

Keuze-architectuur of individuele verantwoordelijkheid? Een filosofische bijdrage

Maureen Sie

1. Inleiding

Het is alsof de enorme berg aan wetenschappelijke informatie over de menselijke cognitie onlangs over een dood punt is geraakt en zich nu als een lawine over de mensen uitstort. “Komt dat zien: wij zijn niet rationeel, autonoom, en moreel. Wij beschikken niet over een vrije wil. Ons brein bepaalt ons doen en laten, wij hebben slechts het nazien!” Het ontstaan en de impact van deze lawine is begrijpelijk. Het wetenschappelijk onderzoek naar het menselijk functioneren en de invloeden daarop omvat al vele decennia en strekt zich uit over diverse wetenschapsgebieden. En hoewel zich onder de cognitie- en neurowetenschappers, gedragseconomen en psychologen langzaam overeenstemming begint af te tekenen over de basale uitgangspunten voor het begrijpen van menselijk gedrag (ik gebruik ‘gedrag’ en ‘handelen’ ook om naar keuzes en oordelen te verwijzen), blijft de common sense opvatting achter: rationaliteit, autonomie en moraliteit blijven centraal staan in ons zelfbeeld terwijl het onderzoek nu juist de beperkingen daarvan aan het licht brengt. Althans zo lijkt het.

In deze bijdrage presenteer en bespreek ik bovengenoemde ontwikkelingen en de vraag wat er uit volgt voor diegenen die zich bezig houden met het bestuur van bedrijven en organisaties en het toezicht daarop. De vraag die daarbij centraal staat is of het verantwoord is staande bedrijfs- en toezichts praktijken te continueren waarin ons vermogen tot zelfreflectie en communicatie en het idee van individuele verantwoordelijkheid centraal staat. Op basis van het wetenschappelijk onderzoek worden immers vraagtekens gezet bij de autonomie, zelfkennis en vrije wil van het *individu*. Zoals ik zal laten zien zijn er wetenschappers die menen dat we het menselijk functioneren vooral moeten begrijpen in termen van automatische en onbewuste reacties op onze omgeving. Een dergelijk begrip staat op gespannen voet met een praktijk waarin wordt uitgegaan van ons individuele vermogen tot zelfreflectie en communicatie en individuele verantwoordelijkheid. Je zou je daarmee kunnen afvragen of gangbare bedrijfs- en toezichtspraktijken niet geherstructureerd moeten worden op basis van wetenschappelijke expertise, bijv. met behulp van wat men ook wel ‘keuze-architectuur’ noemt. Of dat bedrijven, bijvoorbeeld, verplicht zouden moeten worden om wetenschappers te raadplegen. Ik zal argumenteren dat geen van beide het geval is.

In de eerste paragraaf breng ik wat orde aan in het vele en uiteenlopende materiaal over de invloeden en mechanismen die betrokken zijn bij ons alledaagse functioneren. In de tweede paragraaf bespreek ik hoe dit onderzoek het idee van onszelf als rationele wezens lijkt te ondermijnen en hoe het idee zou kunnen postvatten dat wij gebruik moeten maken van keuze-architectuur. In de derde paragraaf bespreek ik waarom het onderzoek volgens sommigen het idee van onszelf als autonome wezens lijkt te ondergraven. Hier bespreek ik de visie dat bijvoorbeeld het expliciteren van gedragsregels en codes niet zinvol is en dat de waarde van communicatie en zelfreflectie wordt overschat. In de vierde paragraaf leg ik uit dat die conclusie voorbarig is. Hier argumenteer ik bovendien dat het toepassen van de wetenschappelijke kennis om bedrijfsprocessen en procedures van toezicht te verbeteren waarschijnlijk minder eenvoudig is dan sommige populair wetenschappelijke literatuur suggereert.

2. Achtergrond: de onderzoeksresultaten

Sinds een aantal decennia zien we in de gedrags-, cognitie- en neurowetenschappen een paradigma opkomen dat naast het ‘adaptief onbewuste,’ ook wel het ‘nieuwe onbewuste’ genoemd wordt of, in eigen land, het ‘slimme onbewuste’.¹ In tegenstelling tot het Freudiaanse onbewuste toont dit adaptief onbewuste zich niet in zeldzame versprekingen en verstoringen van ons bestaan, maar juist in onze alledaagse vermogen adequaat te functioneren. Omdat wij adaptief onbewuste wezens zijn, zijn we in staat allerlei informatie te verwerken zonder dat daar ‘bewuste controle’ en ‘reflectief bewustzijn’ aan te pas komt. Vandaar de naam: het *adaptief* onbewuste.²

Wat we bij ons spreken over het adaptief onbewuste voor ogen moeten houden is dat de concepten ‘onbewust’ en ‘automatisch’ die ermee verbonden zijn verdere uitwerking en bestudering vereisen, net zoals de daaraan nauw verwante concepten ‘bewust’, ‘bewustzijn’ en ‘bewuste controle’ én de daar weer aan gekoppelde maatschappelijke begrippen zoals ‘vrije wil’ en ‘morele verantwoordelijkheid’. De betekenis van die begrippen is onduidelijk. Laat me dit kort uitleggen.

Wanneer een voetballer een goal scoort heeft hij niet eerst nagedacht wat daar voor nodig is. Een goede voetballer kan zonder nadenken doeltreffend schieten: Een goede voetballer weet waar de andere spelers zich bevinden, waar de goal is. Toch zou het vreemd zijn van zo’n voetballer te beweren dat hij ‘slechts’ onbewust scoorde en, bijvoorbeeld, dat hij er niet persoonlijk verantwoordelijk voor is. Behalve dat de voetballer jaren van ervaring en training achter de rug heeft en daarmee allerlei

1 Zie bijvoorbeeld: Wilson, Timothy D. (2002): *Strangers to ourselves. Discovering the adaptive unconscious*. Cambridge, MA: Belknap Press of Harvard University Press. Hassin, Ran R.; James S. Uleman; John A. Bargh (eds.) (2005): *The new unconscious*. Oxford; New York: Oxford University Press. en Dijksterhuis, A. (2007): *Het slimme onbewuste*. Amsterdam: Bert Bakker.

2 Men vindt in de literatuur ook wel verwijzingen naar twee verschillende systemen die betrokken zijn bij ons functioneren; één sneller, automatisch, onbewust, grotendeels op basis van emotionele processen functionerend en één trager, bewust en grotendeels op rationele processen functionerend. Chaiken, S.; Y. Trope (2002): *Dual-process theories in social psychology*. New York: Guilford.

reacties geautomatiseerd heeft, is zijn doelbewustheid, balvaardigheid en spelinzicht nu precies zijn talent. Ons zogenaamde automatisch en onbewust functioneren kan niet zondermeer en simpelweg gecontrasteerd worden met bewuste aansturing, bewustzijn en bewuste controle. Ik kom daar nog op terug. Laten we om te beginnen eerst een paar voorbeelden van het experimentele onderzoek bekijken.

Wat te denken van het feit dat we in een bepaalde situatie voor woordenboek A 15 euro over hebben en voor woordenboek B 25 euro, maar in een iets andere situatie voor datzelfde woordenboek A 25 euro en woordenboek B, 15 euro? Wat we over hebben voor identieke producten hangt af van de context waarin ze ons aangeboden worden. Of we een woordenboek met gescheurde kaft en 20.000 woorden en eentje zonder gescheurde kaft met 10.000 woorden, afzonderlijk beoordelen of in vergelijking met elkaar heeft zo'n effect.³ Dit voorbeeld laat ons wellicht onaangedaan, en enigszins terecht zoals ik in de volgende paragraaf zal uitleggen. Immers, is het niet begrijpelijk dat ons oordeel over een product anders is wanneer we vergelijkingsmateriaal hebben? Maar wat te denken van de volgende voorbeelden?

Wanneer we met zijn allen naar één en hetzelfde filmpje kijken van twee auto's die op elkaar af rijden dan schat de ene helft de snelheid van die auto hoger in dan de andere helft, uitsluitend op grond van het feit dat hen gevraagd werd hoe snel de auto reed op het moment dat hij 'botste', in plaats van de vraag hoe snel de auto reed wanneer deze de andere 'raakte'.⁴ En ook wanneer we moeten kiezen tussen twee identieke maatregelen dan wordt wat we kiezen bepaald door het woordgebruik! Sterk gesimplificeerd: wanneer een arts ons de keus voorlegt een medicijn te gebruiken dat in '35% van de gevallen *effectief*' is of ons mededeelt dat het '65% van de gevallen *ineffectief*' is dan zal dat gevolgen hebben voor het aantal mensen dat besluit het medicijn te gaan gebruiken, ook al is de gecommuniceerde werking van het medicijn hetzelfde. Tot slot nog een voorbeeldje uit de morele sfeer. Wanneer mensen in een laboratorium setting een casus over gedrag moeten beoordelen, is hun morele afkeuring sterker wanneer zij bijvoorbeeld achter een tafel zitten die smerig en onopgeruimd is dan wanneer deze schoon en opgeruimd is.⁵

Binnen het paradigma van het adaptief onbewuste kun je dit soort experimenten als volgt duiden. Woorden en hun geassocieerde betekenis, de achtergrond waartegen we naar een bepaalde optie kijken, en vele ander factoren die zich aan ons bewuste zicht onttrekken, beïnvloeden onze keuzes en oordelen. Dit leidt vaak tot hilarische uitkomsten omdat de experimentele set-up zo is dat onze keuzes voorspelbaar zijn én schijnbaar irrationeel. Bij medicijngebruik zouden we ons moeten laten leiden

3 Hsee, C. K. 'The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives.' *Organizational behavior and Human Decision processes* 67, 1996, p. 247-257.

4 Dit voorbeeld is oorspronkelijk afkomstig uit Aldert Vrij's *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). *Wiley series in the psychology of crime, policing and law*. Vrij, Aldert New York, NY, US: John Wiley & Sons Ltd., 2008.

5 Dit cluster van onderzoek is afkomstig van Jonathan Haidt en collega's. Zie bijvoorbeeld, Schnall, Simone; Jennifer Benton; Sophie Harvey (2008a): 'With a Clean Conscience. Cleanliness Reduces the Severity of Moral Judgments.' In: *Psychological Science* 19. p. 1219-1222.

door de *relevante* informatie, niet door de associaties die de gebruikte woorden bij ons losmaken; en ook als we de snelheid van een auto moeten inschatten doen de gebruikte woorden waarin ons de vraag wordt gesteld er niet toe. De omstandigheden waarin we een oordeel vellen over iemands gedrag of een bepaald product zouden er alleen toe moeten doen wanneer ze dit gedrag of product betreffen. Betekent dit nu dat we de gedachte dat we rationele en morele wezens zijn moeten laten varen? En wat betekent het om deze gedachte te laten varen? Moeten bestuurders zich bewust worden van onze beperkte rationele en morele vermogens en daar bij het bestuur van hun bedrijf en organisatie meer rekening mee gaan houden, bijvoorbeeld door het intensieve gebruik van zogenaamde keuzearchitectuur? Laat me in de volgende paragraaf daar nader op ingaan.

3. Suboptimale resultaten en het adaptief onbewuste

In de vorige paragraaf hebben we een paar voorbeeldjes gezien van ‘onbewuste’ invloeden op ons gedrag die experimenteel onderzoek aan het licht bracht. Het is echter vooral de hoeveelheid en diversiteit van dat soort onderzoek dat suggereert dat de wetenschappers er inmiddels in geslaagd zijn de mechanismen te ontrafelen achter deze invloeden. Het is immers omdat, en in zoverre, ze weten hoe ons keuze- of oordeelsvermogen werkt dat zij een experimentele set-up kunnen ontwerpen die onze menselijke ratio in zekere zin laat ontsporen. ‘Ontsporen’ in de zin dat de uitkomst van ons gedrag in deze experimenten niet correspondeert met dat wat we als de beste, meest rationele of moreel verantwoorde uitkomst zien.

Er is veel wetenschappelijke en filosofische discussie over hoe we de mechanismes die dit ‘ontsporen’ veroorzaken precies moeten begrijpen en wat voor een mensbeeld dat impliceert. Maar wat moeten we er praktisch gezien van leren? In ieder geval *niet* dat we doorgaans irrationeel functioneren. Het onderzoek toont eerder aan op basis van welke ‘onbewuste’ mechanismes we normaliter adequaat functioneren.

Laat me beginnen bij onze gevoeligheid voor gebruikte woorden en termen. Mensen zijn door en door talige wezen, wij communiceren door de keuze van woorden. Het maakt een groot verschil of ik mijn man bij thuiskomst laat weten dat ik ‘een andere auto geraakt heb’ of dat ik ‘betrokken ben in een botsing’. In het eerste geval zal hij als reactie informeren naar de schade aan de auto, in het tweede zal hij eerst bezorgd vragen of ik gewond ben. Onze keuze voor een beschrijving in termen van ‘raken’ of ‘botsen’ communiceert de ernst van de situatie, maar dat wil niet zeggen dat wij die keuze bewust maken. Wij *denken* niet eerst *na* over welke woorden wij zullen gebruiken om een boodschap te communiceren, we doen dat automatisch en onbewust. Het experimentele onderzoek brengt naar de oppervlakte dat we ons niet voldoende bewust zijn van deze gevoeligheid voor taal en dat deze, wellicht mede daarom, ons óók beïnvloedt wanneer dat beter niet het geval zou zijn. Namelijk wanneer ons gevraagd wordt een oordeel te vellen over de snelheid van een auto die we

op een filmpje te zien krijgen. (Ik kom hieronder terug op de vraag waarom dat dan tot de inschatting van een *hogere* snelheid leidt.)

Evenmin is het irrationeel dat ons oordeel beïnvloed wordt door de context waarin een optie ons aangeboden wordt, zoals in het geval van de woordenboeken. Het is immers geen eenvoudige zaak vergelijkende oordelen te vellen over complexe relaties, in dit geval over kwaliteit, de verhouding prijs-kwaliteit, de verhouding nut/functie-prijs-kwaliteit. En daarmee is het niet irrationeel voor deze oordelen gebruik te maken van contextuele informatie, zoals in het geval van het woordenboek. Net zoals in het geval van ons woordgebruik geldt ook hier dat er vele dagelijkse contexten te bedenken zijn waarin we onbewust gebruik maken van informatie die voorhanden is. Sterker nog, waarin we gebruik *moeten* maken van zulke informatie om adequaat te kunnen functioneren.

Ook wat betreft het voorbeeld van het medicijngebruik lijkt een uitleg mogelijk op basis van het inmiddels de revue gepasseerde: woordgebruik en contexten informeren ons doorgaans effectief en stellen ons daarmee in staat oordelen te vellen en keuzes te maken over complexe zaken zonder dat daarvoor heel veel energie en tijd nodig is. Dit kan in sommige gevallen tot een suboptimale uitkomst leiden, maar dat toont niet aan dat we als zodanig irrationeel zijn of functioneren.

Tot slot dan nog het voorbeeld uit het morele domein. Ook op dit gebied lijkt de wetenschap erin geslaagd enkele mechanismes te identificeren die aan ons gedrag ten grondslag lijken te liggen én die in sommige gevallen kunnen ‘ontsporen’. Aan het cluster van onderzoek dat achter het besproken voorbeeld van de smerige tafel ligt, ligt de hypothese ten grondslag dat aan onze morele oordelen niet gebaseerd zijn op denkprocessen maar op *gut-reactions*, directe sterke gevoelens met emotionele basis zoals bijvoorbeeld walging. En die *gut-reactions* kunnen gemakkelijk gemanipuleerd worden (bijvoorbeeld via smerige bureaus). Bovendien suggereert ander onderzoek dat wanneer we een uitkomst als moreel slecht evalueren, dit onze oordelen over een feitelijke gang van zaken beïnvloedt. Om een voorbeeld te geven: een reeks van handelingen die tot de dood van een crimineel leidt wordt door ons als minder opzettelijk beoordeeld dan exact dezelfde reeks handelingen die tot de dood van een onschuldige omstander leidt.⁶ Men noemt dit ‘uitkomst-vooordeel’ en het zou kunnen verklaren waarom wij de snelheid van de auto op een filmpje in het geval van het gebruik van het woord ‘botsing’ hoger inschatten dan in het geval van ‘elkaar rakende auto’s’. Immers een ‘botsing’ evalueren wij als een moreel slechtere uitkomst dan twee elkaar rakende auto’s; naarmate de snelheid van de auto’s hoger is neemt de mogelijke verwijtbaarheid van de chauffeur in onze ogen wellicht toe.

Binnen het paradigma van het adaptief onbewuste zijn ook deze onbewuste invloeden in het morele domein te plaatsen. Moraal is immers van belang voor het

⁶ Zie Thomas Nadelhoffer, ‘Bad Acts, Blameworthy Agents, and Intentional Actions. Some Problems for Juror Impartiality’, in: Knobe, J. and Nichols, S. (eds), *Experimental Philosophy*, Oxford University Press 2008, p. 149-167.

reguleren van gedrag en één van de zaken die het hedendaagse onderzoek aan het licht brengt is dat emoties in de tot stand koming van gedrag een cruciale rol spelen.⁷ Emotionele lichamelijke reflexen stellen ons in staat in een *split second* afwijzend of goedkeurend te reageren. Onze onderbuik-gevoelens van walging zijn een goed voorbeeld van zulke reflexen: als wij iets zien dat wij smerig vinden dan reageert ons lijf onmiddellijk, zelfs als we dat misschien liever niet zouden willen. Wanneer wij adaptief onbewuste wezens zijn die grotendeels automatisch en onbewust reageren op wat ons overkomt, dan is het niet verwonderlijk ons moreel oordeelsvermogen van een dergelijk mechanisme ‘gebruik maakt’, ermee verknoopt is. Doordat wij ons onvoldoende bewust zijn van deze verknoping (in tegenstelling tot de wetenschappers), kunnen wij in situaties gebracht worden waarin onze *gut-reactions* ons misleiden, dat wil zeggen een rol spelen terwijl ze dat niet zouden moeten doen.

Inmiddels wordt er door filosofen en moreel psychologen uitgebreid gediscussieerd over de exacte rol en morele status van onderbuik-gevoelens.⁸ Hoewel het aannemelijk gemaakt is dat we onze morele oordelen, net zoals onze zogenaamde rationele oordelen, niet blind moeten vertrouwen, is de conclusie dat onze alledaagse morele oordelen volstrekt onbetrouwbaar zijn evenmin geoorloofd. Waarom is dit soort onderzoek dan toch van belang voor bestuurders en toezichthouders?

Wanneer we weten waarom we in bepaalde gevallen suboptimale resultaten krijgen, kunnen we onze omgeving zo inrichten dat we die verbeteren en de opties die we voorgeleggen zo formuleren dat wij een zo goed mogelijk oordeel kunnen vellen. Dit noemt men ook wel keuze-architectuur: Leg de gezonde appels op de juiste plaats in de kantine, daar waar een maximaal aantal mensen tot aankoop verleid zal worden; maak datgene tot de *default* optie waarvan je weet dat het tot het gewenste of minst ongewenste resultaat zal leiden etc. De ongezonde snoeprepen mogen wel blijven liggen, maar dan op een plek die het de mensen wat lastiger maakt daarvoor te kiezen.⁹ Daarnaast kan men door betere voorlichting en de verspreiding van wetenschappelijk kennis over onze cognitieve blinde vlekken en onze zogenaamde cognitieve vooroordelen, ons zelfinzicht vergroten en daarmee ons functioneren optimaliseren.¹⁰

Keuzearchitectuur en kennis van onze cognitieve vooroordelen kunnen gebruikt worden om bestuur, bedrijfsprocessen, toezicht en processen van toezicht te structureren. Naar alle waarschijnlijkheid gebeurt dat al. En waarom zouden we niet ons voordeel doen met de toegenomen kennis over omstandigheden die tot suboptimale resultaten leiden? (zie voor enkele kanttekeningen daarbij, de laatste paragraaf van

7 Zie bijvoorbeeld Antonio Damasio *Descartes' error: emotion, reason, and the human brain*, New York G.P. Putnam, 1994. Experimentele bevindingen in de zogenaamde neuro-ethiek tonen aan dat onze emoties zelfs ook altijd betrokken zijn bij het oordelen over morele dilemma's, zie bijvoorbeeld Greene, J. D., R. B. Sommerville, Nystrom, et al. 'An fMRI investigation of emotional engagement in moral judgment.' *Science* 293(14), 2001, p. 2105-2108.

8 Zie bijvoorbeeld Kelly, Daniel, *Yuck! The Nature and Moral Significance of Disgust*, MIT, Cambridge, 2011.

9 Zie de bestseller van Richard Thaler en Cas Sunstein, beiden belangrijke adviseurs in de politieke arena, *Nudge. Improving Decisions About Health, Wealth, and Happiness*, 2008. New Haven, CT: Yale University Press.

10 Zie de bestseller van Daniel Kahneman, *Thinking, Fast and Slow*, Farar Straus and Giroux, 2011.

dit stuk) Onze bespreking van de onderzoeksresultaten in termen van het adaptief onbewuste maakt echter duidelijk dat als bestuurders en toezichhouders gangbare processen *niet* herzien op basis van de nieuwste wetenschappelijke inzichten, zij niet in gebreke blijven. Het onderzoek toont immers niet aan dat wij irrationeel en immoreel functioneren. Er zijn echter ook andere geluiden te vernemen. Volgens sommigen vormt besproken onderzoek aanleiding tot zeer radicale veranderingen: een aanpassing van onze praktijk waarin bewuste aansturing op basis van reflectie niet langer centraal staat. Maar veronderstelt de bestuurspraktijk niet juist dat het expliciteren van gedragsregels en bedrijfscodes zinvol is, dat bestuurders verantwoordelijkheid kunnen dragen voor hun bestuur en dat het zinvol is als zij daarover verantwoording afleggen? Veronderstelt het niet de haalbaarheid en zinvolheid van toezicht onder meer op naleving van geformuleerde regels en codes?

In de volgende paragraaf zal ik laten zien dat deze twijfel aan onze staande praktijken te maken heeft met een ander aspect van het onderzoek, namelijk het feit dat wanneer men ons vraagt waarom wij handelen zoals wij doen wij vrijwel altijd exclusief antwoorden in termen van waar de, spreekwoordelijke, goal staat en waar onze tegenstanders zich bevinden, terwijl wetenschap hele andere invloeden aan het licht brengt.

4. Autonomie: rationalisatie en confabulatie

Neurowetenschappers claimen al decennia lang dat zij op basis van zichtbare breinactiviteit bewuste besluiten of keuzes kunnen voorspellen en de gedachte ligt voor de hand dat onze bewuste rationele overwegingen er (dus) helemaal niet toe doen.¹¹ Recentelijk zijn daar aanwijzingen bijgekomen dat informatie die uit brein-scans is af te leiden zelfs meer voorspellende waarde hebben dan overwegingen of voorkeuren waarvan wij ons bewust zijn. In ons brein is de activiteit van bepaalde gebieden zichtbaar te maken die voor de ‘verwachte beloning’ en ‘kosten van bepaald gedrag’ staan, bijvoorbeeld het gebruik van zonnebrand-crème, of de aankoop van bepaalde producten. Op basis van hoe de activiteit in elk van deze gebieden zich tot elkaar verhoudt, kan men (aldus de wetenschappers) het gedrag van de gescande personen beter voorspellen dan op basis van zelf-rapportage én beter dan de personen dat zelf kunnen. Dat zijn natuurlijk fascinerende bevindingen die suggereren dat wij onszelf slecht kennen en die goed aansluiten bij de reeds besproken experimenten.

Het is de combinatie van deze verschillende clusters van onderzoek die al menig mediagenieke uitspraak en titel inspireerde: Wij zijn ons brein, ons brein beslist

11 Libet, Benjamin et al., ‘Time of conscious intention to act in relation to onset of cerebral activities (readiness-potentials)’. The unconscious initiation of a freely voluntary act. In: *Brain* 106, 1983, p. 623-642. Meer recent Soon, Chun Siong, et al. ‘Unconscious determinants of free decisions in the human brain,’ in: *Nat Neurosci* 11 (5), 2008, p. 543-545.

(niet ik), de vrije wil is een illusie, bewuste aansturing bestaat niet.¹² Een mogelijke reconstructie van de gedachtegang achter deze boude claims is: breinscans tonen aan wat wij zullen doen voordat wij ons daarvan bewust zijn (of zelfs: waar wij ons van bewust zijn speelt nauwelijks een rol in wat we gaan doen); er bestaan allerlei invloeden op ons handelen die aan onze aandacht ontsnappen; wanneer wij redenen geven voor ons gedrag noemen wij daarbij dikwijls/soms niet de oorzaken die volgens wetenschappers doorslaggevend zijn. Ergo, er is geen/veel minder/nauwelijks invloed van rationele overwegingen en ons bewustzijn op ons gedrag dan wij doorgaans denken.

Een aantal wetenschappers gaat zó ver dat zij de redenen die we opgeven voor ons gedrag als ‘louter achteraf gebabbel’ afdoen. Het uitwisselen van redenen, zo kenmerkend voor het publieke domein, de politieke arena, en de dagelijkse menselijke omgang, heeft volgens hen louter een sociale functie: we proberen elkaar te overtuigen van de zinvolheid en legitimiteit van ons handelen. Maar daarbij gaan we te werk als advocaten gericht op het vrijpleiten van onszelf, niet als rechters gericht op de feiten.¹³ We hebben geen idee van de feiten, de werkelijke oorzaken van ons gedrag, maar fantaseren er lustig op los. Dat laatste verschijnsel heet in de wetenschappelijke discussie ook wel ‘confabulatie,’ een term in zwang geraakt door de wetenschapper Michael Gazzaniga.¹⁴

Hoewel de wetenschappelijke inzichten inderdaad aan licht brengen dat wij vaak pas redenen bedenken nadat wij een besluit namen of een oordeel velden én ook dat veel invloeden die op onze keuzes en oordelen werken aan onze aandacht ontsnappen, is het onzinnig om te concluderen dat bewuste aansturing niet bestaat of dat wij geen idee hebben waarom wij handelen en op basis van welke redenen. Het uitwisselen van en nadenken over onze redenen is een uiterst effectieve praktijk en het zou deze effectiviteit ontberen als er geen min of meer betrouwbare relatie zou bestaan tussen onze (uitgewisselde) redenen enerzijds en ons handelen anderzijds. Laat me dit uitleggen.

Dat wat mensen zeggen, denken en geloven over hun eigen gedrag stelt ons over het algemeen aardig in staat op elkaars, en ons eigen toekomstige gedrag, te anticiperen en het desgewenst te beïnvloeden. Daarom heeft het ook zin om iemand te vragen op tijd te zijn omdat de vergadering belangrijk is, erop te vertrouwen dat je je geld krijgt als iemand je dat beloofd heeft, en taart in te slaan wanneer je mensen voor je verjaardag hebt uitgenodigd. We denken voortdurend na over wat wij en anderen doen en wisselen redenen uit die ons informatie verschaft over wat we van elkaar

12 In eigen land: Dick Swaab, *Wij zijn ons brein. Van baarmoeder tot Alzheimer*, Contact, 2010 en Victor Lamme, *De Vrije Wil Bestaat Niet, Over wie er echt de baas is in het brein*, Bert Bakker, 2010. In het buitenland voorafgegaan door ondermeer: Wegner, D.M. *The Illusion of Conscious Will*, Cambridge, MA: MIT Press, 2002.

13 Jonathan Haidt ontwikkelde een *sociaal intuitionistisch model* van moreel oordelen waarin deze functie van het uitwisselen van redenen centraal staat. Zie bijvoorbeeld zijn: Haidt, J., ‘The emotional dog and its rational tail: A social intuitionist approach to moral judgment’, *Psychological Review*, 108, 2001, p. 814–34.

14 Gazzaniga, Michael S.; Joseph E. LeDoux: *The integrated mind*, New York: Plenum, 1978.

kunnen verwachten en wat niet. Dat werkt alleen wanneer we er regelmatig in slagen dat te doen wat we beloofden, denken te moeten doen, planden, ons voornamen, reden hebben te doen; wanneer we met andere woorden in staat zijn tot bewuste aansturing. Natuurlijk gaat dat ook wel eens mis, maar dat is alleen begrijpelijk omdat het meestal goed gaat.

Redenen zijn wat dat betreft de munteenheid van het handelen. Geld wordt vervalst, gestolen, gevraagd voor zaken die het niet waard zijn, is aan de lokale economie onderhevig, en gevoelig voor in- en deflatie; het zou echter niet langer van hand tot hand gaan of misbruikt worden als het niets waard was. Zo is het ook met redenen. We vergissen ons, bedenken soms achteraf redenen, zuigen ze af en toe (of regelmatig) uit onze duim, sommige redenen verliezen met de tijd hun geldigheid etc. maar ze zouden niet langer uitgewisseld worden als ze nergens op sloegen. Hoe moeten we deze observatie nu verenigen met de gedachte onszelf, blijkens besproken onderzoek, zo gebrekkig kennen? Wil ik dan toch beweren dat al dat wetenschappelijke onderzoek niet echt zo interessant is? Geenszins.

De gebrekkige zelfkennis die door het onderzoek aan het licht gebracht wordt is kennis over wat je de details en de mechanismen van het handelen zou kunnen noemen. In de praktijk reageren we op elkaar en de omgeving zonder precies te weten hoe en waarom, automatisch en onbewust. Daarmee blijft er (veel) ruimte voor vergissingen over wat we doen en waarom, vergissingen die niet eenvoudigweg gecorrigeerd kunnen worden door zogenaamde 'introspectie': de blik naar binnen keren om te zien welke gedachtes, verlangens en overwegingen aan ons handelen ten grondslag lagen. Het zelfbeeld dat door hedendaags wetenschappelijk onderzoek op de schop gaat is dat wij 'transparante' en 'deliberatief controlerende' wezens zijn. Dat wil zeggen, wezens die de blik slechts naar binnen hoeven te keren om te weten welke redenen er aan ons gedrag ten grondslag liggen (transparant) en die op basis van reflectie over wat wij moeten doen handelen (deliberatief controlerend). Dat betekent dat wij feilbaar en onvolledig zijn met betrekking tot het antwoord op de vraag wat wij deden en waarom precies, niet dat wij nooit weten waarom we doen wat we doen.

Laat me kort resumeren. Veel invloeden op ons gedrag ontsnappen aan onze aandacht. Binnen het paradigma van het adaptief onbewuste is dat de keerzijde van ons vermogen adequaat op informatie te reageren zonder deze bewust te verwerken en te controleren. Dat tast onze autonomie' aan wanneer dat gedefinieerd is als een feilloos inzicht hebben in onze eigen beweegredenen en ons handelen bewust aan sturen op basis van reflectie voorafgaand aan dat handelen. Besproken onderzoek maakt duidelijk dat wij dit ontberen. Dat wil echter niet zeggen wij geleefd worden door de invloeden die op ons werken, dat deliberatie geen invloed heeft en dat het uitwisselen van redenen een onzinnige praktijk is. Wanneer we het onderzoek in die mate op de spits drijven missen we de leerzame boodschap ervan: enige terughoudendheid ten aanzien van ons zelfinzicht en de redenen die we op basis van dat inzicht met elkaar uitwisselen. Wat impliceert die constatering nu voor de praktijk? Laat me daar tot slot op ingaan.

5. Algemene gevolgen voor de praktijk

Het tot dusver besproken onderzoek rechtvaardigt geen twijfel aan het bestaan van bewuste aansturing of de zinvolheid van het uitwisselen en nadenken over onze redenen. Het onderzoek toont echter wel dat wij onszelf slecht kennen én vooral introspectie (de blik naar binnen richten) geen bron van onfeilbare informatie is. Men zou kunnen menen dat dit voldoende grond biedt om de praktijk te herzien waarin zelfreflectie, de articulatie van motieven en redenen voor gedrag, communicatie en individuele verantwoordelijkheid centraal staan. We zouden bijvoorbeeld kunnen ophouden met elkaar individueel aan te spreken op gedrag, en daarvoor moreel verantwoordelijk te houden. In plaats daarvan zouden we de aandacht kunnen verschuiven naar het collectief en de omgeving. Dat wil zeggen meer aandacht voor de rol van de bedrijfscultuur en processen en de inrichting van de keuze-omgeving door middel van voornoemde keuze-architectuur. Bijvoorbeeld door wanneer mensen voortdurend te laat komen of zich vaak ziek melden allereerst te kijken hoe we de processen zo kunnen herinrichten dat dit verzuim kan worden rechtgetrokken. Hoewel voor deze verschuiving van aandacht wat te zeggen is, zal ik er in deze paragraaf enkele kanttekeningen bij plaatsen. De reden daarvoor is dat het huidige onderzoek, mijns inziens, vooral aanleiding is elkaar vaker, eerder en meer op individuele verantwoordelijkheid aan te spreken. Daarnaast maakt het onderzoek duidelijk dat wij terughoudend moeten zijn met keuze-architectuur omdat het ons leert dat onze praktijk ondoorzichtig en complex is en radicaal ingrijpen daarom lastig. Laat me drie zaken onderscheiden en bespreken: individuele verantwoordelijkheid, noodzaak en gevaren van keuze-architectuur en de complexiteit van de praktijk.

Laat me beginnen met het eerste aspect. Wanneer het paradigma van het adaptief onbewuste juist is doen wij veel op de automatische piloot en op basis van overwegingen die aan onze aandacht ontsnappen. Dit automatische en 'onbewuste' wordt doorbroken wanneer we elkaar aanspreken op wat we doen. Dat aanspreken nodigt uit tot zelfreflectie én tot het uitspreken van hoe je je verhoudt tot de overwegingen die je aanhaalt om het automatische gedrag uit te leggen of te rechtvaardigen. Stel je voor dat een aantal mensen altijd te laat is voor bepaalde vergaderingen, wat gebeurt er dan wanneer je hen vraagt om uitleg? Je maakt daarmee, ten eerste, de personen in kwestie bewust van hun gedrag. Ten tweede, geef je hen een kans uit te leggen waarom zij zich zo gedragen. Ten derde creëer je de mogelijkheid om met elkaar over die genoemde redenen van gedachten te wisselen. In alle drie de gevallen gaat het bij het uitwisselen van redenen om de positionering in de zogenaamde normatieve ruimte.

Dat wil zeggen, met wat wij zeggen geven we aan waarop men ons mag aanspreken, wat wij belangrijk vinden, en welke overwegingen wij van belang achten. Zo maakt het verschil of iemand laat weten vergadering 'niet zo belangrijk te vinden', 'het te druk te hebben', of 'er helemaal geen erg in gehad te hebben dat zij te laat was'. Deze positionering heeft gevolgen voor ons toekomstige omgang met elkaar en geeft daarmee mede gestalte aan de mores die onze omgang reguleert. Als iemand beweert er 'geen erg in

gehad te hebben dat zij altijd te laat was' nadat zij daar kritisch op is aangesproken, zou het heel vreemd zijn wanneer zij nog steeds voortdurend te laat zou komen.

Waarom is dit proces van belang in verband met het onderwerp van dit artikel? Het maakt duidelijk dat wanneer wij redenen noemen voor ons gedrag wij niet slechts, en wellicht zelfs niet primair, bezig zijn met het verslag doen van de overwegingen die feitelijk aan ons gedrag ten grondslag lagen. Overwegingen die door middel van introspectie achterhaald kunnen worden. Als ik zeg 'ik wist niet dat die vergadering zo belangrijk was' dan rapporteer ik niet per sé de gedachte die voorafging aan mijn te laat komen, ik geef eerder de overweging aan die maken dat het op tijd komen voor mij geen prioriteit hadden én ik vraag in één adem of ik me daarin vergist heb. Als de ander die vergadering aanmerkt als wél belangrijk zal ik in het vervolg zorgen dat ik op tijd ben. Het zou ook kunnen dat jij me laat weten dat die vergadering weliswaar niet belangrijk is, maar dat dat geen reden is om te laat te komen. In dat geval zal mijn reactie op die mededeling bepalend zijn voor mijn toekomstig 'punctueel' gedrag. Met het uitwisselen van redenen is, met andere woorden, ook het tot stand komen van toekomstige omgangsvormen gemoeid.

Laat me in aansluiting hierop nog even terug komen op het in de tweede paragraaf aangestipte onderzoek dat aantoonde dat emoties van onmisbaar belang zijn in de regulering van gedrag. Als dat het geval is dan is deze *bottum-up* wijze van regulering, het elkaar voortdurend aanspreken op individuele verantwoordelijkheid, van groot belang. Neem gedragsregels en protocollen, of bepaalde idealen die aan een bedrijf ten grondslag liggen. Deze zullen pas effectief worden wanneer ze (ook) in de emotioneel geladen dagelijkse omgang terecht komen. Een omgang waarin we, bijvoorbeeld met verontwaardiging, reageren op de schending van de normen, waarden gemaakte afspraken of ge-expliciteerde *commitments*. Juist wanneer we voor een groot deel onbewust en automatisch functioneren zijn praktijken waarin we individueel worden aangesproken op ons gedrag -en er zodoende toe gedwongen worden ons er bewust toe te verhouden - van groot belang.

Het zou dan ook heel onfortuinlijk zijn wanneer men elkaar niet langer moreel aan zou spreken vanwege besproken experimentele bevindingen. De functie die het uitwisselen van redenen heeft valt of staat niet bij de mogelijkheid dat wij ons kunnen vergissen over onze redenen. Juist doordat wij op ons gedrag aangesproken worden, kan duidelijk worden wat wij doen en hoe anderen dat gedrag zien en waarderen; door met elkaar over dat gedrag te spreken achterhalen we hoe wij ons tot verschillende overwegingen en gedraging verhouden en welke overwegingen die omgang in de toekomst zullen reguleren.

Laat me nu overgaan tot de bespreking van het tweede punt, de noodzaak en gevaar van keuze-architectuur. Een van de eerste zaken die we ons moeten realiseren is dat veel wetenschappelijke inzichten hun weg naar onze discussies over wat we moeten doen en laten al lang hebben gevonden. Zo is het heel normaal iemand aan te raden geen ingewikkelde besluiten te nemen voordat ze er een nachtje over geslapen

heeft. In het verlengde daarvan zal ook kennis over het bestaan en de werking van cognitieve vooroordelen onze normatieve discussies, discussies over wat het juiste is om te doen of te laten, beïnvloeden. Het zou bijvoorbeeld kunnen dat wij er achter komen dat sommige cognitieve vooroordelen zo hardnekkig, onvermijdelijk én verstorend zijn dat wij hun effectiviteit willen voorkomen door bepaalde maatregelen te treffen (denk daarbij bijvoorbeeld aan zogenaamde impliciete vooroordelen tegen vrouwen en allochtonen, vooroordelen die ook wanneer we ze niet onderschrijven aantoonbaar ons gedrag beïnvloeden).

Het gebruik van keuze-architectuur én de keuze daartoe over te gaan is echter niet eenvoudig. Immers wat het paradigma van het adaptief onbewuste duidelijk maakt is dat informatieverwerking op allerlei niveaus plaatsvindt. Ons gedrag wordt vormgegeven in en door de praktijk: door de redenen die wij expliciet met elkaar uitwisselen, door hoe onze omgeving is ingericht en hoe onze interacties met de wereld en elkaar verlopen. Elke ingreep in de omgeving is daarmee een ingreep in wat we de ‘normatieve ruimte’ zouden kunnen noemen, de ruimte die ons informeert over wat we moeten doen en laten. Zo zou het kunnen dat wanneer je auto’s veiliger, geruislozer, robuuster maakt dit erin resulteert dat mensen minder veilig gaan rijden. Niet omdat ze denken ‘mij kan niets meer gebeuren’ maar omdat de ervaring van een geluidloze, snelle, stevige en veilige auto hen het gevoel geeft dat zij veilig rijden. Als iets dergelijks het geval zou zijn, zullen campagnes met boodschappen over verkeers*on*veiligheid bij deze mensen niet aankomen, ze zouden zich niet aangesproken voelen.

Juist omdat er ook op onbewust niveau allerlei informatie gecommuniceerd wordt die een rol speelt in de regulering van ons gedrag, is het zo moeilijk te overzien wat maatregelen en veranderingen teweeg zullen brengen. En dit geldt ook voor de redenen die we expliciet communiceren, bijvoorbeeld met behulp van opgestelde reglementen, gedragscodes, regels en normen. Als je in een natuurpark bordjes ophangt met daarop ‘ten strengste verboden planten te stelen’ kan het zijn dat het percentage gestolen planten omhoog schiet omdat deze bordjes communiceren dat de planten zeldzaam en gewild zijn en dat je wel een hele brave borst moet zijn om er geen bij je te steken.

Dit brengt mij ten slotte bij het laatste punt. Bij het ingrijpen in en beoordelen van de praktijk is het van belang grip te krijgen op alle informatie die gecommuniceerd wordt, ook op de informatie die niet expliciet geformuleerd is maar wel invloed heeft op de sociale praktijken en fysieke omgeving. Daarnaast is het noodzakelijk oog te blijven houden voor het belang van een praktijk van het uitwisselen van redenen opdat er op natuurlijke wijze mores kunnen ontstaan die de mensen bindt vanwege hun persoonlijke positionering in de normatieve ruimte. Wetenschappelijke inzichten kunnen en zullen in beiden gevallen een bijdrage kunnen leveren, mits niet op de spits gedreven.

6. Conclusie

In deze bijdrage heb ik het wetenschappelijk onderzoek naar (moreel) handelen gepresenteerd en besproken welke gevolgen de bevindingen hebben voor onze alledaagse praktijk. Ik heb het idee van adaptief onbewuste gebruikt om de enorme hoeveelheid experimenteel onderzoek te ordenen en begrijpelijk te maken. Ik heb geargumenteed dat hoewel het onderzoek de grenzen van onze autonomie aan het licht brengt, allesomvattende conclusies over vrije wil en bewuste aansturing wetenschappelijk onverantwoord en filosofisch ongenueanceerd zijn. Bescheidenheid en voorzichtigheid bij het ingrijpen in en beoordelen van de praktijk is op zijn plaats. Bestuurders en toezichthouders doen er goed aan zich te informeren over de wetenschappelijke inzichten in ons functioneren en daar, zo mogelijk, hun voordeel mee doen. Het is uit de lucht gegrepen om te denken dat het op basis van deze inzichten onverantwoordelijk zou zijn om geen gebruik te maken van wetenschappelijk onderzoek of staande praktijken te continueren waarin het vermogen tot zelfreflectie, communicatie, het expliciteren en implementeren van normen en waarden centraal staan. Sterker, deze suggestie doet meer kwaad dan goed. Het onderzoek lijkt immers juist aan te tonen is dat de praktijk veel wijsheid bergt die bewuste beschouwing, reflectie, en wetenschap ontgaat.