

evalueren

Evalueren van rampen en crises

Deel 2. Waar men zoal tegenaanloopt en hoe het wellicht beter kan.

Menno J. van Duin

In het eerste deel (RBOH, afl. 3, 2012) is gekeken naar de volgende vijf vragen:

1. waarom worden er evaluaties verricht van rampen en crises?
2. wie voeren deze evaluaties uit en wie bepalen (opdrachtgever?) dat er onderzoek wordt verricht?
3. wat wordt er geëvalueerd; het soort van gebeurtenissen waar evaluaties naar worden verricht?
4. wanneer worden deze evaluaties verricht (meestal direct erna maar niet altijd)?
5. Ten slotte de vraag hoe die evaluaties geschieden (in de regel schriftelijk maar je ziet hier ontwikkelingen in).

In dit tweede deel wordt eerst een aantal thema's en onderscheiden besproken die onlosmakelijk gekoppeld zijn aan het thema van evalueren. Deze thema's sluiten deels aan bij de hoe-vraag: hoe wordt invulling gegeven aan het evaluatieproces? Wat zijn de achterliggende dilemma's en overwegingen die bij het doen van evaluaties een rol spelen? Vervolgens wordt kort stil gestaan bij het buitenland; hoe doet men het daar?

De bestaande praktijk kent verschillende beperkingen die gepresenteerd worden. Ten slotte wordt in het laatste deel een aantal adviezen besproken die, naar de mening van de auteur, de kwaliteit van het evalueren zou verhogen. Uiteraard is hierover discussie mogelijk. Deze zou dan ook gevoerd moeten worden.

Thema's en onderscheiden

Onafhankelijk of niet?

Er is altijd veel discussie over de wenselijkheid van onafhankelijk onderzoek. Wat de kenmerken van onafhankelijk onderzoek precies zijn, blijft altijd wat lastig. In het verleden werd het COT geprezen om zijn onafhankelijke onderzoek, maar aangezien dit betaald wordt door de opdrachtgevers werd gaandeweg afscheid genomen van het idee dat hier sprake was van onafhankelijkheid. Soms worden dergelijke opdrachtonderzoeken overigens gewoon betiteld als onafhankelijk. Minister Opstelten stelt een onafhankelijk onderzoek in naar de aanbesteding van de aanschaf van politieauto's (NRC Handelsblad, 7 november 2011). De inspecties en de Onderzoeksraad voor Veiligheid (OvV) worden thans vaak als onafhankelijke onderzoeksinstanties gezien, maar zeker over de inspecties is er discussie mogelijk. Hoe onafhankelijk is Inspectie voor openbare orde en veiligheid (IOOV) wanneer zij onderzoek doet naar een incident terwijl de IOOV ook de betreffende regio's visiteert en een rol heeft bij het toezicht op de kwaliteit van het personeel van de hulpdiensten?

Als de strenge norm van onafhankelijkheid wordt gevolgd dan blijft waarschijnlijk alleen nog de OvV over, maar zelfs dan is er vast nog wel iets op te merken. Zonder meer dragen verschillende procedurele aspecten bij aan de onafhankelijkheid. De Raad wordt rechtstreeks gefinancierd, is niet ondergebracht bij een ministerie, bepaalt zelf wat het onderzoekt, kent heldere procedures van hoor- en wederhoor en zo verder. Toch zijn er ook beperkin-

gen. De Raad maakt bij evaluaties vaak gebruik van deskundigen uit het veld. Daarmee is er mogelijk sprake van verwevenheid tussen onderzoekers en het onderzoeksobject. Er is in ieder geval sprake van een klein wereldje van personen die binnen de wereld van fysieke veiligheid (onderzoeksmatig) rondlopen. Overal komt men elkaar tegen. Daarnaast moet ook de OvV zichzelf (blijven) bewijzen om op lange termijn het eigen voortbestaan te garanderen. Als bijvoorbeeld deze organisatie jaren achtereen geen spraakmakend rapport publiceert, zal extern de roep groter worden de financiële bijdragen te verminderen. Intern kan dit er bijvoorbeeld toe leiden dat bij een nieuwe evaluatie meer werk worden gemaakt om maar 'te scoren'. Het is allemaal hypothetisch maar echte, volledige onafhankelijkheid is – in een zo kleine wereld – een illusie.

De vraag is of de onafhankelijkheid van de onderzoeker zo belangrijk is. In plaats van onafhankelijkheid is transparantie, helder maken wat en hoe onderzocht wordt, belangrijker. Feitelijk gaat het om dezelfde kenmerken als voor wetenschappelijk onderzoek; betrouwbaarheid, geldigheid, inzicht in de methodiek en dergelijke, opdat anderen die eenzelfde onderzoek zouden uitvoeren (reproducen) tot vergelijkbare resultaten zouden komen.

Leren of verantwoorden?

Verantwoorden hoort bij een democratische rechtstaat. Bestuurders, de verantwoordelijken, leggen in het parlement (of gemeenteraad) verantwoording af van hun daden en beslissingen. Dat geldt dagelijks (of jaarlijks ten aanzien van de uitgaven), maar zeker na bijzondere gebeurtenissen. Muller (2001, p. 90) onderscheidt vijf vormen van politieke verantwoording:

- non-verantwoording (afwezigheid van verantwoording door ontbreken van contact);
- informeren (inlichten zonder verdere discussie);

- aanspreken (interactie tussen bestuurders en het forum);
- aansprakelijk stellen (sancties verbonden aan functioneren van bestuurders);
- volledige verantwoording (gehele proces doorlopen en aan het einde bepalen op welke wijze de uitkomsten worden geïmplementeerd).

Een onderscheid dat vaak gemaakt wordt bij evaluaties, betreft leren – er beter van worden – en verantwoorden: rapporteren ten behoeve van het verantwoordingsproces (soms zelfs in termen van goed en fout). Het lijkt een duidelijk onderscheid, maar in de praktijk is het buitengewoon fluïde.

Sommige rapporten die – volgens de opstellers – helemaal geschreven zijn vanuit het oogmerk van leren, kennen wel erg stevige conclusies. Bijvoorbeeld het rapport¹ van de IOOV bij haar evaluatie van Moerdijk:

- De Inspectie constateert dat de regio... haar oefenbeleidsprogramma en oefeningen onvoldoende heeft afgestemd ... op de risico's (p. 11).
- De Inspectie komt tot het oordeel dat de gemeente en de regio ... bij de voorbereiding steken heeft laten vallen en onvoldoende invulling heeft gegeven aan hun verantwoordelijkheden (p. 12).
- De Inspectie constateert dat de veiligheidsregio ... onvoldoende adequaat heeft bestreden (p. 17).

Tegelijkertijd lezen wij ook in het betreffende rapport: "doel van de rapportage is niet om schuldigen aan te wijzen, aansprakelijkheden vast te stellen of organisaties of functionarissen af te rekenen" (p. 4).

¹ Brand Chemie-Pack Moerdijk, Een onderzoek naar de bestrijding van (de effecten van) het grootschalig incident. Publicatienr: j-9549

Ook de Onderzoeksraad kan stevige conclusies formuleren, zoals bleek bij hun rapport over de cellenbrand (Schiphol/Haarlemmermeer)².

- De organisatie van het cellencomplex Schiphol-Oost was onvoldoende voorbereid en ingericht op het optreden bij brand (p. 174)
- De hoofddirectie van de DJI draagt de verantwoordelijkheid voor de veiligheid van celbewoners en personeel. In de dagelijkse praktijk is de locatiedirecteur van het cellencomplex, conform de wet, de verantwoordelijke voor veiligheid waaronder brandveiligheid, maar heeft deze taak onvoldoende ingevuld (p. 175)
- Zowel de hoofddirectie van de DJI als de locatiedirecteur hebben onvoldoende aantoonbaar gebruik gemaakt van de aanwezige signalen over de tekortkomingen in brandveiligheid van het cellencomplex, ook naar aanleiding van eerdere branden (p. 176).
- De RGD heeft, vanuit zijn verantwoordelijkheid voor een brandveilig gebouw, zijn taak onvoldoende invulling gegeven ... (p. 177).

Een andere evaluatie van de Onderzoeksraad: de verdronken duiker bij Terneuzen³:

- Op landelijk niveau – over alle duikende korpsen bezien – zijn de NVBR, de IOOV, de AI, de werkgevers en het systeemverantwoordelijk Ministerie van BZK er niet in geslaagd de gemeenten ertoe aan te zetten de gesignaleerde tekortkomingen weg

te nemen en ze daarbij te ondersteunen (p. 83).

- Het duikteam van de brandweer Terneuzen was onvoldoende voorbereid op de duikinzet. Zowel voorafgaande als tijdens de inzet zijn de risico's van de weersomstandigheden, het gebruik van de seinlijn en het gebruik van ademlucht in relatie tot de opdracht en de veiligheid van de duikers onvoldoende onderkend (p. 84).

Ook ad-hoc commissies kunnen stevig uitpakken zoals de parlementaire enquêtecommissie (Bijlmerramp)⁴:

- De berging van elementen van het vliegtuigwrak is onzorgvuldig verlopen (p. 397).
- Het onderzoek door de rijksrecherche naar de mannen in de witte pakken is onzorgvuldig verlopen (p. 397)
- De Commissie constateert dat de Rijkspolitie op Schiphol op de avond van de ramp slecht georganiseerd is (p. 397)
- De regering heeft onvoldoende aandacht gegeven aan al dan niet op waarheid berustende maatschappelijke onrust (p. 398).
- Het aantal malen dat de Tweede Kamer onduidelijk, onvolledig, ontijdig of onjuist is geïnformeerd, is naar de mening van de Commissie te groot, juist ook gezien de nasleep van de ramp (p. 398).

Heldere en duidelijke conclusies die waarschijnlijk veel fundering hebben in de feiten. Tegelijkertijd zijn het wel conclusies die oordelen en organisaties en soms personen afrekenen. Dat roept de vraag op hoe dergelijke conclusies zich nu verhouden tot het primaire doel van veel van deze evaluaties: namelijk er lering uit te trekken. Ongetwijfeld kunnen stevige boodschappen soms een groter effect sorteren dan omfloerste taal. Toch laat dat onverlet dat dergelijke rapporten die primair het oogmerk hebben om van te leren, soms meer verantwoordingsdossiers worden die (mede door de stevige taal) vooral gaan over de vraag of de verantwoordelijken nog wel kun-

2 Eindrapport van het onderzoek van de Onderzoeksraad voor Veiligheid naar de brand in het detentie- en uitzetcentrum Schiphol-Oost in de nacht van 26 op 27 oktober 2005, Projectnummer: M2005CH1026-1, gepubliceerd op 21 september 2006

3 Duikongeval Terneuzen, 12 Maart 2008, oktober 2009 (projectnummer M2008CH031201)

4 Enquête vlieg-ramp Bijlmermeer, Eindrapport van de Parlementaire Enquêtecommissie (Kamerstuk 26241, nr 9, 1999)

nen blijven. Het blijkt lastig om te evalueren zonder te (be/ver)oordelen.

Causale of contextuele betekenisgeving?

In een boeiende verhandeling onderscheidt Hans de Bruijn (2007) een causale en een contextuele benadering bij evaluaties. De OvV hanteert, bijvoorbeeld bij de rapportage van de cellenbrand, een causale benaderingswijze. De Raad constateert dat er een groot aantal fouten is gemaakt die niet gemaakt zouden zijn als de betrokkenen zich hadden gehouden aan de regels, voorschriften en dergelijke, en dus hadden gedaan wat ze moesten doen. Men construeert, zo zegt De Bruijn, een causaal verband tussen deze handelingen en de uiteindelijke gevolgen.

Contextuele betekenisgeving zoekt niet naar causaliteit. Ook hier kan wel geconstateerd worden dat regels niet zijn nagekomen, maar de vraag waarom dat het geval is wordt ook belicht. Waarom deden de betrokkenen zoals ze deden; daarmee wordt de complexiteit van de maatschappelijke werkelijkheid inzichtelijk gemaakt.

Een voorbeeld (De Bruijn, 2007, p. 11). De Raad is in de rapportage niet alleen zeer kritisch richting de ministeries van Justitie en VROM maar ook richting de gemeente. Ten aanzien van de gemeente is echter in een andere rapportage (de commissie-Hendrikx) geconstateerd dat er sprake was van informatie-discongruentie. De gemeente was niet voldoende in staat op te boksen tegen de rijksoverheid bij de controle en handhaving.

Een dergelijke analyse is meer contextueel en wordt vrijwel niet gehanteerd door de OvV. Het zal duidelijk zijn dat het nogal uitmaakt of een meer causale of meer 'verstehende' methode gehanteerd wordt. Thans domineert in de meeste onderzoeken de causale benadering.

Een uitzondering hierop vormt de recente evaluatie door de Politieacademie van de gebeurtenissen in Alphen aan den Rijn na het schietdrama in het winkelcentrum De Ridderhof (Van Duin e.a., 2012). Deze publicatie is feitelijk ook geen evaluatie maar een uitvoerige beschrijving en analyse van een aantal dilemma's, bijvoorbeeld over de relatie en spanning tussen het gemeentelijk beleidsteam en de driehoek (met de hoofdofficier van Justitie, de burgemeester en de korpschef). Hier wordt dichtbij de 'verstehende'/contextuele methode gekomen.

Enkele en dubbele lus leren?

De theorie van Argyris en Schon over organisationeel leren is een zinvolle mogelijkheid om onderscheid aan te brengen in het leren (zie ook Van Duin, 1992).

Enkele lus leren is de meest gangbare vorm van leren, hierbij wordt een bestaande discrepantie tussen feit en norm weggewerkt. Er doet zich een incident voor. De oorzaak van dat incident wordt geanalyseerd en het feilen wordt – waar mogelijk – aangepast.

Dubbele lus leren is verstrekkender. Hierbij wordt niet alleen gekeken naar in het oog springende factoren, maar wordt er ook achter deze factoren gekeken. Een voorbeeld. Na het zoveelste duikincident met een brandweerman kunnen opnieuw aanbevelingen worden gegeven om meer te oefenen, ook onder lastige omstandigheden en dergelijke. Daarnaast kan ook de vraag gesteld worden of de brandweer wel de geëigende organisatie is om die taak te verrichten. Daarmee wordt de taak als zodanig ter discussie gesteld. Veranderingen zijn meer structureel van aard waarbij ook de uitgangspunten zelf ter discussie worden gesteld. In de recente Leerarena naar aanleiding van Moerdijk is ook aan de hand van dit onderscheid gewerkt.

Product of proces?

Bij evaluaties denken wij natuurlijk eerst en vooral aan een product; een rapport dat wordt opgeleverd of openbaar wordt gemaakt. Toch kan het ook goed zijn dat niet per se het product – het rapport – centraal staat maar juist het evaluatieproces zoals dat plaatsvindt. Niet het product maar het evaluatieproces is de opbrengst. Het hoeft geen betoog dat in de nodige gevallen het proces belangrijker is dan het product (De Bruijn e.a., Procesmanagement). Ter vergelijking: niet het rampenplan, maar het nadenken over en het planningsproces als zodanig heeft een belangrijke functie. Op een vergelijkbare wijze kan ook het evaluatieproces als zodanig waarde hebben voor de evaluerende en mogelijk ook voor de te evalueren partij(en). Dat geldt uiteraard in situaties waarin er overlap is tussen beide. Zelf-evaluaties bijvoorbeeld vervullen potentieel zowel een waarde vanwege de opbrengst als vanwege het proces. Dat verklaart mede, naast allerlei wettelijke verplichtingen, dat in vele gevallen bij ongevallen, rampen en crises de betrokken partijen ook zelf (interne) evaluaties verrichten. Dit is ook een reden waarom het waardevol kan zijn bij een evaluatie betrokkenen te betrekken. Voor de doorwerking en het vergroten van het effect van een evaluatie is betrokkenheid – een procesvariabele! – van belang. Ook in situaties waarin deze rollen gescheiden zijn kan de interactie tussen de evaluator en de geëvalueerde waarde hebben voor deze laatste. Uit het gesprek kunnen onderwerpen of ideeën worden gehaald die weer verder kunnen helpen. Als de geëvalueerden een ‘slecht gevoel hebben’ van deze interactie neemt de kans toe dat de effecten van het product gering zullen zijn. Het uiteindelijke rapport wordt niet meer gewaardeerd.

Er zijn verschillende recente voorbeelden van deze nieuwe aanpak. Adang e.a. (zie Adang, 2011) werken met de methode van peer reviews: een beoordeling door gelijken (‘peers’) om bijvoorbeeld grootschalige eve-

nementen en risicovolle voetbalwedstrijden te beschouwen. Zowel de onderzoekers als de onderzochten leren van dat proces. Onder leiding van de vertegenwoordigers van de brancheorganisatie hebben brandweermensen zelf de belangrijkste observaties en lessen van de brand in Moerdijk in kaart gebracht. Zij komen met een lijst van eerste en tweede orde lessen (Leerarena Moerdijk).

Gebeurtenis of evaluatie?

Er bestaat de neiging lessen die betrokken worden uit calamiteiten en andere crises toe te schrijven aan de evaluaties die na afloop hebben plaatsgevonden. Toch is dat vaak maar ten dele het geval.

Allereerst zijn evaluaties niet altijd nodig om te weten wat de belangrijkste lessen van een gebeurtenis zijn. Feitelijk blijkt vaak dat binnen enkele dagen na een calamiteit al veel duidelijk is. Welke dominante fouten er zijn gemaakt? Welke zaken behoeven verbetering? etcetera.

Er waren – om het cru uit te drukken – echt geen duizenden pagina’s nodig om te weten dat vuurwerkopslag en zeker de verwerking van vuurwerk niet in de onmiddellijke nabijheid van de woonomgeving dienen te geschieden. Toch was dat wel de belangrijkste les van de vuurwerkkramp. Eveneens hoeft je geen ‘raketgeleerde’ te zijn om te weten dat aan elkaar geplakte containers niet bedoeld zijn voor bewoning (Cellenbrand). Ook weten de meesten wel dat in een overvol café dat helemaal vol hangt met zeer brandbaar materiaal een klein vlammetje snel enorme gevolgen kan hebben (Volendam).

Vervolgens blijkt uit veel onderzoek (Van Duin, 1992 en 2011, Birkland 2006) dat calamiteiten vaak ‘gelegen’ komen. Dikwijls liggen er al allerlei rapporten waarin voorstellen worden gedaan om bepaalde maatregelen te nemen. In de VS lagen zeer uitgewerkte plannen voor het verhogen van de veiligheid van

de luchtvaart. Pas na 11/9 kregen deze plannen voet aan de grond en werd in zeer korte tijd een heel pakket aan extra beveiligingsmaatregelen uitgerold (Birkland, 2006). De ideeën waren er wel, maar de kans om ze te realiseren moest nog komen. Als zich dan een calamiteit voordoet is dat de ideale trigger om het op dat moment openstaande raam ('policy window': Kingdon, 1984) te gebruiken; een vorm van opportunistische beleidsvorming.

Feiten of normen?

Bij evaluaties draait het in belangrijke mate om feiten. Het begrip feitenrelaas wordt vaak gebezigd als de term die gebruikt wordt om 'feitelijk' aan te geven wat er gebeurd is. Hoe kon het gebeuren?

Er wordt bij evaluaties veel werk gemaakt om de feiten in kaart te brengen. Een groot deel van deze feiten staat redelijk onomstotelijk vast. Beeldmateriaal, metingen en constateringen geven hier bijvoorbeeld de onderbouwing. Deels komen feiten uit een reconstructie die soms letterlijk geschiedt (terugkeer met Breivik naar het eiland Utoja) maar vaker zullen zij komen uit divers bronmateriaal en gesprekken die na afloop plaatsvinden.

Naast de feiten is er vaak ook aandacht voor normen. Bij evaluaties wordt vaak impliciet – maar de laatste jaren ook steeds vaker expliciet – een achterliggende normenkader gehanteerd. Om iets te kunnen zeggen over het gebeurde, en zeker om er in een meer (be)oordeelende vorm iets van te vinden, is een bepaald referentiekader waardevol. Hoe moeten wij de wijze van handelen van de brandweer beoordelen? Wat valt te zeggen over de kwaliteit van het optreden? Op verschillende manieren kan deze normering worden gemaakt. Veelvuldig blijkt de bestaande wet- en regelgeving, de leidraden en convenanten als basis te kunnen fungeren. Het gebeurde wordt dan afgezet tegen datgene wat was voorbereid, afgesproken en vastgelegd. Een andere manier zou kunnen zijn te bekijken hoe anderen dat in een der-

gelijke situatie zouden doen of hoe anderen zich hebben voorbereid. Oosting (Enschede) vergeleek hoe bij 43 andere vuurwerkopslagbedrijven het proces van de vergunning was verlopen. Ook werd Enschede vergeleken met vier andere steden en zelfs werd gekeken naar het buitenland.

Ten slotte kan een meer wetenschappelijke benadering worden gekozen. Wat weten wij uit literatuur en casuïstiek over andere (vergelijkbare) gevallen? Bij sommige evaluaties – bijvoorbeeld door de OvV – wordt vooraf een vrij uitgebreid normenkader ontwikkeld waartegen de werkelijkheid van de betreffende casus wordt afgezet. Hier zou ook plaats kunnen worden ingeruimd voor kennis uit wetenschappelijke literatuur en geanalyseerde casuïstiek.

Evalueren in andere landen

Het voert te ver om in deze notitie uitgebreid aandacht te besteden aan de wijze waarop in ons omringende landen en ook verder weg wordt omgegaan met het fenomeen dat in deze notitie centraal staat. Er is veel internationale literatuur over het lerend vermogen maar veel minder over de specifieke wijzen waarop in landen na ongevallen, rampen en andersoortige crises het evaluatieproces is ingericht. Deze paragraaf is dan ook zeker niet 'af', maar is meer een eerste vingeroefening op basis van enkele ervaringen en voorbeelden. Toch geven deze wel een eerste voorzichtige aanwijzing dat de wijze waarop in Nederland met dit evalueren wordt omgegaan bepaald niet alom gedeeld wordt. Veeleer lijkt in Nederland sprake van een tamelijk uitzonderlijke situatie.

Engeland en de VS

Ons overzeese buurland werd met name in de jaren tachtig regelmatig opgeschrikt door grotere rampen zoals de ramp met de veerboot de Herald of Free Enterprise, voetbalstadionrampen (Hillsborough en Bradford), metrobranden en de explosie op de Piper Alpha (boor-

platform). Van al deze rampen werden door (gewezen?) rechters en prominente wetenschappers gezaghebbende rapporten gemaakt die grote bekendheid kregen zowel daar als buiten Groot-Brittannië. Het Sheen-rapport (Herald), het Taylor-rapport (Hillsborough) en het Cullen-rapport (Piper Alpha) – alle genoemd naar de auteur – werden nadien veel geciteerd en gebruikt in allerlei wetenschappelijke artikelen en rapporten. Hoewel er vast wel discussies naar aanleiding van deze rapporten zullen zijn gevoerd, lijkt het er sterk op dat de uiteindelijke uitkomsten weinig ter discussie hebben gestaan en daarom ook zo bekend zijn geworden. De zin: ‘From top to bottom the body corporate (de rederij van de Herald, MvD) was infected with a disease of sloppiness’, werd bijvoorbeeld een van de meest geciteerde zinnen uit de rampliteratuur.

In de VS zien wij iets vergelijkbaars. Ook hier zijn er na grote crises als het neerstorten van de Challenger, 11/9 en Katrina gezaghebbende onderzoeksrapporten gemaakt. In dit geval zijn het commissies met uiteenlopende namen als Presidentiële commissie of Nationale commissie, die de rapporten opstelden, soms mede op basis van openbare verhoren.⁵ Ook deze rapporten zijn vaak gezaghebbend geworden en worden – net als in Groot-Brittannië – vaak gezien als *de* evaluaties van de betreffende gebeurtenissen.

Duitsland

Ervaringen die ik zelf recent opdeed bij een congres in Duitsland (waar ik sprak over de wijze waarop in Nederland evaluaties worden verricht) en bij de bestudering van met name de ramp van 1998 in Eschede, leerden het vol-

gende: in Duitsland is geen vanzelfsprekende evaluatiecultuur. Van veel ongevallen en calamiteiten is weinig tot niets terug te vinden in termen van evaluaties. Daar waar ze er zijn, zijn ze primair gericht op het detecteren van de oorzaak.

Ik probeerde deze bewerkingen te staven aan de hand van de casus Eschede. Juni 1998 deed zich een van de grootste rampen uit de Duitse naoorlogse periode voor, toen een ICE-trein ontspoorde ter hoogte van het Duitse plaatsje Eschede als gevolg van een gebroken wiel dat enkele kilometers eerder reeds was losgeschoten. De trein botste tegen een viaduct en uiteindelijk kwamen 101 personen om en raakten zo'n 88 personen zwaar gewond.

Uiteraard zijn er evaluaties naar aanleiding van deze ramp gestart. Het belangrijkste onderzoek ging over de oorzaak; hoe had het kunnen gebeuren waarbij vooral de aandacht werd gericht op het gebroken wiel. Dit onderzoek werd uitgevoerd door ‘Diplomingenieur Hans-Heinrich Grauf’, werkzaam bij het ‘Eisenbahn-Bundesamt’ (Wikipedia, Eschede, Duitse versie). Dit rapport vormde de basis voor jaren lopende strafrechtelijke procedures tegen een ingenieur, een medewerker en de directeur van Deutsche Bahn.

Daarnaast zijn er nauwelijks duidelijke evaluaties te traceren. Wel zijn er vele artikelen – vooral over de oorzakenanalyse – geschreven en zijn er verschillende boeken naar aanleiding van deze ramp gemaakt. De meeste boeken zijn door journalisten geschreven. Uitzondering hierop vormt het boek dat is uitgegeven bij Springer Verlag en getiteld is: *Die Katastrophe von Eschede: Erfahrungen und Lehren, Eine interdisziplinäre Analyse*. Dit boek bevat 34 hoofdstukken, overwegend geschreven door practitioners – die deels ook ter plaatse een (leidende) rol gespeeld – en behandelt in de respectievelijke delen de gehele organisatie van de hulpverlening ter plaatse, de organisatie daar en het gehele psychologische management na de ramp voor betrokkenen, hulpver-

⁵ Overigens gaan het benoemen van deze commissies niet altijd probleemloos. Na Katrina zou een onafhankelijke commissie (‘bipartisan committee’ van de House of Representatives) worden benoemd met vertegenwoordigers van zowel de republikeinen als de democraten maar deze laatste deden uiteindelijk niet mee.

leners en anderen. Al lezend wordt een goed beeld verkregen van de gehele hulpverlening. Opvallend is echter wel dat in het gehele boek geen verwijzingen zijn opgenomen naar rapporten. Het lijkt erop dat er verder ook weinig evaluaties zijn gemaakt, want via Google en in de verschillende artikelen waaronder het zeer uitgebreide Wikipedia-document, zijn geen verwijzingen gevonden naar andere, meer substantiële evaluaties.

Frankrijk

Een jaar na de vuurwerkramp in Enschede deed zich een enorme explosie voor in Toulouse die qua omvang nog groter was dan die in Enschede. Tot 80 kilometer van de plaats van de explosie(s), een fabriek van ASE, zo'n vijf kilometer ten zuiden van Toulouse, was de klap hoorbaar⁶. De explosie van grote hoeveelheden ammoniumnitraat bij deze chemische fabriek vond slechts tien dagen na 11 september 2001 plaats, wat de gebeurtenis vanaf het begin verdacht maakte. Hardnekkig waren geruchten over een aanslag van Al Quaida. De explosie kostte dertig mensen het leven; velen raakten gewond en de totale schade werd boven de miljard euro geschat. Na de brand ontstond een lange juridische strijd. Een commissie uit leden van de Assemblée Nationale onder leiding van Francois Loos stelde een onderzoek in dat vooral gebaseerd werd op grote aantallen openbare gesprekken met betrokkenen van het bedrijf, uit de overheid en anderen.

Het kan natuurlijk zijn dat ik er te dicht opzit, maar het heeft er alle schijn van dat er in het buitenland niet zo'n variëteit aan evaluaties na elke verstoring, ongeval en calamiteit bestaat als in ons land. In veel gevallen is er daar één dominant rapport geschreven en is daaromheen niet zo'n evaluatiecultuur (of afrekencultuur?) ontwikkeld als in Nederland. Alleen Engeland kent een duidelijke geschiedenis van rampevaluaties. Wel zien wij in het buitenland na calamiteiten vrij veel verslagen en evaluaties van de direct betrokken organisaties, gemeen-

ten, provincies/regio's en departementen verschijnen. Daarnaast worden er de laatste paar jaar meer evaluaties opgesteld, maar een flink deel daarvan (bijvoorbeeld in België) wordt niet openbaar.

Er is wat voor te zeggen om de rapporten van de commissies Oosting en Alders te zien als vergelijkbare exercities als de Engelse en Amerikaanse voorbeelden van gerenommeerde rapporten. De Nederlandse rapporten zijn uitgebreider en veelal meer integraal dan de buitenlandse voorbeelden. In die laatstgenoemde rapporten draait het vrijwel steeds om de vraag: hoe kon de ramp gebeuren. Aandacht voor rampenbestrijding is er vaak minder.

Beperkingen bij de evaluaties

Overdaad schaadt

Er wordt zo veelvuldig geëvalueerd dat er zeker sprake is van evaluatiemoedheid. Don Berghuijs, voormalig regionaal commandant regio Rotterdam-Rijnmond, verzuchtte na de zoveelste evaluatie: 'stop nu een paar jaar en geef ons daadwerkelijk de tijd en de menskracht om al datgene wat geconstateerd en aanbevolen wordt in de evaluaties, daadwerkelijk te implementeren'. Nu kunnen degenen die de lessen moeten trekken het vaak niet aan.

Overdaad is er niet alleen ten aanzien van de aantallen evaluaties – zoals eerder aangegeven zijn er meermalen van een en hetzelfde incident of ongeval vier of meer evaluaties gemaakt –, deze is er ook ten aanzien van de omvang van de rapporten (enkele duizenden pagina's na Enschede, maar ook meer dan vijfhonderd na een noodlottig duikincident) en ten aanzien van de aantallen conclusies en aanbevelingen die soms geformuleerd wor-

6 Uit onderzoek bleek dat er sprake was van twee grote explosies in een tijdspanne van zo'n 10 seconden, een feit dat de relatie met een aanslag nog vergrootte.

den. Een klein incident in een dierenwinkel in Hoogeveen, een casus die nauwelijks iemand zich zal herinneren, leverde een rapport op met 36 conclusies en een vergelijkbaar aantal aanbevelingen. De commissie Oosting kwam met 92 actiepunten; waarna enkele maanden later Alders er nog eens 55 naar aanleiding van Volendam aan toevoegde. Het COT kwam met 31 aanbevelingen na het drama in Hoek van Holland. Overdaad dus om verschillende redenen: te veel en te omvangrijke rapporten met te veel conclusies en aanbevelingen.

Onduidelijke of onuitvoerbare aanbevelingen

Voor velen is de kern van een evaluatierapport de lijst met aanbevelingen. Daarmee wordt aangegeven wat er geleerd kan worden en hoe betrokkenen een volgende keer kunnen voorkomen of beter beslagen ten ijs zullen komen als zich iets vergelijkbaars nogmaals zou voordoen. Als vele – vooral oudere rapporten – worden bestudeerd valt het op hoe lastig het is goede aanbevelingen op te schrijven en hoe vaak aanbevelingen gewoon het omgekeerde zijn van bepaalde conclusies. Er waren te weinig plannen; breng het aantal plannen op orde. Er werd onvoldoende informatie uitgewisseld; heb oog voor de informatie-uitwisseling.

Voorbeelden:

- De autoriteiten moeten inlichtingen van de kant van de kant van de Inlichtingen Diensten zorgvuldig beoordelen op mogelijke overschatting (COT, ordehandhaving extreem rechts, 1995, p. 68)
- De justitiële afhandeling bij ordehandhaving mag niet langer als een restpost worden beschouwd. De verantwoordelijkheden van de politie en justitie dienen met meer helderheid onderscheiden te worden (idem, p. 69).
- Het systematisch evalueren van daadwerkelijk optreden en van oefeningen dient verder gestimuleerd te worden (IOOV, Brand Harderwijk 1998, p. 31)
- Met name aan de wijze waarop met de overcapaciteit, die vooral in de middelgro-

te vrijwilligerskorpsen vaak op de plaats van het incident optreedt, wordt omgegaan dient (in verband met de span of control van de leiding) aandacht te worden besteed (idem, p. 32).

- *Alertheid.* Alle betrokken diensten dienen structureel alert te zijn op mogelijke gebeurtenissen voorafgaand, tijdens en na evenementen. Deze structurele alertheid komt tot gelding in het ontwikkelen van adequate scenario's, en voldoende, serieus te nemen oefening (COT, Hoek van Holland, 2009, p. 121).
- *Herkenbaarheid politiecommandanten.* Indien zij operationeel actief zijn, moeten politiecommandanten in uniform als zodanig herkenbaar zijn, tenzij dit risico's met zich mee brengt voor het welslagen van de operatie en/of de eigen veiligheid (idem, p. 121).

Op zich geen aanbevelingen waar iets mis mee is, maar vaak zegt het alleen nog zo weinig. De stap van aanbeveling naar een passend handelingsperspectief is nog groot. Vaak wordt een rapport afgerond met een serie aanbevelingen, en daarmee eindigt niet alleen het rapport maar ook de verantwoordelijkheid van de evaluator. De rapportage wordt over de schutting gegooid en daarna zoekt men het uit en gaat de evaluator naar een volgende klus. Het hoe, wat, wanneer en in welke volgorde dat gedaan zou kunnen worden – in termen van implementatie – blijft veelal onbesproken. Inspecties en de OvV hebben een taak ook nog eens nadien te kijken wat er nu gebeurd is. Ook kan de betrokken organisatie (bijvoorbeeld een gemeente of een ministerie) natuurlijk zelf dat proces monitoren. Toch is het aantal voorbeelden hiervan beperkt.

Uniciteit

Rampen en ongevallen en ook andere crises zijn buitengewoon zeldzame gebeurtenissen. Dat betekent dat negen van de tien keer betrokkenen voor het eerst met een derge-

lijke gebeurtenis worden geconfronteerd. In de lectorale rede (Van Duin, 2011, p. 33-47) is een hoofdstuk gewijd aan het thema 'normaal en bijzonder'. Moraal van dat betoog is onder andere weergegeven in onderstaande bevindingen:

- Ervaring is een van de belangrijkste voorwaarden voor succesvol handelen en optreden: dat geldt ook op het terrein van crisisbeheersing. Dagelijkse kwaliteit draagt bij aan goede crisisbeheersing.
- Naarmate de werkwijzen, plannen e.v. voor het bijzondere meer afwijken van die van het normale, neemt de kans toe dat dit ten tijde van crises tot problemen leidt. (p. 45).

Deze constatering leidt ertoe dat het al mooi is als bij ongevallen en crises betrokkenen een zetje weten te scoren. Men moet roeien met de riemen die men heeft, men doet het voor de eerste keer, sleutelpersonen vallen weg, de druk is extreem hoog en de ervaring is juist laag. Een aantal typerende kenmerken van ongevallen en crises waaronder de factor tijdsdruk (altijd te weinig tijd om informatie te vergaren om degelijke beslissingen te nemen), maken dat per definitie in dergelijke situaties suboptimaal gehandeld zal worden.

De verwachtingen zijn hoog, de ogen zijn gericht op alles wat er mis gaat en juist in die situatie moeten betrokkenen proberen er het beste van te maken. Voor betrokkenen zijn dit dus unieke gebeurtenissen. Bij de evaluaties zou daar (meer) rekening moeten worden gehouden, wat echter te weinig gebeurt.

Normering

Bij evaluaties wordt meestentijds een beoordelingskader geformuleerd. Het is goed aan te geven in een evaluatie welke basisuitgangspunten gehanteerd worden om tot een beoordeling te komen. Zonder een dergelijke normering ontbreekt de basis onder een evaluatie. Hoe komt men in een dergelijke situatie dan tot een waardering?

Er zijn verschillende manieren om tot normering te komen. Vaak zijn de aanwezige afspraken en datgene wat in regels, plannen, procedures en wetten is vastgelegd het uitgangspunt van het normenkader. Hoewel dat logisch lijkt, is hier een duidelijk nadeel bij, dat aansluit bij het thema van de uniciteit. Daarnaast passen deze evaluaties in de lijn van wat De Bruijn de causale methode noemt. Juist het unieke van de situatie en het meestal aanwezige gebrek aan ervaring spelen mee. Bekende rampsociologen als Quarantelli (1986) hebben uit en te na benadrukt dat crisismanagement absoluut iets anders is dan crisisplanning en dat dit vaak ten onrechte op één lijn wordt gezet. Papier is geduldig, de hectische werkelijkheid van een ramp niet. Juist de kenmerken van rampen en crises maken dat crisisbeheersing niet zo gemakkelijk is als op basis van de planvorming soms wordt aangenomen. De werkelijkheid is grillig, er raken veel en veel meer partijen betrokken dan waarop vooraf is gerekend, sleutelpersonen vallen weg, cruciale hulpbronnen vallen weg, de betekenis van niet-verwachte en informele hulp en hulpbronnen wordt veel groter (zie voor een uitgebreidere analyse, Van Duin, 2011).

Neutraal of gekleurd?

Er is op het eerste gezicht weinig reden om te veronderstellen dat evaluaties – zeker in Nederland – gekleurd zouden zijn. Daarnaast lijkt dit onderscheid te overlappen met de tweedeling onafhankelijk en niet-onafhankelijk onderzoek. Waarom zou een – zeker onafhankelijk – rapport gekleurd zijn? Toch kan ook de vraag gesteld worden, wanneer is een rapport neutraal? Bij een technische analyse van de oorzaak van bijvoorbeeld een explosie zal er weinig discussie zijn. Toch kan er altijd een discussie ontstaan over het belang van een bepaald feit (een handeling, het nalaten ervan en zo verder). Is sprake van neutraliteit in het rapport van de OvV over de cellenbrand wanneer er duidelijk drie schuldigen worden aangewezen: het ministerie van Justitie (exploit-

tant), het ministerie van VROM (eigenaar) en de gemeente (toezichthouder, verantwoordelijk voor de bestrijding van de brand)? Er kan immers ook worden beweerd dat de rol van de gemeente om verschillende redenen duidelijk ondergeschikt is (zeker als tegenpartij ten opzichte van machtige ministeries). Het rapport dat door de commissie Hendriks werd gemaakt zat veel meer op die lijn, maar was wel weer in opdracht van diezelfde gemeente gemaakt. Hoe neutraal zijn rapporten over de q-koorts, de varkenspest of de aviaire influenza waarin niet of weinig aandacht is voor de intensieve veeteelt, een factor die toch ergens wel een rol zal spelen bij de risico's van besmetting?

Kwaliteit van het onderzoek en de onderzoekers

Een laatste thema betreft de kwaliteit van het onderzoek en de onderzoekers. Hoe weten wij nu dat de kwaliteit van het onderzoek goed is? De evaluerende organisaties hanteren zelf wel allerlei normen als ijkpunt voor hun onderzoek, maar wie bepaalt nu of degenen die de onderzoeken verrichten capabel zijn om dat onderzoek te verrichten? Wordt hier ook een bepaalde norm als ijkpunt voor competentie gehanteerd? In hoeverre hebben de betrokken personen en hun leidinggevendende voldoende competenties om dergelijke evaluaties te verrichten? Als wij de recente discussie volgen over het IOOV-onderzoek naar de brand bij Moerdijk en ook de reacties naar aanleiding van deze evaluatie vanuit Midden- en West-Brabant zien (inclusief de gemaakte film), dan is er op zijn minst discussie over de kwaliteit van het betreffende rapport. Er is geen opleiding voor evaluator van rampen en crises. Er is ook geen begin van een maatlat waarlangs kan worden gemeten welke onderzoekers en onderzoeken nu goed zijn en welke niet.

Ten dele kan dit punt worden opgevangen door bij evaluaties gebruik te maken van specifieke externe deskundigheid. De Onderzoeksraad werkt standaard met een team van externe

deskundigen die gaandeweg meelesen en soms zelfs meeschrijven. Ook de Inspectie heeft trouwens met een dergelijke klankbordgroep gewerkt bij de evaluatie-Moerdijk.

Massamedia en evalueren

Niet alleen onderzoekers stellen evaluaties op. Meer en meer zijn de laatste jaren ook media aan het evalueren geslagen. Onderzoeksjournalistiek was tot recent voorbehouden aan de grotelandelijke (met name 'kwaliteits') kranten. Zo verschenen er uitvoerige reconstructies van kabinetscrises, het dreigende faillissement van de ABN, de bankencrisis enzovoorts. Inmiddels zien wij ook dat regionale kranten deze rol oppakken. Het Dagblad van het Noorden stelde bijvoorbeeld zowel naar aanleiding van de wateroverlast van begin 2012 als na Haren (2012) een reconstructie op. Over het water bleef dat beperkt tot één pagina in de krant; bij Haren werd een hele bijlage gevuld (3 november) met een hoogwaardige reconstructie waar dooreen team van journalisten aan was gewerkt.

Ook de radio en televisie zijn geschikte media voor dergelijke reportages. Zembla scoorde enorm met de documentaire over q-koorts. Niet alleen werd het uitgeroepen tot de beste documentaire van 2010; ook vrijwel onmiddellijk na deze uitzending werden door de overheid allerlei maatregelen afgekondigd. Een jaar later kwam de EO met een mooie documentaire over Kootwijkerbroek waarbij zij – tien jaar na dato – probeerden te reconstrueren waarom juist in dit dorp de protesten tegen de politie en de overheid zo fel waren bij ontruiming in het kader van de MKZ-uitbraak van 2001.

Het is toe te juichen dat ook de media meer aandacht besteden aan deze crises; niet alleen als alle ogen erop gericht zijn, maar ook nadien. Tegelijkertijd betekent dat dan bijvoorbeeld dat nadien journalisten steeds vaker met WOB-verzoeken komen om alle verslagen en stukken die ten tijde van de crisis zijn opge-

steld, op te vragen (verslagen van het RBT en ROT; geluidsbanden e.d.).

Afronding: kansen en mogelijkheden

In dit slot komen verschillende opties aan de orde die zouden kunnen bijdragen aan het verhogen van de kwaliteit, maar ook de effectiviteit van de evaluaties. Effectiviteit is hier opgevat als de mate van doorwerking van een rapport, waarbij doorwerking onder andere betekent dat conclusies en aanbevelingen worden overgenomen en gebruikt. Primair is dit een persoonlijke visie.

Minder incidenteel en meer structureel evalueren

Verantwoordelijken dienen zich de vraag te stellen of er wel zo veelvuldig geëvalueerd zou moeten worden. Nu worden er zoveel evaluaties opgesteld dat vele evaluaties waarschijnlijk alleen al door de overdaad in de lades beladen. Er zouden gewoon minder evaluaties moeten worden verricht, waarmee die evaluaties die geschieden mogelijk ook belangrijker worden. Tegelijkertijd zouden vaker meer trendmatige evaluatiestudies kunnen worden gehouden waar tegelijkertijd een aantal (gelijksoortige) gebeurtenissen zouden worden bekeken (trendanalyse, relevante rode draden, opvallende verschillen e.d.). Daarbij kan ook worden gekeken naar de effecten van de ene gebeurtenis op de andere (case contaminatie).

Voorbeeld:

Direct na de brand bij Moerdijk kwam er ophef over het totale negeren van sociale en nieuwe media. Er werd geen gebruik van gemaakt; informatie via internet kwam met uren vertraging. Enkele weken na Moerdijk was er een forse brand op Kijfhoek. Ongetwijfeld geleerd van Moerdijk werd gelijk een twitter-account aangemaakt en werden sociale media actief ingezet om de omwonenden te informeren. Weer enkele maanden later was in Alphen aan den Rijn onmiddellijk na het drama in het winkelcentrum De Ridderhof ook bij de overheidsorganisaties enorme aandacht voor alles wat via Twitter werd gecommuniceerd. Dergelijke (mogelijke) verbanden worden in afzonderlijke evaluaties niet gelegd.

Minder versnippering

Een kwetsbaarheid van hedendaagse evaluaties is ook de versnippering. Afzonderlijke evaluatoren maken los van elkaar rapporten waarin ze verschillende onderwerpen behandelen. Niet alleen is het omslachtig dat van een en dezelfde calamiteit vijf of soms nog meer rapporten verschijnen, ook kan dit extra negatieve gevolgen oproepen.

Verschillende rapporten over één gebeurtenis kunnen deels strijdig zijn. Daarnaast wordt het lastiger om in dergelijke gevallen voldoende afweging te kunnen maken van het (relatieve) belang van de gedane constatering, conclusies en aanbevelingen. Nu worden bij alle afzonderlijke onderdelen zoveel accenten gelegd en worden deze allemaal zo belangrijk gemaakt, dat een overall-benadering en relativering niet meer plaats kan hebben. Als, zoals bijvoorbeeld bij de evaluatie van Moerdijk door het IOOV het geval is, blijkt dat de brand bij aankomst niet meer te blussen was, en er ter plaatse nooit een dergelijke brand had mogen ontstaan (er mochten buiten – waar de brand begon en snel om zich heen greep – helemaal geen gevaarlijke en brandbare stoffen opgeslagen zijn, laat staan met open vuur gewerkt worden) plaatst dat alle kritische geluiden over de repressie in een ander daglicht. In een integraal rapport zouden conclusies over de repressie dan mogelijk minder prominent zijn geweest en ook niet die aandacht hebben gekregen.⁷

Ook kunnen verschillende evaluatoren (van dezelfde gebeurtenis) met elkaar een soort van concurrentiestrijd aangaan; ‘evaluatie-inflatie’. Men gaat elkaar overtreffen in kritiek over het gebeurde om daarmee zoveel mogelijk aandacht te mobiliseren voor de eigen bevindingen. Aangezien, zoals iedereen weet, positieve

⁷ Uiteraard kunnen anderen dit als een voordeel zien. Alles wat geleerd zou kunnen worden krijgt hiermee uitvoerig aandacht.

rapporten 'niet verkopen' richting media, is dat een extra stimulans om kritisch te zijn.

Afstand en betrokkenheid

Er is vaak een sterke roep om onafhankelijk onderzoek. Tegelijkertijd weten wij uit studies dat personen vaak meer leren van zaken die zij zelf ervaren of van dichtbij zien dan dat zij leren van 'derden' die zeggen hoe zij het beter moeten gaan doen. Er is veel voor te zeggen om personen en organisaties zelf te betrekken bij evaluaties. Al doende leren is een waardevolle en bekende leervariant die vaak beter werkt dan dat anderen jou zeggen wat je (anders) moet doen. Uiteraard staat betrokkenheid (bij je eigen evaluatie) op gespannen voet met voldoende afstand en bestaat er de vrees van beïnvloeding van de inhoud. Transparantie van het onderzoek- en evaluatieproces is in die situaties dus des te meer cruciaal.

Achterliggend is hier natuurlijk ook de vraag belangrijk welk doel een evaluatie heeft. Als het vooral past in het verantwoordingsproces is het logischer dat betrokkenen niet 'mee-evalueren' dan als uiteindelijk leren en verbeteren het primaire oogmerk zijn.

In Nederland zien wij recent op verschillende terreinen en bij verschillende organisaties vormen van zelfevaluatie vorm krijgen. Binnen Politie-Nederland is het instrument van de peer-review populair aan het worden. Zo zijn in opdracht van de raad van de korpschefs peer reviews gemaakt van diverse grote evenementen (Koninginnenach Den Haag, Vierdaagse en de feesten in Nijmegen, TT-Asen. Lowlands in Biddinghuizen/Dronten e.a.). In deze evaluaties wordt – onder begeleiding van Otto Adang, lector Orde en Gevaarsbeheersing van de Politieacademie – juist door leidinggevende politiefunctionarissen op het vlak van grootschalig optreden, bij anderen in de keuken gekeken.

Iets vergelijkbaars is binnen Brandweer-Nederland ontwikkeld onder de noemer Leerarena. Naar aanleiding van Moerdijk hebben

betrokkenen bij deze gebeurtenis gesproken en verteld over hun ervaringen aan collega-brandweerfunctionarissen met vergelijkbare rollen en functies. Wat ging goed en wat heeft aandacht nodig? Wat zijn directe en wat zijn meer indirecte lessen die uit deze gebeurtenis zijn te trekken.

Minder conclusies en aanbevelingen

Er bestaat vaak bij degenen die evalueren een wens om zoveel mogelijk van al datgene wat zij naar boven halen ook weer te geven. Hoe meer er vervolgens geconcludeerd wordt, hoe beter het is. Veel conclusies en aanbevelingen geven al snel de schijn van doorwrochtheid en dragen bij aan het vergroten van het belang van de evaluatie ('het is goed dat wij dit bekeken hebben, hier was zoveel te leren en wij hebben het allemaal boven water gekregen'). Het is echter de vraag of de doelstelling van de evaluator en de geëvalueerde hier wel parallel lopen.

In veel gevallen wordt een evaluatie onderwerp van een politiek debat en in veel gevallen ondersteunt de politiek de conclusies en aanbevelingen en zit de organisatie vervolgens met een fors aantal te nemen maatregelen in de maag. De combinatie van de kort na elkaar verschenen rapporten van de commissie-Oosting (Enschede) en de commissie-Alders leverde ruim 140 actiepunten op die allemaal opgepakt, gemonitord, etcetera moesten worden. Op zich waren dit grote en ingrijpende gebeurtenissen, maar ook veel kleinere incidenten en verstoringen leverden soms forse aantallen conclusies en aanbevelingen op. In het algemeen wordt er in de lijst van de aanbevelingen geen prioritering aangebracht waardoor de ontvangende organisaties een enorme druk voelen om de voorgestelde maatregelen allemaal snel op te pakken en te implementeren. Dat probleem wordt nog groter als er niet één rapport is dat aanbevelingen heeft en een organisatie opdraagt zaken aan te pakken, maar er twee, drie en soms meer dan tien van dergelijke lijsten met aanbevelingen zijn gemaakt.

Preventie	Preparatie en response
– onvoldoende voorschriften gevolgd	– verouderde of niet vastgestelde plannen
– wat strijdigheid in regels en voorschriften	– te weinig geoefend en getraind
– proces erg ingewikkeld geworden in de loop der jaren	– problemen met cruciale verbindingen
– onvoldoende controle op de naleving	– onvoldoende en late externe communicatie
– onderschatting risico's	– onderschatting van uitvoeringsproblemen
– etc.	– etc.

Toenemende sleetsheid en overlap

Een probleem bij de grote aantallen evaluatie-studies naar ongevallen en calamiteiten gedurende de laatste jaren betreft de toenemende sleetsheid van datgene wat gezegd en geconcludeerd wordt. Feitelijk kan je vrij blindelings aangeven wat er bij een volgende calamiteit geconstateerd zal worden. Met het bovenstaand lijstje kom je bij een volgende evaluatie waarschijnlijk al een eind.

Het komt niet zo vaak meer voor dat een evaluatie een volstrekte verrassing laat zien. Meestal bevat de tekst voor degenen die bekend zijn met dergelijke rapporten betrekkelijk weinig nieuws. Hooguit is er verrassing omdat de conclusies wat scherper zijn verwoord dan het rapport en de gebeurtenis liet aanzien. Uiteraard leidt deze sleetsheid er ook toe dat degenen voor wie de rapporten primair bedoeld zijn, vaak als eerste afhaken: 'weer zo'n rapport dat aangeeft wat we allemaal niet goed doen, en constateert dat wij ons veel steviger moeten voorbereiden en beter moeten communiceren'. Verschillende evaluatoren hebben zo ook hun eigen stokpaardjes. Het COT zoomt in op de communicatieprocessen, de OvV roept op tot een integraal veiligheidsplan en IOOV hamert erop dat men doet wat men vooraf heeft vastgelegd en afgesproken.

De toon en enige relativering

Ook de toon van het rapport en de conclusies en aanbevelingen kunnen bijdragen aan het omarmen of juist verwerpen van een evaluatierapport. Rapporten waarin nauwelijks iets

heel blijft van een bepaalde organisatie hebben vaak tot gevolg dat de betreffende organisatie het rapport niet of nauwelijks serieus zal gaan nemen. Uiteraard speelt het prestige van de evaluator daarbij een rol, maar ook van een zeer prestigieuze evaluator wordt niet alles aangenomen. Het komt in dergelijke gevallen ook wel voor dat de betrokkene(n) gelijk al afstand nemen van een aantal bevindingen. Stevigheid dient soms vergezeld te gaan van enige relativering; zaken kunnen wel verkeerd gegaan zijn, maar veelal is daar wel een plausibele reden voor aan te geven.

De kwaliteit van aanbevelingen

Naast de hoeveelheid aanbevelingen speelt ook de kwaliteit – in termen van bruik- en uitvoerbaarheid – van aanbevelingen een rol. Soms zijn aanbevelingen slechts de omdraaiing van een conclusie. De conclusie dat de interne communicatie beter moet, leidt dan tot de aanbeveling dat de betrokkenen oog moeten hebben voor de interne communicatie of dat moet worden toegezien op de verbetering van de onderlinge communicatie. Het formuleren van goede, implementeerbare aanbevelingen is een kunst op zich die ook lang niet altijd beheerst wordt door de evaluatoren. Vaak zijn onderzoekers uitstekend in staat de belangrijkste zaken die in een situatie gespeeld hebben in kaart te brengen en daarmee de situatie te duiden. Dat is echter een andere kunst dan het aandragen van implementeerbare adviezen. Voor dat laatste is andere kennis en expertise nodig dan voor het eerste. De onderzoeker is vaak beter in het opstellen van de (analytische)

rapportage, maar de adviseur of procesbegeleider in het opstellen – en bij voorkeur ook ondersteunen – van het invoeringsproces. Analyseren en adviseren zijn feitelijk twee verschillende soorten van expertise.

Implementatie meenemen

Hierboven is reeds aangegeven dat het formuleren van aanbevelingen moeilijk is en aanbevelingen veel meer zijn dan omgekeerde conclusies. Aanbevelingen zullen mogelijk meer gedragen worden – en daarmee ook eerder geïmplementeerd worden – als de geëvalueerde heeft mee kunnen denken/doen in het proces en is meegenomen in de gedachten en de wereld van de evaluator. Terwijl een forse afstand tussen observator en object (en dus degenen die onderwerp zijn van de evaluatie) vanuit onafhankelijkheid een groot goed wordt gevonden, is hier om andere redenen wel wat tegenin te brengen. Naarmate er in het evaluatieproces meer ruimte en aandacht zijn voor datgene wat de betrokken personen en organisaties in de toekomst zouden kunnen en moeten doen, neemt de kans toe dat er daadwerkelijk iets met de evaluatie en de aanbevelingen gedaan zal worden. Personen en organisaties veranderen niet omdat zij worden gedwongen, maar omdat zij er zelf van overtuigd zijn (geraakt) dat de veranderingen waardevol kunnen zijn.

Tijd

Uiteraard is het lastig onderzoekbaar maar het lijkt erop dat een lange tijdsspanne tussen gebeurtenis en rapportage niet bijdraagt aan de effecten die beoogd worden met de evaluatie. Vooral de OvV heeft veel tijd nodig om te komen tot haar rapporten. Gemiddeld is een jaar niet vreemd. Uiteraard liggen hier factoren als grondigheid, zorgvuldigheid en dergelijke

aan ten grondslag. Toch roept het de vraag op of deze factoren nu opwegen tegen andere factoren als effectiviteit en invloed. Eén van de (strategische) effecten van deze lange doorlooptijd is dat verantwoordelijke autoriteiten mede daarom zelf ook kortlopende onderzoeken entameren met de daarbij behorende risico's (waaronder interne strijdigheid tussen bepaalde zaken en een afname van integraliteit in de rapportages). Recent heeft de OvV hier een koerswijziging aangekondigd.

Ten slotte

Er is – alles overziende – daarom wel wat voor te zeggen om niet hele batterijen aan evaluaties op te starten na ieder incident of crisis, maar die ene evaluatie die wordt gemaakt dan wel zo integraal mogelijk te maken. Dit betekent dat in de rapportage aandacht is voor alle aspecten van de veiligheidsketen en dat er in de rapportage en conclusies een (af)weging komt in de zwaarte van de verschillende aspecten. Deze evaluatie zou kunnen voldoen aan de volgende eigenschappen:

- *snelheid*: binnen maximaal zo'n half jaar een rapport; desgewenst een interim-rapport als met name technische onderzoeken nog niet zijn afgerond;⁸
- *integraal*: door de gehele veiligheidsketen heen, aandacht voor alle fasen i.c. aandacht voor veiligheid en risicobeleid; preventie; oorzakenanalyse; voorbereidingsniveau, crisismanagement en organisatie van/in de nafase;
- *overtuigende feitenrelas (op alle thema's)*: het is goed denkbaar het hier bij te laten en vervolgens de rapportage terug te geven aan degenen die 'er iets mee moeten' (het veld, de politiek en andere betrokkenen);
- *afgewogen*: in de beoordeling – voorzover die wenselijk wordt geacht (zie ook vorige punt) – voldoende aandacht voor de afweging tussen de verschillende stadia en daarbij ook nuchter, realistisch en begripvol zijn;

8 Deze snelheid van rapporteren ontdoet de betrokken autoriteiten ook van een argument om zelfstandig onderzoek te entameren.

- *aandacht voor dilemma's*: uitdragen dat er voor sommige problemen die zich voordoen geen simpele antwoorden voorhanden zijn. Ieder calamiteit of crises kent van deze 'wicked problems';
- *aanbevelingen* (indien opgesteld!) zo dicht mogelijk bij of in overleg met betrokkenen opstellen.

Het is voor de hand liggend dat in de gevallen waarin de Onderzoeksraad evalueert, anderen dat niet doen of louter ten behoeve van en ondergeschikt aan de Onderzoeksraad. In de gevallen waarin de Onderzoeksraad niet evalueert, pakken Inspecties het gezamenlijk op (zoals de laatste jaren al vaker gebeurt).

Naast deze afzonderlijke evaluaties moet aandacht zijn voor meer trendmatige ontwikkelingen die zichtbaar zijn bij ongevallen, rampen en crises. Deze worden dan in afzonderlijke studies opgepakt. Voorbeelden op het onderhavige terrein zouden kunnen zijn:

- snelle ontwikkelingen op het terrein van de sociale media;
- invoering, werking en praktijk van netcentrisch werken;
- ontwikkeling en betekenis van het operationele niveau (operationeel team: tussen-niveau, cruciaal of overbodig?);
- verhouding(en) tussen bestuurders en operationele leiding;
- voorlichting en externe communicatie.

crisis-communicatie

Overheid, crisis en nieuwe media

door Ellen Kamer

Het is de dag dat de Nederlandse overheid NL-alert introduceert, een nieuw aanvullend alarmeringssysteem dat de helft van de bevolking in de directe omgeving van een noodsituatie via een bericht op de mobiele telefoon kan informeren over wat er aan de hand is en wat men kan doen. Het is ook de dag dat inwoners van Nijmegen en omgeving rond 7 uur wakker schrikken van een enorme knal. Een knal die "doet denken aan een aantal straaljagers die tegelijkertijd door de geluidsbarrière denderen", gevolgd door een aanhoudend nabulderen dat te horen is tot in het nabijgelegen Groesbeek, Zevenaar, Berg en Dal en Bemmelen.

Wat volgt verdient geen schoonheidsprijs. Terwijl Twitteraars in heel Nederland binnen 5 minuten weten dat de herrie te maken heeft met een explosie in de elektriciteitscentrale van Electrabel, krijgen 3200 burgers na een half uur via burgernet het sms-advies ramen en deuren gesloten te houden en af te stemmen op Omroep Gelderland. Deze omroep wisselt, bij gebrek aan informatie van het bevoegd gezag, haar berichtgeving rond het incident af met muziek en andere nieuwsitems. Wanneer het NOS-journaal meldt dat Omroep Gelderland live-beelden heeft, volgt een overbelasting van de website die vervolgens anderhalf uur uit de lucht is. De politie meldt op Twitter, ruim drie kwartier na de explosie, dat ze richting Nijmegen vliegen in verband met een explosie in