dat is een verhelderende exercitie. Want dankzij internet is het met behulp van zoekmachines tegenwoordig makkelijk speuren naar wetenschappelijke literatuur.<sup>30</sup> Daarmee is een reusachtige hoeveelheid wetenschappelijke artikelen over een bepaald onderwerp voorhanden. Wie bijvoorbeeld met de zoekterm 'bladder cancer risk in painters' in *Google Scholar* naar vakpublicaties sondeert, treft maar liefst 16 400 wetenschappelijke artikelen aan. Binnen een dergelijke gigantische voorraad is beslist wel iets van eigen gading te vinden en zo'n gegevensbestand van publicaties kan daarom dienen als een grabbelton met voor elk wat wils. Er zijn genoeg casuïstische mededelingen en descriptieve studies van blaaskanker bij schilders. In dat genre zullen er best auteurs te vinden zijn die beweren dat het causale vehikel van blaaskanker bij schilders roken is, net als er auteurs te vinden zullen zijn die beweren dat blootstelling aan toxische verfstoffen de beslissende factor is. Maar de ene deskundige is de andere niet en om erachter te komen op welk soort kennis de deskundige zich beroept, kan de rechter met figuur 2 in de hand deskundigen bevragen over hun ankers. Baseren de deskundigen hun verklaringen op casuïstiek, op quasi-experimenten, op experimenten of op meta-analyses. Dat maakt nogal wat uit.<sup>31</sup>

De rangorde van figuur 2 zegt alleen iets over de mogelijke bewijskracht van het *type* onderzoek. Voor sommige vraagstellingen zijn bepaalde typen onderzoek niet uitvoerbaar en dus niet beschikbaar. Experimenteel onderzoek om het verband tussen schilderen en blaas-

kanker te toetsen is, zoals gezegd, om ethische redenen onmogelijk. Dan is een deskundige aangewezen op quasi-experimentele studies en aangezien blaaskanker niet zo vaak voorkomt, heeft een goede deskundige het liefste een meta-analyse om over de best mogelijke schattingen van het verband tussen toxische blootstelling en blaaskanker te beschikken.<sup>32</sup> De betrouwbaarheid van schriftkundig onderzoek is echter wel langs experimentele weg te onderzoeken en dat is ook gebeurd.<sup>33</sup>

De aard van het probleem dat de inzet van een procedure vormt, bepaalt dus welk type onderzoek idealiter de beste informatie geeft. De deskundige die de aandacht wil vestigen op de foutenmarge van handschriftkundigen kan volstaan met de verwijzing naar een paar goed uitgevoerde experimenten waarin die foutenmarge schattingen werd vastgesteld. De deskundige die wil beweren dat er een hard, causaal verband bestaat tussen toxische blootstelling en blaaskanker in schilders legt idealiter een omvangrijke meta-analyse op tafel.

### 5. Toepasbaar?

Dat een deskundige zich beroept op vakliteratuur waarin casuïstiek, een experiment of zelfs een meta-analyse wordt beschreven en dat de rechter daar op geleide van figuur 2 een waardering aan kan geven, is één ding. Een andere en minstens zo'n belangrijke kwestie is of het onderzoek dat door de deskundige in stelling wordt gebracht relevantie bezit voor de zaak. De rechter die voor bewijswaarderingen en -beslissingen wil weten of door deskundigen terecht een beroep werd gedaan op inzichten uit wetenschappelijk onderzoek, zal daarvoor ijkpunten willen gebruiken. Het formuleren en toepassen van dergelijke ijkpunten is met name in de geneeskunde ontwikkeld en verfijnd.<sup>34</sup> Omdat er een duidelijke gelijkenis bestaat tussen juridisch en medisch redeneren, denken we dat rechters hun voordeel kunnen doen met de ijkpunten die aan de medische context zijn te ontleen.<sup>35</sup> We sommen enkele belangrijke ijkpunten op.

#### 1. Is de aard van het probleem helder?

Het gaat hier om het expliciteren van het probleem.<sup>36</sup> Dat lijkt een open deur, maar een rechtsvraag is wezenlijk anders dan een wetenschappelijke vraag, terwijl die laatste de gids is bij het zoeken naar de juiste verklaring. Bij het probleem van schilders met blaaskanker zal de rechter zich over werkgeversaansprakelijkheid moeten uitspreken. Maar dat kan pas als de deskundige onder verwijzing naar meta-analyses heeft laten zien dat er inderdaad een causaal verband is vastgesteld tussen het schildersberoep en het vaker krijgen van blaaskanker. Dáár ligt de essentie van het probleem – verklaart blootstelling aan verfstoffen afdoende het

24. McCloskey et al. 2009.

25. Giard 2006.

26. = onder overigens gelijke omstandigheden.

27. Ook het zoveel mogelijk uitsluiten van toeval is het domein van de statische toets.

Zie Finkelstein & Levin 2014, hfdst. 5.

28. Cooper 2015.

29. Guha et al. 2010.

30. Voorbeelden zijn algemene zoekprogramma's zoals Scirus, Google Scholar of Web of Science en meer vakgericht zoals PubMed (medisch), Psychnet (psychologie).

31. Zie Giard 2016b.

32. Zie als voorbeeld daarvan de meta-analyse van Guha et al. 2010.

33. Sita et al. 2002. Deze onderzoekers vonden bijvoorbeeld dat schriftkundigen

beter dan leken zijn in de vaststelling of een handtekening echt dan wel nagemaakt is (foutenpercentages van respectievelijk 3.4% en 19.3%). En toch is dat foutenpercentage voor schriftkundigen nog steeds van dien aard dat het uitspraken van het type 'deze handtekening is met een zekere mate grenzende waarschijnlijkheid authentiek' onmogelijk maakt.

34. Een internationale werkgroep heeft daarvoor het systeem van GRADE ontwikkeld: Grading of Recommendations Assessment, Development and Evaluation. Zie hun website [www.gradeworkinggroup.org](http://www.gradeworkinggroup.org) en Alonso-Coello et al. 2016a en 2016b.

35. Samuel 2014.

36. Zie De Haan & De Heer 2015.

## Elke toepassing van empirische kennis in een juridische procedure brengt steeds een *group-to-individual*-probleem met zich mee, want het gaat steeds om de materiële waarheid bij individuele kwesties

ontstaan van blaaskanker? – ook voor het leveren van bewijs. Het vaststellen van een eventuele werkgeversaansprakelijkheid, de rechtsvraag, komt pas daarna en wordt gegrond in de uitkomsten van de onderzoeksvraag.

### 2. Hoe vaak komt dit probleem nu voor?

Observationeel onderzoek levert geen deugdelijk bewijs voor causale relaties, maar is toch van belang omdat het duidelijk kan maken hoe vaak een probleem voorkomt. Anders gezegd: gaat het hier om een zeldzaamheid of is het een structureel probleem? Deze vraag negeren staat in de besliskunde te boek als *base-rate neglect*. Meer precies verwijst het begrip naar het negeren van toepasselijke achtergrondinformatie die de zaak in het juiste perspectief kan plaatsen.<sup>37</sup> De traditionele juridische benadering van te beoordelen problemen maakt dat deze achtergrondinformatie eenvoudigweg meestal niet gezocht wordt.<sup>38</sup>

Een aanschouwelijk voorbeeld bieden aansprakelijkheids- en tuchtzaken die handelen over het beweerdelijk ‘missen’ door radiologen van borstkanker bij het bevolkingsonderzoek naar die ziekte.<sup>39</sup> Bij geen van deze zaken kwam de relevante achtergrondinformatie ter sprake. Die betreft hier de diagnostische onzekerheid – foutenmarge – die voortvloeit uit de beperkingen van de techniek en niet de incompetentie van de individuele arts. Bij ongeveer 30% van de vrouwen, die regelmatig deelnamen aan de borstkankerscreening werd de diagnose van mammacarcinoom *buiten* de screeningsdiagnostiek om gesteld. De foutenmarge van 30% is zodoende de relevante base rate. Dan gaat het niet langer om broddelwerk door een individuele radioloog waarvan iemand slachtoffer werd. Hier is duidelijk sprake van een ander – veel groter – en vooral structureel probleem dat verdisconteerd moet worden in het oordeel over de individuele radioloog die een mammacarcinoom ‘over het hoofd zag’.<sup>40</sup>

### 3. Hoe staat het met de nauwkeurigheid en de precisie van de test?

Deskundigen kunnen in hun rapport verwijzen naar wetenschappelijke literatuur, maar ze kunnen ook verwijzen naar tests die zij zelf hebben afgenomen om de feiten van een zaak op te helderen. Dat kunnen psychologische, medische of forensische tests zijn. Deze tests verwijzen in een afgeleide vorm naar wetenschappelijke literatuur over hun bruikbaarheid. De rechter kan naar die literatuur vragen en meer specifiek naar twee aspecten. Ten eerste: hoe nauwkeurig (=accuraat) is deze test in het herkennen van de gevallen waar het om gaat?<sup>41</sup> Ten tweede: levert herhaling van de test door anderen of dezelfde experts steeds hetzelfde resultaat op? Dat is de reproduceerbaarheid (de consistentie of precisie) van de test.

De kwesties van accuraatheid en reproduceerbaarheid zijn bijvoorbeeld aan de orde – zouden dat althans moeten zijn – in het bestuursrecht bij procedures over de juistheid van de beoordeling van arbeidsvermogen. Die cirkelen steeds om twee vragen: welk onderliggend lijden veroorzaakt de problemen bij de werknemer en in welke mate kan hij nog functioneren? Deskundigen gaan deze vragen te lijf met tests en analysemethoden en leveren dusdoende ‘medisch bewijs’. Maar daarbij blijven accuraatheid en reproduceerbaarheid vaak onbesproken.<sup>42</sup>

### 4. De toepasbaarheid van wetenschappelijk onderzoek op het individu: het *reference class* problem

Wetenschappelijk onderzoek beschrijft vaak een steekproef, die op een bepaalde manier werd samengesteld en die gekenmerkt wordt door diverse parameters: leeftijdsopbouw, verhouding tussen mannen en vrouwen, aard van het werk, geografische factoren, mate van blootstelling, ernst van afwijkingen, en het gelijktijdig vóórkomen van andere factoren die de afloop kunnen beïnvloeden (leefgewoontes, voeding). Maar bezitten de inzichten die de deskundige naar voren schuift op basis van dit onderzoek relevantie voor de individuele casus die de inzet van het rechtsgeschied vormt? Kan in redelijkheid worden aangenomen dat deze casus in cruciale opzichten op de steekproef lijkt?

Een 43-jarige man met duizeligheidsklachten en instabiliteit wordt op een spoedeisende hulp-afdeling van een ziekenhuis onderzocht en weer naar huis gezonden.<sup>43</sup> Drie dagen later blijkt hij een herseninfarct te hebben met duurschade: deze diagnose was dus niet gesteld door de artsen in het ziekenhuis. In het kader van een mogelijke proportionele aansprakelijkheid wordt door deskundigen gezocht naar studies, waaruit valt af te leiden hoe groot de gezondheidsschade is vanwege de te laat gestelde diagnose. Een van de partijdeskundigen baseert zich daarbij op een descriptieve studie, uitgevoerd in de regio Schotland, naar vaatziekten zoals hartinfarcten en beroertes onder personen ouder dan 60 jaar. Hier zien we het *reference class problem*.<sup>44</sup> De Schotse studie betreft namelijk een andere leeftijdsgroep (boven de 60 jaar) dan die waartoe de patiënt behoort (onder de 50 jaar), het doel van de studie was niet het in kaart brengen van foutenmarges bij de diagnostiek van herseninfarcten en al evenmin de effecten van verlate diagnostiek. Bovendien werd deze studie uitgevoerd in een regio van een land met een hogere frequentie van hart- en vaatziekten dan in ons land. De door de deskundige aangehaalde studie past dus niet bij de onderzoeksvraag die in deze casus aan de orde moet zijn.

### 5. De vertaling naar het individu: het *G2i*-probleem

Als het *reference class problem* tot volle tevredenheid is opgelost, duikt er een volgende horde op. Hoe wordt ken-

nis vanuit groepsonderzoeken naar het niveau van de individuele casus gebracht, het specifieke *diagnostische bewijs*.<sup>45</sup> Dit wordt ook wel aangeduid als *group-to-individual* (G2i).<sup>46</sup> Elke toepassing van empirische kennis in een juridische procedure brengt steeds een G2i-probleem met zich mee, want het gaat steeds om de materiële waarheid bij individuele kwesties.

In het genoemde voorbeeld van de schilder met blaaskanker is de eerst te beantwoorden vraag of er wel voldoende betrouwbaar bewijs bestaat dat schilders een groter risico hebben op het krijgen van blaaskanker: bewijst het algemene raamwerk de G van G2i? Vervolgens dient zich de vraag aan of daarvan ook in dit individuele geval sprake is: dus de i van G2i. Hier gaat het om een sprong van het algemene risico op blaaskanker bij schilders naar het individuele geval van deze schilder en zijn blaaskanker. De volgende stap is dan het maken van een inschatting voor het individu in kwestie: die kan nooit lager zijn dan de algemene kans van het algemene raamwerk, het is de ondergrens.<sup>47</sup>

## 6. De rechter op onbekend terrein

De kennisparadox doet zich voor als de rechter moet gaan oordelen over een kwestie waarbij hij, vanwege zijn eigen kennistekort, de hulp van een deskundige nodig heeft maar hij de inbreng van die deskundige niet goed op zijn merites kan beoordelen.<sup>48</sup> Het is voor rechters ondoenlijk maar ook onwenselijk om dat kennistekort via scholing op te heffen, want er moet nu eenmaal binnen een redelijke termijn recht worden gesproken. De rechter hoeft zich echter niet neer te leggen bij de kennisparadox: als de rechter op een methodologisch kompas vaart, kan hij heel goed deskundigen de maat nemen. Het vereist wel dat de rechter zich de algemene principes eigen maakt van hoe wetenschappelijke kennis kan worden vergaard en zodoende begrijpt dat de ene vakpublicatie de andere niet is.<sup>49</sup>

Deskundigen zullen elkaar tegenspreken als de uitkomsten van wetenschappelijke studies waarop zij zich beroepen diametraal tegenover elkaar staan. Dan moet de rechter de deskundigen actief bevragen. Naar welk soort onderzoek verwijst de deskundige en waarom? Welk type bronnen werd geraadpleegd (zie figuur 1)? Wat is de bewijskracht ervan (zie figuur 2)? Is dit alles werkelijk van toepassing op deze casus (zie de vijf punten van par. 5)? Het is goed denkbaar dat de rechter al vragenderachter komt dat de ene deskundige zich baseert op meta-analyses en de ander op casuïstiek. In zo'n geval doet de rechter er verstandig aan om het deskundigenbericht van de eerste expert meer krediet te geven dan dat van de tweede expert.

Een actieve houding van rechters valt te verkiezen, waarbij ze met de bovenbeschreven hiërarchie in staat

## De verschillende manieren waarop je empirische wetenschap kunt bedrijven beogen allemaal *to diminish the odds that we will fool ourselves and fool others*, maar sommige zijn daarin succesvoller dan andere

zullen zijn daaraan vorm te geven door deskundigen uit te horen over hun specialistische kennis en in welke mate die relevantie draagt voor de onderhavige casus. Hoe hard is de deskundige plank over het moeras van rechterlijke onwetendheid? Dat is steeds de vraag en een echte deskundige snapt die vraag en kan haar beantwoorden.

Niet alleen de deskundige moet zijn conclusies expliciet motiveren, dat geldt ook voor de rechter. De Hoge Raad legt tot nog toe in haar uitspraken over deskundigen een epistemisch riskante asymmetrie aan de dag: niet motiveren als de deskundige wordt gevolgd en wel motiveren als van het deskundige oordeel wordt afgeweken.<sup>50</sup> Wat ons betreft motiveert de rechter zijn keuzes in alle gevallen en doet hij dat onder verwijzing naar hoe solide de specialistische kennis is. We begrijpen de juridische redenen voor een vrije bewijsleer, maar stellen dat dit in het geval van empirisch gefundeerde bewijsmiddelen epistemologisch niet vrijblijvend kan en mag zijn. Want de verschillende manieren waarop je empirische wetenschap kan bedrijven beogen allemaal *to diminish the odds that we will fool ourselves and fool others*, maar sommige zijn daarin succesvoller dan andere.<sup>51</sup>

### Verkort aangehaalde literatuur

Alonso-Coello, P. et al., 'GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: introduction', *BMJ* 2016, vol. 353, i2016.

Alonso-Coello, P. et al., 'GRADE evidence to decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: clinical practice guidelines', *BMJ* 2016, vol. 353, i2089.

Amaya, A., 'Coherence, evidence, and legal proof', *Legal Theory* 2013, vol. 19, p. 1-43.

37. Guthrie et al. 2000.

38. Engel 2012.

39. Bijv. Rb. Zwolle-Lelystad 8 september

2004 ECLI:NL:RBZLY:2004:AQ9892,

Rb. Amsterdam 18 mei 2005,

ECLI:NL:RBAMS:2005:AT7785 en Rb. Gel-

derland, locatie Arnhem 26 augustus 2014,

JA 2014/121 (m.nt. R.W.M. Giard).

40. Zie voor de uitwerking daarvan Giard 2016a, hfdst. 8.

41. Dit heet ook wel de gevoeligheid (de conformiteit of sensitiviteit) van de test.

42. Giard 2016a.

43. Rb. Amsterdam 14 juli 2010,

ECLI:NL:RBAMS:2010:BO3288.

44. Zie daarover Roberts 2007 en Colyvan & Regan 2007.

45. Zie Faigman et al., 2014. Hij gebruikt hiervoor respectievelijk de termen 'framework evidence' en 'diagnostic evidence'.

46. Dawid 2015.

47. Zie Dawid 2015, Dawid et al. 2015 en Giard 2016b.

48. De Groot 2008, p. 13-14.

49. Giard 2016a.

50. HR 5 december 2003,

ECLI:NL:PHR:2003:AN8478, r.o. 3.6.

51. Lilienfeld et al. 2016.

- Bock, R.H. de, *Tussen waarheid en onzekerheid: over het vaststellen van feiten in de civiele procedure* (diss. Tilburg) Deventer: Kluwer 2011.
- Colyvan, M., & H.M. Regan, 'Legal decisions and the reference class problem', *The International Journal of Evidence & Proof*, 11 (2007), p. 274-85
- Cooper, H., *Research synthesis and meta-analysis: A Step-by-Step Approach*, (vijfde druk) Sage publications, 2015.
- Crombag, H.F.M. 'Rechters & deskundigen', *NJB* 2000, afl. 33, p. 1659-1665.
- Dawid, A.P. et al., 'From statistical evidence to evidence of causality', *Bayesian Analysis*, 2015, p. 1-28.
- Dawid, A.P., 'The Role of Scientific and Statistical Evidence in Assessing Causality', in: *Perspectives on Causation*, R. Goldberg (red.) Oxford: Hart Publishing 2011, p. 133-147.
- Dawid, A.P., 'On individual risk', *Synthese* 2017, vol. 194(9), p. 3445-3474.
- Engel, Chr., *Neglect the Base Rate: It's the Law!*, Max Planck Institute for Research on Collective Goods, Working Paper, Bonn 2012.
- Ettekhoven, B.J. van, 'De deskundige deskundige. Over registers en de "disclosure statement"', *O&A* 2016/53.
- Faigman, D.L. et al., 'Group to individual (G2i) inference in scientific expert testimony', *The University of Chicago Law Review* 2014, vol. 81, p. 417-480.
- Finkelstein, M.O., and B. Levin, *Statistics for Lawyers*, Dordrecht: Springer 2014.
- Fisher, M& F.C. Keil, 'The illusion of argument justification', *Journal of Experimental Psychology, General*, 2014, vol. 143 p. 425-433.
- Giard, R.W.M., 'Hoe bewijst het recht zichzelf? Een pleidooi voor evidence-based law', in: J.H. Nieuwenhuis & C.J.J.M. Stolker (red.), *Vooruit met het recht! Wat geldt in de rechts-wetenschap als vooruitgang?* Den Haag: Boom Juridische uitgevers 2006, p. 49-59.
- Giard, R.W.M., *Werken aan waarheidsvinding. Over het belang van de juiste onderzoeksmethoden in het aansprakelijkheidsrecht*, Den Haag: Boom Juridisch 2016a.
- Giard, R.W.M., 'Aansprakelijkheid voor personenschade door schadelijke stoffen ("toxic torts") de bijzondere verbintenis tussen toxicologie, epidemiologie en recht', *Expertise en Recht*, 2016b, p. 243-53
- Gigerenzer, G & W. Gaissmaier, 'Heuristic decision making', *Annual Review of Psychology* 2011, vol. 62, p. 451-482.
- Grimes, D.A., K.F. Schulz, 'An overview of clinical research: the lay of the land', *The Lancet* 2002, vol. 359, p. 57-61.
- Groot, G. de, *Het deskundigenadvies in de civiele procedure*, Deventer: Kluwer 2008.
- Guha, N. et al., 'Bladder cancer risk in painters: a meta-analysis', *Occupational and Environmental Medicine* 2010, vol. 67, p. 568-573.
- Guthrie, C. et al., 'Inside the judicial mind', *Cornell L. Rev.* 2000, vol. 86, p.778-829.
- Haan, A. de, & P. de Heer, 'Making the problem explicit', in: *Solving Complex Problems*, Den Haag: Eleven International Publishing 2015, p. 23-79.
- Hoving, R.A., *Deskundigenbewijs in het strafproces* (diss. Groningen), Oisterwijk: Wolf Legal Publishers 2017.
- Keil, F.C., 'Explanation and understanding', *Annual Review of Psychology* 2006, vol. 57, p. 227-254.
- Keil, F.C., 'Getting to the truth', *Brooklyn Law Review* 2008, vol. 73, p. 1035-1052.
- Lilienfeld, S. O. et al., *Scientific research in forensic samples. The Handbook of Forensic Psychopathology and Treatment*, London: Routledge 2016.
- Lipton, P., *Inference to the Best Explanation*, London: Routledge 2004.
- Maccoun, R.J. 'Biases in the Interpretation', *Annual Review of Psychology*, 1998, p. 259-287.
- McCloskey, M.S. et al., 'Effects of acute alcohol intoxication and paroxetine on aggression in men', *Alcoholism: Clinical and Experimental Research* 2002, vol. 33(4), p. 581-590.
- Pardo M.S. & R.J. Allen, 'Juridical proof and the best explanation', *Law and Philosophy* vol. 27, 2008, p. 223-268.
- Roberts, P., 'From theory into practice: introducing the reference class problem', *International Journal of Evidence & Proof*, 2007, p. 243-54
- Samuel, G., 'Is legal reasoning like medical reasoning?' *Legal Studies* 2015, vol. 35(2), p. 323-347.
- Schröder, M.J.C., 'De invloed van de authority bias op het waarden van deskundigenbewijs', *NTBR* 2016, p. 75-84.
- Sita, J. et al. 'Forensic handwriting examiners' expertise for signature comparison', *Journal of Forensic Science* 2002, vol. 47, p. 1-8.
- Strevens, M. 'No understanding without explanation', *Studies in History and Philosophy of Science Part A* 2013 vol. 44, p. 510-515.
- Zemla, J.C., et al., 'Evaluating everyday explanations', *Psychonomic Bulletin & Review* 2017, vol. 2, p. 1-13. •