

De kwaliteit van civielrechtelijke vonnissen

Waarom integrale toetsing daarvan zo belangrijk is

Rechters maken zich zorgen dat ze steeds minder in staat zijn om rechtspraak van goede kwaliteit te leveren. Maar waaruit blijkt dat? Daartoe wil je de deugdelijkheid van rechtspraak op betrouwbare wijze onderzoeken en mogelijke veranderingen kunnen objectiveren. Dat vraagt om een systematische wijze van onderzoek met een juiste methodologie. Er is in ons land al enkele malen onderzoek verricht naar de kwaliteit van civiele rechtspraak. Maar beoordeeld vanuit het perspectief van 'total quality management' waren die evaluaties nog te beperkt van opzet en kunnen er vraagtekens geplaatst worden bij de uitvoering ervan. Kwaliteitstoetsing is de kritische reflectie over het eigen handelen langs de weg van systematisch (zelf)onderzoek. Die benadering vraagt echter een multidisciplinaire aanpak en vooral ook begripsbepaling van wat onder kwalitatief goede rechtspraak moet worden verstaan. Dat laatste ontstaat pas gaandeweg, want kwaliteit is voor velerlei uitleg vatbaar. Kwaliteitszorg moet ingebed zijn in het dagelijkse werk, maar krijgt nu veel te weinig aandacht en ontbeert de benodigde expertise en infrastructuur.

1. Zorgen over toenemende druk op de rechtspraak

In 2006 pleitte de *Commissie evaluatie modernisering rechterlijke organisatie* voor meer structurele aandacht voor onderzoek naar en besteding van kwaliteit van rechtspraak.¹ In september 2013 organiseerde de Tweede Kamer een rondetafelgesprek over de kwaliteit van de rechtspraak. Naar aanleiding daarvan meldde de Raad voor de rechtspraak: 'De sprekers zijn het erover eens dat hoge kwaliteit van rechtspraak van essentieel belang is voor de samenleving en dat de geleverde kwaliteit weinig te wensen overlaat.'² Het belang van het organiseren van goede kwaliteitszorg werd ook benadrukt in het *Rapport Visitatie Gerechten 2014*.³

Door toenemende werkdruk en aanhoudende organisatieveranderingen ontstonden er binnen de magistratuur zorgen over de deugdelijkheid van rechtspraak. Een groep rechters vroeg met klem om meer aandacht voor de – dreigende – neergang van de kwaliteit van het recht met de publicatie van hun *Leeuwarder Manifest*.⁴ Scheidend P-G Fokkens van de Hoge Raad vond in 2016 goede kwaliteit van rechtspraak niet zo vanzelfsprekend.⁵

Ook de voormalig vicepresident van de Raad van State Tjeenk Willink uitte recent zijn zorgen over de overmatige aandacht voor de bestuurlijke en organisatorische aspecten van rechtspraak, die ten koste gaat van de aandacht voor de kwaliteit ervan.⁶ Maar wordt dat gestaafd door feiten of zijn het veronderstellingen? Wat weten we nu eigenlijk *concreet* over de kwaliteit van onze rechtspraak en is die de laatste jaren mogelijk verslechterd? Hoe stel je dan vast dat de kwaliteit ten nadele verandert – maar vooral: wat is nu precies 'kwaliteit'?

Ieder menselijk handelen – of soms het nalaten daarvan – heeft gevolgen. Met behulp van een normatief kader

maakt kwaliteitstoetsing duidelijk of de handelwijze juist was en of het resultaat was zoals bedoeld. Kwaliteitsonderzoek van rechtspraak kun je zowel pro- als retrospectief benaderen. Prospectieve toetsing betekent het ontwikkelen en in praktijk brengen van optimale werkwijzen, dat is vooral een kwestie van methodologie. Een retrospectieve aanpak behelst het onderzoeken en beoordelen van de gehanteerde werkwijzen, evaluatie van de bereikte resultaten om daarna te zien waar het beter zou kunnen. Die pro- en retrospectieve aanpak staan steeds in wisselwerking met elkaar en maken duidelijk waar en hoe bijgestuurd zou moeten worden. Rechtspraak is een elementair sociaal-maatschappelijk systeem dat dagelijks zijn zo belangrijke productie levert. Het functioneren van die rechtspraak is steeds meer onderwerp van zowel een intern als een publiek debat geworden. Aanleiding daartoe was onder meer de onrust over enkele ernstige strafrechtelijke dwalingen. De rechter staat om verschillende redenen bovendien steeds meer onder druk en ook is er maatschappelijk meer behoefte aan transparantie.⁷ Bij al deze kwesties staat de kwaliteit van rechtspraak steeds centraal.

In dit artikel behandel ik de noodzakelijkheid om een meer methodologisch verantwoorde en vooral brede, dus integrale, kwaliteitstoetsing van civielrechtelijke procedures uit te voeren (dergelijke beginselen zijn trouwens ook op andere rechtsgebieden van toepassing). Ik zal daartoe eerst de doelstellingen van rechtspraak en het daarvan afgeleide normatieve kader nader specificeren (par. 2) en me vervolgens richten op de tot nu toe uitgevoerde toetsingsstudies van civiele rechtspraak (par. 3). Dan komen basisprincipes van kwaliteitsonderzoek aan de orde (par. 4). In de slotparagraaf beschouw ik deze uitkomsten (par. 5).

* Prof. dr. dr. R.W.M. Giard is arts, jurist en methodoloog. Hij is emeritus hoogleraar methodologie en aansprakelijkheid, Erasmus School of Law, EUR. Hij was tot 2019 redacteur van dit tijdschrift.

1. W.J. Deetman (voorz.) et al., *Rechtspraak is kwaliteit*, Den Haag: Commissie evaluatie modernisering rechterlijke organisatie 2006.

2. <https://www.rechtspraak.nl/Organisatie-en-contact/Organisatie/Raad-voor-de-rechtspraak/Nieuws/Paginas/Werkdruk-bedeigt-kwaliteit-rechtspraak.aspx>.

3. Raad voor de rechtspraak, *Rapport Visitatie Gerechten 2014*, Den Haag: Raad voor de rechtspraak 2014.

4. 'Raadsheren manifesteren grote zorgen over kwaliteit rechtspraak', *NJB* 2013/45.

5. R.P. van der Pijl & E.J.H.G. van Binnebeke, "Het was een voorrecht procureur-generaal te zijn", *TREMA* 2016, afl. 7.

6. H. Tjeenk Willink, 'De "verbestuurlijke" van de rechterlijke macht en het mislukken van het KEI-project', *NJB* 2018/880.

7. Zie daarover D. Broeders et al. (red.), *Speelruimte voor transparantere rechtspraak*, Verkenningen Wetenschappelijke raad voor het regeringsbeleid nr. 26, Amsterdam: Amsterdam University Press 2013.

2. Het operationaliseren van kwaliteitsmetingen en een normatief kader

Kwaliteitstoetsing van rechtspraak is niets minder dan het normatief beoordelen van het werk van een normatieve discipline. Dat vraagt een omlijning van zowel het doel daarvan als van het begrip kwaliteit, dat weer vanuit verschillende gezichtspunten benaderd kan worden.

Meestal kristalliseert dat begrip kwaliteit zich pas gaandeweg dat proces van onderzoek verder uit. Conceptueel gezien houdt kwalitatief goede rechtspraak het verwerken van de doelstellingen van rechtspleging in langs de weg van optimale werkprocessen; zo worden de gewenste resultaten gerealiseerd. Grofweg kan men voor twee verschillende benaderingen kiezen: voor een bedrijfsmatig perspectief, d.w.z. voor een vlotte en efficiënte afhandeling van een geschil tegen redelijke kosten. Of juist voor een vaktechnische invalshoek: rechtspreken als het nemen van een beslissing, die juridisch-dogmatisch juist is en op de juiste waarheidsvinding berust. Of is dit eigenlijk een vals dilemma?

Omdat zonder goede feitenvaststelling geen goed oordeel mogelijk is, draagt de rechter zowel juridische als epistemische verantwoordelijkheid.⁸ De procedure dient, vooral formeelrechtelijk, goed gevoerd te worden, maar tegelijk vraagt de waarheidsvinding daarbij ook om aandacht, dat is de epistemische verantwoordelijkheid. Bij een civiele procedure is daarom steeds sprake van een *dubbele* normativiteit; ook de kwaliteitstoetsing vraagt dan een gelijktijdige en evenwichtige beoordeling vanuit beide perspectieven.⁹

Waarheidsvinding lijkt op het eerste gezicht gewoon een kwestie van gewoon je gezonde verstand gebruiken. Maar schijn bedriegt: juist hier doet een veelheid van nadelige invloeden zich gelden.¹⁰ Waarheidsvinding vraagt daarom een eigen plaats en een geëigende methode van onderzoek binnen de juridische procedure. Expliciete aandacht en een beproefde werkwijze daarvoor zijn echter nog onvoldoende binnen de rechtspraak ontwikkeld.¹¹ De epistemische verantwoordelijkheid van de rechter vraagt om een breed niet-juridisch referentiekader, juist ook voor de evaluatie van werkprocessen en hun uitkomsten.¹²

Betrouwbare én betekenisvolle kwaliteitsmetingen zijn steeds bij allerlei werkdomeinen methodologisch een lastige klus gebleken. Als onderwerp van kwaliteitsstudies krijgen we bij rechtspraak te maken met het feit dat het een complex sociaal systeem is met uiteenlopende elementen: zijn structuren, zijn organisatie, de instroom

van zijn cliënten, zijn juridische deelnemers, de adviseurs en deskundigen, de werkprocessen en de uitkomsten daarvan en dat alles in voortdurende wisselwerking met elkaar.¹³ Bovendien vindt rechtspraak plaats in een almaar veranderende politieke en maatschappelijke context. Dan moet gewaakt worden voor een al te simplistische benadering van kwaliteitstoetsing.

Het ene juridische probleem is het andere niet dus vraagt kwaliteitstoetsing steeds te worden toegespitst op de aard van het juridisch probleem in kwestie; zo kan het onderzoek per onderdeel dan veel specifiek en fijnmaziger worden ingevuld.¹⁴ Het keuren van rechtspraak is bovendien rechtsdiscipline overstijgend en vraagt derhalve kennis uit andere domeinen. Kwaliteitstoetsing en -bevordering hebben zich inmiddels als een zelfstandig wetenschappelijk gefundeerd vakgebied ontplooid.¹⁵ Het oorspronkelijke ontwikkel- en werkterrein ervan, industriële productie, is inmiddels sterk verbreed en omvat nu bijvoorbeeld ook dienstverlenende instanties.

Er bestaan verschillende normatieve kaders om producten, diensten en systemen te kunnen beoordelen.¹⁶ De al genoemde meerdimensionale toetsing omvat met name een taxatie van de structuur en de cultuur van de organisatie, de inrichting en uitvoering van de werkprocessen, de uitvoerenden, de bredere context waarbinnen dit alles plaatsvindt (organisatorisch, fysiek, maatschappelijk en politiek) en de evaluatie van de uitkomsten. Dat zijn allemaal facetten van een veelomvattend systeem van onderzoek, *total quality management*.¹⁷ Wat is er tot nu toe aan kwaliteitstoetsing van rechtspraak in Nederland gedaan?

3. Ervaringen met kwaliteitstoetsing van civiele rechtspraak

Er werden in Nederland nog maar enkele studies over kwaliteitstoetsing van civiele rechtspraak verricht.¹⁸ Die analyses werden echter niet of nauwelijks opgezet met gebruikmaking van inzichten uit het *total quality management* (TQM; zie hieronder). Evenmin werd kruisbestuiving gezocht vanuit andere disciplines – een kwestie die overigens beslist niet uniek voor ons land is.¹⁹ De beoordeling van kwaliteit vond tot nu toe steeds binnen de rechterlijke organisatie plaats en werd door rechters zelf uitgevoerd.

De Raad voor de rechtspraak benoemde in 2010 de Toetsingscommissie Civiele Vonnissen en vroeg haar een kader te ontwikkelen waarmee de *ambachtelijke*

8. R.H. de Bock, *Tussen waarheid en onzekerheid: over het vaststellen van feiten in de civiele procedure*, Deventer: Kluwer 2011.

9. Daarover A. Amaya, 'Justification, coherence, and epistemic responsibility in legal fact-finding', *Episteme: A Journal of Individual and Social Epistemology* 2008, vol. 5, p. 306-319.

10. R.J. Allen, 'Taming complexity: rationality, the law of evidence and the nature of the legal system', *Law, Probability and Risk* 2013, vol. 12, p. 99-113.

11. R.W.M. Giard, *Werken aan waarheidsvinding. Over het belang van de juiste onderzoeksmethoden in het aansprakelijkheidsrecht*, Den Haag: Boom Juridisch 2016.

12. Deze problematiek wordt behandeld in: P. Hester & K. Adams (eds.), *Systemic decision making: fundamentals for addressing problems and messes*, Cham: Springer 2017.

13. R.L. Ackoff, 'Systems thinking and thinking systems', *System Dynamics Review* 1994, vol. 10(2/3), p. 175-188.

14. Te denken valt aan categorieën als bijzondere overeenkomsten (nader gespecificeerd), faillissementsrecht, fiscaal recht, intellectueel eigendom, onrechtmatige daad, verzekeringsrecht etc. Zie het voorbeeldformulier in de rapportage van Milar & Rouwendaal 2014 (noot 18), p. 41.

15. Zie voor een oriënterende bibliografie bijvoorbeeld <http://spartan.ac.brocku.ca/~pscarbrough/dfca1stmods/dfca1stmods/biblio/tqmb.html> en voor TQM zie bijvoorbeeld <http://www.dtic.mil/dtic/tr/fulltext/u2/a249837.pdf>.

16. Sinds 1947 bestaat de *International Organization for Standardization* (ISO; <https://www.iso.org/home.html>): 'International Standards make things work. They give world-class specifications for products, services and systems, to ensure quality, safety and efficiency.' Hoe een organisatie haar kwaliteit kan waarborgen, is vastgelegd in de ISO 9000-serie van dit instituut. In Nederland bestaat verder het normalisatie-instituut NEN (=Nederlandse Norm); zie www.nen.nl.

17. Voor een inleiding daarover zie: J.S. Oakland, *Total quality management and operational excellence: text with cases*, London: Routledge 2014 (4e ed.).

18. Dat zijn: P. Neleman & P.J. Neijt, *Toetsingscommissie Civiele Vonnissen*, Den Haag: Raad voor de rechtspraak 2010; R.H. de Bock & L.R. Harinxma thoe Slooten, *Pilot Kwaliteitstoetsing Civiele Vonnissen*, Den Haag: Raad voor de rechtspraak 2012; O.G. Milar & J.W. Rouwendaal, *Rechtspraakbrede toetsing van civiele vonnissen*, Den Haag: Raad voor de rechtspraak 2014.

19. W.H. Simon, 'Where is the "quality movement" in law practice?', *Wisconsin Law Review* 2012, vol. 2, p. 387-406.

kwaliteit van civiele vonnissen van rechtbanken kan worden getoetst:

'Het beoordelingskader zou de ambachtelijke kwaliteit van een vonnis centraal moeten stellen, waarbij valt te denken aan: vaktechnisch goed in elkaar gezet, leesbaar en helder, niet langer of diepgravender dan nodig, consistent, procedureel- en materieelrechtelijk foutloos, uitmondend in een overtuigende en aanvaardbare beslissing'.²⁰

Zo geformuleerd, gaat het hier bij een dergelijke toetsing om te realiseren *waarden* – een conceptualisatie waarbij gevraagd wordt of de dingen wel goed werden gedaan. Maar hoe staat het met de operationalisering hiervan? Genoemde kwaliteitstoetsingsstudies richtten zich op het *vonnis*, dus in essentie op de eindfase van de procedure. Er werd een systematiek ontwikkeld voor het toetsen van de kwaliteit van het ambachtelijke proces dat tot de vonnissen leidt aan de hand van vragenlijsten, die later nader werden uitgewerkt en aangescherpt.²¹ De onderzoekers hanteren een gestructureerde vragenlijst met uitsluitend gesloten vragen. De hoofdonderwerpen zijn: (1) vaststelling van de onbetwiste relevante feiten, (2) bewijzen en informatiegaring, (3) comparitie na antwoord, (4) juridische beoordeling, (5) motiveren en beslissen, (6) inkleding van het vonnis/de vonnissen en (7) een tekstueel eindoordeel als slotvraag.²²

Binnen de onderdelen 1 t/m 6 werden vervolgens meerdere subonderwerpen geformuleerd, die moesten worden beoordeeld met een score op de rangordeschaal: slecht – onvoldoende – voldoende – goed – n.v.t. Als aanvulling op de vragenlijst werd een *Toelichting toetsingsformulier civiele vonnissen* gegeven waarop tekst en uitleg werd gegeven.²³ Deze aanwijzing is vooral conceptueel over *wat* er beoordeeld moet worden maar niet of nauwelijks over het *hoe* daarvan.

Bij het onderdeel over feitenvaststelling wordt bijvoorbeeld gevraagd om drie items te willen scoren: (a) de kwaliteit van de feitenvaststelling in het algemeen, (b) de nauwkeurigheid van de feitenvaststelling en (c) de volledigheid van de feitenvaststelling. De toelichting vermeldt dat elk vonnis een opsomming van de *relevante* feiten dient te bevatten, die tussen partijen als vaststaand hebben te gelden. En bij het laatste item gaat het om de volledigheid van de feitenvaststelling: ontbreken er onbetwiste feiten, die relevant zijn voor de beslissing en/of het debat tussen partijen? Maar wat bepaalt de relevantie van feiten? Voor het beoordelen van de waarheidsvinding bestaat een aparte methodologie.²⁴

Hoe zit het met de overige onderwerpen als bewijzen, de uitvoering van een comparitie na antwoord, de juridi-

sche beoordeling, het motiveren en beslissen en de inkleding van het vonnis? Welke maatstaven worden hiervoor gehanteerd? Hoe onderscheidt de kwaliteitsbeoordelaar hier juist van onjuist? Het oordelen hierover lijkt eerder intuïtief dan rationeel te zijn en doet denken aan de impliciete en vooral intuïtieve maatstaf van VS-opperrechter Potter Stewart: 'I know it when I see it'.²⁵

De studie van Milar & Rouwendaal uit 2014 omvatte een selectie die bestond uit 632 civiele vonnissen waartegen hoger beroep was ingesteld. Het oordeel hierover was dat de ambachtelijke kwaliteit in 81% van de gevallen goed of voldoende was. Bij één op de vijf uitspraken had naar hun oordeel de eerste feitenrechter steken laten vallen maar een goede analyse naar de oorzaken daarvan ontbrak. Hun eindconclusie was dat de toetsing van de ambachtelijke kwaliteit van civiele vonnissen, indien zij periodiek wordt herhaald en volgehouden, meerwaarde heeft – maar welke dan? Wat waren de concrete leereffecten? Expliciete verbeterpunten werden niet aangegeven. Bovendien was hier sprake van een evidente selectiebias: alleen die zaken waarvoor appel was gevraagd. Zo'n selectie beperkt de generalisering van de uitkomsten sterk. Advocaten vervullen een essentiële rol bij de procedure en hun rol dient bij kwaliteitstoetsing ook onderzocht te worden, maar dat heeft tot nu toe nauwelijks aandacht gekregen. In het civiele recht wordt partijautonomie als een belangrijk principe gezien.²⁶ De houding, inzet en inbreng van partijen zijn daarmee een belangrijke – wellicht zelfs beslissende – kwaliteitsfactor en die verdienen daarom expliciet aandacht bij kwaliteitstoetsing.²⁷ Dat punt moet worden meegenomen bij evaluaties.

4. De basisprincipes van kwaliteitstoetsing en -bevordering

Kwaliteitstoetsing is de kritische reflectie over eigen en andermans handelen langs de weg van systematisch en methodisch onderzoek. Het is geen doel op zich maar steeds een middel om welomschreven doelen te realiseren, bijvoorbeeld om te kunnen leren van deze ervaring of om, waar nodig, het rechtssysteem te veranderen en te verbeteren. Voor kwaliteitstoetsing is informatie verkregen uit terugkoppeling essentieel. Maar welke terugkoppeling waarop?

4.1. Structurele terugkoppeling

Feedback is bijvoorbeeld te realiseren via het model van *triple-loop learning* met drie aangrijpingspunten (figuur 1).²⁸ Het eerste niveau behelst reflectie over de directe uitvoering van de taak: deden we de dingen wel goed? Het tweede niveau houdt reflectie over de gehanteerde werkmethode in: deden we wel de goede dingen?

20. Zie Neleman & Neijt 2010 (noot 18), p. 5.

21. R.H. de Bock & L.R. Harinxma thoe Slooten, *Pilot Kwaliteitstoetsing Civiele Vonnissen*, Den Haag: Raad voor de rechtspraak 2012.

22. Zie bijlage 3 van De Bock & Harinxma thoe Slooten 2012.

23. Zie bijlage 4 van De Bock & Harinxma thoe Slooten 2012.

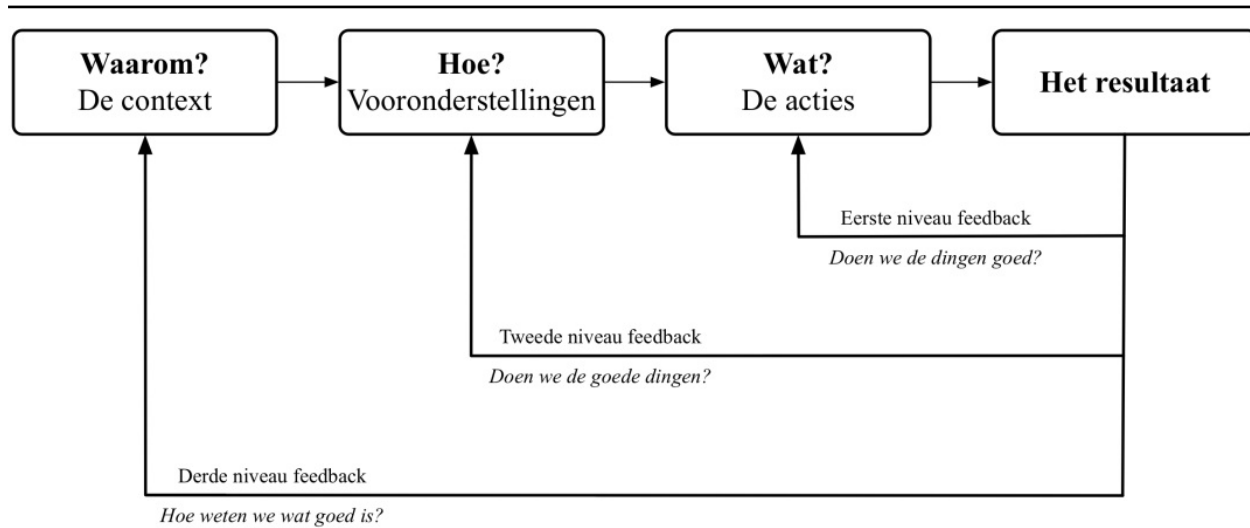
24. Zie over die aspecten R.W.M. Giard, 'Rechtspraak: waarheidsvinding als een (zelf)kritische praktijk', *Rechtstreeks* 2014, afl. 3/4, p. 16-45.

25. In US Supreme Court *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1964). Zie ook hierover P. Gewirtz, 'On "I know it when I see it"', *The Yale Law Journal* 1996, vol. 105, p. 1023-1047.

26. Zie wat dat betreft de artikelen 24 en 149 lid 1 Rv.

27. Zie bijvoorbeeld over dit probleem H.M. Kritzer & N. Vidmar, 'Lawyers on trial: juror hostility to defendants in legal malpractice trials', *Hofstra L. Rev.* 2015, vol. 45, p. 375-399.

28. Zie over deze benadering M.F. Peschl, 'Triple-loop learning as a foundation for profound change, individual cultivation, and radical innovation. Construction processes beyond scientific and rational knowledge', *Constructivist Foundations* 2007, vol. 2(2/3), p. 136-145; P. Tosey et al., 'The origins and conceptualizations of "triple-loop" learning: a critical review', *Management Learning* 2012, vol. 43(3), p. 291-307; M. Reynolds, 'Triple-loop learning and conversing with reality', *Kybernetes* 2014, vol. 43(9/10), p. 1381-1391.



Figuur 1: De 'triple-loop learning'.

Het derde niveau strekt ertoe om vanuit een meta-perspectief naar de rechtspraktijk kijken: hoe beoordelen we die en hoe weten we wat juist is?²⁹ Is bijstelling nodig van de werkmethode of moeten we ook de theoretische grondslagen van de praktijk herijken?

Er zijn verschillende enkel- of meervoudige terugkoppelsystemen beschreven.³⁰ Welk model men ook kiest, steeds is er goed empirisch onderzoek van de praktijk nodig van werkprocessen en de uitkomsten daarvan.³¹

4.2. Uitkomsttoetsing

Een rechter kan in beginsel een vordering ten onrechte toe- of afwijzen. Zo'n 'misser' kan grosso modo berusten op onjuiste feitelijke en/of op onjuiste juridische grondslagen voor het oordeel. Om mislagen in de civiele rechtspraak te herkennen heb je uitkomsttoetsing nodig aan de hand van een onafhankelijke betrouwbare standaard. Zo'n gouden standaard bestaat helaas niet. Er kunnen wel drie benaderingsmogelijkheden worden gebruikt.

De eerste is als er rechtsmiddelen worden aangewend en de resultaten van de eerste rechter worden vergeleken met die van de tweede feitenrechter of eventueel met die van de cassatierechter. Maar kun je met de uitspraak van de appel- en cassatierechter de eerste feitenrechter wel deugdelijk de maat nemen? 'Omdat in hoger beroep echter nieuwe feiten mogen worden aangevoerd en vorderingen mogen worden gewijzigd, wordt in hoger beroep vaak een andere situatie en/of vordering beoordeeld dan de rechtbank heeft gedaan. De vernietiging of bekrachtiging door een hof van een uitspraak van de rechtbank zegt daarom vaak weinig over de kwaliteit van deze uitspraak,' aldus De Bock & Harinxma thoe Slooten.³²

De waarde hiervan lijkt zo bezien dus beperkt en bovendien wordt bij meer dan 90% van de beslissingen van de eerste feitenrechter geen appel ingesteld. Dit is dus geen systematische vorm van toetsing. Ondanks genoemde beperkingen is er echter nog nooit een poging gedaan de uitkomsten van civiele appel- en cassatieprocedures systematisch met die van de eerste feitenrechter te vergelijken en eventuele verschillen te onderzoeken.

De tweede mogelijkheid is om de procedure te repliceren door een andere rechter het dossier voor te leggen maar beslist zonder kennis van het eindoordeel of eventuele tussentijdse beslissingen (outcome bias!). Dergelijke simulaties zijn in de VS uitgevoerd met *mock juries* maar generalisaties naar ons rechtssysteem zijn lastig vanwege de slechte vergelijkbaarheid van hun rechtspraak met de onze. Zo'n type studie is mij niet bekend in Nederland. De vraag is bovendien hoe zulke *mock juries* zouden moeten opereren en hoe valide de conclusies uit een dergelijk onderzoek zijn.³³ Een geheel nieuwe mogelijkheid is om met behulp van kunstmatige intelligentie zaken opnieuw te analyseren en te beoordelen.³⁴

Als laatste bron kunnen annotaties genoemd worden: op welke punten bekritiseert de annotator de beslissing en op welke gronden? Dat is wederom hooguit een incidentele en zeker geen systematische kwaliteitstoetsing. Noodgedwongen zal bij kwaliteitstoetsing van rechtspraak het accent daardoor verschuiven van uitkomsttoetsing naar processtoetsing.

4.3. Total quality management

Een van de belangrijkste grondleggers van kwaliteitstoetsing- en bevordering was de statisticus en organisatie-deskundige William Edwards Deming (1900-1993).³⁵ Hij muntte het reeds ter sprake gekomen begrip *total*

29. A.G.L. Romme & A. van Witteloostuijn, 'Circular organizing and triple loop learning', *Journal of Organizational Change Management* 1999, vol. 12(5), p. 439-454.

30. Zie daarover B.L. Simonin, 'N-loop learning: part I – of hedgehog, fox, dodo bird and sphinx', *The Learning Organization* 2017, vol. 24(3), p. 169-179.

31. B.L. Simonin, 'N-loop learning: part II – an empirical investigation', *The Learning Organization* 2017, vol. 24(4), p. 202-214.

32. De Bock & Harinxma thoe Slooten 2012, p. 4.

33. B.H. Bornstein, 'The ecological validity of jury simulation: is the jury still out?', *Law and Human Behavior* 1999, vol. 23(1), p. 75-91.

34. Zie als voorbeeld daarvan D.A. Martin et al., 'Supreme Court decision making', *Perspectives on Politics* 2004, vol. 2(4), p. 761-767.

35. William Edwards Deming (1900-1993). Zijn gedachtegoed zette hij uiteen in *The new economics for industry, government, education*, Cambridge (MA): MIT Press 1993 en in *Out of the crisis*, Cambridge (MA): MIT Press 2000.

quality management (TQM).³⁶ Bij TQM dient het functioneren van het systeem diepgaand en in al zijn facetten te worden doorgrond.³⁷ Daarvoor wordt de term *system of profound knowledge* gebruikt, een systeem dat rust op vier pijlers. Die benadering is ook toepasbaar voor rechtspraak.³⁸

In de eerste plaats zal er onderzoek gedaan worden naar het functioneren van het systeem van rechtspraak in zijn totaliteit.³⁹ De toetsing van kwaliteit mag zich dus niet beperken tot het handelen van de rechter of tot diens vonnis. De volgende stap is inzicht te krijgen in de reproduceerbaarheid van waarnemingen en oordelen, zowel inter- als intra-individueel. Dat is belangrijk om te weten of het oordeel over hetzelfde probleem bij herhaling bestendig is of juist niet – de eis van rechtseenheid. Daarnaast impliceert het uitvoeren van werk interacties tussen mensen en de manier waarop de informatie wordt verwerkt. Dat is het terrein van respectievelijk de sociale en de cognitieve psychologie. Bij elk menselijk handelen liggen talrijke valkuilen en dwaalwegen op de loer.⁴⁰ Dat is het probleemgebied van heuristieken en biases, die de uitkomsten van onderzoeken en oordelen kunnen vertekenen.⁴¹ Een cruciaal punt bij werkprocestoetsing is bijvoorbeeld de outcome bias: hoe werkt kennis van de rechterlijke beslissing door op de werkprocestoetsing door rechters van het hof?⁴²

Is er bij het vaststellen van de feiten en omstandigheden van het geval bijvoorbeeld gelet op de invloeden van

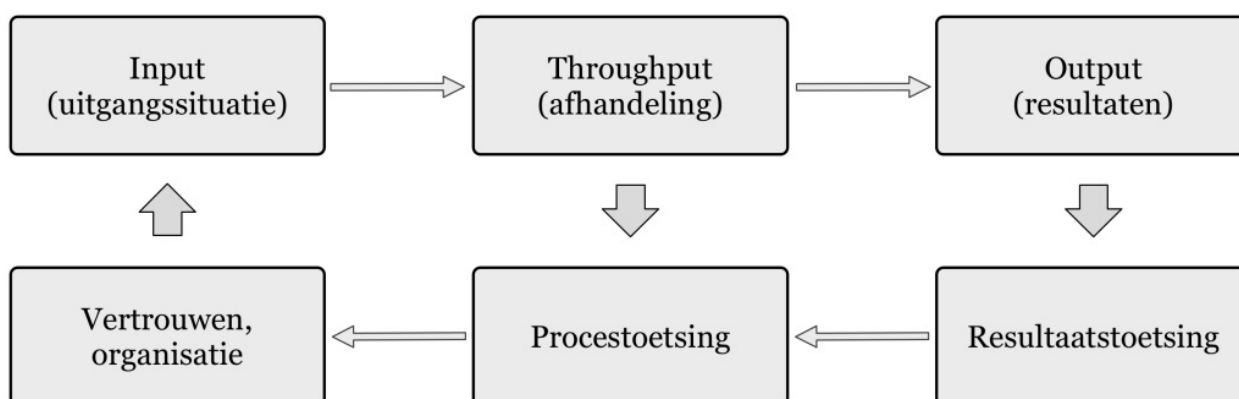
een eventuele bias? Als dat niet zo is, bestaat de kans dat feiten te selectief werden verzameld (work-up bias), de oorzaak verkeerd werd ingeschat (hindsight bias), dat er tunnelvisie optreedt (bevestigingsbias) en niet was nagedacht over de vraag welke informatie men hier echt nodig had (information bias).

Ten slotte zijn er nog de kennistheoretische aspecten, zowel die wat betreft de fundering van de werkprocessen als de toetsing daarvan. Juristen onderschatten bij procedures vaak de complexiteit van de realiteit die zij moeten onderzoeken en waarover de rechter moet beslissen.⁴³ Een juiste inschatting van het probleem vraagt de uitkomsten van empirisch onderzoek, die als achtergrondkennis nodig zijn voor het opzetten van het onderzoek, de interpretatie ervan en het formuleren van een oordeel.⁴⁴

4.4. Onderzoek heel het systeem

Voor het onderzoeken van rechtspraak als *systeem* kan de driedeling *input – throughput – output* worden gehanteerd (figuur 2).⁴⁵

Wat de *input* betreft is het belangrijk te realiseren dat het grootste deel van de problemen buitengerechtelijk wordt opgelost.⁴⁶ Door dit selectieproces ontstaat het fenomeen *spectrum bias*: de rechter krijgt bij een omschreven probleem juist te maken met een complexere problematiek en/of zich lastiger gedragende procespartijen.⁴⁷



Figuur 2: De hoofdonderdelen van kwaliteitstoetsing.

36. B. Aquilani et al., 'A systematic literature review on total quality management critical success factors and the identification of new avenues of research', *The TQM Journal*, 29 nr. 1, 2017: 184-213.

37. W.E. Deming, *The new economics: for industry, government, education*, Cambridge (MA): MIT Press 1993, m.n. hfdst. 3 en 4.

38. R.D. Moen & N.L. Clifford, 'Always applicable', *Quality Progress* 2016, vol. 49(6), p. 46-53.

39. In productieprocessen worden daarbij ook de toeleveranciers van producten betrokken; in meer immateriële zin geldt dat ook voor dit type analyse – de toeleveranciers van professionals (universiteiten, beroepsopleidingen) en kennis vanuit observationeel of experimenteel onderzoek.

40. Een bias is een intuïtieve vooringenomenheid. Zie voor een overzicht https://en.wikipedia.org/wiki/List_of_cognitive_biases. Er zijn inmiddels meer dan 200 vormen van biases gesignaleerd. Daarover D. Chavalarias & J.P.A. Ioannidis, 'Science mapping analysis characterizes 235 biases in biomedical research', *Journal of Clinical Epidemiology* 2010, vol. 63(11), p. 1205-1215.

41. Zie daarover algemeen A. Tversky & D. Kahneman, 'Judgment under uncertainty: heuristics and biases', *Science* 1974, vol. 185, p. 1124-1131 en voor rechtspraak J.J. Rachlinski, 'Heuristics and biases in the courts: ignorance or adaptation', *Or. L. Rev.* 2000, vol. 79, p. 61-102; E. Peer & E. Gamliel, 'Heuristics and biases in judicial decisions', *Court Review* 2013, vol. 49(2), p. 114-118.

42. J. Baron & J.C.C. Hershey, 'Outcome bias in decision evaluation', *Journal of personality and social psychology* 1988, vol. 54(4), p. 569-579.

43. Zie daarover o.a. R.J. Allen, 'Taming complexity: rationality, the law of evidence and the nature of the legal system', *Law, Probability and Risk* 2007, vol. 63, p. 59-61 en L.P. Moyer, 'The role of case complexity in judicial decision making', *Law and Policy* 2012, vol. 34(3), p. 291-312.

44. Zie in W.H. van Boom et al. (red.), *Capita civilologie. Handboek empirie en privaatrecht*, Den Haag: Boom Juridische uitgeverij 2013 bijvoorbeeld de hoofdstukken over 'Informed consent' (hfdst. 9) of 'Bewijswaardering en psychologische inzichten' (hfdst. 35).

45. Zie over deze systeembenadering K.M. Adams et al., 'Systems theory as the foundation for understanding systems', *Systems Engineering* 2014, vol. 17, p. 112-123 en M.C. Edson et al. (eds.), *A guide to systems research philosophy, processes and practice*, Singapore: Springer Science+Business Media 2017.

46. M.J. ter Voert & C.M. Klein Haarhuis, *Geschiedbeslechtingdelta 2014*, Den Haag: Boom Lemma uitgeverij 2015.

47. Spectrum bias betekent selectie vanuit een brede waaier van verschijningsvormen van hetzelfde probleem tot een beperkt deel van het spectrum daarvan.

Procestoetsing is de fase van *throughput*. Wat staat nu bij procestoetsing centraal? Eerst is conceptuele klaarheid nodig, dan volgt de fase van operationalisering. Wat is bij de gegeven casus nu precies het probleem en waarom dan deze procedure? Om welke verschillende werkprocessen gaat het hierbij en hoe verhouden die zich tot elkaar? Dan volgt het nauwkeurige onderzoek daarvan. Bestaat daarvoor een *one-size-fits-all* benadering of moeten we voortdurend maatwerk leveren? Welk normatief kader staat ons voor procestoetsing ter beschikking? Daarvoor kunnen we de funderende vierslag van TQM gebruiken zoals die reeds in paragraaf 4.3 is weergegeven.

De evaluatie van de *output* kwam reeds in paragraaf 4.2 ter sprake. Hoe moet het nu verder gaan met de kwaliteitstoetsing en -bevordering van civiele rechtspraak?

5. Beschouwing

De essentie van kwaliteitstoetsing is het doen van kritisch (zelf)onderzoek van werkprocessen en uitkomsten. Met zo'n evaluatie kan rechtspraak zich legitimeren langs de wegen van procestoetsing (wordt er wel goed gewerkt?) en resultaatstoetsing (wat is het resultaat nu waard?).⁴⁸ Tegelijk is kwaliteit een weerbaarstig begrip en inzichten daarover ontwikkelen zich vaak pas 'werkende weg'. Het is zeker te waarderen dat er initiatieven zijn genomen om kwaliteitstoetsing en -bevordering van civiele rechtspraak te gaan exploreren, maar dat was pas het begin. Die activiteiten hebben na 2014, zoals door Tjeenk Willink benadrukt, jammer genoeg geen vervolg gekregen.

Bij opzet en uitvoering van die toetsingsprojecten van rechtspraak zijn bovendien enkele kritische kanttekeningen te plaatsen. De activiteiten richtten zich voornamelijk op de inhoudelijke aspecten van de rechterlijke beslissing – de uitkomst.⁴⁹ De benadering was voornamelijk conceptueel en men heeft onvoldoende stilgestaan bij de zo noodzakelijke gedetailleerde operationalisering. Zo'n onderzoek vraagt bovendien steeds om een helder – dus expliciet – normatief kader.

Men heeft gekozen voor een serie van diverse civiele zaken. Een meer thematisch gericht onderzoek van omschreven kwesties is waarschijnlijk effectiever dan de toetsing van een breed scala van zaken, bovendien geselecteerd omdat er appel was ingesteld – een duidelijke selectiebias. Zo'n pluraliteit van vraagstukken maakt het lastig om specifieke conclusies te trekken.

De evaluatiestudies verwijzen nauwelijks naar een wetenschappelijke onderbouwing van hun werkwijze. Als er al verwijzingen zijn, betreft dat uitsluitend Nederlandse juridische publicaties maar niet naar ervaringen in andere landen. Waar meer aandacht voor zou mogen bestaan, is het gebruik van kennis uit andere disciplines – met name uit het kennisdomein van de *quality move-*

ment. Zo wordt er te veel rondgedraaid in eigen juridische kring. Tegelijk dringt zich dan het beeld op van de slager die het eigen vlees keurt.⁵⁰ Juist bij procestoetsing bestaat het gevaar dat tekortkomingen in de werkwijze tot fouten leiden, maar omdat men het gewend is om het nu eenmaal zo te doen, worden diezelfde onvolkomenheden nadien niet opgemerkt – en dat is een 'dubbele vloek'.⁵¹

De besproken Nederlandse kwaliteitstoetsingsstudies zijn door de toetsers aangemerkt als 'peer review'. Zo'n type opzet vraagt echter zowel een doordachte en beproefde methodologie als een goede training van de beoordelaars.⁵² Maar beoordeling door gelijken mag niet het enige instrument zijn voor de kwaliteitsbeoordeling. Juist onderzoek van het reilen en zeilen van de rechtspraak vraagt om een multidisciplinaire benadering en is zeker ook gebaat met een goed georganiseerde kritische blik van buitenaf.

Een belangrijk verworven inzicht is dat kwaliteitstoetsing alomvattend – dus integraal – dient te zijn. Er wordt een *systeem* getoetst en daarbij is het wezen van dat systeem ook belangrijk: dat is complex van aard en dat vraagt daarom weer om een gepaste benadering. Er werd hoege-naamd geen aandacht geschonken aan mogelijke systematische fouten bij het uitvoeren van de Nederlandse toetsingsstudies; er zijn wat dat betreft goed onderbouwde en waardevolle inzichten uit de cognitieve psychologie, die genoemd noch gebruikt worden. Met name de evaluatie van het cruciale onderdeel waarheidsvinding is in mijn ogen daarom op methodologische gronden beslist insufficiënt.

Het succes van kwaliteitstoetsing staat of valt ook met de kunst van het stellen van de juiste vragen voor het onderzoek, de erotetiek. De gehanteerde vragenlijsten voor het onderzoek zijn te veel geïnspireerd door te realiseren waarden en te weinig gericht op de praktische uitvoering van het onderzoek zelf. Bij de gehanteerde methodologie ligt de nadruk op conceptualisatie, niet op operationalisering: er wordt wel geschetst *wat* moet worden onderzocht maar niet goed *hoe* dat moet. Het ontwikkelen van een op praktische toetsing toegesneden methodologie is dus hard nodig en die vraagt weer gespecialiseerde niet-juridische kennis.⁵³

Tot nu is er alleen aandacht geweest voor retrospectieve studies. Die hebben hun inherente beperkingen, vooral de kwaliteit van de dossiers is daarbij sterk bepalend. Ze vragen ook veel tijd. Als voorbereiding op toekomstige retrospectieve studies zal meer aandacht besteed mogen worden aan een goed georganiseerde kennisinfrastructuur voor kwaliteitstoetsing en dat ontbreekt tot nu toe. Ook dient nagedacht te worden over de inrichting van dossiers en de opbouw en motivering van vonnissen met het doel die door hun eenvoudigheid toegankelijker te maken voor kwaliteitstoetsing. Kleinschalige experimen-

48. Zie daarover, met name vanuit het perspectief van systeembenadering, het essay van M. Bokhorst & W. Witteveen, 'Als gezag verdiend moet worden...', in: D. Broeders et al. (red.), *Speelruimte voor transparantere rechtspraak*, WRR Verkenningen nr. 26, Amsterdam: Amsterdam University Press 2013, p. 127-164.

49. Zoals bepleit door R.H. de Bock, 'Grip op kwaliteit; een model voor inhoudelijke kwaliteit van rechterlijke beslissingen', in: *Handelingen NJV*, Deventer: Wolters Kluwer 2015, p. 27-140.

50. Zie hfdst. 4, 'Binnen en buiten: over transparantiewensen en het aanbod van de rechtspraak', in: D. Broeders et al. (red.), *Speelruimte voor transparantere rechtspraak*, WRR Verkenningen nr. 26, Amsterdam: Amsterdam University Press 2013, p. 69-70.

51. Zie over die dubbele last van onwetendheid D. Dunning, 'The Dunning-Kruger Effect. On being ignorant of one's own ignorance', *Advances in Experimental Social Psychology* 2011, vol. 44, p. 247-296.

52. Zie voor de praktische problemen D.E. Ho, 'An experiment of experimentalism: does peer review work?', *Stanford Law Review* 2017, vol. 69, p. 1-121.

53. Zie daarvoor de publicaties van methodoloog Piet Verschuren zoals *De probleemstelling voor een onderzoek*, Houten: Het Spectrum 2011; *Praktijkgericht onderzoek*, Den Haag: Boom Lemma uitgeverij 2011 en *Kernthema's in de methodologie. Op weg naar beter onderzoek*, Amsterdam: Amsterdam University Press 2017.

tele studies kunnen helpen de operationele aspecten beter vorm te geven.

Juist het doen van prospectieve studies is wenselijk: die zijn dan op omschreven problemen gericht. Dan heeft men de gelegenheid om ex ante te formuleren welke gegevens men nodig heeft, hoe die vastgelegd gaan worden, welk operationeel beoordelingskader men daarvoor wil gaan gebruiken. Ook is dan duidelijk welke kennis hierop van toepassing is en wat daaruit voortvloeit bijvoorbeeld wat betreft waarheidsvinding en oordelen. Daarbij zal, nogmaals, niet geschroomd moeten worden de hulp in te roepen van andere disciplines zodat het moderne recht als kruispuntwetenschap wordt gerealiseerd.⁵⁴

Wat betreft het verder ontwikkelen van kwaliteitstoetsing en -bevordering van civiele rechtspraak moet nog veel gedaan worden; we staan eigenlijk pas aan het begin daarvan. De politiek zou het belang hiervan meer mogen steunen door meer middelen beschikbaar te stellen voor kwaliteitsonderzoek van deze belangrijke maatschappelijke functie.

54. Zie daarover W.H. van Boom, *Door meten tot weten – over rechtswetenschap als kruispunt*, Den Haag: Boom Juridische uitgevers 2015.