

Likelihood ratio

De betrouwbaarheid van de verklaring is in de methode van de alternatieve scenario's gedefinieerd als de verhouding tussen de waarschijnlijkheid van de verklaring binnen het primaire scenario en die binnen het alternatieve scenario. In andere woorden: als de verklaring net zo goed te verklaren is door het alternatieve scenario als door het primaire scenario, dan is de betrouwbaarheid laag. Dit betekent dat er gezocht moet worden naar zinnige alternatieve scenario's die het ontstaan van de verklaring zouden kunnen verklaren. Veelal betreft dit een kwalitatieve analyse. Soms is er op basis van wetenschappelijk onderzoek daadwerkelijk een waarschijnlijkheid te schatten. Als de waarschijnlijkheid van de getuigenverklaring onder het primaire en alternatieve scenario is te schatten, wordt het mogelijk om een *likelihood ratio* (ofwel diagnostische waarde) te berekenen. Neem als voorbeeld de vraag naar de betrouwbaarheid van een herkenning door een getuige van een verdachte die aan de getuige reeds eerder bekend was. De getuige heeft een hem/haar bekend persoon dus een strafbaar feit zien plegen. Althans dat denkt de getuige, maar de vraag is hoe betrouwbaar die herkenning is. Stel dat bekend is dat de getuige de dader heeft gezien bij daglicht (3000 lux) en op slechts drie meter afstand, dan leert het onderzoek van De Jong et al.¹ het volgende. Als het daadwerkelijk om de persoon gaat van wie de getuige denkt dat het gaat (het primaire scenario), dan is de kans op een herkenning onder deze waarnemingsomstandigheden zeer groot, namelijk naar schatting 97%. Onder deze waarnemingsomstandigheden komen onterechte herkenningen (de getuige denkt dat de dader de hem/haar bekende persoon is, maar in werkelijkheid is het iemand anders die er alleen maar op lijkt) niet of nauwelijks voor (afgerond op een heel getal: 0%). De betrouwbaarheid (uitgedrukt in een *likelihood ratio*) van deze herkenning is dan (97% / 0% =) oneindig groot, of voorzichtiger: tenminste 97.

Als de afstand echter geen drie meter, maar 30 meter was, dan neemt de kans op een terechte herkenning af van 97% naar 40%, terwijl die op een valse herkenning stijgt van 0% naar 5%, blijkens het onderzoek van De Jong et al. De diagnostische waarde van deze herkenning ligt dan rond de (40% / 5% =) 8.

¹ De Jong, M., Wagenaar, W.A., Wolters, G. & Verstijnen, I.M. (2005). Familiar face recognition as a function of distance and illumination: A practical tool for use in the courtroom. *Psychology, Crime & Law*, 11, 87-97.

TABLE 2 Hits (bold) and false alarms (not bold) in the recognition tests (percentage scores).

<i>Distance (m)</i>	<i>Illumination level (lux)</i>								
	<i>0.3</i>	<i>2</i>	<i>3</i>	<i>5</i>	<i>10</i>	<i>30</i>	<i>150</i>	<i>300</i>	<i>3000</i>
3	0	79	100	100	93	96	94	97	97
	0	0	5	0	8	3	0	0	0
5	0	68	89	94	94	97	92	92	100
	0	14	8	14	3	6	0	7	3
7		63	63	89	93	93	97	96	90
		3	10	3	6	3	6	3	0
12			60	70	76	81	80	92	100
			3	6	6	5	3	5	2
20				32	52	65	77	79	100
				6	14	9	13	6	14
30					20	29	32	45	40
					10	8	11	8	5
40						7	17	17	30
						3	5	10	2

Bedacht moet worden dat een diagnostische waarde een reductie van de werkelijkheid impliceert. In een echte situatie spelen bijvoorbeeld meer factoren dan alleen lichtsterkte en afstand een rol (denk aan stress en verwachtingseffecten). (Gedrags)wetenschappelijk onderzoek is per definitie een gereduceerde simulatie van de werkelijkheid. In onderzoek wordt een factor geïsoleerd, gemanipuleerd en vervolgens bestudeerd. Die reductie is inherent aan wetenschappelijke inzichten. Die reductie zit overigens ook verpakt in uitspraken van het type “vandaag is de kans op regen 80%”, “de kans om te overlijden na inname van deze drug is 99%”, “die combinatie van omstandigheden lijkt me al te toevallig” en zelfs in ogenschijnlijk niet-kwantitatieve uitspraken als “deze verklaring bevat meerdere kenmerken die duiden op geloofwaardigheid”. Ondertussen geeft de diagnostische waarde een numerieke indicatie van de bewijskracht van het bewijsmiddel. Ik heb enige tijd geleden een poging gedaan om de literatuur over een aantal niet-technische bewijsmiddelen samen te vatten in diagnostische waarden.² Internationaal gezien is het voor technische bewijsmiddelen (behorend tot het domein van de *sciences*) standaard geworden om de betrouwbaarheid uit te drukken in een *likelihood ratio*.³ Het NFI drukt de betrouwbaarheid van technische bewijsmiddelen sinds 2009 overigens ook vaak uit in

² Rassin, E. (2015). *De diagnostische waarde van bewijs*. Deventer: Wolters Kluwer.

³ Association of Forensic Science Providers (AFSP) (2009). Standards for the formulation of evaluative forensic science expert opinion. *Science and Justice*, 49, 161-164.

likelihood ratios. Het NFI heeft ook onderstaand overzicht gepubliceerd van verschillende *likelihood ratios* en hun verbale omschrijving.⁴

Tabel 2. Numerieke definitie van waarschijnlijkheidstermen NFI

Verbale term	Ordegrootte bewijskracht (LR)
ongeveer even waarschijnlijk	1 – 2
iets waarschijnlijker	2 – 10
waarschijnlijker	10 – 100
veel waarschijnlijker	100 – 10.000
zeer veel waarschijnlijker	10.000 – 1.000.000
extreem veel waarschijnlijker	> 1.000.000

Ik merk daarbij op dat de *likelihood ratios* van niet-technische bewijsmiddelen veelal veel lager uitpakken dan die van technische bewijsmiddelen. Zo impliceren De Jong et al. dat een herkenning met een diagnostische waarde van 15 sterk is, terwijl 15 in bovenstaand overzicht in de minder sterk klinkende categorie “waarschijnlijk” valt. In andere, niet technisch-forensische contexten wordt gewerkt met minder hoge waarden, zoals valt te zien in onderstaand schema.⁵ In die benadering zou een bewijsmiddel met een diagnostische waarde van 15 *positive* of *strong* heten.

Table 4.
Interpretation of Bayes Factors as Evidence for Alternative Hypotheses

Statistic		Support for H ₁	
Bayes Factor	Inverse of Bayes Factor	Raftery	Jeffreys
1–.33	1–3	Weak	Anecdotal
.33–.10	3–10	Positive	Substantial
.10–.05	10–20	Positive	Strong
.05–.03	20–30	Strong	Strong
.03–.01	30–100	Strong	Very Strong
.01–.0067	100–150	Strong	Decisive
<.0067	>150	Very Strong	Decisive

⁴ NFI (2014). *De reeks waarschijnlijkheidstermen van het NFI en het Bayesiaanse model voor interpretatie van bewijs*. Den Haag: NFI.

⁵ Jarosz, A.F. & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *Journal of Problem Solving*, 7, 2-9.