

# Bewijs en overtuiging: Een helder zicht op valkuilen

Charles Berger\*, Diederik Aben\*\*

Met dit derde deel sluiten we het drieluik bewijs en overtuiging af. Als u de eerste twee delen (nog) niet gelezen hebt, kunt u dit deel niettemin als een zelfstandig artikel lezen, doordat we het Bayesiaans redeneerschema voor u samenvatten. Gegeven het belang van a-priorikansverhoudingen voor het correct gebruik van het redeneerschema besteden we hieraan enige aandacht. Daarna zullen we ingaan op een aantal veel voorkomende fouten in het redeneren met bewijs, en het risico van het missen van de samenhang tussen bewijsmiddelen. Afsluitend spreken we de verwachting uit dat het redeneerschema een waardevolle bijdrage zal leveren aan een verbetering van het redeneren met onzekere informatie.

*Dit artikel is het laatste uit een reeks van drie, waarin de auteurs langer stilstaan bij de mogelijkheden van het gebruik van het zogeheten Bayesiaanse redeneerschema bij het evalueren van bewijs, en veel voorkomende denkfouten. In dit derde deel gaat het vooral over die denkfouten.*

## **De logica van bewijsevaluatie**

Tot het moment waarop de rechter beslist of het bewijsmateriaal voldoende is voor een bewezenverklaring van de feiten in de tenlastelegging ('the ultimate issue'), moet hij redeneren met gegevens waarvan de juistheid onzeker is. Wij vinden dat inzichten uit de niet-juridische wetenschap de rechter hierbij behulpzaam kunnen zijn.

Een vermaard forensisch wetenschapper, Ian Evett, schreef dat ondanks alle ontwikkeling van technologie de belangrijkste vooruitgang van de afgelopen decennia schuilt in een inzicht omtrent de logica van bewijsevaluatie: voor de bewijskracht van bevindingen dient de deskundige de waarschijnlijkheid van de bevindingen te beschouwen, gegeven de hypothese, en dus niet andersom: de waarschijnlijkheid van

de hypothesen, gegeven de bevindingen.<sup>1</sup> In dit drieluik hebben we dat inzicht ook gepropageerd. Niettemin is de rechter juist geïnteresseerd in de waarschijnlijkheid van de hypothese, en dat in het licht van *alle* bevindingen. De deskundige zal daarover echter, op grond van enkel de bevindingen binnen zijn vakgebied, geen uitspraak kunnen doen. De rechter zal zich terdege bewust moeten zijn van het verschil tussen beide typen van uitspraken.

Het Bayesiaanse redeneerschema voldoet aan deze eisen van de logica, en kan binnen de strafprocessuele waarheidsvinding goede diensten bewijzen. Het sluit nauw aan bij de wijze waarop forensische wetenschappers de bewijskracht bepalen van de resultaten van hun analyse. Alleen al om die reden moet de strafrechter een zekere kennis hebben van dit redeneerschema. Die kennis is bovendien minstens zo belangrijk voor het redeneren van de juridische procesdeelnemers zelf. Naar wij in het tweede deel van het drieluik menen te hebben aangetoond hangt het gebruik ervan niet af van de beschikbaarheid van wetenschappelijke gegevens en cijfermateriaal.

Voor we ingaan op de valkuilen van het redeneren met bewijs, vatten we het Bayesiaanse redeneerschema eerst nog eens samen.

1. De bewijskracht van bevindingen ('evidence') is gerelateerd aan hypothesen. Een hypothese behelst een verklaring (de oorzaak) voor de bevinding (het gevolg), en is gerelateerd aan de relevante vragen in een zaak. Er zullen ten minste twee elkaar uitsluitende hypothesen in aanmerking moeten worden genomen, waarvan de waarschijnlijkheden in hun verhouding tot elkaar worden beschouwd. Het Bayesiaanse redeneerschema dwingt dus alternatieven te onderzoeken op hun waarschijnlijkheid, en daarbij in een strafrechtelijke context niet te volstaan met de lezing van bijvoorbeeld het openbaar ministerie.

Wat zijn dan die hypothesen zoal? Eerst geven we een voorbeeld met hypothesen over de vraag van wie aangetroffen celmateriaal afkomstig is. Dit zijn hypothesen op het zogeheten bronniveau<sup>2</sup>. Hoeveel waarde aan een DNA-match kan worden toegekend is gerelateerd aan de hypothesen in het licht waarvan dit

---

\* Dr. ir. C.E.H. Berger is principal scientist bij het Nederlands Forensisch Instituut en lid van het College van het Nederlands Register Gerechtelijk Deskundigen.

\*\* Mr. D.J.C. Aben is advocaat-generaal bij de Hoge Raad en redacteur van dit tijdschrift.

<sup>1</sup> I. Evett, Evaluation and professionalism, *Science and Justice* 49 (2009) 159-160. Hij las hierover voor het eerst in: M.O. Finkelstein, W.B. Fairley, A Bayesian Approach to Identification Evidence, *Harvard Law Review*, 83 (1970) 489-517.

<sup>2</sup> R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: deciding which level to address in casework, *Science & Justice* 38 (1998) 231-239.

bewijsmateriaal moet worden gewaardeerd. Één hypothese is dat de verdachte de bron is van het op de plaats delict aangetroffen celmateriaal. Een andere hypothese is dat de broer van de verdachte de bron is, en een derde hypothese is dat een nog onbekende, willekeurige persoon het celmateriaal heeft achtergelaten.

Een tweede voorbeeld gaat over de vraag naar de relevantie van de overeenkomst tussen dactyloscopische sporen en de vingerafdrukken van een verdachte. De bewijskracht van de overeenkomst met dactyloscopische sporen die op een schilderijlijst zijn aangetroffen hangt af van de lezingen (hypothesen) die door het openbaar ministerie en de verdediging worden aangedragen over de reden van het aantreffen van die vingersporen. De officier van justitie verklaart de aangetroffen “dacty” met de hypothese dat de verdachte het schilderij uit het museum heeft gestolen. De verdachte werpt als verklarende hypothese op dat hij de beheerder is van het museum waaruit het schilderij is gestolen, en in die functie de lijst heeft aangeraakt. Die hypothese leent zich vanzelfsprekend voor nader onderzoek.

2. Een bevinding (‘evidence’) wijst bij beschouwing van een aantal hypothesen in de richting van de juistheid van de hypothese waaronder die bevinding het meest waarschijnlijk is, met andere woorden: in de richting van de hypothese waarin de bevinding het beste past.

Een DNA-match bijvoorbeeld valt zonder meer te verwachten onder de hypothese dat het DNA-spoor afkomstig is van de verdachte. Die DNA-match is echter zeer onwaarschijnlijk als het DNA-spoor afkomstig is van een willekeurige derde.

In het tweede voorbeeld laat de overeenkomst van een vingerspoor op een schilderijlijst met de vingerafdruk van de verdachte zich goed verklaren indien hij de dief zou zijn van het schilderij. Maar diezelfde overeenkomst laat zich mogelijk even goed verenigen met de genoemde hypothese dat de verdachte die sporen achterliet bij werkzaamheden als beheerder van het museum. Op het bronniveau (wie was de bron van het spoor) is er geen verschil van mening. Maar op het zogeheten handelingsniveau<sup>2</sup> (bij welke handeling werd het spoor achtergelaten) wel. In deze situatie kan aan het dactyloscopische spoor vermoedelijk geen informatie worden ontleend, tenzij bijvoorbeeld duidelijk wordt dat de verdachte niet de beheerder was of dat beheerders hun schilderijen uitsluitend met witte handschoenen vastpakken.

Het ‘wijzen’ door bevindingen in de richting van de juistheid van een hypothese is relatief: de verhouding van de waarschijnlijkheden van de bevinding onder de uiteenlopende hypothesen geeft de relatieve toename of afname van onze mate van overtuiging, hoe groot of klein die ook was voorafgaand aan de bevinding. Die verhouding van waarschijnlijkheden wordt de ‘likelihood ratio’ (LR) genoemd, en in een medische context veelal de ‘diagnostische waarde’. Bewijsmateriaal dat diagnostisch is voor een bepaalde hypothese vergroot dus de kans dat die hypothese juist is, maar we weten daarmee nog niet hoe groot die kans is.

3. De bewijskracht van bewijsmateriaal (bevindingen) bepaalt dus niet de waarschijnlijkheid van die hypothesen *na* het doen van de bevindingen. Voor het bepalen van die waarschijnlijkheid moet ook een inschatting worden gemaakt van de waarschijnlijkheid van de hypothesen *voorafgaand aan* kennisneming van de bevindingen.

Het theorema<sup>3</sup> van Bayes is de wiskundige formule waaruit de voorgaande gedachten voortkomen, en die ze heel compact weergeeft. Een erg behulpzame vorm van dit theorema maakt gebruik van ‘odds’. Dat is het Engelse begrip voor kansverhoudingen dat wij kennen van het gokken op de uitslag van paardenrennen. Deze formule laat goed zien hoe de bevinding  $E$  van een gevolg (het bewijsmateriaal) ons helpt bij het beschouwen van mogelijke oorzaken daarvan (hypothesen  $H_1$  en  $H_2$ ):

$$\underbrace{\frac{P(H_1)}{P(H_2)}}_{\text{prior odds}} \times \underbrace{\frac{P(E | H_1)}{P(E | H_2)}}_{\text{likelihood ratio}} = \underbrace{\frac{P(H_1 | E)}{P(H_2 | E)}}_{\text{posterior odds}}$$

In deze notatie betekent  $P(H_1|E)$  bijvoorbeeld: de waarschijnlijkheid (‘probability’) van hypothese  $H_1$ , gegeven dat  $E$  (‘evidence’) is waargenomen. De linker term wordt de ‘prior odds’ genoemd. Het is een kansverhouding: de waarschijnlijkheid van hypothese  $H_1$ , gedeeld door de waarschijnlijkheid van de alternatieve hypothese  $H_2$ .

De middelste term is de likelihood ratio, een maat voor de bewijskracht gedefinieerd als de verhouding van de waarschijnlijkheid van de bevindingen  $E$ ,

---

<sup>3</sup> ‘Theorema’ wil zeggen dat deze formule deductief (noodzakelijkerwijs) volgt uit de basiswetten van de kansrekening.

gegeven beide hypothesen. Die LR is dus geen constante, maar hangt af van *welke* hypothesen worden betrokken in het onderzoek. In ons voorbeeld van de diefstal van het schilderij waren de dactyloscopische overeenkomsten sterk bewijs voor de hypothesen op bronniveau: de LR was veel groter dan 1. Het bewijsmateriaal ondersteunde dus sterk de hypothese dat de verdachte de bron is van het dactyloscopische spoor. Voor de genoemde hypothesen op handelingsniveau hadden die overeenkomsten echter geen bewijswaarde: de LR was gelijk aan 1. Het bewijsmateriaal als zodanig discrimineerde dus niet tussen de twee onderzochte hypothesen op handelingsniveau.

De uitkomst van de vermenigvuldiging levert wederom een kansverhouding op. Deze wordt de ‘posterior odds’ genoemd. Net als de prior odds is dit een maat voor onze overtuiging omtrent de juistheid van de hypothesen, maar nu geüpdate met de bewijskracht van de bevindingen. Het is de verhouding van de waarschijnlijkheden van de hypothesen, *na* kennisneming van de bevindingen. Deze is dus afhankelijk van zowel de prior odds als de likelihood ratio.

Wij pretenderen niet dat het redeneerschema een eenvoudige oplossing is voor alle problemen met probabilistisch redeneren. Het maakt namelijk ook duidelijk hoe ingewikkeld het is om consequent rationeel met bewijs te redeneren. Met de toenemende aandacht voor dat redeneren met bewijs ontdekken we in toenemende mate fouten in dat redeneren. Juist door het gebruik van het redeneerschema worden die fouten beter zichtbaar gemaakt. We zijn dan kennis rijker en een illusie armer.

In dit laatste deel van het drieluik over bewijs en overtuiging zullen we een aantal typen fouten in het redeneren nader beschouwen. Een groot deel van de voorbeelden is gerelateerd aan het DNA-onderzoek, dat al sinds zijn ontstaan in de jaren tachtig van de vorige eeuw omgaat met bewijsmateriaal op een wetenschappelijke wijze die analoog is aan het besproken Bayesiaanse redeneerschema. Ook door het grote aantal zaken waarin DNA-onderzoek wordt gedaan, speelt dit type onderzoek in onze voorbeelden een grote rol. Diezelfde fouten kunnen echter ook worden gemaakt bij de waardering van andere soorten van bewijsmateriaal: glas- en vezelsporen, patroonsporen, maar ook de waarde van herkenningen door getuigen.

Bij de meest voorkomende redeneerfouten worden de prior odds geheel niet of (impliciet) verkeerd ingeschat. Daarom willen we, voordat we het over die fouten gaan hebben, eerst wat uitgebreider stilstaan bij de prior odds.

### **Het beredeneerd schatten van de prior odds**

Terug naar de eerste term in het theorema van Bayes: de prior odds. Het Bayesiaanse redeneerschema maakt duidelijk dat de waarschijnlijkheid van de hypothesen *voorafgaande aan* de kennisneming van het bewijsmateriaal van essentieel belang is. Die waarschijnlijkheid correspondeert met de mate waarin geloof kan worden gehecht aan de juistheid van die hypothese. Aangezien die kans dus een graad van subjectieve overtuiging belichaamt, wordt dit kansbegrip ook wel het subjectieve kansbegrip genoemd. Een misverstand zou zijn om te menen dat de subjectiviteit van het kansbegrip meebrengt dat niet langs rationele weg over de grootte van de desbetreffende kansen kan worden gediscussieerd. Zogenaemde “frequentistische” statistici zien het begrip ‘waarschijnlijkheid’ eerder als een objectieve eigenschap van een systeem, die samenhangt met uitkomsten van oneindig herhaalde experimenten. Voor de problemen waarover wij hier spreken biedt zo’n kansbegrip geen uitkomst.

In de praktijk zijn mensen gewend te denken in termen van subjectieve kansen: “de kans dat ik aangereden word is groot als ik nu de straat oversteek”, “er is 70% kans dat het morgen regent”, “het is zeer waarschijnlijk dat de verdachte schuldig is.” Deze geschatte kansen hebben ook concrete gevolgen voor de beslissingen die we mede daarop baseren. Beide ‘soorten’ kansen moeten voldoen aan de basiswetten van de kansrekening, waardoor er langs dezelfde lijnen mee kan worden geredeneerd en gerekend. Deze kansen hoeven bovendien niet in getallen te worden weergegeven. Zoals al eerder betoogd, kan meestal met een globale niet-gekwantificeerde notie van de waarschijnlijkheid van een hypothese worden volstaan.

De rechter zal met de aanklager, de deskundige, en liefst ook de verdachte en zijn advocaat moeten zoeken naar de meest relevante hypothesen, want niet iedere hypothese is zinnig. De hypothese dat de verdachte het DNA-spoor achterliet hoeft bijvoorbeeld niet te worden vergeleken met de hypothese dat kaboutertjes zijn DNA naar de plaats delict hebben gebracht. Het maakt wellicht de LR gelijk aan 1, maar een dergelijke – voor zover wij weten onzinnige – hypothese benadeelt toch de

verdachte. Dat de verdachte schuldig is, is immers ook zonder enig bewijs vele malen waarschijnlijker dan de tussenkomst van kaboutertjes.

Nauw verbonden met het formuleren van de hypothesen is dus de vraag welke mate van plausibiliteit daaraan kan worden toegekend. Denk bijvoorbeeld aan zogenoemde ‘Meer en Vaart’-verweren.<sup>4</sup> De verdachte of zijn advocaat werpt een hypothese op. Vrijwel onmiddellijk zal de rechter zich afvragen hoe waarschijnlijk deze hypothese is. Indien hoogst onwaarschijnlijk kan de rechter dit soort hypothesen onbesproken laten, maar buiten dat geval zal de rechter bij verwerping van het verweer ten minste moeten motiveren op grond waarvan hij meent dat die hypothese onwaarschijnlijk (“onaannemelijk”) is. Daarbij kan hij putten uit feitenmateriaal in het strafdossier, althans voor zover ter terechtzitting besproken en onder nauwkeurige verwijzing naar de vindplaats ervan.<sup>5</sup> ‘Feiten van algemene bekendheid’ spelen hierbij eveneens een rol, als zij tenminste daadwerkelijk als zodanig kunnen worden bestempeld.

Ook in andere situaties zal de rechter de waarschijnlijkheid van de hypothesen beredeneerd moeten schatten. Als we op grond van het type en de locatie van een zaak bijvoorbeeld verwachten dat de dader afkomstig is uit de populatie van mannen in de omgeving van Den Haag, dan kan de kans dat een verdachte (een man uit Den Haag) die dader is bijvoorbeeld worden geschat op 1:250.000, of nog voorzichtiger 1:1.000.000. Voorafgaande aan de kennisneming van enig bewijsmateriaal in deze zaak is de kans dat de verdachte de dader is, gedeeld door de kans dat een willekeurige andere man uit die populatie de dader is, een heel klein getal.

Prior odds inschatten is lastig, maar het komt ons voor dat in de strafzaken waarin het ertoe doet, kwesties als deze ter terechtzitting aan de orde zouden moeten worden gesteld, opdat alle procesdeelnemers zich over de eventuele aannamen kunnen uitlaten. Overbodig is het beredeneerd schatten van de prior odds zeker niet. Het door deskundigen aangedragen bewijs moet immers worden gewaardeerd tegen de achtergrond van hypothesen, waarvan de waarschijnlijkheid los van het bewijs ook meegewogen moet worden.

Wanneer er slechts 2 hypothesen zijn en geen overige informatie, zijn die hypothesen daarmee in ieder geval niet automatisch even waarschijnlijk. De prior odds kunnen dus niet om die reden gelijk aan 1 (50%/50%) gesteld worden. Juist

---

<sup>4</sup> De verdachte voert daarbij een verweer dat niet in strijd is met de bewijsmiddelen maar wel met het ten laste gelegde.

<sup>5</sup> Zie onder meer Hoge Raad 23 oktober 2007, LJN BA5851, NJ 2008, 69.

wanneer er geen informatie is zouden de onbekende prior odds dan immers exact bepaald zijn. Toch werd deze willekeurige aanname in het verleden regelmatig onterecht gedaan.<sup>6</sup>

We zullen nu een aantal valkuilen beschouwen. Omdat er veel manieren zijn om te redeneren op een wijze die indruist tegen de logica hebben we een selectie moeten maken. We zullen ons beperken tot enkele veel voorkomende denkfouten, in het Engels ‘fallacies’ geheten.

### **Defense attorney’s fallacy**

Een valkuil die ‘defense attorney’s fallacy’<sup>7</sup> wordt genoemd doet het bewijs minder belastend lijken dan het in werkelijkheid is, door (impliciet) uit te gaan van veel lagere prior odds dan op basis van de feiten gerechtvaardigd is. Zo’n redenering verloopt bijvoorbeeld als volgt:

*Het DNA-profiel van de verdachte, met een frequentie van 1 op de 1 miljard, matcht met dat van het spoor. Maar er zijn op de wereld naar verwachting nog 6 mensen met datzelfde DNA-profiel. De kans dat de verdachte de donor is van het spoor is dus slechts 1/7.*

Die redenering is misleidend, omdat het impliciet en onterecht veronderstelt dat iedere wereldbewoner evenveel kans heeft om de donor te zijn. De prior odds worden dan gesteld op 1 op de 6 miljard. In werkelijkheid zal er altijd informatie zijn die het aantal mogelijke daders beperkt.

Meer subtiele varianten van deze misleidende redenering zijn vaak lastiger te herkennen. In het volgende voorbeeld lijkt de bewijskracht van schoensporen onderschat te worden op een wijze die doet denken aan de defense attorney’s fallacy.

*“De schoensporen veilig gesteld aan weerszijden van het slachtoffer vertonen hetzelfde profiel als zolen van de muilen en slippers in de maat 46 van het merk Damage te koop bij Bristol. De fabrikant van het schoeisel met deze zolen heeft bevestigd dat deze muilen en slippers uitsluitend geleverd zijn aan Bristol. Uit de gegevens van pintransacties*

---

<sup>6</sup> I.M. Ellman, D. Kaye, Probabilities and proof: can HLA and blood group testing prove paternity? *New York University Law School* 54 (1979) 1131, en A. Biedermann, F. Taroni, P. Garbolino, Equal prior probabilities: Can one do any better? *Forensic Science International* 172 (2007) 85.

<sup>7</sup> W. Thompson, E. Schumann, Interpretation of statistical evidence in criminal trials: the prosecutor’s fallacy and the defense attorney’s fallacy, *Law and Human Behaviour* 11 (1987) 167.



*bij Bristol te [P] blijkt dat de vriendin van verdachte op 9 juni 2006 een paar slippers van het merk Damage, maat 46 heeft gekocht. (...)*

*Uit de door Bristol ter beschikking gestelde gegevens blijkt dat alleen al in Nederland in de maten 40 tot en met 46 respectievelijk 3.152 en 2.698 exemplaren slippers en muilen van het merk Damage zijn verkocht. (...)*

*De rechtbank is van oordeel dat dit betekent dat de mogelijkheid dat de schoensporen zijn achtergelaten door ander schoeisel zodanig reëel is dat de schoensporen niet eenduidig kunnen worden herleid tot de verdachte. (...)*

*Daar waar de officier opmerkt, dat al deze sporen in verband kunnen worden gebracht met verdachte, gaat dat verband feitelijk niet verder dan dat verdachte niet kan worden uitgesloten als donor.”<sup>8</sup>*

Stel dat er 1660 paar slippers in maat 46 in Nederland (16,6 miljoen inwoners) zijn verkocht. Dit betekent dat gemiddeld ongeveer 1 op de 10.000 Nederlanders in het bezit is van die specifieke slippers. Als we vanwege de nogal forse maat 46 aannemen dat deze alleen door mannen worden gedragen, zal dat grofweg een factor 2 schelen. Is de verdachte dan hoogstens niet uitgesloten wanneer hij zulke slippers blijkt te bezitten, of is deze bevinding juist krachtig bewijs?

Laten we deze casuspositie beschouwen op de wijze die we in dit drieluik propageren. De kans om dergelijke sporen aan te treffen is ongeveer 10.000 keer groter indien de verdachte de sporen op de plaats delict heeft achtergelaten dan wanneer een willekeurige andere Nederlander de sporen heeft achtergelaten (dat wil zeggen, de LR is 10.000). Met andere woorden, van iedere 10.000 willekeurige Nederlanders zijn er gemiddeld 9.999 wel uitgesloten en één niet. Het is dus wel degelijk bijzonder om ‘niet uitgesloten’ te zijn. De LR van 10.000 is dan ook groot. Het is krachtig bewijs voor de hypothesen op bronniveau.

### **Transposed conditional**

Een andere valkuil hangt samen met de tweede term in de door ons gepresenteerde variant van het theorema van Bayes: de likelihood ratio (LR). Deze fout is die van de ‘transposed conditional’, een erg veel voorkomende fout die ook

---

<sup>8</sup> Rechtbank ‘s-Gravenhage 15 juli 2008, LJN BD7186.

wel wordt bestempeld als de ‘prosecutor’s fallacy’.<sup>9</sup> Die laatste benaming is echter minder gelukkig gekozen omdat deze fout in werkelijkheid door vrijwel iedereen regelmatig wordt begaan, en deze fout ook niet per definitie in het ‘voordeel’ van het openbaar ministerie is.

Bij deze denkfout vindt, zoals de naam probeert duidelijk te maken, een niet toegestane verwisseling plaats binnen de context van uitspraken over voorwaardelijke kansen. Een eenvoudig voorbeeld betreft de kans dat een dier vier poten heeft als het een koe is. Na het transponeren van de voorwaarde doet de spreker een uitspraak over de kans dat een dier een koe is als het vier poten heeft. Dat is een heel andere kans! Zo is ook de kans dat een man rijk is zeer verschillend van de kans dat een rijk persoon man is. In formulevorm:  $P(A|B)$  wordt verwisseld voor  $P(B|A)$ . Die kansen zijn in het algemeen dus niet gelijk. De voorwaarde is getransponeerd. Al is het onderscheid in deze simpele voorbeelden vrij evident, het risico deze fout te maken neemt toe wanneer de situatie minder inzichtelijk is.

Je zou ook kunnen zeggen dat men bij deze denkfout veronderstelt dat de posterior odds gelijk zijn aan de likelihood ratio, terwijl dat alleen zo is wanneer de prior odds toevallig gelijk zijn aan 1. De posterior odds worden daardoor te hoog ingeschat wanneer de prior odds in werkelijkheid kleiner zijn dan 1, of te laag wanneer de prior odds juist groter zijn dan 1.

Wij hebben op de website ‘rechtspraak.nl’ een aantal voorbeelden van deze denkfout met betrekking tot DNA-bewijs gezocht, en vrij eenvoudig gevonden.<sup>10</sup> Waar de deskundige telkens uitspraken deed over de kans op de bevindingen gegeven de hypothese  $P(E|H)$ , werden die door de rechter gelezen als de waarschijnlijkheid van de hypothese  $P(H|E)$  (waarbij de bevindingen meestal niet meer genoemd werden).

In een eerste voorbeeld overwoog de rechtbank:<sup>11</sup>

*“Het scenario dat het gaat om celmateriaal afkomstig van het slachtoffer en van verdachte, afgezet tegen het scenario dat het gaat om*

---

<sup>9</sup> P. Diaconis, D. Freedman, The persistence of cognitive illusions, *The Behavioral and Brain Sciences* 4 (1981) 333-334. I.W. Evett, Avoiding the transposed conditional, *Science & Justice* 35 (1995) 127-131. W. Thompson, E. Schumann, Interpretation of statistical evidence in criminal trials: the prosecutor’s fallacy and the defense attorney’s fallacy, *Law and Human Behaviour* 11 (1987) 167.

<sup>10</sup> Google search met als zoektermen: “de kans” “kleiner dan” DNA site:rechtspraak.nl.

<sup>11</sup> Rechtbank ‘s-Gravenhage 3 april 2009, LJN BH9948. Zie ook: H.W.J. de Groot, Schoten aan de Bezuidenhoutseweg in Den Haag, *Expertise en Recht* (2010) 91-98.

*celmateriaal afkomstig van het slachtoffer en een onbekende leidt bij het NFI tot de conclusie dat het eerste scenario 840 miljoen maal waarschijnlijker is dan het tweede”*

In het NFI-rapport werd echter de relatieve waarschijnlijkheid van de *bevindingen* onder twee hypothesen gegeven:

*“Onder de aanname dat van de DNA-loci die zijn meegenomen in de beschouwing van de twee hierboven weergegeven scenario’s alle DNA-kenmerken van de celdonoren in het DNA-mengprofiel zichtbaar zijn, zijn de verkregen resultaten ongeveer 840 miljoen maal waarschijnlijker onder scenario I dan onder scenario II.”*

In een tweede voorbeeld<sup>12</sup> werd de conclusie van de deskundige door het hof als volgt beschreven:

*“Volgens de onderzoeker in zijn brief van 7 juli 2000 is de kans dat het bloed op de glasscherf van een ander dan verdachte afkomstig is, kleiner dan 1:53.000.000.”*

Dat is een uitspraak over de waarschijnlijkheid van de *herkomst* van het celmateriaal. Maar de deskundige rapporteerde in werkelijkheid over de *zeldzaamheid* van het profiel:

*“De kans dat een willekeurig individu hetzelfde DNA-profiel heeft als welke gevonden is in het onderzochte bloed op de glasscherf ABG758, bedraagt vele malen minder dan één op de miljoen. Het actuele getal dat gevonden wordt bij statistische berekening in ons referentiedatabestand van blanke Nederlanders is één op 53 miljoen.”*

Ook in de volgende voorbeelden deed de deskundige telkens een uitspraak over zeldzaamheid, te weten de waarschijnlijkheid van de bevindingen wanneer een willekeurige persoon de donor van het sporenmateriaal zou zijn. Vervolgens zijn deze uitspraken telkens (foutief) getransponeerd in de waarschijnlijkheid dat de donor een willekeurige (andere) persoon is, en dat is een herkomstuitspraak.

*“Uit de technische en forensische onderzoeken en rapportages leidt de rechtbank af (...)*

---

<sup>12</sup> Zie HR 10 december 2002, LJV AE6863. De passage is afkomstig uit het verkorte arrest van het gerechtshof. Overigens is de hier gesignaleerde fout hersteld in de aanvulling op het verkorte arrest.

*dat op zeven van de acht enveloppen DNA materiaal is gevonden dat afkomstig kan zijn van verdachte, waarbij ten aanzien van vijf (feiten 2, 3, 6, 7 en 8) de kans dat de donor van de biologische sporen een ander is dan verdachte kleiner wordt geacht dan 1 op 1 miljard”<sup>13</sup>*

*“- Bemonstering van een vloermat van de bijrijderzijde: betreft bloed dat afkomstig kan zijn van het slachtoffer. De kans dat het bloed niet aan het slachtoffer toebehoort, is kleiner dan 1 op 1 miljard. De rechtbank gaat er van uit, gelet op deze kansberekening, dat deze bemonstering het bloed van het slachtoffer betreft;*

*- Bemonstering van de deurbekleding van het deurvak van de bestuurderzijde: betreft bloed dat afkomstig kan zijn van het slachtoffer. De kans dat het bloed niet aan het slachtoffer toebehoort, is kleiner dan 1 op 1 miljard.”<sup>14</sup>*

*“Uit deze vergelijking kwam naar voren dat het DNA materiaal dat op de jas aangetroffen was matcht met het DNA van verdachte. De kans dat het DNA materiaal dat op de jas aangetroffen was afkomstig is van een willekeurig gekozen ander persoon, is kleiner dan één op één miljard.”<sup>15</sup>*

*“De kans dat het onderzochte DNA-materiaal daadwerkelijk afkomstig is van verdachte en aangeefster en één ander onbekend persoon is 884.000 maal waarschijnlijker dan dat het onderzochte DNA-materiaal afkomstig is van verdachte en twee willekeurig gekozen personen.”<sup>16</sup>*

In de onderstaande voorbeelden verwordt de ontorechte herkomstuitspraak zelfs tot een soort categorisch oordeel:

*“Op 21 november 2008 heeft het NFI gerapporteerd dat op grond van vergelijkend DNA-onderzoek het bloed/celmateriaal: (...)*

*- bij genoemde stoel afkomstig is van verdachte en dat de kans dat het van iemand anders afkomstig is, kleiner is dan 1 op 1 miljard. (...)*

---

<sup>13</sup> Rechtbank Groningen 1 december 2008, LJN BG5720.

<sup>14</sup> Rechtbank Amsterdam 4 februari 2009, LJN BH1795. Doch zo ook het gerechtshof Amsterdam 2 februari 2010 in hoger beroep van dezelfde zaken, LJN BL9992 en BL9993 (moord garagehouder Amsterdam-Noord).

<sup>15</sup> Rechtbank Dordrecht 16 juni 2009, LJN BI8230.

<sup>16</sup> Rechtbank Utrecht 15 juli 2009, LJN BK7518.

- op de binnenzijde van genoemde toegangsdeur afkomstig is van [slachtoffer] en dat de kans dat het van iemand anders afkomstig is, kleiner is dan 1 op 1 miljard.”<sup>17</sup>

“In de handschoen die is aangetroffen op 70 cm van het hoofd van het slachtoffer is DNA gevonden van verdachte. De kans dat dit DNA afkomstig is van een ander dan verdachte is kleiner dan één op een miljard.”<sup>18</sup>

“Op twee peuken is de kans dat het aangetroffen DNA-materiaal afkomstig is van iemand anders dan van [medeverdachte 2], kleiner dan 1 op 1 miljard. Op vier peuken is de kans dat het aangetroffen DNA-materiaal afkomstig is van iemand anders dan [medeverdachte 6], kleiner dan 1 op 1 miljard.”<sup>19</sup>

“De kans dat dit bloed van een ander dan verdachte afkomstig is, is kleiner dan één op één miljard.”<sup>20</sup>

“De verklaringen van aangeefster en de getuigen over het signalement van de dader, in combinatie met het op de hondenriem en tailleband aangetroffen DNA-materiaal, waarvan de kans kleiner is dan één op één miljard dat het van een willekeurige andere gekozen persoon afkomstig is dan van verdachte, leiden met inachtneming van al het voorgaande tot de conclusie dat wettig en overtuigend bewezen is dat verdachte het ten laste gelegde heeft begaan.”<sup>21</sup>

“Gegeven de hierboven genoemde uitkomsten van het DNA-onderzoek, kan worden geconcludeerd dat de kans verwaarloosbaar klein is dat het aangetroffen bloed afkomstig is van een willekeurige andere man dan verdachte.”<sup>22</sup>

Wij beweren niet dat de denkfouten in deze zaken hebben geleid tot rechterlijke dwalingen. Met name bij zeer onderscheidend bewijsmateriaal, zoals een DNA-match kan zijn, en niet te verwaarlozen prior odds, hoeft deze denkfout nog niet

---

<sup>17</sup> Rechtbank Zutphen 8 juli 2009, LJN BJ1981.

<sup>18</sup> Rechtbank 's-Hertogenbosch 4 maart 2009, LJN BH4598.

<sup>19</sup> Rechtbank Utrecht 23 oktober 2009, LJN BK1686.

<sup>20</sup> Rechtbank Zutphen 4 november 2008, LJN BG3285.

<sup>21</sup> Rechtbank Arnhem 12 april 2010, LJN BM0871.

<sup>22</sup> Rechtbank Groningen 18 juni 2008, LJN BA7826.

tot ongelukken te leiden. Het gevaar is echter levensgroot dat dezelfde denkfout ook wordt gemaakt bij de evaluatie van bewijsmateriaal dat een minder groot onderscheidend vermogen heeft, of wanneer de prior odds erg klein zijn.

De zogeheten ‘false positive fallacy’<sup>23</sup> ten slotte, is ook een voorbeeld van de ‘transposed conditional’. Ditmaal betreft de denkfout niet een verwisseling van de zeldzaamheid van een DNA-profiel met de waarschijnlijkheid van de herkomst van het betreffende celmateriaal. De thans bedoelde denkfout heeft betrekking op het onderscheid tussen het feitelijk overeenkomen van twee DNA-profielen en het *rapporteren* daarvan. Daarbij spelen de mogelijkheid van laboratorium- of administratieve fouten en andere fouten in de ‘chain of custody’ een rol. De ‘false positive fallacy’ bestaat uit de verwarring van twee kansen, namelijk (1) de kans op een gerapporteerde DNA-match in het geval DNA-profielen van sporenmateriaal en celmateriaal van de verdachte (feitelijk) *niet* overeenkomen, en (2) de kans dat DNA-profielen van sporenmateriaal en celmateriaal van de verdachte (feitelijk) *niet* overeenkomen in het geval een DNA-match is gerapporteerd.

### **Base rate fallacy**

Een denkfout die nauwe samenhang vertoont met de hiervoor beschreven ‘transposed conditional’ is de ‘base rate fallacy’.<sup>24</sup> Hiervoor hebben wij al gewaarschuwd toen wij in de samenvatting van het Bayesiaanse redeneerschema meedeelden dat de ‘posterior odds’ mede afhankelijk zijn van de ‘prior odds’. Degene die deze denkfout begaat negeert de ‘prior odds’, hetgeen alleen mag als de ‘prior odds’ toevallig gelijk zouden zijn aan 1.

Als we bijvoorbeeld stellen dat redeneerfouten veel voorkomen zou u al gauw kunnen denken dat er dus ook veel onterechte veroordelingen zijn. Maar zelfs wanneer u onterecht zou aannemen dat iedere denkfout een veroordeling tot gevolg heeft, dient u ook in aanmerking te nemen welk percentage onschuldigen er überhaupt voor de rechter komt te staan.

Zelfs een grote bewijskracht kan lage posterior odds opleveren wanneer de prior odds erg klein zijn. Als de bewijskracht (LR) van een positieve HIV-test bijvoorbeeld 100 is, dan heeft een gemiddelde volwassen Nederlander op grond van alleen die test ruimschoots meer kans om *niet* besmet te zijn. Dat komt doordat de

---

<sup>23</sup> W.C. Thompson, F. Taroni, C.G.G. Aitken, How the Probability of a False Positive Affects the Value of DNA Evidence, *Journal of Forensic Science*, 48 (2003) 1-8.

<sup>24</sup> M. Bar-Hillel, The base-rate fallacy in probability judgments, *Acta Psychologica*, 44 (1980) 211-233.

prior odds erg klein zijn vanwege de lage prevalentie van HIV-infectie onder volwassen Nederlanders.

In een forensische context bestaat het gevaar van een base rate fallacy bijvoorbeeld bij het waarderen van een DNA-match die door een ‘database search’ is verkregen. De likelihood ratio van de DNA-match verandert binnen deze context niet. Daar staat echter tegenover dat bij eventueel gebrek aan enig ander bewijs de ‘prior odds’ voor de hypothese dat de persoon wiens identiteit door de database search is achterhaald ook werkelijk de bron van het spoor is zeer klein zijn.<sup>25</sup>

### **De samenhang missen: onafhankelijkheid van bewijsmateriaal**

Onafhankelijk bewijsmateriaal dat wijst in dezelfde richting, versterkt elkaar. We zijn geneigd te denken dat het bewijs zo sterk is als zijn zwakste schakel, maar onafhankelijk bewijsmateriaal ondersteunt elkaar eerder als parallelle kettingen. Ook de zwakste delen kunnen dan helpen de ketting sterker te maken.

Een voorbeeld van een strafzaak waarin dit ogenschijnlijk niet werd ingezien, is recent beschreven in dit tijdschrift.<sup>26</sup> In die zaak kwamen onder meer drie bewijskwesities aan de orde waaromtrent het volgende werd betoogd: (1) Een getuige kan ten onrechte menen de verdachte te hebben herkend van de PD, (2) celmateriaal van de verdachte kan indirect zijn overgedragen op het lichaam van een moordslachtoffer en is dus mogelijk niet delictgerelateerd, en (3) een voorwerp dat als moordwapen is gehanteerd kan voordien van de verdachte zijn gestolen. Dat deze drie voorvallen onafhankelijk van elkaar plaatsvinden ten nadele van één verdachte is echter nog veel onwaarschijnlijker dan de (on)waarschijnlijkheid van de afzonderlijke gebeurtenissen. Onafhankelijk bewijs versterkt elkaar op die manier. Door *afhankelijkheid* zou het bewijs daarentegen een stuk zwakker kunnen worden. In het geval een derde de bedoeling had de verdachte voor moord op te laten draaien zou het verhaal immers kunnen zijn dat die persoon vals getuigde, verdachtes DNA bewust overbracht, en verdachtes wapen stal om dat zelf bij de moord achter te laten.

De vraag naar de onderlinge (on)afhankelijkheid van verschillende bevindingen is dus buitengewoon belangrijk. In het Bayesiaanse redeneerschema is die notie terug te vinden. Onder de voorwaarde dat de hypothesen gelijkkluidend zijn,

---

<sup>25</sup> K. van der Beek, A. Kloosterman, M. Sjerps, De interpretatie van een DNA-databankmatch, *Delikt en Delinkwent* (2010) 138-155, en M. Sjerps, R. Meester, Selection effects and database screening in forensic science, *Forensic Science International*, 192 (2009) 56-61.

<sup>26</sup> H.W.J. de Groot, Schoten aan de Bezuidenhoutseweg in Den Haag, *Expertise en Recht* (2010) 91-98.

mogen de likelihood ratio's van *onafhankelijk* bewijsmateriaal met elkaar worden vermenigvuldigd om tot de gecombineerde bewijskracht te komen.

Een bekend buitenlands voorbeeld van een zaak waarin juist te gemakkelijk werd aangenomen dat bevindingen onderling onafhankelijk waren, is de zaak van Sally Clark in het Verenigd Koninkrijk. Haar beide zoontjes overleden enkele maanden na hun geboorte. Een expert op het gebied van wiegendood liet weten dat de kans op wiegendood voor een gezin als het hare 1 op de 8543 was. Hij getuigde dat de kans op tweemaal wiegendood dus slechts 1 op de  $8543 \times 8543 = 73$  miljoen was. Daarmee zou zo'n gebeurtenis volgens hem landelijk maar één keer per eeuw plaatsvinden. De jury veroordeelde Sally Clark vervolgens tot levenslang voor dubbele moord. In werkelijkheid komt dubbele wiegendood echter veel vaker voor: De FSID (Foundation for the Study of Infant Death) liet weten dat dit fenomeen gemiddeld één of twee keer per jaar plaatsvindt. Twee gevallen van wiegendood binnen één gezin zijn namelijk niet onafhankelijk van elkaar: het eerste geval doet de kans op een tweede aanzienlijk toenemen.

Wij sluiten bovendien niet uit dat de jury nog een denkfout heeft gemaakt, door wellicht te redeneren dat de kans op onschuld van Sally Clark onder de geschetste omstandigheden zo'n slordige 1 op 73 miljoen zou zijn. Dat zou weer een voorbeeld van de 'transposed conditional' zijn.<sup>27</sup> Sally Clark werd uiteindelijk na een verblijf van ruim drie jaren in de gevangenis vrijgesproken.

### **De samenhang missen: bewijs vóór het combineren al wegstrepen**

Ten slotte willen wij kort wijzen op een redeneerfout die bij het optuigen van bewijsconstructies en vrijspraakmotiveringen kan worden gemaakt. Vrijspraken worden nogal eens toegelicht door het bewijsmateriaal afzonderlijk onder de loep te nemen en de samenhang buiten beschouwing te laten. Ieder bewijsonderdeel afzonderlijk wordt dan door de rechter onvoldoende sterk bevonden om daarop een bewijsconstructie te bouwen en 'weggestreept'.<sup>28</sup> De structuur van de bewijsredenering kan inderdaad meebrengen dat bij het wegvallen van een onderdeel daarvan een bewezenverklaring onvoldoende gefundeerd is in de bewijsmiddelen. In

---

<sup>27</sup> Er ging nog veel meer fout in die strafzaak. Zo was de kans op wiegendood voor *jongetjes* meer specifiek 1 op de 1300. Verder had niemand geclaimd dat de eerste overleden baby een geval van wiegendood was. Ook was medische informatie over infecties bij de kinderen achtergehouden. Zie verder: R. Hill, Multiple sudden infant deaths – coincidence or beyond coincidence?, *Paediatric and Perinatal Epidemiology*, 18 (2004) 320–326.

<sup>28</sup> Zie gerechtshof 's-Hertogenbosch 24 september 2008, LJN BF2188; rechtbank 's-Gravenhage 15 juli 2008, LJN BD7186.



veel gevallen echter wijzen verschillende – onafhankelijke – bewijsmiddelen in dezelfde richting. Één bevinding mag dan onvoldoende zijn voor de sprong naar een bewezenverklaring, maar de onderlinge combinatie en samenhang van het geheel zou uiteindelijk een bewezenverklaring mogelijk kunnen dragen. Het gevaar is dus de bewijsstandaard op een eerder moment toe te passen dan op het moment waarop het bij bewijsvragen aankomt: ‘the ultimate issue’, de bewijsbeslissing die aan de feitenrechter is voorbehouden. Deze redeneerfout heeft veel weg van het afronden van getallen, om daar vervolgens mee door te rekenen. Dat kan tot absurde resultaten leiden (“*druppel dan mijn benzinetank maar vol!*”).

Dezelfde redeneerfout kan ook tot veroordelingen leiden. Als de bewijsstandaard op de afzonderlijke bewijsmiddelen wordt toegepast, en die maatstaf wordt telkens gehaald, dan zou het lijken alsof die conclusie ook voor de gehele bewijsconstructie zou gelden. De redenering verliest dan echter uit het oog dat *onzekerheden* elkaar onder omstandigheden ook kunnen versterken. De afzonderlijke redeneerstappen mogen begrijpelijk zijn, de uitkomst van het geheel is dat dan misschien niet meer.

## **Conclusie**

Dit was het laatste deel van ons drieluik. Wij hebben hierin stilgestaan bij de valkuilen van het redeneren met bewijs. Het door ons gepropageerde redeneerschema brengt de denkfouten onzes inziens eerder aan het licht en helpt ook om zulke fouten te voorkomen. Wij zijn ons ervan bewust dat het niet gemakkelijk is, en dat kansrekening, en daarmee ook dit redeneerschema, zich binnen het domein van juristen nog niet mag verheugen in een hoge mate van populariteit. Pogingen om dit redeneerschema toe te passen op bewijskwesities en de vraag of bepaald bewijsmateriaal de enige getuigenverklaring voldoende ondersteunt, worden wel ervaren als “*problematiserend*”.<sup>29</sup> Maar de maatschappij stelt steeds hogere eisen, en het wetenschappelijke niveau van het forensisch onderzoek blijft toenemen. Een begin zal dus toch moeten worden gemaakt, zo is ons gezamenlijk oordeel, dat stoelt op ervaringen binnen de wetenschappelijke wereld en binnen de wereld van de rechtspraak. Wij zijn ervan overtuigd dat die werelden elkaar kunnen naderen zonder ten onder te gaan aan de zwaartekrachten van het eigen domein.

---

<sup>29</sup> Zie T. Kooijmans, *Bewijzen door de strafrechter*, *Ars Aequi* (2010) 456-465.

Het drieluik “Bewijs en overtuiging” gaat over rationeel redeneren voor wetenschappers en juristen. Het Bayesiaanse denkraam geeft daarvoor een aantal regels en biedt een helder zicht op valkuilen.

1. Bewijs en overtuiging: Rationeel redeneren sinds Aristoteles
2. Bewijs en overtuiging: Redeneren in de rechtszaal
3. Bewijs en overtuiging: Een helder zicht op valkuilen