

JAN VERPLAETSE<sup>a</sup>

# *De controversiële erfenis van Benjamin Libet*

## *Vrije wilsceptici en hun critici over het belang van Libet-type experimenten in het debat over vrije wil en verantwoordelijkheid*



*Panopticon*, 35 (2), 106-119  
© 2014 MAKLU | ISSN 0771-1409 | MAART 2014

- <sup>a</sup> Hoofddocent, Vakgroep Grondslagen en Geschiedenis van het Recht, Faculteit Rechtsgeleerdheid, Universiteit Gent (Corresp.: jan.verplaetse@ugent.be).

### ABSTRACT

**The controversial legacy of Benjamin LIBET. Free will skeptics and their critics on the importance of LIBET-type experiments in the free will and responsibility debate.**

To most legal scholars free will is an indispensable philosophical assumption underlying our penal law system. Since the 1980s LIBET-type experiments (LT) aim to criticize this assumption by showing that unconscious brain processes instead of conscious and deliberative thoughts actually initiate our decisions. Here, after summarizing the main findings, we review three branches of criticism: (a) methodological considerations, (b) objections made by philosophy of science, and (c) metaphysical remarks relating to moral and legal consequences. At the end of this critical review we discuss the possible impact of these neuropsychological investigations on the foundations of penal law.

**Keywords:** free will, LIBET-type experiments, responsibility

**Kernwoorden:** vrije wil, LIBET-type experimenten, verantwoordelijkheid

### 1. INLEIDING

Vermoedelijk delen veel Belgische strafrechtsspecialisten de volgende mening van hun Nederlandse collega Anton MOOIJ: 'de strafrechtpraktijk (...) kan niet functioneren zonder het uitgangspunt van de vrije wil als ordeningsprincipe te aanvaarden' (MOOIJ, 2012, 37). De vrije wil blijft één der fundamenten van ons huidige strafrecht. Wat de vrije wil en het vrije wildebat precies is, is minder duidelijk. Er zijn verschillende definities en debatten. De kans is niettemin groot dat MOOIJ iets als volgt bedoelt. Het strafrecht vooronderstelt dat een normaal, volwassen persoon *keuzevrijheid* heeft om een delict wel of niet te plegen. Hij of zij kan het delict vermijden. Ik noem deze invulling 'vrije wil als keuzevrijheid'.

Van oudsher staat die invulling op gespannen voet met de filosofische doctrine van het *determinisme*. Determinisme houdt in dat elke gebeurtenis het gevolg is van voorafgaande gebeurtenissen waaruit die gebeurtenis onherroepelijk volgt. Voor een determinist is er maar één toekomst. Keuzevrijheid is er niet. Komt wetenschap vroeg

of laat tot de slotsom dat menselijke beslissingen ook deterministisch werken dan valt het doek over 'vrije wil als keuzevrijheid'. Momenteel is het determinisme niet meer dan een – overigens onwaarschijnlijke – hypothese. Wie in 'vrije wil als keuzevrijheid' gelooft, hoeft niet te vrezen dat deze doctrine snel werkelijkheid wordt.

Determinisme is niet de enige bedreiging voor 'vrije wil als keuzevrijheid'. Mocht blijken dat al onze bewuste beslissingen veroorzaakt zijn door voorafgaande en onbewust verlopende breinprocessen waarop we geen controle kunnen uitoefenen, dan is er van 'vrije wil als keuzevrijheid' opnieuw weinig sprake. Wat hebben we aan keuzevrijheid als onze beslissing al vooraf is genomen en we die niet ongedaan kunnen maken omdat ze buiten ons bewustzijn tot stand is gekomen? Die bedreiging is ernstiger dan die van het determinisme. Ze is niet gebaseerd op abstracte filosofische theorieën, maar wordt onderbouwd door neuropsychologische experimenten. Deze experimenten gaan terug op de originele studies van Benjamin LIBET (1916-2007) en noemt men LIBET-type experimenten (LT). Deze LT-experimenten zouden aantonen dat al onze beslissingen al genomen zijn vooraleer we ons van deze beslissing bewust worden en dat onze keuzevrijheid bijgevolg een illusie is (SPENCE, 1996, 83-84).

Wie dacht dat 30 jaar experimenteel onderzoek voor een wetenschappelijk consensus zorgde, vergist zich. De LT-experimenten blijven controversieel. Gezien het enorme maatschappelijke belang is dit niet verwonderlijk. De vrije wil is een gebied waar wetenschap zich vermengt met ideologie en filosofie. Voorstanders willen niet dat de resultaten van de LT-experimenten waar zijn omdat ze 'social nihilism' vrezen (KLEMM, 2010). Vrije wilsceptici verwelkomen deze bevindingen als argumenten in hun pleidooi voor een 'verwijtloze samenleving' en een 'schuldloos strafrecht' (PEREBOOM, 2001; WALLER, 1991; HARRIS, 2012). In deze bijdrage geef ik eerst een overzicht van de oorspronkelijke LT-experimenten en hun latere variaties. Vervolgens ga ik in op de kritieken en geef ik enkele replieken. Ten slotte bespreek ik de diepere betekenis van dit hele debat op de fundamenten van het strafrecht. Vragen de bevindingen van de LT-experimenten om een nieuwe strafrechtfilosofie zonder schuld en verantwoordelijkheid? Ik geef straks antwoord.

## 2. DE LIBET-TYPE EXPERIMENTEN

De experimenten van LIBET vind je terug in drie papers waarvan de *Brain*-paper de beroemdste is (LIBET *et al.*, 1982; LIBET *et al.*, 1983; LIBET *et al.*, 1983). LIBET wist reeds dat ons brein bij een actieve en een passieve reactie een ander hersengolfpatroon laat zien. Bij een actieve reactie vraag je proefpersonen om bijvoorbeeld hun rechterhand te buigen op een vooraf meegedeeld tijdstip. Komt een roterende klokwijzer op 10 uur, sluit dan je hand. Bij passieve reacties reageren proefpersonen op een onverwachte prikkel. Buig je hand na een prik op de rug van je rechterhand. Terwijl een actieve reactie aandachtige voorbereiding vergt, wacht je bij een passieve reactie tot de prik komt. Een hersenfysioloog registreert dit alles. De handbuiging meet hij met een elektromyograaf of EMG die de spierspanning meet. Het moment waarop je de hand van de proefpersoon prikt, leest hij af van een klok. Het hersengolfpatroon ten slotte meet hij met een elektro-encefalograaf (EEG) waarbij je elektroden bevestigt op een gebied vlak voor de motorische cortex en dit op beide hersenhelften. Bij de passieve reactie vertoont de EEG een vlak patroon. Er is niets bijzonders te zien. Tijdens de actieve reactie zie je een stijgend patroon dat circa één seconde voor het buigen van de hand begint. Dat stijgend patroon werd ontdekt door KORNHUBER en DEECKE (1964)

die dit de *Bereitschaftspotential* of *readiness potential* (gereedheidspotential of twel RP) noemden.

LIBET ging een stap verder en liet zijn proefpersonen het bewegingsmoment zelf kiezen. Geen vooraf meegedeeld tijdstip meer, maar je kiest zelf wanneer je je hand buigt. Onder 'zelf kiezen' verstond LIBET twee dingen. Hij vroeg zijn proefpersonen om toe te geven aan de plotse neiging om je hand te buigen. Die neiging borrelt als het ware uit het niets op en je geeft je eraan over. Die beslissingen noemde LIBET 'spontaan'. Dit lukte niet altijd. Vaak waren de beslissingen 'gepland'. Dit was het geval wanneer ze vooraf het tijdstip van bewegen hadden gekozen of wanneer ze de beslissing om te bewegen uitstelden naar een later ogenblik. Na elke trial vroeg LIBET de proefpersonen of ze een spontane dan wel geplande beslissing hadden genomen. Weinig verrassend vond LIBET dat geplande beslissingen een RP laten zien die vroeger begint (700 ms voor de beweging), vergeleken met de RP van de spontane beslissingen (tussen 400 ms en 700 ms voor de beweging). Ten slotte vroeg LIBET aan zijn proefpersonen om mee te delen wanneer ze zich voor het eerst bewust waren van hun wil om te bewegen. Hij vroeg ze naar hun eerste ervaring van een wil of intentie om te bewegen, of anders gezegd, naar het tijdstip van het zogenaamde W-moment. Voor de tijdsregistratie gebruikte LIBET een snel ronddraaiende klok, een *cathode ray oscilloscope* (CRO) die binnen 2,54 seconden een volledige toer cirkelde en precieze timing toeliet.

LIBET vond dat bij spontane beslissingen de RP al op 350 milliseconden voor het W-moment begint. Bij geplande beslissingen is dat nog vroeger, namelijk 800 milliseconden voor het W-moment. Vooraleer je bewust bent van wat je wilt, is je RP al actief. Dit resultaat vat men kernachtig samen met  $(RP)t_1 < (W)t_2 < (M)t_3$ , waarbij  $t_1 < t_2 < t_3$ , oftewel gereedheidspotential komt voor het bewuste wilsmoment W dat op zijn beurt voor de motorische reactie M komt. Zelf schreef hij:

*'It is concluded that cerebral initiation of a spontaneous, freely voluntary act can begin unconsciously, that is, before that there is (at least recallable) subjective awareness that a 'decision' to act has already been initiated cerebrally. This introduces some constraints on the potentiality for conscious initiation and control of voluntary acts'* (1983, 623).

Uit dit citaat blijkt meteen hoe scherp LIBET zich bewust was van de filosofische implicaties van zijn vondst. Sindsdien hebben onderzoekers allerlei replicaties (KELLER & HECKHAUSEN, 1990; HAGGARD & EIMER, 1999, TREVENA & MILLER, 1999; MILLER, SHEPHERDSON & TREVANA, 2011) en variaties van dit basisexperiment bedacht en uitprobeerde. Zo vroeg men aan proefpersonen om de linker- of rechterwijsvinger te bewegen (HAGGARD & EIMER, 1999), links of rechts af te drukken (SOON *et al.*, 2008; HERRMANN *et al.*, 2008), getallen op te tellen of af te trekken (SOON *et al.*, 2013). Voor de tijdsregistratie van het W-moment gebruikt men behalve klokwijzers ook voorbijglijdende letters (SOON *et al.*, 2008; SOON *et al.*, 2013). Participanten onthouden bij welke letter ze de eerste aandrang voelden om te beslissen. Naast de oorspronkelijke bilaterale EEG-schedelregistratie zijn er ook studies die gebruik maken van ERP (event-related potentials), subdurale EEG (YAZAWA *et al.*, 2000), MRI (magnetic resonance imaging) (SOON *et al.*, 2008; SOON *et al.*, 2013) en MEG (magnetoencephalography) (ERDLER *et al.*, 2000). Sinds kort meet

men ook rechtstreeks de hersenactiviteit van aparte neuronen tijdens breinoperaties (FRIED, MUKAMEL & KREIMAN, 2011).

### 3. KRITIEKEN OP DE LIBET-TYPE EXPERIMENTEN

Ondanks al deze replicaties en variaties blijven de LT-experimenten voorwerp van kritiek en repliek. Je kunt drie soorten kritieken op de LT-experimenten onderscheiden. Vooreerst zijn er de *methodologische bezwaren* waarbij men de juistheid van het bekomen resultaat,  $(RP)t_1 < (W)t_2 < (M)t_3$ , in twijfel trekt. Ik concentreer me hier op  $(RP)t_1 < (W)t_2$ . In de tweede plaats is er de *wetenschapsfilosofische kritiek* die dit resultaat niet betwist, maar vraagtekens plaatst bij de interpretatie en extrapolatie. Wat is het verband tussen RP en W? Kun je dit resultaat veralgemenen naar alle vrijwillig genomen beslissingen? Ten slotte is er de *metafysische kritiek* die de pijlen richt op de vermeende implicaties voor ons mens- en maatschappijbeeld. Zelfs indien het resultaat juist is, RP causaal verbonden is met W en dit resultaat en verband gelden voor alle menselijke beslissingen, is dit dan het einde van de vrije wil, schuld, verantwoordelijkheid? De metafysische kritiek betwist dit. Ik begin met de methodologische kritiek.

#### 3.1. Methodologische kritiek

Bepaalde methodologische bezwaren zijn achterhaald. Aanvankelijk stoorden critici zich aan het gering aantal proefpersonen of aan het gebrek aan keuze in de oorspronkelijke taak daar men alleen het tijdstip kon kiezen. Door de vele replicaties en variaties, waaronder die met meer proefpersonen (bijvoorbeeld  $N = 19$  in TREVANA & MILLER, 2002) en die met keuzetaken (keuze tussen linker- en rechterwijsvinger bewegen in HAGGARD & EIMER, 1999), vervallen beide kritieken.

Ook een stuk achterhaald is de kritiek op de timing van de RP-aanvang. Die timing berust op de gemiddelde RP over de trials heen en niet op individuele RP's per trial. EEG laat niet toe om de individuele RP per trial weer te geven. Daarvoor is die hersengolf te zwak. Je krijgt maar een duidelijk resultaat wanneer je de hersengolven over de verschillende trials samenneemt tot een gemiddelde hersengolf. Dan is het mogelijk dat voor elke individuele trial de volgorde  $(W)t_1 < (RP)t_2$  of de volgorde  $(RP)t_1 = (W)t_2$  geldt, maar omdat sommige RP's en W's erg vroeg beginnen krijg je de valse indruk van  $(RP)t_1 < (W)t_2$ . RADDER en MEYNER (2012) illustreren dit effect met de volgende vergelijking. Stel dat er zich in een ruimte 40 filosofen bevinden en 120 kinderen en dat elk kind één van deze filosofen als ouder heeft, dan zou het fout zijn te concluderen dat elke filosoof één of meerdere kinderen heeft. Immers, één of meerdere filosofen kunnen heel veel kinderen hebben, terwijl andere er geen hebben. Hetzelfde kan zich voordoen bij de berekening van de gemiddelde RP-aanvang ten opzichte van het gemiddelde W-moment.

Er zijn verschillende manieren om te controleren voor dit vertekenend effect. Ofwel elimineer je de vroegste RP's en W-momenten uit je analyse en kijk je of dit de volgorde  $(RP)t_1 < (W)t_2$  aantast. Ofwel selecteer je net die vroege trials en kijk je of die volgorde ook dan gehandhaafd blijft. TREVANA en MILLER (2002) hebben beide technieken toegepast en telkens bleef de volgorde  $(RP)t_1 < (W)t_2$  overeind. Recent onderzoek op cellulair niveau holt deze kritiek verder uit (FRIED *et al.*, 2011). Dit onderzoek werkt niet met gemiddelden en meet de aanvang van de hersenactiviteit in de individuele hersencellen ten opzichte van het W-moment per individuele trial. In het gebied waar met EEG de RP's worden gemeten, de *supplementary motor area* (SMA), waren 23% van de geselecteerde

neuronen voor W-moment actief. Die neurale activiteit zag je bij elke trial en vanaf circa 1 seconde voor het W-moment. Hoewel niemand weet of deze cellulaire activiteit een weerspiegeling is van de EEG-gereedheidspotentiaal, is nu wel zeker dat aan het W-moment neurale activiteit voorafgaat die op één of andere manier verband houdt met vrijwillig genomen beslissingen. Over dit verband heb ik het straks.

De felste kritiek is er op de subjectiviteit van het getimede W-moment. Deelnemers zeggen zelf wanneer ze voor het eerst de keuze of de beslissing namen. Dit is geen eenvoudige taak. Tegelijkertijd moet je de beslissing ervaren en ze je precies herinneren. Klokwijsers of voorbijglijdende getallen helpen om die timing te objectiveren, maar het blijven hulpmiddelen. Stel dat de deelnemers manifest liegen over dit W-moment (wat nogal onwaarschijnlijk is) of stel dat de deelnemers een zekere neiging hebben om dat W-moment later te dateren dan in werkelijkheid het geval is (wat al waarschijnlijker is), dan is er geen enkele reden om aan te nemen dat  $(RP)t_1 < (W)t_2$  de juiste volgorde is. Vast staat dat er een grote variatie in W-momenten is, van ruim 1 seconde voor de beweging tot 1 seconde na (!) de beweging (KELLER & HECKHAUSEN, 1991; TREVENA & MILLER, 2002). Bovendien zijn er wel degelijk proefpersonen met een volgorde  $(W)t_1 < (RP)t_2$  (HAGGARD & EIMERS, 1999; MATSUHASHI & HALLETT, 2008), waarbij men gemakshalve aanneemt dat zij hun W-moment per vergissing te vroeg timeden, maar meer dan een vooronderstelling is dit niet.

Om deze vertekening beter in kaart te brengen vergeleken MATSUHASHI & HALLETT (2008) de klassieke trials *met* opdracht om het W-moment te onthouden met trials *zonder* zulke opdracht. Wanneer de deelnemers toevallig een toon hoorden tijdens het ervaren W-moment werd hen gevraagd om van de intentie af te zien om hun wijsvinger te buigen. Het tijdstip van de toon tijdens de intenties vergeleken ze met het tijdstip van het meegeedeelde W-moment in de LIBET-type klokstudies. Ze stelden vast dat die eerste bijna 1 seconde vroeger komt dan de laatste. Is dit systematisch het geval, dan haalt dit de basisbevinding – RP komt 350 to 800 milliseconden voor het W-moment – onderuit. Maar dit is een te snelle conclusie. Vooreerst kwam de RP-aanvang in deze studie nog altijd voor het vervroegde W-moment. In de tweede plaats verschilt deze taak van de taak in het klassieke LT-experiment. Er wordt deelnemers gevraagd van een beslissing af te zien wanneer ze een toon horen. Die hogere complexiteit kan ertoe leiden dat het W-moment vroeger ligt. Ten slotte is ook deze studie afhankelijk van subjectieve rapportering. Als de deelnemers de toon te laat of te vroeg horen, dan vertekent dit evenzeer de resultaten.

Hoe dan ook, de timing van het W-moment is gevoelig voor vertekening. Parallel psychologisch onderzoek bevestigt dit (DANQUAH *et al.*, 2008; MILLER *et al.*, 2010; JOORDENS *et al.*, 2002). Tot het ogenblik waarop wetenschap een objectievere maat voor de timing van het W-moment heeft, is het verstandig om rekening te houden met een voldoende grote foutmarge. De jongste generatie voorspellende studies doen dit ook. FRIED *et al.* (2011) maten dezelfde cellulaire activiteit bij hun hersenpatiënten wanneer ze het W-moment 200 ms vervroegden, maar plaatsten zij het W-moment nog vroeger, dan tastte dit de voorspellende kracht van die hersenactiviteit aan. SOON *et al.* (2008) hadden nog meer marge. Zij vonden hersenactiviteit die toeliet om liefst tot 10 seconden voor het W-moment uit te maken of de proefpersoon links dan wel rechts zou gaan drukken. Het precieze tijdstip van kiezen was tot 5 seconden voor de beslissing af te leiden uit activiteit in de SMA. Deze studies zijn duidelijk minder kwetsbaar voor een verkeerde timing van het W-moment, al roepen zij tal van andere vragen op, waaronder enkele wetenschapsfilosofische waarop ik nu vervolgens inga.

### 3.2. Wetenschapsfilosofische kritiek

Deze kritiek aanvaardt de juistheid van  $(RP)_t_1 < (W)_t_2$ , maar bekritiseert de causale interpretatie van deze volgorde, alsook de extrapolatie van dit resultaat naar alle vrijwillige beslissingen. Ik begin met de eerste kritiek.

Als iemand beweert dat RP de (volledige) oorzaak is van W, dan is deze kritiek volkomen juist. LIBETS oorspronkelijke experiment maakt dit al duidelijk. Herinner je dat RP ook optreedt bij actieve reacties die vooraf bepaald zijn. Je kunt dus niet beweren dat RP alleen bij vrij genomen beslissingen voorkomt. Voortgezet onderzoek toont bovendien aan dat RP identiek is voor alle beslissingen, en dus niet specifiek voor een bepaalde beslissing. Zo vroeg de Duitse breinonderzoeker HERRMANN zijn proefpersonen om met de linkervinger af te drukken wanneer ze een vierkant op het scherm zagen (HERRMANN, 2008). Zagen ze een cirkel, dan drukten ze rechts af. De figuren werden lukraak getoond. De gemiddelde tijd die nodig was om af te drukken bedroeg circa 400 ms na de vertoning van de figuur. Niettemin vond HERRMANN RP's die 1300 ms voor het afdrukken begonnen. Dus: 900 ms vooraleer je weet of je links of rechts moet afdrukken!

Is RP de oorzaak van W, dan hoeft RP geen voldoende en geen noodzakelijke voorwaarde te zijn. Een brandende sigaret zorgt niet op zijn eentje voor een huisbrand en huisbranden ontstaan ook zonder brandende sigaret. Is er sprake van een complexere causale relatie tussen RP en W waarbij RP toch de oorzaak van W is? RADDER & MEYNEN (2011) hebben dit onderzocht en geconcludeerd dat de LT-experimenten hier geen bewijs voor leveren. Samengevat komt hun terechte kritiek hierop neer. We weten niet of RP *plus allerlei andere hersenprocessen* voldoende is om W tot stand te brengen. De LT-experimenten geven ons enkel informatie over de prioriteit van RP voor W, maar niet over het causale effect dat RP gecombineerd met allerlei hersenprocessen op W heeft. Aangezien RP niet specifiek is voor vrijwillige beslissingen, is RP – wil het een oorzaak zijn – afhankelijk van die specifieke hersenprocessen die gecombineerd met RP W opleveren. Over die hersenprocessen bieden de LT-experimenten ons geen informatie. Bijgevolg weten we niet of RP inderdaad een oorzaak is in de totstandkoming van W.

Breinmodulatie kan ons meer zekerheid verschaffen. Neem de brandende sigaret weg (*kunstmatige onderbreking*) of steek een brandende sigaret aan (*kunstmatige stimulatie*) en kijk of er (nog) huisbrand volgt. Helaas staat men vandaag nog niet zover. Kunstmatige onderbreking is om ethische en methodologische redenen niet gemakkelijk om uit te voeren. Je kunt het menselijke beslissen niet (tijdelijk) uitschakelen bij gezonde proefpersonen. Kunstmatige stimulatie is gebeurd in twee studies waarbij rechtstreekse elektrische stimulatie van SMA en PPC (posterieure pariëtaal cortex) inderdaad leidde tot een verlangen om een specifiek lichaamsdeel te bewegen (FRIED, 1991; DESMURGET, 2009). Bij kunstmatige SMA stimulatie was dit verlangen om te bewegen specifiek. De proefpersonen wilden hun linkerarm of tong kort bewegen. Bij PPC stimulatie bleef dit algemener. Ze wilden hun voet bewegen, maar wanneer de experimentleider vroeg hoe de patiënt zijn voet wilde bewegen, antwoordde die dat hij het niet wist. Maar bewijst dit dat deze voorafgaande hersenactiviteit essentieel is voor het nemen van vrijwillige beslissingen? Neen, je kunt hier tegenin brengen dat stimulatie van RP-relevante gebieden (SMA en PPC) een nog onbekend beslissingsgebied activeert dat ook alleen in staat is om W te veroorzaken. Door een toevallige connectie met SMA of PCC wordt dit beslissingscentrum actief.

Tot nu toe is het causale verband tussen RP en W onbewezen. Eén studie suggereert overigens dat het verband tussen beide louter toevallig is. MILLER, SHEPHERDSON & TREVENA

(2011) vergeleken de RP in een klassiek LT-experiment met en zonder kloktiming. In de laatste conditie vroeg men proefpersonen gewoon om hun hand te bewegen telkens wanneer ze dat wilden, maar vroegen ze niet om dit tijdstip aan de hand van een klok te timen. De RP's in deze conditie waren veel minder uitgesproken, terwijl de proefpersonen nochtans een identieke beslissingstaak uitvoerden. Heeft RP dan niets te maken met het vrijwillig beslissen maar eerder met het timen van deze beslissing met een roterende klok? Een minder vergaande interpretatie van dit 'klokeffect' is dat het timen van het W-moment de RP van het beslissen eenvoudig versterkt. Helaas laat EEG niet toe om het aandeel van elke cognitieve activiteit in de gezamenlijk RP afzonderlijk te bepalen.

Opnieuw zorgt de timing van het W-moment voor problemen. Die timing vraagt om extra cognitieve activiteiten die interfereren met het loutere beslissen. Een oplossing hiervoor biedt onderzoek dat minder afhankelijk is van deze timing. Precies daarin voorzien predictiestudies die zich (1) niet beperken tot het voorspellen van het W-moment maar ook de eigenlijke keuze van de proefpersoon voorspellen en (2) zich ook niet beperken tot meting van de RP of hersenactiviteit ten hoogte van SMA, maar zoeken naar voorspellende breinactiviteit in andere hersengebieden, en dus lang voor het moeilijk correct te dateren W-moment. De studies van SOON *et al.* (2008; 2013) en in mindere mate FRIED *et al.* (2011) zijn hierin toonaangevend, maar opnieuw niet vrij van kritiek. De studies van SOON *et al.* (2008; 2013) voorspelden motorische (links of rechts afdrukken) en abstracte keuzes (getallen aftrekken of optellen) met een accuraatheid van 60% op basis van hersenactiviteit uit diverse hersengebieden. Maar is 60% voldoende accuraat? Dit is niet meer dan 10% boven kansniveau. En wat met 40% inaccuraat voorspelde keuzes? Zijn dit onvoorspelbare, vrije keuzes of keuzes die met fijnere technieken ook te voorspellen zijn? Bovendien bewijst accuraat voorspellen een onderliggende correlatie, maar geen causaliteit. Neerstrijkende ganzen kondigen de winter aan maar veroorzaken deze natuurlijk niet.

Niemand kent op dit ogenblik het inhoudelijk verband tussen RP en W. Met zekerheid is geweten dat aan het W-moment en aan de W-keuze allerlei hersenprocessen, waaronder RP maar ook andere, voorafgaan die het tijdstip en de inhoud van de beslissing met soms erg bescheiden accuraatheid voorspellen. Hieruit concluderen dat deze hersenprocessen deel uitmaken van het neurale circuit dat vrijwillige beslissingen veroorzaakt is voorlopig een brug te ver. De vrije wilscepticus zal opwerpen dat zo'n neuraal beslissingscircuit hoe dan ook moet bestaan. Of dit circuit nu gevormd wordt door deze hersenprocessen of door andere maakt niet zoveel uit. Elke cognitieve handeling gaat terug op hersenprocessen die aan die handeling voorafgaan en die handeling veroorzaken. Alleen is dit een metafysisch standpunt dat niet door wetenschappelijke feiten aangetoond is en derhalve voorwerp blijft van filosofische controverse (zie *metafysische kritiek*). Meer kennis van het beslissende brein zal uitwijzen of die sprong van correlatie naar causaliteit gerechtvaardigd is.

Zelfs indien de causale interpretatie de juiste is, dan hoeft dit niet voor alle vrije beslissingen te gelden. Opdat deze kritiek niet triviaal zou zijn – je kunt de causaliteit niet voor elke individuele beslissing controleren – moet je kandidaat-uitzonderingen goed motiveren. Twee goede kandidaten bleken toch niet zo uitzonderlijk te zijn. *Abstracte* beslissingen (getallen optellen of aftrekken) kun je even goed voorspellen als motorische keuzes (links of rechts afdrukken), al is het tijdsvenster waarbinnen de voorspelling succesvol kleiner (4 seconden in plaats van 10) en is het niet mogelijk om het tijdstip van de abstracte keuze accuraat te voorspellen (SOON *et al.*, 2013).



Ook *vetobeslissingen* vormen geen uitzonderingen die de causale dans ontspringen. Bij een *vetobeslissing* beslis je om een keuze uiteindelijk af te breken. Net voordat je een boze mail wilt verzenden, blaas je het verzenden toch af. LIBET had zijn proefpersonen in een aparte conditie om zulke *vetobeslissingen* gevraagd (LIBET, 1983). Tot de ronddraaiende klok op 10 uur stond moesten ze de intentie hebben om hand te buigen, maar net voor 10 uur moesten ze dit voornemen laten varen. Ook bij *vetobeslissingen* treden er RP's op. Je ziet een stijgend EEG-patroon dat tussen 1 seconde en 700 ms voor de (niet-uitgevoerde) beweging begint. Maar er is ook een interessant verschil. Tussen 200 en 100 ms voor het afgesproken tijdstip om de beweging niet uit te voeren, daalt het patroon. Op het ogenblik dat de proefpersoon zijn intentie afbreekt om zijn hand te buigen, dooft de elektrische activiteit waardoor het lijkt dat deze *vetobeslissingen* uit het niets ontstaan. LIBET was hier erg tevreden mee. We zijn misschien niet vrij om te beslissen, maar wel om een beslissing te verhinderen. We hebben een *free won't* in plaats van een *free will*.

25 jaar later schrijven onderzoekers onomwonden: 'the hypothesis of a special, non-neural veto process could therefore become unnecessary' (BRASS & HAGGARD, 2007). Met MRI en ERP is hersenactiviteit gevonden die aan *vetobeslissingen* voorafgaat (BRASS & HAGGARD, 2007; KÜHN, HAGGARD & BRASS, 2009). In een recent ERP-experiment (FILEVICH, KÜHN & HAGGARD, 2013) kregen proefpersonen de opdracht om na een signaal ofwel snel af te drukken ofwel hun afdrukken uit te stellen. Dit signaal deelde mee of ze een vrije trial dan wel een onvrije trial speelden. Bij onvrije trials was de keuze vooraf bepaald. Bij vrije trials mochten ze zelf kiezen. Tot 150 milliseconden voor dit signaal vonden onderzoekers hersenactiviteit die verschilt tussen vrije trials om het afdrukken uit te stellen en vrije trials om snel te reageren. Vooraleer proefpersonen weten of zij vrij mogen beslissen, zie je hersenactiviteit die samenhangt met de specifieke keuze die later volgt. *Vetobeslissingen* verschillen niet van beslissingen tot uitvoering van een motorische handeling. Ook aan deze beslissingen gaat voorspellende hersenactiviteit vooraf.

Andere kandidaat-uitzonderingen vallen moeilijk te bewijzen. Niemand ontkent dat de beslissingstaken in LT-experimenten zoals links of rechts drukken, optellen of aftrekken te eenvoudig zijn. Mensen beslissen ook over vakantieplannen, studierichting, huwelijk of echtscheiding. Dit zijn keuzes op grond van redenen en argumenten en bijgevolg keuzes die er toe doen en die je alleen neemt als je motivering sterk genoeg is. Voorspelt neurale activiteit deze gemotiveerde, lange termijnbeslissingen? Kun je op voorhand zien of iemand beslist om zijn vrouw of man te verlaten? Het probleem bij dit soort keuzes is dat ze erg persoonlijk zijn, ze zich niet op een tijdstip laten vastpinnen, je ook niet weet of ze niet op voorhand al vastliggen en ze interfereren met allerlei cognitieve en emotionele processen. Daarom zijn de simpele LT-experimenten net zo elegant. Je hebt van al deze complicaties geen last. Maar is eenvoudig een fundamenteel tekort? Hebben we redenen om aan te nemen dat complexe beslissingen ook anders verlopen? Waarom zou je bij complexe beslissingen geen voorafgaande hersenactiviteit observeren die deze beslissing voorspelt en mogelijk maakt? Al zijn er theoretische gronden om dit te geloven, experimenten met abstracte beslissingen en *vetobeslissingen* manen aan tot voorzichtigheid.

### 3.3. *Metafysische kritiek*

Deze kritiek verwerpt de implicaties van de LT-experimenten op ons mens- en maatschappijbeeld. Zelfs indien de bevinding  $(RP)t_1 < (W)t_2$  juist is, RP causaal verbonden is met W en dit resultaat en verband gelden voor alle menselijke beslissingen, dan heeft



dit geen effect op het al dan niet bestaan van de vrije wil, schuld, verantwoordelijkheid. Centraal mikpunt van deze kritiek is de invulling van 'vrije wil als keuzevrijheid'. Verwerp je die invulling, dan hoef je niets te vrezen van wetenschappelijke resultaten die 'vrije wil als keuzevrijheid' ondermijnen. Dit is precies wat de zogenaamde compatibilisten zoals Daniel DENNETT (1986) beogen. Zij beweren dat 'keuzevrijheid' niet nodig is om ons een 'vrije wil die de moeite waard is' (DENNETT, 1986) te geven.

'Keuzevrijheid' vraagt twee zaken. Aan de ene kant moet je keuze hebben. Dit betekent dat *indeterminisme* op het gebied van menselijke beslissingen een metafysische realiteit moet zijn. Als het universum deterministisch werkt, dan heb je die keuze nooit. Maar zoals in de inleiding gesteld, zeggen de LT-experimenten niets over dit aspect van keuzevrijheid. Aan de andere kant moet je controle hebben over de keuze. Je keuze mag niet door allerlei onbewuste hersenprocessen al genomen zijn. Wil de beslissing jouw keuze zijn, dan moet je in staat zijn om ondanks alle onbewuste oorzaken, zelf de oorzaak te blijven van die keuze. De beslissing is *up to you*. De LT-experimenten bedreigen dit element van *ultieme controle*. Heeft ons brein, lang voor onze 'vrije' beslissing, al een keuze gemaakt, dan is van ultieme controle geen sprake.

Niet elke filosoof vindt ultieme controle noodzakelijk voor een vrije wil. Dit is alleen het geval voor (*agent-causation*) libertaristen en harde incompatibilisten. Die eersten, waaronder Robert KANE (1998) en Timothy O'CONNOR (2000), geloven dat vrije wilbeslissingen genomen worden door een agent die causaal onafhankelijk is van allerlei sturende hersenprocessen. Wat ons brein ook wil, onze geest, ons Ik of een andere agent beslist om die wil uit te voeren of net niet. Zij verwerpen de LT-experimenten, vooral een causale interpretatie van de volgorde  $(RP)_t_1 < (W)_t_2$ . Zij houden vol dat breinprocessen een beslissing weliswaar materieel mogelijk maken, maar formeel nog niet veroorzaken. Harde incompatibilisten zoals Derk PEREBOOM (2001) verwerpen het bestaan van zo'n causaal onafhankelijke agent. Niettemin zijn zij ervan overtuigd dat een echte vrije wil zo'n onafhankelijke agent vereist. Hierin verschillen ze niet van die libertaristen. Alleen trekken zij de conclusie dat aangezien zo'n agent niet bestaat, een vrije wil niet mogelijk is. Harde incompatibilisten zijn dus vrije wilsceptici.

Voor deze twee metafysische posities in het vrije wildebat hebben de LT-experimenten grote implicaties. De harde incompatibilisten omarmen ze; de (*agent-causation*) libertaristen verwerpen ze. Voor de andere posities zijn de LT-experimenten irrelevant. Een compatibilist zal de schouders ophalen en niet begrijpen waarom ze het einde betekenen van vrije wil, schuld en verantwoordelijkheid. Noch het *indeterminisme*, noch *ultieme controle* zijn noodzakelijk. Wanneer iemand zich vrij voelt om keuzes te maken die hij zelf wilt, dan volstaat dit voor het compatibilisme. In tegenstelling tot een alcoholicus die wil stoppen met drinken maar het niet kan, is een levensgenieter die stopt wanneer hij vindt dat hij voldoende heeft genoten een vrij mens. Al is zijn keuze om te stoppen of door te gaan wellicht even gedetermineerd als die van de alcoholicus, in tegenstelling tot een verslaafde ervaart hij keuzevrijheid. Geen objectieve keuzevrijheid, maar wel subjectieve keuzevrijheid. Voor een compatibilist is dit voldoende. Voor een compatibilist bevinden vrije wil, schuld en verantwoordelijkheid zich in onze subjectieve belevingswereld.

Wie heeft gelijk? Dit hangt af van het aanvaarden van filosofische argumenten waar we hier niet op in kunnen gaan. Ik beperk me tot de vaststelling dat, zoals zo vaak in filosofie, er geen eensgezindheid is over de waarde van deze filosofische argumenten. Sommigen vinden ze erg overtuigend, terwijl anderen weigeren om er zich mee in te laten. Uit dit algemene gebrek aan filosofische consensus volgt dat vrije wil, schuld en verantwoordelijkheid niet meteen hoeven te verdwijnen als de

bevindingen van LT-experimenten correct zijn. Dit hangt van je metafysische positie af. Voor een libertarist zijn ze dodelijk, maar voor een compatibilist zijn ze irrelevant. Een bijkomend probleem hierbij is dat hersenwetenschappers soms weinig afweten van het filosofische debat of gemakshalve – maar geheel ten onrechte – aannemen dat hun intuïtie ook de algemeen geldende intuïtie is. Op die manier ontstaat zowel bij hersenwetenschappers als bij leken de oppervlakkige indruk dat de LT-experimenten enorme maatschappelijke implicaties hebben, terwijl dat niet automatisch zo hoeft te zijn. Of dit ook zo is, bepaalt je positie in het wijsgerige debat. Compatibilisten geven een andere invulling aan de vrije wil die verenigbaar is met een causaal-deterministisch mens- en wereldbeeld. Al ben je het met deze positie niet eens, er zijn geen definitieve argumenten die toelaten die positie te negeren.

#### 4. IMPLICATIES VOOR HET STRAFRECHT

Vragen de LT-experimenten een ander strafrecht? Een strafrecht zonder vrije wil, schuld, verantwoordelijkheid? Neen. Vooreerst blijft het onduidelijk wat de LT-experimenten precies betekenen. Deze experimenten tonen aan dat aan elke beslissing hersenactiviteit voorafgaat op basis waarvan je met bescheiden accuraatheid de genomen beslissing bij eenvoudige taken kunt voorspellen. Maar elke claim die verder reikt dan dit, is discutabel. Maakt die voorspellende hersenactiviteit deel uit van een neurale beslissingscircuit? Wie een causaal-deterministisch mens- en wereldbeeld omarmt, zal die stap zetten, maar op dit ogenblik is hier geen wetenschappelijk bewijs voor en blijft er ruimte voor alternatieve standpunten.

Daarnaast bepaalt je filosofische positie hoe verrijkend de maatschappelijke gevolgen zijn. Wat is de filosofische positie van het Belgisch strafrecht? Hoewel ik niet meteen over statistieken beschik, heb ik de indruk dat de meeste juristen en magistraten in een vrije wil geloven en die aan keuzevrijheid verbinden. Maar begrijpen zij deze keuzevrijheid objectief (indeterminisme, ultieme controle) of eerder subjectief (autonomie); is hun invulling van vrije wil eerder libertaristisch dan wel compatibilistisch? In het eerste geval zijn de gevolgen voor het strafrecht enorm. Een schuldloos strafrecht maakt geen onderscheid meer tussen toerekeningsvatbare en ontoerekeningsvatbare daders en kan daders niet langer straffen omdat ze straf verdienen (hoewel andere strafdoelen zoals afschrikking weliswaar acceptabel blijven). Schuldloos strafrecht wordt zuiver maatregelenrecht. In het tweede geval is de schade onbeduidend want schuld, toerekening en verantwoordelijkheid krijgen een compatibilistische invulling. Echter, net als hersenwetenschappers nemen juristen en magistraten geen duidelijke metafysische positie in. We weten dus niet wat de precieze juridische implicaties van de LT-experimenten zijn. Overigens kan je dit positief duiden. Een te expliciete positionering kan in strijd zijn met de filosofische neutraliteit die je van het strafrecht verwacht.

Niettemin wil ik toch twee bemerkingen maken die wat subjectiever en persoonlijker zijn. Ik geef ze de lezer mee ter reflectie. Een eerste bemerking heeft betrekking op het *compatibilistische strafrecht*. Hoe ziet dit strafrecht eruit waarin alle elementen die verband houden met objectieve keuzevrijheid, dus met indeterminisme en ultieme controle, verwijderd zijn? Denk hierbij aan de (on)toerekeningsvatbaarheid. Wie deze beoordeling wat van nabij volgt, weet hoe snel psychiaters en rechters hun toevlucht nemen tot uitdrukkingen zoals 'X twijfelde dus X kon anders gekozen hebben', 'X kon zich ook niet agressief gedragen want in gelijkaardige omstandigheden kon hij het ook', 'X heeft zichzelf de opdracht gegeven om dit misdrijf te plegen dus hij had vrije wil'.

In een compatibilistische beoordeling verdwijnen deze uitdrukkingen. Maar is dit realistisch? Ik heb mijn twijfels en beperk me tot één voorbeeld. De Amerikaanse rechtsfilosoof Stephen MORSE deelt mijn mening niet en heeft het volgende compatibilistische voorstel in petto (MORSE, 2011). Morse stelt terecht dat het aartsmoelijk is om het onderscheid te maken tussen een verdachte die beweert dat hij zich 'niet kon controleren' en een verdachte die zich eigenlijk 'niet wilde controleren'. De eerste is niet toerekenbaar, de laatste wel. Bij dwang lukt dat misschien nog wel, maar bij onweerstaanbare drang moet je onderscheid maken tussen iemands eigen wil en een vreemde wil en aannemelijk maken dat die vreemde wil dominant werd over die eigen wil op het ogenblik van het delict. Hoe toon je dit aan? MORSE stelt voor om in deze gevallen alleen voort te gaan op het 'wetens' element en het 'willens' element te laten vallen. Wist de verdachte dat het delict niet kon? Indien hij het wist, dan is hij toerekenbaar. Wist hij het niet, dan maakt hij kans op vrijspraak wegens ontoerekenbaarheid. Maar wat dan met de verdachte die oprecht weet en eerlijk gelooft dat een vreemde wil zijn eigen wil buitenspel zette? Hij wordt net als een leugenaar die zulke vreemde wil veinst gestraft. Dit voelt onrechtvaardig aan.

Wil MORSE geen discussie meer over 'niet willen controleren' versus 'niet kunnen controleren' omdat dit uitmondt in een debat over de (objectieve) keuzevrijheid van de verdachte? Het strafrecht gaat ervan uit dat wie zich oprecht beroept op onweerstaanbare dwang geen keuzevrijheid heeft, terwijl wie dit veinst wel nog keuzevrijheid heeft. Maar omdat keuzevrijheid in de metafysische betekenis niet thuishoort in het compatibilisme wil MORSE deze discussie liever vermijden. Het gevolg is dat sommige oordelen over toerekenbaarheid ethisch onverdedigbaar worden want indruisen tegen ons rechtvaardigheidsgevoel. Ik ben dus behoorlijk sceptisch over de mogelijkheid van een compatibilistisch strafrecht. Ofwel veronderstel je indeterminisme en ultieme controle wat theoretisch niet mag, ofwel kom je tot uitspraken die ethisch niet verdedigbaar zijn.

Een tweede bemerking houdt verband met vermeende *sociaal nihilisme* dat uit de LT-experimenten zou voortvloeien. Stel dat de bevinding  $(RP)_t, < (W)_t$  klopt, de causale interpretatie juist is en de implicaties met zich meebrengen dat vrije wil, schuld en verantwoordelijkheid illusies zijn. Is dit maatschappelijk een ramp? Velen menen van wel waardoor kritiek op de LT-experimenten een soort strijd voor de goede zaak wordt. Zo eindigt William KLEMM zijn kritisch overzicht met een apocalyptische schets waarin elke misdadiger vrijuit gaat, de samenleving niet meer (met doodstraf) mag straffen en jongeren massaal aan de drugs gaan, dit omdat een onbestaande vrije wil alles verontschuldigt en eigen initiatief volledig verlamt. KLEMM besluit dat 'the monstrous magnitude of zombian nihilism requires us to reject cavalier acceptance of research that purports to show that there is no free will' (KLEMM, 2010, 61). De bevindingen van de LT-experimenten mogen gewoon niet waar zijn. Wanneer die aantonen dat al onze beslissingen voortkomen uit voorafgaande, onbewuste processen, dan komt de rechtstaat in gevaar.

Is dit zo? Harde incompatibilisten ontkennen dat dit het geval moet zijn (PEREBOOM, 2002; WALLER, 1991; VERPLAETSE, 2011). Die indruk ontstaat door allerlei misvattingen. Eén van die misvattingen is dat een samenleving zonder vrije wil, verantwoordelijkheid en schuld meteen een samenleving is zonder plichten, straf, afkeuring of eigen initiatief. Verantwoordelijkheid en vrijheid zijn begrippen met meerdere betekenissen. Met het verdwijnen van één betekenis verdwijnt de andere betekenis niet. Zo verdwijnt verantwoordelijkheid in de betekenis van schuld maar niet in de betekenis van plicht. Zo verdwijnt keuzevrijheid in zijn objectieve, metafysische betekenis, maar niet in zijn

subjectieve, ethische betekenis van autonomie. Er is geen enkele theoretische reden om te geloven dat het afscheid van een vrije wil meteen ook het afscheid betekent van allerlei zaken zonder dewelke een samenleving inderdaad niet kan functioneren. Wie kritisch is ten aanzien van de LT-experimenten en terecht waarschuwt voor onterechte conclusies en interpretaties, hoort ook kritisch te zijn voor allerlei vermeende sociale implicaties die een onbestaande vrije wil met zich zou meebrengen.

Je kunt opwerpen dat de grote massa deze subtiele onderscheiden niet begrijpt. Zij hebben de illusie van de vrije wil nodig. Bovendien zijn er sociaalpsychologische experimenten die tonen dat proefpersonen meer geneigd zijn om normen sneller te overtreden en minder wilskracht te tonen wanneer ze twijfelen aan het bestaan van de vrije wil (VOHS & SCHOOLER, 2008; BAUMEISTER *et al.*, 2009; ALQUIST, AINSWORTH & BAUMEISTER, 2013).<sup>1</sup> Echter, die opmerking en die experimenten zijn beide erg kwetsbaar voor het volgende probleem. Stel dat wetenschap definitief aantoont dat de vrije wil niet bestaat, wat is dan de beste optie? Ofwel leg je aan de massa uit dat een onbestaande vrije wil geen maatschappelijke ramp hoeft te zijn. Je ontkracht de misvattingen en benadrukt dat iedereen plichten blijft hebben en straffen mogelijk blijft. Ofwel houd je verborgen dat de vrije wil niet bestaat omdat de massa beter gelooft in een illusie, al weet je dat die theoretisch niet strikt noodzakelijk is. Mij lijkt deze tweede optie ethisch verwerpelijk. Enerzijds is het raadselachtig hoe je bewezen feiten in een democratische kennismaatschappij zomaar kunt verzwijgen. Anderzijds is het bedenkelijk dat je mensen zou aanmoedigen aan een illusie vast te houden die maatschappelijk eigenlijk niet hoeft. Als iets waar blijkt te zijn, dan heb je de opdracht om met die waarheid te leren leven. Liegen en illusies zijn geen optie.

Bovendien biedt leven zonder vrije wil zelfs perspectieven. Dit geldt ook voor het strafrecht. Mochten de bevindingen van de LT-experimenten correct blijken te zijn en het compatibilistische alternatief geen lange toekomst beschoren zijn, dan is dit nog niet het einde van het strafrecht, zoals de Nederlandse specialist MOOIJ aan het begin van deze bijdrage liet uitschijnen. Toegegeven, zonder vrije wil zal het strafrecht er anders uitzien. Maar anders staat niet gelijk aan onmogelijk en betekent ook niet slechter. Met de algemene justitiële crisis op de achtergrond, de polemieken over de kwaliteit van de psychiatrische rapportering over de toerekening, de overbevolking van de gevangenis en de ontbrekende effectiviteit van (gevangenis)straffen, is een radicale reflectie op de grondslagen van het strafrecht zelfs welkom. Hoewel de huidige resultaten van de LT-experimenten vaak overdreven worden, en ze zeker niet aantonen dat mensen geen vrije wil hebben, stimuleren ze deze radicale reflectie over een toekomstige samenleving en een strafrecht zonder vrije wil.

## REFERENTIES

- ALQUIST, J.L., AINSWORTH, S.E. & BAUMEISTER, R.F. (2013). Determined to conform: Disbelief in free will increases conformity. *Journal of Experimental Social Psychology*, 49, 80-86.
- BAUMEISTER, R.F., MASICAMPO, E.J. & DEWALL, C.N. (2009) Prosocial benefits of feeling free: disbelief in free will increases aggression and reduces helpfulness. *Personality and Social Psychology Bulletin*, 35, 260-268.
- BRASS, M. & HAGGARD, P. (2007). To do or not to do: the neural signature of self-control. *Journal of Neuroscience*, 27, 9141-9145.

<sup>1</sup> Voor een erg pittige discussie over deze experimenten, zie polemiek tussen James L. MILES (MILES, 2013) en Andrew J. VONASCH en Roy F. BAUMEISTER (VONASCH & BAUMEISTER, 2013) in *British Journal of Social Psychology*.

- DANQUAH, A.N., FARRELL, M.J. & O'BOYLE, D.J. (2008). Biases in the subjective timing of perceptual events: LIBET *et al.* (1983) revisited. *Consciousness & Cognition*, *17*, 616-627.
- DENNETT, D.C. (1986). *Elbow Room. Varieties on free will worth wanting*. Cambridge, Massachusetts: MIT press.
- DESMURGET, M., REILLY, K.T., RICHARD, N., SZATHMARI, A., MOTTOLESE, C. & SIRIGU, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science*, *324*, 811-813.
- ERDLER, M., BEISTEINER, R., MAYER, D., KAINDL, T., EDWARD, V., WINDISCHBERGER, C., LINDINGER, G. & DEECKE, L. (2000). Supplementary motor area activation preceding voluntary movement is detectable with a whole-scalp magnetoencephalography system. *Neuroimage*, *11*, 697-707.
- FILEVICH, E., KÜHN, S. & HAGGARD, P. (2013). There is no free won't: antecedent brain activity predicts decisions to inhibit. *PLoS ONE*, *8*(2), e53053.
- FRIED, I., KATZ, A., MCCARTHY, G., SASS, K.J., WILLIAMSON, P., SPENCER, S.S. & SPENCER, D. (1991). Functional organization of human supplementary motor cortex studies by electrical stimulation. *Journal of Neuroscience*, *11*, 3656-3666.
- FRIED, I., MUKAMEL, R. & KREIMAN, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron*, *69*, 548-562.
- HAGGARD, P. & EIMER, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, *126*, 128-133.
- HARRIS, S. (2012). *Free Will*. Houston: Free Press.
- HERRMANN, C.S., PAUEN, M., MIN, B., BUSCH, N.A. & RIEGER, J.W. (2008). Analysis of a choice-reaction task yields a new interpretation of LIBET's experiments. *International Journal of Psychophysiology*, *67*, 151-157.
- KANE, R. (1998). *The significance of free will*. Oxford: Oxford University Press.
- KORNHUBER, H.H. & DEECKE, L. (1964). Hirnpotentialänderungen beim Menschen vor und nach Willkürbewegungen, dargestellt mit Magnetbandspeicherung und Rückwärtsanalyse. *Pflügers Archiv für die gesamte Physiologie des Menschen und der Tiere*, *281*, 52.
- LIBET, B., WRIGHT, E.W. JR. & GLEASON, C.A. (1982). Readiness-potentials preceding unrestricted 'spontaneous' vs. preplanned voluntary acts. *Electroencephalography and Clinical Neurophysiology*, *54*, 322-335.
- LIBET, B., GLEASON, C.A., WRIGHT, E.W. & PEARL, D.K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*, *106*, 623-642.
- LIBET, B., WRIGHT, E.W. JR. & GLEASON, C.A. (1983). Preparation- or intention-to-act, in relation to pre-event potentials recorded at the vertex. *Electroencephalography and Clinical Neurophysiology*, *56*, 367-372.
- LIBET, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*, 529-566.
- LIBET, B., FREEMAN, A. & SUTHERLAND, J.K.B. (eds.) (1999). *The volitional brain: towards a neuroscience of free will*. Exeter: Imprint Academic.
- LIBET, B. (2004). *Mind time: the temporal factor in consciousness*. Cambridge, MA: Harvard University Press.
- JOORDENS, S., VAN DUJIN, M. & SPALEK, T. M. (2002). When timing the mind one should also mind the timing: biases in the measurement of voluntary actions. *Consciousness and Cognition*, *11*(2), 231-40.
- KELLER, I. & HECKHAUSEN, H. (1990). Readiness potentials preceding spontaneous motor acts: voluntary vs. involuntary control. *Electroencephalography and Clinical Neurophysiology*, *76*, 351-361.

- KLEMM, W. (2010). Free will debates: Simple experiments are not so simple. *Advances in Cognitive Psychology*, 6(1), 47-65.
- KÜHN, S., HAGGARD, P. & BRASS, M. (2009). Intentional inhibition: how the 'veto-area' exerts control. *Human Brain Mapping*, 30, 2834-3843.
- MATSUHASHI, M. & HALLETT M. (2008). The timing of the conscious intention to move. *European Journal of Neuroscience*, 28, 2344-2351.
- MILES, J.B. (2013). Irresponsible and a disservice: the integrity of social psychology turns on the free-will dilemma. *British Journal of Social Psychology*, 52, 205-218.
- MILLER, J.O., VIEWEG, P., KRUIZE, N. & McLEA, B. (2010). Subjective reports of stimulus, response, and decision times in speeded tasks: How accurate are decision time reports? *Consciousness and Cognition*, 19, 1013-36.
- MILLER, J., SHEPHERDSON, P. & TREVENA, J. (2011). Effects of clock monitoring on electroencephalographic activity: is unconscious movement initiation an artifact of the clock? *Psychological Science*, 22(1), 103-9.
- MOOU, A.W.M. (2012). De toerekeningsvatbaarheid. Hoe verder? *Delikt en Delinkwent*, 42(1), 36-53.
- MORSE, S.J. (2011). Mental disorder and criminal law. *The journal of criminal law & criminology*, 101, 885-968.
- O'CONNOR, T. (2000). *Persons and causes. The metaphysics of free will*. Oxford: Oxford University Press.
- PEREBOOM, D. (2001), *Living without free will*. Cambridge: Cambridge University Press.
- RADDER, H. & MEYNEN, G. (2013). Does the brain 'initiate' freely willed processes? A philosophy of science critique of LIBET-type experiments and their interpretation. *Theory & Psychology*, 23(1), 3-21.
- SIRIGU, A., DAPRATI, E., CIANCIA, S., GIRAUX, P., NIGHOGHOSSIAN, N., POSADA, A. & HAGGARD, P. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nature neuroscience*, 7, 80-84.
- SOON, C.S., BRASS, M., HEINZE, H.-J. & HAYNES, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature neuroscience*, 543-545.
- SOON, C.S., HE, A.H., BODE, S. & HAYNES, J.-D. (2013). Predicting free choices for abstract intentions. *Proceedings of the National Academy of Sciences of the U.S.A.*, 110(15), 6217-6222.
- SPENCE, S.A. (1996). Free will in the light of neuropsychiatry. *Philosophy, Psychiatry, & Psychology*, 3, 75-90.
- TREVENA, J.A. & MILLER, J. (2002). Cortical movements preparation before and after a conscious decision to move. *Consciousness and Cognition*, 11, 162-190.
- VERPLAETSE, J. (2011). *Zonder vrije wil. Een filosofische essay over verantwoordelijkheid*. Amsterdam: Nieuwezijds.
- VOHS, K.D. & SCHOOLER, J.W. (2008). The value of believing in free will. Encouraging a belief in determinism increases cheating. *Psychological Science*, 19, 49-54.
- VONASCH, A.J. & BAUMEISTER, R.F. (2013). Implications of free will beliefs for basic theory and societal benefit: Critique and implications for social psychology. *British Journal of Social Psychology*, 52(2), 219-27.
- WALLER, B.N. (1991). *Freedom without responsibility*. Philadelphia: Temple University Press.
- YAZAWA, S., IKEDA, A., KUNIEDA, T., OHARA, S., MIMA, T., NAGAMINE, T., TAKI, W., KIMURA, J., HORI, T. & SHIBASAKI, H. (2000). Human presupplementary motor area is active before voluntary movement: subdural recording of Bereitschaftspotential from medial frontal cortex. *Experimental Brain Research*, 131, 165-177.