

Foutendetectie in de menselijke hersenen

Sander Nieuwenhuis

[intro]

Electrofysiologisch onderzoek heeft aangetoond dat onze hersenen zijn uitgerust met een mechanisme voor het herkennen van fouten. Op het moment dat mensen onder tijdsdruk een verkeerde reactie op een stimulus geven, laat het EEG een scherpe negatieve potentiaal zien: de “error-related negativity”. Deze potentiaal weerspiegelt waarschijnlijk een interactie tussen het dopamine-systeem en de mediale frontale cortex. De amplitude van de error-related negativity is abnormaal in verscheidene neuropsychiatrische stoornissen waaronder obsessief-compulsief stoornis, Gilles de la Tourette syndroom, schizofrenie en de ziekte van Parkinson. Dit levert mogelijk interessante aanwijzingen voor de theorievorming over de functionele en neurale oorzaak van deze aandoeningen.

Menselijke fouten kunnen worden geclasseerd in twee categorieën. Sommige fouten (*vergissingen* of *mistakes*) worden begaan door een gebrek aan kennis over de correcte actie. Dit kan bijvoorbeeld gebeuren als iemand een situatie verkeerd interpreert en vervolgens actie onderneemt op basis van die verkeerde interpretatie. Meestal worden mensen zich pas bewust van dit soort fouten door een signaal uit de omgeving. Zulke externe feedback vormt vaak de basis voor leren. Een ander type fout (*slips*, goede Nederlandse vertaling ontbreekt) wordt gekenmerkt door het uitvoeren van een verkeerde actie ondanks dat de juiste kennis en intentie aanwezig is. Een bekend voorbeeld uit het dagelijks leven vormen ‘slips of the tongue’: er komen andere woorden uit je mond dan de bedoeling was. Een ander voorbeeld vormt het (mis)slaan met de honkbalknuppel naar een bal die ‘wijd’ wordt gegooid door de pitcher. In beide gevallen is de juiste kennis en intentie aanwezig om geen fouten te maken (zoals niet slaan naar de bal). Alleen omdat er haastig wordt of moet worden gehandeld gaat het mis; Meestal realiseert de persoon zich een fractie van een seconde later dat hij een fout heeft gemaakt. In de jaren ’60 heeft Patrick Rabbitt met behulp van een serie meerkeuze-reactietijdexperimenten aangetoond dat mensen dit soort fouten zeer betrouwbaar kunnen detecteren en rapporteren. Ook liet Rabbitt zien dat mensen zich goed aanpassen na fouten: ze gaan bijvoorbeeld langzamer (= voorzichtiger) reageren (Rabbitt, 1966). Dit suggereert dat ons cognitieve systeem op efficiënte wijze reageert op fouten.

Een neurale systeem voor het detecteren van fouten

Ruim tien jaar geleden vonden onderzoekers het eerste bewijs dat de menselijke hersenen razendsnel reageren wanneer er een fout (slip) wordt gemaakt. Michael Falkenstein en collega’s lieten proefpersonen een tweekeuze-reactietijdtaak uitvoeren (zie figuur 1a) en vergeleken het EEG-signal volgend op correcte reacties met het EEG-signal volgend op foute reacties (Falkenstein et al., 1991). Foute reacties werden gekenmerkt door een negatieve potentiaal die ongeveer 60 ms na de knopindruk zijn minimum bereikte (zie figuur 1b). Deze fouten-gerelateerde potentiaal is inmiddels in tientallen laboratoria gerepliceerd en is bekend geworden als de *error-related negativity* (ERN). De ERN treedt op na foute reacties met de handen, voeten, ogen, en stem, en zowel als de proefpersoon de verkeerde overte reactie geeft (zoals bij slips of the tongue) als wanneer deze per ongeluk een reactie geeft op een “nogo-signaal” (zoals in het voorbeeld van de honkballer; Falkenstein et al., 2000). Verder is de potentiaal groter in amplitude naarmate de fout groter is (bijvoorbeeld bij knopindrukken met de verkeerde vinger van de verkeerde hand), en naarmate de proefpersoon meer waarde hecht aan accuratesse. Samenvattend kan gezegd worden dat de

ERN optreedt na allerlei typen fouten, en dat de amplitude van de potentiaal groter is naarmate de fout saillanter is.

Een ander belangrijk aspect van de ERN is dat deze zo bijzonder snel optreedt na het maken van de fout. De *onset* van de potentiaal vindt meestal al plaats vóór de registratie van de foute knopindruk van de proefpersoon (zie figuur 1b). Dit suggereert dat de ERN het resultaat is van een vergelijking tussen de bedoelde reactie en een “efference copy” van de werkelijke reactie - dat wil zeggen, een kopie van de opdracht die door centrale hersengebieden naar perifere motorstructuren is gestuurd. Het idee dat de ERN niet afhankelijk is van lichamelijke feedback is onlangs bevestigd door de observatie van een ERN in een (unieke) patiënt wier hersenen volkomen afgesloten zijn van afferente signalen (door de gevolgen van Guillain-Barré Syndroom; Allain et al., 2004). Verscheidene onderzoekers hebben gesuggereerd dat de hersenactiviteit die tot uitdrukking komt in de ERN fungeert als een alarmsignaal dat andere delen van de hersenen aanzet tot het corrigeren van de fout, of tot bijstellingen die verdere fouten moeten voorkomen (bijvoorbeeld het verhogen van de concentratie). Een voorspelling die voortvloeit uit dit idee is dat de amplitude van de ERN samen zou moeten hangen met de grootte van aanpassingen in het gedrag na fouten. Sommige (maar niet alle) onderzoeken hebben deze voorspelling bevestigd (Gehring et al., 1993). Een aansprekend voorbeeld hiervan is gerapporteerd door Ridderinkhof en collega’s die vonden dat alcohol-inname (in doses vergelijkbaar met 2+ glazen) leidt tot een verkleining van de ERN en minder grote gedragsaanpassingen na fouten (Ridderinkhof et al., 2002).

Er is sterke evidentie dat de ERN wordt gegenereerd in de “rostral cingulate zone” in het posteriore deel van de mediale frontale cortex (Ridderinkhof et al., 2004). Dit hersengebied beslaat de caudale anterieure cingulate schors en delen van omliggende hersengebieden. Het gebied is niet alleen betrokken bij het detecteren van fouten maar wordt ook geactiveerd bij onzekere of moeilijke keuzes, straf en beloning, en pijn (Ridderinkhof et al., 2004). Deze activiteit is ook te meten op de schedel: Er treden ERN-achtige potentialen op in het EEG wanneer proefpersonen een moeilijke keuze moeten maken, of wanneer ze een feedback-stimulus ontvangen die aangeeft dat ze een fout hebben gemaakt of iets van waarde hebben verloren (Nieuwenhuis et al., 2004). Het is op dit moment nog niet duidelijk of al deze gebeurtenissen ook door dezelfde neuronen gecodeerd worden. Een andere mogelijkheid is dat naburige neuronen ieder een eigen, doch aanverwante functie hebben. In elk geval lijkt de activiteit van de posteriore mediale frontale cortex gemoduleerd te worden door allerlei soorten plezierige en onplezierige gebeurtenissen. Een belangrijk doel voor toekomstig onderzoek is het ontwikkelen van theorieën die alle functies van dit gebied onder één noemer kan brengen (zie Brown & Braver, 2005, voor een stap in die richting).

Vergrote ERN in obsessief-compulsief stoornis en aanverwante stoornissen

Obsessief-compulsief stoornis (OCS), ook wel bekend als dwangstoornis, is een stoornis waarbij de kwaliteit van het leven in aanzienlijke mate wordt aangetast door angst, dwanggedachten en dwanghandelingen. Gehring en collega’s vergeleken het EEG van patiënten met OCS met dat van een controlegroep (gematched met betrekking tot leeftijd, geslacht en opleiding) tijdens de uitvoering van een Stroop-taak. De twee groepen verschilden niet veel wat betreft taakprestatie: De patiënten maakten even veel fouten als de controlegroep en reageerden iets sneller. Het belangrijkste verschil tussen de twee groepen betrof de ERN, die sterk in amplitude vergroot was bij de OCS-patiënten (Gehring et al., 2000). De vondst van een grotere ERN bij OCS is consistent met het gegeven dat OCS samenhangt met verhoogde activiteit in de mediale frontale cortex (met name de anterieure cingulate schors en orbitofrontale cortex) en basale ganglia, onder andere in reactie op fouten (Saxena et al., 1998; Ursu et al., 2003). Deze structuren worden verondersteld een “circuit” te

vormen dat verantwoordelijk is voor het in de gaten houden (*monitoring*) van onze informatieverwerking en acties.

De sterke electrofisiologische reactie op fouten bij deze patiënten is ook consistent met een theorie van Pitman (1987) over de functionele oorzaak van OCS. Volgens deze theorie wordt ons gedrag gecontroleerd door een cognitief systeem dat externe input vergelijkt met een interne standaard (bijvoorbeeld een te bereiken doel). Dit *comparator*-systeem berekent het verschil tussen het input-signalen en de interne standaard, en wanneer deze niet overeenkomen wordt er een *error* signaal gegenereerd. Dit signaal leidt tot een gevoel van onbehagen en brengt cognitieve, motorische, en affectieve systemen op de hoogte van de noodzaak de ervaren discrepantie te corrigeren. Volgens Pitman hebben OCS-patiënten een intrinsiek defect aan het comparator-systeem, zodat dit systeem ongepaste en versterkte *error* signalen genereert. Dit leidt tot de typische symptomen van OCS: een aanhoudend gevoel dat er iets niet goed is, en de drang om te corrigeren.

Overigens is de vergrote ERN bij fouten niet volledig specifiek voor OCS. Mensen met het syndroom van Gilles de la Tourette laten ook een vergrote ERN zien (Johannes et al., 2002). Dit syndroom is een neurologische aandoening die gekenmerkt wordt door motorische en/of vocale tics en komt vaak in combinatie voor met OCS. Er zijn aanwijzingen dat OCS en Gilles de la Tourette syndroom een gedeeltelijk gemeenschappelijke neurale basis hebben, waaronder een verhoogde activiteit van de mediale frontale cortex (zie Johannes et al., 2002). Mensen die bijzonder gespannen (*anxious*) zijn laten tevens een grotere ERN zien (Hajcak et al., 2003). Vervolgonderzoek heeft laten zien dat ERN-amplitude met name samenhangt met gespannenheid als persoonlijkheidstrekk (Trait anxiety) en niet met gespannenheid als toestand (State anxiety): Moser en collega's lieten mensen met een spinnenfobie een flankertaak (zie figuur 1a) uitvoeren onder twee condities en meetten de ERN volgend op foute reacties. In de experimentele conditie zat de proefleider naast het computerscherm met een tarantula in zijn handen te spelen. In de controle-conditie speelde de proefleider met een bal in zijn handen. Hoewel vragenlijsten (en informele observaties!) aangaven dat de proefpersonen in de experimentele conditie veel angstiger waren, verschilderde de amplitude van de ERN niet tussen de twee condities (Moser et al., in press).

De ERN en dopamine-deficiënte populaties

Volgens een invloedrijke theorie wordt de ERN veroorzaakt door fasische dopaminerge modulatie van neuronale activiteit in de mediale frontale cortex (Holroyd & Coles, 2002). Dit idee is consistent met de sterke betrokkenheid van het dopamine-systeem bij het coderen van plezierige en onplezierige gebeurtenissen. In de theorie van Holroyd en Coles fungeert fasische dopaminerge activiteit ten gevolge van plezierige of onplezierige gebeurtenissen als een signaal dat door de mediale frontale cortex wordt gebruikt om gedrag dat tot die gebeurtenissen heeft geleid (respectievelijk positief of negatief) te bekraftigen. Op deze manier leert het systeem gedragingen aan die tot plezierige gebeurtenissen leiden, en leert het onplezierige gebeurtenissen te vermijden. Twee psychofarmacologische studies met gezonde proefpersonen hebben bewijs geleverd voor een rol van het dopamine-systeem in het genereren van de ERN. In een studie werd gevonden dat inname van de dopamine-antagonist haloperidol leidde tot een kleinere ERN en een verslechterde leerprestatie (Zirnheld et al., 2004). In een andere studie werd gevonden dat de dopamine-agonist amfetamine leidt tot een grotere ERN (De Bruijn et al., 2004).

Het is bekend dat disfunctioneren van het dopamine-systeem een belangrijke rol speelt in het veroorzaken van schizofrenie, de ziekte van Parkinson, en de mentale achteruitgang tijdens normale veroudering. Verscheidene studies hebben aangetoond dat de ERN kleiner is bij patiënten met schizofrenie en de ziekte van Parkinson, en in gezonde ouderen (Falkenstein et al., 2001; Mathalon et al., 2002; Nieuwenhuis et al., 2002). Deze

resultaten lijken goed aan te sluiten bij de theorie van Holroyd en Coles (2002). Echter, een belangrijke en grotendeels onbeantwoorde vraag is of en hoe de kleinere ERN gerelateerd kan worden aan de specifieke symptomen in ieder van deze populaties. In ons eigen onderzoek hebben we aanwijzingen gevonden dat de kleinere ERN in gezonde ouderen nauw samenhangt met een lagere snelheid van feedback-gestuurd leren (Nieuwenhuis et al., 2002). Met behulp van een computermodel van de theorie van Holroyd en Coles hebben we laten zien dat de geobserveerde samenhang tussen kleinere ERN-amplitudes en lagere leersnelheid kan worden verklaard door een verminderde sterkte van dopaminerige bekrachtigingssignalen in ouderen (Nieuwenhuis et al., 2002).

De abnormale ERN in schizofreniepatiënten is door enkele auteurs in verband gebracht met de positieve symptomen van de stoornis, en met name met wanen dat de eigen gedachten en acties niet zelf gecontroleerd maar van buitenaf ingegeven zijn. Chris Frith heeft opgeworpen dat deze positieve symptomen mogelijk veroorzaakt worden door problemen met het monitoren van de eigen handelingen: De waarneming van de eigen handelingen wordt niet op juiste wijze in verband gebracht met de *efference copy* van deze handelingen waardoor het lijkt alsof ze niet zelf-geïnitieerd zijn (Frith, 1987). De slechte werking van het monitoring-systeem komt tot uiting in een verkleinde ERN. Als ondersteuning voor deze hypothese heeft Frith aangetoond dat schizofreniepatiënten met illusoire externe controle van hun gedachten en handelingen minder fouten corrigeren in een motorische taak dan controleproefpersonen en schizofreniepatiënten zonder deze symptomen (Frith & Done, 1989).

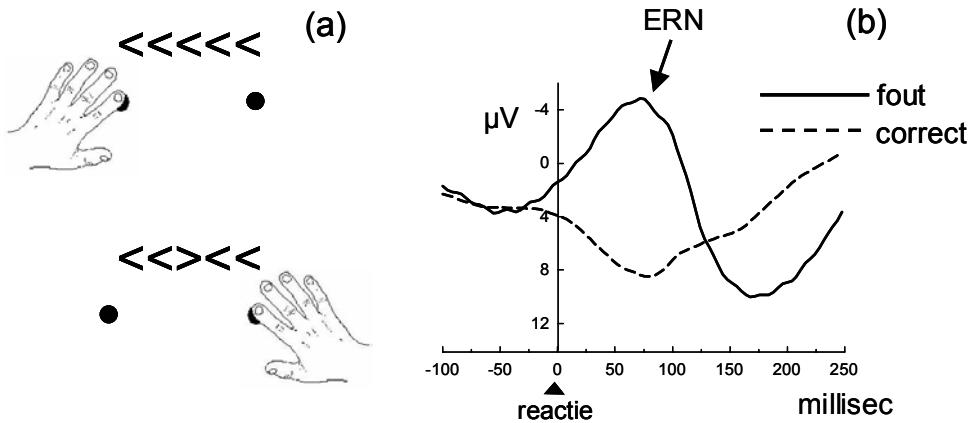
Tot slot

Ongeveer vijftien jaar na het ontdekken van de ERN is het onderzoek naar deze potentiaal en de relatie met cognitieve controle in een behoorlijk ver stadium. Dit maakt het interessant om groepsverschillen in de ERN die samenhangen met neurologische en neuropsychiatrische stoornissen te relateren aan de etiologie van die stoornissen. Aan de ene kant biedt de bestaande kennis over de neurale basis van de ERN mogelijke aanwijzingen met betrekking tot de lichamelijke oorzaken van een stoornis. Aan de andere kant biedt de relatie tussen de ERN, het coderen van onplezierige gebeurtenissen, en foutenverwerking in het bijzonder, een theoretisch raamwerk waarin mogelijk sommige van de symptomen van de stoornis beschouwd kunnen worden.

Literatuur

- Allain, S., Hasbroucq, T., Burle, B., Grapperon, J., & Vidal, F. (2004). Response monitoring without sensory feedback. *Clinical Neurophysiology*, 115, 2014-2020.
- Brown, J.W., & Braver, T.S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, 307, 1118-1121.
- De Brujin, E.R.A., Hulstijn, W., Verkes, R.J., Ruigt, G.S.F., Sabbe, B.G.C. (2004). Drug-induced stimulation and suppression of action monitoring in healthy volunteers. *Psychopharmacology*, 177, 151-160.
- Falkenstein, M., Hielscher, H., Dziobek, I., Schwarzenau, P., Hoormann, J., Sunderman, B., Hohnsbein J. (2001). Action monitoring, error detection, and the basal ganglia: an ERP study. *Neuroreport*, 12, 157-161.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, 78, 447-455.
- Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, 51, 87-107.

- Frith, C.D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological Medicine*, 17, 631-648.
- Frith, C. D., & Done, D. J. (1989). Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychological Medicine*, 19, 359–363.
- Gehring, W.J., Goss, B., Coles, M.G.H., Meyer, D.E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science* 4, 385-390.
- Gehring, W.J., Himle, J., & Nisenson, L.G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science*, 11, 1-6.
- Hajcak, G., McDonald, N., Simons, R.F. (2003). Anxiety and error-related brain activity. *Biological Psychology*, 64, 77-90.
- Holroyd, C.B., Coles, M.G.H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679-709.
- Johannes, S., Wieringa, B.M., Nager, W., Muller-Vahl, K.R., Dengler, R., & Munte, T.F. (2002). Excessive action monitoring in Tourette syndrome. *Journal of Neurology*, 249, 961-966.
- Mathalon, D.H., Fedor, M., Faustman, W.O., Gray, M., Askari, N., Ford, J.M. (2002). Response-monitoring dysfunction in schizophrenia: An event-related brain potential study. *Journal of Abnormal Psychology*, 111, 22-41.
- Moser, J.S., Hajcak, G., & Simons, R.F. (in press). Performance monitoring and attentional allocation during spider phobic symptom provocation. *Psychophysiology*.
- Nieuwenhuis, S., Holroyd C.B., Mol, N., & Coles, M.G.H. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neuroscience & Biobehavioral Reviews*, 28, 441-448.
- Nieuwenhuis, S., Ridderinkhof, K.R., Talsma, D., Coles, M.G.H., Holroyd, C.B., Kok, A., & Van der Molen (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. *Cognitive, Affective & Behavioral Neuroscience*, 2, 19-36.
- Pitman, R.K. (1987). A cybernetic model of obsessive-compulsive psychopathology. *Comprehensive Psychiatry*, 28, 334-343.
- Rabbitt, P.M.A. (1966). Errors and error correction in choice reaction tasks. *Journal of Experimental Psychology*, 71, 264-272.
- Ridderinkhof, K.R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G.P. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science*, 298, 2209-2211.
- Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., & Nieuwenhuis, S. (2004). The role of medial frontal cortex in cognitive control. *Science*, 306, 443-447.
- Saxena, S., Brody, A.L., Schwartz, J.M., Baxter, L.R. (1998). Neuroimaging and frontal-subcortical circuitry in obsessive-compulsive disorder. *British Journal of Psychiatry*, 173 (Suppl. 35), 26-37.
- Ursu, S., Stenger, V.A., Shear, M.K., Jones, M.R., Carter, C.S. (2003). Overactive action monitoring in obsessive-compulsive disorder: Evidence from functional magnetic resonance imaging. *Psychological Science*, 14, 347-353.
- Zirnheld, P.J., Carroll, C.A., Kieffaber, P.D., O'Donnell, B.F., Shekhar, A., & Hetrick, W.P. (2004). Haloperidol impairs learning and error-related negativity in humans. *Journal of Cognitive Neuroscience*, 16, 1098-1112.



Legenda

Figuur 1. (a) Voorbeelden van stimuli en bijbehorende reacties in de flanker-taak, een veelgebruikte taak om de ERN te onderzoeken. De opdracht aan proefpersonen is om zo snel mogelijk de knop in te drukken aan de kant waar de middelste pijl heen wijst. De flankerende pijlen moeten genegeerd worden. Proefpersonen maken veel fouten wanneer de middelste pijl en de flankerende pijlen in tegenovergestelde richtingen wijzen. (b) Voorbeeld van de ERN, de negatieve potentiaal volgend op foute reacties. Op de x-as staat de tijd vanaf het registeren van de (correcte of foute) knopindruk. Op de y-as staat de sterkte van de potentiaal in microvolt. Volgens conventie staan negatieve waarden naar boven geplot.