

Ronald Meester – Lonneke Stevens

REDENEREN ONDER ONZEKERHEID IN STRAFZAKEN



Hoe een
Bayesiaanse
benadering
behulpzaam
kan zijn



Redeneren onder onzekerheid in strafzaken

Ronald Meester en Lonneke Stevens

REDENEREN ONDER ONZEKERHEID IN STRAFZAKEN

Hoe een Bayesiaanse benadering
behulpzaam kan zijn

Opmaak binnenwerk: Crius Group, Hulshout
Omslagontwerp: Nick van Silfhout, Den Haag

© 2025 De auteurs | WJS uitgevers

Behoudens de in of krachtens de Auteurswet gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige ander manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van reprografische verveelvoudigingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht (Postbus 3051, 2130 KB Hoofddorp, www.reprorecht.nl). Voor het overnemen van (een) gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (art. 16 Auteurswet) kan men zich wenden tot de Stichting PRO (Stichting Publicatie- en Reproductierechten Organisatie, Postbus 3060, 2130 KB Hoofddorp, www.stichting-pro.nl).

No part of this book may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission from the publisher.

ISBN 978-94-93458-03-1
NUR 820

wjs-uitgevers.nl

INHOUD

Hoofdstuk 1. Inleiding	7
Casuïstiek	11
Hoe zien wij kansrekening in strafzaken?	11
Modellen van de werkelijkheid	13
Redeneermodellen	15
Een overzicht van het boek	18
 Hoofdstuk 2. Het Bayesiaanse redeneermodel – theorie en voorbeelden	 23
Bayesiaans redeneren; een mens begint niet bij nul	28
Bewijs moet relatief worden beoordeeld	30
Oordeelsvorming	42
Het redeneermodel samengevat in gangbare terminologie	50
Praktische uitvoering – voorbeelden	51
 Hoofdstuk 3. De rol van getallen en modellen	 59
Kansen en informatie	60
Modellen	65
Kansen in modellen – voorbeelden	68
Numerieke waarden voor de prior?	77
Combineren van bewijs	85
De praktijk	93
Kennisleer	95

Hoofdstuk 4. Bayesiaans redeneren en strafrechtelijk bewijsrecht	101
De onschuldpresumptie	101
Persoonlijke overtuigingen	105
Schakelbewijs	109
Zwijgrecht, proceshouding en Bayes	113
De bewijsminimumregel en Bayes	117
Hoofdstuk 5. Misverstanden en problemen	123
De stand van de wetenschap en Bayes	124
De mythe dat er maar één likelihood ratio in een zaak bestaat	128
Een kleine kans zegt weinig	131
Is een aangetroffen DNA-profiel beslissend?	134
Moet een kansuitspraak in een categorische stelling worden omgezet?	136
De prosecutor's fallacy	138
Drie visies op wat Bayesiaans redeneren is	140
De aard van onze onzekerheden	142
Hoofdstuk 6. Analyse van een aantal zaken	147
De zaak ECLI:NL:RBZWB:2021:2638; een levensdelict met tegenstrijdige aanwijzingen en vrijspraak	148
De zaak ECLI:NL:GHARL:2020:10710; de 'Thialf zaak'	163
De zaken ECLI:NL:RBNHO:2014:7224 en ECLI:NL:GHAMS:2017:1801; de overleden vrouw in het trappenhuis	175
Ten slotte	185

HOOFDSTUK 1

Inleiding

Rechtszaken gaan van nature gepaard met onzekerheid. In de meeste gevallen is het op zijn minst voor enkele personen in een proces onduidelijk wat zich precies heeft afgespeeld, en dient het proces onder meer om daar juist helderheid over te krijgen. In die zin lijkt het schrijven van een boek met de titel “Redeneren onder onzekerheid in strafzaken” op het schrijven over ‘schaatsen op ijs’ – het ijs en de onzekerheid zijn eigenlijk overbodig. Het feit dat dit boek toch voor u ligt kan echter niets anders betekenen dan dat het volgens ons toch nodig is om erover te reflecteren. Onze eerste taak is om de lezer daarvan te overtuigen.

Als er geen onzekerheden in het spel zijn in een strafzaak, dan is het trekken van conclusies met behulp van huis-tuinen-keukenlogica niet zo ingewikkeld. Als het bijvoorbeeld zeker is dat de overvaller een man was, dan is het ook zeker dat elke vrouw automatisch vrijgepleit is. Als het zeker is dat de verdachte dezelfde persoon is als de inbreker die op camerabeelden is vastgelegd, dan is het ook zeker dat de verdachte schuldig is. Als het zeker is dat een verdachte een alibi heeft, dan is het ook zeker dat hij of zij niet schuldig kan zijn. Als een getuige zeker de waarheid spreekt en persoon A aanwijst als de inbreker, dan is A de inbreker.¹

Maar helaas, zo zeker is het vrijwel nooit, en de rechter of raadsheer zal het in de rechtszaal vaak moeten doen met

¹ Wij zullen niet altijd ‘hij of zij’ schrijven maar ons beperken tot ‘hij’ waarmee elke andere gender nadrukkelijk geïmpliceerd is.

beweringen waarvan hij niet zeker kan weten of ze waar zijn of niet. Er kan, bijvoorbeeld, twijfel bestaan over het waarheidsgehalte van een getuigenverklaring. Er bestaan, uiteraard, ook gradaties van onzekerheid, en het is wellicht goed om over die gradaties nu alvast iets te zeggen.

Het lijkt misschien aantrekkelijk om te denken dat onzekerheid altijd in een getal of percentage uit te drukken zou zijn. Je zou er bijvoorbeeld voor 90% van overtuigd kunnen zijn dat een bepaalde getuige de waarheid spreekt. Hoe je tot een dergelijk percentage zou moeten komen, is echter niet zo duidelijk, en wat het precies betekent misschien nog wel minder. Onze ervaring leert in ieder geval dat dergelijke getallen zelden zinvol zijn en meer verwarring veroorzaken dan ze oplossen; we komen daar nog uitgebreid over te spreken. Veel meer dan met getallen zullen we op een kwalitatieve manier spreken over onzekerheid; zo kunnen we bijvoorbeeld denken dat een bepaalde gebeurtenis ‘waarschijnlijk’ waar is, waarbij we noodgedwongen in het midden laten wat dit exact betekent voor degene die het uitspreekt of opschrijft.

Redeneren onder onzekerheid behelst dus meer dan het manipuleren met getallen of formules – die laatste spelen vrijwel nooit een rol, en u zult in dit boek dan ook geen enkele formule aantreffen. Redeneren onder onzekerheid betekent niet dat we kansen gaan uitrekenen op bepaalde gebeurtenissen. Het gaat erom uit te leggen welke gevolgtrekkingen wel of juist niet af te leiden zijn van de data als je over bepaalde aspecten daarvan onzeker bent. Hoe vreemd het ook klinkt, het is soms helemaal niet zo duidelijk of een bepaald bewijsmiddel belastend of ontlastend is, en onze intuïtie kan ons daarbij behoorlijk in de steek laten. Het is daarom belangrijk om het redeneren onder onzekerheid te structureren, en dat is precies wat we in dit boek zullen doen. We willen daarbij vanaf het eerste moment benadrukken dat dit niet betekent

dat we rechtspreken willen of kunnen vervangen door kansrekening. Zoals we uitgebreid zullen betogen is dat onmogelijk en ook onwenselijk.

Waar kan een rechter² in dit verband zoal tegenaan lopen? We geven twee voorbeelden. Sommige gegevens (vaak ‘data’ genoemd) zijn simpelweg onmogelijk onder bepaalde veronderstellingen. Als bijvoorbeeld op de plaats delict een DNA-profiel wordt aangetroffen dat niet matcht met verdachte A, dan kunnen we concluderen dat A niet de donor van het profiel kan zijn. Dat is elementaire logica: als A de donor is, dan is deze data gewoon onmogelijk en dus kan A de donor niet zijn.³

Maar stel nu eens dat we alleen maar zouden weten dat het DNA-profiel *onwaarschijnlijk* is als A de donor is. Dat lijkt misschien vreemd in de context van DNA, want is het niet zo dat DNA min of meer uniek is? Hoe kan er dan onzekerheid bestaan over het donorschap van A? Welnu, een eenvoudige situatie waarbij dat wel degelijk het geval kan zijn, is dat we op de plaats delict een zogenaamd mengprofiel aantreffen waar DNA van zowel het slachtoffer als de dader in zit. Omdat je niet zomaar kunt zien welk DNA van wie afkomstig is, kan dit extra onzekerheid genereren en kan het donorschap van A onzeker zijn.

Maar hoe dan ook, als het profiel onwaarschijnlijk is als A de donor zou zijn, kunnen we dan ook concluderen dat het onwaarschijnlijk is dat A de donor is? Het antwoord, u raadt het al, is ‘nee’. Hoewel ‘onwaarschijnlijk’ op een bepaalde manier dicht bij ‘onmogelijk’ ligt, verandert de logica werkelijk totaal als we overstappen van ‘onmogelijk’ naar ‘onwaar-

² Wanneer we ‘rechter’ schrijven bedoelen we vaak ‘rechter of raadsheer’. Dit bevordert de leesbaarheid.

³ Als we mee zouden nemen dat DNA kan muteren, dan is zelfs dit niet helemaal waar, maar deze mogelijkheid negeren we hier.

schijnlijk'. We denken dat de meeste lezers zich nu afvragen hoe dat precies zit, en het feit dat het antwoord op die vraag niet evident is, is al voldoende reden voor ons om het boek te schrijven en voor u om het te lezen.

Het tweede voorbeeld is van een heel andere aard. Het is bekend dat getuigen zich kunnen vergissen, en dat dit globaal gesproken met een bepaalde frequentie gebeurt. Stel dat uit onderzoek zou blijken dat 20% van de getuigenverklaringen onbetrouwbaar zou zijn, hoe moet de rechter met dit gegeven omgaan bij een specifieke getuigenverklaring? Speelt die 20% dan nog een rol en zo ja, welke dan precies?⁴

Dit boek gaat over dit soort dilemma's en vragen. Het wordt daarmee, zoals gezegd, zeker geen technisch boek over kansrekening en statistiek. Beide vakken hebben een bepaalde reputatie – ze zouden moeilijk zijn, en niet geschikt voor alfa's zoals juristen, die er ook bepaald niet voor warmlopen. Iedereen begrijpt ook dat rechtspreken iets anders is dan alleen maar kansrekening en statistiek, dus we zullen goed moeten nadenken op welke manier de kanstheoretische logica succesvol ingezet kan worden in rechtszaken.

Uiteindelijk is de wiskunde die achter kansrekening en statistiek zit, niets anders dan een formele weergave van behoorlijk intuïtieve patronen; patronen die we helder kunnen krijgen zonder de ballast van formules en rekenpartijen te gebruiken. Het heeft geen enkele zin om in dit boek de formele regels van de kansrekening te bespreken, en we gaan het ook niet hebben over technisch statistische bespiegelingen. En het aardige is: dat is ook helemaal niet nodig.

⁴ Die 20% is fantasie, we noemen een willekeurig getal om ons punt te kunnen maken.

CASUÏSTIEK

The proof of the pudding is in the eating, dus de beste manier om ons punt te maken is door te laten zien dat wat wij te zeggen hebben in de praktijk werkt. Casuïstiek dus. Dat betekent dat u in dit boek veel concrete voorbeelden tegen zult komen. En met ‘concrete voorbeelden’ bedoelen we echt concreet in de zin van reële zaken. Maar al te vaak worden bespiegelingen over de logica die men in de rechtszaal geacht wordt te gebruiken, gepresenteerd aan de hand van irreële en kunstmatige voorbeelden. Maar als de voorbeelden kunstmatig zijn, dan kan onmogelijk duidelijk worden hoe je de besproken onderwerpen en concepten in de werkelijkheid zou kunnen gebruiken.

Het is niet voor niets dat het schrijversduo van dit boek bestaat uit een wiskundige en een juriste. We bevinden ons nadrukkelijk buiten de spreekwoordelijke ivoren toren en hebben oog voor wat er in de rechtszaal echt gebeurt.

HOE ZIEN WIJ KANSREKENING IN STRAFZAKEN?

Zonder filosofie vaart niemand wel, je moet weten waar je het over hebt. Maar filosofie wordt zinloos als deze niet uiteindelijk bevredigend toegepast kan worden. Geen enkele bespreking van een bepaalde methodiek of bepaald gezichtspunt kan zonder een fatsoenlijke filosofie.. Wij zijn ons scherp bewust van het feit dat rechtsfilosofen niet altijd gericht zijn op praktische toepasbaarheid – dat zijn wij echter nadrukkelijk wel. Het is simpelweg zo dat je moet weten waar je over praat wanneer je woorden als ‘(on)waarschijnlijk’, ‘kans’ en ‘onzekerheid’ in de mond neemt. Wij vinden het belangrijk om te reflecteren op wat het betekent om wetenschappelijke bevindingen in je uitspraak te verwerken,

en dat vraagt om een filosofie van de kansrekening. Wat betekenen kansuitspraken, en hoe kan een rechter er op de juiste manier gebruik van maken? Het begrip ‘kans’ kan heel verschillende dingen betekenen, en het is belangrijk om die verschillende betekenissen goed van elkaar te kunnen onderscheiden.

Ooit kreeg een van ons (RM) van een bekende advocaat te horen dat hij kansrekening en statistiek tot op het bot wantrouwde. Deze advocaat dacht dat je met statistiek zo ongeveer alles kon aantonen wat je wilde, en daar wilde hij zijn cliënt niet aan blootstellen. Een beter argument voor de noodzaak van dit boek is moeilijk te vinden. De advocaat had namelijk ten dele gelijk: statistiek kan voor veel dingen worden gebruikt en misbruikt. Maar het is juist de bedoeling dat een rechter dit kan doorzien, en dat hij in staat is om te begrijpen wat er aan de hand zou kunnen zijn. Met het juiste inzicht kun je de juiste vragen stellen. Dit vraagt om kennis over het gebruik van redeneringen waarbij onzekerheid een rol speelt.

Nu we het toch over statistiek en statistische gegevens hebben, is het goed om ook even stil te staan bij de interpretatie ervan. Duiding is vaak lastig en is niet terug te brengen tot een algoritmische benadering: het vergt ervaring, veel kennis en vooral wijsheid om numerieke gegevens te interpreteren. We zullen daarom de verschillende statistische paradigma’s langslopen en nagaan hoe die zich verhouden tot rechtszaken. Het is een spannend onderdeel van het metier, waarin wetenschapsfilosofie ook een rol speelt.

We schreven het al eerder: redeneren onder onzekerheid en de bijbehorende kansrekening is niet altijd even eenvoudig. Maar uiteindelijk is dat onder de knie te krijgen door het volgen van een bepaalde manier van redeneren: niet omdat iemand je dat geleerd heeft maar omdat je de logica ervan

inziet. Deze vaardigheid komt niemand vanzelf aanwaaien, je moet er iets voor doen, maar wiskunde is geen magie: de wiskunde en de manier waarop wij mensen redeneren is natuurlijk en intuïtief. Bedenk dat de wiskunde altijd de realiteit volgt, en niet andersom. De wiskunde is geen magisch domein waar de werkelijkheid zich aan dient te onderwerpen.

MODELLEN VAN DE WERKELIJKHEID

De belangstelling voor het onderwerp van dit boek is de afgelopen decennia gegroeid in Nederland en zeker ook in het buitenland. Deze groeiende belangstelling vindt, zo denken wij, haar oorsprong in het debacle in de zaak tegen de verpleegster Lucia de Berk, die twee decennia geleden verdacht werd van het ombrengen van enkele van haar patiëntjes. Een statisticus rekende uit dat de kans dat een onschuldige verpleegster zo veel sterfgevallen zou meemaken als Lucia de Berk in feite had meegemaakt, ongeveer 1 op de 342 miljoen zou zijn. De rechtbank accepteerde deze conclusie.⁵ Later werd echter duidelijk dat dit getal van 1 op 342 miljoen ten eerste volkomen arbitrair was en ten tweede weinig verband hield met de vraag of Lucia de Berk schuldig zou zijn of niet. Om dat in te zien is het nodig om te begrijpen hoe de redeneringen in deze casus precies werkten, en na het lezen van dit boek zal de lezer dat zelf (hopelijk) kunnen begrijpen en zelf kunnen nagaan.

Maar hoe dan ook, na deze dwaling groeide de belangstelling voor het redeneren onder onzekerheid. Als we het

⁵ Rb. 's-Gravenhage 24 maart 2003, ECLI:NL:RBSGR:2003:AF6172. In de latere veroordeling door het hof speelde de kansrekening geen expliciete rol in de bewijsvoering.

daarover hebben, onderscheiden we vier deelonderwerpen die we hierboven al aanstipten:

1. Hoe kunnen we ervoor zorgen dat we geen logische fouten maken als we proberen te redeneren onder onzekerheid?
2. Hoe kunnen we ervoor zorgen dat we numerieke gegevens op de juiste manier interpreteren?
3. Als we uitspraken doen over onzekere situaties, hoe moeten we die uitspraken zélf verstaan?
4. Hoe moeten we modellen van de werkelijkheid gebruiken en interpreteren?

Bij het vierde punt moeten we even stilstaan. Wat bedoelen we precies met een model van de werkelijkheid? Een model is een poging om een aantal facetten en grootheden in een bepaalde situatie aan elkaar te relateren met behulp van wiskunde. In de zaak Lucia de Berk kun je bijvoorbeeld een model maken waarin het zo is dat er in elke dienst van 8 uur een vaste kans is dat er een levensbedreigend incident plaatsvindt. Je kunt in je model bijvoorbeeld ook opnemen dat het optreden van een incident in de ene dienst niet betekent dat zo'n incident in een andere dienst meer of minder waarschijnlijk wordt. Dat zijn allemaal modelaannames.

In zo'n wiskundig model kun je dan gaan rekenen, en dat is precies wat er in deze casus gebeurde. Het zal (en daar zullen we uitgebreid op ingaan) voor de meeste mensen wel duidelijk zijn dat de getallen die uit zo'n berekening komen altijd met een flinke korrel zout genomen moeten worden – het model is immers de werkelijkheid niet, en de uitkomst zal in de meeste gevallen cruciaal afhangen van de keuzes die de modelmaker heeft gemaakt.

Het is daarom goed om je te realiseren dat élk getal, élk kans die daadwerkelijk uitgerekend wordt, een model veronderstelt – er zijn geen uitzonderingen. Zonder wiskundig

model kun je eenvoudigweg niet rekenen omdat elke berekening per definitie in een model plaatsvindt. De wereld is geen wiskundige arena, en als je wilt rekenen, dan zul je eerst moeten zorgen dat dat kan – de enige optie is dan een wiskundig model.

Deze overwegingen betreffen nadrukkelijk ook berekeningen aan DNA en de bewijskracht van DNA-matches. Als het NFI (Nederlands Forensisch Instituut) gevraagd wordt om een bepaalde DNA-match te interpreteren, rapporteert het NFI vaak iets in de trant van: “De kans dat een willekeurig gekozen persoon dit DNA-profiel heeft, is kleiner dan 1 op de miljard.” Het simpele feit dat hier een getal gerapporteerd wordt, betekent automatisch dat er modelberekeningen zijn gemaakt.

Het aardige van DNA is dat de modellen die we daarvoor gebruiken werkelijk uitstekend zijn, en dat betekent dat de resultaten van dergelijke berekeningen ook heel betrouwbaar zijn. Dat wil overigens niet zeggen dat de uitkomst of de rapportage van het NFI eenvoudig te begrijpen is; daar komen we nog op terug. Vergelijk dit met de modellen die in de zaak Lucia de Berk werden gebruikt. Daar waren de modellen zeer omstreden, en de uitkomst van de berekeningen dus ook. Het ene model is het andere niet.

REDENEERMODELLEN

We vinden modellen echter ook op een ander niveau, namelijk op het niveau van onze menselijke redeneerlogica. Modellen die daarvoor gebruikt worden, pogen niet om een bepaald gedeelte van de fysieke werkelijkheid te beschrijven, maar proberen helder te krijgen hoe we redeneren. Het belang hiervan kan moeilijk overschat worden. De rechter

heeft de taak om een begrijpelijk en navolgbaar vonnis te ontwerpen. Begrijpelijk en navolgbaar betekent echter niet dat elke rechter in dezelfde zaak tot dezelfde conclusie zou moeten komen – dat is heel wat anders. Het redeneren onder onzekerheid brengt nu eenmaal met zich mee dat bepaalde gegevens door verschillende mensen ook verschillend worden gebruikt en geïnterpreteerd.

Maar hoe kan de rechter zijn vonnis uitleggen? Hoe kan hij ervoor zorgen dat andere mensen zijn uitleg begrijpen en kunnen navolgen, zelfs als ze het er niet mee eens zijn? Het is precies hier dat redeneerschema's en modellen opgeld doen. Als we kunnen begrijpen hoe mensen (en rechters in het bijzonder) tot hun opvattingen komen, dan zijn die opvattingen daarmee communiceerbaar geworden. Het mes snijdt dan aan twee kanten, want als de rechter begrijpt hoe hij zijn eigen denkpatroon kan formaliseren, dan kan hij ook beter inzien welke vragen hij aan getuigen of deskundigen kan of moet stellen. Dit alles leidt tot minder fouten, dat kan haast niet anders.

We moeten echter niet vergeten dat elk redeneermodel ook maar een model is. Hoe mensen tot hun opvattingen komen, zal nooit helemaal geformaliseerd kunnen worden, en datzelfde geldt voor de manier waarop mensen bijvoorbeeld van mening veranderen. De menselijk psyche is buitengewoon complex en geen enkel model zal daar helemaal recht aan kunnen doen. Een objectieve beoordeling van een situatie bestaat simpelweg niet. Intuïtief is dat wellicht al duidelijk, en de psychologie zal hier ook wel het een en ander over te zeggen hebben. Het is een simpel gegeven dat iemand meer bewijs verlangt als de hypothese waar het om gaat hem erg vreemd voorkomt. De subjectiviteit spat er als het ware vanaf. Dat betekent dat wat wij in dit boek te zeggen hebben, ook gevolgen heeft voor de vraag naar automatisering

van bepaalde aspecten van het rechtspreken. De AI-hype die ons tegenwoordig overspoelt, maakt de vraag naar de reikwijdte ervan in het recht onvermijdelijk. We denken dat een diepgaande analyse over hoe rechters tot hun oordelen komen ook zal impliceren dat we weg moeten blijven van AI-toepassingen in het recht, hoe aantrekkelijk die er op het eerste gezicht ook uit kunnen zien.

Het is overigens interessant om op te merken dat de logica van het redeneren onder onzekerheid, die we uit de doeken zullen doen, volkomen congruent is aan bepaalde psychologische fenomenen. Het hebben van een vooroordeel bijvoorbeeld, wordt vaak gezien als iets dat je zou moeten vermijden, en zeker een rechter. Rechters worden zelfs gewraakt als blijkt dat ze voorafgaand aan het proces zich te nadrukkelijk hebben uitgelaten over aspecten die voor de zaak van belang zijn.

Maar iedereen heeft vooroordelen, wij als auteurs, u als lezer, en rechters in een proces ook. Dat is simpelweg niet te vermijden – je bent geen machine. Het begrip ‘vooordeel’ heeft een negatieve connotatie, maar omdat vooroordelen niet te vermijden zijn, hebben we dringend behoefte aan een methodiek van redeneren waarin duidelijk wordt welke plaats vooroordelen precies hebben. Waar gebruikt de rechter zijn algemene kennis en visie op de wereld precies? In plaats van te roepen dat rechters geen vooroordelen mogen hebben (iets wat onmogelijk is), kun je beter een redeneerschema hanteren waardoor volkomen transparant wordt hoe de rechter tot zijn besluit is gekomen. Sterker nog, een paradigma waarin geclaimd wordt dat het tot objectieve beslissingen komt, kan om die reden niet adequaat zijn.

We beweren dat al deze aspecten, problemen en mogelijke oplossingen ervan, allemaal zichtbaar gemaakt kunnen worden aan de hand van de praktijk. Sterker nog, we denken dat

weinig rechters of andere juristen behoefte hebben aan een tekst waarin de problematiek op een abstracte manier wordt benaderd. Zo'n abstracte benadering gaat al snel over formele rekenregels, formules die wel iets betekenen misschien, maar wat dan precies? Het heeft meestal ook weinig zin om te praten over extreme hypothetische situaties die wel erg ver afstaan van de dagelijkse praktijk van elke strafrechter.

EEN OVERZICHT VAN HET BOEK⁶

Het boek is bedoeld om bewustwording over het redeneren onder onzekerheid te stimuleren, en om het redeneerproces zelf te verbeteren. We hopen dat het redeneren onder onzekerheid en het denken over kansen en statistiek tot het 'gewone' gereedschap van elke strafjurist zal gaan behoren, en niet meer als 'eng', 'iets voor nerds' of 'niets voor alfa's' gezien gaat worden. We hopen dat juristen gaan begrijpen hoe ze van de inzichten uit dit vakgebied gebruik kunnen maken om hun uitspraken beter te maken.

Laten we deze inleiding afsluiten met een globaal overzicht van de hoofdstukken die voor u in het verschiet liggen. We beginnen in hoofdstuk 2 met het rustig beschouwen en bespreken van de manier waarop mensen tot hun oordelen en beslissingen komen, zoals beloofd vooral aan de hand van concrete zaken. We zullen daarin onderscheid maken tussen de gegevens en data die specifiek zijn voor de onderhavige casus, en gegevens die algemeen zijn. Deze laatste worden doorgaans impliciet maar soms ook expliciet gebruikt in

⁶ Dit boek is het resultaat van jarenlang denkwerk en het bouwt voort op eerdere publicaties. Het is daarmee onvermijdelijk dat bepaalde passages uit het boek overeenkomsten vertonen met passages uit eerder (gezamenlijk) werk.

zaken, en de vraag is hoe we daar precies mee om moeten gaan. Al met al bouwen we het redeneermodel in dat hoofdstuk op, op een manier die hopelijk aansluiting vindt bij de belevingswereld van juristen. De benadering die wij kiezen staat bekend als het ‘Bayesiaanse redeneermodel’. Onze bespreking ervan leidt automatisch tot de conclusie dat een dergelijk redeneerschema *an sich* al impliceert dat objectiviteit niet bereikt kan worden. Sterker nog, we zullen duidelijk maken dat dit ook niet de bedoeling kán zijn, en dat pogingen daartoe alleen maar tot ongelukken zullen leiden.

In hoofdstuk 3 gaan we dieper in op de rol die kansen, getallen en modellen hebben in onze benadering. Allereerst reflecteren we op de begrippen ‘kans’ en ‘onzekerheid’. Wat bedoelen we precies met die begrippen? Dat is een subtiele kwestie die we niet uit de weg kunnen gaan en die na de eraan voorafgaande hoofdstukken hopelijk niet meer abstract aanvoelt. Deze overwegingen leiden dan bijna automatisch naar een filosofische bespiegeling over wetenschap en statistiek voor zover relevant in strafzaken. Het gaat dan niet alleen over de vraag hoe we wetenschappelijke bevindingen moet zien, maar ook over de grenzen van de wetenschap zelf. Sommige dingen zijn niet te weten en vaak ook niet te meten, en als dat zo is moeten we dat accepteren en niet doen alsof we het wel weten of kunnen meten. Aantonen dat een bepaald effect het gevolg is van een bepaalde oorzaak is simpelweg niet altijd mogelijk, en dat is in het recht een belangrijk gegeven. Hierbij speelt de discussie over het gebruik van algemene statistieken, *naked statistics*, een belangrijke rol. Natuurlijk proberen we onze argumenten zo veel mogelijk te illustreren met sprekende voorbeelden.

In hoofdstuk 4 gaan we in op de relatie tussen het Bayesiaanse redeneermodel en de wet. Het is natuurlijk belangrijk dat we correct redeneren, maar staat ‘Bayes’ niet op gespan-

nen voet met elementaire juridische beginselen? Het feit dat we het hebben van vooroordelen expliciet opnemen in onze redeneerstructuur lijkt bijvoorbeeld voor veel mensen op gespannen voet te staan met de onschuldpresumptie. We zullen zien dat dit vermeende probleem op een misverstand berust en dat ons redeneervoorstel volkomen congruent is met de bestaande praktijk én met de onderliggende rechtsfilosofische gedachten omtrent de onschuldpresumptie. In dat hoofdstuk bespreken, en weerleggen, we de kritiek dat Bayesiaans redeneren persoonlijke overtuigingen in het vonnis ‘binnenhaalt’, en staan we verder stil bij het zwijgrecht, bij schakelbewijs en ten slotte bij de bewijsminimumregel – allemaal met een Bayesiaanse bril.

Hoofdstuk 5 behandelt een aantal ‘rode vlaggen’; fouten die vaak worden gemaakt en hoe we die kunnen voorkomen. We bespreken bijvoorbeeld het feit dat wanneer een bepaalde bevinding een heel kleine kans had onder een bepaalde hypothese, dat helemaal niet betekent dat die hypothese daardoor onaannemelijk wordt. Ook besteden we aandacht aan de mythe dat het bewijs in een zaak maar op één manier bepaald zou kunnen worden. Kortom, een hoofdstuk waarin we waarschuwen voor allerlei gevaren en mogelijke obstakels, uiteraard weer met de nodige voorbeelden.

Ten slotte bespreken we in hoofdstuk 6 een aantal zaken die we op een Bayesiaanse manier zullen analyseren. Soms leidt dat tot de conclusie dat het in een bepaalde zaak goed is gegaan, maar soms lijkt het dat een juiste analyse tot een andere uitspraak had kunnen leiden.

Hopelijk is het inmiddels duidelijk dat we graag een leesbaar boek schrijven. Saai gaat het vast niet worden, denken we. De gedachten die erin besproken worden, vinden hun oorsprong in diverse wetenschapsgebieden, maar we schrij-

ven dit boek meer als een praktisch leerboek dan als een wetenschappelijke tekst. We zullen ons dan ook niet houden aan de gewoonte om wetenschappelijke juridische teksten van eindeloze voetnoten te voorzien. We zullen ook spaarzaam zijn met verwijzingen naar de literatuur, maar als het nodig is doen we dat natuurlijk wel. Voor ons is het vooral van belang dat het een prettig en herkenbaar boek is om te lezen. Alleen dan, zo stellen wij ons voor, wordt het ook daadwerkelijk gelezen.

HOOFDSTUK 2

Het Bayesiaanse redeneermodel – theorie en voorbeelden

Redeneren onder onzekerheid is moeilijk, en fouten liggen altijd op de loer. Redeneren appelleert aan wiskunde en in het bijzonder aan de mathematische logica. We moeten daarbij wel goed in gedachte houden dat een wiskundige benadering altijd een beperking en vereenvoudiging van de werkelijkheid impliceert. Hoe mensen tot hun oordelen en beslissingen komen is een ingewikkelde kwestie, en het is maar zeer de vraag tot welk punt zich zoiets laat formaliseren. Ook de vraag of een rechter een getuige gelooft kan niet tot pure logica, kansrekening en statistiek worden teruggebracht. Er spelen dingen mee die zich maar moeilijk of helemaal niet laten kwantificeren: lichaamstaal, manier van spreken, oogopslag, de psychologische conditie en houding van de getuige, et cetera. De rechter zal dit – veelal onbewust – allemaal meenemen in zijn beoordeling van de getuige en zijn getuigenis. In zichzelf betekent dit al dat eventuele statistische kennis over de vraag hoe vaak getuigen de waarheid spreken, nauwelijks zinvol is in een specifieke zaak, en hooguit achtergrondinformatie kan zijn die de rechter natuurlijk wel ter harte moet nemen.

Achtergrondkennis heeft dus zeker wel betekenis, en geen enkele rechter kan doen alsof dat niet zo is. Mensen hebben kennis van de wereld, op allerlei formele maar vooral ook informele manieren: ervaring, intuïtie, gesprekken met andere mensen, de media. Soms zal die kennis een wetenschappelijke basis hebben, maar heel vaak ook niet. Niet-

wetenschappelijke kennis is ook kennis, zelfs als je niet zeker weet hoe je eraan komt.

Iets dergelijks geldt ook voor vaardigheden. Als vaardigheden algoritmisch zouden zijn en in principe door iedereen aan te leren, dan zouden er niet zulke grote verschillen tussen mensen bestaan. We zeggen vaak dat iemand ergens aanleg of talent voor heeft. Wat dat exact betekent is niet helemaal duidelijk, maar dat hoeft ook niet om de begrippen toch te kunnen gebruiken. Je kunt tot op zekere hoogte leren koken, maar zoals een bekende chef-kok onlangs bij zijn afscheid zei: koken is niet het volgen van een recept, het is gevoel.¹ Waar dat gevoel vandaan komt, laten we gemakshalve maar ook noodzakelijkerwijs in het midden. Zo is het dirigeren of uitvoeren van een symfonie ook veel meer dan het spelen van de noten uit de partituur. Wat dat ‘veel meer’ nu echter precies is, is niet helemaal duidelijk, maar dat er zeer grote verschillen bestaan tussen dirigenten is evident als je luistert naar hun respectievelijke uitvoeringen.

Er zijn kennelijk een heleboel ‘dingen’ in de wereld die zich niet zo makkelijk laten beargumenteren. Op dat vlak zijn ‘begrip’ en ‘overtuiging’ misschien wel het ingewikkeldst. Het is heel moeilijk om uit te leggen waarom je iets denkt te begrijpen: het fenomeen is niet conceptualiseerbaar en uiteraard persoonsafhankelijk. Je kunt iemand of jezelf niet dwingen iets te begrijpen – het overkomt je. Natuurlijk kun je zorgen voor ideale omstandigheden: je kunt in de leer gaan bij goede leermeesters, je kunt goede literatuur lezen, je kunt de juiste ervaring gaan opdoen. Maar het laatste stukje van de weg naar begrip moet je helemaal alleen afleggen. Als alle argumenten de revue zijn gepasseerd en alle leermeesters hun zegje hebben gedaan komt begrip wel of niet. Dat begrip

¹ *Leidsch Dagblad*, 25 mei 2024.

is daarom uiteindelijk ook niet terug te brengen tot een afvinklijstje of tot een strikt logische redenering. Als dat zo zou zijn, zouden er op dat punt ook geen verschillen bestaan tussen mensen.

Een overtuiging lijkt op het eerste gezicht beter uitlegbaar te zijn dan begrip. Je bent ergens van overtuigd omdat je (denk je) redenen hebt om iets te geloven. Je kunt een persoonlijke overtuiging dus beargumenteren. Het probleem echter is dat de praktijk uitwijst dat verschillende mensen met dezelfde gegevens toch verschillende overtuigingen hebben. Kennelijk waarderen verschillende mensen gegevens op verschillende manieren. Hoe dat precies werkt, is waarschijnlijk meer een vraag voor psychologen, maar wat evident is, is dat een overtuiging niet zomaar gereduceerd kan worden tot een aantal argumenten. Er zijn immers altijd ook nog andere argumenten, dus moet je vervolgens uitleggen waarom juist deze argumenten voor jou doorslaggevend zijn en andere niet. En dan begint het proces opnieuw, en voor je het weet kom je in een oneindige regressie terecht, en lijkt het erop dat geen enkele overtuiging een basis kan hebben. Je kunt immers altijd doorvragen.

Filosofen hebben zich hier het hoofd over gebroken, en natuurlijk bestaan er veel verschillende opvattingen over overtuigingen. Die zijn op zich zeker interessant, maar het gaat ons hier vooral om de vaststelling dat het bepalen van een overtuiging niet eenvoudig kán zijn. Proberen uit te leggen waarom je ergens van overtuigd bent, kan niet anders dan een heel persoonlijk proces zijn. Geen enkele overtuiging kan het gevolg zijn van alleen maar objectieve en simpele logica, zo werkt het zeker niet.

Rechters dienen in elke strafzaak aan te geven of ze overtuigd zijn van de schuld van de verdachte of niet, en deze overtuiging dient navolgbaar te zijn. Navolgbaar bete-

kent nadrukkelijk niet hetzelfde als objectief, want men kan navolgbaar toch op verschillende antwoorden uitkomen.

Deze korte beschouwing leidt al snel tot de conclusie dat overtuigingen niet via een geautomatiseerd proces tot stand kunnen komen. Dat is principieel onmogelijk.

Wij kunnen als auteurs assisteren bij het vormen van een oordeel, maar niet op een manier dat we alle rechters tot dezelfde conclusie kunnen laten komen. Dat zou de hele idee van het hebben van een persoonlijke overtuiging natuurlijk ook volledig teniet doen. Als we vanuit de logica en de wetkunde een bijdrage kunnen leveren, dan moet die bijdrage op een ander vlak liggen. Wat we hier zullen beargumenteren is dat onze menselijke logica tot een denkmodel kan leiden dat een rechter in staat stelt om zijn eigen opvattingen, ideeën, voorkennis en vooroordelen, alsook de gegevens in een specifieke zaak op een coherente manier te communiceren zodat zijn uiteindelijke conclusie navolgbaar wordt. En niet alleen dat: het denkmodel kan ook gebruikt worden om het eigen denken te corrigeren, daar waar het soms makkelijk is om fouten te maken.

Want laten we wel zijn, het is ontzettend moeilijk om bij logisch redeneren fouten te voorkomen. Fouten die je misschien niet maakt als je de grote lijnen volgt van het denkmodel dat we aanreiken. Het volgen van een bepaalde procedure maakt argumenten niet alleen navolgbaar en communiceerbaar, maar vormt tot op zekere hoogte diezelfde gedachten. Het is een wisselwerking, ongeveer zoals het schrijven van een boek dat ook is: je schrijft op wat je in gedachten hebt, maar het feit dat je het opschrijft, vormt in zichzelf die gedachten ook weer. Een goed denkmodel doet meer dan de puzzelstukjes in elkaar passen: een goed denkmodel zal de puzzelstukjes tot op zekere hoogte ook zelf aanleveren.

Een dergelijk denkmodel kan niet anders dan een model op een tamelijk hoog abstractieniveau zijn, en kan zeker geen model zijn dat ons vertelt hoe we onze gedachten in elke concrete situatie moeten bepalen. Het dient een model te zijn dat in zichzelf een bepaalde gedeelde waarde heeft, maar waarbinnen iedereen terecht kan met zijn eigen gedachten en overtuigingen. Het moet een model zijn waarbinnen aangegeven wordt hoe we moeten omgaan met redeneren onder onzekerheid, en moet tegelijkertijd ruimte bieden aan ieders persoonlijkheid en aan ieders specifieke kennis en rol in een proces.

Bayesiaans denken is precies dat: een denkmodel op een tamelijk hoog abstractieniveau. Het is niet dé waarheid – geen enkel model is dat – maar het kan zeer behulpzaam zijn bij het vormen van je gedachten en overtuigingen. Sommige mensen vinden dat Bayesiaans redeneren moeilijk is, of zelfs tegenintuïtief, of stellen dat het niet zou passen binnen ons rechtssysteem. We hebben echter ook al heel vaak gehoord dat wanneer men Bayesiaans redeneren eenmaal begrijpt, het eigenlijk heel logisch is en ook heel dicht staat bij wat men altijd al deed. Er is dan ook helemaal geen sprake van dat wij van plan zijn om een bepaalde methodologie op te dringen. In tegendeel: wij willen alleen maar laten zien hoe verhelderend het denken in Bayesiaanse termen kan zijn.

Het is misschien wel aardig om iets te zeggen over de reden dat het denkmodel dat we hier gaan uitleggen deze wat vreemde naam heeft: Bayesiaans redeneren. De naam komt van Thomas Bayes, een achttiende-eeuwse Engelse geestelijke, filosoof en statisticus. Dat laatste was hij vooral als hobbyist, maar dat bedoelen we absoluut niet denigrerend. Hoewel de kansrekening en de statistiek in die tijd nog niet erg ontwikkeld waren, formuleerde Bayes een regel binnen de kansrekening die we nu aan eerstejaars studenten wiskunde

heel makkelijk uit kunnen leggen, en die wiskundig gezien bijzonder eenvoudig is. De regel zegt iets over zogenaamde conditionele kansen, kansen op gebeurtenissen die veranderen als je nieuwe informatie krijgt. Deze elementaire kansrekening is op geen enkele manier controversieel, en er bestaat dan ook helemaal niet zoiets als ‘Bayesiaanse kansrekening’, iets wat veel mensen schijnen te denken. Er bestaat alleen een simpele rekenregel die de regel van Bayes wordt genoemd, en die door geen enkele wiskundige in twijfel wordt getrokken.

Dat Bayesiaans redeneren desondanks door sommigen als controversieel wordt gezien, heeft te maken met de interpretatie van de wiskunde en niet met de wiskunde zelf. Daarover zullen we in het volgende hoofdstuk meer schrijven.

BAYESIAANS REDENEREN; EEN MENS BEGINT NIET BIJ NUL

Het is dan nu tijd om uit te leggen wat Bayesiaans redeneren behelst. Dat doen we uiteraard niet in wiskundige termen. Veel auteurs doen dat wel, maar daarmee geven ze precies het verkeerde signaal af: Bayesiaans denken is namelijk nadrukkelijk geen wiskunde die zich opdringt aan juristen. Het is precies andersom: het is een denkmodel dat je desgevraagd (ook) formeel wiskundig kunt beschrijven. De relatie tussen het redeneermodel en de wiskunde is dus precies omgekeerd aan wat veel auteurs beweren. Een wiskundige beschrijving in formules is nadrukkelijk niet de kern van Bayesiaans denken; die beschrijving is niets meer dan een wiskundige formalisering ervan. Je hebt die wiskunde soms nodig om bepaalde aspecten van Bayesiaans redeneren te ontdekken of te communiceren, maar we kunnen niet genoeg benadrukken dat onze benadering absoluut niet betekent dat wij juristen wiskundig naar strafzaken willen laten kij-

ken. Een wiskundige formulering van een bepaald concept betekent altijd dat je iets verliest, want de wiskunde kan de werkelijkheid niet ‘zijn’. Het is daarom niet juist om met de wiskunde te beginnen en dan te claimen dat we moeten denken zoals de wiskunde dat kennelijk voorschrijft. Het moet precies andersom: eerst moeten we duidelijk maken wat de gedachtegang van Bayesiaans redeneren eigenlijk is, en als je wilt kun je die daarna proberen te formuleren in wiskundige termen – over dat laatste zullen we in het volgende hoofdstuk nog het een en ander schrijven.

De gedachte achter Bayesiaans redeneren is in zichzelf eenvoudig. De Bayesiaanse benadering accepteert dat er geen objectieve overtuiging kan bestaan, en dat er dus ruimte moet zijn voor persoonlijke invulling. Bayesiaans denken begint daarom met de vaststelling dat we allemaal een bepaalde kijk op de wereld hebben: vooroordelen, kennis, achtergrond en ervaring. Een verdachte die voor de rechter verschijnt, begint niet bij nul: er is een reden dat deze persoon daar staat en niet iemand anders, en de rechter is zich daarvan bewust. Geen mens kan volhouden dat hij ergens absoluut geen mening over heeft, want ‘geen mening’ hebben is ook een standpunt. Je kunt net zo goed een mens vragen om zich niet te gedragen: dat gaat niet want als je bestaat en aanwezig bent, dan gedraag je je automatisch op een bepaalde manier. Het hebben van een vooroordeel is voor sommigen eigenlijk al niet acceptabel en een reden om Bayesiaans redeneren niet te accepteren. Maar geen mens komt daaraan in het proces naar besluitvorming, en dus kun je maar beter een denkmodel hanteren waar dit een rol krijgt en waarin dit transparant wordt. De verdachte die voor de rechter staat, is geen ‘gemiddeld’ persoon, wat dat ook precies moge betekenen. We zullen in hoofdstuk 4 zien hoe dit

gegeven zich verhoudt tot de onschuldpresumptie, maar nu concentreren we ons eerst op het denkmodel zelf.

In het ruwe denkmodel dat we hier beschrijven, beginnen we dus niet bij nul: dat kan ook helemaal niet en het is dus goed dat het model daar ruimte voor schept. De gedachte is dat we vervolgens bekijken in hoeverre nieuwe informatie onze opvattingen – die we dus tot op zekere hoogte al moeten hebben – verandert. Bijvoorbeeld, als een verdachte van een zedendelict in beeld is gekomen door verdacht gedrag, dan zal nieuwe informatie in de vorm van een waterdicht alibi hem vrijpleiten. Met andere woorden: als je nieuwe informatie krijgt, dan zal je mening over een persoon (of ergens anders over) kunnen en soms zelfs moeten veranderen.

Dat nieuwe informatie iemands mening of overtuiging kan veranderen, is natuurlijk helemaal niet controversieel. Wat minder voor de hand ligt, is het feit dat deze meningsverandering altijd relatief moet zijn. We leggen nu eerst uit wat we daarmee bedoelen. We doen dat aan de hand van een aantal voorbeelden: eerst een fictief voorbeeld en daarna enkele concrete voorbeelden uit de rechtspraak.

BEWIJS MOET RELATIEF WORDEN BEOORDEELD

Stel je voor dat de vraag voorligt of twee mensen op een bepaalde manier gerelateerd zijn, bijvoorbeeld bij een asiel-aanvraag waarbij de vraag of twee mensen familie van elkaar zijn belangrijk is. Stel vervolgens dat de DNA-profielen van deze twee mensen een bepaalde overeenkomst hebben. Is die overeenkomst bewijs dat ze inderdaad familie van elkaar zijn?

Om die vraag te beantwoorden hebben we aan deze informatie niet genoeg. Wat we moeten weten is of de specifieke

overeenkomst in het DNA die we hier vinden, gebruikelijk is of niet. Als deze specifieke overeenkomst in de helft van de gevallen waarin we twee willekeurige personen vergelijken voorkomt, dan is het duidelijk dat ze weinig informatie geeft over het al dan niet familie zijn van deze twee specifieke personen. Als de overeenkomst echter praktisch niet voorkomt bij niet-gerelateerde mensen, dan zegt ze veel meer. Met andere woorden, we moeten weten of de data (de overeenkomst tussen het DNA van de twee personen) vaak voorkomen onder mensen die juist niet familiair gerelateerd zijn. Die informatie is nodig om de overeenkomst op waarde te schatten. Je kunt dit ook als volgt formuleren: we moeten weten of de overeenkomst beter past bij twee mensen die gerelateerd zijn, of bij twee mensen die niet (familiair) gerelateerd zijn. Als de specifieke overeenkomst (veel) gebruikelijker is onder familieleden dan onder vreemden, dan is de overeenkomst een aanwijzing dat de twee personen inderdaad gerelateerd zijn.

We schrijven hier ‘bewijs’ maar daarmee bedoelen we niet dat de familierelatie hiermee onomstotelijk is bewezen. We bedoelen ermee dat het een aanwijzing in die richting is. Je kunt over de vraag of het een aanwijzing is niets zeggen als je niet weet of de match in een andere omstandigheid al dan niet gebruikelijk is. Dat betekent dat bewijs altijd relatief zal zijn.

Hier is een tweede voorbeeld om het punt van de relativiteit van bewijs te illustreren. Dat voorbeeld ontlene wij aan een (zeden)zaak.² Tijdens een familiebarbecue vergeleek een 6-jarig meisje, schijnbaar uit het niets, een frikandel met de penis van de verdachte, haar oudoom. Haar moeder heeft haar daarna vragen gesteld, waaronder de vraag of

² ECLI:NL:RBLIM:2020:2348.

ze ook haar “muske” (hiermee bedoelt ze de vagina) heeft laten bekijken. Moeder heeft op 29 augustus 2017 aangifte gedaan en heeft onder meer verklaard dat het gebeurd moet zijn op 12 augustus 2017 tijdens een verjaardagsfeest in het huis van de tante van het slachtoffer. Het slachtoffer is daarna op 11 september 2017 in de kindvriendelijke studio gehoord. Naar de betrouwbaarheid van de tijdens dit – woordelijk uitgewerkte – studioverhoor afgelegde verklaring van het slachtoffer heeft een rechtspsycholoog op verzoek van de verdediging onderzoek gedaan. De rechtbank schrijft dat dit gaat over “de waarschijnlijkheid dat deze verklaring het gevolg is van een alternatief scenario, waarin geen sprake is geweest van ontuchtige handelingen”. De waarschijnlijkheid van een gesuggereerd (maar dus incorrect) antwoord kan volgens de deskundige toenemen tot 60%. “Dat betekent niet dat de kans op een valse verklaring in dit geval 60% is, maar het betekent dat er een niet te verwaarlozen kans bestaat dat het slachtoffer belastend zou verklaren, ook als er geen sprake is geweest van onzedelijke handelingen.”

Hoe moeten we deze 60% nu interpreteren? De meeste studenten en rechters zeggen desgevraagd dat ze denken dat dit getal van 60% ontlastend is. Er lijkt immers een best grote kans te bestaan dat het kind op deze manier zou verklaren ook als er niets gebeurd is. Wat echter doorgaans vergeten wordt, is dat die 60% alleen maar slaat op het scenario dat zich niets onoorbaars heeft voorgedaan. Het enige wat zou kunnen maken dat dit überhaupt zinvol bewijsmateriaal zou kunnen zijn, is dat de kans op een dergelijke verklaring onder het schuldscenario anders is. Immers, als een dergelijke verklaring onder zowel het schuld- als het onschuldscenario even vaak voorkomt, of even verrassend is, dan zegt de verklaring van het kind niets als het gaat om de vraag welk scenario waar is en welk niet.

Tijdens deze zaak is dit volledig over het hoofd gezien. We zijn geen psychologen en wij weten niets over de betrouwbaarheid van het type verklaring zoals die hier door het kind is afgelegd. Maar laten we voor het maken van ons punt stellen dat als het kind wél onzedelijk is behandeld, de kans op een dergelijke verklaring 90% zou zijn. Dat zou betekenen dat het feit dat de verklaring er nu eenmaal is, beter past bij het schuld- dan bij het onschuldscenario, en zouden we tot de conclusie moeten komen dat de verklaring nog steeds belastend is, en niet ontlastend.

Uit de uitspraak wordt duidelijk dat de rechter de verklaring van de expert als ontlastend beschouwde, maar daar is dus helemaal geen reden toe. En net als bij het eerdere DNA-voorbeeld zien we dat we de gegevens niet goed kunnen waarderen zonder ook de alternatieve scenario's mee te laten doen. Om te weten of een bepaald bewijsmiddel belastend is of niet moeten we weten in hoeverre het aanwezig zijn van dit bewijsmiddel beter door de schuld- dan door de onschuldhypothese verklaard kan worden. Daarvoor moeten we dus noodzakelijkerwijs beide scenario's beschouwen. Hiermee wordt duidelijk dat deze manier van redeneren ook bescherming biedt tegen het hebben van een tunnelvisie. Je wordt immers gedwongen om naar alternatieve verklaringen voor de gegevens te kijken.

Ons derde voorbeeld betreft de beruchte strafzaak tegen Sally Clark in Groot-Brittannië aan het einde van de vorige eeuw. We hoeven hier niet alle details van die zaak uit de doeken te doen; een korte samenvatting volstaat. Sally Clark was moeder van een zoontje van enkele maanden oud, toen dit zoontje onverwacht en zonder duidelijke oorzaak overleed. Sally Clark was op het moment van overlijden alleen thuis met haar zoontje. Enkele jaren later gebeurde met haar tweede zoontje precies hetzelfde: het overleed toen het enkele

maanden oud was en zijn moeder alleen met hem thuis was. Sally Clark werd aangeklaagd voor dubbele moord en tot twee keer toe veroordeeld.

Het voornaamste argument dat tot deze veroordelingen leidde, was een berekening: men probeerde uit te rekenen hoe groot de kans is dat in een willekeurig gezin twee kinderen aan wiegendood overlijden. Immers, hoewel het niet vaak gebeurt, komt het voor dat kinderen van die leeftijd om onverklaarbare reden sterven. Men berekende die kans en kwam uit op een astronomisch klein getal. Dit getal was zo klein dat men concludeerde dat een dubbele wiegendood niet meer geloofwaardig was.

Dit is opnieuw een voorbeeld van een zaak waarin men vergat dat bewijs relatief is. In dit geval waren er twee scenario's waar men rekening mee wilde houden: ofwel Sally Clark was onschuldig en de dood van de twee kinderen een tragisch voorval, ofwel Sally Clark had haar twee zoontjes zelf omgebracht. Het verbijsterende aan deze zaak wordt duidelijk als je je af gaat vragen in hoeverre we twee overleden kinderen 'verwachten', onder elk van de twee scenario's. Dit is precies wat we bij de eerdere voorbeelden ook deden: we kijken in hoeverre de gegevens (hier de twee gestorven kinderen) beter 'passen' of minder verrassend zijn bij een van de scenario's. In dit geval is het antwoord eenvoudig: bij beide scenario's, dubbele moord of dubbele wiegendood, verwachten we twee gestorven kinderen. We concluderen, enigszins verrassend wellicht, dat het feit dat er twee kinderen in deze omstandigheden stierven, op geen enkele manier onderscheidend is voor de vraag of het een dubbele wiegendood of een dubbele moord was. Als iemand correct redeneert, dan is zijn mening over de vraag of Sally Clark een dubbele moordenaar is of dat het een dubbele wiegendood betreft door de aanwezigheid van twee gestorven kinderen niet veranderd.

Een betere illustratie voor de noodzaak om correct te redeneren kan bijna niet gegeven worden. Wat we uit dit en de daaraan voorafgaande voorbeelden leren, is dat het helemaal niet relevant is hoe groot de kans is dat we een bepaalde aanwijzing zouden vinden onder bijvoorbeeld de onschuldhypothese. Waar het om gaat is of de gegevens beter passen bij het ene of bij het andere scenario. Het is nu eenmaal zo dat de feiten zich hebben voorgedaan. Dat de feiten zelf zeldzaam zijn, is niet van belang. Nu de feiten zich hebben voorgedaan, is de enige vraag wie of wat voor deze feiten verantwoordelijk waren. Als de feiten veel logischer zijn onder de eerste hypothese dan onder de tweede, dan vormen die feiten bewijs voor de eerste hypothese.

Ons volgende voorbeeld is van een heel andere aard, maar de systematiek is ook in deze zaak van toepassing.³ Op vrijdag 28 juni 2018 omstreeks 03.30 uur werd een vrouw door een haar onbekende man met een mes of scherp voorwerp in haar arm gestoken toen zij samen met een vriendin naar huis liep. Deze vriendin verklaarde dat de persoon die haar vriendin had gestoken een vermoedelijk Turkse man was van ongeveer 35 jaar oud met donker kort haar, wat dunner bovenop, brede, borstelige wenkbrauwen, een lichtbeige jas en donkere kleding. De man had tevens een opvallend ‘pinguïn’-loopje. Hij had hen voor hoer uitgemaakt en toen het slachtoffer hierop reageerde, haar direct gestoken. Ook zei hij iets als: “Ik steek jullie neer.” Er wordt een verdachte aangehouden. Bij de politie en bij Meld Misdaad Anoniem komen na het incident diverse anonieme meldingen binnen waarin verdachte wordt genoemd als dader van het steekincident. Ook wordt verdachte door meerdere getuigen herkend wan-

³ ECLI:NL:RBOBR:2020:4543.

neer beelden van de vermoedelijke dader worden getoond bij het televisieprogramma Opsporing Verzocht.

Op basis van het signalement van de persoon die zichtbaar is op de beelden, in combinatie met de locatie en het tijdstip van deze beelden, acht de rechtbank bewezen dat de persoon op deze beelden ook de dader van het tenlastegelegde steekincident en de tenlastegelegde bedreiging is. Maar de vraag is of dit inderdaad de verdachte is.

De rechtbank constateert dat, voor zover op basis van het dossier kan worden nagegaan, alle herkenningen geheel of in belangrijke mate zijn gebaseerd op het loopje van de man zoals te zien op de beelden. Naar de overeenkomsten tussen het gangbeeld van de dader van de feiten en het gangbeeld van verdachte is onderzoek gedaan door deskundigen van het Universitair Medisch Centrum en de Rijksuniversiteit Groningen. De conclusie van dit onderzoek was dat de kans op het vinden van overeenkomst in gangbeeld van de dader van het steekincident en de verdachte, uitgaand van de hypothese dat het om dezelfde persoon gaat, 1,1 keer zo groot is als de kans uitgaande van de hypothese dat het om verschillende personen gaat. Deze onderzoekers rapporteren dus een expliciet getal: het loopje is volgens deze onderzoekers nauwelijks minder verrassend als de dader en de verdachte dezelfde personen zijn dan wanneer zij verschillende personen zijn. Daarmee draagt, althans volgens deze onderzoekers, het loopje niet of nauwelijks bij aan het onderscheiden van de twee hypothesen waar het hier om gaat. Met andere woorden, volgens de deskundigen is het gangbeeld dus vrijwel neutraal bewijs, en is het gangbeeld bijna even goed te verklaren als de man op de beelden de verdachte is, als wanneer de man op de beelden dat niet is.

Laten we deze zaak eens van enig commentaar voorzien. Om te beginnen is de logica van het Bayesiaans redeneren

correct toegepast: men heeft zich afgevraagd onder welke van twee rivaliserende hypothesen het bewijs (het loopje) het beste te verklaren was. Daarna volgt natuurlijk de vraag waar dat getal van 1,1 vandaan komt. Het vonnis geeft daar geen enkel inzicht in, en dat is niet acceptabel. Het getal is niet navolgbaar. Het getal bevreemdt ons ook: het loopje werd als bijzonder gezien door de getuigen en het slachtoffer, en dat bijzondere zou zich moeten vertalen in een zeker onderscheidend vermogen. Het gerapporteerde getal van 1,1 is in essentie neutraal bewijs, precies zoals in de zaak Sally Clark de twee gestorven kinderen neutraal bewijs vormden voor de twee hypothesen die we daar bekeken. Neutraal bewijs betekent dat je eigenlijk niets wijzer geworden bent na het beoordelen van het bewijs voor de twee specifieke hypothesen waar je naar kijkt. In het geval van het loopje lijkt dit eigenlijk onmogelijk want het loopje is kennelijk bijzonder genoeg om op te vallen. Als iets opvalt, dan zal het ook enig onderscheidend vermogen moeten hebben – anders was het immers niet opgevallen.

We zien nu iets interessants: het correct toepassen van het redeneermodel betekent kennelijk niet dat alle conclusies navolgbaar zijn. Maar we kunnen nu wel scherp zien welk gedeelte van de redenering nadere uitleg behoeft. Hoewel we kritisch zijn over het specifieke gebruik van het model hier, is de transparantie absoluut winst en belangrijk.

In het eerdere voorbeeld in de zedenzaak rond de familiebarbecue was ook sprake van een getal: de 60% die door de psycholoog werd genoemd. In dat geval was die 60% ongetwijfeld gebaseerd op de ervaring en vakkennis van de psycholoog. Het is niet zo lastig om die 60% te interpreteren, ook al schreven we eerder dat dergelijke statistieken niet zomaar in een specifieke casus gebruikt kunnen worden. In ons laatste voorbeeld was dat anders. Het getal 1,1 dat daar

gebruikt werd, heeft een onduidelijke herkomst. Hoe zou je onderzoek kunnen doen naar het onderscheidend vermogen van een loopje? Het feit dat er een getal uitkomt, betekent dat er iets uitgerekend moet zijn: dat getal komt ergens vandaan. Maar wanneer iets uitgerekend wordt, moet er een wiskundig model op de achtergrond bestaan waarbinnen dat getal een bepaalde betekenis krijgt. Je kunt zelfs stellen dat élk getal dat uitgerekend wordt, een wiskundig model vereist. Als dat model de relevante werkelijkheid niet goed beschrijft, dan heeft dat getal weinig betekenis. Dat betekent andermaal dat de rechter het hier genoemde getal van 1,1 niet zomaar had mogen accepteren.

Over het gebruik van getallen en modellen hebben we in het volgende hoofdstuk nog veel meer te zeggen, maar we introduceren nu de terminologie die hierbij doorgaans wordt gebruikt. Het getal van 1,1 uit het laatste voorbeeld wordt een ‘likelihood ratio’ genoemd. Het getal drukt uit onder welke van de twee hypotheses de gegevens het beste verklaard kunnen worden, en met welke factor. In het geval van Sally Clark hadden we een likelihood ratio van 1, dat is dus neutraal bewijs waar je niets van leert. In het voorbeeld van de zedenzaak rond de familiebarbeque werd door de deskundige het getal van 60% genoemd als de waarschijnlijkheid dat het kind op deze manier zou verklaren als er niets zou zijn voorgevallen. Als de 90% die we daarna zelf introduceerden reëel zou zijn, dan zou dat leiden tot een likelihood ratio van $90/60 = 1,5$. Dit is helemaal geen overtuigend getal, maar ontlastend wordt de verklaring dus zeker niet.

Hier is een volgend voorbeeld waarin getallen gebruikt worden die niet verder geduid of uitgelegd worden.⁴ De zaak draait om een zogenaamde ‘tipgevertelefoon’. De verdachte

⁴ ECLI:NL:RBOBR:2017:6556.

werd verdacht van het via een sms-bericht doorseinen van de komst van het slachtoffer. Het slachtoffer werd door een schutter opgewacht en meerdere keren beschoten, waarna het slachtoffer overleed. In deze zaak staat niet de schutter maar de tipgever voor de rechter.

Een van de bewijsmiddelen in deze zaak bestaat eruit dat men heeft onderzocht of de telefoon van verdachte en de tipgevertelefoon zich op dezelfde locatie bevonden gedurende de periode voorafgaand aan de moord. Over de locatie van telefoons kan informatie verkregen worden door na te gaan welke masten door de telefoons worden ‘aangestraald’. Men formuleerde daartoe, zoals het hoort, twee hypothesen: ófwel de twee telefoons (die van de verdachte en de tipgevertelefoon) bewogen onafhankelijk van elkaar, ófwel ze bevonden zich gedurende de relevante tijd bij elkaar.

We citeren nu uit de uitspraak: “Het NFI heeft berekend dat de mate van overeenkomst in locatiegegevens 19 tot 105 keer waarschijnlijker is als de tipgevertelefoon en de telefoon [van verdachte] zich bij dezelfde gebruiker bevonden dan als ze zich bij verschillende, zich onafhankelijk van elkaar bewegende gebruikers bevonden.” Het NFI heeft kennelijk een manier gevonden om hier aan te kunnen rekenen. Dat moet, opnieuw, in een wiskundig model zijn gedaan want anders is er niets te berekenen. Welk model hiervoor gebruikt is, kunnen we niet weten, en de rechter heeft er voor zover wij dat kunnen nagaan ook niet naar gevraagd. Opnieuw hebben we hier een situatie waarin de gehanteerde logica correct is, maar waarin de uitvoering toch onnavolgbaar is omdat informatie over de berekeningen ontbreekt.

Overigens kent deze zaak nog een interessant aspect in de zin dat er niet één getal genoemd wordt, maar twee; wat de grenzen van 19 en 105 precies betekenen wordt echter niet helemaal duidelijk. De formulering suggereert dat er een

‘werkelijke’ of ‘objectieve’ waarde te vinden is die zich tussen deze twee uitersten bevindt. In het volgende hoofdstuk zullen we zien dat die interpretatie onjuist moet zijn, waarmee de betekenis van de getallen toch enigszins mistig blijft.

We hebben nu met een aantal praktijkvoorbeelden laten zien dat we de waarde van een bewijsmiddel alleen kunnen vaststellen als er ten minste twee scenario's met elkaar worden vergeleken. Als we maar één scenario bekijken hebben we geen manier om te bepalen of de gegevens bijzonder of zeldzaam zijn: je moet altijd vergelijken met een andere mogelijke situatie waarmee de gegevens wellicht ook verklaard kunnen worden. Dit wordt bevestigd door de algemeen aanvaarde richtlijnen in de wereld van de forensische wetenschap. Het belang hiervan kan moeilijk overschat worden, dus we citeren hier graag uit de aanbevelingen in een rapport van het ENFSI, het European Network of Forensic Science Institutes.⁵ Op pagina 10 lezen we:

“The findings should be evaluated given at least one pair of propositions: usually one based upon one party’s account of the events and one based upon an alternative (opposing party’s) account of the events. If no alternative can be formulated, the value of the findings cannot be assessed. In that case, forensic practitioners should state clearly that they are not reporting upon the value of the findings.”

Even later, op pagina 13, wordt deze boodschap nog eens herhaald:

⁵ ENFSI Guideline for Evaluative Reporting in Forensic Science, https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf.

“The report should stress that in the absence of an alternative proposition, it is impossible to evaluate the findings.”

In het licht van de voorbeelden die we gaven, komt deze aanbeveling natuurlijk niet meer uit de lucht vallen. Als iets bijzonder lijkt en dus een aanwijzing kan zijn, moet je altijd kijken naar de context waarbinnen dit eventueel bijzonder is. Je moet de gegevens (de ‘findings’) altijd afzetten tegen wat je in een andere omstandigheid verwacht, anders kun je onmogelijk vaststellen in hoeverre de gegevens bijzonder zijn.

Maar het ENFSI-rapport zegt nog veel meer. Op pagina 6 lezen we bijvoorbeeld:

“Evaluation [...] is based on the assignment of a likelihood ratio. Reporting practice should conform to these logical principles. This framework for evaluative reporting applies to all forensic science disciplines. The likelihood ratio measures the strength of support the findings provide to discriminate between propositions of interest. It is scientifically accepted, providing a logically defensible way to deal with inferential reasoning. [...] Forensic findings as such need to be distinguished from their evaluation in the context of the case. For the latter evaluative part only a likelihood ratio based approach is considered.”

Dit is een buitengewoon belangrijke les: niet alleen moet er altijd een alternatief geformuleerd worden, maar de likelihood ratio die we hiervoor introduceerden, is de enige manier om de sterkte van forensisch bewijs te meten. Als het gaat om het evalueren van bewijs zijn we aangewezen op de likelihood ratio en daarmee impliciet op Bayesiaans redeneren.

OORDEELSVORMING

De vraag die zich vervolgens opdringt is hoe de rechter het bewijs moet of kan gebruiken om tot een uiteindelijke conclusie of tot een bepaald oordeel te komen. Dat is de derde stap in ons redeneermodel: na de oorspronkelijke inschatting en het bewijsmateriaal is de vraag wat je er eigenlijk van moet vinden als alle gegevens de revue zijn gepasseerd. In feite vragen we dus naar een manier om binnen het redeneermodel tot een oordeel of uiteindelijke overtuiging te komen. Dat kan in principe geen heel eenvoudige zaak zijn, en om dat te illustreren beginnen we opnieuw met een hypothetisch voorbeeld.

Stel je voor dat een bepaalde persoon al maanden wordt vermist. Deze persoon leidt normaal gesproken een teruggetrokken leven, maar zelfs in die omstandigheid begint zijn afwezigheid na enkele maanden wel op te vallen. Uiteindelijk besluit het bevoegd gezag om het huis van deze vermiste persoon binnen te gaan, en men treft daar een lichaam aan in verregaande staat van ontbinding dat niet meer visueel te identificeren is. Bijna iedereen zal denken dat dit het lichaam van de vermiste persoon moet zijn. Maar toch, je wilt dat wel zeker weten, en je verlangt daarvoor een bepaalde hoeveelheid bewijs.

Wat is er nodig om overtuigd te worden van het feit dat dit het lichaam van de vermiste persoon is? Niet veel, want je bent door de omstandigheden al bijna overtuigd. Als je bijvoorbeeld weet dat de vermiste persoon ooit allebei de onderarmen heeft gebroken en je ziet de breuklijnen terug in het lichaam, dan zal dat waarschijnlijk alle twijfel wegnemen. Bedenk dat twee gebroken onderarmen weliswaar niet heel vaak voorkomen, maar ook weer niet extreem uitzonderlijk zullen zijn.

Maar stel je nu voor dat we een lichaam vinden in Zuid-Duitsland, ver weg van de woonplaats van de vermiste persoon. Er is nu geen enkele aanleiding om te denken dat het hier om deze vermiste persoon gaat, en ook de ontdekking van twee breuklijnen in de onderarmen zal ons daar waarschijnlijk niet van overtuigen. Dat is interessant omdat hetzelfde bewijs (twee gebroken onderarmen) in het ene geval wel maar in het andere geval niet overtuigend is.

De achterliggende reden hiervan is niet ingewikkeld: een hypothese die je al waarschijnlijk acht behoeft veel minder bewijs dan een hypothese die je tamelijk bizar acht. Om van een bizarre hypothese toch overtuigd te raken heb je heel veel bewijs nodig, terwijl een geloofwaardige hypothese veel minder bewijs vraagt. Deze gedachtegang impliceert dat overtuiging afhangt van twee dingen: de sterkte van het bewijs én de oorspronkelijke geloofwaardigheid van de hypothese waar het om gaat.

We benadrukken (opnieuw) dat de oorspronkelijke geloofwaardigheid in juridische zin ‘gewoon’ bewijs is of kan zijn. Wij gebruiken het begrip ‘bewijs’ hier om te verwijzen naar gegevens uit de specifieke casus waar het om gaat. We concluderen nu echter dat een overtuiging in een zaak uiteindelijk altijd subjectief en dus persoonlijk moet zijn. Immers, zelfs als twee rechters het bewijsmateriaal op precies dezelfde wijze waarderen, dan nog hebben ze normaal gesproken een andere oorspronkelijke opvatting. We illustreren dit nu eerst met enkele situaties uit de praktijk.

We beginnen met een levensdelict.⁶ Op 29 juli 2012 werd in een kajuitbootje aan de Hobbemakade te Amsterdam het levenloze lichaam van een man aangetroffen. Het slachtoffer lag op dat moment op zijn rug. Zijn onderlichaam was

⁶ ECLI:NL:GHAMS:2020:818.

ontbloot en zijn broek tot de enkels naar beneden getrokken. Aan de binnenzijde van zijn bovenarmen, in zijn mond en naast zijn strottenhoofd zijn bloeditstortingen aangetroffen, naast andere letsels aan onder andere zijn ribben, lever en oogleden.

De verdachte wordt ervan beschuldigd dat hij het slachtoffer met geweld om het leven heeft gebracht. Om tot een bewezenverklaring te kunnen komen, is in de eerste plaats vereist dat kan worden vastgesteld dat de dood van het slachtoffer het gevolg is geweest van geweldstoepassing door een ander. Over de doodsoorzaak van het slachtoffer is in eerste aanleg gerapporteerd door twee forensisch pathologen en door een forensisch arts. In hoger beroep is gerapporteerd door een forensisch patholoog.

In de uitspraak van het hof zijn twee passages aan te wijzen waar het om achtergrondkennis gaat. We citeren deze passages hier helemaal:

“Het hof onderkent dat diverse omstandigheden wijzen in de richting van de verdachte als mogelijke betrokkene bij het overlijden van [slachtoffer]. Zo kan als vaststaand worden aangenomen dat de verdachte en [slachtoffer] elkaar kenden en elkaar frequent zagen, dat de verdachte geregeld op de boot (waar [slachtoffer] verbleef) kwam en dat zij kort voor het overlijden van [slachtoffer] samen waren. Daarnaast hebben getuigen verklaard dat de verdachte geweld jegens het slachtoffer niet schuwde en heeft de verdachte aanvankelijk gelogen dat hij [slachtoffer] niet kende en nog nooit op diens boot was geweest. Deze omstandigheden maken echter niet dat alsnog kan worden vastgesteld waaraan [slachtoffer] is komen te overlijden of dat de verdachte op strafbare wijze betrokken was bij diens overlijden.

[...]

Het hof kan zich weliswaar vinden in het oordeel van de diverse deskundigen dat de combinatie van bij [slachtoffer] aangetroffen letsels verdacht is te noemen, maar daar staat tegenover dat verschillende letsels een plausibele niet-criminele oorzaak kunnen hebben en dat bepaald niet vaststaat dat [slachtoffer] alle letsels op of rond de tenlastegelegde datum, 29 juli 2012, heeft opgelopen. Bij dit laatste betreft het hof dat [slachtoffer] verkeerde in een circuit van daklozen en (alcohol)verslaafden en hij volgens getuigen wel vaker klappen kreeg van personen uit dat circuit.”

Uit het eerste deel van het citaat concluderen we dat het hof de omstandigheden en de achtergrond heel serieus meeweegt. Er wordt een aantal zaken benoemd die niet rechtstreeks met het mogelijke misdrijf te maken hebben, maar die wel belangrijk worden geacht als context. In dit verband is het woord ‘alsnog’ heel betekenisvol. Het woord suggereert – terecht – dat de context ervoor zorgt dat het hof waarschijnlijk sneller geneigd zou zijn de verdachte te veroordelen dan wanneer de gehele context er niet was geweest. Om tot een veroordeling te komen moet er echter, zo lezen wij de uitspraak, wel iets zijn dat de verdachte verbindt met plaats en tijd van het misdrijf.

Aan het einde van het tweede deel van het citaat zien we ook een verwijzing naar de achtergrond en de context van de gebeurtenissen. Je zou kunnen stellen dat de hier benoemde achtergrondkennis juist nadruk legt op de gedachte dat het verleden heeft laten zien dat verdachte wel vaker het slachtoffer werd van een vechtpartij en dat het dus best aannemelijk is dat de verwondingen die hij hier heeft opgelopen op die manier te verklaren zijn.

Met andere woorden, de context en achtergrond spelen bij de overwegingen van het hof een belangrijke rol. Overwegingen die niet direct betrekking hebben op de gebeurtenissen zelf maar die wel van belang zijn bij de oordeelsvorming hierover. Precies zoals het volgens het Bayesiaanse schema zou moeten.

Ons tweede voorbeeld gaat over een inbraak in Rotterdam.⁷ Een buurman heeft gezien dat drie Marokkaanse mannen de woning waar de inbraak plaatsvond door een ingeslagen raam verlieten en heeft hen enige tijd achtervolgd. Deze en andere getuigen zijn gehoord, kennelijk zonder dat daaruit aanknopingspunten voor verder onderzoek naar voren kwamen. Tevens stelde de technische recherche sporen veilig, waaronder een bloedspoor op een gordijn in de woonkamer. Het NFI stelde vast dat dit spoor overeenkomt met eerder ontvangen materiaal en leidt tot een verdachte. De verdachte, die ten tijde van het feit in Rotterdam werkt, verklaart een twee-eiige tweelingbroer te hebben. Hij weet niet hoe zijn DNA op de plaats delict is terechtgekomen en beroept zich bij zijn derde verhoor vooral op zijn zwijgrecht. We citeren nu uit de uitspraak:

“Het enige bewijs dat de verdachte in deze zaak rechtstreeks met de tenlastegelegde inbraak verbindt is het aangetroffen bloedspoor – kennelijk een daderspoor – waarvan het DNA-profiel met dat van de verdachte overeenstemt en aldus een zeer aanzienlijke objectieve bewijswaarde heeft. Elke aanwijzing ontbreekt dat in de bewijsketen, dan wel bij het onderzoek van het (bloed) spoor fouten zijn gemaakt. De verdachte heeft bovendien zijn betrokkenheid bij de inbraak weliswaar ontkend,

⁷ ECLI:NL:GHSGR:2011:BQ9005.

maar die ontkenning op generlei wijze onderbouwd of ter zake gespecificeerd verweer gevoerd; het hof onderkent overigens dat zulks door het tijdsverloop bepaald niet eenvoudig zou zijn geweest. Bovendien stelt het hof vast dat de verdachte, die van Marokkaanse origine is, in zoverre voldoet aan de door de eerder genoemde getuige gegeven beschrijving van de mannen die de woning verlieten. Hij is daarnaast eerder met Justitie ter zake inbraak in aanraking gekomen, werkte in dezelfde gemeente als waar werd ingebroken en hem wordt in de onderhavige zaak een tweede (poging tot) inbraak verweten, ten aanzien waarvan het hof – zoals hiervoor uiteen is gezet – tot een bewezenverklaring komt; in zoverre wekken de beschuldiging én de match ook geen verbazing.”

Vooraf de laatste zin is voor ons hier interessant. Er is zeer sterk bewijs, namelijk een DNA-match. De verwijzing van verdachte naar zijn twee-eiige tweelingbroer is natuurlijk bedoeld om te suggereren dat het DNA-profiel ook van die broer afkomstig kan zijn. Het DNA van de broer kan niet zomaar bepaald worden, maar je kunt wel uitrekenen wat de kans is dat het profiel van de broer overeenkomt met dat van de verdachte. Twee-eiige broers of zussen zijn genetisch net zo gerelateerd als normale broers of zussen, en deze berekening laat zien dat deze broer met een zekerheid grenzende waarschijnlijkheid niet hetzelfde profiel zal hebben als de verdachte. In die omstandigheden maakt het hof de zaak af door te stellen dat er achtergrondinformatie is die erop wijst dat verdachte ook in algemene zin binnen een categorie van potentiële verdachten kan vallen. Het hof zegt met zoveel woorden dat het van zichzelf al zeer sterke DNA-bewijs misschien in absolute isolatie niet doorslaggevend kan zijn, maar dat de omstandigheden deze verdachte zeker niet a priori uitsluiten. Voor het hof was dat voldoende voor een veroordeling.

We zagen eerder bij het voorbeeld van het lichaam in de woning van de vermiste persoon dat voor een voor de hand liggende hypothese weinig bewijs nodig is, terwijl een bizar geachte hypothese veel meer bewijs verlangt. Het voorbeeld van de inbraak zit hier tussenin. Er is extreem sterk bewijs in de vorm van een passend DNA-profiel, en de omstandigheden zijn zo dat de hypothese dat deze verdachte inderdaad de dader is dan wel niet waarschijnlijk is, maar zeker niet uitgesloten kan worden. Samen met het zeer sterke bewijs is dat voor het hof voldoende.

We zullen later, in hoofdstuk 4, terugkomen op de vraag of deze redenering in strijd is met de onschuldpresumptie – wij zullen beargumenteren van niet. Voor nu stellen we vast dat het Bayesiaanse redeneerschema context en bewijs als het ware combineert om tot een overtuiging te komen. Bij een zeer sterke oorspronkelijke verdenking is er niet zo veel verder bewijs nodig, maar bij een situatie waarin er geen reden is om de verdachte te verdenken is veel meer nodig. Tussen deze twee uitersten zitten allerlei situaties waarbij afgewogen moet worden of context en bewijs samen genoeg overtuigingskracht hebben. Deze manier van denken is correct vanuit het perspectief dat we hier hanteren, en zorgt voor een navolgbare uitspraak.

We vervolgen met een derde voorbeeld, een winkeldiefstal uit de Media Markt in Zoetermeer.⁸ We citeren direct uit de uitspraak – dit citaat geeft voldoende context van de zaak.

“Uit de camerabeelden blijkt dat de verdachte en medeverdachte samen de winkel binnenkomen en direct naar de fotoafdeling gaan. Daar is te zien dat de verdachte producten aanwijst, welke de medeverdachte vervol-

⁸ ECLI:NL:PHR:2022:1191.

gens pakt en uit de verpakking haalt. Deze handelingen worden duidelijk omschreven door de verbalisant die de camerabeelden heeft bekeken. Direct na het bezoek aan de fotoafdeling verlaten zij de Media Markt zonder iets af te rekenen. De aangever herkent de verdachte bij het verlaten van de winkel van een foto van een winkeldiefstal bij de Media Markt in Utrecht. Ongeveer zeven minuten later komt de aangever aan op de fotoafdeling en vindt diverse lege verpakkingen. De Media Markt heeft omschreven welke lege verpakkingen zijn aangetroffen. Daarmee is naar het oordeel van de politierechter voldoende vast komen te staan dat de verdachte samen met de medeverdachte de genoemde producten heeft weggenomen. Aan de gedragingen van de medeverdachte heeft verdachte bijgedragen, reeds door kennelijk met de auto die zij op haar naam had gehuurd naar de winkel te gaan (de tijd die daarvoor nodig is matcht met de waarnemingen van de APNR beelden terwijl daaruit ook blijkt dat met de auto niet naar winkels in de buurt is gereden en geen andere winkels zijn bezocht: verdachte is uitsluitend naar de Media Markt gereden en weer terug), toe te staan dat medeverdachte daarin kennelijk meereed dan wel alleen naar die winkel van de Media Markt te gaan, daarin mee te lopen met de medeverdachte, een product aan te wijzen dat de medeverdachte vervolgens uitpakt en meeneemt, door met hem mee door de winkel te lopen en samen langs de kassa de winkel verlaat zonder iets af te rekenen, terwijl zij daar niet voor enig ander gebleken redelijk doel naar de Media Markt was gegaan. Aan de overtuiging draagt bij dat zij ook bekend is van een diefstal bij een vestiging van de Media Markt met medeverdachte in Utrecht.”

Het gaat ons natuurlijk om de laatste zin, waarin onomwonden wordt uitgesproken dat eerdere vergelijkbare criminele activiteiten van de verdachte bijdragen aan de overtuiging van de rechtbank dat verdachte schuldig is. En zo zou het ook moeten gaan. Die eerdere criminele activiteiten dragen niets bij aan concreet bewijs rond de diefstal in Zoetermeer, maar zijn kennelijk wel van invloed op de algemene waardering van de rechtbank. We hebben in dit hoofdstuk tot nu toe gezien dat dit precies is hoe het zou moeten, in ieder geval vanuit een logisch redeneerperspectief.

HET REDENEERMODEL SAMENGEVAT IN GANGBARE TERMINOLOGIE

We hebben het Bayesiaanse redeneermodel uitvoerig beschreven. In principe bestaat het uit drie onderdelen.

Allereerst hebben we vastgesteld dat een mens bij zijn oordeelsvorming simpelweg niet bij nul kan beginnen. We hebben al een idee of een mening voorafgaand aan de verwerking van aanwijzingen in een zaak. In de literatuur spreken we in dit verband over ‘a priori’ kansen of kennis. Deze a priori kennis is niet objectief, en hangt af van de informatie of het wereldbeeld van de uitdrager ervan.

Deze a priori kennis of mening kan bijgesteld worden door aanwijzingen in een zaak. We hebben gezien dat we hierbij minimaal twee hypothesen nodig hebben, en dat we moeten kijken naar de mate waarin de aanwijzingen in een zaak al dan niet verrassend zijn onder elk van die hypothesen. Dit wordt uitgedrukt met een likelihood ratio, die ons in principe vertelt welke van de hypothesen de aanwijzingen (de data dus) het beste verklaart, en hoeveel keer beter. Dit kan bij gebrek aan een goed wiskundig model doorgaans niet numeriek, maar dat hoeft het gebruik van een likelihood

ratio niet in de weg te staan. Als er een deskundige wordt ingeschakeld, dan zal deze iets over de likelihood ratio zeggen, zoals we hiervoor al verschillende keren hebben gezien.

Ten slotte bepalen de a priori kennis en de likelihood ratio samen onze overtuiging na het verwerken van alle informatie. Dit wordt dan de ‘a posteriori’ overtuiging genoemd.

In een meer wiskundig-formele benadering kun je deze concepten in formulevorm opschrijven en aan elkaar relateren; men spreekt dan meestal over ‘prior odds’ en ‘posterior odds’ die met de likelihood ratio in een formule aan elkaar gerelateerd worden. We benadrukken dat deze formule niet de kern van Bayesiaans denken is – ze is slechts de wiskundige weergave van de logica zoals we die in dit hoofdstuk hebben besproken. Om deze logica te begrijpen en om ermee te werken is die wiskundige formulering niet cruciaal, en we gebruiken deze dan ook niet in dit boek.

PRAKTISCHE UITVOERING – VOORBEELDEN

Het is nog niet zo gemakkelijk om zelf op de hier geschetste manier argumenten te ontwikkelen. Hoewel de logica, zo vinden wij, helemaal niet zo ingewikkeld is, is er zoals bij alles oefening en routine nodig om deze je eigen te maken. Maar als je dat doet, zal de situatie zeker helderder worden en wordt een uitspraak waarschijnlijk beter navolgbaar. We kunnen dat illustreren aan de hand van enkele voorbeelden die we hierboven bespraken.

In de zaak Sally Clark had de rechtbank zelf twee scenario’s moeten formuleren en de vraag naar de oorspronkelijke inschatting daarvan aan deskundigen moeten overlaten. Deze zaak was in zoverre bijzonder dat de waarde van het bewijs op een volkomen triviale manier neutraal bleek te zijn: de

dode baby's zouden hoe dan ook onder beide scenario's (dubbele moord of dubbele wiegendood) aangetroffen zijn, en in dit geval vroeg de rechter de deskundige dus eigenlijk naar de oorspronkelijke inschatting: hoe vaak kunnen we überhaupt twee gestorven kinderen verwachten als er geen sprake is van moord? Als de rechter Bayesiaans had geredeneerd, dan zou de vraag naar de oorspronkelijke inschatting voor dubbele moord ook gesteld zijn, en zou de zaak zonder enige twijfel heel anders afgelopen zijn.

In de zaak van de Hobbemakade ligt het allemaal wat subtieler. We citeren uit het arrest wat er gezegd wordt over de bijdragen van de geraadpleegde deskundigen.

“[deskundige 1] heeft niet kunnen concluderen dat een accidentele oorzaak meer of minder waarschijnlijk is dan een niet-accidentele oorzaak. Verder heeft zij weliswaar gesteld dat sommige letsels bij [slachtoffer] een ‘suspecte combinatie’ vormen, maar een verdergaande conclusie heeft zij niet kunnen trekken.

[...]

[deskundige 2] heeft geconcludeerd dat geen zekere doodsoorzaak is aan te wijzen. Zijns inziens is niet uitgesloten dat verwurging heeft plaatsgevonden, maar bepaalde letsels kunnen ook zonder meer worden verklaard door herhaaldelijk vallen (mede gelet op de hoge alcoholconcentratie in het lichaam van [slachtoffer]).

[...]

[deskundige 3] heeft geconcludeerd dat het overlijden van [slachtoffer] veel waarschijnlijker het gevolg is van samen-

drukkend en/of samensnoerend geweld op de hals zoals door wurghandelingen (in de vorm van manuele strangulatie of ligatuurstrangulatie) dan van een val.

[...]

[deskundige 4] heeft betoogd dat de bloedingen bij het strottenhoofd kunnen zijn ontstaan bij de sectie, door bloedstuwning of door krachtsinwerking, bijv. door een val of wurging, maar dat die bloedingen niet onderscheidend genoeg zijn voor een duidelijke conclusie.”

In deze zaak is deskundige 4 de enige die helder maakt dat zijn oordeel moet gaan over het onderscheidend vermogen van het sporenbewijs. Bij de andere deskundigen is of onduidelijk wat ze rapporteren, of lijken ze te rapporteren over een vraag die ze niet kunnen beantwoorden. Zo heeft deskundige 1 niet kunnen concluderen dat een accidentele oorzaak meer of minder waarschijnlijk is dan een niet-accidentele oorzaak, en rapporteert deskundige 2 dat er volgens hem geen zekere doodsoorzaak is aan te wijzen. Deskundige 3 concludeert dat de verwondingen veel waarschijnlijker het gevolg zijn van geweld dan van een val. Bij deskundige 1 en 2 is duidelijk dat zij rapporteren over de waarschijnlijkheid van de val dan wel van het geweld. Bij deskundige 3 is onduidelijk of hij bedoelt dat de verwondingen onder een aanname van geweld waarschijnlijker zijn dan onder een aanname van een val, of dat hij bedoelt dat het geweld waarschijnlijker is. In het eerste geval rapporteert hij over het onderscheidend vermogen van het bewijs gezien de beide hypothesen en dat is precies de rol van de deskundige. In het tweede geval laat hij zich net als deskundige 1 en 2 uit over de verschillende hypothesen zelf (hoe waarschijnlijk is het geweld; hoe waarschijnlijk is een val).

Voor het beantwoorden van die vraag is immers contextuele informatie nodig, zoals we hebben gezien.

De uiteindelijke vraag die het hof moet beantwoorden is: “Heeft de verdachte geweld uitgeoefend dat tot de dood van het slachtoffer heeft geleid?” Het arrest wekt de indruk dat de vraag aan de deskundigen ook in die termen is gesteld. Het zou beter geweest zijn als de vraag anders zou zijn geformuleerd: “In hoeverre zijn de sporen onderscheidend voor de verschillende hypothesen?” Wellicht is op dit punt betere afstemming tussen hof en deskundigen mogelijk, maar dan moeten alle partijen wel bekend zijn met Bayesiaans redeneren.

Ons volgende voorbeeld is een geval waarin een moeder verdacht wordt van het toebrengen van ernstig letsel aan haar baby.⁹ Op zondag 28 augustus 2016 aan het einde van de middag is verdachte samen met haar dochtertje en de baby alleen thuis. De man van verdachte is op dat moment bij een jarige buurvrouw op bezoek. Verdachte is niet meegegaan naar de verjaardag omdat de baby wat onrustig was. Andere bijzonderheden aan de baby zijn zowel verdachte als haar man die dag niet opgevallen. Op enig moment komt verdachte, met de baby slap in haar armen, in paniek bij de woning van de buurvrouw. De baby wordt met een ambulance naar het ziekenhuis gebracht waar hij wordt opgenomen. In het ziekenhuis worden onder meer bloeduitstorting en onder het harde hersenvlies van de grote en de kleine hersenen en meerdere netvliesbloedingen in beide ogen geconstateerd. De baby overleeft de gebeurtenissen wel.

Een geraadpleegde deskundige concludeerde in zijn rapport dat het ontstaan van de hiervoor genoemde letsels bij de baby los van elkaar bezien “waarschijnlijker” zijn en in com-

⁹ ECLI:NL:GHARL:2023:9286.

binatie met elkaar bezien “veel waarschijnlijker” zijn bij een letselmechanisme met impact (botsend geweld op of tegen het hoofd), bij een heftig schudincident (repeterend acceleratie-deceleratie-trauma) of bij een combinatie van beide, dan bij een andere oorzaak (medisch of traumatisch). De moeder van het kind ontkent elke beschuldiging en verklaart dat het kind spontaan in de problemen is gekomen, en dat ze in paniek hulp heeft gezocht door naar haar man bij de buurvrouw te rennen.

Het hof doet nu iets opmerkelijks waar we even bij stil moeten staan. We citeren uit de uitspraak en leggen deze daarna uit.

“Daarnaast betreft het hof bij zijn oordeel dat [de deskundige] in zijn rapport met betrekking tot de gestelde hypothesen heeft geconcludeerd dat het ontstaan van de hiervoor genoemde letsels bij [naam kind] in combinatie met elkaar bezien ‘veel waarschijnlijker’ is bij een letselmechanisme met impact dan bij een andere oorzaak en dat deze conclusie een bewijskracht heeft van een orde-grootte van 100 tot 10.000. Het hof stelt vast dat de door de deskundige aangegeven orde-grootte zich bevindt op de vierde schaal binnen de door het NFI gehanteerde numerieke definitie van waarschijnlijkheidstermen. Boven deze schaal bevinden zich nog twee schalen, namelijk ‘zeer veel waarschijnlijker’ of ‘extreem veel waarschijnlijker’. Deze laatste twee termen uit de gebezigde waarschijnlijkheidsterminologie hebben een bewijskracht van respectievelijk 10.000 tot 1.000.000 en groter dan 1.000.000. De door het NFI aan de verschillende waarschijnlijkheidstermen toegekende getallen als geheel genomen, maken naar het oordeel van het hof dat aan de door de deskundige in dit geval getrokken conclusie op zichzelf genomen relatief

beperkte bewijskracht toekomt. Voor een bewezenverklaring is nodig dat – in juridische termen – met aan zekerheid grenzende waarschijnlijkheid is komen vast te staan dat het letsel bij [naam kind] door verdachte is toegebracht.”

We zien dat er in dit citaat numerieke waarden gegeven worden voor de bewijskracht – de likelihood ratio. Die numerieke waarden zijn we eerder tegengekomen bij de zaak met het ‘loopje’ en bij de tipgevertelefoonzaak. Wat hier echter nieuw is, is de relatie tussen de getallen en de voorgestelde verbale vertaling ervan. In feite gaat het hof van de bewoordingen van de verdachte naar een getallenrange die door het NFI is opgesteld. Deze gang van zaken is al enigszins bevreemdend, want het suggereert dat de bevindingen van de expert op die manier objectief in getallen kunnen worden uitgedrukt. Dat is natuurlijk niet zo: hoe zou je kunnen onderzoeken wat het verschil is tussen twee experts waarvan er een zegt dat het bewijs “zeer veel waarschijnlijker” onder de ene dan onder de andere hypothese is, versus een expert die dat “extreem veel waarschijnlijker” acht? De woorden zijn zonder meer subjectief en niet zomaar met elkaar te vergelijken.

Het hof doet dat toch en stelt vervolgens dat de bevindingen van de expert een “relatief beperkte bewijskracht” opleveren. Dat is allemaal goed en wel, maar de laatste zin uit het citaat is een conclusie die op basis van de overwegingen niet zomaar kan volgen. Immers, het hof springt hier van de sterkte van het bewijs zonder enige toelichting over op het feit dat ze niet overtuigd is van de schuld van verdachte. In zichzelf volgt dit niet uit de redenering omdat we hebben gezien dat voor een overtuiging altijd contextuele argumenten een rol moeten spelen.

Nu zou een criticus kunnen aanvoeren dat we ons hier wel erg formeel opstellen, maar wij willen toch graag een

lans breken voor exacte logische redeneringen. De vraag die hierin (ook) beantwoord moet worden, is of de context van de situatie aanleiding geeft om de schuldhypothese al heel begrijpelijk of acceptabel te vinden. Waarschijnlijk zou het hof anders concluderen als deze moeder al eerder voor een dergelijk vergrijp zou zijn veroordeeld, of als er bijvoorbeeld bedreigingen richting het kind zouden zijn geuit. In dit geval lezen we niets in de uitspraak dat erop zou wijzen dat de schuldhypothese al van tevoren enigszins redelijk of acceptabel gevonden zou kunnen worden. Situaties als deze komen in Nederland niet veel voor: het aantal per jaar is op één hand te tellen dus als er geen bijzondere reden is om te denken dat hier zoiets aan de hand zou kunnen zijn, kan het niet anders dan dat de schuldhypothese in beginsel weinig waarschijnlijk lijkt. Als het bewijs vervolgens ook niet voldoende onderscheidend is, dan is een vrijspraak begrijpelijk.

Als het hof haar conclusie op deze manier zou hebben verwoord, dan zou de conclusie navolgbaar en transparant geweest zijn. Nu is dat eigenlijk niet zo, omdat de precieze bewoordingen ruimte laten voor twijfel over de vraag of het hof eigenlijk wel correct redeneert.

Hopelijk heeft de lezer inmiddels een idee over de toegevoegde waarde van het Bayesiaanse redeneermodel. Zoals we al eerder schreven: het werken met dit model impliceert geenszins dat we bij een zaak moeten gaan rekenen. Daarover gaat het volgende hoofdstuk.

HOOFDSTUK 3

De rol van getallen en modellen

In 2022 verscheen bij uitgeverij Ambo Anthos het boek *Het verraad van Anne Frank*, geschreven door Rosemary Sullivan. Zoals de titel al doet vermoeden ging het boek over de vraag wie Anne Frank heeft verraden. De zaak werd voorgesteld als een coldcaseonderzoek zoals je dat bij een recent misdrijf ook zou kunnen doen.

In het boek zelf wordt niet gerekend, maar de presentatie ervan ging vergezeld met de publieke uitspraak dat de kans 85% zou zijn dat ‘de notaris’ schuldig zou zijn aan het verraad. Dit getal was het resultaat van een berekening. Deze berekening werd uitgevoerd door klimaatfysicus Frans Alkemade. Desgevraagd weigerde hij om ons inzage te geven in de manier waarop hij tot dat getal gekomen was.

Het boek – maar vooral dit getal – zorgde voor grote nationale opwinding, die ertoe leidde dat het boek zelfs uit de handel werd genomen – in Nederland een zeer uitzonderlijke gebeurtenis. Bijna iedereen voelde intuïtief wel aan dat een kans als dit uitrekenen tamelijk bizar is, maar het is nog niet eens zo makkelijk om de vinger precies op de zere plek te leggen en aan te geven waarom dat dan zo is.

Precies dát willen we in dit hoofdstuk doen: de vinger op de zere plek leggen. In het vorige hoofdstuk hebben we uitgelegd dat wiskunde en rekenen ondergeschikt zijn aan dat wat die wiskunde geacht wordt te modelleren en te beschrijven. In het bijzonder hebben we uitgelegd dat overtuiging een persoonlijke aangelegenheid is, en dat daaraan geen objectief getal gekoppeld kan worden. Dat wil niet zeggen dat getallen

helemaal nutteloos zijn, maar wel dat ze met grote terughoudendheid en vooral met wijsheid ingezet moeten worden. We zullen uitleggen waarom dit getal van 85% buiten de persoonlijke overtuiging van de rekenaar betekenisloos is, en we illustreren onze argumenten met een aantal verdere praktijkvoorbeelden.

KANSEN EN INFORMATIE

We beginnen met uit te leggen hoe we een percentage of een kans in de context van een rechtszaak dienen te interpreteren. Wat betekent het eigenlijk om te spreken van een kans van 85% dat de notaris schuldig is aan het verraad van Anne Frank?

Kansen kunnen nogal veel verschillende dingen betekenen. Als we naar een casino gaan om roulette te spelen, dan kun je spreken over de kans dat het balletje bij de volgende draai aan het roulettewiel op nummer 7 terechtkomt. Als er geen reden is om binnen alle mogelijke uitkomsten onderscheid te maken, kennen we alle uitkomsten dezelfde kans toe. Als je een dobbelsteen gooit zal iedereen het ermee eens zijn om de kans op bijvoorbeeld 5 ogen op $1/6$ te zetten; er zijn nu eenmaal zes mogelijke uitkomsten.

Het aardige van deze interpretatie is dat ze goed overeenkomt met onze ervaring, in de volgende zin. Als we heel vaak met een dobbelsteen gooien, dan zul je zien dat de fractie van de keren dat je, bijvoorbeeld, 5 ogen gooit, steeds dichter in de buurt van $1/6$ komt te liggen. Dat is een ervaringsfeit, en voor veel mensen is dat ervaringsfeit in zichzelf al reden genoeg om $1/6$ als de kans op 5 ogen te nemen. Sterker nog, als je een asymmetrisch voorwerp vaak zou gooien, bijvoorbeeld een punaise, en je wilt weten wat de kans is dat

de punaise met de pin naar boven eindigt, dan heb je niets anders dan de ervaring. Er is geen symmetrie meer zoals bij de dobbelsteen, dus geen reden om de twee mogelijke uitkomsten (pin naar boven of naar beneden) dezelfde kans te geven. De ervaring moet dan uitwijzen hoe vaak de twee mogelijkheden optreden bij het herhaald gooien van een punaise.

De reden dat deze interpretatie van kans voor ons intuïtief heel goed werkt, heeft ermee te maken dat de verschillende draaien aan het roulettewiel, of de verschillende worpen van de dobbelsteen of punaise, elke keer min of meer gelijk zijn. De omstandigheden veranderen niet of nauwelijks. Dat is cruciaal. Als we het roulettewiel de ene keer recht houden en de andere keer schuin, dan zien we wel in dat dat niet eerlijk is, en dat je de uitkomsten dan niet meer zo makkelijk met elkaar kunt vergelijken. Het is dus belangrijk dat we het elke keer over min of meer dezelfde draai of dezelfde worp hebben.

Tot zover is er geen controverse. De kans op een bepaalde uitkomst heeft een bevredigende interpretatie: deze kans beschrijft hoe vaak een bepaalde uitkomst voorkomt bij het vaak herhalen van het experiment, en ze is daarmee in essentie een voorspellende kans.

Maar hoe zit het dan met de kans op schuld van een bepaalde verdachte? Elke rechtszaak is uniek, en het is duidelijk dat twee verschillende zaken weinig met elkaar te maken hebben en zeker geen vergelijkbare serie vormen zoals bij roulette of dobbelstenen. De gebeurtenis waarop we de kans willen bepalen, het al dan niet schuldig zijn van de verdachte, heeft bovendien al plaatsgevonden. Kunnen we dan überhaupt wel spreken over de kans dat een specifieke verdachte schuldig is zoals we bij dobbelen konden spreken over de

kans dat de volgende worp 5 ogen zal zijn? Het kan zeker niet om een voorspellende kans gaan, zoveel is duidelijk.

Dit is een subtiële kwestie en we moeten dit zorgvuldig bespreken. Alles hangt af van wat je wilt uitdrukken met een kans. Hier is een voorbeeld. Stel dat onderzoek uitgewezen heeft dat 80% van de verdachten die de laatste tien jaar voor de rechter hebben gestaan daadwerkelijk schuldig is.¹ Je kunt nu (in ieder geval in gedachten) een willekeurig persoon kiezen uit alle mensen die de afgelopen tien jaar voor de rechter hebben gestaan. De kans dat deze persoon schuldig is, is dan 80%. Dit is gewoon weer een voorspellende kans zoals bij roulette: als je dit maar vaak genoeg herhaalt, dan zul je zien dat je in ongeveer 80% van de keren een daadwerkelijk schuldige verdachte kiest.

Het moge echter duidelijk zijn dat deze interpretatie van een kans in een concrete rechtszaak geen rol speelt. In een specifieke rechtszaak dienen we te kijken naar de omstandigheden en bewijsmiddelen in die zaak. Net als de eerder genoemde kans van 85% dat de notaris de verrader was van Anne Frank, moet de kans dat een verdachte schuldig is iets anders betekenen; deze kans moet gerelateerd worden aan de karakteristieken en gegevens uit die specifieke zaak, en geen andere. Kansen in het recht kunnen geen voorspellende kansen zijn zoals bij het werpen van een dobbelsteen.

Maar wat zou die 85% dan wél kunnen betekenen? Wat is een kans in het recht dan wel? Feit is dat de verdachte of wel of niet schuldig is – dus de uitspraak dat een verdachte met een bepaalde kans schuldig is, lijkt in eerste instantie een beetje vreemd. Echter, als je er even over nadenkt, dan zie je dat kansrekening in het recht geen kwestie is van voorspel-

¹ Deze 80% is geen feit. Wij kennen het juiste percentage niet maar dat is voor ons verhaal niet belangrijk.

len, maar van informatie. Een kansuitspraak in het recht, over een gebeurtenis die zich in het verleden heeft afgespeeld, is een uitspraak over de geloofwaardigheid van een hypothese, gegeven de informatie die beschikbaar is.

Dit belangrijke punt werken we hieronder uit, en illustreren we eerst met dobbelsteenwerpen. Stel je voor dat Ronald een ondoorzichtige beker heeft met één dobbelsteen. Hij schudt de beker en zet hem ondersteboven op tafel, met de dobbelsteen eronder. Als iemand nu gevraagd wordt naar de kans dat er 5 ogen liggen, dan lijkt $1/6$ het enige redelijke antwoord, hoewel de gebeurtenis ‘5 ogen’ of wel of niet heeft plaatsgevonden. Het punt is dat we niet weten welke uitkomst opgetreden is, en daarom vinden we die kans van $1/6$ redelijk.

Maar stel nu dat Ronald de beker een klein beetje optilt en zonder dat Lonneke het kan zien de uitkomst bekijkt. Voor Ronald verandert alles nu: de kans op 5 ogen is ofwel 0% ofwel 100%, al naar gelang de uitkomst die hij heeft gezien. Voor Lonneke echter is de kans dat er 5 ogen gegooid zijn nog steeds $1/6$. De twee deelnemers aan dit experiment, Ronald en Lonneke, hebben nu verschillende informatie en daarmee verandert kennelijk hun kansinschatting. Er is dus niet zoiets als dé kans dat er 5 ogen liggen: deze kans is persoonlijk en hangt af van de informatie die de verschillende personen hebben. Als Ronald vervolgens prijsgeeft dat het aantal ogen oneven is, dan zal Lonneke haar kans op 5 ogen op basis van die informatie bijstellen van $1/6$ naar $1/3$: er zijn nu immers nog maar drie mogelijke uitkomsten voor haar.

Deze interpretatie van kans is zeker niet voorspellend meer. Immers, de gebeurtenis waarin we geïnteresseerd zijn, heeft al plaatsgevonden. Het is een kansuitspraak over een gebeurtenis in het verleden waarover we misschien niet alles

weten en waarover we onze onzekerheid kunnen uitdrukken in een kans.

De analogie met rechtszaken is zeer sterk hier. Het misdrijf is wel of niet door de verdachte gepleegd, net zoals de dobbelsteen onder de beker wel of niet 5 ogen boven heeft liggen. De verdachte weet doorgaans of hij het misdrijf heeft gepleegd, en voor hem is de kans op schuld dus 0% of 100%, net zoals dat voor Ronald het geval was, die immers ook volledige informatie had. Andere betrokkenen bij het proces weten minder, en zullen aan het schuldscenario een andere kans toekennen, net zoals Lonneke dat moest doen op basis van de gegevens die zij had. Verder kunnen kansinschattingen in een proces veranderen op basis van de gegevens die besproken of ontdekt worden, net zoals Lonneke dat deed toen zij nieuwe informatie van Ronald ontving.

Niet iedereen volgt deze interpretatie van kansen in strafzaken. Zo stelt statisticus Henk Elffers:

“Er bestaat geen zinvolle interpretatie van subjectieve waarschijnlijkheid in de rechtbank. Ofwel heeft de verdachte het gedaan, of niet. Ik vind het onaanvaardbaar dat een rechter de a priori waarschijnlijkheid zou evalueren van mijzelf die een bepaalde moord begaat, als iets anders dan nul. Ik kan oprecht zeggen dat de waarschijnlijkheid dat ik zo’n moord heb gepleegd, nul is.”²

Natuurlijk heeft Elffers gelijk dat de verdachte het wel of niet heeft gedaan. Maar ook als Elffers van zichzelf weet dat hij het misdrijf niet begaan heeft, is het nog steeds zo dat iemand anders dat niet zeker kan weten. Dat kun je

² Henk Elffers, “Dienststroosterdata in het strafproces”, in M.J. Sjerps en J.A. Coster van Voorhout (red.), *Het onzekere bewijs. Gebruik van statistiek en kansrekening in het strafrecht*, Deventer: Kluwer 2005, p. 148.

onaanvaardbaar noemen, maar eigenlijk ontkent Elffers hier de realiteit en heeft een rechter geen andere keus.

Een kansuitspraak is dus persoonlijk, en afhankelijk van de informatie die iemand heeft en zijn of haar kennis van zaken. Algemene of specialistische kennis kan en zal hierbij inderdaad ook een rol spelen, want bijzondere kennis (of juist de afwezigheid ervan) kan ertoe leiden dat informatie door iedereen op een andere manier wordt verwerkt en geïnterpreteerd.

Er bestaat dus helemaal geen ‘objectieve’ kans dat de notaris Anne Frank heeft verraden. Feitelijk heeft hij dat wel of niet gedaan, maar iemands mening of overtuiging hierover kan niet anders dan gebaseerd zijn op zijn eigen verwerking van de gegevens die hij in de overwegingen mee wil of kan nemen. In het allerbeste geval is de gerapporteerde 85% de kans die de rekenaar voorstaat, maar er is geen enkele reden dat iemand anders dat ook zou vinden. Dat maakt het niet minder dan bizar om dat getal te publiceren, zeker omdat getallen makkelijk een eigen leven kunnen gaan leiden – zoals in de casus van Anne Frank dan ook prompt gebeurde.

MODELLEN

Hiermee zijn de problemen nog lang niet allemaal besproken. Naast het feit dat getallen persoonlijk zijn, kan het ook niet anders zijn dan dat getallen het gevolg zijn van berekeningen in een wiskundig model van de relevante werkelijkheid. Elke kans, zonder enige uitzondering, is het resultaat van een berekening in zo’n model. Deze opmerking is zo belangrijk dat we er even bij stilstaan.

Op het eerste gezicht lijkt de claim misschien onjuist. Immers, bij bijvoorbeeld het gooien van een munt kan de

kans op ‘kop’ toch niet anders dan $1/2$ zijn? Waar is dan het model? Welnu, het model bestaat hieruit dat we aannemen dat elke worp hetzelfde is, en dat we bij gebrek aan informatie over de uitkomst geen reden hebben om de ene uitkomst vaker te verwachten dan de andere – we bespraken dat zojuist al. In de praktijk is echter geen enkele munt zuiver, en bij heel veel herhalingen zul je dat op een gegeven moment gaan ontdekken. Maar het model waarin we doen alsof dat wel zo is, werkt heel goed, en stelt bijvoorbeeld casino’s in staat om te voorspellen hoeveel winst ze zullen gaan maken.

Ook sommige fysische processen zijn goed te modeleren, juist omdat er veel regelmaat in het spel is. Zo is er bijvoorbeeld een buitengewoon succesvol fysisch model waarin de bewegingen van de planeten in ons zonnestelsel worden beschreven. Het grote brein achter dat model, de Engelse wiskundige en fysicus Isaac Newton, begreep echter verdraaid goed dat niet alles zinvol te modelleren valt. Hij schreef hierover het volgende: “I can model the movements of the planets, but not the madness of people.”³ Een wiskundig model voor het gedrag van mensen kan onmogelijk een goede beschrijving geven van zaken als intentie, verantwoordelijkheid, reden om iets te doen, reden om iets te laten, proportionaliteit, et cetera. Een model getuigt van de gedachten en ideeën van de maker ervan, maar is geen vervanging voor de realiteit.⁴ Een model stelt de maker in staat om zijn aannames op een duidelijke manier te communiceren – je legt met een model als het ware je kaarten op tafel. In sommige situaties, zoals bij een casino en bij de fysica van Newton zijn

³ Geciteerd in Adam Kucharski, *The Rules of Contagion: Why Things Spread – And Why They Stop*, Londen: Profile Books 2020.

⁴ Zie bijvoorbeeld Marc Jacobs en Ronald Meester, *Van aardbeving tot zoönose. Over de inzet van modellen voor beleid*, Zutphen: Mazirel Pers 2023.

de modellen zeer realistisch en zullen er weinig verschillen bestaan tussen modelmakers. Bij modellen waarbij menselijk gedrag cruciaal is, is dat anders. Modellen zijn – altijd! – het begin van een gesprek ergens over, en nooit het eindstation.

De neiging om modellen tot realiteit te verheffen heeft maatschappelijk veel schade opgeleverd in bijna alle grote maatschappelijke dossiers. Deze neiging maakt dat wij de werkelijkheid steeds technocratischer gaan opvatten, waarbij we de menselijkheid en de specifiek menselijke aspecten ervan steeds verder uit het oog verliezen. De toeslagenaffaire, bijvoorbeeld, was mede het gevolg van een te groot vertrouwen in modellen.

Zodra in een strafzaak een kans wordt uitgerekend, kan het dus niet anders dan dat daarbij een model wordt gehanteerd voor de realiteit die voor de specifieke rechtszaak van belang is. Zo'n keuze van een model vereist vele, vele keuzes, en geen twee mensen zullen onafhankelijk van elkaar tot dezelfde modellen of uitkomsten komen. Meestal lijken ze in werkelijk niets op elkaar.⁵ Newton zou zo'n model nooit accepteren als objectief model voor de werkelijkheid, en dat is het natuurlijk ook niet.

Bayesiaans redeneren zoals we dat in hoofdstuk 2 hebben uiteengezet is zelf ook een model. Geen model van werkelijke gebeurtenissen maar een model over de manier waarop wij mensen denken en tot conclusies en overtuigingen komen. Bayesiaans redeneren biedt een acceptabel kwalitatief kader om bewijsmiddelen en gegevens in een zaak op de juiste wijze te verwerken. Als je het denkkader van deze benadering

⁵ Zie ook bijvoorbeeld *Topics in Cognitive Science* 12(4) (2020) voor verschillende bijdragen waaruit duidelijk wordt dat onafhankelijke teams dezelfde situatie op totaal verschillende manieren modelleren, en daarmee ook op totaal verschillende kansen uitkomen dat de verdachte daadwerkelijk schuldig is.

volgt, dan ben je transparant over je aannames en vooronderstellingen, ben je gewapend tegen tunnelvisies en kun je correct besluiten of bepaalde aanwijzingen belastend of juist ontlastend zijn. Je kunt dit denkkader echter niet gebruiken om concrete kansen uit te rekenen. De kans dat de verdachte schuldig is, is geen objectieve grootte maar een persoonlijke uitspraak die bovendien doorgaans niet numeriek uit te drukken zal zijn. Een getal uitrekenen zoals naar aanleiding van *Het verraad van Anne Frank* gebeurde, is daarom bijna hallucinant, zeker omdat veel mensen zullen denken dat dit ‘het’ juiste getal is. Dat is niet alleen onjuist maar evident gevaarlijk en ethisch niet te verantwoorden.

Om deze beweringen verder te onderbouwen bespreken we enkele voorbeelden waarin wetenschappers toch hebben geprobeerd om concrete kansen op te voeren in rechtszaken. Uiteindelijk is het toch het beste om aan de hand van voorbeelden te laten zien waar doorgeschoten rekenen toe kan leiden.

KANSEN IN MODELLEN – VOORBEELDEN

Misschien is het goed om te beginnen met een situatie waarin kansen juist wél zinvol kunnen zijn. Voor sommige onderdelen van een eventuele bewijsvoering bestaan immers goede wiskundige modellen, bijvoorbeeld voor DNA op bron-niveau. Als we kunnen inzien waarom getallen daar wel nuttig zijn, dan is het misschien ook beter te begrijpen waarom ze dat elders niet zijn.

DNA op bronniveau

We weten redelijk goed hoe vaak bepaalde DNA-profielen in de bevolking voorkomen. Misschien weten we het minder precies dan de meeste mensen denken; er zijn nog wel wat problemen zoals verschillen in profiel frequenties tussen verschillende etnische groepen en sowieso zijn de schattingen gebaseerd op een steekproef en daarmee tot op zekere hoogte onzeker. Desondanks kunnen we redelijke uitspraken doen over bijvoorbeeld de kans dat een willekeurig gekozen persoon een bepaald profiel heeft. Dat betekent dat we de bewijskracht van een DNA-match op bronniveau goed kunnen inschatten. Het betekent niet dat we onmiddellijk kansuitspraken kunnen doen over de donor van een profiel, want zoals in het vorige hoofdstuk is besproken hebben we daar ook onze voorkennis en de achtergrond voor nodig.

Tegenwoordig zijn DNA-profielen vaak zo goed dat de herkomst ook niet door de verdediging wordt betwist. In plaats daarvan gaat het dan vooral over de vraag naar de manier waarop het profiel op de plaats delict is geraakt. De verdediging zal bijvoorbeeld aanvoeren dat het weliswaar het DNA van de verdachte is, maar dat ze geen idee hebben hoe het daar verzeild is geraakt. Het kan ook zijn dat de verdediging een alternatief scenario inbrengt waarmee de aanwezigheid van het DNA verklaard kan worden. Dergelijke scenario's gaan ook over menselijk gedrag, en daar is veel minder zinvol aan te rekenen dan aan DNA op bronniveau: immers, een berekening vereist een model, en modellen voor menselijk gedrag zijn nu eenmaal niet eens bij benadering realistisch genoeg om in een specifieke zaak zinvol te zijn.

Een tweede probleem bij het beoordelen van een DNA-profiel heeft ermee te maken dat vaak zogenaamde mengprofielen worden aangetroffen, waarin DNA van meerdere

personen zit. Omdat je niet precies kunt zien welke persoon welk deel heeft bijgedragen, geeft dat een extra onzekerheid bij de numerieke berekeningen. We illustreren dat met een casus waarvan we uit de uitspraak citeren:⁶

“Het touw waarmee [slachtoffer] was vastgebonden is bemonsterd (AAOP9780NL#03) en door een deskundige van het Nederlands Forensisch Instituut (NFI) is deze bemonstering onderworpen aan een DNA-onderzoek. De deskundige concludeert dat het DNA mengprofiel AAOP9780NL#03 *meer dan 1 miljard keer waarschijnlijk* is wanneer de bemonstering DNA bevat van [slachtoffer], [verdachte] en één willekeurig onbekende persoon, dan wanneer de bemonstering DNA bevat van [slachtoffer] en twee willekeurige onbekende personen. De rechtbank stelt dan ook vast dat het DNA van [verdachte] in de vorm van een mengprofiel is aangetroffen op het touw. De rechtbank kan op basis van het dossier niet vaststellen waar dit touw vandaan komt en merkt het touw aan als een verplaatsbare sporendrager.”

Dit citaat verdient enige uitleg. Allereerst is daar het getal van 1 miljard. Dat is een extreem groot getal; het NFI rapporteert eigenlijk nooit getallen groter dan dat. Interessanter is het om te kijken op welke hypothesen dat getal slaat. We zagen immers in het vorige hoofdstuk dat bewijskracht van gegevens altijd relatief ten opzichte van twee hypothesen. In dit geval is kennelijk uit de context te verwachten dat het DNA een mengprofiel van drie personen is, en wordt er gekeken of het profiel beter verklaard kan worden met verdachte als een van deze drie personen, of zonder de verdachte. Het antwoord is duidelijk. We hebben geen

⁶ ECLI:NL:RBMNE:2023:6176, nadruk origineel.

inzicht in het DNA-profiel zelf; het is goed mogelijk dat dit profiel niet van minder dan drie personen afkomstig kán zijn.⁷ Dat zou mede kunnen verklaren waarom de hypothesen op deze manier zijn geformuleerd. Ten slotte zien we uit de laatste zin van het citaat dat hoewel de bijdrage van de verdachte aan het DNA-spoor niet of nauwelijks wordt betwist, er wel onzekerheid is over de rol van de sporendrager (hier een touw) die niet zomaar aan het misdrijf gekoppeld kan worden. Ondanks de DNA-match volgde hier dan ook vrijspraak.

Het getal van 1 miljard was gezien de omstandigheden niet eens beslissend, ondanks het feit dat dit getal betekenisvol berekend kon worden. Er zijn echter ook heel andere situaties waarin getallen en kansen expliciet benoemd worden. Soms komen voorstellen voor kansen uit algemene statistieken.

Algemene statistieken

Stel je voor dat de rechter de betekenis van een bepaalde getuigenverklaring moet inschatten. Kan hij de getuige geloven of niet? Natuurlijk zijn er (al dan niet betrouwbare) statistische gegevens te vinden over de algemene betrouwbaarheid van getuigen in strafzaken. Stel (deze getallen verzinnen we zelf) dat getuigen in 70% van de keren correct vaststellen dat ze de verdachte op de plaats delict hebben gezien, maar in de gevallen dat dit niet zo was, toch nog in 35% van de gevallen zeggen dat het wél de verdachte was. Stel verder dat we nu worden geconfronteerd met een getuige die zegt dat hij de verdachte heeft zien weglopen van de plaats delict.

⁷ Dat is bijvoorbeeld het geval wanneer op een bepaalde locus meer dan vier verschillende allelen gevonden worden.

Volgens sommigen kunnen bovengenoemde percentages gebruikt worden in deze specifieke casus. Maar dat lijkt ons evident onredelijk. Immers, de rechters moeten de betrouwbaarheid van déze specifieke getuige inschatten. Daarbij spelen aspecten mee die onmogelijk zomaar in getallen zijn uit te drukken: lichaamstaal bijvoorbeeld, of de manier waarop de getuige zijn of haar verhaal vertelt. Ook andere kennis over de getuige kan een rol spelen. Hoe het ook zij: er staat hier geen gemiddelde getuige zijn verhaal te doen, maar een specifieke. Er bestaan geen gemiddelde mensen en dus zijn algemeenheden op dat moment niet meer van belang.

Het getal van 35% is echter ook niet helemaal nutteloos. Het kan niet zomaar gebruikt worden in een specifieke zaak, maar het is wel goed als de rechter zich van dit getal bewust is. De rechter weet dan dat getuigen zich regelmatig vergissen en dat zal hem er hopelijk toe bewegen om zorgvuldig na te gaan hoe het met de betrouwbaarheid van deze getuige staat. Hoe hij dat doet, is in elke zaak natuurlijk anders.

Over de vraag in hoeverre algemene statistieken gebruikt mogen worden in de bewijsvoering, woedt in de literatuur nog steeds een stevige discussie.⁸ We willen er hier daarom graag wat meer over zeggen. Hierbij valt het op dat er weliswaar geen algemene overeenstemming is, maar dat geen enkele serieuze wetenschapper het gebruik van dergelijke statistieken zonder omhaal accepteert of niet-problematisch vindt.

⁸ Bijvoorbeeld in D.T. Wasserman, “The Morality of Statistical Proof and the Risk of Mistaken Liability”, *Cardozo Law Review* 13 (1991): 935-976; Frederick F. Schauer, *Profiles, Probabilities, and Stereotypes*, Cambridge, MA: Harvard University Press 2003; Amit Pundik, “Rethinking the Use of Statistical Evidence to Prove Causation in Criminal Cases: A Tale of (im)Probability and Free Will”, *Law and Philosophy* 40(2) (2021): 97-128.

Allereerst zijn er verschillende soorten statistische gegevens, en het is goed enkele voorbeelden te beschouwen. Wanneer het bijvoorbeeld gaat om de betrouwbaarheid van een getuige zoals zojuist, blijft het een feit dat er een getuige was die iets heeft gezien in de onderhavige zaak. Ook zonder het gebruik van statistieken blijft dat laatste feit overeind en relevant voor de zaak. Er is, kortom, nog steeds bewijs in strafrechtelijke zin.

Een andere statistiek is de algemene statistiek dat meer dan de helft van omgebrachte, getrouwde vrouwen in Nederland door hun echtgenoot om het leven is gebracht.⁹ Is deze algemene statistiek bruikbaar in een specifieke zaak? Merk op dat deze statistiek net even anders is dan die in het vorige voorbeeld. Immers, als je deze laatste statistiek weglaat, dan blijft er niets over (als bewijs) in een concrete zaak waarin een vrouw om het leven is gebracht (behalve natuurlijk het feit dat iemand overleden is), in tegenstelling tot de zaak van de getuige van zojuist, waar in ieder geval de getuigenis nog blijft bestaan, ook zonder statistieken. De echtgenoot van een omgebrachte vrouw heeft echter in het geheel geen boodschap aan wat andere mannen hebben gedaan. Waarom zou het gedrag van anderen op hun rechtszaak van toepassing zijn? Daar is simpelweg geen reden toe.

Ook vanuit de filosofie wordt wel bezwaar gemaakt tegen het gebruik van algemene statistieken. Zo is bijvoorbeeld gesteld dat dergelijk gebruik het unieke, individuele karakter van elke persoon niet respecteert. In een recent artikel van filosoof Amit Pundik¹⁰ wordt in het verlengde hiervan zelfs

⁹ Paul Nieuwbeerta en Gerlof Leistra, *Dodelijk geweld. Moord en doodslag in Nederland*, Amsterdam: Balans 2007.

¹⁰ Pundik, *Rethinking the Use of Statistical Evidence to Prove Causation in Criminal Cases: A tale of (im)Probability and Free Will*, Law & Philosophy 40, p. 97-128.

een link gelegd met het bestaan van de vrije wil. Het gebruik van statistieken op bovengenoemde manier is volgens Pundik niet verenigbaar met de aanname van vrije wil. Hij concludeert: “It follows that using statistical evidence on the low probability of an alternative cause is objectionable, regardless of how reliable the statistical analysis is.” Het gebruik van algemene statistieken komt neer op speculatie, en op het invullen van gegevens die er simpelweg niet zijn. Ze kunnen niet worden ingezet als wettig bewijsmiddel.

Er zijn overigens situaties te bedenken – buiten het strafrecht – waarin het gebruik van dergelijke statistieken niet problematisch is. Het is goed om daar ook wat over te zeggen, omdat het onze argumenten verheldert. Als iemand zijn of haar vertrouwen in een bepaalde gebeurtenis wil kwantificeren met het oog op, bijvoorbeeld, een weddenschap, dan zien we geen reden waarom je algemene statistieken niet zou kunnen gebruiken. Als iemand gedwongen wordt om te kiezen tussen twee alternatieven en er zijn weinig gegevens voorhanden, dan zou het zeker te verdedigen zijn om de beslissing dan maar op basis van statistieken te nemen. Als iemand moet kiezen tussen twee medicijnen, dan kiezen de meeste mensen bij gebrek aan verdere informatie liever voor het medicijn met de beste statistieken. Als iemand echter meer zou weten van het medicijn, bijvoorbeeld als iemand meer informatie heeft over de omstandigheden of bijzonderheden waaronder het medicijn succesvol is, dan zou die persoon ervoor kunnen kiezen om een besluit te nemen dat tegen de algemene statistieken ingaat, omdat hij meer informatie heeft dan waarop die statistieken zijn gebaseerd. Bij medicijnen is de context bovendien heel anders. Geen twee mensen zijn hetzelfde, maar je kunt wellicht aanvoeren dat de chemie van mensen op een bepaalde manier minder van elkaar verschilt dan de contexten bij verschillende rechtszaken. Algemeenhe-

den zijn daar dus relevanter dan bij rechtszaken die allemaal uniek zijn.

Het nemen van een beslissing omdat je moet kiezen, is heel wat anders dan overwegen of er voor een bepaalde, unieke, gebeurtenis bewijs is of niet, zoals de rechter in de context van een strafzaak moet doen. In het strafrecht gaat het om uniek bewijs, en daar hoort speculatie simpelweg niet bij. Als er geen gegevens zijn, dan kun je die niet verzinnen op basis van algemeenheden, maar moet je simpelweg accepteren dat het bewijs er niet is. En dat betekent dus ook dat, wil een Bayesiaanse analyse van een (moeilijke) strafzaak zinvol kunnen zijn voor de overtuiging van de rechter, deze analyse gebruik moet maken van het aanwezige bewijs, en niet van algemene gegevens die niet expliciet bij deze zaak horen. Deze op zich relevante gegevens ‘verplaatsen’ we dan liever naar de context waarbinnen de rechter het bewijs evalueert en waardeert.

Ad-hockansen

Tot zover de discussie van algemene statistieken. Er zijn echter ook andere manieren waarop kansen soms ingeschat worden; een kans kan uit een specifieke ad-hocaanname volgen. We geven hiervan een praktijkvoorbeeld.¹¹

Geconfronteerd met een verdachte van een zedenmisdrijf werd nagegaan hoe vaak deze verdachte een dergelijk misdrijf had begaan in het verleden. Het bleek dat de verdachte vijf jaar geleden voor het laatst voor een dergelijk vergrijp was veroordeeld. Men stelde nu dat de kans dat de verdachte

¹¹ Dit voorbeeld komt uit een zaak waarbij een van ons (RM) als contraexpert optrad. De methode is zoals deze in die zaak werd gebruikt, maar de getallen zijn fictief.

in een bepaald gegeven jaar een zedendelict zou begaan, gelijk aan 20%. Deze 20% zou dan gezien moeten worden als context, op de manier zoals we dat in het vorige hoofdstuk hebben beschreven.

Slaat het getal eigenlijk wel ergens op? Laten we dat eens onderzoeken. De gedachte erachter was (denken wij) dat de periode van vijf jaar werd aangegrepen om te kijken hoe vaak de verdachte zedenmisdrijven zou begaan: op basis van de afgelopen vijf jaar werd gesteld dat de verdachte zoiets kennelijk één keer in de vijf jaar doet.

Het behoeft denk ik geen betoog dat een dergelijke reductie van een zeer weerbarstige praktijk tot een enkel getal weinig of geen betekenis heeft. Onze wiskundige en statistische collega's wilden aanvankelijk niet geloven dat deze model-aanname en dit percentage een rol kan hebben gespeeld in een zaak, zo volledig buiten elke realiteit als de aanname is. Men doet net alsof het plegen van zedenmisdrijven te vergelijken is met het gooien van een munt. Voor muntworpen, die in zeer vergelijkbare omstandigheden plaatsvinden, is het model waarin de kans elke keer gelijk is heel goed. Maar voor menselijk gedrag dat al dan niet tot zedenmisdrijven leidt, geldt dat natuurlijk niet. Bovendien vergeet men in dit geval dat een steekproef van vijf jaar veel te klein is om met enige nauwkeurigheid de kans te bepalen, zelfs al geloof je in het model (wat wij niet doen). Sterker nog, dat geldt ook voor de muntworpen: op basis van vijf of zes worpen kan de kans op succes helemaal niet met enige nauwkeurigheid vastgesteld worden, zoals elke inleiding in de statistiek laat zien.

Toch werd bovenstaande aanname gewoon gebruikt. Het is een klassiek voorbeeld waarin de noodzaak om de context mee te nemen (het zedenverleden van de verdachte) vertaald wordt in een onzinnig getal. De interpretatie van het feit dat de verdachte vijf jaar geleden voor het laatst veroordeeld was

voor een zedenmisdrijf hangt nogal van de omstandigheden af. Was hij daarvoor actiever op dit gebied? Heeft hij therapie of misschien wel TBS gehad? Wat zegt hij er zelf eigenlijk over? Is het bijvoorbeeld geloofwaardig dat hij zelf aangeeft de problemen te boven zijn gekomen, en op basis waarvan is dat al dan niet het geval? De lijst met mogelijke en zeer relevante vragen is eindeloos. Die lijst kun je niet platslaan met een getal.

Misschien wel het beroemdste voorbeeld van een getal uit de Nederlandse jurisprudentie komt uit de zaak Lucia de Berk. In deze zaak was berekend dat de kans dat de verpleegster Lucia de Berk, die naar het zich liet aanzien vaker te maken had met incidenten en gestorven patiëntjes, in geval van onschuld zoveel incidenten zou meemaken 1 op de 342 miljoen was. Nadat dat getal werd bekritiseerd op basis van het feit dat het voortkwam uit een zinloos model, vroeg het hof naar het ‘juiste getal’. Een betere illustratie voor de problemen is moeilijk te geven: er is niet een juist getal, zo kan het niet werken.

NUMERIEKE WAARDEN VOOR DE PRIOR?

Hoewel we inmiddels duidelijk hebben gemaakt dat Bayesiaans redeneren niet betekent dat er gerekend zou moeten worden in rechtszaken, willen we op deze plek stilstaan bij pogingen van diverse auteurs om de oorspronkelijke inschatting, de prior, wel numeriek te bepalen. We doen dit enerzijds vanuit historisch perspectief maar anderzijds vooral ook ter illustratie van onze argumenten. Het is namelijk meer dan een veeg teken dat we de suggesties die we hier de revue zullen laten passeren niet kunnen illustreren met concrete voorbeelden in rechtszaken. Dat kan bijna niets anders betekenen

dan dat numerieke priors vooral gepropageerd worden door onderzoekers die niet zelf in de rechtszaal staan: het lijkt een tamelijk academische aangelegenheid. Dat kan voor juristen een reden zijn om Bayesiaans redeneren niet te gebruiken, we schreven dat al eerder, en dat willen we graag voorkomen. We kunnen er bovendien toch ook nog wel iets van leren, zoals we zullen zien.

In de literatuur zijn verschillende suggesties verschenen over hoe een prior te berekenen. In de eerste benadering wordt de verdachte behandeld als willekeurig gekozen uit een groep van mogelijke daders. Als deze groep de grootte N heeft, dan wordt de a priori kans dat de verdachte de dader is vastgesteld op $1/N$. Onder de vele problemen van deze benadering noemen we de voor de hand liggende vraag waaruit die groep van mogelijke daders eigenlijk bestaat. Als een misdaad in Amsterdam is gepleegd, hoe zou je dan N kiezen? De bevolking van Amsterdam? De bevolking van Nederland? Er lijkt geen canoniek antwoord op deze vraag te zijn, en dat maakt deze benadering problematisch. Bovendien is het natuurlijk ook niet zo dat iedereen binnen een bepaalde bevolkingsgroep dezelfde relatie tot het misdrijf heeft: sommige mensen wonen of zijn vlakbij, andere niet. Deze benadering kunnen we daarom niet erg serieus nemen.

In de tweede benadering wordt onzekerheid benaderd door de zaak te behandelen als willekeurig gekozen uit situaties van dezelfde 'soort'. Dus, bijvoorbeeld, als de vrouw van een bepaalde man vermoord wordt aangetroffen, dan wordt de prior kans dat de man de moordenaar is, vastgesteld op 52%, zijnde het aandeel van bekende gevallen waarin de man inderdaad de moordenaar was van een vermoorde, getrouwde vrouw, in Nederland.¹²

¹² Nieuwbeerta en Leistra, *Dodelijk geweld*.

We hebben in het vorige hoofdstuk al uitgelegd dat deze benadering problematisch is, omdat ze uitgaat van algemene statistieken, die met de specifieke zaak in kwestie niet veel te maken hebben. We zullen de argumenten die we daar gaven hier niet herhalen en slechts de conclusie vermelden dat we dit soort algemene informatie niet zo maar cijfermatig in een bepaalde casus kunnen gebruiken.

Een derde benadering om numerieke betekenis te geven aan een prior komt neer op de afspraak dat men de prior kans van de hypothese van de aanklager op een vast getal stelt, bijvoorbeeld 1% of iets dergelijks. Zo schrijven Dahlman en Koflaath:¹³

“Een andere benadering van het probleem van de prior in strafzaken is om een juridische regel te hebben die de feitenonderzoeker opdraagt om aan te nemen dat de hypothese van de aanklager een bepaalde waarschijnlijkheid (p_0) heeft voordat het bewijs tijdens het proces is gepresenteerd. Als voorbeeld zou zo’n regel een prior kans van 1% ($p_0 = 0,01$) kunnen postuleren en zeggen dat, in elke strafzaak, de feitenonderzoeker de evaluatie van het bewijs moet beginnen onder de aanname dat de waarschijnlijkheid van de hypothese van de aanklager 1% is voordat het bewijs tijdens het proces is gepresenteerd. We zullen deze oplossing een veronderstelde prior noemen.”

Volgens Dahlman en Koflaath is een belangrijk voordeel van de veronderstelde prior “dat het niet de willekeur heeft die andere manieren om de prior vast te stellen teisteren, noch in principe noch in de praktijk”. Uiteraard is dit waar,

¹³ C. Dahlman en E. Koflaath, “The Problem of the Prior in Criminal Trials”, in C. Dahlman, G. Tuzet en A. Stein (red.), *Philosophical Foundations of Evidence Law*, Oxford: Oxford University Press 2021 (onze vertaling).

maar men moet zich afvragen of dit echt is wat nodig is. Het zou impliceren dat de veronderstelde prior in zekere zin universeel is. De werkelijke numerieke waarde van de veronderstelde prior is, volgens Dahlman en Koflaath open voor discussie, en zij citeren waarden als 0,01, 0,001 en 0,0001. Tegen het argument dat elke keuze willekeurig zou zijn, stellen zij dat een dergelijk argument onjuist is “[...] aangezien de veronderstelde prior niet gericht is op een feitelijke kwestie [...]. Wanneer de wetgever een numerieke waarde kiest voor de veronderstelde prior, kiest de wetgever hoe zwaar de bewijslast is die de wet op de aanklager legt.” Bovendien stellen zij dat “de keuze van de wet voor een veronderstelde prior niet meer willekeurig is dan de keuze van de wet voor een numerieke waarde voor de drempelwaarschijnlijkheid die vereist is voor veroordeling”.

Vooraf dit laatste argument is voor ons interessant, omdat eruit blijkt dat de auteurs Bayesiaans redeneren als een berekening zien. Met een zogenaamde veronderstelde prior en een regel die een drempel stelt voor de vereiste uiteindelijke of ‘a posteriori’ kans voor veroordeling, biedt dit volgens hen “een normatief kader voor bewijsbeoordeling”.

Het is vanuit een bepaald perspectief ook wel aantrekkelijk om altijd dezelfde prior te nemen, maar het probleem is dat je daarmee de context geheel negeert. We hebben echter al gezien dat dat niet kan: de context bepaalt in hoeverre het aanwezige bewijs daadwerkelijk tot de overtuiging leidt dat een verdachte al dan niet schuldig is.

Wat kunnen we hier nu verder van leren? Welnu, de gedachte dat je altijd dezelfde prior zou moeten gebruiken, heeft bepaalde consequenties die we nu zullen bespreken. Wij vinden deze consequenties onaanvaardbaar, waarmee we impliciet ook zeggen dat de benadering van de ‘veronderstelde prior’ onaanvaardbaar is voor ons.

Stel dat bij het onderzoek op een plaats delict een onvolledig DNA-profiel wordt veiliggesteld, en stel dat er een bepaalde verdachte is in deze zaak. Het NFI onderzoekt het DNA-profiel en komt tot de conclusie dat dit gedeeltelijke profiel 20.000 keer beter past bij de verdachte dan bij een willekeurig, niet verwant persoon: het NFI zegt daarmee dat de kans op het aantreffen van dit profiel 20.000 keer zo waarschijnlijk is wanneer de verdachte de donor is, dan wanneer een willekeurig, ongerelateerd persoon de donor is.

Stel nu dat de veronderstelde prior zo hoog is dat deze prior samen met de bewijskracht van 20.000 zorgt voor een uiteindelijke overtuiging dat verdachte schuldig is. Als dit alle gegevens in de zaak zouden zijn, dan zou de verdachte dus veroordeeld moeten worden. Maar dan vergeet men dat er naar schatting zo'n 900 mensen in Nederland zijn die evengoed bij het profiel passen als de verdachte. Immers, 1 op de 20.000 mensen heeft dat profiel, en zijn ongeveer 18 miljoen mensen in Nederland. Moeten we nu al die 900 mensen ook gaan veroordelen?

De overgrote meerderheid binnen die groep van 900 mensen zal een extreem kleine a priori waarschijnlijkheid hebben om de donor van het DNA-profiel te zijn, en dit feit kun je simpelweg niet negeren: het moet op de een of andere manier in aanmerking worden genomen. Met een veronderstelde a priori kan er geen onderscheid worden gemaakt tussen alle overeenkomende personen, iets wat duidelijk onaanvaardbaar lijkt.

Maar er is nog een tweede, zo mogelijk nog overtuigender argument, en we moeten hiervoor even rustig de tijd nemen. Om ons punt te maken gaan we terug naar de zaak Sally Clark die we in hoofdstuk 2 al bespraken. Als service voor de lezer herhalen we de relevante details van de zaak hier, en geven iets meer numerieke informatie dan in hoofdstuk 2.

In december 1996 werd Sally Clarks eerstgeboren zoon Christopher dood in zijn bed gevonden, op de leeftijd van ruim 2 maanden. Hij stierf toen Sally alleen in huis was met hem. In januari 1998 stierf haar tweede kind Harry, ook 2 maanden oud, onder soortgelijke omstandigheden. Kort daarna werd Sally beschuldigd van het vermoorden van haar twee kinderen door verstikking. Naast de mogelijkheid van dubbele moord werd ook de mogelijkheid van een dubbele wiegendood overwogen.

Het belangrijkste bewijs van de aanklager kwam van een kinderarts, Sir Roy Meadow, die getuigde dat de verwachte incidentie van een kind dat sterft aan wiegendood ongeveer 1 op 8500 was. Hij getuigde ook dat een tweede optreden van wiegendood, gegeven een eerste, even waarschijnlijk zou zijn als een eerste, resulterend in een algehele kans op dubbele wiegendood van ongeveer 1 op 72 miljoen.

We kunnen hier verschillende sets van hypotheses formuleren. Bijvoorbeeld, men kan als eerste hypothese nemen dat Sally Clark haar kinderen heeft vermoord, en als tweede hypothese dat ze beide zijn gestorven aan wiegendood:

H_1 : Sally Clark heeft haar twee kinderen vermoord.

H_2 : De twee kinderen stierven zonder aanwijsbare oorzaak (en dit werd als wiegendood bestempeld).

Dit is wat we in hoofdstuk 2 deden. We merkten daar al op dat de gegevens (twee gestorven kinderen) onder beide hypotheses even waarschijnlijk waren: het is onder beide hypotheses zeker dat we twee gestorven kinderen aantreffen, en dit leidt dus tot neutraal bewijs met een likelihood ratio van 1. Het bewijs verandert iemands mening over de verhouding tussen deze twee hypotheses dus niet, hoe vreemd dat ook in eerste instantie moge klinken.

De precieze bewoording van de hypotheses kan echter net even anders zijn. Je kunt bij dezelfde eerste hypothese (Sally Clark heeft haar twee kinderen vermoord) als tweede hypothese nemen dat ze haar kinderen níét heeft vermoord:

H_1 : Sally Clark heeft haar twee kinderen vermoord.

H_3 : Sally Clark heeft haar twee kinderen niet vermoord.

Dat lijkt een onbetekenende verandering ten opzichte van de eerdere keuze, en natuurlijk zou het ook niet uit moeten maken voor de uiteindelijke conclusie. Maar kijk wat er gebeurt: de kans om twee gestorven kinderen aan te treffen als Sally Clark ze heeft vermoord, is natuurlijk 100%, net als hiervoor. Maar de kans om twee dode kinderen aan te treffen als Sally Clark ze niet heeft vermoord, moet wel heel klein zijn: kinderen gaan meestal niet dood. Dat betekent dat de likelihood ratio voor het tweede paar hypotheses erg groot is: de data (twee gestorven kinderen) passen nu eenmaal veel beter bij het moordscenario H_1 dan bij het scenario H_3 waarin Sally Clark geen moordenaar is.

Wat is hier nu precies aan de hand? In het eerste geval is het bewijs neutraal, maar in het tweede geval is het bewijs helemaal niet neutraal en wijst het sterk naar de eerste hypothese van dubbele moord. Bedenk dat er geen reden lijkt te zijn om juist de eerste of de tweede set van hypotheses te kiezen. Maar, en dat is de cruciale observatie, de aanvankelijke a priori kans op H_3 is veel groter dan die op H_2 . Dus de oorspronkelijke inschattingen van de hypotheses zijn totaal verschillend. Het feit dat in het tweede geval de likelihood ratio (veel) groter was dan in het eerste geval, wordt gecompenseerd door het feit dat de oorspronkelijke inschatting op H_3 logischerwijs groter moet zijn dan die op H_2 .

Het is op dit soort momenten dat de wiskundige formulering van Bayesiaans redeneren behulpzaam is. Je kunt namelijk laten zien dat het voor de uiteindelijke overtuiging niet uitmaakt of je voor H_2 of voor H_3 kiest als tweede hypothese. Het feit dat het bewijs voor H_1 ten opzichte van H_3 een stuk sterker is, wordt exact gecompenseerd door het feit dat de oorspronkelijke inschatting veel meer in de richting van H_3 dan H_2 wijst – logischerwijs.

Het is daarom onaanvaardbaar dat we in beide gevallen dezelfde prior zouden kiezen, want de logica van de situatie verbiedt dat feitelijk. We hebben een bepaalde vrijheid nodig om informatie als het ware heen en weer te kunnen schuiven tussen concreet bewijs en context, waarbij we nogmaals benadrukken dat beide in juridische zin gebruikt kunnen worden als bewijs in een zaak. Filosoof Richard Swinburne verwoordt het als volgt: “Elke verdeling van bewijs tussen nieuw observatiebewijs en achtergrondbewijs is willekeurig.”¹⁴

Deze uitspraak vat de problemen met een ‘veronderstelde prior’ redelijk goed samen. Het is een onmogelijke procedure omdat er in dezelfde zaak verschillende keuzes te maken zijn over de hypothesen die je wilt beschouwen. Er is soms (zoals in het geval dat we zojuist bespraken) geen duidelijke voorkeur aan te geven, en je hebt dan de vrijheid nodig om de prior te kunnen veranderen.

We concluderen dat een numerieke prior onoverkomelijke problemen met zich meebrengt, en dat is vanuit onze filosofische en praktische positie geen verrassing te noemen.¹⁵ Redeneren met Bayes betekent helemaal niet dat je moet rekenen. Het is zoals we al schreven een veeg teken dat numerieke

¹⁴ Richard Swinburne, *Bayes's Theorem*, British Academy 2002.

¹⁵ Zie ook hoofdstuk 4 voor een discussie over de prior in relatie tot de onschuldpresumptie.

priors eigenlijk alleen maar in academische kringen worden bestudeerd en gesuggereerd. Het zou wat ons betreft jammer zijn als juristen Bayes niet zouden gebruiken vanwege het feit dat ze denken dat het alleen met rekenen te maken heeft.

COMBINEREN VAN BEWIJS

Een van de argumenten die wel eens aangevoerd worden om toch maar te gaan rekenen, is het feit dat je op die manier de bewijswaarde van verschillende gegevens met elkaar kunt vergelijken en ook kunt combineren. Het komt natuurlijk vaak voor dat bewijsmiddelen elkaar tegenspreken: sommige lijken belastend, en andere lijken juist weer ontlastend. Het zou natuurlijk handig zijn als je die twee bewijsmiddelen kunt combineren tot één getal waaruit dan de gecombineerde waarde van het bewijs berekend zou kunnen worden.

Nu is het zeker zo dat Bayesiaans redeneren impliciet laat zien dat vele kleine aanwijzingen bij elkaar toch heel overtuigend kunnen zijn. Als er tien aanwijzingen zijn die allemaal een beetje belastend zijn, dan kan het ‘hele plaatje’ er zeer belastend uitzien. De bekende ‘salamitactiek’ waarbij de verdediging elk middel ontkracht door te stellen dat de bewijswaarde ervan gering is, is dan ook principieel onjuist.

Dit fenomeen kan uitstekend geïllustreerd worden door in een eenvoudig geval te rekenen. Een DNA-profiel bestaat uit een aantal allelen. Als je naar een DNA-profiel kijkt dan is een overeenkomst van een enkel allel tussen het spoor op de plaats delict en het DNA van een verdachte weliswaar belastend, maar zeker niet beslissend. Dat geldt voor elk individueel allel, maar het spoor als geheel is wel zeer belastend. In Bayesiaanse termen is dat goed te begrijpen. De context waarbinnen elk volgend allel wordt bekeken wordt simpel-

weg steeds belastender. Na vijf overeenkomsten is de situatie voor de verdachte anders dan aan het begin van de evaluatie. Het bouwt als het ware steeds meer op, en dat vele kleintjes één grote maken is hier meer dan waar.

Je kunt dit ook wiskundig goed begrijpen. Als de bewijsmiddelen voldoende onafhankelijk zijn,¹⁶ dan mag je de bewijskracht van de verschillende onderdelen met elkaar vermenigvuldigen. We benadrukken dat dit alleen maar betekenis heeft binnen het model waarin je dan rekent – hier komen we na het volgende voorbeeld nog op terug.

Hier is een voorbeeld waarin deze salamitactiek door de rechter zelf werd bedreven.¹⁷ De zaak draait om een woningoverval waarin verschillende bewijsmiddelen besproken worden: een DNA-spoor op een wapen, een DNA-spoor op een handschoen die notabene in de slaapkamer van de moeder van de verdachte werd aangetroffen (!), en een schoenspoor ter plaatse. Dit is wat we in de uitspraak lezen:

“Het eerste is een DNA-spoor van verdachte dat op de slede van een wapen die in de slaapkamer van aangevers [benadeelde partijen] is aangetroffen. De aanwezigheid van dat spoor is weliswaar belastend voor verdachte, maar is nog geen direct bewijs voor verdachtes aanwezigheid tijdens de overval.

[...]

Het hof kan niet in redelijkheid uitsluiten dat een ander dan verdachte de sporen op de handschoen heeft achtergelaten en ook niet dat een ander de handschoen in die

¹⁶ Wat “voldoende onafhankelijk” precies betekent is helemaal niet zo makkelijk uit te leggen, en we gaan hier niet in detail op dit punt in.

¹⁷ ECLI:NL:GHARL:2015:5572.

slaapkamer heeft achtergelaten. De aanwezigheid van de DNA-nevenkenmerken die overeenkomen met het profiel van verdachte zeggen niks over de wijze waarop die nevenkenmerken daar zijn gekomen en voorts blijkt uit voorgaande dat die nevenkenmerken daar ook door anderen kunnen zijn achtergelaten. Bovendien zijn er ook nog nevenkenmerken van twee andere personen aangetroffen. Het hof acht dan ook de aangetroffen DNA-nevenkenmerken die overeenkomen met die van verdachte in het licht van het voorgaande zodanig zwak dat het niet kan worden aangemerkt als bewijs van verdachtes daderschap.

[...]

Het derde belastende bewijs wordt gevormd door schoensporen. [...] Ook de aanwezigheid van dat spoor is weliswaar belastend voor verdachte, maar is nog geen direct bewijs voor verdachtes aanwezigheid tijdens de overval.”

We zien dat bij elk van de drie bewijsmiddelen gezegd wordt dat het specifieke bewijsmiddel in zichzelf niet sterk genoeg is om tot een veroordeling te komen. Afgezien van de wel zeer bijzondere vindplaats van de handschoen is dat begrijpelijk. Maar de logica waarbij het hof elk bewijsmiddel apart beziet, en niet in Bayesiaanse samenhang, is incorrect. Na elk belastend bewijsmiddel is immers, als het goed is, de mening van het hof veranderd, en moet het volgende bewijsmiddel begrepen worden vanuit die veranderde positie. Op die manier is het werkelijk zo dat vele kleintjes één grote maken. Dat volgt niet alleen onverbiddelijk uit de wiskunde die deze procedure beschrijft, maar is ook op intuïtief niveau onontkoombaar.

We zeggen tot slot nog iets meer over het combineren van bewijsmiddelen, want hierover bestaan enkele hardnekkige

misverstanden. We beginnen daartoe met een uitgebreid citaat uit een artikel van Alkemade en Stikkelbroeck:¹⁸

“Zo kunnen onafhankelijke bewijsmiddelen in een strafzaak dus op een correcte manier gecombineerd worden: Aan elke bevinding wordt een likelihood ratio toegekend, een getal dat de *bewijskracht* uitdrukt, en door die getallen met elkaar te vermenigvuldigen verkrijgt men de gecombineerde bewijskracht.

[...]

Het feit dat vermenigvuldigen zoveel ‘sneller’ tot zeer grote of zeer kleine getallen kan leiden dan optellen of aftrekken, is een van de redenen dat de uitkomst soms contra-intuïtief kan overkomen.

[...]

Dat komt doordat de bewijsmiddelen nog te makkelijk intuïtief en vanuit ‘gezond verstand’ optellend worden gewogen, in plaats van dat expliciet zowel de aparte likelihood ratio’s als ook de gecombineerde bewijskracht vermenigvuldigend aan bod komen.”

We geven dit citaat omdat hieruit blijkt hoe ver de modelwereld van Alkemade en Stikkelbroeck (hierna: AS) staat van de realiteit. Het vermeende verschil tussen vermenigvuldigen en optellen is een modelconstruct, en heeft in

¹⁸ F. Alkemade en H. Stikkelbroeck, “In onderling verband en samenhang bezien (een praktische uitleg van de Bayesiaanse aanpak in het strafrecht). Deel 1: Het vaststellen van de bewijskracht van bevindingen en de regel van Bayes”, *Trema* (november 2015): 285-297.

tegenstelling tot wat AS schrijven geen enkele fundamentele betekenis. Dat gaan we nu uitleggen.

Allereerst stellen we vast dat de vermenigvuldiging van bewijskracht zoals beschreven door AS in het wiskundige model dat zij hanteren, correct is. Bij voldoende onafhankelijkheid is dat wat het wiskundig model zegt dat je moet doen. Maar er is wel iets heel vreemds aan de hand, ook als we binnen het model blijven zoals AS dat hanteren. Er wordt in het citaat namelijk nogal nadruk gelegd op precies dat vermenigvuldigen. Het wordt vergeleken met optellen en aftrekken in de zin dat vermenigvuldigen sneller tot heel kleine of heel grote getallen kan leiden dan optellen en aftrekken, en dat zou contra-intuïtief zijn. Is dat zo? Nee.

We hebben gezien dat correct redeneren vereist dat we kijken of bepaalde gegevens beter bij de ene dan bij de andere hypothese passen. Een manier om dat uit te drukken is om te kijken naar de kans dat we deze gegevens zouden vinden als de ene hypothese waar is, en dat vergelijken met de kans onder de rivaliserende hypothese. We schreven echter al uitgebreid over de subjectiviteit van het begrip kans: dat is niet een gegeven getal maar afhankelijk van veel persoonlijke factoren. In plaats van naar de kans te kijken, zou je ook kunnen kijken naar hoe verrast je bent om bepaalde gegevens aan te treffen onder de diverse hypothesen. Hoe verrassender, hoe minder de gegevens erbij passen.

Hoe dan ook, het is dus geen wet van Meden en Perzen dat we naar de kans op de gegevens kijken. In de literatuur wordt veelvuldig gewezen naar een andere grootheid: niet de likelihood ratio zelf maar de *logaritme* van de likelihood ratio. Voor de niet zo wiskundig aangelegde lezers moeten we daar wat over zeggen.

Als we iets in getallen uitdrukken, en twee metingen of berekeningen komen uit op bijvoorbeeld 100 en 1000, dan is

het duidelijk dat 1000 groter is dan 100. Het is zelfs zo dat we kunnen zeggen hoeveel keer groter 1000 is dan 100: er zit precies een factor 10 tussen. Maar je kunt de getallen 100 en 1000 ook schrijven als $100 = 10^2$ en $1000 = 10^3$. Het feit dat 1000 groter is dan 100 wordt weerspiegeld door het feit dat de exponent 3 groter is dan de exponent 2. Maar natuurlijk is het niet zo dat de exponent 3 tien maal zo groot is als de exponent 2. Als we in plaats van de getallen 100 en 1000 de bijbehorende exponenten hadden gemeten, dan zou je geen factor 10 hebben gemeten, maar een factor 1,5. Belangrijker is echter dat als ik de getallen 100 en 1000 met elkaar vermenigvuldig, de exponenten 2 en 3 bij elkaar opgeteld moeten worden. Immers $100 \times 1000 = 100.000$ en dat kan ik schrijven als $10^{2+3} = 10^5$. Als we de uitkomsten niet zelf meten maar juist de exponenten, dan werken we op een zogenaamde logaritmische schaal. Op die schaal verandert vermenigvuldigen dus in optellen.

Veel metingen en grootheden die we meten, kennen een logaritmische schaal: geluidsmetingen in decibellen, kracht van aardbevingen op de (energie)schaal van Richter, de kracht van stormen uitgedrukt in windkracht, en nog veel meer. Deze schalen zijn niet lineair: geluid van 40 decibel is niet twee maal zo hard als geluid van 20 decibel. De schaal is zo ontworpen dat elke extra 3 decibel een verdubbeling (!) van het geluidsniveau weergeeft. Zo is een aardbeving met kracht 8 op de schaal van Richter veel zwaarder dan het dubbele van een aardbeving met kracht 4. De reden dat men dit soort logaritmische schalen gebruikt, heeft te maken met het feit dat veel natuurlijke fenomenen zich niet lineair gedragen, en dan is het beter om ze ook niet lineair te meten.

Een ander voorbeeld is dat een extra inkomen van 1000 euro voor een miljonair niet dezelfde betekenis heeft als voor iemand in de bijstand. Als je de reële betekenis daar-

van wilt weten, dan kom je ook op een logaritmische en dus niet-lineaire schaal uit. Immers, voor iemand die op of onder het bestaansminimum leeft is 1000 euro echt heel veel geld. Een miljonair echter zal er zijn schouders over ophalen want 1000 euro betekent voor hem niets.

Ook in de forensische wereld is het gebruik van een logaritmische schaal uitgebreid besproken.¹⁹ Het blijkt zo te zijn dat de logaritme van de likelihood ratio op het bewijsmateriaal rekenkundige eigenschappen heeft die prettig zijn, en die suggereren dat het misschien beter is om daarnaar te kijken.²⁰ Als je dat doet, is een likelihood ratio van 100 niet tien keer zo sterk als een likelihood ratio van 10, maar slechts twee keer zo sterk.²¹ Dit komt precies overeen met het eerste voorbeeld van de 100 en de 1000: immers, we kunnen schrijven $10 = 10^1$ en $100 = 10^2$. De werkelijke betekenis van de likelihood ratio wordt volgens dit gezichtspunt beter beschreven door de exponent.

Als je dus met logaritmen van likelihood ratio's werkt, dan moet je de numerieke uitkomsten niet met elkaar vermenigvuldigen maar, u raadt het al, bij elkaar optellen; kijk maar naar de exponenten in de voorbeelden hierboven. Overigens is deze manier van werken, op de logaritmische schaal dus, heel natuurlijk in onze context. Als ik een bepaald bewijsmiddel twee keer onafhankelijk van elkaar verkrijg,

¹⁹ Zie bijvoorbeeld Ronald Meester en Klaas Slooten, *Probability and Forensic Evidence: Theory, Philosophy, and Applications*, Cambridge: Cambridge University Press 2021.

²⁰ Deze logaritme van de kans heeft bijvoorbeeld de eigenschap dat onder de eerste hypothese, extra bewijs 'gemiddeld' gesproken tot een meer belastende situatie leidt, en onder de tweede hypothese 'gemiddeld' tot een minder belastende situatie – allemaal binnen het formele wiskundige model natuurlijk. Bewijs wijst dus 'gemiddeld' in de juiste richting als je het zo kwantificeert.

²¹ Als je tenminste met grondtal 10 voor de logaritme werkt.

dan is het redelijk te zeggen dat je twee keer zo veel bewijs hebt als wanneer je het maar één keer verkrijgt. Dat is precies wat er gebeurt als je op deze manier rekent.

Wat betekent dit nu voor de claim van AS? Het betekent dat het door AS gesignaleerde verschil tussen optellen en aftrekken, een verschil dat zelfs werd inhoudelijk werd geïnterpreteerd, onzinnig en zonder enige betekenis is. Bayesiaans denken zegt helemaal niet dat je bewijswaarden moet vermenigvuldigen. Je kiest voor een bepaalde maat om dingen te meten, en afhankelijk van de maat die je kiest leidt dat tot een theorie waarin je hetzij moet vermenigvuldigen, hetzij moet optellen. Er is niets principieels aan.

Als je de realiteit verwacht met het model, dan is een dergelijk misverstand zo geboren. Binnen die verwarring lijkt het dan net alsof er een belangrijk principieel punt aan het licht komt, maar dat is dus niet zo. Het is gewoon een misverstand. AS doen voorkomen dat Bayesiaans redeneren ons gezond verstand zou overrulen: wij zouden van nature geneigd zijn op te tellen maar de theorie zou ons vertellen dat dat niet zo is. Maar zo zit het dus helemaal niet. Het is andersom: al naar gelang de meetmethode blijkt het een of het ander het juiste te zijn om te doen. Er is dus al met al geen reden om de denken dat Bayesiaans denken tegenintuïtief zou zijn.

Er is nog een andere reden om buitengewoon voorzichtig te zijn met getallen in relatie tot het combineren van bewijsmiddelen, namelijk de afhankelijkheden tussen bewijsmiddelen. Hoewel (wiskundige) afhankelijkheid een gecompliceerde theorie is, is de kern van de zaak is niet ingewikkeld: als de ene getuige beweert dat de dader een snor heeft en de andere dat hij een baard heeft, dan hebben die twee getuigenissen iets met elkaar te maken. Wat dat 'iets' is, is nog niet zo eenvoudig te bepalen maar de meeste mensen voelen wel

aan dat je die twee bevindingen op de een of andere manier aan elkaar moet relateren. Als een eerste getuige zegt dat de dader een baard heeft, dan is de kans dat hij ook een snor heeft groter geworden dan die zou zijn geweest zonder de informatie over de baard. Dat betekent dat de waarde van een tweede getuige die zegt dat de dader een snor heeft, daardoor minder informatie toevoegt in vergelijking met de situatie dat we nog niets wisten van de dader.

Dit soort samenhang laat zich in de praktijk niet makkelijk numeriek uitdrukken. Wat je natuurlijk kunt proberen te doen is het probleem als het ware weg te definiëren door een gecompliceerder model te maken waarin de afhankelijkheden ook numeriek uitgedrukt worden. Dat lost het probleem echter niet op, want je zult daarvoor allerlei modelkeuzes moeten maken die nog veel lastiger zijn dan de kansen op individuele bevindingen.

DE PRAKTIJK

Het is interessant om te zien hoe er zoal wordt gereageerd op het gebruik van berekeningen in de rechtspraak. We beginnen met twee citaten:

“De berekening van Derksen²² 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 kansrekening²³ gebruikt voor de strafrechtelijke waarheidsvinding waarvan de uitkomst in hoge mate

²² Filosoof Ton Derksen.

²³ De rechtbank bedoelt natuurlijk niet dat de regel binnen de wiskunde omstreden is, maar dat de toepassing ervan in het recht dat zou zijn.

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 en verwerpt het op deze rapporten gebaseerde verweer van de verdediging.”²⁴

“Het gebruik van getallen wekt een schijn van objectiviteit [...] en deze subjectieve kans kan niet objectief getoetst worden.”²⁵

Laten we deze citaten eens onder de loep nemen. In het eerste citaat wordt korte metten gemaakt met het uitrekenen van een schuldkans, en dat is precies wat wij ook betogen. Het is overigens interessant dat er zich onder de voorstanders van het rekenen in strafzaken niet of nauwelijks wiskundigen en statistici bevinden. Dat is vast geen toeval. Maar hoe dan ook, de rechter in het eerste citaat is buitengewoon verstandig omdat hij duidelijk maakt dat rechtspreken geen kansrekening is, en dat waardering en selectie van bewijsmiddelen aan de rechter is en aan niemand anders. Tegelijkertijd wordt de mogelijkheid open gehouden dat Bayesiaans redeneren wel degelijk behulpzaam kan zijn. In

²⁴ ECLI:NL:RBZWB:2016:3060.

²⁵ ECLI:NL:GHSHE:2018:421.

het tweede citaat wordt duidelijk gemaakt dat de rechters zich bewust zijn van de schijn van objectiviteit die getallen met zich brengen.

Als dit de typische reactie is in concrete zaken, waarom maken wij ons dan toch zorgen over rekenen in strafzaken? Kennelijk wordt de soep niet zo heet gegeten. De reden om toch bezorgd te zijn is het feit dat, zo is onze indruk, veel rechters (en andere juristen) denken dat Bayesiaans redeneren impliceert dat je berekeningen maakt. En omdat ze intuïtief wel aanvoelen dat de oplossing niet in rekenen kan liggen, concluderen ze vervolgens dat het hele Bayesiaanse denken wel problematisch moet zijn. Met andere woorden, ze gooien het kind met het badwater weg.

Af en toe wordt dit pijnlijk geïllustreerd in de praktijk. We schreven in hoofdstuk 1 al over de strafpleiter die tegenover een van ons (RM) stelde dat hij in ‘zijn’ zaak zeker geen statistici wilde: die zouden met getallen alles kunnen bewijzen wat ze maar wilden. Deze strafpleiter was afgeschrikt door statistische blunders zoals in de zaak Lucia de Berk. Voor hem was een Bayesiaanse analyse equivalent met het doorrekenen van een zaak, en daar had hij geen vertrouwen in.

De reden dat we bezorgd zijn, is dat het evident is dat rechters afhoudend zijn als iemand hen komt uitleggen hoe je een kans op schuld zou moeten uitrekenen. Die houding is terecht, zoals we hebben uitgelegd, maar we hopen dat de overwegingen in dit hoofdstuk handvatten bieden voor een beter geïnformeerde positiebepaling.

KENNISLEER

De overwegingen in dit hoofdstuk zijn zowel praktisch als filosofisch van aard. Zoals we al schreven voelen de meeste mensen wel aan dat je voorzichtig moet zijn met rekenen in

strafzaken, maar uitleggen waarom dat zo is, is nog niet zo eenvoudig: deze uitleg moet enigszins filosofisch van aard zijn. Filosofie houdt ons bij de les omdat we gedwongen worden om na te denken over wat we aan het doen zijn.

We zagen al enkele filosofische aspecten de revue passerend: we schreven over de relatie tussen model en werkelijkheid, over het feit dat je gebrek aan kennis niet zomaar kunt vervangen door algemeenheden, en natuurlijk ook over de filosofie van wat een kans of een percentage eigenlijk uitdrukt. Zijn er ook filosofische posities die juist pleiten vóór het gebruik van rekenen in strafzaken? Die zijn er, en het is interessant om aan het eind van dit hoofdstuk zo'n positie te schetsen en die vervolgens te becommentariëren. In het volgende hoofdstuk zullen ook nog enkele argumenten voor het gebruik van getallen besproken worden. Het volgende citaat komt uit het eerder genoemde artikel van Alkemade en Stikkelbroeck:

“Een tweede punt wat hier alvast besproken moet worden is het feit dat we in deze aanpak zo veel mogelijk getalsmatig willen gaan werken. Dat zal voor veel lezers erg onwennig, en misschien wel ongemakkelijk overkomen. En er schuilt een gevaar in: Een getal kan een onterechte illusie van exactheid oproepen. De voordelen van het getalsmatig werken wegen daar volgens ons echter tegen op: Door expliciet na te denken over wat je eigenlijk bedoelt als je denkt in termen als ‘soms’, ‘zo goed als uitgesloten’, ‘veel waarschijnlijker’, etc., ondervraag je als het ware je eigen intuïtie, en breng je expliciet in kaart wat je wel en niet weet of denkt te weten. Het getalsmatig werken moet dus niet gezien worden als een poging om ‘valse zekerheid’ te creëren, maar juist als een voorzet tot gericht nadenken en (innerlijke dan wel externe) discussie: Ben ik

het met dit getal eens? Waarom? Waarom niet? En als het op geen enkele zinnige manier lukt om een getal (of een getalsmatige marge van onzekerheid) toe te kennen aan een bewijskracht, wees dan gewaarschuwd: Waarom lukt dat niet? Als we *niets* getalsmatigs over deze bevinding kunnen zeggen, mogen we hem dan wel betrekken in ons oordeel? Bedenk ook – meer in het algemeen – dat *zonder* een getal het voor twee mensen erg moeilijk wordt om te weten of ze het al dan niet met elkaar eens zijn: De een kan bijvoorbeeld met het woordje ‘soms’ echt een totaal andere orde van grootte bedoelen dan de ander.”

Laten we dit citaat eens onder de loep nemen. We beginnen bij de claim dat als je geen numerieke getalswaardige uitspraak kunt doen over bepaalde bevindingen, je die bevinding misschien helemaal niet kunt betrekken in je oordeel. Welnu, een sterker pleidooi voor het paradigma ‘meten is weten’ is moeilijk te geven. Volgens AS is het zo dat iets pas meegenomen kan worden in de overwegingen als er een getal opgeplakt kan worden. Dat is een heel bijzondere en waarschijnlijk onhoudbare bewering die we nader moeten onderzoeken.

Er is geen enkele reden om te denken dat alles wat van belang is ook meetbaar of zinvol kwantificeerbaar is. De al eerder besproken verpleegkundige Lucia de Berk schreef in haar dagboek dat ze “toegegeven had aan haar neigingen”. Dat werd begrijpelijk opgevat als een aanwijzing voor haar schuld, want waar zou het citaat anders over kunnen gaan? Maar hoe kwantificeer je een aanwijzing zoals deze? Elke poging daartoe lijkt volkomen uit de lucht gegrepen. De dagboek aantekening van Lucia de Berk is evident interessant in de strafzaak tegen haar, maar dat wil niet zeggen dat ze kwantificeerbaar is.

Het beoordelen van menselijk gedrag is een buitengewoon complexe zaak waar meestal helemaal niet aan gerekend kan worden. We zien dat we hier dus te maken hebben met een bepaalde kennistheoretisch filosofische positie dat ‘alles’ kwantificeerbaar is. Het is alleen binnen dat paradigma dat de rekenmodellen voor dit soort bewijs zinvol zouden kunnen zijn.

En dan is daar de bewering in het citaat dat het moeilijk is voor twee mensen om zonder getallen te gebruiken te weten of ze het met elkaar eens zijn. De vraag is in hoeverre het nodig is om dat altijd te kunnen weten, maar hoe dan ook gaat het gebruik van getallen dit niet oplossen. We zagen al dat getallen persoonlijke zaken zijn en dat wat de een overtuigend vindt voor een ander niet overtuigend is. Er is ook geen reden om te denken dat twee mensen die hetzelfde getal produceren, het wél met elkaar eens zijn. Die persoonsgebonden getallen kunnen niet met elkaar vergeleken worden, uitzonderingen daargelaten. Het is verder ook niet nodig dat iemand zich afvraagt of hij het met een getal al dan niet eens is. Als je niet denkt dat de waarde van bevindingen in getallen kunnen worden uitgedrukt, dan heeft het weinig zin om je af te vragen of je het met een bepaald getal wel of niet eens bent. We zijn dan ook helemaal niet overtuigd dat het werken met getallen de voordelen met zich meebrengt waar AS het over hebben.

Om terug te keren bij het dagboekvoorbeeld van Lucia de Berk: er is gesuggereerd dat deze bevinding een likelihood ratio van 5 op zou leveren.²⁶ Stel nu eens dat er een andere, onafhankelijke, bevinding is waarvan de likelihood ratio 1/5 is. Omdat we in het geval van onafhankelijkheid de likeli-

²⁶ Zie Ronald Meester en Klaas Slooten, *Probability and Forensic Evidence*, CUP 2021, bladzijde p. 124.

hood ratio's met elkaar mogen vermenigvuldigen zouden deze bevinding en het dagboek samen neutraal zijn. Dat zou toch een bizarre conclusie zijn. Het zou de keuze zijn van de rekenaar om die bevindingen elkaar op te laten heffen en niet het gevolg zijn van de bevindingen zelf. Daar is alles mee gezegd: de numerieke waarde van de bevindingen in een rekenmodel zijn het gevolg van persoonlijke en daarmee tamelijk arbitraire keuzes, en nergens anders van. Het is, met andere woorden, niet zo dat je op basis van het model tot de conclusie komt dat de bewijsmiddelen elkaar 'opheffen', het is de maker van het model die zo besloten heeft. Het is belangrijk om dat altijd voor ogen te houden.

Voorstanders van rekenen, zie ook AS in het citaat hierboven geven soms als argument dat je door het rekenen wordt gedwongen je eigen gedachten kritisch te beoordelen. Het ontgaat ons echter waarom bij een dergelijke kritische zelfevaluatie getallen nodig zijn. Als de getallen namelijk uit zouden wijzen dat er een onverwachte of vreemde conclusie uit volgt, welke conclusie moet je dan trekken? Dat je gedachtegang niet correct is, of dat de keuzes van de getallen niet deugen?

Een ander nadeel bij het maken van berekeningen is dat de verdediging werkelijk elke gemaakte berekening kan bekritisieren en vragen kan stellen bij de reden waarom voor een bepaald getal gekozen is. Dit is een praktisch punt en niet een principieel bezwaar. Een uitspraak gebaseerd op dit soort getallen zou om die reden al onhoudbaar zijn in de behandeling van een zaak. Als dat zo is, waarom zou je er dan eigenlijk aan beginnen?

Er is, al met al, niet veel reden om getallen en berekeningen in rechtszaken heel serieus te nemen. Dat betekent nadrukkelijk niet dat Bayesiaans denken niet serieus genomen zou moeten worden. Bayesiaans denken komt in rechts-

zaken uitstekend tot zijn recht, niet omdat we een bepaalde manier van denken zouden willen opdringen maar omdat ons menselijk denkproces door dit paradigma op zo'n manier beschreven wordt dat het ons behoedt voor het maken van logische fouten. Maar het blijft een model.

HOOFDSTUK 4

Bayesiaans redeneren en strafrechtelijk bewijsrecht

In de afgelopen drie hoofdstukken hebben we met veel praktijkvoorbeelden, enige theorie en een beetje filosofie stilgestaan bij Bayesiaans redeneren. We weten nu hoe het werkt. Dit laat onverlet dat we nog moeten onderzoeken in hoeverre deze manier van redeneren past binnen het (Nederlandse) strafrechtelijke bewijsrecht. Dit is zeker geen uitgemakke zaak. Er zijn in de literatuur een aantal bezwaren naar voren gebracht waar we ons toe moeten verhouden. In dit hoofdstuk bespreken we achtereenvolgens de onschuld-presumptie, de gedachte dat Bayesiaans redeneren leidt tot het ‘binnenhalen’ van persoonlijke overtuigingen van de rechter, schakelbewijs, het zwijgrecht van de verdachte, en ten slotte de bewijsminimumregel. De aard van het onderwerp van dit hoofdstuk maakt dat het iets theoretischer is dan de voorgaande drie hoofdstukken. Dat maken we later weer goed wanneer we meer zaken gaan bespreken, maar het is nu eenmaal zo dat we moeten onderzoeken in hoeverre de specifieke manier van redeneren onder onzekerheid die we hier bepleiten congruent is met wat juridisch kan en mag.

DE ONSCHULDPRESUMPTIE

We hebben gezien dat Bayesiaans redeneren erkent dat een rechter niet bij ‘nul’ kan beginnen met zijn evaluatie van het bewijsmateriaal in een strafzaak; er is hoe dan ook iets als

een oorspronkelijke overtuiging of oorspronkelijke kijk op de zaak. Vanuit juridisch oogpunt kan de vraag worden gesteld of zo'n oorspronkelijke overtuiging over schuld of onschuld niet in strijd is met de onschuldpresumptie.

De onschuldpresumptie is vastgelegd in verschillende (internationale) juridische documenten, en houdt kort gezegd in dat verdachten en beschuldigen als onschuldig worden beschouwd totdat hun schuld volgens de wet is bewezen. Zo stelt bijvoorbeeld artikel 11 van de Universele Verklaring van de Rechten van de Mens het volgende:

“Een ieder, die wegens een strafbaar feit wordt vervolgd, heeft er recht op voor onschuldig gehouden te worden, totdat zijn schuld krachtens de wet bewezen wordt in een openbare rechtzitting, waarbij hem alle waarborgen, nodig voor zijn verdediging, zijn toegekend.”

Deze algemene bewoordingen geven nog vrij weinig prijs over waar het bij de onschuldpresumptie nu precies om draait. Op het eerste gezicht lijkt het concept zelfs tegenstrijdig. Veronderstellen dat een verdachte of beklaagde onschuldig is: hoe kan dat? Er bestaat immers een redelijke verdenking dat hij of zij een misdaad heeft begaan, anders zou de verdachte daar niet staan. Het lijkt op zijn minst nogal naïef om onschuld te veronderstellen onder belastende omstandigheden; de verdachte is niet zomaar iemand en hij staat niet voor niets voor de rechter.

Het is daarom belangrijk te beseffen wat de onschuldpresumptie níét is. Het is nadrukkelijk geen feitelijke veronderstelling. Het zegt niets over de werkelijke schuld van de verdachte of beklaagde in welke fase van de strafprocedure dan ook. Geen enkele politieagent, aanklager of rechter zal ooit veronderstellen dat de verdachte die is gearresteerd voor bijvoorbeeld het bezit van kinderporno, onschuldig is. Inte-

gendeel, een verdachte wordt gearresteerd en berecht, precies omdat hij wordt verdacht van deze specifieke misdaad. Dat wil zeggen, op basis van een feitelijke ‘redelijke verdenking’ (bijvoorbeeld, de computer van de verdachte bevat pornografische beelden) die tijdens het proces verder onderbouwd moet worden (hoe zijn deze beelden op de computer terechtgekomen?). Een ‘oorspronkelijke’ verdenking aan het begin van een proces is daarom redelijk. De oplettende lezer zal inzien dat dit in lijn is met het Bayesiaans denken waarin het gaat om het steeds bijstellen van overtuigingen op basis van nieuwe informatie.

Maar wat is de onschuldpresumptie dan wel? De betekenis van onschuldpresumptie is onderwerp geweest van uitgebreide discussie en verschillende interpretaties blijven naast elkaar bestaan. Zo wordt het principe op het Europese continent ook geassocieerd met hoe de beschuldigde behandeld moet worden tijdens het strafproces, terwijl in de Angelsaksische traditie juist de bewijsdimensie van groot belang is. Dit boek is niet de plaats om die discussie uit te diepen, maar wij leiden eruit af dat de onschuldpresumptie een gemeenschappelijke basis heeft die vier (gerelateerde) standaarden omvat met betrekking tot bewijs:

1. De rechter of jury mag de schuld van de verdachte niet als uitgangspunt nemen.
2. Bewijs van schuld moet buiten redelijke twijfel zijn voor een veroordeling.
3. De bewijslast ligt bij de aanklager.
4. Twijfel werkt in het voordeel van de verdachte (in dubio pro reo).

Deze standaarden betreffen algemene normen die een aanklager of rechter moet onderbouwen op basis van de omstandigheden in een specifieke zaak. Hun interpreta-

tie hangt verder af van vele factoren, zoals de specifieke bewijsregels binnen een rechtsgebied, de kenmerken van een rechtssysteem, het politieke klimaat, et cetera. De rationale en het doel zijn echter duidelijk: het voorkomen van de veroordeling van onschuldigen. Als zodanig is de onschuldpresumptie een principe dat functioneert op een vrij abstract niveau. Hiermee bedoelen we dat bijvoorbeeld de bewijslast, op wetgevend of normatief niveau, in wezen een wederzijdse rolverdeling is tussen de aanklager en de beschuldigde. De aanklager moet de beschuldiging bewijzen, terwijl het aan de beschuldigde is om (redelijke) twijfel te zaaien. De verantwoordelijkheid van de beschuldigde eindigt op het punt waar zijn rol zou verschuiven naar het bewijzen van onschuld (en vice versa, waar de aanklager niet zou voldoen aan de bewijslast). Hoe dat evenwicht precies uitpakt, varieert van geval tot geval, afhankelijk van de feiten en omstandigheden van een zaak, maar is ook afhankelijk van hoe strafrechtelijke aansprakelijkheid precies is gedefinieerd. Hoewel dat misschien vreemd klinkt, zijn de bewijslast voor enerzijds doodslag en anderzijds snelheidsovertredingen *in principe* gelijk. Maar voor doodslag betekent dit in concreto dat een daad van geweld moet worden bewezen, terwijl ‘eigenaar van de auto zijn’ – aanvankelijk – voldoende is om strafrechtelijke aansprakelijkheid voor snelheidsovertredingen vast te stellen.

Samenvattend is de onschuldpresumptie een abstract normatief bewijsrechtelijk uitgangspunt dat nadere uitwerking krijgt in – nog steeds abstracte – bewijsregels als ‘veroordeeling buiten redelijke twijfel’, waarbij die bewijsregels in een specifieke zaak concreet moeten worden ingevuld op basis van zaakspecifieke informatie. De onschuldpresumptie is geen feitelijke veronderstelling, maar een fundamentele, hoewel vrij algemene, waarschuwing aan de rechters. Misschien zou je kunnen zeggen dat ze functioneert als het geweten van het strafproces. Haar taak is ons eraan te herinneren dat het

veroordelen van een mens gerechtvaardigd moet zijn binnen de context van een specifieke strafprocedure.

Binnen deze betekenis kan de onschuldpresumptie dan ook nooit in strijd zijn met het Bayesiaanse uitgangspunt dat er een oorspronkelijke kijk op een zaak moet worden geëxpliciteerd op basis van de context die dan voorhanden is. Het wordt echter wel problematisch als je denkt dat dat startpunt kan worden uitgedrukt in een bepaald vast getal als uitdrukking van de onschuldpresumptie, zoals we analyseerden in hoofdstuk 3. Deze benadering is in onze ogen niet alleen vanuit logisch oogpunt maar ook juridisch absurd omdat zij erop neer komt dat de onschuldpresumptie wordt gezien als een feitelijke presumptie van (on)schuld die in elke zaak gelijk is. Het zou immers betekenen dat voor elke verdachte van welk misdrijf dan ook aanvankelijk dezelfde belastende feiten gelden, en het behoeft geen betoog dat dit evident onredelijk is. Er zijn immers immense verschillen tussen verschillende zaken en verschillende verdachten.

PERSOONLIJKE OVERTUIGINGEN

De vraag waar het in deze sectie om draait is de volgende: als een rechter veroordelen heeft, dan worden deze gerealiseerd in de prior overtuiging van de rechter, en dat kan tot gevolg hebben dat iemand die aan die veroordelen voldoet, sneller wordt veroordeeld dan iemand die dat niet doet. Dat is een serieuze bedenking waar we niet omheen kunnen. We bespreken het probleem aan de hand van een concreet voorbeeld.

Rechtsgeleerde Sean Sullivan bespreekt in een van zijn publicaties precies dit mogelijke gevaar dat persoonlijke over-

tuigingen invloed kunnen hebben op het vonnis.¹ Hij doet dit aan de hand van een zaak van het Hooggerechtshof in de Verenigde Staten: *Pena-Rodriguez v. Colorado* (137 S. Ct. at 862). Deze zaak betrof een jurylid dat tijdens de beraadslagingen uitlegde dat de Latijns-Amerikaanse verdachte, die terechtstond voor seksuele intimidatie, schuldig kon worden geacht omdat “in [zijn] ervaring als ex-wetshandhavingsfunctionaris Mexicaanse mannen een bravoure hadden die hen deed geloven dat ze alles met vrouwen konden doen wat ze wilden”. Het jurylid deelde verder zijn mening dat Mexicaanse mannen fysiek controlerend waren over vrouwen: “Ik denk dat hij het deed omdat hij Mexicaan is en Mexicaanse mannen nemen wat ze willen.” Alsof hij alle twijfel weg wilde nemen dat deze uitspraken bedoeld waren om de aanvankelijke kans van de schuld van de verdachte te beschrijven, concludeerde het jurylid dat “in zijn ervaring, negen van de tien keer Mexicaanse mannen schuldig waren aan agressief gedrag tegenover vrouwen en jonge meisjes”. Sullivan concludeert:

“Stel dat deze overtuigingen oprecht werden gedeeld door het jurylid. Elk feitenonderzoek dat gebaseerd is op de persoonlijke overtuigingen van deze feitenonderzoeker over de feiten, begint noodzakelijkerwijs vanuit de eerdere overtuigingen van de feitenonderzoeker over de feiten—waaronder eerdere overtuigingen die op hun beurt gebaseerd zijn op raciale vooroordelen, genderstereotypen, aannames over religieuze groepen, en dergelijke. Een Bayesiaanse benadering van feitenonderzoek gedooft niet alleen, maar vereist eigenlijk het gebruik van deze persoonlijke eerdere

¹ Sullivan, S. P. (2019) *A Likelihood Story: The Theory of Legal Fact-Finding*, University of Colorado Law Review, 1 (onze vertaling).

overtuigingen. Op deze manier schrijft het een differentiële behandeling van verdachten voor. Toegepast op de situatie in *Pena-Rodriguez*, zou een kans- of geloofsgebaseerde benadering van feitenonderzoek het bevooroordeelde jurylid instrueren om alleen het gepresenteerde bewijs tijdens het proces in overweging te nemen, en om een kleinere hoeveelheid bewijs te eisen om een Mexicaanse verdachte te veroordelen dan om een anders identieke verdachte van een andere bevolkingsgroep te veroordelen.”

Heeft Sullivan hier een punt? Wij denken uiteindelijk van niet, en we zullen uitleggen hoe we tot die conclusie komen. Het is duidelijk dat het in overweging nemen van de vermeende aard van Mexicaanse mannen op de manier die Sullivan beschrijft ongepast is, daar bestaat volgens ons geen twijfel over. We hebben de problemen rondom algemene statistieken al eerder besproken, maar de kritiek met betrekking tot deze zaak gaat verder. Het is niet zozeer een statistiek die ten grondslag ligt aan de overtuiging van het jurylid, maar eerder een persoonlijk vooroordeel over Mexicaanse mannen; men zou kunnen zeggen dat het een vermeende statistiek is.

Maar we zien ook dat het ongepaste gebruik van persoonlijke overtuigingen in deze zaak duidelijk werd, juist doordat het jurylid zich gedwongen voelde om er commentaar op te geven. Blijkbaar vond hij het nodig om zijn uiteindelijke standpunt uit te leggen. Het werd daarom overduidelijk wat er aan de hand was.

Een Bayesiaanse benadering dwingt iemand om rekening te houden met de context, maar kan op zichzelf niet voorkomen dat dit op een onredelijke manier gebeurt. Het feit dat ‘Bayes’ misbruikt kan worden, maakt de benadering in zichzelf niet ongeldig of ongepast.

Bayesiaans denken is uiteraard niet immuun voor misbruik, geen enkele methodologie is dat, maar dit betekent niet dat het systeem zelf ‘fout’ is. Integendeel, we hebben betoogd dat juridisch feitenonderzoek zonder rekening te houden met de context tot systematische problemen leidt: het negeren van de context is een principiële fout. Dit betekent niet dat de context altijd op een passende manier in overweging wordt genomen. Het voorbeeld dat door Sullivan wordt aangehaald, is daar een duidelijk voorbeeld van. Uiteraard betekent dit niet dat we moeten overstappen op een manier van bewijsverwerking die vanaf het begin gebrekkig is. In plaats daarvan moeten we een rechtssysteem ontwerpen waarin het moeilijk of onmogelijk wordt om Bayesiaans denken op de manier te gebruiken zoals het jurylid doet in Sullivans voorbeeld.

Wellicht zijn niet alle rechtssystemen even kwetsbaar voor dergelijk misbruik van Bayesiaans denken. In bepaalde rechtsgebieden (bijvoorbeeld in Nederland) wordt de rechter wettelijk gedwongen om een begrijpelijk vonnis te schrijven. In dat geval zouden de argumenten van het jurylid in de bovengenoemde zaak simpelweg niet acceptabel zijn. In het geval van juryrechtspraak is het vermijden van ongewenste argumenten wellicht moeilijker, maar dit is enigszins speculatief. Tegelijkertijd is het natuurlijk ook zo dat motiverings-eisen niet geheel kunnen voorkomen dat rechters persoonlijke overtuigingen en vooroordelen meenemen in hun beslissingen maar deze niet opschrijven.

In ieder geval is het zeker niet waar, zoals Sullivan beweert, dat Bayesiaans denken een kleinere hoeveelheid bewijs zou vereisen om een Mexicaanse verdachte te veroordelen dan om een anders identieke verdachte van een andere bevolkingsgroep te veroordelen. Dat zou namelijk misbruik van het systeem zijn, en zeker geen noodzakelijke conse-

quentie. Bayes nodigt ons uit om de context op een ‘wet-tige’ manier in overweging te nemen. Het is geen vrijbrief om argumenten te gebruiken die niet direct verband houden met de zaak in kwestie, zoals we al concludeerden toen we algemene statistieken bespraken. Evenzo is het geen vrijbrief om persoonlijke overtuigingen te gebruiken als vermeende statistische ‘feiten’.

Zelfs wanneer een jurylid er persoonlijk van overtuigd is dat een verdachte schuldig is, moet hij zich bewust zijn van het feit dat zijn eerdere overtuiging mogelijk niet wettig is. Daarom kan het gevaar van misbruik geen reden zijn om Bayesiaans denken niet te gebruiken. Nogmaals: er is geen systeem, geen methodologie die niet op de een of andere manier vatbaar is voor misbruik of ondeskundig gebruik. Daar komt bij dat in Nederland het nu eenmaal zo is dat de rechter overtuigd moet zijn dat er voldoende wettig bewijs is om een verdachte te veroordelen. Dit brengt, we hebben dit uitgelegd in hoofdstuk 2, noodzakelijkerwijs een bepaalde mate van subjectiviteit met zich mee. Binnen de marges die de wet geeft, is dat niet problematisch, en zelfs onvermijdelijk. Iemand die dat anders ziet, zegt in feite dat we rechtspraak in principe zouden kunnen automatiseren, en wij kennen eigenlijk geen wetenschappers of juristen die dat voor hun rekening zouden willen nemen.

SCHAKELBEWIJS

Hoe zou een Bayesiaanse interpretatie van schakelbewijs eruit zien? ‘Schakelen’ van bewijs is het gebruiken voor het bewijs van de overeenkomsten tussen de *modus operandi* in verschillende zaken. Schakelen komt voor toepassing in aanmerking indien bij de vergelijking van delicten specifieke,

onderscheidende overeenkomsten worden waargenomen, terwijl significante verschillen ontbreken. Twee voorbeelden kunnen het concept wat vlees op de botten geven.

Een man werd verdacht van driemaal doodslag op kwetsbare jonge vrouwen, verslaafde prostituees. Twee van die zaken – één van een vrouw die naakt en met ingeslagen hoofd op een talud werd aangetroffen en één van een vrouw wier lichaam in delen werd aangetroffen in een vuilniszak – leken sterk op elkaar. Niet alleen kon de verdachte op basis van DNA worden gekoppeld aan (seks met) het eerste slachtoffer en aan (het dichtknopen van) de vuilniszakken waarin de lichaamsdelen van het tweede slachtoffer zaten, ook kon het eerste slachtoffer door middel van DNA op de vuilniszak met lichaamsdelen van slachtoffer 2 en DNA in de auto van slachtoffer 2, worden gekoppeld aan dat tweede slachtoffer. De derde zaak leek veel zwakker: vast stond dat de verdachte een seksrelatie had met het slachtoffer en haar als laatste had gezien. Daarnaast wees zijn gedrag erop dat hij na het laatste contact ‘bewijs’ creëerde waaruit zou moeten blijken dat ze nog leefde. Ook heeft hij zich dagenlang onbereikbaar gehouden voor de politie. Maar een lijk is nooit aangetroffen. Ze zou ook zelfmoord kunnen hebben gepleegd of door een noodlottig ongeval om het leven kunnen zijn gekomen. Geen lijk, geen doodsoorzaak. Hij werd toch veroordeeld, met behulp van die andere twee zaken, die overigens vele jaren eerder speelden.² De derde zaak paste immers in het bijzondere patroon van misbruik maken van jonge vrouwen in een zeer kwetsbare positie waarbij de verdachte zich vervolgens van hen onttreedt.

² Zie HR 20 december 2022, ECLI:NL:HR:2022:1902, NJ 2023/102, m.nt. Reijntjes.

Misschien wel de meest bekende toepassing van een schakelconstructie in de recente geschiedenis betreft de zaak tegen Jos B. Deze man werd verdacht van het misbruiken en doden van de elfjarige Nicky Verstappen toen deze in de zomer op kamp was op de Limburgse hei.³ Voor het bewijzen van het misbruik kon gebruik worden gemaakt van het DNA van Jos B. dat op het lichaam en in de onderbroek van Nicky was gevonden; ook zaten de kleren van Nicky achterstevoren en binnenstebuiten. Bewijzen dat B. verantwoordelijk was voor het doden van het slachtoffer was moeilijker omdat geen doodsoorzaak kon worden vastgesteld bij Nicky. Hier werd het schakelbewijs cruciaal. Het hof gebruikte twee oude geseponeerde zaken waarin B. twee jongens op de hei benaderde, vastgreep, de mond dichthield en betastte. Het was vooral dat snoeren van de mond dat uiteindelijk bijdroeg aan de conclusie dat Jos B. schuldig was aan doodslag.

Welnu, wat kan een Bayesiaans perspectief zeggen over deze twee schakelzaken in het bijzonder en over schakelbewijs in het algemeen? Wat dat laatste betreft is het vanuit Bayesiaans oogpunt evident – logisch – dat feiten uit de ene zaak als zinvolle informatie kunnen dienen in een andere zaak. Afhankelijk van de aard van de informatie geeft Bayes soms de keuze of we bepaalde aanwijzingen willen opvatten als achtergrond of context enerzijds, of als zaakspecifiek bewijs dat direct ziet op de ten laste gelegde handelingen anderzijds. Beide vormen van informatie kunnen als juridisch bewijs dienen, zoals we hiervoor hebben gezien. Het schakelbewijs dat in de zaak Nicky Verstappen werd gebruikt vinden wij eerder geschikt om te dienen als context en achtergrond, en dan ook nog eens alleen als het gaat om de beschuldiging van het zedendelict. De hei waar B. vlakbij

³ HR 26 september 2023, ECLI:NL:HR:2023:1303.

woont, de zomer en het vastpakken van jonge jongens, het is allemaal relevante informatie maar het zegt helemaal niets over het doden van Nicky. In de oude zaken is immers juist niemand omgebracht en ten aanzien van de dood van Nicky hebben we geen idee wat zich precies heeft afgespeeld, over welke handelingen Jos B. zou hebben verricht tasten we op basis van het bewijs totaal in het duister. De vergelijking loopt daar dus volledig mank. Eigenlijk zien we hier niet zozeer een schakelconstructie, maar eerder een opbouw van een prior tegen de verdachte Jos B. Daarmee willen we niets afdoen aan de moeilijkheid van deze beslissing, en we zeggen ook niet dat het oordeel van het hof onjuist is geweest of dat de Hoge Raad strenger zou moeten oordelen over schakelconstructies. Er zijn juridische en rechtspolitieke overwegingen die deze beslissingen begrijpelijk maken. Wat we wel zeggen is dat Bayes hier helpt in te zien dat de schakelconstructie vanuit logisch oogpunt eigenlijk geen schakelconstructie mag heten. Met een Bayesiaanse bril op wordt duidelijk dat het schakelbewijs niet onderscheidend is voor een veroordeling voor doodslag. Strikt genomen voldoet dit schakelbewijs dan ook niet aan de juridische definitie die we aan het begin van deze paragraaf gaven.

Bij de zaak van de drie verslaafde prostituees ligt dat anders. Daar is sprake van een bonafide schakelconstructie. Hoewel er geen lichaam is gevonden in de derde zaak werd de verdachte toch veroordeeld, op basis van de bewezen geachte doodslag op de andere vrouwen en de vergelijkbare omstandigheden. Het interessante hier is dat Bayes laat zien dat juist het ontbreken van het lichaam van de derde vrouw, wat in eerste instantie het zwakke punt lijkt, de schakelconstructie zeer onderscheidend maakt. Dat wordt duidelijk als we de twee bewezen zaken als uitgangspunt nemen en in de derde zaak de hypothesen beschouwen dat ofwel de verdachte

haar heeft vermoord en het lichaam heeft kwijtgemaakt, ofwel dat het slachtoffer is overleden door een ongeval of door zelfdoding. Onder de eerste hypothese is het op basis van de andere twee zaken volkomen begrijpelijk dat het lichaam van de derde vrouw onvindbaar is, terwijl dat onder de tweede en derde hypothese veel minder te verwachten is. Bij de meeste ongevallen en doorgaans ook in geval van zelfdodingen, wordt er immers wel een lichaam aangetroffen. Bij een man die twee kwetsbare vrouwen doodt en zodanig verminkt dat ze onherkenbaar dan wel moeilijk vindbaar zijn, past een derde kwetsbare en onvindbare vrouw veel beter.

Wat leren we nu uit deze twee verschillende voorbeelden? Wij denken het volgende: schakelbewijs is een juridische constructie die in een Bayesiaans kader zowel in de context, als in het zaaksspecifieke bewijs gebruikt kan worden, afhankelijk van de details van de zaak. Er bestaat ons inziens niet zoiets als een theorie van Bayesiaans schakelbewijs. De relatie is andersom: we kunnen schakelbewijs in een Bayesiaans kader begrijpen. En net als in ‘gewoon’ bewijs komt schakelbewijs en zijn onderscheidend vermogen in gradaties. Er is geen enkele logische reden om algemeen kritisch te zijn over schakelbewijs, en de jurisprudentie laat zien dat het ook juridisch gezien toegestaan is, onder de juiste omstandigheden en met de juiste argumenten. Een Bayesiaanse analyse van het gebruik van schakelbewijs in een concrete zaak kan daarbij laten zien hoe onderscheidend dat bewijs is.

ZWIJGRECHT, PROCESHOUDING EN BAYES

Een verdachte heeft het recht om te zwijgen, en het zwijgen mag niet worden gebruikt als bewijsmiddel voor een veroordeling. Maar zwijgen kan wel degelijk bijdragen aan het

bewijzen van een strafbaar feit en dat geldt ook voor andere proceshoudingen – denk aan liegen, of ongeloofwaardig en onaannemelijk verklaren door de verdachte.

Hoe verhoudt het gebruiken van de proceshouding van de verdachte zich tot het redeneren onder onzekerheid en Bayes? Laten we met die vraag in ons achterhoofd eens kijken naar de rechtspraak inzake witwassen, waarin de proceshouding van de verdachte een grote rol speelt. Het beslismodel dat in die rechtspraak is geformuleerd, komt erop neer dat zodra de door het openbaar ministerie aangedragen feiten en omstandigheden een vermoeden rechtvaardigen dat het niet anders kan zijn dan dat het voorwerp uit enig misdrijf afkomstig is, van de verdachte mag worden verlangd dat hij een concrete, verifieerbare en niet op voorhand hoogst onwaarschijnlijke verklaring geeft dat een voorwerp niet van misdrijf afkomstig is. Geeft de verdachte een dergelijke verklaring, dan is het aan het openbaar ministerie om daar onderzoek naar te doen. De rechter zal vervolgens, mede op basis van de resultaten van dat onderzoek, moeten beoordelen of ondanks de verklaring van de verdachte het witwassen bewezen kan worden op de grond dat (het niet anders kan zijn dan dat) het voorwerp uit enig misdrijf afkomstig is. In het geval dat de verdachte niet een dergelijke verklaring geeft, mag de rechter die omstandigheid betrekken in zijn overwegingen omtrent het bewijs. Datzelfde geldt voor een verklaring die uiteindelijk niet aannemelijk is geworden.

Juist vanwege de rol van de proceshouding bij het bewijzen wordt het beslisschema vanuit de strafrechtswetenschap en de strafrechtspraktijk met enige argwaan bekeken. Wat het witwasschema naar ons idee echter vooral laat zien is dat het volledig in lijn is met de Bayesiaanse uitgangspunten (1) dat bewijskracht relatief is en dat derhalve in alternatieve scenario's of hypotheses moet worden gedacht, en (2) dat een

redenering niet bij nul begint en context van belang is om die scenario's op hun eigen merites te kunnen beoordelen. Het is in die zin een redeneermodel dat niet exclusief is voor witwassen maar bij het bewijzen van alle delicten kan worden toegepast. Het meenemen van (de kwaliteit van) een alternatieve verklaring van de verdachte is daarbij een logisch onderdeel van een bewijsredenering – en van nader onderzoek – en niet per definitie iets dat ten nadele van de verdachte werkt. Het kan immers net zo goed de andere kant op werken. Een niet onwaarschijnlijk verhaal van de verdachte, dat het bewijs ongeveer even goed verklaart als het verhaal van het openbaar ministerie, kan ervoor zorgen dat de drempel voor veroordeling niet wordt gehaald.

Het Bayesiaanse perspectief maakt hiermee inzichtelijk dat de bewijsrisico's bij witwassen niet zozeer zitten bij het mogen meenemen van de proceshouding van de verdachte bij het bewijzen van het feit. De bewijsrisico's ten aanzien van witwassen – een omkering van het bewijs of een inbreuk op het zwijgrecht en de verklaringsvrijheid – worden vooral gecreëerd doordat materieelrechtelijk van het bewijs niet veel wordt verwacht. Het 'prototypescenario' waar het bestanddeel 'afkomstig uit enig misdrijf' om vraagt is immers niet meer dan 'verdachte bezit heel dure spullen en heeft geen legaal inkomen'. Daarmee ontstaat vervolgens de verleiding voor het openbaar ministerie om dat prototypescenario weinig concreet te onderbouwen en de bal bij de verdachte te leggen.

De oplettende lezer heeft wellicht opgemerkt dat we in het voorgaande spreken over 'de proceshouding van de verdachte' en dat daarmee geen onderscheid wordt gemaakt tussen de zwijgende verdachte en de verdachte die een verklaring over een alternatief scenario aflegt. Hiermee sluiten we aan bij de Hoge Raad, die in zijn witwasrechtspraak

de op voorhand hoogst onwaarschijnlijke verklaring en de niet-aannemelijk geworden verklaring op één lijn plaatst met het zwijgen (het niet geven van een niet op voorhand hoogst onwaarschijnlijke verklaring) voor wat betreft het kunnen meenemen in de redenering en het eindoordeel van de rechter.⁴ Dit suggereert dat volgens de Hoge Raad het zwijgen in de bewijsconstructie een met het alternatieve scenario vergelijkbare rol heeft, en dat roept een vraag op. Want is het gebruik van het zwijgen niet iets heel anders dan het gebruik van een alternatief scenario dat niet wordt geloofd? Is het gebruik van het zwijgen niet simpelweg het gebruiken van de ervaringsregel ‘de verdachte zwijgt omdat hij het gedaan heeft’?

Ook hier kan Bayes duidelijk maken dat het vooral een toepassingsrisico betreft en kan de Bayesiaanse redeneerstructuur bovendien helpen om dat risico zich niet te laten verwezenlijken. Vanuit het uitgangspunt ‘bewijs is relatief’ kan het zwijgen van de verdachte namelijk worden gezien als een ‘alternatief scenario dat zich kennelijk niet heeft laten formuleren’. Als de verdachte niet met een redelijk alternatief komt, en de rechter heeft zelf ook kunnen constateren dat er op basis van de beschikbare informatie geen enkel redelijk alternatief bestaat dat de feiten kan verklaren, dan heeft die constatering zelfstandige waarde. Bayes zegt dan dat het scenario van het openbaar ministerie, de tenlastelegging, daardoor wint aan geloofwaardigheid. Meer Bayesiaans gezegd: de a priori kans op het alternatieve scenario is erg klein. In de terminologie van het witwasmodel zou je ook kunnen zeggen dat er sprake is van een situatie vergelijkbaar met de situatie waarin het alternatieve scenario op voorhand

⁴ Zie bijv. de laatste zin van r.o. 2.3 in HR 9 juli 2019, ECLI:NL:HR:2019:1137, NJ 2019/350.

hoogst onwaarschijnlijk is. Dit alles betekent ook dat zwijgen niet pas wordt beoordeeld als het bewijs al is geleverd. Het zwijgen draagt als niet te formuleren alternatief en in relatie tot het bewijs bij aan de bewijsbaarheid van het ten laste gelegde. De hamvraag of gevolgtrekkingen uit het zwijgen te ver gaan, loopt daarmee parallel aan de vraag of het bewijs te weinig zegt. Het antwoord is iedere keer weer afhankelijk van de omstandigheden van het geval en dat antwoord is aan de rechter; Bayes gaat daar niet over.

DE BEWIJSMINIMUMREGEL EN BAYES

De bewijsminimumregel van artikel 342 lid 2 Wetboek van Strafvordering vertelt ons dat een verdachte niet mag worden veroordeeld op basis van de verklaring van één getuige. Er moet een tweede bewijsmiddel zijn. De Hoge Raad geeft invulling aan deze regel op basis van de ratio van de deugdelijkheid van de bewijsbeslissing en met behulp van het criterium dat het tweede bewijsmiddel ‘voldoende steun’ moet bieden aan de getuigenverklaring opdat die verklaring niet op zichzelf staat. Daarbij geldt dat het steunbewijs de bewezenverklaring in haar geheel betreft en niet een onderdeel daarvan, en ook dat het steunbewijs geen betrekking hoeft te hebben op de gedragingen zoals letterlijk ten laste gelegd.⁵ Concreter dan dit kunnen we de regel niet maken, zo geeft de Hoge Raad in zijn rechtspraak aan. Wat de regel precies

⁵ Zie bijvoorbeeld HR 13 juli 2010, ECLI:NL:HR:2010:BM2452, *NJ* 2010/515, m.nt. M.J. Borgers; HR 15 mei 2018, *NJ* 2019/298, m.nt. Rozemond; HR 2 november 2021, ECLI:NL:HR:2021:1594, *NJ* 2022/32, m.nt. Vellinga; HR 10 september 2024, *NJ* 2025/8, m.nt. Vellinga.

behelst, moet duidelijk worden door de afzonderlijke beslissingen van de Hoge Raad in concrete gevallen.

Op zoek naar meer houvast wordt vervolgens door commentatoren van de rechtspraak geprobeerd vuistregels te formuleren en patronen te ontdekken. En dat blijkt bepaald geen sinecure. Een veronderstelde vuistregel is bijvoorbeeld ‘de emotie van de aangever die direct na het vermeende feit door een getuige wordt waargenomen’. Maar het is niet altijd direct duidelijk waarom in het ene geval deze ‘vuistregel’ wel voldoende steun biedt (buurman ziet dat zwangere vrouw huilend aan de deur staat direct na vermeende mishandeling),⁶ en in het andere geval niet (vriendin hoort meisje huilen aan de telefoon direct nadat zij in haar billen is geknepen).⁷ Het waarom lijkt bij nadere bestudering van de zaken onder andere gelegen in de unieke combinatie van omstandigheden van elke zaak (bijvoorbeeld: naast emotie is er ook nog pijn bij de zwangere vrouw waargenomen).⁸

Wij denken dat de Bayesiaanse redeneerstructuur kan helpen de bewijsminimumrechtspraak beter te begrijpen, namelijk door er een logisch geheel van te maken met een rode draad in plaats van vuistregels die geen vuistregels lijken te zijn. Zodoende hopen wij ook te kunnen bijdragen

⁶ Zie HR 22 april 2014, ECLI:NL:HR:2014:957, *NJ* 2014/328, m.nt. Rozemond.

⁷ Zie HR 2 november 2021, ECLI:NL:HR:2021:1594, *NJ* 2022/32, m.nt. Vellinga.

⁸ In HR 22 april 2014, ECLI:NL:HR:2014:957, *NJ* 2014/328, m.nt. Rozemond was naast de emotie ook nog sprake van een verkrampte houding van de mishandelde vrouw en had zij haar handen op haar buik. Vooral dat laatste herkent iedereen met enige kennis van zwangere vrouwen als een bij zwangere vrouwen passend gebaar – zeker bij pijn aan hun buik. Zie over de casuïstische aard van de rechtspraak en de beperkte waarde van de vuistregels, ook Borgers in zijn annotatie onder HR 7 juli 2015, ECLI:NL:HR:2015:1817, *NJ* 2015/488.

een stevige(r) onderbouwing van bewijsminimumzaken. Wij lichten dit toe.

Rechters (mensen) denken in verhalen. In de bewijsminimumrechtspraak is dat terug te zien in de aard van het (potentiële) steunbewijs. Een acute emotie van een meisje is iets dat mensen herkennen als passend in een verhaal waarin net iets vervelends is gebeurd. Dat wat in de juridische commentaren ‘context’ wordt genoemd – zoals de wijze waarop een verdachte zijn slachtoffer in een zwembad heeft afgezonderd van andere mensen⁹ – is bijvoorbeeld het ‘hoe’ dat standaard onderdeel is van een verhaal met een ‘wie, wat, hoe, waar, wanneer en waarom’. Bayes verklaart waarom binnen een bewijsconstructie zo’n verhaalselement of bewijsgrond geen algemene vuistregel kan zijn voor steunbewijs. De waarde ervan wordt immers bepaald door het unieke verhaal van de aangifte in relatie tot de ontkenning van de verdachte. Bijvoorbeeld: een afzonderingsstrategie tijdens het zwemmen met familie en vrienden past beter bij het verhaal van de aangever dan bij het verhaal van de verdachte. Het gaat er met andere, of Bayesiaanse, woorden om hoe precies en – ten opzichte van elkaar – onderscheidend de hypothesen worden onderbouwd. Bayes laat hiermee ook zien dat het vooral in bewijsminimumzaken belangrijk is dat bewijs – in relatie tot een verhaal en een tegensprekend verhaal – zo scherp mogelijk wordt opgeschreven. Een specifieke ‘hoe’ of een precieze ‘wanneer’ dat gevonden wordt in een tweede bron, kan immers het verschil maken.

Winst is evenwel ook te behalen vanuit het tweede uitgangspunt, de a priori waardering van het scenario. In zaken

⁹ Zie HR 13 februari 2018, ECLI:NL:HR:2018:189, *NJ* 2018/297. Die strategie bleek in deze zaak alleen uit de verklaring van de getuige en niet uit andere bronnen. Zie de annotatie van Rozemond bij *NJ* 2018/298.

met weinig bewijs, waarbij het in veel gevallen gaat om het verhaal van de aangever tegenover de ontkenning of het zwijgen van de verdachte, zal de rechter zich doorgaans de vraag stellen: “Geloof ik de getuige?” In de motiveringen van de bewijsminimumzaken is het antwoord op deze vraag terug te zien in de nadruk die de rechter legt op de betrouwbaarheid van de getuige. De beoordelingsfactoren die in dat verband worden aangedragen – zoals ‘spontane verklaring’, ‘niet onder druk afgelegd’ – bepalen tezamen de betrouwbaarheid van de getuige, of meer Bayesiaans gezegd: bepalen de mate van geloof die aan de verklaring van deze getuige (dit scenario) kan worden gehecht. Deze beoordeling van de getuige is niets anders dan het toepassen van het Bayesiaanse uitgangspunt dat bij het beoordelen van bewijs de context niet moet worden vergeten, ofwel dat de kwaliteit van een scenario ook moet worden beoordeeld los van het bewijs. Juist in zaken waarin het bewijs beperkt is en het steunbewijs niet of nauwelijks discriminerend, zal deze a priori overtuiging een belangrijke rol spelen. Een geloofwaardig verhaal van de getuige, dat het bewijs evengoed verklaart als het weinig geloofwaardige ontkennende verhaal van de verdachte, kan ervoor zorgen dat de drempel voor veroordeling net wordt gehaald.

Deze wisselwerking tussen de prior en de kracht van het bewijs is overigens ook te herkennen in hoe de Hoge Raad de bewijsminimumregel benadert. De onderbouwing van de betrouwbaarheid van de getuigenverklaring lijkt namelijk een rol te mogen spelen bij de beoordeling van het inhoudelijke verband tussen de getuigenverklaring en de tweede onafhankelijke bron. Die relatie lijkt erop neer te komen dat er een omgekeerd evenredig verband bestaat tussen de mate waarin een getuigenverklaring betrouwbaar en dus overtuigend wordt geacht, en de eisen die worden gesteld aan de

‘voldoende steun’ die de tweede bewijsgrond biedt.¹⁰ Vanuit de Bayesiaanse redeneerstructuur is deze wisselwerking volstrekt logisch.

Overigens lijkt de rechtspraak te laten zien dat rechters doorgaans vooral aandacht besteden aan de betrouwbaarheid of geloofwaardigheid van de verklaring van de getuige. Vanuit het Bayesiaanse gezichtspunt dat bewijs relatief moet worden beoordeeld en dat dus (minstens) twee hypothesen nodig zijn, moet daar altijd een beoordeling van de verklaring van de verdachte tegenover worden gezet. In de hoofden van de rechters speelt dit oordeel onvermijdelijk een rol, durven wij te stellen. Wellicht dat rechters denken dat deze informatie niet relevant mag zijn. Iets als een algemeen gedragspatroon is inderdaad niet overduidelijk redengevend bewijs voor het desbetreffende feit. Of misschien wordt gedacht dat een dergelijke motivering een beslissing kwetsbaar maakt, of er is geen tijd om het allemaal goed op te schrijven. Het expliciteren van deze Bayesiaanse logica is niettemin juridisch mogelijk en ook wenselijk, zo menen wij. Zo zou een rechter op grond van Bayes kunnen beredeneren dat de belastende verklaring van een meisje dat geen reden heeft om te liegen, geloofwaardiger (of: betrouwbaarder) wordt bevonden dan een ontkenning van een verdachte van exhibitionisme die aangeeft langdurig last te hebben gehad van neigingen inzake exhibitionisme, en dat tegen die achtergrond het bewijs voldoende wordt geacht voor een veroordeling.

Om misverstanden te voorkomen: ons betoog hier houdt niet in dat rechters hun vooroordelen of onderbuikgevoelens maar moeten gaan opschrijven om een veroordeling steviger te maken, we schreven hier al eerder over in dit hoofdstuk.

¹⁰ Corstens/Borgers/Kooijmans 2021, p. 854-855.

Ook betogen we niet dat dat er soepeler gekeken zou moeten worden naar de redengevendheid van bewijs. We denken wel dat met een klein beetje moeite de blijkens de rechtspraak reeds aanwezige Bayesiaanse redeneerstructuur bewuster, vollediger en explicieter kan worden toegepast door de feitenrechter in bewijsminimumzaken. Op die manier kan zowel het uitgangspunt ‘bewijs moet relatief worden beoordeeld’ als het uitgangspunt ‘een mens begint niet bij nul’ een – controleerbare – rode draad vormen bij het invullen van het vereiste van ‘voldoende steun’.

HOOFDSTUK 5

Misverstanden en problemen

Over het gebruik van het Bayesiaans denkmodel doen nogal wat misverstanden de ronde. We bespraken er al een aantal in de eerdere hoofdstukken, impliciet en expliciet, maar we denken dat het nuttig is om hier nog wat meer aandacht aan te besteden. In dit hoofdstuk bespreken we een aantal misverstanden en mogelijke problemen. Hopelijk helpt dit om eventuele koudwatervrees voor Bayesiaans denken weg te nemen. Achtereenvolgens bespreken we:

- de incorrecte gedachte dat correct toepassen van Bayes iets met de stand van de wetenschap te maken heeft;
- de mythe dat er maar één likelihood ratio is in een zaak;
- de incorrecte gedachte dat een zeldzame gebeurtenis automatisch verdacht is;
- de incorrecte gedachte dat een volledig DNA-profiel altijd beslissend is;
- de incorrecte gedachte dat een kansuitspraak omgezet dient te worden in een categorische uitspraak;
- de *prosecutor's fallacy*;
- verschillende visies van wat Bayesiaans redeneren eigenlijk is.

Voor de filosofisch geïnteresseerde liefhebbers sluiten we dit hoofdstuk af met een korte beschouwing over de vraag in hoeverre onze menselijke onzekerheden überhaupt beschreven kunnen worden met kansen.

DE STAND VAN DE WETENSCHAP EN BAYES

De vraag naar de ‘wetenschappelijkheid’ van Bayesiaans denken wordt regelmatig gesteld. Het hof Den Haag schreef bijvoorbeeld het volgende over de Bayesiaanse methode:¹

“[...] dat 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.”

Wat kan het hof hiermee bedoelen? Het citaat lijkt in tegenspraak te zijn met bijna alles wat we in dit boek hebben geschreven; we hebben immers beargumenteerd dat de methode uiterst nuttig en toepasbaar is. Is het succesvol toepassen van Bayesiaans redeneren werkelijk afhankelijk van de huidige stand van de wetenschap? Wat betekent dat eigenlijk precies, het ‘wetenschappelijk’ zijn van een bepaalde methode, gedachte of model?

Bayesiaanse statistiek is binnen de toegepaste statistiek een erkend paradigma dat op enkele belangrijke punten afwijkt van de meer ‘klassieke’ statistiek. Binnen het klassieke paradigma verwerpt men een bepaalde hypothese zodra de data die gevonden zijn, voldoende onwaarschijnlijk zijn onder die hypothese. We zagen echter al bij de bespreking van de zaak Sally Clark, dat dit paradigma binnen het recht niet goed functioneert. De kans om, uitgaande van de onschuld van Sally Clark, twee gestorven baby’s te zien, was natuurlijk erg klein, maar we hebben gezien dat dit helemaal geen reden was om de onschuldhypothese te verwerpen. In het recht is het zaak om te besluiten welk van verschillende hypothesen de beste verklaring biedt voor de geobserveerde

¹ ECLI:NL:GHDHA:2015:2860.

data. Dat die data een kleine kans hebben om op te treden, onder welke hypothese dan ook, is daarbij niet zo relevant. De data zijn er nu eenmaal, en nu moeten we die verklaren.

De gedachte is, zoals we uitgebreid hebben besproken, dat we gaan kijken in hoeverre de data onderscheidend zijn voor de relevante hypothesen, en in hoeverre de data onze mening over die hypothesen dus eventueel verandert. Om een mening te veranderen moet je echter al wel een mening hebben.

Dit statistische Bayesiaanse paradigma is, zoals gezegd, een gevestigd onderdeel van de statistiek, en geen mens zou het in zijn hoofd halen om dit onwetenschappelijk te vinden. Hoe kwam het hof Den Haag dan tot de gedachte dat de betrouwbaarheid van ‘Bayes’ “naar de huidige stand van de wetenschap te onzeker is”?

Om dit te begrijpen is meer context van deze specifieke zaak nodig. Klimaatfysicus Frans Alkemade had in deze zaak een ‘doorrekening’ gemaakt zoals hij dat ook had gedaan in de Anne Frank-zaak waar we in hoofdstuk 3 over schreven. En net zoals in de Anne Frank-zaak kwam Alkemade tot een bepaalde conclusie over de kans dat een verdachte schuldig is. We hebben uitgebreid besproken dat zo’n doorrekening van een zaak een tamelijk nutteloze en betekenisloze exercitie is, en dat de uitkomst slechts iets kan zeggen over hoe de rekenaar er zelf tegenaan kijkt. Die argumenten gaan we hier niet herhalen, maar uit het volgende citaat uit dezelfde uitspraak maken we op dat het hof dit tot op zekere hoogte goed begrijpt:

“Het hof heeft echter op basis van het onderzoek ter terechtzitting in herziening geconstateerd dat er wetenschappelijk onvoldoende overeenstemming bestaat over de vraag of de door [deskundige1] gehanteerde methode van (klassieke) Bayesiaanse analyse betrouwbaar kan worden

toegepast op en in complexe strafzaken als de onderhavige, waarin sprake is van vele, en bovendien ten opzichte van elkaar niet immer als onafhankelijk te beschouwen, bevindingen. In dit kader heeft het hof mede in zijn beoordeling betrokken dat, blijkens zijn rapport en toelichting ter terechtzitting, [deskundige 1] niet de beschikking had over het gehele dossier, doch slechts over een hem door de advocaat-generaal ter beschikking gestelde selectie daaruit. Er blijken bovendien aanzienlijke verschillen van inzicht tussen deskundigen te bestaan over welke (en op welke wijze) bevindingen in complexe strafzaken zaken in de analyse dienen te worden betrokken, hoe de (*a priori*) kansen van de bevindingen (betrouwbaar) kunnen worden vastgesteld dan wel ingeschat en welke expertise voor die vaststelling dan wel inschatting noodzakelijk is.”

Het hof is zich goed bewust van het feit dat alle berekeningen subjectief zijn en dat een besluit om bepaalde informatie wel of niet mee te nemen wordt genomen op basis van persoonlijke inschattingen. Het hof begrijpt, met andere woorden, dat zo’n berekening niet zomaar gevolgd kan worden.

Het hof lijkt echter van mening dat de huidige berekening van Alkemade dan weliswaar niet betrouwbaar is, maar dat er *in principe* wel degelijk een betrouwbare berekening mogelijk zou kunnen zijn, als de deskundigen het maar eens zouden worden over hoe bevindingen betrokken moeten worden in de analyse en als de (*a priori*) kansen maar “betrouwbaar” zouden kunnen worden vastgesteld. Dat lijkt de reden te zijn dat het hof in het eerste citaat beweert dat het “de huidige stand van de wetenschap” is die toepassing van Bayesiaans denken op dit moment zou verhinderen.

We hebben echter in hoofdstuk 3 gezien dat de kansen die iemand aan bepaalde scenario’s, hypotheses of gebeur-

tenissen toekent, een kwestie is van informatie. Als mensen andere informatie hebben, dan worden hun kansinschattingen ook anders. Deze kansen, immers, zijn gebaseerd op de kennis die je hebt. Verschillende inschattingen zijn voor de verschillende personen die ze opschrijven allemaal correct, omdat ze elk de informatiepositie van deze persoon weerspiegelen.

Wat betreft de selectie van het bewijsmateriaal: ook daarvoor geldt dat er geen algoritme kan bestaan dat besluit welke gegevens wel en welke niet meegenomen worden in de overwegingen van de rechtbank of het hof. Dat zijn namelijk geen objectieve keuzes; de wetenschap heeft daar weinig tot niets over te zeggen. Dat betekent dat het niet van de stand van de wetenschap afhangt in hoeverre een Bayesiaanse benadering van een zaak adequaat is of niet. De Bayesiaanse methode geeft een algemeen erkende manier om met informatie om te gaan; de rol van de wetenschap is daarin beperkt. Er komt dus helemaal geen moment dat de wetenschap voldoende ontwikkeld is om dit wél te kunnen doen.

Natuurlijk kan de wetenschap soms wel iets zeggen over de inschatting van kansen, maar dat hangt nogal van de context en het type informatie af. Als de progressie in de wetenschap er bijvoorbeeld voor zorgt dat we nu bij DNA de allelen op meer loci kunnen bepalen dan vroeger, dan betekent dat echter niet dat de inschattingen van vroeger onjuist waren. Ze waren niet onjuist, maar ze waren gebaseerd op minder informatie dan we nu tot onze beschikking hebben. Het hele bouwwerk van een Bayesiaanse analyse kon vroeger, met minder loci en dus met minder informatie, uitstekend gebruikt worden, en het was zeker nooit onwetenschappelijk om dat te doen. Het enige verschil is dat we nu meer informatie tot onze beschikking hebben. Het zou het onethisch zijn deze niet te gebruiken.

De stand van de wetenschap gaat daarom ook het probleem van elkaar tegensprekende deskundigen niet oplossen. Het is nu eenmaal zo dat mensen verschillend over dingen denken, ook deskundigen; vooral deskundigen. Dit feit hebben we maar te accepteren. Rechters kunnen we proberen zo op te leiden dat ze in staat zijn om op basis van de juiste vragen beargumenteerd te besluiten welke deskundige ze wel en welke ze niet willen volgen. Daar heeft de stand van de wetenschap niets mee te maken.

DE MYTHE DAT ER MAAR ÉÉN LIKELIHOOD RATIO IN EEN ZAAK BESTAAT

Een likelihood ratio vertelt ons bij welke van twee hypothesen het bewijs het beste past, en ook hoeveel keer beter. Een likelihood ratio van 100 vertelt ons dat de data 100 keer beter bij de eerste dan bij de tweede hypothese passen. We moeten echter niet vergeten dat zo'n likelihood ratio van een aantal dingen afhangt, namelijk van:

1. de data;
2. de keuze van de hypothesen;
3. het model waarbinnen berekeningen worden gedaan.

Dat betekent dat je niet zomaar kunt spreken over dé likelihood ratio, en dat er ook in een en dezelfde zaak sprake kan zijn van heel verschillende likelihood ratio's, die elk hun eigen zeggingskracht hebben. Het gebruik van een likelihood ratio vereist dus sowieso een goed inzicht in de betekenis ervan en in de manier waarop ze worden bepaald. We illustreren dit met een voorbeeld waarin daadwerkelijk gerekend kan worden.

In een wetenschappelijk artikel uit 2013 wordt een DNA-mengprofiel besproken dat van zowel twee als van drie personen afkomstig kan zijn.² De hypothesen die we kunnen opstellen gaan dan niet alleen over de vraag wie heeft bijgedragen aan het mengsel, maar impliciet ook over de vraag of het een twee- of driepersoonsmengsel is. De vraag waar het om draait is of er aanwijzingen zijn dat de verdachte V heeft bijgedragen aan het mengsel. Om dit te onderzoeken kunnen we verschillende hypothesen opstellen en ook verschillende modellen waaronder we rekenen. In de expositie die nu volgt, verwijzen we naar tabel 1, waarin we de uitkomsten van de berekeningen hebben samengevat. De details van de zaak en van de berekening doen er hier niet toe.

Tabel 1 Verschillende hypothesen en bijbehorende likeliheid ratio's. Zie de tekst voor verdere uitleg.

H_1	H_2	likeliheid ratio
S + V + O	S + O + O	0,63
S + V	S + O	4,17
V + O	O + O	0,0000016

Naast de verdachte V hebben we hier ook te maken met het slachtoffer S, en is het van belang om te kunnen spreken over een onbekende persoon O. In de eerste regel van de tabel beschouwen we de hypothesen dat het DNA-mengprofiel ofwel afkomstig is van slachtoffer, verdachte en een onbekend persoon, of van het slachtoffer en twee onbekende personen. De eerste regel gaat over hypothesen waarbij het mengsel van drie personen afkomstig is, de

² H. Haned, G. Dørum, T. Egeland en P. Gill, "On the Meaning of the Likelihood Ratio: Is a Large Number Always an Indication of Strength of Evidence?", *Forensic Sci. Int.: Genet. Suppl. Ser.* 4(1) (2013): e176–e177, <http://dx.doi.org/10.1016/j.fsigs.2013.10.091>.

andere twee gaan over hypothesen waarbij het mengsel van twee personen afkomstig is.

Het gebruikte rekenmodel is voor ons niet echt van belang, maar voor de volledigheid vermelden we dat het afkomstig is van het forensim-softwarepakket.³ In dat model worden zoals in elk model keuzes gemaakt. Zo is er een kans dat een allel dat wel in het fysieke mengsel zit, toch niet wordt opgemerkt. Andersom is het ook mogelijk dat er een allel opduikt dat er in werkelijkheid niet is.

Misschien wel het meest in het oog springende feit zien we als we de eerste en de tweede regel met elkaar vergelijken. Het enige verschil tussen de vergeleken hypothesen is dat in regel 1 uitgegaan wordt van een driepersoonsmengsel, terwijl in regel 2 een tweepersoonsmengsel wordt aangenomen. Het verschil in de gerapporteerde likelihood ratio is opmerkelijk: in regel 1 is het bewijs ontlastend voor de verdachte V (want de likelihood ratio is kleiner dan 1 en dus is het bewijs onder de tweede hypothese waarschijnlijker dan onder de eerste) terwijl in regel 2 het bewijs juist belastend is. Er is dus helemaal geen sprake van dé likelihood ratio, want de keuze van de hypothesen heeft sterke invloed op de numerieke waarde ervan.

Dat is allemaal best lastig om te begrijpen, hoewel de verschillen te verklaren zijn als je je gaat verdiepen in de populatiefrequenties van de waargenomen allelen. Maar het punt dat we willen maken is hopelijk duidelijk: zelfs in situaties waarin een niet-controversieel rekenmodel wordt gebruikt, is het mogelijk dat verschillende paren hypothesen verschillende resultaten opleveren wat betreft de waarden van de bijbehorende likelihood ratio. De keuze voor de scenario's moet uit de context van de specifieke zaak komen.

³ <https://forensim.r-forge.r-project.org/>.

Natuurlijk is dit vervelend voor een rechter. Het zou veel eenvoudiger zijn als dit fenomeen niet zou bestaan en alles eenduidig zou zijn. Advocaten maken ook graag gebruik van de mogelijke verwarring die dit oproept, en dus is het zaak dat rechters precies weten hoe een en ander zit.

Het is dus zo dat ‘de’ likelihood ratio niet bestaat. Een inschatting daarvan is altijd afhankelijk van keuzes. Dit betekent andermaal – en nu op een iets andere manier gemotiveerd – dat er niet zoiets bestaat als een zinvolle doorrekening van een zaak. De gegevens in samenhang bezien is nog niet zo eenvoudig, maar het is zeker niet algoritmisch: alles wat erbij komt kijken moet worden beoordeeld door rechters of raadsheren. Het Bayesiaanse denkmodel kan hierbij zeker helpen, op de manieren zoals we dat hebben besproken, maar zonder kennis van zaken is toepassing van datzelfde denkmodel niet goed mogelijk.

EEN KLEINE KANS ZEGT WEINIG

Stel dat Lonneke een loterij wint. Ronald is jaloers en probeert Lonneke verdacht te maken. Hij beweert dat de kans dat Lonneke zonder frauduleuze handelingen de loterij wint, kleiner is dan 1 op de miljoen. Omdat deze kans zo klein is, gelooft hij niet meer in de hypothese dat Lonneke de loterij op een eerlijke manier heeft gewonnen.

De meeste lezers zullen wel in de gaten hebben dat er iets vreemds is aan deze redenering. Je kunt een dergelijke berekening immers altijd maken, bij elke winnaar van een loterij. De kans dat de winnaar zou winnen is altijd klein. Toch gebeurt het zelden dat een winnaar verdachte wordt van fraude. Hoe zit dat precies? Bedenk dat een berekening als

deze de reden was dat Sally Clark en Lucia de Berk vervolgd werden.

Het feit dat een gebeurtenis met een kleine kans toch is opgetreden, is niet interessant en ook niet onverwacht. Elke winnaar van een loterij is onverwacht. Het probleem is dat de kleine kans een zogenaamde voorspellende kans is. Je berekent de kans dat precies deze persoon in een loterij zal gaan winnen. Na de trekking is er altijd een winnaar, en alle aandacht gaat dan naar hem of haar uit. Het feit dat deze persoon van tevoren een kleine kans had om te winnen, is niet meer relevant als hij eenmaal gewonnen heeft. De relevante vraag is nu: gegeven dat deze persoon heeft gewonnen, welke hypothese over de reden ervan is het meest aannemelijk? Met andere woorden, als een feit zich eenmaal heeft voorgedaan, dan is de juiste vraag om te stellen wie of wat voor dat feit verantwoordelijk was. In dit geval zou je bijvoorbeeld kunnen kijken naar twee mogelijke hypothesen:

H_1 : Lonneke heeft eerlijk aan de loterij meegedaan;

H_2 : Lonneke heeft de trekking op de een of andere manier gemanipuleerd.

Onder H_1 is de kans dat Lonneke een hoofdprijs wint heel klein, laten we zeggen 1 op de miljoen. Onder H_2 is de kans veel groter. Laten we aannemen dat onder deze hypothese Lonneke een uitstekende manipulator is, en dat ze door haar manipulatie zeker wint. De kans op winst is dan dus 1, of 100%. Dit betekent dat de likelihood ratio ten gunste van H_2 erg groot is: deze is 1 miljoen. De data (het feit dat Lonneke de loterij wint) is onder H_2 miljoen keer waarschijnlijker dan onder H_1 .

We moeten echter niet vergeten dat a priori, dus voordat we naar de uitslag van de loterij gingen kijken, hypothese H_1

veel aannemelijker was dan hypothese H_2 . De meeste mensen manipuleren geen loterijen. Dat betekent dat we extreem veel bewijs nodig hebben om hypothese H_2 te gaan geloven. Hoeveel bewijs hangt af van de kans die we oorspronkelijk aan H_2 toekennen. Een klein rekenvoorbeeld kan dit verduidelijken, we verwijzen naar tabel 2.

Tabel 2 Een rekenvoorbeeld bij de loterij.

	winst in loterij	geen winst in loterij	totaal
manipulanten	1	0	1
eerlijke mensen	1	999.999	1.000.000
totaal	2	999.999	1.000.001

Stel je voor dat we een populatie hebben van 1.000.001 mensen. Stel je bovendien voor (om het rekenwerk makkelijker te maken) dat bij eerlijk spel elke deelnemer een kans van 1 op de miljoen heeft om te winnen en dat er meerdere (of geen) winnaars kunnen zijn. We denken dat er maar één manipulant is, de rest van de mensen is eerlijk. Onder de miljoen eerlijke mensen verwachten we dan één winnaar, en bovendien verwachten we dat een manipulator altijd wint. Als we dus van tevoren denken dat er één manipulator is, dan verwachten we twee winnaars. De kans dat een gegeven winnaar gemanipuleerd heeft is dan 1 op 2, ofwel 50%.

Misschien vindt de lezer dat best een grote kans. Maar bedenk dat dit het gevolg is van onze aanname dat ongeveer 1 op de miljoen mensen loterijen manipuleert, en de aanname dat meerdere mensen kunnen winnen. Als we dat getal (1 op de miljoen) kleiner maken, dan wordt de kans dat een gegeven winnaar gemanipuleerd heeft, automatisch ook kleiner, zoals je makkelijk kunt narekenen met een tabel zoals hierboven.

Het optreden van een gebeurtenis met een kleine kans is dus in principe geen reden tot zorg of een reden om een nader onderzoek in te stellen. Maar dit feit wordt vaak vergeeten, en als je even niet oplet gaat het fout. De meeste mensen herinneren zich wel dat eind 2012 een grensrechter bij het amateurvoetbal aangevallen en tegen zijn hoofd geschopt werd, de volgende dag onwel werd en overleed. Hij was een gezonde veertiger. Natuurlijk ging het bij deze zaak om de vraag of het overlijden van de grensrechter (redelijkerwijs) toegeschreven kon worden aan de aanval. De verdediging bracht het volgende argument in. Slechts ongeveer 1 op de 10.000 vechtpartijen zonder wapens heeft een overlijden tot gevolg. Daarom is het niet aannemelijk dat het overlijden van de grensrechter het gevolg is van de vechtpartij.

De fout die hierbij werd gemaakt, is dat men zich niet realiseerde dat deze 1 op 10.000 een voorspellende kans is. Waar het werkelijk om gaat is dat er iemand is overleden en we willen weten wie of wat daarvoor verantwoordelijk was. Het feit dat een vechtpartij meestal goed afloopt, is niet meer relevant zodra er na de vechtpartij iemand overlijdt.

De fout die we hier bespreken komt vrij algemeen voor. Het gaat in het recht niet over voorspellende kansen. Kansen in het recht kijken altijd terug: er is iets gebeurd en we willen weten wat de oorzaak of reden daarvoor was. Het feit dat de gebeurtenis in kwestie zeldzaam is, is bij die overwegingen niet relevant.

IS EEN AANGETROFFEN DNA-PROFIEL BESLISSEND?

De technologie om vanuit biologische sporen een DNA-profiel te maken is inmiddels dermate ontwikkeld dat je rustig kunt zeggen dat een volledig DNA-profiel op bronniveau

uniek is, afgezien van eenenige tweelingen. Dat roept de vraag op of een aangetroffen (volledig) profiel altijd beslissend bewijs is in een zaak.

Het antwoord op die vraag is doorgaans ‘nee’. Om dit punt te maken kijken we nog eens naar de zaak Nicky Verstappen.⁴ Het op het lichaam en kleding van het slachtoffer gevonden DNA was van de verdachte Jos B., en dit werd door de verdediging ook niet betwist. De vraag naar de bron van het DNA werd in het proces echter vervangen door de vraag naar hoe het DNA daar terecht was gekomen. Omdat wij allemaal behoorlijk wat DNA achterlaten op plekken waar we zijn geweest, is ons DNA op veel plekken natuurlijk-kerwijs en totaal onschuldig aanwezig.

Het gaat dus bij het evalueren van een (volledig) DNA-profiel vaak om verschillende hypothesen die moeten verklaren hoe het DNA op die specifieke plaats terecht is gekomen, en niet zozeer om de bron ervan. In forensisch jargon zeggen we dat we (zoals bij de zaak Nicky Verstappen) geen hypothesen op bronniveau beschouwen, maar op activiteiten-niveau. Dat maakt een analyse veel lastiger. De genetische modellen die ons informatie kunnen geven over de bewijswaarde van een bepaald DNA-profiel, zeggen niets over dat activiteiten-niveau, en doorgaans kan er op dat niveau niet zinvol gerekend worden – dat zagen we al verschillende malen in dit boek.

Een rechter moet tot een overtuiging komen voordat hij iemand kan veroordelen, en deze overtuiging wordt opgebouwd door middel van oorspronkelijke achtergrondkennis en context, samen met de bewijsmiddelen in een zaak. Deze bewijsmiddelen hebben een bepaalde bewijswaarde die meestal niet maar soms wel in getallen kan worden uitge-

⁴ ECLI:NL:HR:2023:1303 in hoger beroep.

drukt. Bij DNA-bewijs zijn vaak zinvolle getallen mogelijk, maar alleen op bronniveau. De vraag of DNA-bewijs voldoende is in zichzelf, is dus eigenlijk geen zinvolle vraag omdat het DNA op bronniveau, hoe sterk ook, altijd maar een (klein) stukje van de puzzel geeft. De vraag voor een rechter blijft simpelweg of hij op grond van alle informatie ervan overtuigd is dat de verdachte het ten laste gelegde heeft begaan. Het antwoord op die vraag is niet of nauwelijks in getallen uit te drukken is, zoals we ook bespraken in hoofdstuk 3.

MOET EEN KANSUITSPRAAK IN EEN CATEGORISCHE STELLING WORDEN OMGEZET?

Raadsheer De Groot stelde het volgende:⁵

“Rechters redeneren echter in vaststaande feiten. Zij kunnen pas aan het werk met een waarschijnlijkheidsuitspraak als ze deze eerst in een categorische stelling hebben omgezet. Dat zou met het oog op de controleerbaarheid van de rechterlijke redenering eigenlijk moeten gebeuren volgens een vooraf bepaalde sleutel, een grenswaarde: bijvoorbeeld een herkomstbepaling is niet meer mogelijk als de frequentie van het voorkomen van het DNA-profiel groter is dan 1:10.000. Zover is de discussie over de ratio-nale bewijswaardering in kringen van strafrechtbeoefenaren echter nog lang niet.”

Wat zegt De Groot hier precies en waarom zijn wij het daar niet mee eens? Hij stelt (terecht) dat herkomstbepaling niet

⁵ J. de Groot, *DNA in het strafrecht: Hoe moet het niet? Hoe moet het wel?*, Zutphen: Paris 2018, p. 151.

mogelijk is als een bepaald DNA-profiel maar vaak genoeg voorkomt. Maar dat doet niets af aan het feit dat een profiel dat bij 1 op de 10.000 mensen voorkomt, een bepaalde bewijskracht heeft. Er is geen reden om een ‘grenswaarde’ in te stellen: ook als het profiel bij 20% van de mensen voor zou komen, dan nog steeds heeft het precies die bewijswaarde die bij die frequentie past, en het zou onjuist zijn om die bewijswaarde te negeren. Er is, met andere woorden, geen reden om een waarschijnlijkheidsuitspraak om te zetten in een categorische stelling. Sterker nog, het zou onjuist zijn om dat te doen, omdat de gegevens op die manier incorrect worden gebruikt.

Uitspraken vanuit de forensische wetenschap zijn kanstheoretische en statistische uitspraken, uitgedrukt in bepaalde bewoordingen van onzekerheid. DNA speelt hierbij geen bijzondere rol. Natuurlijk is het zo dat DNA-bewijs vaak grote bewijskracht heeft op bronniveau, maar als dat onverhoopt niet zo is (bijvoorbeeld vanwege een slecht profiel), dan is de beperkte bewijskracht gewoon ook relevant.

Vanuit forensisch perspectief is de grootte van de likelihood ratio belangrijk en niet het bewijsmiddel waar deze op is gebaseerd. Een voorbeeld kan dit punt wellicht nog wat verhelderen. Vingerafdrukken zijn doorgaans minder onderscheidend dan een DNA-profiel. Met de gangbare methoden zal een volledige vingerafdruk nooit leiden tot een likelihood ratio van meer dan enkele duizenden. Voor een DNA-profiel echter, kan de likelihood ratio in de miljarden lopen. Een profiel dat tot een likelihood ratio van slechts enkele duizenden leidt, kan niet anders dan een slecht profiel zijn. Maar het punt is dat twee bevindingen met dezelfde likelihood ratio even onderscheidend zijn: een volledige vingerafdruk heeft dus dezelfde bewijskracht als dat slechte DNA-profiel. Dat heeft met de herkomst van het profiel of van de vingerafdruk eigenlijk niet zoveel te maken.

Ten slotte merken we op dat De Groot kennelijk a posteriori uitspraken wil doen aan de hand van een bepaalde DNA-match; hij wil immers de herkomstvraag relateren aan het DNA-materiaal. Daarbij lijkt hij zich schuldig te maken aan de *prosecutor's fallacy* die we hierna bespreken.

DE PROSECUTOR'S FALLACY

Waarschijnlijk is de meest gemaakt fout bij het redeneren onder onzekerheid de zogenaamde *prosecutor's fallacy*. De fout bestaat hierin dat een likelihood ratio wordt aangezien voor een kansuitspraak over de hypothesen zelf. We illustreren dit met een casus.⁶ Vergelijk de volgende twee passages uit de uitspraak:

“Door het NFI is een vergelijkend onderzoek tussen proefafdrukken gemaakt met de linkerschoen van [verdachte] en de in bloed gestempelde schoensporen uitgevoerd. De conclusie van het NFI luidt dat het zeer veel waarschijnlijker is wanneer deze schoensporen veroorzaakt zijn met de linkerschoen van [verdachte] dan wanneer ze veroorzaakt zijn door een willekeurige andere linkerschoen van verdachten van misdrijven in Nederland. Deze onderzoeksresultaten ondersteunen de verklaring van [medeverdachte] over de aanwezigheid en betrokkenheid van [verdachte].

[...]

Zoals hiervoor overwogen over de jas en de kabeltrui van [verdachte] heeft hij zich gedurende de avond omgekleed, hetgeen van invloed kan zijn geweest op het aangetrof-

⁶ ECLI:NL:RBZWB:2015:3291.

fen (bloed)sporenbeeld. Bovendien wordt in de epicrise,⁷ waarin het aangetroffen sporenbeeld op de kleding van [verdachte] is betrokken, geconcludeerd dat het scenario van het openbaar ministerie inhoudende dat [verdachte] betrokken was bij de dood van [slachtoffer] (veel) waarschijnlijker is dan het scenario van [verdachte] dat hij daar niet betrokken bij was.”

De details van de zaak zijn voor ons nu niet erg van belang, de citaten spreken voor zich en kunnen op zichzelf worden beoordeeld. Het eerste deel van het citaat is logisch gesproken helemaal correct. Er wordt gekeken naar welke hypothese het beste past bij het bewijs, in dit geval een schoenspoor. Het gevonden spoor past veel beter bij de hypothese dat het is veroorzaakt door de verdachte dan bij de hypothese dat ze door een onbekend persoon is achtergelaten. Wat wel bevreemdt is dat de rechtbank spreekt over “verdachten van misdrijven in Nederland”; dat moet wel haast een verschrijving zijn en wordt hiermee gewoon een willekeurige andere linkerschoen bedoeld.

Het tweede deel van het citaat is echter problematisch. Het gaat hier kennelijk om een ander bewijsmiddel, namelijk bloedsporen op kleding. Waarschijnlijk is het de bedoeling om ook hier te zeggen dat het bloedspoor beter bij de eerste dan bij de tweede hypothese past. In plaats daarvan wordt er plotseling gesproken over de waarschijnlijkheid van de hypothesen zelf. Als dat gebeurt spreken we van een prosecutor’s fallacy. We hebben gezien dat je op basis van de bewijsmiddelen nooit zomaar een uitspraak kunt doen over de waarschijnlijkheid van de hypothesen zelf, je moet immers altijd de context in overweging nemen. Zelfs bij bewijs dat sterk in de richting van bijvoorbeeld de eerste hypothese wijst, is het

⁷ Beschouwing, of eindoordeel.

helemaal niet gezegd dat die hypothese daarmee waarschijnlijk wordt. Dat hangt ervan af wat men van die hypothese vond voorafgaand aan het bewijs.

In onze ervaring is het voor veel juristen heel moeilijk om deze prosecutor's fallacy te vermijden. Waarschijnlijk is de enige manier om hier bedreven in te worden herhaling: als je het vaak genoeg gezien of gehoord hebt, dan lukt het uiteindelijk wel om er een soort zintuig voor te ontwikkelen dat waarschuwt wanneer de fallacy dreigt.

DRIE VISIES OP WAT BAYESIAANS REDENEREN IS

We hebben in dit boek al verschillende keren erkend dat er wat betreft de Bayesiaanse methode sprake kan zijn van koudwatervrees. Voor een aantal mensen lijkt die methode vooral te gaan over getallen en statistiek, en dan kan terecht de vrees ontstaan dat rechtspreken overgenomen wordt door kansrekening. Hopelijk hebben wij die vrees enigszins weg kunnen nemen. Het lijkt ons goed om in dit hoofdstuk over misverstanden de verschillende visies of interpretaties van wat Bayesiaans redeneren is, op een rijtje te zetten. Ruwweg onderscheiden we de volgende visies:

1. Bayesiaans redeneren is het rapporten van de bewijskracht van bewijsmateriaal via een likelihood ratio – numeriek als dat kan en anders verbaal.
2. Bayesiaans redeneren staat voor het doorrekenen van een casus om tot een uiteindelijke kans op schuld te komen.
3. Bayesiaans redeneren is niets anders dan een bepaalde structuur aanbrengen in iemands redeneerschema, en daarbij goed onderscheid te maken tussen context en zaaksspecifieke aspecten.

De koudwatervrees waar we net over schreven betreft vooral visie 2: de gedachte dat je een zaak kunt terugbrengen tot een rekenexercitie, waarbij terecht de vraag wordt gesteld of dat eigenlijk wel zinvol is. We hebben uitgebreid beargumenteerd dat deze visie op Bayesiaans redeneren tekortschiet. Het doorrekenen van een zaak is geen verstandig idee.

Wat visie 1 betreft hebben we in hoofdstuk 2 gezien dat volgens de richtlijnen van het ENFSI, bewijsmiddelen altijd relatief beoordeeld moeten worden, en dat er maar één manier is om de sterkte ervan uit te drukken: via de likelihood ratio. Als dit niet kan, om welke reden dan ook, dan is de kracht van het bewijs niet (numeriek of verbaal) uit te drukken. Dit betekent dat er altijd, zonder enige uitzondering, naar minimaal twee elkaar beconcurrerende hypothesen moet worden gekeken. Bewijskracht heeft nu eenmaal geen objectieve op zichzelf staande waarde, maar hangt af van welke hypothesen er überhaupt beschouwd worden.

Dit betekent dat zelfs de grootste criticus van Bayesiaans redeneren niet om de likelihood ratio heen kan. De gedachte hierachter is, zoals we hebben gezien, dat we in het recht helemaal niet geïnteresseerd zijn in de vraag of een bepaalde gebeurtenis zeldzaam is of niet. Dat doet er helemaal niet toe. Wat er wel toe doet is de vraag, als een gebeurtenis zich eenmaal heeft voorgedaan, wie of wat daar de oorzaak van was. Bij het onderzoek naar het antwoord op deze vraag zijn we aangewezen op de likelihood ratio. Dit betekent dat we minimaal visie 1 dienen te omarmen.

Visie 3 gaat verder, en legt uit hoe de onontkoombare likelihood ratio uit visie 1 ingebed moet of kan worden in het grotere geheel. We hebben beargumenteerd dat dit ‘grotere plaatje’ nodig is om de likelihood ratio op een goede manier in te kunnen zetten. We hebben in hoofdstuk 4 ook laten

zien dat deze inbedding van Bayesiaans redeneren heel goed is in te passen in de juridische omgeving waarin we hier werken.

Samengevat, het afwijzen van Bayesiaans redeneren gebeurt doorgaans op onterechte gronden, omdat Bayesiaans redeneren abusievelijk geïnterpreteerd wordt als het doorrekenen van een zaak, en het zich uitleveren aan statistiek.

DE AARD VAN ONZE ONZEKERHEDEN

De slotparagraaf van dit hoofdstuk is enigszins filosofisch van aard, omdat we een probleem adresseren dat in de praktijk moeilijk op te lossen zal zijn. Het gaat om de ‘natuur’ van onzekerheid die wij mensen kunnen hebben, en de beperkingen die een wiskundige beschrijving ervan heeft. Ons startpunt is het volgende citaat uit een uitspraak:⁸

“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.*”

⁸ ECLI:NL:GHSE:2018:421, onze nadruk.

Hoewel het ons hier primair om de laatste zin gaat, is de context waarin het hof dit schrijft wel van belang. Het hof reageert met dit citaat op een rapport van filosoof Ton Derksen die voor de onderhavige zaak een Bayesiaanse ‘doorrekening’ had gemaakt. In deze doorrekening rekende Derksen allerlei kansen uit, en daar ging het hof niet in mee.

Maar wat kan het hof met die laatste zin bedoelen? Om dat te begrijpen moeten we even terug naar de basis: het begrijpen en verwerken van onzekerheid in een zaak. We zijn er tot nu toe van uitgegaan dat we deze onzekerheid in principe kunnen beschrijven met kansrekening. Ongeacht de manier waarop wij kansen interpreteren, moeten deze wel aan bepaalde redelijke en logische regels voldoen. Wij hebben daar in dit boek tot nog toe weinig aandacht aan besteed – dat was niet nodig – maar dat wordt nu anders. Wat zijn die ‘redelijke en logische regels van de kansrekening? Welnu, dat zijn er slechts twee:

1. de totale kans moet 100% zijn;
2. kansen op gebeurtenissen die elkaar uitsluiten, mogen we bij elkaar optellen.

Ter illustratie: als we een dobbelsteen gooien dan is de kans op een even uitkomst $1/2$, en de kans op een 3 is $1/6$. De kans op een uitkomst in de verzameling $\{2, 3, 4, 6\}$ is dan natuurlijk $1/2$ (de kans op een even uitkomst) + $1/6$ (de kans op een 3) = $2/3$. Natuurlijk hadden we dat ook anders uit kunnen rekenen: elke uitkomst heeft kans $1/6$, dus de kans op $\{2, 3, 4, 6\}$ is dan $4 \times 1/6 = 2/3$.

Op dezelfde manier is de kans dat de verdachte schuldig is plus de kans dat hij onschuldig is gelijk aan 100%, of 1 als je liever in getallen dan in percentages denkt.

Deze kans wordt ook wel beschreven als de ‘mate van geloof’, van een rechter, in een bepaalde hypothese. Een groot persoonlijk geloof van een rechter leidt niet noodzakelijk tot een veroordeling als dat geloof niet door toelaatbare bewijsmiddelen wordt geschraagd. Daarom is er wel voorgesteld om ‘mate van geloof’ te vervangen door ‘mate van bewijsbaarheid’.⁹ Natuurlijk heeft ‘bewijsbaarheid’ nog steeds dezelfde subjectiviteit als geloof, maar in ieder geval wordt met dat woord wel beter omschreven dat de argumenten juridisch steek moeten houden.

Toch is deze andere manier van formuleren niet helemaal onschuldig. Als we ‘geloof’ als ‘bewijsbaar’ opvatten, dan is het maar zeer de vraag of deze ‘bewijsbaarheid’ aan de twee genoemde regels van de kansrekening voldoen. Het kan immers zo zijn dat voor een bepaalde hypothese geen of nauwelijks bewijs is, maar voor zijn ontkenning evenmin. Dat betekent dat de opgetelde bewijsbaarheid van deze twee hypothesen geen 100% kan zijn.

Een eenvoudig voorbeeld is de situatie waarin een auto een ongeluk veroorzaakt, en de auto twee inzittenden heeft. Soms is het onduidelijk wie van de twee aan het stuur zat tijdens het ongeluk – bijvoorbeeld als de twee inzittenden elkaar aanwijzen. In dat geval is de bewijsbaarheid dat de eerste persoon aan het stuur zat klein (ervan uitgaande dat er verder geen aanwijzingen zijn) maar de bewijsbaarheid dat de tweede persoon stuurde, is ook klein. Dat zou betekenen dat we in zo’n geval onze daadwerkelijke onzekerheid niet met de ‘klassieke’ kansrekening kunnen uitdrukken.

In zichzelf is het niet verrassend dat zoiets kan gebeuren. We hebben herhaaldelijk gesteld dat het Bayesiaanse denk-

⁹ Bijv. in Glenn Shafer, *A Mathematical Theory of Evidence*, Princeton, NJ: Princeton University Press 1976.

model ook maar een model is, en modellen zijn nu eenmaal niet de werkelijkheid. We lopen hier tegen de grenzen van zo'n model aan – dat moet vroeg of laat een keer gebeuren. Het is dan niet nodig om het hele model overboord te gooien, maar we moeten ons wel realiseren dat we nu in een situatie zitten die speciale aandacht vereist.

We kunnen in principe twee kanten op. We kunnen binnen het bestaande model kijken of we de situatie toch nog kunnen redden, of we kunnen het model verfijnen zodat de geschetste situatie er wel in beschreven kan worden. Pragmatisch gezien kiest men bijna altijd voor de eerste optie. Men kan er in de situatie van de twee inzittenden voor kiezen om de context te vertalen in a priori fiftyfiftykansen dat elk van de inzittenden de bestuurder was. Als er verder geen enkele aanwijzing is die hier meer duidelijkheid in kan scheppen, dan zal op basis hiervan geen van beide inzittenden veroordeeld kunnen worden. Dat lijkt congruent met de gedachte dat de bewijsbaarheid voor beide inzittenden te klein is om tot een veroordeling te komen.

De tweede optie is vanuit een meer filosofisch perspectief interessanter: kunnen we het model dat wij voor onze onzekerheid hanteren zo verfijnen dat situaties als deze er wel bevredigend mee kunnen worden beschreven? Het antwoord is ja, dat kan – het kan zelfs op verschillende manieren. Dit boek is niet de plaats om hier uitgebreid op in te gaan, want de verschillende benaderingen zijn niet zo eenvoudig en niet onmiddellijk intuïtief. De populairste methode werd geïntroduceerd door Glenn Shafer. Hij introduceerde zogenaamde *belief functions*, een generalisatie van wat wij kansen hebben genoemd, die beter passen bij het concept 'bewijsbaarheid'. De theorie van Shafer was niet helemaal bevredigend op het

gebied van het verwerken van nieuwe informatie, maar juist op dat gebied is er enige jaren geleden vooruitgang geboekt.¹⁰

Vooralsnog verhindert de toch wat grotere complexiteit van Shafers concept brede toepassing in het recht. Dat neemt niet weg dat het bestaan van zo'n theorie van grote waarde kan zijn: ze herinnert ons eraan dat we geen enkel model en geen enkele theorie als dé waarheid mogen gaan zien. Elke theorie heeft haar grenzen, en wat vandaag nog te gecompliceerd is om toegepast te kunnen worden, kan met nieuwe inzichten morgen gemeengoed zijn.

¹⁰ Zie Ronald Meester en Timber Kerkvliet, “Waarom onzekerheid in het recht beter met belief functies dan met klassieke kansrekening benaderd kan worden: theorie en praktijk”, *Expertise en Recht* (april 2017): 62-68; hier zijn ook verdere referenties naar belief-functies te vinden.

HOOFDSTUK 6

Analyse van een aantal zaken

The proof of the pudding is in the eating, zeggen de Engelsen, en gelijk hebben ze. Natuurlijk hebben we in dit boek al een groot aantal voorbeelden gegeven van het gebruik van Bayesiaans redeneren in de rechtszaal. Veelal betrof dit commentaar op de gang van zaken, het vonnis of een ander aspect van een zaak. In dit hoofdstuk willen we naar een aantal zaken kijken en die in hun geheel analyseren. We bestuderen de situatie op basis van het vonnis of het arrest, en kijken er vervolgens met een Bayesiaanse bril naar.

Het is nadrukkelijk niet de bedoeling om een uitspraak goed of fout te vinden – we gaan niet op de stoel van rechter zitten. We zullen wel opbouwend kritisch zijn over de argumentatie, in de hoop dat we daar allemaal iets van leren. Hierbij moeten we wel een voorbehoud maken. Wij zijn op geen enkele manier betrokken geweest bij de zaken die we bespreken. We hebben niet het hele dossier kunnen bestuderen en baseren onze analyse op wat er in de uitspraken over geschreven wordt. Doorgaans geven die uitspraken een goed beeld van de situatie en de overwegingen van de rechtbank of het hof, maar het is goed mogelijk dat er informatie in de dossiers stond die wel gebruikt is maar niet in de uitspraak is opgenomen. We zullen het moeten doen met de informatie die we hebben.

DE ZAAK ECLI:NL:RBZWB:2021:2638; EEN LEVENSDELICT MET TEGENSTRIJDIGE AANWIJZINGEN EN VRIJSPRAAK

Tenlastelegging en achtergrond

In deze zaak was de verdachte kort gezegd ten laste gelegd dat hij het slachtoffer om het leven heeft gebracht door haar te steken met een mes. De rechtbank sprak hem vrij. We beginnen met het schetsen van de situatie en de context. De verdachte en het slachtoffer wonen in hetzelfde gebouw. De verdachte woonde op de begane grond, het slachtoffer op de tweede verdieping, boven de verdachte. Uit het onderzoek ter terechtzitting is naar voren gekomen dat verdachte met sommige bewoners op gespannen voet leefde. Nadat op een vergadering van de Vereniging van Eigenaren in december 2018 aan verdachte een ultimatum was gesteld om zijn koopappartement te verlaten, namen de spanningen toe. De verdachte heeft zich in die periode ten opzichte van sommige bewoners stevig en ook dreigend uitgelaten, zowel verbaal als schriftelijk. Ook in e-mails aan de wijkagent over de situatie heeft hij zich in vergaande bewoordingen geuit. In de vroege avond van 24 mei 2019, de dag van het delict, heeft verdachte gebeld naar het radioprogramma *Zegt u het maar* van Omroep Zeeland, waarin hij heeft verteld dat hij werd gepest en wat dat met hem doet. Uit het dossier komt naar voren dat politieagenten en bewoners hebben verklaard dat inderdaad sprake was van pestgedrag richting verdachte.

Uit telefoongegevens van de vaste telefoonlijn van het slachtoffer is gebleken dat verdachte op 22 mei 2019 rond 22.30 uur twee keer naar het slachtoffer heeft gebeld. Verdachte heeft hierover verklaard dat hij eerder die avond al aan de deur was geweest bij het slachtoffer om te praten over parkeerproblemen. Omdat niet werd opengedaan heeft hij later twee keer gebeld. Hij was boos dat zij de deur niet

open had gedaan, en wilde haar laten weten hoe irritant zij was. Er werd niet opgenomen waarna hij de voicemail heeft ingesproken.

In de ochtend van 24 mei 2019 vond een incident plaats waarbij de jongste hond van verdachte over de reling van zijn balkon op de begane grond was gesprongen. De hond was daarna tegen een medebewoonster opgesprongen die daardoor viel. Diverse bewoners hadden het incident gezien en reageerden daar verbaal op richting verdachte, waarop verdachte verbaal terug reageerde. Ook het slachtoffer had het incident gezien. Verdachte heeft de medebewoonster overeind geholpen en een andere bewoner heeft haar naar de zorgvoorziening in het complex gebracht. Door dit incident is bij verdachte paniek en stress ontstaan, omdat dit incident in zijn ogen mogelijk aanleiding kon zijn om de hond af te laten maken. Ook vreesde hij dat het incident gevolgen voor hemzelf zou kunnen hebben. Verdachte heeft daarom zijn advocaat gebeld en per e-mail contact gezocht met de wijkagent. De wijkagent was die dag niet in de gelegenheid om terug te mailen. Verdachte heeft verklaard dat hij daarna is gaan slapen en rond half vier wakker werd. Vervolgens is hij naar de fitnessruimte gelopen, maar toen hij daar was, besliste hij om toch niet te gaan sporten. Om 16.14 uur heeft een medebewoonster verdachte gemaild over de steekpartij en haar vermoeden geuit dat het slachtoffer was overleden.

De verdachte heeft, nadat de ambulance er was, zijn honden uitgelaten in het park. De echtgenoot van het slachtoffer heeft verklaard dat hij in de woonkamer zat toen hij zijn vrouw om hulp hoorde roepen. Hij is naar de hal in hun eigen appartement gelopen waar hij haar zag liggen in een plas bloed. Hij is eerst naar de burens gelopen. Die waren niet thuis, waarna hij met de lift naar beneden is gegaan. Daar heeft hij om hulp gevraagd. De zorgcoördinator van het

appartementencomplex is naar boven gegaan en toen hij de situatie zag heeft hij om 16.02 uur 112 gebeld. Ondertussen was een verpleegkundige vanuit een andere woning gekomen. Zij is gestart met reanimeren. Het gebruik van een AED had geen effect. Daarna is het ambulancepersoneel gearriveerd. Het slachtoffer bleek overleden te zijn.

Het forensisch onderzoek

Op verschillende plaatsen is DNA-onderzoek verricht. Op de deurbel, aan de buitenzijde van de voordeur van het slachtoffer, werd geen bruikbaar DNA aangetroffen. Wel vond men een ‘vette veeg’ aan de buitenkant van deze voordeur, op ongeveer 120 cm hoogte en 2 à 3 cm breed. Onderzoek wees uit dat deze veeg DNA-materiaal van één man bevatte, overeenkomend met het profiel van de verdachte. De matchkans van het spoor werd kleiner dan 1 op 1 miljard bevonden, en dit was voor de rechtbank voldoende om te concluderen dat het DNA in de veeg van de verdachte afkomstig was.

Aan de binnenzijde van dezelfde voordeur, dus in het huis van het slachtoffer, werd op de deurklink DNA gevonden van minimaal vier personen. Dit mengprofiel was volgens de forensisch deskundigen meer dan één miljard keer waarschijnlijker als verdachte een donor was dan wanneer dit niet zo was. De rechtbank concludeert dan ook dat het inderdaad het DNA van verdachte is dat in het mengsel is aangetroffen; het gaat daarbij zelfs om een relatief groot deel van het aangetroffen DNA. Dit gegeven wordt nog interessanter als de verdachte beweert nooit in het huis van het slachtoffer te zijn geweest.

Op het lichaam en kleding van het slachtoffer werd geen DNA van verdachte aangetroffen. Dat is relevant omdat het slachtoffer overleed als gevolg van een messteek en er aanwij-

zingen waren dat het slachtoffer zich enigszins had verweerd. In de woning van de verdachte zijn messen aangetroffen en in beslag genomen. Op geen van de messen werd het DNA van het slachtoffer aangetroffen.

De hypothesen en de sterkte van het bewijs

In de uitspraak worden allereerst de volgende hypothesen besproken:

H_1 : De verdachte heeft het slachtoffer gestoken.

H_2 : Een ander persoon (niet zijnde haar echtgenoot) heeft het slachtoffer gestoken.

Deze twee hypothesen werden uitsluitend in relatie tot de verwondingen van het slachtoffer beschouwd, het overige bewijs speelde hierbij geen rol. De rechtbank ziet in (snij)letsels aan de rechterhand en blauwe plekken op de linker handrug aanwijzingen dat het slachtoffer zich heeft afgeweed bij het steekincident. Bovendien heeft de echtgenoot van het slachtoffer verklaard dat hij haar een gil hoorde slaken. De rechtbank gaat dan ook uit van het scenario waarin een worsteling met afweer heeft plaatsgevonden waarbij het slachtoffer afwerend heeft gereageerd op haar aanvaller en er meer fysiek contact tussen dader en slachtoffer heeft plaatsgevonden dan slechts de steekhandelingen. De deskundige concludeert in dat geval dat de DNA-sporen waarschijnlijker zijn als een ander persoon het slachtoffer heeft gestoken (H_2) dan als verdachte dat heeft gedaan (H_1). Hierbij speelt het gebrek aan DNA-sporen van verdachte op het lichaam en de kleding van het slachtoffer een grote rol. Dit resultaat merkt de rechtbank aan als ontlastend voor verdachte.

Voor het DNA-materiaal dat op de binnenste deurklink is aangetroffen, stelt de rechtbank twee verdere hypothesen op, namelijk:

H_3 : De verdachte heeft de deurklink vastgepakt op 24 mei 2019.

H_4 : De verdachte heeft de deurklink nooit vastgepakt of aangeraakt.

De bespreking die daarop volgt is het waard om hier helemaal te citeren:

“Bij de evaluatie van deze hypothesen speelde de eerder besproken vette veeg op de buitenzijde van de voordeur een prominente rol. Of deze veeg is ontstaan *tijdens* het incident, of *daarvóór* al aanwezig was op de voordeur, heeft invloed op de te verwachten DNA-sporen. Er is daarom gekozen voor het evalueren van de DNA-resultaten onder beide mogelijkheden. De rechtbank ziet zich gesteld voor de vraag of zij in het dossier aanwijzingen ziet die betrekking hebben op het ontstaansmoment van de veeg. Zij betreft hierbij onder andere het feit dat de veeg door verbalisanten als ononderbroken wordt omschreven. Deze informatie acht de rechtbank niet van doorslaggevende betekenis in deze kwestie. Het is ook mogelijk dat de veeg al eerder op de deur zat en dat de dader deze tijdens het incident niet heeft aangeraakt (of niet zodanig dat deze verstoord is geraakt). De rechtbank betreft hierbij ook het feit dat de deskundige van het TMFI heeft verklaard geen tekenen van afbraak te zien in de DNA-kenmerken van verdachte in de bemonstering van de veeg. Deze verklaring kan de rechtbank echter niet vertalen naar een concreet moment van ontstaan van de

veeg op de deur. De rechtbank moet concluderen dat het dossier onvoldoende informatie bevat om een onderbouwd oordeel te geven over het ontstaansmoment van de vette veeg op de buitenzijde van de voordeur. Zo valt bijvoorbeeld niet uit te sluiten dat de veeg is ontstaan toen verdachte twee dagen eerder aan de deur van het slachtoffer stond. Zij zal daarom uitgaan van zowel de mogelijkheid dat de veeg tijdens het incident is ontstaan, als dat deze al vooraf aanwezig was, door uit te gaan van de gezamenlijke reikwijdte van de getrokken conclusies.

De deskundige concludeert in dat geval dat de DNA-resultaten waarschijnlijker tot en met veel waarschijnlijker zijn als verdachte de deurklink op 24 mei 2019 heeft vastgepakt (Hypothese 3), dan als verdachte de deurklink nooit heeft vastgepakt of aangeraakt (Hypothese 4). Hierbij speelt het gevonden DNA van verdachte op de binnenste deurklink van de voordeur én op de vette veeg aan de buitenzijde van de voordeur een grote rol. Dit resultaat merkt de rechtbank aan als belastend voor de verdachte, mede gelet op de verklaring van verdachte dat hij nooit in de woning van het slachtoffer is geweest.”

Kennelijk is de rechtbank van mening dat de veeg aan de buitenkant op de een of andere manier gerelateerd is aan het gevonden DNA aan de deurklink aan de binnenkant van de deur. Hoe de logica van de rechtbank hier precies werkt weten we niet, maar het lijkt erop dat de rechtbank naar haar eigen zeggen conservatief oordeelt: ook als de veeg eerder is ontstaan, dan nog is het DNA op de deurklink onder H_3 veel waarschijnlijker dan onder H_4 . Sterker nog, de verdachte beweert nooit in de woning van het slachtoffer te zijn geweest. In die omstandigheid is het DNA op de deurklink onder H_4 niet te verklaren. Dat lijkt ons

een sterk argument dat tegen de onschuld van de verdachte pleit.

In een ‘tussenconclusie’ oordeelt de rechtbank als volgt:

“De rechtbank ziet zich aldus geconfronteerd met zowel ontlastend als belastend bewijs in de resultaten van het DNA-onderzoek op activiteitsniveau. Bij de weging van deze voor de bewijsvraag conflicterende onderzoeksresultaten betreft de rechtbank naast de door de deskundige berekende (reikwijdtes van de) bewijskrachten ook de mate waarin de resultaten in de ogen van de rechtbank delictgerelateerd zijn. Zij overweegt dat DNA-sporen op het lichaam en de kleding van het slachtoffer een sterker en directer verband hebben met het steekincident dan DNA-sporen op en rondom de voordeur van het slachtoffer. De rechtbank hecht voor haar eindoordeel met betrekking tot het DNA-onderzoek in brede zin daarom relatief meer waarde aan de voor de verdachte ontlastende conclusies over het steekincident zelf, dan aan de voor de verdachte belastende conclusies over het aanraken van de deurklink.

Samenvattend concludeert de rechtbank na een weging van het deels tegenstrijdige forensische bewijs dat deze meer steun geeft aan het scenario dat een ander persoon het slachtoffer gestoken heeft, dan dat verdachte dat heeft gedaan.”

We willen hier toch nog wat dieper op ingaan, omdat het niet onmogelijk is dat de rechtbank een principieel punt over het hoofd heeft gezien. We leggen dat hierna uit.

Delictgerelateerdheid van likelihood ratio's

De beoordeling van de bewijswaarde van een bewijsmiddel ten opzichte van twee hypothesen wordt verbaal of numeriek uitgedrukt in een likelihood ratio. De gedachte is, zoals altijd, om te bepalen bij welke hypothese de gegevens het beste passen. Dat kun je soms numeriek uitdrukken. Het punt dat we nu willen maken is dat bij de beoordeling of gegevens beter bij de ene of bij de andere hypothese passen, je al gebruikmaakt van de context waarin die gegevens verkregen zijn. Het feit dat er geen DNA op de kleding van de verdachte gevonden werd, leidde tot een bepaalde likelihood ratio bij H_1 en H_2 waarin de context al gebruikt was: de kleding werd gedragen tijdens het doden, en dat is de reden om er een bepaalde likelihood ratio aan toe te kennen. Als die kleding uit de kledingkast van de verdachte zou komen, dan zou eventueel DNA-bewijs irrelevant zijn: die kleding zou immers niet gedragen zijn en de bijbehorende likelihood ratio zou neutraal zijn. Het is daarom in principe onjuist om daarna nog te zeggen dat je aan de ene likelihood ratio meer waarde toekent dan aan een andere, omdat de gegevens duidelijk delictgerelateerd zouden zijn. Die informatie had je immers al gebruikt bij het vaststellen van de likelihood ratio. Je kunt, met andere woorden, de ene likelihood ratio niet belangrijker vinden dan een andere: het belang ervan wordt namelijk precies uitgedrukt door de grootte die de likelihood ratio heeft.

Hier is een analogie dat het punt dat we willen maken wellicht verheldert. Stel er zijn in een casus twee likelihood ratio's; één die van een vingerafdruk afkomstig is en één die van een DNA-profiel afkomstig is. Als deze aanwezig onverhoopt ongeveer even sterk zijn maar in een andere richting wijzen, dan kun je niet kiezen welke van de twee je

belangrijker vindt; ze zijn er nu eenmaal allebei. Een argument dat DNA-bewijs betrouwbaarder zou zijn dan vingerafdrukken kan niet juist zijn, want als dat zo was dan zou de likelihood ratio van de vingerafdruk simpelweg minder overtuigend geweest zijn.

Hier is een ander voorbeeld. Stel dat we willen vaststellen of Ceciel al dan niet vals speelt bij het werpen van een munt. Zonder manipulatie heeft iemand 50% kans om kop te gooien, maar we vermoeden dat Ceciel de boel manipuleert en 75% kans heeft op kop. Stel vervolgens dat we een worp van Ceciel waarnemen waarbij munt wordt gegooid. Deze observatie is ontlastend voor Ceciel, want munt is waarschijnlijker onder het scenario dat ze eerlijk is dan onder het scenario dat ze manipuleert.

Maar stel je nu voor dat we een muntworp waarnemen waarvan we niet zeker weten of deze door Ceciel is uitgevoerd; we zijn er 50% zeker van dat zij dat heeft gedaan, maar met kans 50% heeft iemand anders de worp uitgevoerd. Nu is de uitkomst ‘munt’ nog steeds ontlastend voor Ceciel, maar wel minder ontlastend dan voorheen, omdat we niet zeker weten wie de worp heeft uitgevoerd. We weten, met andere woorden, niet zeker of deze muntworp eigenlijk wel ‘delictgerelateerd’ is.

Als we, in het meest extreme geval, zeker weten dat de muntworp niet door Ceciel is uitgevoerd, dan is het bewijs helemaal niet meer delictgerelateerd en zal de likelihood ratio gelijk worden aan 1: we leren er niets van als het gaat om de eventuele manipulatie van Ceciel.

De moraal van dit verhaal is dat bij het berekenen van de likelihood ratio al gebruikgemaakt wordt van het al dan niet direct delictgerelateerd zijn van de gegevens. Hoe meer gerelateerd, hoe belangrijker en onderscheidender het bewijs zal zijn, uitgedrukt in een hogere likelihood ratio. Om terug

te keren naar de zaak die we hier bespreken: het komt er dus op neer dat de rechtbank een principiële vergissing lijkt te begaan door de ene likelihood ratio belangrijker te maken dan een andere op basis van de delictgerelateerdheid. Zo kan het niet werken.

Verschillende hypotheseparen

De twee likelihood ratio's hebben betrekking op verschillende hypotheseparen. Dat maakt een combinatie natuurlijk lastig, zo niet onmogelijk. De rechtbank lost dat op door te stellen dat de ene belangrijker is dan de andere omdat de eerste zeker delictgerelateerd is en de andere wellicht niet, maar we zagen zojuist dat dit niet juist is. De twee hypotheseparen zijn allebei op activiteitenniveau maar hebben betrekking op verschillende activiteiten: de ene gaat over het steekincident, de andere over het al dan niet vastpakken van de deurklink aan de binnenkant van de voordeur. Het zou overzichtelijker zijn geweest om het hele scenario in één paar hypotheses te formuleren, ongeveer als volgt:

H_5 : De verdachte heeft aangebond, is binnengelaten, heeft het slachtoffer met messteken om het leven gebracht en is daarna naar buiten gegaan.

H_6 : Een ander, onbekend persoon, is verantwoordelijk voor de gebeurtenissen zoals in H_5 omschreven.

Nu we alle activiteiten samen hebben gebracht, kunnen we vervolgens onderzoeken in hoeverre het beschreven DNA-bewijs past bij deze twee hypotheses. Welnu, analyse van het DNA op de deurklink bewijst dat de verdachte hieraan (fors) heeft bijgedragen. Deze bevinding past zeer goed bij H_5 , maar is onder H_6 vrijwel niet te verklaren. Bedenk hierbij ook dat

de verdachte ontkent ooit in het huis van het slachtoffer te zijn geweest. De bijbehorende likelihood ratio is lastig numeriek in te schatten, maar als we aannemen dat iemand die naar buiten loopt de deurklink moet vastpakken en dat bij een dergelijke aanraking normaal gesproken DNA-materiaal achterblijft, dan zien we dat de bewijskracht van die match zeer aanzienlijk is.

De bewijskracht van het afwezig zijn van DNA-materiaal op de kleding en het lichaam van het slachtoffer moet haast wel veel kleiner zijn. De rechtbank kent echter veel waarde toe aan het feit dat er geen DNA van de verdachte op het lichaam en kleding van het slachtoffer is aangetroffen. Het is voor ons niet zo duidelijk hoe ontlastend dat feit eigenlijk is. De uitdrukking “waarschijnlijker” die de deskundige gebruikt, geeft in dit geval niet veel informatie, en we zouden graag meer willen weten over de reden van deze kwalificatie, maar sowieso lijkt ze ons niet heel sterk. Als we naar de numerieke waarde kijken die volgens het NFI zelf overeenkomt met de kwalificatie “waarschijnlijker”, dan zien we dat we moeten denken aan een likelihood ratio van 10 à 100.¹ Dat is in vergelijking met de orde van grootte van de likelihood ratio die hoort bij het DNA op de deurklink, bijna te verwaarlozen.

Op deze manier geformuleerd betreft het slechts één hypothesepaar. Het ontlastende bewijs is licht, het belastende bewijs is sterk.

Overig bewijs

We citeren voor wat betreft het overige bewijs in de zaak opnieuw uit de uitspraak, omdat het citaat alle relevante informatie bevat.

¹ Zie Ronald Meester en Klaas Slooten, *Kan dat geen toeval zijn*, Amsterdam: Amsterdam University Press 2022, p. 91.

“Het niet-forensische deel van het dossier bevat geen direct bewijs dat in de richting van verdachte als dader wijst. Er waren geen getuigen die hebben gezien dat het slachtoffer werd neergestoken. Het steekwapen waarmee het slachtoffer is gestoken is niet gevonden. Verdachte heeft steeds ontkend dat hij het slachtoffer heeft gestoken.

Een onderbuurvrouw van het slachtoffer, heeft tussen 15:00 uur en 16:00 uur iemand gehaast de trap af naar beneden zien lopen. Over de omschrijving van die persoon heeft zij tegen diverse getuigen, de politie en de rechter-commissaris wisselend verklaard.

Door getuigen wordt gesuggereerd dat zij niet heeft durven zeggen dat zij verdachte had herkend als de persoon op de trap omdat zij erg bang van hem is. De onderbuurvrouw heeft daar zelf over verklaard dat zij inderdaad bang is voor verdachte, maar dat zij het had gezegd als ze verdachte had herkend omdat dat haar burgerplicht is. Het enige dat zij nog zeker zegt te weten is dat die persoon een lichtblauw overhemd of shirt droeg. Ook heeft zij verklaard dat zij in het genoemde tijdvak nog meer personen de trap op en af heeft zien gaan. Over het blauwe shirt overweegt de rechtbank dat het dossier foto's bevat van verdachte op 24 mei 2019 en dat op die foto's is te zien dat hij een lichtblauw shirt draagt. Zelfs als daaraan al de conclusie zou kunnen worden verbonden dat het verdachte is geweest die de onderbuurvrouw op de trap heeft gezien, dan nog plaatst dat hem niet in het appartement van het slachtoffer. Enig bewijs voor het daderschap van verdachte kan er dan ook niet uit worden afgeleid.

Gelet op deze omstandigheden ziet de rechtbank in de verklaring van getuige geen belastend bewijs.”

We hebben te maken met twee observaties van de getuige: ze heeft iemand gehaast de trap af zien lopen, en ze herinnert zich dat deze persoon een lichtblauw overhemd droeg. De rechtbank ziet hier geen belastend bewijs in, omdat er geen bewijs voor daderschap van de verdachte uit kan worden afgeleid.

Wat dat laatste betreft heeft de rechtbank natuurlijk gelijk, maar bewijs kan belastend zijn zonder dat het doorslaggevend is. Dat lijkt hier toch wel aan de hand te zijn. Laten we om te beginnen kijken naar de vraag of de gehaaste persoon inderdaad de verdachte was. Hiertoe stellen we twee hypothesen op:

H_7 : De gehaaste persoon was de verdachte.

H_8 : De gehaaste persoon was een ander, onbekend persoon.

Het bewijs bestaat nu uit de observatie dat de gehaaste persoon een lichtblauw overhemd droeg. Onder H_7 is de kans daarop 100%, afgezien van de mogelijkheid dat iemand zich overdag kan verkleden. Immers, uit foto's blijkt dat de verdachte die dag een lichtblauw overhemd droeg. Onder H_8 is de kans kleiner; de meeste mensen – een ambtenaren-omgeving daargelaten – dragen immers op een gegeven dag geen lichtblauw overhemd. Dat betekent dat de aanwijzing dat de gehaaste persoon een lichtblauw overhemd droeg, bewijs vormt voor H_7 . Hoe sterk dat bewijs is, is moeilijk vast te stellen of te kwantificeren, maar het is ontegenzeggelijk een aanwijzing.

De vraag is vervolgens of er reden is om aan te nemen dat de gehaaste persoon (die dus heel goed de verdachte zou kunnen zijn) degene is die het slachtoffer heeft omgebracht. Deze vraag kunnen we niet zonder meer beantwoorden omdat het antwoord afhangt van contextuele informatie die we niet

hebben. Zoals altijd bij een Bayesiaanse analyse hangt het antwoord immers af van de vraag hoe uitzonderlijk het is dat de verdachte boven is (hij woont daar immers niet) en haast lijkt te hebben. Als hij nooit boven komt, dan is het belastender dan wanneer hij dagelijks boven komt. Verder heeft de rechtbank natuurlijk gelijk door te stellen dat zelfs indien vastgesteld kan worden dat deze persoon de verdachte is, deze vaststelling hem niet onmiddellijk in het huis van het slachtoffer plaatst. Het DNA-bewijs op de binnenkant van de deurklink doet dat echter wél. Dit in combinatie met de bewering van de verdachte dat hij daar nooit binnen is geweest, maakt het allemaal behoorlijk belastend. Daarmee lijkt ons de laatste zin uit het citaat waarin de rechtbank aangeeft in de verklaring van getuige geen belastend bewijs te zien, toch wat voorbarig.

Verdere analyse en conclusie

De rechtbank schrijft bijna terloops dat de verdachte heeft verklaard nooit in de woning van het slachtoffer te zijn geweest. Het DNA op de deurklink aan de binnenkant van de voordeur vormt extreem sterk bewijs voor het tegendeel. Als we er, zoals de rechtbank doet, van uitgaan dat het inderdaad DNA-materiaal van verdachte is, dan mag de rechtbank er ook van uitgaan dat de verdachte onwaarheid spreekt. Het spreken van onwaarheid kan zélf ook Bayesiaans geduid worden: de kans dat verdachte onwaarheid spreekt is onder H_5 vele malen groter dan onder H_6 , en deze leugenachtigheid is dus sterk belastend bewijs.

De kern van de zaak lijkt ons dus uiteindelijk niet dat er tegenstrijdig forensisch bewijs zou zijn. Als we de zaak als geheel bekijken, dan stellen we allereerst vast dat de prior overtuiging dat de verdachte inderdaad de dader is,

niet extreem klein is. We weten niets over de verdere sociale context van het slachtoffer, maar ze heeft onmin met de verdachte. Uit het hele verhaal blijkt verder ook dat de verdachte met grote zekerheid in het huis van het slachtoffer geplaatst kan worden en dat hij daarover onwaarheden vertelt. Dat is zeer belastend. Bovendien zijn er verdere aanwijzingen die weliswaar in zichzelf niet doorslaggevend kunnen zijn, maar wel degelijk bijdragen aan het belastende bewijs.

De kernvraag waarvoor de rechtbank zich gesteld ziet, is of er in het licht van al deze aanwijzingen voldoende bewijs bestaat voor H_5 . Dit is uiteindelijk een kwestie van persoonlijke beoordeling. Het belangrijkste bewijs, het DNA op de klink van de voordeur, lijkt echter niet of nauwelijks te verklaren te zijn onder de hypothese dat de verdachte niets met de dood van het slachtoffer te maken heeft. Het belangrijkste ontlastende bewijs past juist weer wel beter bij de hypothese dat het een ander was, maar die aanwijzing was minder sterk.

Uiteindelijk is in onze benadering van de zaak nog steeds sprake van een bepaalde mate van onzekerheid die ervoor kan zorgen dat een rechter vrijspreekt. Wat we hebben willen aantonen is dat die onzekerheid net anders ligt dan blijkt uit het vonnis. Interessant is dat de contextuele informatie en het 'indirecte' bewijs (de verdachte kan geplaatst worden in het appartement) wijzen richting de schuld van de verdachte terwijl het beeld van de steekpartij die moet hebben plaatsgevonden in het duister blijft. Onzekerheid over juridisch cruciale ten laste gelegde handelingen zullen we in de hierna te bespreken zaken ook terugzien.

DE ZAAK ECLI:NL:GHARL:2020:10710; DE 'THIALF ZAAK'

Tenlastelegging, context en tegenstrijdige verklaringen over het gebeurde

In deze zaak draait het feitelijk om de vraag of de verwondingen van het slachtoffer toe te schrijven zijn aan het afvuuren van een wapen door de verdachte. Het eerste primair ten laste gelegde feit is een poging tot moord/doodslag, wat in de tenlastelegging vervolgens wordt afgezwakt naar zware mishandeling en mishandeling. Steeds staan daarbij het wapen en de verwondingen centraal. Het tweede ten laste gelegde feit is bedreiging, wederom met behulp van een wapen. Door de onzekerheid over de vraag of er überhaupt een schietpartij heeft plaatsgevonden volgt vrijspraak. De zaak is voor ons interessant omdat hier geredeneerd moet worden op het scherpst van de snede. We beginnen met een schets van de context en wat er gebeurd zou zijn volgens de aangever dan wel de verdachte.

Op 7 januari 2017 rond 19.00 uur is verdachte aanwezig op het parkeerterrein van het Thialf-stadion in Heerenveen. Hij staat daar op het slachtoffer te wachten, die naar zijn verwachting daar zal komen na afloop van de schaatswedstrijden in het stadion. Zij kennen elkaar vanwege zakelijke contacten en hebben een zakelijk geschil. Verdachte is eerder op de dag naar het stadion gekomen om de auto van het slachtoffer op te zoeken, een Tesla. Nadat hij de vermoedelijke auto heeft aangetroffen, is hij weer vertrokken en is teruggekomen tegen de tijd dat de wedstrijden zijn afgelopen. Slachtoffer verlaat het stadion en stapt op het parkeerterrein in zijn auto en maakt aanstalten om weg te rijden van het terrein. Verdachte loopt vervolgens in de richting van de auto van slachtoffer die op hem af komt rijden. Slachtoffer stopt de auto ter hoogte van de verdachte, die zich schuin voor de

auto van slachtoffer bevindt terwijl hij een zwaaiende of wenkende beweging naar slachtoffer maakt. Slachtoffer heeft zijn raam geopend. Vervolgens probeert verdachte tevergeefs het portier achter de bestuurder te openen. Hij staat dan naast de auto ter hoogte van het bestuurdersportier. Het raam van het bestuurdersportier is nog steeds geopend. Op enig moment heeft verdachte een wapen, of een daarop lijkend voorwerp, in de hand. Slachtoffer geeft dan gas en rijdt weg. Tegelijkertijd is er een knal te horen en valt verdachte op de grond. Slachtoffer ramt een paar auto's. Even verderop stopt hij, stapt uit de auto en roept dat hij is beschoten. Slachtoffer is dan gewond aan zijn rechterhand. Later blijkt dat de aan de handrug grenzende kootjes van de middel- en ringvinger in meerdere stukken gebroken of versplinterd zijn. Verdachte heeft inmiddels, met medeneming van het wapen, het parkeerterrein verlaten. Het wapen is niet gevonden.

Dat er een vuurwapen was en dat verdachte dit wapen op enig moment in de hand had op het moment dat het slachtoffer gas gaf met zijn auto, wordt door zowel slachtoffer als verdachte verklaard. Zij hebben echter verschillend verklaard over wie dat wapen mee heeft gebracht, wat voor wapen het was, hoe het in de hand van de verdachte is gekomen en of daarmee gericht op het slachtoffer is geschoten.

We kijken nu eerst naar het scenario van de verdachte. Deze heeft een zakelijk geschil met slachtoffer. Hij zegt veel geld tegoed te hebben van het slachtoffer vanuit een overeenkomst met betrekking tot de overname van zijn (gefaillleerd) bedrijf. Eerder heeft hij tevergeefs gepoogd om slachtoffer te bewegen hem te betalen. Hij woont inmiddels in Brazilië en is naar Nederland afgereisd om nogmaals te proberen om het slachtoffer te laten betalen. Hij verwacht dat slachtoffer bij het EK schaatsen in het Thialf-stadion in Heerenveen op 7 januari 2017 aanwezig is, en wacht daar op het slachtoffer.

Hij weet dat het slachtoffer in een Tesla rijdt en zoekt op het parkeerterrein naar die Tesla. Op het moment dat het slachtoffer in zijn Tesla weggrijdt, gaat verdachte erop af. Als het slachtoffer stopt, checkt verdachte of het de juiste persoon is die in de auto zit. Het slachtoffer doet het raampje open, verdachte klopt het slachtoffer op zijn borst, geeft hem ‘de beste wensen’ in verband met het nieuwe jaar en probeert links achterin in de auto te stappen. Hij krijgt de deur niet open en kijkt weer op. Op dat moment ziet hij een wapen in de hand van het slachtoffer en pakt meteen dat wapen uit zijn hand. Het slachtoffer drukt vervolgens op iets op het dashboard en de auto schiet weg. Meteen valt verdachte met dat wapen in zijn hand op de grond en hoort hij een knal. Hij raakt in paniek en vlucht. Het wapen gooit hij – gewikkeld in een petje – in een sloot. Hij vlucht het land uit en houdt zich schuil totdat hij in Spanje wordt aangehouden.

Tot zover het scenario van de verdachte. Het slachtoffer verklaart anders. Hij heeft naar eigen zeggen zakelijke contacten gehad met verdachte in verband met een voorgenomen overname van het bedrijf van verdachte. Van deze overname heeft hij op enig moment afgezien omdat het bedrijf niet gezond was. Na het faillissement van het bedrijf van verdachte heeft hij contact gehad met de curator en onderdelen van het bedrijf overgenomen. Vervolgens stuurde verdachte mails en appjes naar het slachtoffer dat hij geld van hem wilde. Hij liet zijn jurist daar altijd op reageren en had zelf geen contact met verdachte. Op 7 januari 2017 rond 19.00 uur verlaat hij het stadion na afloop van het EK schaatsen en stapt daar op het parkeerterrein in zijn Tesla om eten te gaan halen voor zijn gezin. Hij start de auto en rijdt weg. Even verderop staat een man die op zijn auto af komt lopen. Hij stopt en doet het raampje open en wil de man aanspreken. De man staat aan de bestuurderskant en probeert de

linker achterdeur van de auto te openen. Hij zegt “wat doe je” tegen de man. De man kijkt hem aan en vraagt of hij hem niet meer kent. Het slachtoffer herkent dan verdachte. Hij ziet een pistool in de hand van verdachte en drukt direct op het gaspedaal. De auto schiet weg. Hij hoort meteen een knal en ramt met zijn auto enkele andere auto's. Hij ziet dan dat hij gewond is aan zijn hand. Hij stapt uit en roept dat hij beschoten is. Al schreeuwend rent hij weg.

Het belang van de relatie tussen letsel en vuurwapen

Het hof is van mening dat het scenario van de verdachte “niet aannemelijk geworden is”, om verschillende redenen. Waarom zou het slachtoffer bijvoorbeeld een wapen meenemen naar het stadion? Het hof stelt vast dat verdachte kennelijk iets wilde van het slachtoffer, daarvoor helemaal vanuit Brazilië naar Nederland kwam, het slachtoffer onverhoeds 's avonds op een parkeerterrein benaderde, op de vlucht sloeg na het incident, het wapen wegmaakte, naar Spanje is gevlucht en na aanhouding de eerste tijd bij zijn verhoren heeft gezwegen. Dit past niet bij het door verdachte geschetste scenario waarin hij zich plotseling geconfronteerd zag met een wapen in de hand van het slachtoffer in zijn auto die verdachte daarmee bedreigde. Het hof schuift dit scenario gelet hierop terzijde en gaat bij de verdere beoordeling van de feiten uit van de verklaring van het slachtoffer dat verdachte degene was die het wapen heeft meegebracht. Bovenstaande suggereert dat het hof het verhaal van het slachtoffer eerder zou geloven.

Toch komt het wat betreft het vermeende schot gericht op het slachtoffer niet tot een veroordeling, omdat de zaak een bijzondere wending neemt. Niet alleen wordt het wapen niet gevonden op de plek waar verdachte het naar eigen zeggen

heeft weggegooid, maar belangrijker: op het parkeerterrein worden geen hulzen of kogels aangetroffen en aan en in de auto worden geen beschadigingen of schotopeningen aangetroffen. In de auto vindt men alleen bloedsporen maar geen enkel spoor van een projectiel, een inslag of een beschadiging. Niets. Tegen die achtergrond wordt het van groot belang of het letsel van de aangever kan worden beschouwd als een gevolg van het schieten met het wapen. Dat brengt ons bij de verklaringen van de verschillende deskundigen.

De forensisch arts komt tot de conclusie dat “het letsel kan passen bij de door het slachtoffer aangegeven toedracht”. Een andere deskundige, een forensisch patholoog, verklaart sterker. Hij constateert op basis van de preoperatieve röntgenfoto van de gewonde hand van het slachtoffer dat dergelijk letsel niet kan optreden wanneer er geen sprake zou zijn van een afgevuurd metalen projectiel.

Wanneer we deze twee uitspraken op de gebruikelijke manier Bayesiaans interpreteren, dan concluderen we dat de forensisch arts slechts één hypothese beschouwt en aangeeft dat de verwondingen daarbij passen. Dat is niet voldoende om een bewijssterkte te bepalen want er ontbreekt een tweede scenario. De patholoog stelt dat onder elke hypothese waarbij geen sprake is van een afgevuurd projectiel, deze verwondingen niet kunnen optreden. Als we dat geloven, dan sluit deze bevinding scenario's waarbij niet is geschoten dus uit en mogen we veronderstellen dat het slachtoffer inderdaad beschoten is.

Hier blijft het niet bij wat betreft de analyse van de precieze verwonding van de hand en de deeltjes die in de hand van het slachtoffer werden aangetroffen. Enkele jaren na het incident, in januari 2020, wordt het slachtoffer opnieuw geopereerd en daarbij wordt lichaamsvreemd materiaal veiliggesteld. Er is wel een kwestie dat dit materiaal enige tijd

in een niet-verzegeld potje is bewaard zodat er geen absolute zekerheid bestaat dat er niets met de inhoud is gebeurd. Hoe dit ook zij, de inhoud van dit potje wordt door het NFI onderzocht, met daarbij de volgende vraagstelling aan het NFI:

1. Kunt u aangeven wat de aard van deze deeltjes is?
2. Kunnen deze deeltjes afkomstig zijn van kruit- of patroonresten c.q. van een schietproces?
3. Indien vraag 2 bevestigend wordt beantwoord: met welk soort (vuur)wapen kan een dergelijk patroon verschoten zijn?
4. Wat kunt u vanuit uw deskundigheid in het kader van dit onderzoek voor het overige opmerken?

We merken op dat in vraag 2 opnieuw sprake is van het vragen naar bevindingen onder slechts één hypothese en dat het antwoord erop dus niet onmiddellijk informatief is. Omdat de vraag niet helemaal zuiver is, is het antwoord dat ook niet. De deskundige van het NFI antwoordt namelijk op 22 mei 2020 als volgt: “De looddeeltjes kunnen afkomstig zijn van een kogel (schietproces), maar kunnen ook afkomstig zijn van een ander loodhoudend voorwerp.”

De vraag die dit antwoord oproept, is natuurlijk onder welke hypothese een ander loodhoudend voorwerp de oorsprong van de loodhoudende deeltjes zou kunnen zijn. Het zou beter geweest zijn als hier de volgende twee hypothesen zouden zijn geformuleerd:

H_1 : De deeltjes zijn afkomstig van een schietproces.

H_2 : De deeltjes hebben een andere oorsprong.

De verwarring hierover neemt enige maanden na het deskundigenrapport van 22 mei 2020 toe, wanneer er op

20 november 2020 aanvullend gerapporteerd wordt. Kennelijk beschikte het NFI bij het eerste onderzoek niet over bepaalde medische informatie en was een aanvulling noodzakelijk.

We citeren nu uit de uitspraak (nadruk en cursivering origineel):

“Op 20 november 2020 is aanvullend gerapporteerd door [deskundige] van het NFI nadat gebleken was dat het NFI bij het eerste onderzoek niet beschikte over de medische informatie. Daarin staat onder meer het volgende

Vraag

Naar aanleiding van de conclusie 2 in het rapport van 22 mei 2020: hoe groot is de waarschijnlijkheid dat de loden deeltjes en het deeltje met lood, barium en antimoon tezamen afkomstig zijn van een schietproces?”

Het is niet helemaal duidelijk wie deze vraag heeft gesteld, maar we merken op dat een deskundige deze vraag helemaal niet kan beantwoorden. De deskundige kan helpen bij het onderzoeken of bepaald bewijsmateriaal beter bij de ene dan bij de andere hypothese past, maar kan geen uitspraak doen over de waarschijnlijkheid van de hypothesen zelf. Toch wordt die vraag hier gesteld.

Het antwoord van de deskundige maakt de verwarring alleen maar groter. We citeren:

“Met andere woorden, de mogelijk bron(nen) van de aangetroffen deeltjes kunnen afzonderlijk geduid worden zoals in de rapportage van 22 mei 2020 reeds is gedaan. Het totale sporenbeeld is te beperkt om hier verdere invulling aan te geven. Vanwege dit beperkte sporenbeeld, de verschillende mogelijke verklaringen voor het sporenbeeld

die niet van elkaar onderscheiden kunnen worden en de afwezigheid van referentiemogelijkheden zoals stukken van overtuiging die een dergelijk proces hebben doorgelopen, kan geen inschatting van de bewijskracht (waarschijnlijkheid) gedaan worden.”

De deskundige heeft gedeeltelijk gelijk: er kan geen bewijskracht worden vastgesteld, maar dat is vooral omdat er geen twee expliciete hypothesen naast elkaar worden gezet. Het feit dat de deskundige “bewijskracht” identificeert met “waarschijnlijkheid” doet vermoeden dat er verwarring bestaat over wat er nu precies wel of niet waarschijnlijk is. Waar doet ze precies een uitspraak over? Het gevaar bestaat dat gematigde conclusies over het bewijsmateriaal, zoals bijvoorbeeld de bewering dat de verwondingen afkomstig “kunnen” zijn van een kogel, als ontlastend worden geïnterpreteerd, terwijl dat bij gebrek aan een alternatief scenario helemaal niet zo hoeft te zijn. Als een dergelijke verwonding *kán* optreden onder dit scenario, en er tegelijkertijd geen ander redelijk scenario te bedenken is waarbij dat ook mogelijk zou zijn, dan is de bewijskracht van deze aanwijzing alsnog erg groot. Hoe een en ander hier geïnterpreteerd wordt, blijft echter enigszins in het vage, precies omdat men niet Bayesiaans redeneert.

De deskundige vakgebied Wet wapens en munitie

Naast de medische kant van de zaak is er ook advies gezocht van een deskundige op het gebied van wapens en munitie. Hem werd de volgende vraag gesteld: “Kan het letsel van het slachtoffer zijn veroorzaakt door een vuurwapen met een scheikundige ontploffing?” Ter zitting van het hof op 25 november 2020 is deze wapendeskundige gehoord over zijn bevindingen in de onderhavige zaak. Hij is daarbij geïn-

formeerd over het schotrestenonderzoek van het NFI, en zijn verklaring luidde als volgt:

“Binnen de geschetste scenario’s door verdachte en slachtoffer kunnen de verwondingen van het slachtoffer niet door een vuurwapen zijn veroorzaakt. Als er met een ‘scherp’ wapen (met kogels) was geschoten, dan was er zeker schade in de auto gevonden in de vorm van een kogelgat in het dashboard, de deur of de ruit of schampschade aan dashboard, stuur of de kleding van het slachtoffer en dat is niet het geval. Als er met een gas-/alarmrevolver zou zijn geschoten (losse flodders, geen kogels) dan ontstaan er – binnen de geschetste scenario’s – niet de verwondingen zoals het slachtoffer die heeft bekomen. Een dergelijke revolver produceert een mondingsvlam en een knal; het maakt geen kogelgaten, hoogstens schroeiplekken.”

Deze verklaring is buitengewoon: volgens deze expert zijn de bevindingen namelijk onder *beide* scenario’s onmogelijk. Dat gebeurt niet vaak. Het hof noemt deze getuigenis “navolgbaar”, maar wij hebben daar toch vragen over. Het is immers een feit dat het slachtoffer ter plekke gewond is geraakt, en wel op zo’n manier dat deskundigen dat volkomen consistent achtten met een schotwond. Volgens de forensisch patholoog kunnen de verwondingen zelfs op geen andere manier ontstaan zijn. Tel daarbij op dat het hof de verklaring van de verdachte al vanaf het begin niet geloofde. Dan rijst de vraag toch op hoe een en ander dan verklaard kan worden. Het slachtoffer heeft de schotwond niet verzonnen, dus welke alternatieve verklaring is er dan nog te bedenken?

Het hof redeneert niet op Bayesiaanse wijze – dat is zichtbaar door de vragen die het stelt – en dat leidt tot verwar-

ring. De verklaring van de wapendeskundige suggereert dat er een derde scenario bekeken moet worden: een scenario waarin helemaal niet is geschoten. We hebben dan de situatie waarin de verwondingen volgens de patholoog niet ontstaan kunnen zijn onder dat derde scenario en waarin de patholoog beweert dat de afwezigheid van schade aan de auto onder de andere twee scenario's onmogelijk is. Het is dan aan het hof om de tegenstrijdige verklaringen van de verschillende deskundigen wel of niet te accepteren.

Een waarschuwing is hier op zijn plaats. Het is zeer goed mogelijk dat het afwezig zijn van schade aan de auto heel onwaarschijnlijk is onder beide scenario's van de verdachte en het slachtoffer. Echter, als daar al een zinvolle uitspraak over gedaan kan worden, dan zal deze noodzakelijkerwijs niet erg onderscheidend zijn – daarvoor ontbreken immers de nodige aanwijzingen.

Conclusie van het hof

Het hof besluit dat het niet goed weet wat zich heeft afgespeeld en spreekt verdachte vrij van het eerste ten laste gelegde feit. De verdachte wordt wel veroordeeld voor bedreiging met een wapen. Hierbij zijn er nog enkele details uit de conclusie van het hof waar we graag de aandacht op willen vestigen.

Allereerst schrijft het hof het volgende over de getuigenis van de eerder genoemde patholoog (onze nadruk):

“Ten tweede betreft het onderzoek van de deskundige [...] alleen het geconstateerde letsel *en niet meer dan dat*. De bevindingen van deze deskundige op het terrein van de forensische pathologie zijn – met inachtneming van zijn deskundigheidsgebied – gebaseerd op onderzoek van de

letsels zonder hierbij omstandigheden van een mogelijke toedracht te betrekken. Het gebruiken van deze bevindingen zonder dat deze zijn geplaatst binnen het kader naar de ontstaansoorzaak van de letsels zijn *in die zin gemankeerd*.”

De diskwalificatie dat het onderzoek van de patholoog niet meer dan het letsel betreft, kunnen wij niet goed volgen. Dat is namelijk precies wat een patholoog moet doen. Dat het gebruiken van deze bevindingen “gemankeerd” is als ze niet in context zijn geplaatst, is enigszins misleidend: dat is immers altijd het geval met bevindingen van experts. De Bayesiaanse redeneerstructuur geeft precies aan hoe een getuigenis van een expert past in het geheel, en in de woorden van het hof is dat altijd “gemankeerd”. Sterker nog, dat móét het ook zijn, want dat is precies de taak van de patholoog. De patholoog beperkt zich tot zijn domein, en het is de taak van het hof om dat in te passen in het grotere geheel – dat kan en mag je de patholoog niet kwalijk nemen. Integendeel, het pleit juist vóór hem.

In zijn overwegingen schrijft het hof ook nog dit:

“Het NFI concludeert dat de aangetroffen looddeeltjes afkomstig kunnen zijn van een kogel, maar ook afkomstig kunnen zijn van een ander loodhoudend object.”

Het is evident dat het hof deze conclusie vrij letterlijk opvat, maar vanuit het NFI is nauwelijks iets anders te verwachten. Het NFI sluit een kogel niet uit, en vanuit het perspectief van het NFI is dat logisch en navolgbaar. Immers, het NFI kent de context niet, of liever gezegd, het hoeft zich met die context niet bezig te houden.

Het hof moet dat wel, dus het feit dat het NFI een andere bron niet uitsluit moet door het hof in context gezien worden. Het hof zou dan kunnen concluderen dat de bevindingen

gen van het NFI prima passen bij bijvoorbeeld het scenario van het slachtoffer. Het hof echter, concludeert uiteindelijk als volgt:

“Hoewel het hof zich, gezien al het voorgaande, geconfronteerd ziet met verklaringen van slachtoffer en getuigen die wijzen op een schietincident en iets dat in zijn aard past bij schieten met een (scherp) wapen is er, met name gezien de beperkte resultaten van de verrichte forensische onderzoeken teveel onzekerheid om tot de vergaande conclusie te komen dat verdachte heeft geschoten.”

Onze conclusie

Het hof heeft in deze zaak niet expliciet met scenario's gewerkt. In deze zaak werd tegenstrijdig gerapporteerd door de deskundigen: de één verklaarde dat de verwonding alleen maar door een schietproces veroorzaakt kon zijn, de ander dat het onmogelijk is dat er geen schade aan de auto wordt gevonden als er geschoten is. Het is begrijpelijk dat dit voor de nodige onzekerheid heeft gezorgd en dat het hof heeft overwogen dat het onvoldoende duidelijk is wat zich precies heeft afgespeeld. Er zijn te veel onzekerheden om tot een veroordeling te komen.

Het is speculeren wat er gebeurd zou zijn als het hof de verklaring van de patholoog had omarmd en dus had geconcludeerd dat er wel geschoten móét zijn. In die omstandigheid is de verklaring van de wapendeskundige noch ontlastend, noch belastend omdat de bevindingen onder beide overgebleven scenario's erg onverwacht zijn.

Een Bayesiaanse benadering zou hier erg nuttig geweest zijn, niet alleen voor het denkproces over de zaak zelf maar ook voor de navolgbaarheid van de conclusies van het hof.

Het hof zou dan andere vragen hebben gesteld en zou antwoorden die niet Bayesiaans zijn ook als zodanig hebben herkend. Dat zou alles veel overzichtelijker en waarschijnlijk ook beter gemotiveerd hebben gemaakt. Het is zelfs mogelijk dat dit tot een andere conclusie zou hebben geleid, maar dat is niet het punt dat we hier willen maken.

DE ZAKEN ECLI:NL:RBNHO:2014:7224 EN
ECLI:NL:GHAMS:2017:1801; DE OVERLEDEN VROUW IN
HET TRAPPENHUIS

Tenlastelegging, achtergrond en verklaringen van de
verdachte

Deze zaak draait om een man die verdacht wordt van betrokkenheid bij de dood van zijn zwangere ex-vriendin, moeder van hun beider zoon van vier. Verdachte en slachtoffer hebben al enige tijd een ernstig verstoorde relatie als op een dag het slachtoffer wordt aangetroffen in het trappenhuis van de flat waar ze woont, liggend op haar rug en besmeurd met bloed. De medebewoner die haar vindt belt 112, een andere medebewoner start reanimatie. Ambulancemedewerkers die snel ter plaatse zijn, constateren dat er geen hartslag is maar krijgen het hart weer op gang. Een klein halfuur later wordt het slachtoffer naar het ziekenhuis gebracht. Na onderzoek in het ziekenhuis blijkt het kind in de baarmoeder te zijn overleden. Na een spoedkeizersnede overlijdt ook de moeder.

De rechtbank Noord-Holland veroordeelt de verdachte – mede op basis van zijn eigen verklaring waaruit blijkt dat hij bij haar is geweest in het trappenhuis en nog geprobeerd heeft haar te reanimeren maar uiteindelijk is vertrokken – voor het in hulpbehoevende toestand achterlaten van het slachtoffer ten gevolge waarvan zij is overleden (“het met

mishandeling gelijkgestelde opzettelijk benadelen van de gezondheid, terwijl het feit de dood ten gevolge heeft”).

In hoger beroep echter trekt de verdachte zijn eerdere verklaring in, zegt niet in het trappenhuis te zijn geweest en ontkent hij alle betrokkenheid bij de dood van zijn ex. Het eerste feit op de tenlastelegging (die vrijwel gelijk is aan die uit de eerste aanleg) beschuldigt de verdachte ervan zijn ex-vriendin van de trap te hebben geduwd en/of haar te hebben verstikt, en is ten laste gelegd in drie varianten: doodslag van moeder en ongeboren kind, zware mishandeling de dood ten gevolge hebbend en mishandeling de dood ten gevolge hebbend. Het tweede ten laste gelegde feit houdt in dat de verdachte het slachtoffer gewond onder aan de trap heeft achtergelaten en heeft nagelaten adequate hulp te verlenen, ten gevolge waarvan zij is overleden. Het hof spreekt verdachte vrij van alle beschuldigingen.

Medische overwegingen

Als de vrouw wordt gevonden in het trappenhuis, ligt zij op haar rug met haar bovenlichaam nog op de trap en haar benen op het tussenstuk tussen de trappen van de tweede en derde verdieping. Op de handen en het gezicht van het slachtoffer zit bloed en zij ligt met haar hoofd in een plas vocht en bloed. Verschillende (typen) bloedsporen worden ook aangetroffen rondom het trapbordes op de derde verdieping waar de voordeur van het slachtoffer aan grenst. Het hof stelt vast dat het slachtoffer in de periode tussen 11.00 uur en 12.28 uur gewond moet zijn geraakt en terecht is gekomen op de plaats waar ze is gevonden.

Het lichaam van het slachtoffer is door respectievelijk de gemeentelijke lijkschouwer, een patholoog en een forensisch arts onderzocht. Wat na al dat onderzoek vaststaat, is

dat zij verschillende (niet heel zware) verwondingen had op haar lichaam en een scheur op haar kruin. Het letsel past bij een val van de trap maar is voor de dood niet van betekenis geweest; letsels door derden toegebracht kunnen niet worden uitgesloten.

De forensisch arts rapporteert dat slechts in zeer beperkte mate iets van de doodsoorzaak is te zeggen, maar dat het gezien de hoofdwond van het slachtoffer en de lichtstijve pupillen voor de hand ligt om te veronderstellen dat het hoofdletsel heeft geleid tot een ademhalings- en circulatiestilstand. Voor wat betreft de doodsoorzaak van het ongeboren kind wordt geconstateerd dat het kind naar alle waarschijnlijkheid is overleden als gevolg van de instabiele bloedcirculatie bij de moeder.

Gangen en gedrag van de verdachte

Vast staat dat de verdachte geen alibi heeft voor het grootste deel (namelijk tussen ca. 11.00 en iets na 12.00 uur) van de ongeveer anderhalf uur waarin het slachtoffer gewond moet zijn geraakt. Gezien zijn gangen die dag, kan hij in die periode in de flat zijn geweest. Uit camerabeelden is naar voren gekomen dat hij op twee tijdstippen in de ochtend in zijn auto enige tijd achter de auto van zijn ex-vriendin heeft gereden. In de middag is hij anders dan gebruikelijk, een uur te laat voor een werkafspraak om 12.30 uur en meldt hij dit pas tien minuten na het afgesproken tijdstip. Daarnaast valt zijn afwijkende gedrag, sterke geur (aftershave) en kortere aanwezigheid dan anders op bij zijn collega's.

Om 12.08 en 12.09 uur belt de zus van verdachte hem tevergeefs om te zeggen dat zijn zoon niet uit school is opgehaald door zijn moeder. De verdachte belt zijn zus om 12.14 uur terug en bespreekt dat zij hun zoon mee zal nemen. Ver-

volgens zoekt de verdachte die dag op meerdere momenten telefonisch en via e-mail contact met zijn ex, over het niet-ophalen van hun zoon van school. Ook stuurt hij via Facebook een privéchatbericht naar zijn schoonmoeder, waarin onder meer staat “die troela is niet komen opdagen”.

Tas en telefoon van slachtoffer

Het slachtoffer droeg altijd een tas bij zich waarin zij haar sleutels en portemonnee bewaarde. Ook had zij altijd haar mobiele telefoon bij zich. Bij het aantreffen van het slachtoffer in het trappenhuis, en ook bij onderzoek nadien, zijn geen persoonlijke spullen van haar aangetroffen. Er zijn na 09.54 uur geen contacten in de historische verkeersgegevens van het telefoonnummer van het slachtoffer geregistreerd. Wel heeft het WhatsApp-account van slachtoffer (voor het laatst) contact gemaakt met de WhatsApp-server om 12:19 uur. Het kan niet worden vastgesteld van welke zendmast daarbij gebruik is gemaakt of dat dit contact met WhatsApp via een wifi-netwerk is gelegd.

Relatie tussen verdachte en slachtoffer

De verdachte en het slachtoffer hadden enige tijd een ‘los-vast’-relatie. Ze kregen een zoon, woonden samen, slachtoffer vertrok, en raakte toch weer zwanger. Uit sms-verkeer wordt duidelijk dat de verdachte en het slachtoffer in de circa vier maanden voorafgaand aan de dood van het slachtoffer een gespannen verhouding hadden. Zij maakten ruzie over de omgang met hun zoon en over zijn paspoort dat de verdachte kennelijk niet aan het slachtoffer wilde afgeven. Dit leidde ertoe dat de verdachte door de rechtbank Den Haag op

straffe van een dwangsom werd veroordeeld tot onder meer het nakomen van de omgangsregeling en tot het afgeven van het paspoort. De verdachte heeft dwangsommen verbeurd tot het maximale bedrag, waarna de deurwaarder loonbeslag legde onder de werkgever van de verdachte, en de verdachte meerdere malen heel boos de deurwaarder heeft gebeld. Voorts blijkt uit de verklaringen van de moeder en de broer van het slachtoffer en uit sms-berichten dat de verdachte het slachtoffer ‘begluurde’: hij zat in zijn auto voor haar flat.

Verdere gang van zaken

In de eerste politieverhoren zegt de verdachte het slachtoffer al een tijd niet te hebben gezien. Geconfronteerd met de camerabeelden waaruit blijkt dat hij achter haar aanrijdt, verandert hij in het zevende verhoor zijn verklaring. Deze houdt in dat hij op 30 oktober naar de flat van zijn ex is gegaan om met haar te praten over hun zoontje, dat hij haar in het trappenhuis tegenkomt en dat zij daar een woordenwisseling krijgen. Zijn ex-vriendin wordt emotioneel, zwaait met een plastic tasje naar hem waarop dat tasje achter hem op de grond terecht komt. Daarop holt zijn ex naar boven en gaat hij naar beneden met de bedoeling weg te gaan. Als hij beneden is hoort hij een doffe klap en een kreun. Hij blijft even wachten in de verwachting dat hij de deur van de woning van het slachtoffer open en dicht hoort gaan maar gaat weer naar boven als hij niets hoort. Hij treft zijn ex bewegingloos aan op het plateau voor haar woning met haar hoofd in een plas bloed en met een been over de trap. Hij haalt haar hoofd weg bij de plas bloed en begint met reanimeren. Tijdens het reanimeren begint ze echter van de trap te glijden. Hij heeft geprobeerd de reanimatie op de trap door te zetten. Toen hij iemand hoorde bellen is hij in paniek het

flatgebouw uitgerend. Hij ging ervan uit dat het slachtoffer toen al was overleden. Vervolgens heeft hij een knop omgezet en heeft hij maar net gedaan of er niets gebeurd was en of hij van niets wist. Hij was bang dat hij anders zou worden opgepakt en dit zou – samen met het overlijden van zijn moeder – te veel zijn voor zijn zoontje. Hij nam aan dat er inmiddels hulpdiensten ter plaatse waren, zodat er voor hem geen noodzaak was om hulp in te roepen. Bovendien had dit in zijn visie toch geen zin meer gehad, nu hij ervan uitging dat het slachtoffer al was overleden. Tijdens de behandeling ter zitting voert de raadsman van de verdachte aan dat de verdachte wel degelijk heeft getracht adequate medische hulp te verlenen door direct over te gaan tot reanimatie. Er is geen causaal verband tussen het feit dat de verdachte de flat verliet zonder medische hulp in te schakelen en het overlijden van het slachtoffer. De verdachte kon er redelijkerwijs van uitgaan dat er op dat moment geen succesvolle levensreddende handelingen meer zouden kunnen worden toegepast door hulpdiensten.

Voorafgaand aan de eerste regiezitting in hoger beroep laat de raadsman van verdachte aan het hof weten dat de verdachte afstand neemt van zijn verklaring dat hij in het trappenhuis is geweest. Als reden voor die eerdere verklaring geeft de verdachte op dat hij in zijn cel bezoek had gekregen van de Raad voor de Kinderbescherming over een uithuisplaatsing van zijn zoon. Hij dacht dat hij snel naar huis zou kunnen en weer voor zijn zoon zou kunnen zorgen als hij duidelijkheid zou verschaffen over wat er gebeurd was in het trappenhuis. De verklaringen over het gebeurde in het trappenhuis zouden geheel zijn gebaseerd op het voorgeleidsdossier dat hij ten tijde van de politieverhoren tot zijn beschikking had. De verdachte ontkent nu elke betrokkenheid bij het overlijden van het slachtoffer. Namens de ver-

dachte wordt aangevoerd dat het zeer waarschijnlijk is dat het slachtoffer pas rond 12.25 uur uit haar huis is vertrokken, terwijl de verdachte toen allang met de auto op weg was naar zijn werk in Alkmaar. Verder heeft de verdediging aangevoerd dat de dood waarschijnlijk niet door van buiten komend geweld veroorzaakt is en dat een alternatief scenario van natuurlijk overlijden beter bij het dossier past dan de stelling van het openbaar ministerie dat de verdachte de dood moet hebben veroorzaakt. Het slachtoffer is volgens de verdediging overleden als gevolg van hartfalen en dat is hoogstwaarschijnlijk veroorzaakt door een val van de trap.

Maar, zo stelt het hof vast, behalve de door hem later ingetrokken verklaringen over zijn aanwezigheid in het trappenhuis is er geen enkel technisch bewijs dat hij daar was. Nader DNA-onderzoek dat in hoger beroep is gedaan naar de aanwezigheid van DNA van de verdachte op de kleren en onder de nagels van het slachtoffer, heeft niets opgeleverd. Het hof gaat ervan uit dat de verklaring van de verdachte over de reanimatiepogingen leugenachtig is, maar de reden daarvoor kan gelegen zijn in een andere omstandigheid dan dat hij enig strafbaar handelen in het trappenhuis heeft willen verbloemen. Zelfs als ervan uit zou worden gegaan dat de verdachte op de desbetreffende dag in het trappenhuis is geweest tussen grofweg 11.00 en 12.00 uur, is dit onvoldoende om tot een veroordeling te komen omdat onvoldoende duidelijk is wat daar precies is gebeurd. Kortom, zo sluit het hof af, als alle bewijsmiddelen in onderling verband en in samenhang worden beschouwd, blijven er nog te veel onzekerheden over die aan de hand van de inhoud van het dossier niet kunnen worden weggenomen. De stappen die moeten worden gezet voor enige bewezenverklaring, vereisen te veel speculatie over wat er gebeurd moet zijn.

Een Bayesiaanse analyse

Zou Bayesiaans redeneren in deze zaak de rechter nog helpen kunnen helpen bij het vormen van zijn overtuiging, en zo ja hoe? Laten we beginnen met te zeggen dat wat ons betreft het hof duidelijk uiteen heeft gezet waarom een veroordeling er in deze zaak toch niet in zit. Anders gezegd: selectie en waardering van het bewijs zijn inzichtelijk gemaakt, en begrijpelijk. Een Bayesiaanse analyse van een zaak kan niettemin ondersteuning bieden bij het benoemen, ordenen en waarderen van informatie – en kan in deze zaak laten zien waarom de motivering van het hof vanuit het perspectief van logisch redeneren een goede is.

De eerste stap in elke toepassing van de regel van Bayes is dan om alle denkbare relevante hypothesen op te stellen. Hiervoor gaven we reeds aan dat Bayesiaans redenen inhoudt dat onderzocht wordt in hoeverre de bewijsmiddelen *onderscheidend* zijn voor de verschillende scenario's of hypothesen. De regel van Bayes is simpelweg niet te gebruiken zonder het opstellen van relevante hypothesen. De te formuleren hypothesen beschrijven niet noodzakelijk alle gebeurtenissen, maar vatten de belangrijkste gebeurtenissen samen. In deze zaak zouden we bijvoorbeeld de volgende hypothesen kunnen onderscheiden:

H_1 : Verdachte heeft slachtoffer van de trap geduwd en is daarna weggegaan. Slachtoffer is vervolgens aan haar verwondingen overleden.

H_2 : Het slachtoffer is in aanwezigheid van de verdachte van de trap gevallen. De verdachte heeft haar daarna achtergelaten, al dan niet na haar eerst te hebben gereanimeerd. Slachtoffer is aan hartfalen overleden als gevolg van de val van de trap.

H_3 : Verdachte is helemaal niet aanwezig geweest in het portiek ten tijde van de gebeurtenissen. Slachtoffer is aan hartfalen overleden als gevolg van een val van de trap.

Er zijn natuurlijk meer scenario's te bedenken. Zo is het niet onmogelijk dat de verdachte het slachtoffer van de trap heeft geduwd, en dat hij daarna heeft geprobeerd haar te reanimeren. Ook is het niet uitgesloten dat iemand anders dan de verdachte het slachtoffer van de trap heeft geduwd. Deze scenario's lijken ons echter dermate vergezocht dat ze niet serieus in aanmerking komen.

Duidelijk is alvast dat de bewijsmiddelen in deze zaak schaars zijn. Eigenlijk hebben we alleen de verwondingen van het slachtoffer, de slechte relatie tussen de verdachte en het slachtoffer, en de verklaringen van de verdachte.

Er ging al lange tijd een behoorlijke dreiging uit van de verdachte richting slachtoffer. Binnen die context is het a priori niet uitgesloten, en zelfs tot op zekere hoogte te begrijpen, dat H_1 waar zou zijn. A priori is dat scenario zeker niet volkomen ongeloofwaardig. Daar komt bij dat de andere hypothesen H_2 en H_3 een val van de trap – zomaar uit het niets – vereisen, en dat is hoe dan ook een behoorlijk opmerkelijke gebeurtenis.

Na het evalueren van prior kansen op de hypothesen, kijken we naar het verdere bewijsmateriaal. Op basis van het bewijs dat aanwezig is, komen we dan al snel tot een wellicht opmerkelijke conclusie. De verwondingen van het slachtoffer en de manier waarop ze is aangetroffen passen even goed bij elk van de drie genoemde hypothesen. De verwondingen van het slachtoffer passen net zo goed bij een ongelukkige val als bij een duw van de trap. Dat betekent dat de verwondingen en de omstandigheden waaronder het slachtoffer aangetroffen werd, simpelweg niet onderscheidend zijn voor de drie genoemde hypothesen, en dus geen bewijs leveren voor een

van drie genoemde hypothesen ten opzichte van elkaar. De advocaat van de verdachte beweerde zelfs dat het scenario van een natuurlijke dood beter past bij de bevindingen dan het scenario waarin de verdachte het slachtoffer van de trap heeft geduwd, maar deze stelling lijkt gratis en wordt ook niet onderbouwd. Wel gebruikt de verdediging hierbij de juiste redeneerstructuur door te kijken in hoeverre bewijsmiddelen passen bij de diverse hypothesen.

Verder hebben we ook nog de verklaringen van de verdachte. Deze zijn leugenachtig, zoals het hof ook aangeeft, en er lijkt geen reden te zijn om aan welke verklaring dan ook geloof te hechten. Het is op dit punt dat het verschil tussen bewijs en overtuiging belangrijk wordt. Het feit dat de verdachte zijn verhaal steeds verandert, maakt dat wij (en wij vast niet alleen) er persoonlijk van overtuigd zijn dat hij op zijn minst iets met de hele gang van zaken te maken heeft. Deze overtuiging leidt echter niet tot een bewezenverklaring – hoewel de grens tussen persoonlijke overtuiging en bewezenverklaring hier flinterdun lijkt.

We zagen al dat er goede argumenten zijn om de prior overtuiging van H_1 niet heel klein te nemen, gezien de slechte relatie tussen verdachte en slachtoffer. In een geval waarin de prior verdenking al groot is, is weinig nodig om tot een wettige bewijsverklaring te komen. Bij een veel lagere prior is dat veel lastiger.

In de onderhavige casus zal voor veel mensen de prior verdenking van de verdachte groot zijn en zal er weinig verder bewijs nodig zijn om de verdachte aan de val te kunnen koppelen en te kunnen veroordelen voor het eerste feit op de tenlastelegging. Maar dan moet dat kleine beetje wettig bewijs er wel zijn. Om tot de overtuiging te komen dat verdachte niet de waarheid vertelt en in ieder geval ergens schuldig aan zal zijn, is zijn gedrag waarschijnlijk wel voldoende. Het wis-

selende en leugenachtige verklaren van de verdachte draagt bij aan de prior verdenking maar levert niet dat vereiste bewijs. Om tot een wettige bewezenverklaring te komen, zal er iets moeten zijn dat op een wettige manier bijdraagt aan de prior verdenking.

Hier dringt zich een vergelijking op met de eerste casus uit dit hoofdstuk. Ook daar is het onduidelijk gebleven wat zich precies heeft afgespeeld, maar in tegenstelling tot de huidige casus was daar wel bewijs dat de verdachte verbond met de plaats delict: het DNA op de deurknop van de voordeur van het slachtoffer. Bij afwezigheid van dergelijk bewijs hebben we hier te maken met een hoge a priori verdenking, maar verder geen enkele concrete aanwijzing die de verdachte verbindt met het overlijden van zijn ex. In dat licht bezien is de beslissing van het hof zeker navolgbaar, ook vanuit Bayesiaans perspectief.

TEN SLOTTE

De keuze voor juist deze drie zaken is tamelijk willekeurig tot stand gekomen. Ons archief bevat veel meer zaken die vanuit een Bayesiaans perspectief interessant zijn maar we denken dat we met deze drie ons punt wel gemaakt hebben.

We hopen dat dit boek het eventuele koudwatervrees bij de lezer weg heeft kunnen nemen. Het is misschien een geruststelling dat het geenszins onze bedoeling is om oordelen te mathematiseren. We willen ze wel graag structureren, en dat is heel wat anders.

We pleiten voor het opnemen van deze materie in de opleiding van elke jurist. Het is uiteraard geen wondermiddel maar Bayesiaans redeneren kan, zoals we hebben laten zien, bijdragen aan betere en beter onderbouwde beslissingen.

Strafzaken gaan van nature gepaard met onzekerheid. In dit boek leggen Ronald Meester en Lonneke Stevens uit hoe een zogenaamd Bayesiaans gezichtspunt nuttig kan zijn bij het correct redeneren onder die onzekerheid. Een Bayesiaanse benadering betekent zeker niet dat de rechter zich uitlevert aan statistiek, getallen en berekeningen. De essentie van Bayesiaans redeneren is een denkmodel dat desgewenst ook formeel wiskundig benaderd kan worden. Voor de praktijk is die wiskunde echter niet nodig en vaak zelfs contraproductief. Dit boek bevat dus ook geen formules. Het is een boek voor de juridische praktijk waarin de auteurs aan de hand van casuïstiek laten zien hoe een Bayesiaanse benadering tot beter gemotiveerde uitspraken kan leiden en hoe logische fouten kunnen worden voorkomen. Als je wel eens van deze benadering hebt gehoord maar misschien koudwatervrees of gewoon twijfels hebt, dan is dit boek een must. De auteurs nemen de lezer bij de hand en laten zien dat Bayesiaans redeneren een natuurlijke manier is om bewijs in strafzaken te verwerken. Bovendien is het een denkwijze die goed aansluit bij de strafrechtelijke bewijsregels. Als je deze Bayesiaanse bril eenmaal opgezet hebt, wil je hem nooit meer af doen.

Ronald Meester is hoogleraar Waarschijnlijkheidsrekening aan de Vrije Universiteit Amsterdam en doet onder meer veel onderzoek naar forensische kansrekening en statistiek. Hij wordt regelmatig als expert ingezet bij strafzaken en doceert bij SSR. Hij publiceert regelmatig over de rol van modellen in de wetenschap en de wereld, de filosofie van kansrekening en statistiek, en over AI.

Lonneke Stevens werkt bij de Nationale Politie en richt zich daar op de juridische inbedding van datagedreven werken. In haar vorige betrekking als hoogleraar Straf(proces)recht aan de Vrije Universiteit Amsterdam deed zij onderzoek naar diverse strafprocesrechtelijke vraagstukken, waaronder bewijzen in het strafrecht.

ISBN 978-94-93458-03-1



9 789493 458031 >

 **WJS**
UITGEVERS