

De kenniskloof verkend: een onderzoek naar de waardering van (complex) DNA-bewijs in strafzaken

De waardering van bewijs vormt al jaren een heet hangijzer in strafzaken zoals blijkt uit onderzoek dat laat zien dat rechters en andere procesdeelnemers grote moeite hebben bewijsmateriaal op juiste wijze te interpreteren. Uit de literatuur komt naar voren welke redeneerfouten de rechter daarbij kan maken, maar niet duidelijk is hoe vaak deze nu eigenlijk (nog) voorkomen. De afgelopen jaren zijn al de nodige initiatieven genomen ter verbetering van de kwaliteit van rapportages en het begrip hiervan door partijen in het strafproces. In dit onderzoek is gekeken naar de kwaliteit van de rechterlijke uitspraken ten aanzien van de waardering van DNA-mengprofielen. Daarbij is in 27 van de 78 onderzochte zaken (35%) de bewijskracht van het gerapporteerde DNA-onderzoek verkeerd geïnterpreteerd en is er sprake van de zogenaamde prosecutor's fallacy. In nog eens 28 van de 78 zaken is geconstateerd dat de rechter zich hult in vage of onvolledige formuleringen (36%). Dit roept de vraag op of, en zo ja hoe, het begrip van rechters en de kwaliteit van de rechterlijke motiveringen van uitspraken (verder) kunnen worden verbeterd.

1. Inleiding

Sinds de introductie in de jaren tachtig van de vorige eeuw heeft DNA-bewijs in strafzaken een hoge vlucht genomen.¹ DNA-bewijs kan – in vergelijking met andersoortig, meer traditioneel bewijsmateriaal (zoals getuigenbewijs) – een heel hoge bewijskracht hebben en wordt gezien als relatief ‘hard’ bewijs² vanwege de natuurwetenschappelijke methode die aan het vergelijkend onderzoek ten grondslag ligt en de lage frequenties waarmee combinaties van DNA-kenmerken in populaties voorkomen.³ Onderzoek uit 2009 laat echter zien dat verschillende partijen in het strafproces DNA-rapportages en de wijze waarop hierin wordt geconcludeerd niet altijd goed begrijpen en bij de interpretatie fouten maken.⁴ In de afgelopen jaren zijn een aantal initiatieven genomen om de kwaliteit van rechterlijke uitspraken op dit punt te verbeteren.⁵ De vraag is of deze initiatieven voldoende vruchten hebben afgeworpen. In de literatuur zijn immers nog steeds voorbeelden te vinden waarin kritiek wordt geleverd op de wijze waarop de strafrechter het DNA-bewijs heeft gewaardeerd en in dat opzicht gemaakte bewijsredeneringen worden blootgelegd.⁶ Een breder beeld van de omvang van de kenniskloof thans ontbreekt evenwel.

In het kader van een scriptietraject aan het Instituut voor Strafrecht en Criminologie van de Universiteit Leiden is door de eerste auteur van dit artikel onderzoek gedaan naar de kwaliteit van rechterlijke bewijsoverwegingen met betrekking tot complex DNA-bewijs. Daarbij zijn 78

rechterlijke uitspraken waarbij het ging om DNA-mengprofielen, geanalyseerd.⁷ In het bijzonder is gekeken naar het type redeneerfouten dat door rechters in hun uitspraken wordt gemaakt bij dit soort bewijs en de frequentie waarmee deze fouten voorkomen. In dit artikel wordt nader gereflecteerd op de bevindingen van dat onderzoek.

Het onderzoek waarvan dit artikel verslag doet, heeft zich als gezegd primair gericht op de waardering van DNA-mengprofielen. Een DNA-mengprofiel is een DNA-profiel dat resulteert uit één spoor maar waarbij dat spoor bestaat uit celmateriaal van meerdere personen. Door technologische ontwikkelingen kunnen steeds kleinere sporen DNA-profielen opleveren, waarbij het vaak gaat om mengprofielen.⁸ Hoewel die ontwikkelingen de mogelijkheden voor DNA-onderzoek uitbreiden, maken die nieuwe mogelijkheden de interpretatie van de

* Mr. K. Twisk is masterstudent Forensische Criminologie aan de Faculteit der Rechtsgeleerdheid, Universiteit Leiden.

** Mr. dr. M.J. Dubelaar is universitair hoofddocent bij de sectie Strafrecht & Criminologie aan de Faculteit der Rechtsgeleerdheid van de Radboud Universiteit.

*** Prof. dr. ir. C.E.H. Berger is Hoogleraar Criminalistiek bij het Instituut voor Strafrecht & Criminologie, Universiteit Leiden, en Principal Scientist bij het Nederlands Forensisch Instituut.

1. A.J. Meulenbroek, *De Essenties van forensisch biologisch onderzoek. Humane biologische sporen en DNA* (5e druk), Zutphen: Uitgeverij Paris 2009. Zie ook: D.J.C. Aben, 'Het geluid van stille getuigen. Over de toepassing van forensisch bewijs in de Nederlandse strafrechtspleging', in: S. Van Overbeke & D. Aben, *Forensisch bewijs in het Nederlandse en Belgische strafrecht* (Preadvies NVVS 2018), Oosterwijk: Wolf Legal Publishers 2018, p. 1-127, meer in het bijzonder p. 33-57.
2. M. Malsch e.a., 'Hoe hard is DNA-bewijs? Internationaal-vergelijkend onderzoek naar de interpretatie van DNA-profielen', *NJB* 2016, afl. 18, p. 1261-1266.
3. Die hoge bewijskracht geldt echter alleen voor de vraag van wie het DNA afkomstig is en in veel mindere mate voor de vraag hoe het DNA-materiaal daar is terechtgekomen.
4. J.W. de Keijser e.a., *Bijkans begrepen? Feitelijke en vermeend begrip van forensische deskundigenrapportages onder rechters, advocaten en deskundigen*, Den Haag: Boom Juridische uitgevers 2009.
5. Zie hierover in meer detail paragraaf 5.
6. Zie onder meer de uitspraken die zijn becommentarieerd door Jaap de Groot in het tijdschrift *Expertise en Recht* (in de rubriek 'Uit de praktijk'). Een deel daarvan is recent gebundeld verschenen. Zie J. de Groot, *DNA in het strafrecht: Hoe moet het niet? Hoe moet het wel?*, Zutphen: Uitgeverij Paris 2018.
7. Zie voor een toelichting op de selectie paragraaf 3.
8. A.J. Meulenbroek, *De Essenties van forensisch biologisch onderzoek. Humane biologische sporen en DNA* (5e druk), Zutphen: Uitgeverij Paris 2009, p. 125.

resultaten niet eenvoudiger.⁹ De keuze om het onderhavige onderzoek te beperken tot DNA-mengprofielen was echter vooral ingegeven door praktische overwegingen, nu met behulp van de zoekterm 'DNA-mengprofiel' eenvoudigweg alle zaken in 2017 konden worden geselecteerd en daarmee tevens in een duidelijke afbakening werd voorzien. De redeneerfouten die in het verrichte onderzoek zijn blootgelegd kunnen zich echter eveneens voordoen bij enkelvoudige DNA-profielen of bij andere vormen van bewijs.

2. Redeneerfouten

Alvorens nader wordt ingegaan op de methode van onderzoek en de bevindingen uit het onderzoek, zal kort worden stilgestaan bij het onderscheid tussen DNA-onderzoek op bronniveau en DNA-onderzoek op activiteitsniveau. Vervolgens zal aandacht worden besteed aan het theorema van Bayes en het uitdrukken van de bewijskracht in de vorm van een aannemelijkheidsverhouding. Tot slot zal aandacht worden besteed aan de belangrijkste valkuilen bij de interpretatie van DNA-bewijs en het gebruik hiervan in het strafproces.

2.1. Rapporteren van waarschijnlijkheden

Wanneer in een strafrechtelijk onderzoek celmateriaal wordt aangetroffen, kunnen verschillende vragen een rol spelen. Vaak zal het in eerste instantie gaan om de vraag van wie het aangetroffen celmateriaal afkomstig is. Dit is een onderzoeksvraag op bronniveau. De DNA-deskundige kan, door een vergelijkend DNA-onderzoek uit te voeren en hierover te rapporteren, helpen bij het beantwoorden van die vraag. Dit onderzoek vindt plaats in een geaccrediteerd laboratorium van een forensisch instituut¹⁰ en bestaat uit verschillende fasen.¹¹ Wanneer deze zijn doorlopen, kan worden geconcludeerd dat de verdachte géén donor is van het spoor, of dat de verdachte een donor kan zijn.¹² Indien de aangetroffen kenmerken in het DNA-profiel matchen met de kenmerken in het DNA-profiel van de verdachte dan kan de verdachte de donor zijn. Een DNA-match doet de kans dat het aangetroffen celmateriaal afkomstig is van de verdachte enorm toenemen. Maar hoe groot die kans uiteindelijk is, hangt niet alleen af van het DNA-bewijs. De kans dat de verdachte de donor is, hangt mede af van het overige bewijs in de zaak. DNA-deskundigen doen daarom geen uitspraken over de kans dat de verdachte de donor van het spoor is. De deskundigen doen uitspraken over de bewijskracht die aan de resultaten van het uitgevoerde DNA-onderzoek kan worden toegekend.¹³

Deze bewijskracht wordt uitgedrukt in een aannemelijkheidsverhouding die ook wel de *likelihood ratio* wordt

genoemd.¹⁴ Hierbij worden twee relevante hypothesen tegen elkaar afgezet, die niet tegelijkertijd waar kunnen zijn.¹⁵ De likelihood ratio is gelijk aan de kans op de onderzoeksresultaten wanneer de eerste hypothese waar is, gedeeld door de kans op de onderzoeksresultaten wanneer de tweede, alternatieve hypothese waar is. De hypothesen kunnen op verschillende niveaus worden geformuleerd: zo zagen we eerder een vraag op bronniveau. Meestal is de eerste hypothese onderdeel van het scenario zoals door het OM in de tenlastelegging is uitgedrukt, bijvoorbeeld dat celmateriaal dat nabij een slachtoffer is aangetroffen afkomstig is van de verdachte. Een alternatieve hypothese op het eerder genoemde bronniveau¹⁶ kan zijn dat het celmateriaal afkomstig is van een willekeurig andere persoon. Het enkel achterlaten van celmateriaal is niet strafbaar. De handeling waarbij dit celmateriaal wordt achtergelaten mogelijk wel. Daarom zijn voor het strafproces vragen op activiteitsniveau¹⁷ veel meer relevant, omdat de stap van bronniveau naar activiteitsniveau meestal niet triviaal is. Deze vragen op activiteitsniveau richten zich op de handeling waarbij celmateriaal is overgedragen, dan wel degene die een specifieke handeling uitvoerde.¹⁸ In één van de uitspraken die is meegenomen in het onderzoek dat in het kader van het scriptietraject is uitgevoerd, wordt bijvoorbeeld op een kled in de kelder van een verdachte celmateriaal aangetroffen waarvan het DNA-profiel matcht met het profiel van een minderjarige aangeefster.¹⁹ In dit onderzoek wordt de eerste hypothese op basis van het scenario van het OM geformuleerd, namelijk dat het materiaal daar terechtgekomen is toen de verdachte het slachtoffer seksueel misbruikte. Deze hypothese wordt afgezet tegen de hypothese dat het materiaal daar terechtgekomen is toen aangeefster de kleden pakte toen zij verdachte hielp bij het opruimen van zijn kelder. Een deskundige heeft voor de interpretatie van onderzoeksresultaten op activiteitsniveau meer informatie nodig over de context van een zaak dan bij de evaluatie van resultaten op bronniveau. Het is van belang dat de deskundige niet alleen DNA-profielen kan vergelijken, maar ook kennis heeft van de omstandigheden in een zaak, van onder andere de kans dat DNA onder die omstandigheden wordt overgedragen en hoe lang celmateriaal onder die omstandigheden op bepaalde oppervlakken aanwezig blijft.²⁰ Bij een DNA-onderzoek op bronniveau kan de bewijskracht in veel gevallen met een op een berekening gebaseerd getal worden uitgedrukt, vanwege de kennis van frequenties waarmee bepaalde DNA-kenmerken in de populatie voorkomen. Wanneer hypothesen op activiteitsniveau zijn geformuleerd, lijkt het in veel gevallen meer passend om de bewijskracht te rapporteren door middel

9. Zie voor een uiteenzetting over de bewijswaarde van DNA-mengprofielen, K. Slooten, 'Meer onzekerheid, maar niet minder zekerheid: de bewijswaarde en ontrafeling van DNA-mengprofielen', *EeR* 2018, afl. 5, p. 172-192.

10. Art. 7 lid 1 Besluit DNA-onderzoek in strafzaken.

11. B. Kokshoorn e.a., 'Bewijskracht van onderzoek naar biologische sporen en DNA. Deel 2. Bronniveau', *EeR* 2014a, afl. 6, p. 204-205.

12. Kokshoorn e.a. 2014a, p. 205.

13. Nederlands Forensisch Instituut, *De reeks waarschijnlijkheidstermen van het NFI en het Bayesiaanse model voor interpretatie van bewijs* (versie 2.1), Den Haag: Nederlands Forensisch Instituut 2014. Zie ook C.E.H. Berger, 'Criminalistiek is terugredeneren', *NJB* 2010, afl. 13, p. 784-789.

14. Meulenbroek 2009, p. 198-202.

15. B. Robertson, G.A. Vignaux & C.E.H. Berger, *Interpreting Evidence. Evaluating Forensic Science in the Courtroom* (second edition), John Wiley & Sons 2016, p. 33-35.

16. Kokshoorn e.a. 2014a, p. 204-212.

17. B. Kokshoorn e.a., 'Bewijskracht van onderzoek naar biologische sporen en DNA. Deel 3. Activiteitsniveau', *EeR* 2014b, afl. 6, p. 213-219.

18. B. Kokshoorn, B.J. Blankers, J. de Zoete, C.E.H. Berger, 'Activity level DNA evidence evaluation: on propositions addressing the actor or the activity', *Forensic Science International* 2017, 278, p. 115-124.

19. Rb. Den Haag 4 mei 2017, ECLI:NL:RBDHA:2017:4664.

20. Kokshoorn e.a. 2014b.

van een verbale waarschijnlijkheidsterm, waarmee uitdrukking wordt gegeven aan een deskundige mening.²¹

Elke term correspondeert met een bepaalde ordegraote van de likelihood ratio. Bij een likelihood ratio van 2-10 zijn de onderzoeksresultaten bijvoorbeeld 'iets waarschijnlijker' onder de eerste hypothese dan onder de tweede hypothese.

De likelihood ratio heeft een belangrijke plaats in het theorema van Bayes.²² Het theorema biedt inzichten voor het gebruik en de interpretatie van forensisch bewijs in het strafproces²³ en ziet er in woorden als volgt uit: *prior odds* \times *likelihood ratio* = *posterior odds*.²⁴ De odds zijn een kansverhouding: de kans dat de ene hypothese waar is, gedeeld door de kans dat de alternatieve hypothese waar is. Met de prior odds wordt de kansverhouding van de hypothesen voorafgaand aan een bepaalde waarneming uitgedrukt en met de posterior odds wordt de kansverhouding na kennisname van deze waarneming uitgedrukt. De odds zijn het domein van de rechter en worden mede bepaald door al het overige bewijs in een strafzaak. Door kennis te nemen van de door de DNA-deskundige gerapporteerde bewijskracht, veranderen de odds. Uit het Bayesiaanse model²⁵ volgt dat een bewijsmiddel waarvan de gerapporteerde likelihood ratio ongelijk is aan 1, of waarbij de resultaten onder de ene hypothese minimaal 'waarschijnlijker' zijn dan onder de andere hypothese, de kansverhouding van de hypothesen verandert.²⁶ Een bewijsmiddel waarbij een likelihood ratio van 1 gerapporteerd is, of waarbij de resultaten onder de ene hypothese 'even waarschijnlijk' zijn als onder de andere hypothese, heeft geen invloed op de kansverhouding van de hypothesen. Dit bewijsmiddel discrimineert namelijk niet tussen de verschillende hypothesen. Hoewel het gebruik van het Bayesiaanse model in strafzaken nog regelmatig onderwerp van discussie is,²⁷ maken deskundigen hier bij het opstellen van deskundigenrapportages steeds vaker gebruik van. Het model verschaft namelijk inzicht in wat de onderzoeksresultaten precies zeggen over de meest relevante vragen in de zaak. Het kan eveneens inzicht verschaffen in de gewenste rolverdeling tussen de deskundige en de rechter. De deskundige kan zich alleen uitlaten over de waarschijnlijkheid van de bevindingen gegeven de hypothesen, terwijl de rechter zich uiteindelijk een oordeel moet vormen over de waarschijnlijkheid van de hypothesen gegeven de bevindingen. Het is van belang dat strafrechters de wijze waarop deskundigen rapporteren begrijpen en dat zij in de bewijsmotivering geen fouten

maken bij het weergeven van de conclusies van deskundigen en bij het trekken van hun eigen conclusies.

2.2. Fouten en onvolkomenheden in het redeneren met kansen

Zoals eerder al uit onderzoek is gebleken worden door juristen fouten gemaakt bij het interpreteren van bewijs, en dat wordt het meest duidelijk bij het interpreteren van conclusies geformuleerd in termen van de likelihood ratio. De vraag naar de waarschijnlijkheid van de hypothesen vereist een beschouwing van al het relevante bewijs. De wens die vraag (snel) te beantwoorden leidt echter nog weleens tot het lezen van dingen die niet in de deskundigenrapportage staan, wat kan leiden tot redeneerfouten bij de bewijswaardering. Een aantal fouten die uit strafrechtelijke uitspraken kan blijken, zonder dat hiervoor het onderliggende deskundigenrapport geraadpleegd hoeft te worden, wordt in deze paragraaf besproken.

De meest bekende redeneerfout is de *prosecutor's fallacy* of de *transposed conditional*, waarbij voorwaardelijke kansen ten onrechte worden omgedraaid.²⁸ Een simpel voorbeeld betreft de kans dat een voorwerp rechthoekig is, gegeven dat dit voorwerp een boek is. Deze kans is niet gelijk aan de kans dat een voorwerp een boek is, gegeven dat dit voorwerp rechthoekig is. In het geval van DNA-bewijs concludeert de deskundige bijvoorbeeld dat de resultaten van het DNA-onderzoek²⁹ 100.000 keer waarschijnlijker zijn wanneer het celmateriaal afkomstig is van de verdachte (hypothese 1) dan wanneer het celmateriaal afkomstig is van een willekeurig andere persoon (hypothese 2). Wanneer de rechter op basis hiervan concludeert dat het dus 100.000 keer waarschijnlijker is dat het celmateriaal van de verdachte afkomstig is dan van een willekeurig andere persoon, maakt hij een redeneerfout. De deskundige doet namelijk een uitspraak over de kans om bepaalde onderzoeksresultaten te verkrijgen, gegeven de hypothesen. De rechter verdraait dit naar een uitspraak over de kans dat een bepaalde hypothese juist is, gegeven de onderzoeksresultaten. Dat is een andere kans. De conclusie van de deskundige is gebaseerd op de zeldzaamheid van DNA-profielen. Op basis hiervan kan geen uitspraak over de herkomst van celmateriaal worden gedaan. Daarvoor zijn ook de beschikbare andere bewijsmiddelen en informatie in een strafzaak (voor zover relevant) nodig.³⁰

Een andere redeneerfout is de *defense attorney's fallacy*.³¹ Op basis van de frequentie waarmee DNA-kenmerken voorkomen, kan een DNA-deskundige bijvoorbeeld rap-

21. *Vakbijlage. De reeks waarschijnlijkheidstermen van het NFI en het Bayesiaanse model voor interpretatie van bewijs* (versie 2.2), Den Haag: Nederlands Forensisch Instituut 2017.

22. Een theorema is een wetmatigheid die rechtstreeks volgt (te bewijzen is) uit de basiswetten van de logica en kansrekening.

23. Zie onder meer M.O. Finkelstein & W.B. Fairley, 'A Bayesian Approach to Identification Evidence', *Harvard Law Review* 1970, afl. 3, p. 489-517; C.E.H. Berger & D. Aben, 'Bewijs en overtuiging: Redeneren in de rechtszaal', *EeR* 2010a, afl. 3, p. 86-90; Robertson, Vignaux & Berger 2016.

24. Robertson, Vignaux & Berger 2016, p. 15.

25. Strikt genomen is er geen sprake van een model maar van een toepassing van een theorema van de logica.

26. Robertson, Vignaux & Berger 2016, p. 19-20. In de literatuur wordt vaak gesproken van invloed op 'de overtuiging' van de strafrechter. Omdat dit juridisch gezien een complex begrip is, is ervoor gekozen om in plaats hiervan te spreken van kansen en kansverhoudingen.

27. Zie onder meer P. van Koppen, *Overtuigend bewijs. Indammen van rechterlijke dwalingen*, Amsterdam: Nieuw Amsterdam Uitgevers 2011, p. 200-243;

H. Prakken, 'Strafrechtelijk bewijzen: met Bayes of met verhalen? Of is er een derde weg?', *EeR* 2014, afl. 1, p. 4-19; P.J. van Koppen, 'De Tengelhamer en het Schedeldak. Over Rapportages door het Nederlands Forensisch Instituut', *NJB* 2017, afl. 21, p. 1444-1452.

28. W.C. Thompson & E.L. Schumann, 'Interpretation of Statistical Evidence in Criminal Trials. The Prosecutor's Fallacy and the Defense Attorney's Fallacy', *Law and Human Behavior* 1987, afl. 3, p. 167-187; C.E.H. Berger & D. Aben, 'Bewijs en overtuiging: Een helder zicht op valkuilen', *EeR* 2010b, afl. 5/6, p. 162-164; Robertson, Vignaux & Berger 2016, p. 129-133.

29. Bijvoorbeeld wanneer op basis van een vergelijkend onderzoek een match tussen het DNA-mengprofiel aangetroffen op de plaats delict en het DNA-profiel verkregen uit het referentiemateriaal van de verdachte wordt vastgesteld.

30. M. Sjerps, A. Kloosterman & K. van der Beek, 'De interpretatie van een DNA-databankmatch', *DD* 2010/2, p. 138-155.

31. Thompson & Schumann 1987; Berger & Aben 2010b, p. 161-162.

porteren dat de combinatie van DNA-kenmerken die hij in een spoor heeft aangetroffen bij één op de 500 mensen voorkomt. Uit het DNA-onderzoek blijkt dat ook de verdachte deze combinatie van kenmerken heeft. In een populatie van 20.000 personen zouden op basis van de zeldzaamheid van het profiel gemiddeld 40 mensen dit profiel hebben. De *defense attorney's fallacy* wordt gemaakt als op basis van deze gegevens wordt beweerd dat de kans dat de verdachte de bron van het spoor is, slechts 1/40 is. Deze kans hangt namelijk niet alleen af van de bewijskracht van de bevindingen van het DNA-onderzoek, maar ook van de overige beschikbare informatie en bewijs (de prior odds). Bij de *defense attorney's fallacy* wordt impliciet verondersteld dat overig bewijs en informatie geen enkel onderscheid maken tussen de genoemde 40 mensen. Wanneer zulk bewijs en informatie er wel degelijk is, wordt dat hierdoor ten onrechte in één keer van tafel geveegd.

Hoewel de rechter vrij is in de selectie en waardering van bewijs, kan het missen van de samenhang tussen bewijsmiddelen tot een ongewenste uitkomst van het strafproces leiden.³² Ook dit is aangemerkt als redeneerfout. Het kan zijn dat de rechter oordeelt dat de bewijskracht van een afzonderlijk bewijsmiddel niet sterk genoeg is en om deze reden het bewijsmiddel wegstreept. Na dit wegstrepen kijkt de rechter naar een volgend bewijsmiddel, dat wellicht ook als niet sterk genoeg wordt beoordeeld. Een dergelijke werkwijze kan uiteindelijk tot een vrijspraak leiden. Wanneer de bewijsmiddelen in samenhang met elkaar zouden zijn gezien, had het oordeel wellicht anders geluid. De rechter dient de bewijsstandaard pas toe te passen bij de bewijsvraag op het totaal aan bewijs, en niet bij het beoordelen van een individueel bewijsmiddel.³³ De samenhang van verschillende bewijsmiddelen kan gemist worden wanneer niet wordt ingezien dat onafhankelijk bewijs dat in dezelfde richting wijst, elkaar versterkt. Het is mogelijk dat afzonderlijke bewijsmiddelen geen grote bewijskracht hebben, maar gecombineerd wél.³⁴ Om deze reden is het van belang om een zwak bewijsmiddel niet al vóór de uiteindelijke bewijsbeslissing af te schrijven.

3. Methode en bevindingen

In het kader van het eerdergenoemde scriptietraject is een onderzoek verricht op rechtspraak.nl naar in 2017 gepubliceerde rechtspraak waarin de term 'mengprofiel' voorkwam. Dit leverde in totaal 111 uitspraken op: 84 vonnissen van rechtbanken en 27 arresten van gerechtshoven. Van de 111 uitspraken zijn er 33 om diverse redenen niet voor een verdere analyse meegenomen in dit onderzoek, bijvoorbeeld omdat het bij nadere bestudering toch niet bleek te gaan om een DNA-mengprofiel of omdat het desbetreffende gerecht tegen meerdere verdachten uitspraak had gedaan en het meenemen van die uitspraken zou leiden tot een dubbeltelling.³⁵ Uiteindelijk

zijn 60 vonnissen van rechtbanken en 18 arresten van gerechtshoven verder meegenomen in het onderzoek. Eén van de doelen van het onderzoek was het in kaart brengen van de kwaliteit van de bewijswaardering. Deze kwaliteit is beoordeeld aan de hand van de aan- of afwezigheid van redeneerfouten in de bewijsoverwegingen. Tevens is gekeken naar vaagheden of onduidelijkheden in de bewijsmotivering die raken aan de kwaliteit/transparantie van de desbetreffende overweging en eventueel kunnen duiden op een gebrek aan inzicht in de problematiek. Voor dit onderzoek had de onderzoeker niet de beschikking over de onderliggende rapportages. Een enkele keer werd in de uitspraak onder het kopje bewijsmiddelen letterlijk het onderliggende rapport aangehaald, waardoor de eventueel aangetroffen redeneerfout of de onvolledige formulering met zekerheid aan de rechter kon worden toegeschreven. In de meeste zaken was het echter niet mogelijk te herleiden wat de deskundige precies in zijn rapport had opgenomen. Gelet op de gestandaardiseerde wijze van rapporteren kan er evenwel van uit worden gegaan dat de hieronder beschreven redeneerfouten niet door de deskundige zelf zijn begaan. Op basis van de aangetroffen formuleringen is een indeling gemaakt in verschillende typen redeneerfouten en vage formuleringen.

3.1. Redeneerfouten

De *prosecutor's fallacy* is in 27 van de 78 zaken aangetroffen (35%). De *prosecutor's fallacy* is in verschillende vormen en formuleringen aangetroffen. In 17 zaken (22%) worden de voorwaardelijke kansen letterlijk omgedraaid en wordt een uitspraak gedaan over de waarschijnlijkheid van de hypothesen in plaats van over de waarschijnlijkheid van de onderzoeksresultaten. Dit gebeurt onder meer in een uitspraak waarin de volgende conclusie van de deskundige letterlijk is overgenomen: 'De bevindingen van het vergelijkend DNA-onderzoek zijn ongeveer 9 miljoen keer waarschijnlijker als hypothese 1 waar is, dan als hypothese 2 waar is.'³⁶ Dat de zittingsrechters zich schuldig maken aan de *prosecutor's fallacy* blijkt vervolgens uit de volgende overweging die in de uitspraak is opgenomen: 'Volgens de bevindingen van het rapport van het NFI is de hypothese dat het mengprofiel het DNA van verdachte en slachtoffer bevat, ongeveer 9 miljoen keer waarschijnlijker dan de hypothese dat het van verdachte en een onbekend gebleven persoon is.'³⁷ Een ander voorbeeld is de volgende formulering: 'Het NFI concludeert daarbij dat het extreem veel waarschijnlijker is dat de bemonstering DNA bevat van beide broers en een niet verwante persoon, dan dat de bemonstering DNA bevat van de broer van de verdachte en twee onbekende, niet verwante personen.'³⁸

In de overige 10 van die 27 zaken is vervolgens een overweging aangetroffen waarbij het wellicht minder duidelijk is dat het een *prosecutor's fallacy* betreft. Een voorbeeld hiervan is de volgende overweging: 'Uit het

32. Berger & Aben 2010b, p. 164-165.

33. Berger & Aben 2010b, p. 164.

34. Robertson, Vignaux & Berger 2016, p. 72-73.

35. De laatste zoekslag dateert van 28 maart 2018. Nadien gedane gepubliceerde jurisprudentie uit 2017 is niet meer meegenomen. Zie voor een nadere toelichting op de zoektermen en de nadien verrichte selectie paragraaf 4.1 van de scriptie de eerste auteur van dit artikel die in de zomer van 2019 zal worden uitgegeven bij Uitgeverij Paris onder de titel *DNA-mengprofielen in strafzaken: een jurisprudentieonderzoek naar de kwaliteit van de waardering van DNA-bewijs*.

36. Rb. Noord-Nederland 20 september 2017, ECLI:NL:RBNNE:2017:4349.

37. Rb. Noord-Nederland 20 september 2017, ECLI:NL:RBNNE:2017:4349.

38. Rb. Rotterdam 29 november 2017, ECLI:NL:RBROT:2017:10427.

onderzoek aan de handschoen, die is achtergelaten door de dader, blijkt dat daarop aangetroffen DNA-materiaal afkomstig is van verdachte.³⁹ Beargumenteerd kan worden dat dergelijke formuleringen in een 'donkergruis' gebied vallen. Het staat strafrechters namelijk vrij om zelf op basis van het DNA-onderzoek – in het licht van de hele strafzaak – te stellen dat voldoende zeker is wie de donor van het celmateriaal is. Echter, door de wijze van formuleren, namelijk dat 'uit het onderzoek blijkt' is hier wel degelijk sprake van een *prosecutor's fallacy*. De DNA-deskundige zal namelijk geen uitspraak doen over de herkomst van celmateriaal en dit zal dan ook niet uit enkel het onderzoek aan de handschoen zijn gebleken. De formulering is in deze zaak dan ook te kort door de bocht.

De *defense attorney's fallacy* is in geen van de uitspraken aangetroffen en in slechts één zaak wordt naast de *prosecutor's fallacy* eveneens de samenhang van het bewijs gemist.⁴⁰ In deze zaak concludeert de rechter dat een aantal bewijsmiddelen 'de verdenking jegens de verdachte kan versterken maar dat dit geen bewijs oplevert dat hij de dader is van de verkrachting.' Dit betekent 'dat de uitkomsten van het DNA-onderzoek van doorslaggevend belang zijn voor de beoordeling.' De rechtbank concludeert vervolgens dat de resultaten van het DNA-onderzoek: 'wijzen in de richting van betrokkenheid van verdachte, maar daarvoor op zichzelf geen doorslaggevend bewijs vormen'. Doordat de rechtbank eerst een conclusie trekt op basis van het overige bewijs en vervolgens ook het DNA-bewijs als niet doorslaggevend voor de betrokkenheid van de verdachte afdoet, maakt zij duidelijk dat zij de bewijsmiddelen niet in samenhang heeft bekeken. Deze uitspraak is in de literatuur eveneens bekritiseerd.⁴¹

3.2. Het grijze gebied: onduidelijke of onvolledige overwegingen

In het jurisprudentieonderzoek zijn naast de 27 uitspraken waarin een redeneerfout is aangetroffen, nog eens 28 uitspraken gevonden (36%), met daarin vage of onvolledige redeneringen.

Allereerst is in 14 zaken (18%) in de uitspraak opgenomen dat DNA van een bepaald persoon in het mengprofiel 'is aangetroffen'. Hierbij wordt niet duidelijk of deze conclusie alleen op het deskundigenonderzoek gebaseerd is, op een combinatie van het deskundigenbewijs en overige bewijsmiddelen, of op een verklaring van een verdachte die het niet betwist. Een dergelijke absolute herkomstuitspraak zou een *prosecutor's fallacy* zijn wanneer deze uitspraak alleen op basis van het DNA-onderzoek wordt gedaan. Ten tweede concludeert de zittingsrechter in zes zaken (8%) enkel dat er sprake is van een 'match' en in één zaak (1,3%) dat de verdachte de donor van een spoor 'kan zijn'. Zonder bewijskracht discrimineren deze formuleringen niet tussen verschillende scenario's.

Tot slot zijn er zes zaken (8%), waarin de redenering niet geheel sluitend lijkt, maar niet helemaal duidelijk wordt

wat de reden hiervoor is. Zo wordt in drie zaken uit de uitspraak weliswaar duidelijk dat een DNA-onderzoek heeft plaatsgevonden, maar blijft het geheel onduidelijk welke rol het DNA-bewijs speelt in de strafzaak, of op welke wijze het DNA-bewijs redengevend is voor de bezeugenverklaring.⁴² In de drie andere zaken lijkt uit de bewijsoverwegingen te volgen dat de strafrechter het DNA-bewijs niet goed weet te waarderen. Een voorbeeld hiervan is de volgende formulering: 'Uit het vergelijkend DNA-onderzoek volgt echter wel dat het 100 tot 10.000 keer waarschijnlijker is als de bemonstering celmateriaal van aangeefster, verdachte en één of twee willekeurige onbekende personen bevat, dan dat het celmateriaal bevat van aangeefster en drie andere willekeurige personen.'⁴³ Hierbij is het onduidelijk waar 'het' naar verwijst. Als dit verwijst naar het vergelijkend onderzoek is het op zijn minst een kromme zin, maar als 'het' slaat op de hypothesen is er sprake van een *prosecutor's fallacy*.

3.3. Verweren op activiteitsniveau

Zoals hierboven uiteengezet kunnen er op verschillende niveaus hypothesen worden geformuleerd. Door de vaak grote gerapporteerde bewijskracht van de onderzoeksresultaten op bronniveau, is het logisch dat de verdediging vaker verweer voert ten aanzien van de handeling waarbij het celmateriaal op een bepaalde plaats is terechtgekomen dan dat zij betwist dat de verdachte de bron is van het celmateriaal. Dat zien we ook terug in het onderhavige onderzoek. In 55 van de 78 geanalyseerde zaken (70%) heeft de verdediging een of meerdere verweren gevoerd die zich richten op het DNA-bewijs. Hoewel de resultaten van het DNA-onderzoek door de deskundige in 75 van de 78 zaken enkel op bronniveau geëvalueerd zijn, werd het leeuwendeel van de bewijsverweren op activiteitsniveau gevoerd. In 48 van de 55 zaken (87%) waarin verweer gevoerd is dat zich richt op het DNA-bewijs, is ten minste een van deze verweren gemotiveerd weerlegd in de uitspraak. In slechts drie zaken verwijst de zittingsrechter bij deze weerlegging naar de inhoud van het deskundigenrapport of de verklaring van de deskundige (ter terechtzitting of op een eerder moment). In het overige en dus grootste deel van de zaken wordt gereageerd op het verweer, zonder hierbij een deskundige te raadplegen of te verwijzen naar een deskundigenbijdrage. De zittingsrechter lijkt in deze zaken van oordeel te zijn dat nadere raadpleging van een deskundige niet noodzakelijk is om inhoudelijk op de verweren te kunnen reageren, ondanks dat het grootste deel van de verweren betrekking had op activiteitsniveau (en de deskundige zich hierover niet heeft uitgelaten).

In minstens twee zaken formuleren de zittingsrechters ook daadwerkelijk een conclusie op activiteitsniveau, terwijl in het onderliggende deskundigenrapport uitsluitend op bronniveau is gerapporteerd. Zo wordt op basis van een vergelijkend DNA-onderzoek op bronniveau vastgesteld dat 'het masker door verdachte gedragen is'⁴⁴ en wordt geoordeeld 'dat uit de onderzoeken van

39. Rb. Noord-Nederland 4 mei 2017, ECLI:NL:RBNNE:2017:1686.

40. Rb. Amsterdam 13 april 2017, ECLI:NL:RBAMS:2017:2365.

41. H.W.J. de Groot, 'Uit de praktijk. Verkrachting op het Brettenpad', *EeR* 2017, afl. 5, p. 205-213.

42. Onder meer Rb. Noord-Holland 21 december 2017, ECLI:NL:RBNHO:2017:11440.

43. Hof Arnhem-Leeuwarden 14 september 2017, ECLI:NL:GHARL:2017:8015.

44. Rb. Oost-Brabant 12 juli 2017, ECLI:NL:RBOBR:2017:3711.

het NFI kan worden geconcludeerd dat het de verdachte is geweest die de handschoenen als laatste heeft gedragen.⁴⁵ Dit is problematisch, omdat op basis van een onderzoek waarbij de resultaten op bronniveau geëvalueerd worden, geen conclusies op activiteitsniveau kunnen worden getrokken. Een conclusie op activiteitsniveau vereist in het algemeen nog meer informatie en deskundigheid dan een conclusie op bronniveau. Of wellicht heeft de rechter die conclusie aan ander bewijs en informatie uit het dossier ontleend, maar dan kan die conclusie niet in de mond van de deskundige worden gelegd.

4. Reflectie

Het voorgaande vraagt om nadere reflectie. Hoe erg is het dat bovenstaande redeneerfouten worden gemaakt? Wat betekent dat voor de kwaliteit van de genomen beslissing? En wat kan en mag in termen van de deugdelijkheid en transparantie van de bewijsmotivering van de strafrechter worden verlangd? Vooropgesteld moet worden dat het aantreffen van een *prosecutor's fallacy* (of een andere redeneerfout) niet betekent dat de genomen beslissing over het tenlastegelegde onjuist is. Een *prosecutor's fallacy* houdt over het algemeen een overschatting van het bewijs in. Die overschatting hoeft echter geen desastreuze gevolgen te hebben in gevallen waarin het totaal aan bewijs zo groot was dat het ook zonder die overschatting tot dezelfde beslissing geleid had. Bovendien wordt de herkomst van DNA-materiaal vaak door de verdediging niet betwist. De vraag is dan hoe erg het is als de rechter in zijn uitspraak zegt dat het DNA van de verdachte 'is aangetroffen'. De rechter kan die conclusie voor zijn rekening nemen wanneer het bij voorbaat al niet bijzonder onwaarschijnlijk was om het DNA van de verdachte aan te treffen en daar vervolgens nog eens de enorme bewijskracht van een overeenkomend DNA-profiel bijkomt. Let wel: die bewijskracht is zo groot voor de vraag van wie dat DNA komt, de vraag op bronniveau.

Afgezien van mogelijke consequenties voor de juistheid van hun oordeel mag ook van rechters worden verlangd dat zij de resultaten van onderzoek in hun uitspraak correct weergeven met het oog op de transparantie van het vonnis of arrest. De resultaten van dit onderzoek laten zien dat de strafrechter worstelt met zijn rol in dezen en zich soms hult in vage formuleringen, zoals een spoor 'past bij' of is 'consistent met' een bepaalde hypothese. Wanneer slechts geconcludeerd kan worden dat twee zaken niet tegenstrijdig zijn, heeft dit op zichzelf geen bewijskracht. Dergelijke formuleringen discrimineren namelijk niet tussen verschillende hypothesen. De rechter hoeft zich weliswaar ambtshalve niet uit te laten over de vraag waarom hij bepaald bewijsmateriaal redeneerend heeft geacht voor zijn bewijsbeslissing, maar als hij dat wel doet dan moet dat ook voldoende transparant zijn.

Niettemin valt hier wel een zekere spanning te ontwaren met de traditionele wijze van redeneren in het strafproces. De Groot, oud-raadsheer bij het Hof Amsterdam, schrijft hierover in zijn boek *'DNA in het strafrecht: Hoe moet het niet? Hoe moet het wel?'* het volgende.

'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 rationele bewijswaardering in kringen van strafrechtbeoefenaren echter nog lang niet.'⁴⁶

Het idee van een sleutel zou niet leiden tot een rationele bewijswaardering, maar de constatering dat rechters redeneren in vaststaande feiten geeft wel een indicatie van waar het probleem ligt. De vaststaande of vastgestelde feiten vormen inderdaad vaak het beginpunt van een juridische redenering, zeker als het gaat om het toepassen van het recht op de 'feiten'. Ook is het veel eenvoudiger om te redeneren met stellingen waarvan we weten dat ze waar of onwaar zijn. Zo is de redenering 'Als A en B, dan C' veel eenvoudiger dan 'Als hoogstwaarschijnlijk A en misschien B, dan...'. Maar voor de rationele bewijswaardering is het noodzakelijk onzekerheid mee te nemen in de oordeelsvorming. Het bestempelen van stellingen waarvan de waarheid onzeker is als noodzakelijk waar of onmogelijk waar kan gemakkelijk leiden tot verkeerde beslissingen. We noemden eerder al de salami denkfout, waarbij telkens een onderdeel van het bewijs in isolatie wordt bekeken en als onvoldoende wordt weggestreept. De wettelijk vereiste overtuiging van de rechter geldt voor het totaal aan bewijs, en niet voor alle afzonderlijke onderdelen daarvan.

Van rechters mag worden verlangd dat zij dit onderkennen. Er bestaat evenmin een grenswaarde voor bewijsmiddelen. Veel stukjes bewijs met ieder een beperkte bewijskracht tezamen kunnen toch een sterke bewijsconstructie opleveren. Het risico van het presenteren van informatie als vaststaand is dat het probabilistisch karakter van die informatie wordt genegeerd en aan het bewijsmateriaal niet het juiste gewicht wordt toegekend. Dat speelt niet alleen bij DNA-bewijs maar bij alle soorten bewijsmateriaal. Dat de rechter bij dit type feitelijke beslissingen moet werken met categorische stellingen (en er daarmee noodzaak zou zijn voor een grenswaarde), lijkt ons een onjuiste voorstelling van zaken.

Een ander punt dat nadere reflectie behoeft, zijn de resultaten met betrekking tot conclusies op activiteitsniveau. De stap van de vraag wie de bron was van een spoor naar de vraag wie welke handeling uitvoerde is zelden triviaal. Het beantwoorden van vragen op activiteitsniveau vereist meer deskundigheid en informatie dan het beantwoorden van vragen op bronniveau. Toch is dit van oudsher het domein van de rechter, vooral omdat de deskundige zich er vroeger liever niet aan waagde. Tegenwoordig is dit een actief gebied van onderzoek in de forensische wetenschap, en doet een rechter er goed aan zich juist ook op dit punt bij te laten staan door een deskundige. Bij gebrek aan deskundigheid ontstaan er

45. Hof Den Haag 23 mei 2017, ECLI:NL:GHDHA:2017:1468.

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

al gauw nogal absurde vuistregels zoals die volgens welke een spoor op een 'verplaatsbaar voorwerp' van minder of geen waarde zou zijn. Dat komt de rationele bewijswaardering zeker niet ten goede, en dit onderdeel van bewijswaardering laat zich nu eenmaal niet in vuistregels vangen.

5. Hoe verder?

Nederland loopt voorop in het besef dat de kwaliteit van redeneren met en het waarderen van bewijs van groot belang is voor de kwaliteit en de rechtvaardigheid van onze rechtspraak. Dat heeft te maken met zowel het relatief hoge niveau van de forensische wetenschap als met de progressieve houding van veel juristen in ons land. De vooruitgang in de forensische wetenschap heeft een aantal problemen aan het licht gebracht die wij hier besproken hebben. In menig ander land wordt de forensische wetenschap als brenger van die wat ontvullende boodschap er onterecht van beschuldigd de veroorzaker te zijn van die problemen.

Zoals reeds genoemd zijn de afgelopen jaren verschillende initiatieven genomen ter verbetering van de kwaliteit van deskundigenrapporten en het begrip hiervan door partijen in de strafrechtketen.⁴⁷ Zo is het *Handboek deskundigen voor de strafrechter* ontwikkeld⁴⁸ en biedt alleen al het Studiecentrum Rechtspleging (SSR) verschillende cursussen op dit gebied aan. Ook het Nederlands Forensisch Instituut biedt de nodige cursussen, en werkte met zogenoemde gebruikerspanels om breed feedback te krijgen van de gebruikers van zijn rapporten. Daarnaast maakt het NFI vakbijlages beschikbaar waarin het onderzoek wordt toegelicht en is er een specifieke vakbijlage over welke conclusies de strafrechter op basis van het rapport kan en mag trekken.⁴⁹ Tot slot zijn diverse (positief geëvalueerde) pilots uitgevoerd met forensisch ondersteunend personeel in rechtbanken en gerechtshoven.

Het is belangrijk om te benadrukken dat het rationeel waarderen van bewijs niet betekent dat het tot een reken-som leidt, waarbij de rechter een wiskundige moet zijn en waarbij eenieder gegarandeerd ook tot dezelfde conclusie komt. Het waarderen van bewijs is in belangrijke mate subjectief, en competente mensen kunnen daarbij van mening verschillen. Maar het doel is wel om redeneerfouten zoveel mogelijk te beperken. De vraag is of bovengenoemde en andere initiatieven voldoende hun vruchten hebben afgeworpen. Het voorgaande lijkt erop te duiden dat rechters nog steeds moeite hebben met het probabilistisch karakter van bewijs. Hoe kan het begrip van rechters en de kwaliteit van de rechterlijke motiveringen van uitspraken (verder) worden verbeterd? Wij denken dat de nodige winst valt te behalen in de universitaire opleiding van rechters en andere juristen. Al in dat vroege stadium is het van belang te onderkennen dat het niet alleen gaat over het toepassen van het recht, maar zeker ook over waarheidsvinding. Een beperkte aanpassing van het curriculum zou al een flinke kwaliteitsverbetering kunnen opleveren. Toepassing van elementaire kennis uit de logica wil niet zeggen dat een

jurist per se wiskunde in zijn of haar pakket moet hebben gehad, maar het kan geen kwaad als het recht ook een aantrekkelijke studierichting is voor mensen met talent op dat vlak. Ook zou het goed kunnen zijn om deskundigen mee te laten kijken in strafzaken (waarin zij zelf geen rol spelen): het helpt als we elkaars taal spreken en elkaars uitdagingen kennen.

Tot slot beseffen we dat er ook tijd voor nodig is. Per slot van rekening is de prosecutor's fallacy een redeneer-fout die veel deskundigen (buiten het DNA-onderzoek) tien jaar geleden zelf nog maakten. Wanneer de deskundige die fout niet meer maakt, maar de rechter alsnog, dan is dat een gemiste kans op verbetering. Die verbetering is van groot belang om de kansen op een terecht oordeel, en daarmee rechtvaardigheid en vertrouwen in het rechtssysteem, te optimaliseren.

47. Zie voor een meer uitgebreide beschrijving van de initiatieven paragraaf 2.5.3 van de scriptie (zie noot 35).

48. P.C.T. van Dam, *Handboek deskundigen voor de strafrechter*, Den Haag: Raad voor de rechtspraak 2009.

49. Zie hieromtrent ook de website van het NFI: <https://www.forensischinstituut.nl/publicaties/publicaties/2017/10/18/vakbijlage-waarschijnlijkheidstermen>.