

Rechters en Raadsheren over Bayes¹

Henry Prakken

Faculteit der Rechtsgeleerdheid, Rijksuniversiteit Groningen &
Faculteit Bètawetenschappen, Departement Informatica, Universiteit Utrecht

1 januari 2020

1 Inleiding

In zijn hele loopbaan is Peter van Koppen een criticus geweest van de toepassing van Bayesiaans denken over bewijskwesities in strafzaken. Dit geldt zowel voor de interpretatie van de bewijskracht van individuele bevindingen als het gebruik van Bayes als een algemeen denkmodel voor strafrechtelijk bewijzen. Wat betreft het eerste bekritiseert Van Koppen onder meer de strikte scheiding die forensisch deskundigen aanhouden tussen het bepalen van de bewijskracht van een bevinding, wat de deskundige volgens hen moet doen, en het bepalen van de priorikansen, wat de rechter volgens hen moet doen (Van Koppen 2017a, 2017b; Sjerps et al. 2017). Mijn artikel gaat niet hierover, maar over het gebruik van de Bayesiaanse kansrekening als algeheel denkmodel voor bewijsbeslissingen in het strafrecht. Ook dit vindt Van Koppen geen goed idee, onder meer omdat de kansrekening tegenintuïtief is voor juristen, omdat a priorikansen alleen subjectief te schatten zouden zijn en omdat de Bayesiaanse kansrekening niet goed met statistische afhankelijkheden zou kunnen omgaan (zie in meer detail Hoofdstuk 9 van Van Koppen, 2011).

Recent lijkt hij wat betreft het laatste in een aantal rechtbanken en gerechtshoven medestanders te hebben gevonden. In twee zaken ging het om door het OM ingebrachte Bayesiaanse analyses van de gehele zaak door de klimaatfysicus dr. Frans Alkemade. In de herzieningszaak *Zes van Breda* (Gerechtshof Den Haag 14 oktober 2015, ECLI:NL:GHDHA:2015:2860) oordeelde het hof dat de Bayesiaanse methode naar de huidige stand van de wetenschap onvoldoende betrouwbaar is voor toepassing in complexe strafzaken. In het hoger beroep van de zaak van de Oosterlandse brandstichtingen (Gerechtshof 's-Hertogenbosch 22 november 2016, ECLI:NL:GHSHE:2016:5165) was het hof het hier kennelijk mee eens. In twee andere zaken ging het om door de verdediging ingebrachte analyses van de gehele zaak van de wetenschapsfilosoof prof. Ton Derksen. De rechtbank Zeeland-West-Brabant oordeelde op 23 mei 2016 (ECLI:NL:RBZWB:2016:3060) dat de vraag of een tenlastegelegd feit wettig en overtuigend kan worden bewezen, geen vraag is die zich leent voor kansrekening. Tenslotte zette het Gerechtshof 's-Hertogenbosch op 5 februari 2018 (ECLI:NL:GHSHE:2018:421) Derksens analyse met een gedetailleerde motivering opzij.

In dit artikel analyseer ik deze oordelen van de rechters en raadsheren. Ik onderzoek of hun argumenten steekhoudend zijn en zo ja, of ze alleen voor de Bayesiaanse

¹ Dit artikel bouwt voort op Prakken (2014, 2018) en Prakken & Meester (2017). Verschillende fragmenten en inzichten zijn ontleend aan die artikelen.

kansrekening opgaan of misschien ook voor andere theorieën over rationeel bewijzen, en wat dat zegt over hun toepasbaarheid in rechtszaken. Om de context te bepalen vat ik eerst de drie voornaamste benaderingen van rationeel juridisch bewijzen kort samen.

2 Rationeel bewijzen: drie benaderingen

De meeste wetenschappers die rationeel juridisch bewijzen bestuderen, propageren een van drie typen benaderingen of een combinatie daarvan: een argumentatieve, een scenario- of een Bayesiaanse benadering.

De **argumentatieve benadering** gaat terug op de Amerikaanse strafrechtsgeleerde John Henry Wigmore (1931) en is in de Angelsaksische wereld herontdekt door de zogenaamde *New Evidence Scholars* (o.a. Anderson et al. 2005). AI-onderzoekers (Verheij 2000; Bex et al. 2003; Prakken 2004) hebben een formeel-logische variant van deze benadering ontwikkeld in termen van logische argumentatiesystemen. In de argumentatieve benadering heeft juridisch bewijzen twee kanten, een constructieve en een kritische kant:

- (1) het formuleren van een argument voor de bewijsbeslissing door stap voor stap, beginnend bij de bewijsmiddelen, gevolgtrekkingen te maken, gebruikmakend van algemene wereldkennis (generalisaties);
- (2) het kritisch onderzoeken van het zo geconstrueerde bewijsargument op zwakke plekken door bij elke stap te zoeken naar tegenargumenten.

Deze vorm van redeneren wordt 'weerlegbaar' (in het Engels '*defeasible*') redeneren genoemd, in tegenstelling tot deductief redeneren, dat in de klassieke formele logica bestudeerd wordt. Een bewijsbeslissing is in deze benadering gerechtvaardigd als het argument ten eerste intern deugdelijk is en ten tweede succesvol verdedigd kan worden tegen tegenargumenten.

De **scenariobenadering** (door Van Koppen ook wel *verhaalsbenadering* genoemd) is geïntroduceerd door de Amerikaanse psychologen Bennett & Feldman (1981) en Pennington & Hastie (1993) en in Nederland bekend geworden door de (rechts)psychologen Crombag, Van Koppen & Wagenaar (1994). Een recente uiteenzetting van de scenariobenadering is Van Koppen (2011). In deze benadering vindt strafrechtelijk bewijzen plaats door het construeren en vergelijken van scenario's over wat er gebeurd zou kunnen zijn gezien het bewijs. De scenariobenadering modelleert een vorm van redeneren die in de logica en filosofie *inference to the best explanation* wordt genoemd (Lipton 2004). Net als in de argumentatieve benadering heeft bewijzen via scenario's een constructieve en een kritische kant; Van Koppen (2011) noemt ze 'verificatie' en 'falsificatie'. *Verificatie* moet gebeuren door het construeren van een scenario dat *welgevormd* is (d.w.z. compleet en niet ambigu is en een centrale actie heeft in een context die de actie aannemelijk maakt), de bewijsmiddelen voorspelt en via de bewijsmiddelen *verankerd* is in onze gedeelde kennis van de wereld. *Falsificatie* moet gebeuren door het scenario paarsgewijs te vergelijken met redelijke alternatieve scenario's. Een

scenario is beter dan een alternatief scenario als het de bewijsmiddelen beter voorspelt.

De **Bayesiaanse benadering** (Fenton et al. 2016; Sjerps & Berger 2011) ziet juridisch bewijzen als denken over kansen. Uitgangspunt is net als bij de andere benaderingen dat de waarheid nooit met absolute zekerheid vast te stellen is. Bayesianen zeggen dan dat de rechter de *kans* dat de bewijsbeslissing met de waarheid correspondeert zo groot mogelijk moet maken. Zij trekken daaruit de conclusie dat juridisch bewijzen aan de axioma's van de kansrekening moet voldoen. De term 'Bayesiaans' duidt op een bepaald gebruik van het zogenaamde theorema van Bayes, een formule die wiskundig uit de axioma's van de kansrekening af te leiden is. Kort gezegd geeft deze formule een methode om de kans op een hypothese gegeven het beschikbare bewijs (de *a posteriorikans*) te berekenen uit de kans op de hypothese voordat het bewijs bekend is (de *a priorikans*) en de voorwaardelijke kansen op het bewijs gegeven de hypothese en gegeven zijn tegendeel (de bewijskracht of *likelihood ratio*).

De state-of-the art in Bayesiaans denken wordt gevormd door zogenaamde Bayesiaanse netwerken. Zulke netwerken zijn op minstens twee punten een vooruitgang ten opzichte van een rechtstreekse toepassing van het theorema van Bayes: ze ondersteunen een grafische weergave van statistische afhankelijkheden en ze staan toe om tussenstappen te formuleren in de relatie tussen hypothesen en bewijsmiddel, zodat expliciet gemaakt kan worden *waarom* er bepaalde statistische invloeden tussen hypothesen en bewijsmiddelen bestaan. Verder zijn er efficiënte computerprogramma's die de constructie van en het rekenen met Bayesiaanse netwerken ondersteunen.

3 Rechters en raadsheren over gebruik van de Bayesiaanse kansrekening in complexe strafzaken

In de rechtspraak wordt de Bayesiaanse kansrekening afgezien van de optredens van Alkemade en Derksen tot nu toe vrijwel uitsluitend gebruikt op deelaspecten en vooral op zogenaamd bronniveau, waarbij het de vraag is van welke bron een bepaald spoor (bijvoorbeeld DNA, voetsporen, vingerafdrukken) afkomstig is. Ook wordt ze wel gebruikt bij mishandelings-, doodslag- en moordzaken om iets te zeggen over mogelijke oorzaken van verwondingen. In civiele rechtsgedingen wordt de kansrekening met enige regelmaat gebruikt bij DNA-bewijs in verwantschapszaken. Al dit gebruik beperkt zich tot het vaststellen en rapporteren van de bewijskracht (*likelihood ratio*) van een specifieke bevinding of groep bevindingen.

Zoals genoemd in de inleiding, is in vier recente strafzaken de toelaatbaarheid van Bayesiaanse analyses van complexe strafzaken aan de orde gekomen. Ik geef nu de oordelen van de rechtbanken en gerechtshoven weer en bespreek hun argumenten tegen Bayesiaanse deskundigenrapportages.

Eerst de Zes van Breda. Hierin had de raadsheer-commissaris mij gevraagd Alkemades rapport en zijn gebruik van de Bayesiaanse methode te beoordelen. Het hof was het met mij eens dat (in de woorden van het hof)

“... de betrouwbaarheid van voormelde methode en daarmee van de uit het gebruik van de methode voortvloeiende uitkomsten, naar de huidige stand van de wetenschap te onzeker is voor daadwerkelijke toepassing in een complexe strafzaak als de onderhavige. Om deze reden heeft het hof de conclusie(s) uit voormelde rapportage van dr. Alkemade bij de beoordeling van het bewijs dan ook geheel buiten beschouwing gelaten.”

Ik kan niet anders dan speculeren over de precieze redenen voor dit oordeel. Ik had in mijn rapport onder meer het volgende geconcludeerd:

De Bayesiaanse methode is, hoewel niet onomstreden, in zijn algemeenheid in ieder geval een plausibele en door een aantal wetenschappers aanbevolen wetenschappelijk methode voor redeneren in bewijskwesities. Gebruik door getuige-deskundigen van likelihood ratio's bij deelaspecten zoals DNA-bewijs komt met enige regelmaat voor in rechtszaken. Maar als algemeen denkmodel voor bewijskwesities in strafzaken is de Bayesiaanse denkwijze controversieel en nog niet of nauwelijks in de praktijk toegepast: er is discussie of ze wel geschikt is als algemeen denkmodel, of dat ze slechts bruikbaar is bij bepaalde deelaspecten van een zaak. Zo is het nog onduidelijk hoe a priori kansen bepaald kunnen worden en hoe Bayesiaanse netwerken gebruikt kunnen worden om goed met statistische afhankelijkheden om te gaan. Er kan dus niet geconcludeerd worden dat Alkemades methode een beproefde wetenschappelijke methode is. Normaal in de wetenschap is dat nieuwe ideeën eerst in een wetenschappelijke context getest en aan *peer review* onderworpen worden, voor ze in de praktijk worden toegepast. Alkemade heeft deze fases van testen en *peer review* overgeslagen en zijn nieuwe ideeën meteen in de praktijk toegepast.

Ik vermoed dat het hof het hier mee eens was. Maar het hof oordeelde ook dat Alkemade “voor de onderhavige procedure als deskundige kan worden aangemerkt” hoewel ik had geconcludeerd dat de kansschattingen van Alkemade niet meer waarde hebben dan die van ieder willekeurig intelligent persoon die zich over de zaak zou kunnen buigen, omdat het voor de betrouwbaarheid van kansschattingen essentieel is dat men deskundigheid heeft op het gebied waarop de kansen geschat moeten worden. Het is opvallend dat het Hof oordeelt dat iemand deskundige is maar dat zijn methode niet deugt.

Belangrijk is hier dat de criteria voor de betrouwbaarheid van kansschattingen buiten de Bayesiaanse kansrekening liggen. Om betrouwbare conclusies uit een Bayesiaanse analyse te kunnen trekken is meer nodig dan alleen wiskundige berekeningen. Als de kansen waarmee wordt gerekend geen objectieve basis hebben, zullen de conclusies ook geen objectieve basis hebben. Het zegt dan niets dat ze wiskundig uit de gegeven kansen volgen. Hier geldt: *garbage in, garbage out*. Dit geldt niet alleen voor kansschattingen maar ook voor de beslissingen die daaraan vooraf gaan: het opstellen van de hypothesen, de selectie van mogelijk relevante bevindingen, de bepaling van statistische afhankelijkheidsrelaties tussen de verschillende bevindingen gegeven de hypothesen en het bepalen of verdere potentiële bevindingen relevant zouden kunnen zijn. Om deze redenen had het hof de claim dat de betrouwbaarheid van de Bayesiaanse methode *naar de huidige stand van de wetenschap* te onzeker is, beter kunnen vervangen door de claim dat de Bayesiaanse methode per definitie

onvoldoende is om de waarheid in een strafzaak vast te stellen; op zijn minst moet ze gecombineerd worden met betrouwbare manieren om het model waarmee Bayesiaans gerekend wordt op te stellen.

Bij toepassing van de Bayesiaanse kansrekening op complexe strafzaken is het doorgaans onmogelijk om kansen op betrouwbare statistieken te baseren (anders dan bij bijvoorbeeld de bepaling van de bewijswaarde van DNA-sporen), vandaar het belang van inhoudelijke deskundigheid. In de Zes van Breda ging het onder meer over de betrouwbaarheid van verklaringen van getuigen of verdachten, de relevantie van haar- of bloedsporen of vingerafdrukken, en medische kwesties met betrekking tot het tijdstip van sterven. En Alkemade is, als klimaatfysicus, op geen enkele van deze zaken deskundig.

Ook in de zaak van de Oosterlandse brandstichtingen vroeg het hof mij om Alkemades gebruik van de Bayesiaanse kansrekening te beoordelen. Ook dit hof was het kennelijk eens met mijn kritiek op Alkemades rapport:

“Het hof heeft het rapport van dr. Alkemade, gelet op de tegenwerpingen van prof.dr.mr. H. Prakken in het tegenonderzoek, niet laten bijdragen aan het oordeel van het hof omtrent de bewezenverklaring.”

Ik vermoed dat deze instemming ook voor mijn conclusies over de waarde van Alkemades kansschattingen gold, maar zeker weten doe ik dat niet.

De rechtbank Zeeland-West-Brabant merkte in haar vonnis in de zaak ECLI:NL:RBZWB:2016:3060 het volgende op over een deskundigenrapport van prof. Ton Derksen waarin de Bayesiaanse kansrekening werd gebruikt:

“De berekening van Derksen waaruit zou blijken dat verdachte – kort gezegd – waarschijnlijk onschuldig is, is gemaakt volgens de theorema van Bayes. Daarmee is een – naar het oordeel van de rechtbank – niet onomstreden regel uit de kansrekening gebruikt voor de strafrechtelijke waarheidsvinding waarvan de uitkomst in hoge mate af hangt van de selectie en waardering van de bewijsmiddelen die door Derksen is gemaakt.

De rechtbank sluit niet uit dat deze methode, mits juist toegepast, een hulpmiddel kan zijn bij de waarheidsvinding, maar de uiteindelijke beantwoording van de vraag of een tenlastegelegd feit wettig en overtuigend kan worden bewezen, is geen vraag die zich leent voor kansrekening. De beantwoording van die vraag is en blijft voorbehouden aan de rechter. Daarbij komt dat de rechtbank in deze zaak een andere selectie en waardering van de bewijsmiddelen maakt dan Derksen.

De rechtbank laat om bovengenoemde redenen de rapporten van Derksen buiten beschouwing bij de beantwoording van de vragen van de artikelen 348 en 350 Sv ...”

De nadruk die de rechtbank legt op de afhankelijkheid van de uitkomst van een Bayesiaanse analyse van de selectie en waardering van de bewijsmiddelen doet vermoeden dat de rechtbank goed begrepen heeft dat de criteria voor de

betrouwbaarheid van kansschattingen buiten de Bayesiaanse kansrekening liggen. Op de opmerking van de rechtbank dat Bayes, mits juist gebruikt, een hulpmiddel kan zijn, kom ik in paragraaf 5 terug.

Tenslotte zet het Gerechtshof 's-Hertogenbosch in ECLI:NL:GHSHE:2018:421 een door de verdediging ingebracht rapport van Derksen met een stevige motivering opzij:

Vervolgens worden in het rapport drie punten genoemd die het waarschijnlijker maken dat het onschuld-scenario juist is en aan het einde van beide delen van het onderzoeksrapport worden de verschillende elementen van beide rapporten gecombineerd en wordt er met gebruik van de Bayesiaanse methode geconcludeerd dat de kans dat de verdachte onschuldig is 98,3 procent is en de kans dat de verdachte schuldig is 1,7 procent.

Echter, de Bayesiaanse methode – een kwantitatieve onderzoeksmethode die het juridisch bewijzen ziet als denken over kansen – is in de literatuur aan veel kritiek onderworpen. Allereerst wordt in deze methode verondersteld dat, gegeven de hypothesen, alle bewijsmiddelen statistisch onafhankelijk van elkaar zijn. Daarnaast is twijfelachtig of de kansen van de onderdelen van de input in de Bayesiaanse methode in betrouwbare getallen zijn uit te drukken. Het gebruik van getallen wekt een schijn van objectiviteit. Het schatten van de a-priori kans van een in te voeren onderdeel kan alleen subjectief ingeschat worden en deze subjectieve kans kan niet objectief getoetst worden. Ten slotte is de vraag of graden van geloof wel aan de axioma's van de kansrekening moeten voldoen. Mede deze argumenten wegende, is het hof in de onderhavige zaak van oordeel, zoals hierboven besproken, dat de invoer van gegevens van de onderhavige zaak zo willekeurig is geweest en de (subjectieve) uitgangspunten vaak onjuist dan wel onnavolgbaar, dat uit deze invoer en derhalve uit de rapportage zoals overgelegd, geen betrouwbare conclusies kunnen worden getrokken.

Het hof laat dan ook om bovengenoemde redenen de rapporten van prof. dr. T. Derksen buiten beschouwing en verwerpt het op deze rapporten gebaseerd, dan wel het daaraan gerefereerde verweer van de verdediging.

Ik bespreek nu de argumenten van het hof tegen Derksens gebruik van Bayesiaans denken. Om te beginnen valt op dat het hofs argument over statistische afhankelijkheid alleen Derksens specifieke gebruik van het theorema van Bayes treft (ook wel *naïve Bayes* genoemd) maar niet bijvoorbeeld het gebruik van Bayesiaanse netwerken, die juist afhankelijkheden kunnen representeren.

Ten tweede, de twijfel van het hof of kansen in betrouwbare getallen zijn uit te drukken is in casu gerechtvaardigd nu Derksen geen specifieke expertise heeft op de punten waarop hij kansen schat. Als wetenschapsfilosoof heeft hij alleen expertise in rationele modellen van redeneren en kennisverwerving maar niet over bijvoorbeeld betrouwbaarheid van verklaringen van een verdachte of omstandigheden bij diefstal en doodslag.

Maar toch is hier meer over te zeggen, in combinatie met de opmerking van het hof dat de priorikansen subjectief ingeschat worden en niet objectief getoetst kunnen

worden. De vraag is wat ‘subjectief’ en ‘objectief’ hier betekent. Het is niet een kwestie van of-of maar van meer-minder. In essentie zijn er drie manieren om kansen te bepalen: op grond van *statistieken*, op basis van *expertkennis* en met een beroep op het *gezond verstand* (of in juridische termen: de algemene ervaringsregels). Alle drie hebben ze enige mate van objectiviteit, waarbij statistieken (mits juist opgesteld en geïnterpreteerd) het meest objectief zijn en het gezond verstand het minst. Maar geen van de drie bronnen is puur subjectief. Bovendien is er onderzoek naar manieren waarop de prior op redelijke gronden bepaald kan worden. Zo hebben Fenton et al. (2017) het idee van de gelegenheids- of ‘*opportunity*’ prior gelanceerd: als bekend is dat de verdachte tijdens het misdrijf op de plaats van het misdrijf was, en een redelijke schatting vertelt dat op dat moment niet meer dan een bepaald aantal (bijvoorbeeld 10) mensen aanwezig kan zijn geweest, dan vinden zij het redelijk om de a priorikans op schuld op 1 op de 10 (dus 10%) te zetten. Hoewel deze methode geen wet van Meden en Perzen is, is ze ook niet puur subjectief. Dus op dit punt lijkt de kritiek van het hof in zijn algemeenheid te ongenueanceerd.

Kan dan niet gezegd worden dat de rechtbank kan profiteren van het gezond verstand van Derksen, dat ongetwijfeld goed ontwikkeld is? Hier rijst een juridisch probleem, want volgens artikel 332 lid 2 van het Wetboek van Strafvordering zijn de algemene ervaringsregels het domein van de rechter. En omdat deskundigen geacht worden alleen over het gebied van hun expertise te rapporteren, valt de analyse van Derksen juridisch tussen de wal en het schip.

Tenslotte, met de vraag of graden van geloof wel aan de axioma’s van de kansrekening moeten voldoen raakt het hof aan een fundamentele filosofische discussie. Deze discussie kent eminente voor- en tegenstanders en het debat is zeker niet beslist (Meester & Kerkvliet 2016; Talbott 2016). Het algemene criterium dat het hof hier kennelijk hanteert is dat als er over een theorie van rationeel bewijzen filosofische of wetenschappelijke onenigheid bestaat, de door deze theorie gepropageerde denkwijze niet in rechtszaken gebruikt mag worden. Maar dan is geen enkele rationale denkwijze toegestaan, want er bestaat nog geen algemeen aanvaarde theorie over rationeel juridisch bewijzen. Dat geldt net zo goed voor argumentatie- en scenariogebaseerde theorieën als voor Bayesiaanse benaderingen. In Paragraaf 4 kom ik op dit punt terug.

Bovendien brengt het hof zich met dit oordeel in de problemen, want in dezelfde zaak heeft het hof wel een NFI-rapportage over een DNA-match gebruikt, waarbij de NFI-deskundige de ‘*random-match probability*’ rapporteerde, oftewel de kans dat het gevonden DNA matcht met dat van de verdachte als het DNA-spoor van een willekeurige persoon afkomstig is. En daarvan is de relevantie voor de bewijsbeslissing niet te begrijpen zonder kennis van de Bayesiaanse kansrekening (Meester & Sjerps 2004). Sterker nog, NFI-rapporten bevatten tegenwoordig doorgaans een verwijzing naar een publicatie *De reeks waarschijnlijkheidstermen van het NFI²*, die uitlegt hoe de door de NFI-deskundigen gebruikte

² Zie

<https://www.forensischinstituut.nl/publicaties/publicaties/2017/10/18/vakbijlage-waarschijnlijkheidstermen>.

waarschijnlijkheidstermen Bayesiaans geïnterpreteerd moeten worden. Vermoedelijk stond deze verwijzing ook in het rapport in casu.

In dat rapport rapporteerde de NFI-deskundige als volgt:

De bemonsteringen AACQ1456NL#01 en #02 van de hals van het slachtoffer [slachtoffer] bevatten celmateriaal dat afkomstig kan zijn van het slachtoffer [slachtoffer] en [verdachte]. De berekende frequentie of matchkans van het DNA-profiel is kleiner dan één op één miljard.

Het hof maakt hiervan:

het feit dat er DNA van de verdachte in de halsstreek van het slachtoffer is aangetroffen

Strikt genomen is dit, nu het hof niets zegt over de a priori kans dat het DNA van de verdachte afkomstig is, een voorbeeld van de *prosecutor fallacy* (zie hierover o.a. de Keijser et al. 2009; Prakken 2018).

Samengevat is het meest overtuigende argument van de rechters en raadsheren dat gebruik van de Bayesiaanse methode niets impliceert over de betrouwbaarheid van de kansschattingen en de inhoudelijke oordelen die daaraan voorafgaan. Het met Bayesiaans denken trekken van conclusies over de schuldvraag door deskundigen past daarom niet in de rechtszaal. Dit betekent overigens niet dat Bayesiaans denken helemaal geen plaats heeft in de rechtszaal. Op dit punt kom ik in Paragraaf 5 terug.

4 Toepassing van de kritiek op andere benaderingen

Een andere vraag is in hoeverre de argumenten van de rechters en raadsheren tegen Bayes ook opgaan voor andere denkmethoden, met name voor argumentatie- en scenario-gebaseerde methodes of combinaties daarvan. Hierover kunnen we kort zijn: hun meest overtuigende argument gaat op voor elke denkmethode: ook in een argumentatie- of scenario-gebaseerde analyse is de betrouwbaarheid van de uitkomst afhankelijk van de betrouwbaarheid wat er in gaat, dus van de generalisaties en bevindingen waarmee argumenten of scenario's geformuleerd en vergeleken worden. Dus de rechters en raadsheren zouden hetzelfde oordeel moeten vellen over analyses van een gehele zaak gemaakt met andere methodes. Maar dat zou vreemd zijn, want dan zou er in de rechtszaal helemaal niet meer rationeel gedacht mogen worden. Deze conclusie is natuurlijk absurd; de juiste conclusie is dat zowel argumentatie als scenario- en Bayesiaans denken redeneermethoden zijn, dat wil zeggen, methodes om conclusies te trekken uit bepaalde informatie (bevindingen, wereldkennis, wetenschappelijke kennis, statistieken ...). En elke manier van redeneren moet gecombineerd worden met methodes om de uitgangspunten van de redeningen op betrouwbare wijze vast te stellen.

De waarde van rationele redeneermethodes is gelegen in het bevorderen van rationeel denken en het voorkomen van denkfouten. Alle drie benaderingen hebben op dit punt hun waarde. Hierover merkt Van Koppen (2011, p. 245) op: "Welke benadering ook

gekozen wordt (verhaal, argumentatie, statistiek), in al die benaderingen heeft het afwegen van alternatieven een prominente plaats, en zo hoort het ook.” (p. 245). Dit is inderdaad een sterk punt van alle drie benaderingen, onder meer bij het voorkomen van tunnelvisie. Verder hebben alle drie benaderingen hun eigen sterke en zwakke punten. De argumentatieaanpak is sterk in het expliciet maken hoe bewijsconclusies gebaseerd worden op bevindingen en wat daarin de zwakke plekken zijn. De scenarioaanpak is sterk in het geven van samenhang en context aan een grote hoeveelheid bewijs. En de Bayesiaanse kansrekening is onmisbaar bij het interpreteren van door deskundigen gerapporteerde *likelihood ratios* en *random-match probabilities*, en bij het vermijden van statistische denkfouten bij het interpreteren van kansen en statistieken. Om deze redenen is de vraag niet welke redeneermethode bij uitsluiting van de andere methodes gebruikt moet worden maar hoe de methodes het beste gecombineerd kunnen worden. Mijn huidige mening is dat rechters in complexe zaken het beste kunnen beginnen met het schetsen van scenario's, en dan bij de kritische analyse daarvan kunnen inzoomen op deelaspecten met argumentatie of Bayes. Het gebruik van Bayes als algemeen denkmodel zou voor de meeste juristen niet te begrijpen zijn en brengt het gevaar met zich mee dat het gebruik van wiskunde de ontorechte schijn wekt van objectiviteit en wetenschappelijkheid. En argumentatie als overkoepelend denkmodel is ongeschikt om overzicht over grote hoeveelheden bewijs te bieden.

Tot slot van deze paragraaf, als het meest overtuigende argument van de rechters en raadsheren ook opgaat voor andere denkwijzen, waarom zijn er dan geen soortgelijke rechterlijke oordelen over argumentatie- en scenario gebaseerde denkwijzen? Een reden kan zijn dat de rechters mijn zorgen over de onbegrijpelijkheid en het gevaar van de ontorechte schijn van objectiviteit van Bayesiaanse rapportages delen. Een andere reden kan zijn dat de opmerkelijke zelfverzekerdheid van Alkemade en Derksen bij het schatten van kansen en dergelijke de rechters en raadsheren niet bevalt. Maar de meest voor de hand liggende reden is dat we die zelfverzekerdheid niet zien bij proponenten van de andere benaderingen, tenminste niet als deskundige in de rechtszaal. Voor zover mij bekend heeft nog nooit iemand geprobeerd om als deskundige in een rechtszaak een complexe zaak in zijn geheel met argumentatie- of scenario-denken te analyseren. Bij deze benaderingen komt kennelijk minder snel het onjuiste idee op dat expertise in deze denkwijzen expertise in de aspecten van een strafzaak impliceert (vgl. Prakken & Meester 2017).

Dit alles laat onverlet dat filosofen, statistici, rechtspsychologen en anderen vrij zijn in het geven van analyses van complexe zaken buiten de rechtszaal. En bovendien kan dat belangrijke resultaten opleveren. Van Koppens *finest hour* was misschien wel zijn analyse van de Schiedammer parkmoord in Van Koppen (2003), die tot het herstellen van een geruchtmakende gerechtelijke dwaling heeft geleid. En Derksen heeft een belangrijke rol gespeeld bij het herstellen van een andere geruchtmakende gerechtelijke dwaling, in de zaak Lucia de Berk (Derksen 2006). Momenteel ben ik betrokken bij een experiment waarin een groep wetenschappers uit verschillende landen (waaronder Peter van Koppen) hun eigen favoriete benadering toepassen op een afgesloten rechtszaak, het hoger beroep in de Simonshavenzaak (Gerechtshof Den Haag 18 februari 2015, ECLI:NL:HR:2016:2201). Dit doen we niet om de waarheid in de zaak vast te stellen maar om meer inzicht te krijgen in de sterke en zwakke punten van de verschillende benaderingen. De analyses zullen verschijnen in een

speciaal nummer van het tijdschrift *Topics in Cognitive Science* (Prakken, Bex & Mackor 2020) en zijn al online beschikbaar.³

5 Toepasbaarheid van Bayesiaans denken in de rechtszaal

Laten we teruggaan naar rechtbank Zeeland-West-Brabant, die opmerkte “De rechtbank sluit niet uit dat deze methode, mits juist toegepast, een hulpmiddel kan zijn bij de waarheidsvinding.” Het kan zijn dat de rechtbank hier doelt op het gebruik van de Bayesiaanse kansrekening op deelaspecten, zoals het bepalen van de bewijswaarde van DNA- of ander sporenbewijs bij de vraag wat de bron is van een bepaald spoor. Daar kan ik het alleen maar mee eens zijn. Uit meerdere onderzoeken, onder andere van mijzelf, is gebleken dat rechters en andere juristen kansinformatie in dit soort kwesties op grote schaal verkeerd interpreteren (bijv. Lindsey et al. 2003; de Keijser et al. 2009; Prakken 2018) en hierboven hebben we daar weer een voorbeeld van gezien. Het kan ook zijn dat de rechter doelt op de educatieve of zelfs ‘therapeutische’ waarde van Bayesiaans denken bij het voorkomen van veel voorkomende statistische drogredenen tegen te gaan. Zo hebben veel specialisten in de kansrekening of statistiek, bijvoorbeeld Derksen 2006, Meester, Gill (Meester et al. 2006), Grünwald (2011) en in het buitenland Thompson & Schumann (1987), Fenton & Neil (2011) en Dahlman et al. (2016), binnen of buiten de rechtszaal veel gedaan om statistische drogredenen te ontmaskeren. Om deze redenen is onderwijs en training van Bayesiaans denken aan rechtenstudenten en praktijkjuristen erg belangrijk. Met kleine voorbeeldjes kan betrekkelijk eenvoudig getoond worden dat bepaalde verleidelijke denkpatronen inzake kansen incorrect zijn (Berger & Aben 2010).

Maar er is nog een vorm van toepassing van Bayesiaans denken in de rechtszaal, namelijk het op een praktische manier gebruiken van aan wetenschappelijk onderzoek ontleende inzichten. Een klassieker is het inzicht dat de betrouwbaarheid van verklaringen van ooggetuigen afneemt naarmate de gebeurtenis waarover verklaard wordt zeldzamer is. In Bayesiaanse termen betekent is dat de a priorikans dat de gebeurtenis waar is, afneemt. Een recent voorbeeld is het werk van de Zoete et al. (2014) aan *crime linkage*, het voorkomen van meerdere soortgelijke gevallen die nabij zijn in plaats en/of tijd. Een voorbeeld zijn de meerdere sterfgevallen van doodzieke kinderen op de intensive care wanneer Lucia de Berk dienst had. Door de rechtbank is hier de constructie van ketenbewijs toegepast, waarbij het (verondersteld) bewezen zijn dat één van de kinderen door Lucia de Berk vermoord was, gebruikt werd om datzelfde ook in de andere gevallen bewezen te achten. De Zoete et al. laten (in het algemeen, niet in de context van de zaak Lucia de Berk) overtuigend met Bayesiaanse netwerken zien dat ketenbewijs twee kanten opwerkt: niet alleen kan belastend bewijs in één geval de schuldconclusie in de gelinkte gevallen aannemelijker maken, maar ook kan *ontlastend* bewijs in één geval de *onschuldconclusie* in de gelinkte gevallen aannemelijker maken. Zo’n Bayesiaans inzicht kan door rechters in nieuwe gevallen gebruikt worden zonder dat ze daarvoor expliciet Bayesiaans hoeven redeneren of rekenen.

³ <https://onlinelibrary.wiley.com/toc/17568765/0/0>

Een laatste voorbeeld heeft te maken met de soms voorkomende twijfel bij rechters als het bewijs niet alleen door een schuldscenario maar ook door een door de verdediging onschuldscenario verklaard wordt. Als de verdediging het onschuldscenario laat inbrengen, kan er vaak niet meer naar discriminerend bewijs gezocht worden. In een aantal zaken was dat aanleiding voor de rechter om de verdachten vrij te spreken (zie Stevens 2017 voor een bespreking van deze zaken). Maar kennis van de Bayesiaanse kansrekening had de rechters kunnen vertellen dat niet alleen relevant is of een scenario het bewijs verklaart (onderdeel van de *likelihood ratio*) maar ook hoe plausibel het scenario is (onderdeel van de a-priorikansverhouding). Zelfs als een schuld- en onschuldscenario het beschikbare bewijs evengoed verklaren, kan het schuldscenario na beschouwing van het bewijs veel waarschijnlijker zijn dan het onschuldscenario, omdat het a priori veel waarschijnlijker is dan het onschuldscenario. Ook dit Bayesiaanse inzicht vereist niet dat de rechter expliciet Bayesiaans denkt of rekent.

Voor de wetenschap betekent dit dat wetenschappelijk onderzoek naar juridische toepassingen van Bayesiaans denken zeer waardevol kan zijn: niet zozeer om de rechter een denkmethode aan de hand te doen maar om theoretische inzichten in 'hapklare vorm' aan de rechtspraak aan te bieden, zodat rechters en raadsheren daar ook als ze argumentatief of scenariogebaseerd denken hun voordeel mee kunnen doen.

6 Conclusie

In dit artikel heb ik vier uitspraken van rechtbanken en gerechtshoven geanalyseerd waarin gebruik van Bayesiaans denken door deskundigen om een strafzaak in zijn volle omvang te analyseren bekritiseerd wordt. Ik heb betoogd dat niet alle argumenten van de rechters en raadsheren even sterk zijn, maar dat hun sterkste argument, dat gebruik van de Bayesiaanse methode niets impliceert over de betrouwbaarheid van de kansschattingen en de inhoudelijke oordelen die daaraan voorafgaan, hun kritiek op de bewuste rapportages in hun zaken rechtvaardigt. Maar ook heb ik betoogd dat dit sterkste argument ook opgaat voor alternatieve redeneermethoden, zoals argumentatieve en scenariobenaderingen. Dit argument treft daarom niet de Bayesiaanse methode op zich maar het specifieke gebruik daarvan door sommige deskundigen in rechtsgedingen. Tot slot heb ik drie manieren gesuggereerd waarop Bayesiaans denken nuttig kan zijn in of voor rechtszaken: voor een juist begrip van kansrapportages omtrent specifieke (groepen) bevindingen door forensische deskundigen, als 'antiserum' tegen statische redeneerfouten, en als wetenschappelijke bron van praktisch relevante inzichten met betrekking tot specifieke bewijsaspecten.

Personalia

Henry Prakken is hoogleraar rechtsinformatica en juridische argumentatie aan de Faculteit der Rechtsgeleerdheid, Rijksuniversiteit Groningen en universitair docent kunstmatige intelligentie aan de Faculteit Bètawetenschappen, Departement Informatica, Universiteit Utrecht. Hij was samen met Peter van Koppen promotor van Floris Bex in Groningen. Dit resulteerde in hun tot dusver enige gezamenlijke

publicatie, als co-auteurs van Bex et al. (2010). De laatste jaren werkten Prakken en Van Koppen met anderen samen in het case-studyproject van Prakken et al. (2020).

Literatuur

Anderson, T.J., Schum, D.A. & Twining, W.L. (2005). *Analysis of evidence*, second edition. Cambridge: Cambridge University Press.

Bennett, W.L. & Feldman, M.S. (1981). *Reconstructing reality in the courtroom: Justice and judgment in American culture*. London: Methuen–Tavistock.

Berger, C.E.H. & Aben, D.J.C. (2010). Bewijs en overtuiging: Een helder zicht op valkuilen. *Expertise en Recht* 2010-5/6, 159-165.

Bex, F.J., Koppen, P.J. van, Prakken, H. & Verheij, B. (2010). A Hybrid formal theory of arguments, stories and criminal evidence. *Artificial Intelligence and Law* 18, 123–152.

Bex, F.J., Prakken, H., Reed, C. & Walton, D.N. (2003). Towards a formal account of reasoning about evidence: argumentation schemes and generalisations. *Artificial Intelligence and Law* 11, 125-165.

Crombag, H.F.M., Koppen, P.J. van & Wagenaar, W.A. (1994). *Dubieuze zaken: de psychologie van strafrechtelijk bewijs* 2^e druk, Contact, Amsterdam.

Dahlman, C., Zenker, F., & Sarwar, F. (2016). Miss rate neglect in legal evidence. *Law, Probability & Risk* 15, 239-250.

Derksen, T. (2006). *Lucia de Berk: reconstructie van een gerechtelijke dwaling*. Diemen: Uitgeverij Veen Magazines B.V.

Fenton, N.E. & Neil, M. (2011). Avoiding legal fallacies in practice using Bayesian networks, *Australian Journal of Legal Philosophy* 36, 114-151.

Fenton, N. E., Neil, M., & Berger, D. (2016). Bayes and the law. *Annual Review of Statistics and Its Application* 3 (1), 51-77.

Fenton, N. E., Lagnado, D. A., Dahlman, C., & Neil, M. (2017). The opportunity prior: a simple and practical solution to the prior probability problem for legal cases. In *Proceedings of the 16th International Conference on Artificial Intelligence and Law* (pp. 69-76). New York: ACM Press.

Grünwald, P. (2011). Over het bedrijven van statistiek in kansloze situaties. Voordracht Zwolle, 18 mei 2011. <http://www.luciadeb.nl/metta/Grunwalds-Metta-analyse.pdf>.

Keijser, J.W. de, Elffers, H., Kok, R.M. & Sjerps, M.J. (2009). *Bijkans begrepen? – feitelijk en vermeend begrip van forensische deskundigenrapportages onder rechters, advocaten en deskundigen*, Den Haag: Boom Juridische Uitgevers.

Koppen, P.J. van (2003). *De Schiedammer parkmoord. Een rechtspsychologische reconstructie*. Juridische Uitgeverij Ars Aequi.

Koppen, P.J. van (2011). *Overtuigend bewijs. Indammen van rechterlijke dwalingen*. Amsterdam: Nieuw Amsterdam 2011.

- Koppen, P.J. van (2017a). De tengelhamer en het schedeldak. Over rapportages door het Nederlands Forensisch Instituut, *Nederlands Juristenblad* 2017/1121.
- Koppen, P.J. van (2017b). Hoe onlogisch werken toch logisch kan worden gevonden, *Nederlands Juristenblad* 2017/2141.
- Lindsey, S. Hertwig, R. & Gigerenzer, G. (2003). Communicating statistical DNA evidence. *Jurimetrics* 43, 147-163.
- Lipton, P. (2004). *Inference to the best explanation*. London: Routledge, 2nd edition.
- Meester, R., Collins, M., Gill, R. & Lambalgen, M. van (2006). On the (ab)use of statistics in the legal case against the nurse Lucia de B. *Law, Probability & Risk* 5, 233-250.
- Meester, R. & Kerkvliet, T. (2016) Assessing forensic evidence by computing belief functions. *Law, Probability & Risk* 15, 127-153.
- Meester, R. & Sjerps, M. (2004). Why the effect of prior odds should accompany the likelihood ratio when reporting DNA evidence. *Law, Probability & Risk* 3, 51-62.
- Pennington, N., and Hastie, R. (1993). The story model for juror decision making. In Hastie, R. (eds.), *Inside the juror, the psychology of juror decision making*. Cambridge: Cambridge University Press.
- Prakken, H. (2004). Analysing reasoning about evidence with formal models of argumentation. *Law, Probability & Risk* 3, 33-50.
- Prakken, H. (2014) Strafrechtelijk bewijzen: met Bayes of met verhalen? Of is er een derde weg? *Expertise en Recht* 2014-1, 4-19
- Prakken, H. (2018). Kansoordelen door deskundigen: over ‘logisch’ rapporteren en wat daarbij mis kan gaan. *Ars Aequi* 67, 740-747.
- Prakken, H., Bex, F.J. & Mackor, A.R. (red.) (2020). *Models of rational proof in criminal law*. Speciaal nummer van *Topics in Cognitive Science*, te verschijnen.
- H. Prakken & R. Meester (2017), Bayesiaanse analyses van complexe strafzaken door deskundigen. Betrouwbaar en zo ja: nuttig? *Expertise en Recht* 2017-5, 185-197.
- Sjerps, M. & Berger, C.E.H. (2011). Het Bayesiaanse model biedt een helder zicht op een complexe werkelijkheid. Den Haag: Nederlands Forensisch Instituut.
- Sjerps, M., Kloosterman, A & Berger, C.E.H. (2017). Over de rapportage van het NFI: een weerwoord. *Nederlands Juristenblad* 2017/2140.
- Stevens, L. (2017). *Onterechte vrijspraken? Over hoe de rechter redeneert met een alternatief scenario van de verdachte*. Den Haag: Boom Juridisch.
- Thompson, W.C. & Schumann, E.L. (1987). Interpretation of statistical evidence in criminal trials. The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior* 11, 167-187.
- Talbott, W. (2016). Bayesian epistemology. In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.
- Verheij, B. (2000), Dialectical argumentation as a heuristic for courtroom decision-making. In P.J. van Koppen and N.H.M. Roos (eds.): *Rationality, information and*

progress in law and psychology. Liber amoricum Hans F. Crombag (203-226).
Maastricht: Metajuridica Publications.

Wigmore, J.H. (1931). *The principles of judicial proof or the process of proof as given by logic, psychology, and general Experience, and illustrated in judicial trials*, 2nd edn. Boston: Little Brown and Company.

Zoete, J.C. de, Sjerps, M.J., Lagnado, D.A. & Fenton, N.E. (2014). Modelling crime linkage with Bayesian networks. *Science & Justice* 56, 209-217.