

# Computationele psychiatrie: een toekomst voor wiskundige modellen in de classificatie en behandeling van psychopathologie?

Zsuzsika Sjoerds · Hanneke E.M. den Ouden

## Samenvatting

Het huidige systeem voor psychiatrische diagnostiek en nosologie is voornamelijk gebaseerd op extern waarneembare symptomen en een categorische classificatie. Dit leidt tot heterogeniteit en comorbiditeit tussen diagnoses. Om classificatie te verbeteren en individuele behandeling te bevorderen, is er behoefte aan een meer dimensionale en kwantitatieve benadering, waarmee onderliggende (niet direct waarneembare) processen en mechanismen worden gedefinieerd. Een dergelijke benadering zal leiden tot toepasbare diagnostische tests die zich richten op pathofysiologische mechanismen die ten grondslag liggen aan verstoorde observeerbare cognitieve en emotionele processen en de daaruit voortkomende psychopathologie. Computationele psychiatrie biedt een handvat tot een dergelijke mechanistische benadering. Door middel van non-lineaire wiskundige modellen wordt informatie geïntegreerd over latente processen die ten grondslag liggen aan (verstoorde) gedrag, simultaan gemeten brainactiviteit, en zelfs effecten van interventies, zoals hersenstimulatie en farmacologie. De hoop is dat deze benadering zal leiden tot een beter begrip van psychiatrische stoornissen op het niveau van (latente) cognitieve processen en de onderliggende neurobiologie en, daaruit volgend, een verbetering van diagnose en behandeling.

In dit artikel introduceren wij eerst de rationale en werkwijze van de computationele psychiatrie, om vervolgens de stappen te bespreken die naar onze mening genomen moeten worden om een succesvolle bijdrage te leveren aan de psychiatrie en gerelateerde specialismen.

**Trefwoorden** psychiatrie · cognitieve neurowetenschappen · computationele modellen · neurobiologie · computationele psychiatrie

---

Dr. Z. Sjoerds (✉)  
Max-Planck fellow-group 'Cognitive and Affective Control of Behavioural Adaptation', Max-Planck Institute for Human Cognitive and Brain Sciences,  
Stephanstraße 1A,  
04103 Leipzig, Duitsland  
e-mail: sjoerds.zs@gmail.com

Dr. H. E.M. den Ouden  
Donders Institute for Brain Cognition and Behaviour,  
Radboud Universiteit,  
Kapittelweg 29,  
6525 EN Nijmegen, Nederland

## Inleiding

Binnen de huidige psychiatrie bestaan helaas nog geen goede, eenduidige pathofysiologische verklaringmodellen voor de descriptief geïnclassificeerde stoornissen. Dit lijkt een contrast te vormen met veel andere specialismen, en hindert een aanpak waarin een voortschrijdend inzicht in de pathofysiologie het ontwikkelen van behandeling aanstuurt. Mede hierdoor sluiten farmacologische bedrijven hun deuren voor verder neuropsychiatrisch onderzoek [1, 2]. Hoewel het *gemiddeld* behandelingsucces binnen de psychiatrie niet onderdoet voor vele andere specialismen, kan succes op *individuele* basis zeer moeilijk voorspeld worden door de enorme variantie binnen en tussen symptoomprofielen, en is behandeling voornamelijk gebaseerd op *trial and error* van meerdere behandelstrategieën. Door het gebrek aan pathofysiologisch begrip gaan verbetering van diagnose, prognose en behandeling langzaam. Hier en daar klinken er zelfs geluiden dat de psychiatrie ‘in crisis’ is [3, 4], en dat de strategie van classificatie en onderzoek moet veranderen [5].

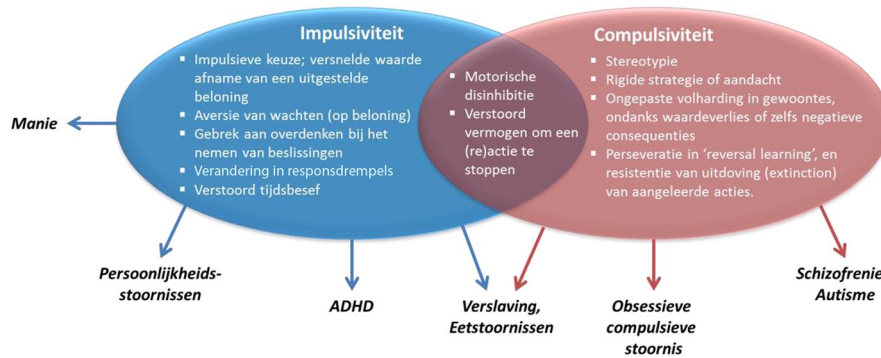
‘Computationale psychiatrie’ is een nieuw veld, dat inzichten en methoden uit de computationele neurowetenschappen [6] vertaalt naar de psychiatrie [7]. Computationale psychiatrie streeft ernaar om de relatie tussen waarneembare psychiatrische symptomatologie en onderliggende (niet direct waarneembare, of latente) cognitieve processen en neurobiologische mechanismen te begrijpen en verklaren door gebruik van wiskundige modellen [8]. Getracht wordt om psychiatrische classificatie op basis van etiologie en pathofysiologie te verbeteren, met als doel een nauwkeuriger diagnose, een betere prognose en een effectievere behandeling [8, 9].

Hoewel het veld relatief nieuw is, duikt de term computationele psychiatrie steeds vaker op binnen de cognitieve en neuropsychiatrische wetenschappen. Sinds 2011 is een aanzienlijk aantal overzichts- en opiniestukken geschreven in toonaangevende internationale vaktijdschriften over deze veelbelovende toepassing [8–18]. In 2013 werd in Florida het eerste congres georganiseerd dat volledig gewijd was aan computationele psychiatrie [19]. Recentelijk zijn er aan de universiteiten van Zürich, Oxford, Londen en Nijmegen workshops georganiseerd om wetenschappers op te leiden in het gebruik van wiskundige modellen binnen hun neuropsychiatrische onderzoek. Als klap op de vuurpijl werd in 2014, in samenwerking tussen het Max-Planck Gesellschaft en University College London, het Centre for Computational Psychiatry geopend, met afdelingen in Londen en Berlijn [20, 21].

Kortom, het bruist binnen de wereld van de computationele psychiatrie. Echter, buiten het directe veld van de cognitieve en neuropsychiatrische wetenschappen is kennis over deze benadering nog niet wijd verspreid. In dit artikel willen wij daarom dit opkomende onderzoeksveld onder de aandacht brengen van Nederlandse klinici, wetenschappers, en belangstellenden. Wij zullen de kernideeën van de computationele psychiatrie bespreken, en stappen schetsen die naar onze mening gezet moeten worden om tot daadwerkelijke klinische toepassingen te komen. Wij zullen vooral een globaal overzicht geven, en dit illustreren aan de hand van enkele voorbeelden. Details over specifieke stoornissen en modellen laten wij hier buiten beschouwing; voor de geïnteresseerde lezer zullen wij waar mogelijk verwijzen naar recent onderzoek en overzichten in de literatuur.

## Problemen binnen de huidige psychiatrie

De psychiatrische diagnostiek en daaropvolgende behandeling zijn sinds jaar en dag gebaseerd op de categoriale indeling van de DSM [22] en ICD [23]. Deze indeling is gefundeerd in de beschrijving van klinische fenomenologie. Dit classificatiesysteem was doelbewust a-theoretisch van opzet, om de psychiatrie los te maken van de sterktheoretische psychoanalyse. Deze benadering was succesvol in de zin dat ze heeft geleid tot meer eenduidigheid, en wetenschappelijk onderzoek mogelijk maakt dat is gericht op welomschreven groepen aandoeningen. Echter, de grenzen van deze nosologie zijn, al sinds de ontwikkeling van de DSM-III, onderwerp van debat. De huidige symptoomgebaseerde classificatie kan ertoe leiden dat patiënten van wie symptomen niet overlappen, wel dezelfde diagnose krijgen. Tegelijkertijd kunnen patiënten die dezelfde symptomen vertonen een verschillend ziektepatroon ontwikkelen en anders op behandeling reageren. Een voorbeeld vanuit de borderlinepersoonlijkheidsstoornis (BPS): volgens de DSM-IV [24] moet een patiënt ten minste vijf van de negen (in de DSM-5: 6 van 11) fenomenologische criteria vervullen om te worden gediagnosticeerd met BPS. Hierdoor zijn er meer dan 250 symptoomcombinaties mogelijk, en bestaat er de mogelijkheid dat twee BPS-patiënten slechts één van negen symptomen gemeen hebben. Deze berekening onderschat bovendien heterogeniteit, omdat ze geen rekening houdt met veel voorkomende comorbiditeit, zoals depressie of verslaving. Deze grote heterogeniteit binnen, alsook comorbiditeit



**Figuur 1** Overlappend symptoomprofiel voor compulsiviteit- en impulsiviteitstoornissen. Een voorbeeld van meetbare cognitieve processen die geassocieerd zijn met het symptoomprofiel van een scala van psychiatrische stoornissen die (gedeeltelijke) overlap vertonen in hun fenomenologie. Met computationele modellen kunnen dergelijke meetbare subprocessen kwantitatief benaderd worden, en gekoppeld worden aan onderliggende neurale mechanismen. Zo kan een

cognitief profiel opgebouwd worden van processen die verstoord zijn in de individuele patiënt, wat dimensionale classificatie kan bevorderen. In dit voorbeeld is een spectrum aan psychiatrische stoornissen weergegeven, dat in te delen is van zeer impulsief naar heel compulsief, en stoornissen die een combinatie van beide laten zien. (Bron: [71] Figuur gereproduceerd en vertaald met toestemming van de auteurs en uitgever (Elsevier)).

tussen diagnostische categorieën, maakt het moeilijk om een pathofysiologisch model te ontwikkelen en goed te voorspellen bij wie welke behandeling succesvol zal zijn.

Er heerst grootschalige overeenstemming over het feit dat het huidige symptoomgebaseerde classificatiesysteem uiteindelijk uitgebreid of zelfs vervangen dient te worden door een systeem dat is gebaseerd op pathofysiologische mechanismen [25], waarbij een meerdimensionale benadering waarschijnlijk meer recht doet aan de complexiteit van psychopathologie dan een categoriale indeling. De dimensionale benadering helpt om over te schakelen van overlappende heterogene categorieën naar een beschrijving van latente subprocessen die (pathologisch) gedrag aansturen, en gezamenlijk ten grondslag liggen aan extern waarneembare symptoomclusters. Een voorbeeld hiervan is weergegeven in fig. 1, waarin geïllustreerd is hoe psychiatrische aandoeningen uit een spectrum van impulsieve-compulsieve pathologie een overlappend profiel vormen van meetbare cognitieve processen. Een mechanistische benadering op basis van pathofysiologie wordt vereenvoudigd door een dergelijke dimensionale benadering. Kwantitatieve diagnostische tests, zoals scans of bloedafname, worden binnen andere medische specialismen standaard uitgevoerd. Nu de neurobiologie steeds belangrijker wordt in de conceptualisering van ziektemechanismen binnen de psychiatrie, wordt het noodzakelijk om ook binnen dit specialisme waarneembaar gedrag te koppelen aan pathofysiologische processen,

en objectieve kwantitatieve tests toe te passen. Er ontbreekt echter een arsenaal aan dergelijke diagnostische tests binnen de psychiatrie. Een belangrijke oorzaak hiervan is dat momenteel onvoldoende kennis beschikbaar is over de neurobiologische basis van psychiatrische aandoeningen om op grond daarvan een pathofysiologisch classificatiesysteem te ontwikkelen [26].

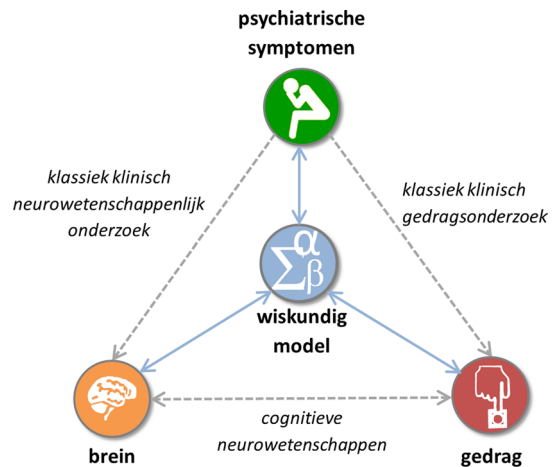
Deze beperkte kennis over de neurobiologische basis van psychiatrische aandoeningen reflecteert echter geen gebrek aan interesse: in de afgelopen decennia is wereldwijd een enorme hoeveelheid studies uitgevoerd waarmee verbanden tussen psychiatrische aandoeningen en cognitieve en neurobiologische maten werden aangetoond. Echter, een vertaling van deze (lineaire) correlaties naar een verbeterde behandeling blijft vooralsnog uit. Dit komt deels omdat deze correlaties gebaseerd zijn op DSM-classificaties, en dus per definitie niet tot andere diagnoses en behandelstrategieën zullen leiden. Daarnaast proberen bestaande studies doorgaans extern waarneembaar gedrag (responspatroon, reactiesnelheid) te koppelen aan psychopathologie, terwijl deze pathologie zeer waarschijnlijk eerder voortkomt uit latente processen die ten grondslag liggen aan (pathologische) variaties in dit waarneembaar gedrag. Ten slotte betreft het altijd correlaties op groepsniveau, aangezien effecten op individueel niveau niet sterk genoeg zijn. In de volgende sectie gaan wij in op hoe de computationele psychiatrie deze impasse kan doorbreken.

## Hoe kan computationele psychiatrie helpen?

Het belangrijkste doel van computationele psychiatrie is de kloof te overbruggen tussen neurobiologie en gedrag, en zo te komen tot een geïntegreerd profiel van het *individu* op deze verschillende niveaus (zie fig. 2). Deze benadering komt voort uit de cognitieve neurowetenschappen, waar computationele modellen met veel succes zijn toegepast in de afgelopen twee decennia. Een model is een schematische voorstelling van de werkelijkheid, dat ontwikkeld wordt aan de hand van bepaalde te meten gedragsfenomenen, met als doel breed toepasbaar te zijn in verschillende onderzoekspopulaties. In een model worden wiskundige formules gebruikt om waarneembaar gedrag aan de hand van onderliggende subprocessen te beschrijven. Om u een voorstelling te geven van hoe een dergelijk model werkt, kunt u aan de dagelijkse weersverwachting denken. Combinaties van verschillende variabelen (luchtdruk, luchtvochtigheid, watertemperatuur, geografische ligging, enz.) voorspellen, aan de hand van wiskundige algoritmen, het weer voor de komende dagen. Computationele modellen binnen de cognitieve neurowetenschappen doen precies hetzelfde door kennis uit de neurowetenschappen en experimentele psychologie te integreren aan de hand van non-lineaire wiskundige algoritmen. Dit heeft als doel zowel neurale activiteit als latente cognitieve processen en waarneembaar gedrag te verklaren en te voorspellen [27]. In fig. 3 is een schematische voorstelling opgenomen van de toepassing van computationele elementen binnen de psychiatrie.

Er bestaat een wisselwerking tussen neurale mechanismen en psychiatrische fenomenologie. Hogere orde cognitieve en emotionele processen zijn afhankelijk van de neurale en computationele bouwstenen, die van een lagere orde zijn [28]. Elke verstoring op het lage niveau van neurale mechanismen (bijvoorbeeld een verlaagde dopamineproductie) heeft ingrijpende gevolgen op de cognitieve processen van een hogere orde, en leiden daardoor tot verandering in gedrag. Modellen variëren in abstractieniveau: van neurobiologische modellen, bijvoorbeeld biofysische en neurale netwerkmodellen, tot meer abstracte modellen, zoals algoritmische modellen van latente cognitieve processen, bijvoorbeeld leersnelheid, beloningsgevoeligheid of voorspellingsfouten (zie box 1), die daaruit voortkomend observeerbaar gedrag beschrijven.

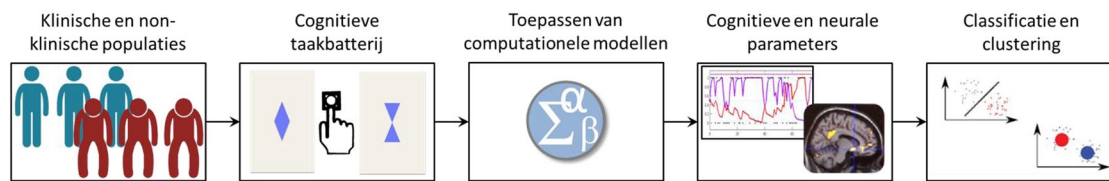
De modellen die worden gebruikt in de computationele psychiatrie onderscheiden zich op twee manieren van andere veelgebruikte modellen. Ten eerste zijn deze modellen



**Figuur 2** De drie voornaamste niveaus die door computationele modellen verbonden kunnen worden. Deze niveaus betreffen psychiatrische symptomatologie, neurobiologie en kwantificeerbaar gedrag in een experimentele omgeving. In dit kader worden klassiek neurowetenschappelijk en gedragsonderzoek samengebracht met de cognitieve neurowetenschappen in een wiskundige benadering, die haar oorsprong vindt in de computerwetenschappen.

kwantitatief. Dat wil zeggen dat ze, in tegenstelling tot *box-and-arrow* modellen, niet alleen verbanden tussen verschillende factoren veronderstellen, maar ook daadwerkelijk de kwantitatieve of wiskundige relatie tussen deze factoren specificeren. Dit dwingt ons impliciete aannames concreet te maken, en maakt modellen experimenteel toetsbaar en vergelijkbaar. Ten tweede zijn deze modellen ‘generatief’ [14, 29]. Een generatief model voorspelt hoe de data geproduceerd worden, bijvoorbeeld hoe je keuzes maakt in een leertaak, om vervolgens te kijken hoe goed de voorspellingen overeenkomen met de daadwerkelijk geobserveerde data [30]. Generatieve modellen staan tegenover zogeheten ‘discriminatieve’ modellen, die slechts proberen om data te classificeren, bijvoorbeeld regressieanalyses, die eenzijdige statistische afhankelijkheid tussen variabelen beschrijven (een vergroting van de ventrikels bij schizofrenie, of verhoogde prevalentie van depressieve stoornissen bij dragers van een bepaald gen).

Deze generatieve computationele modellen dienen ten eerste om gezond gedrag en mechanismen te voorspellen en te begrijpen, aan de hand van normatieve wiskundige algoritmes van neurale en cognitieve functies. Normatief wil hier zeggen dat het model inschat hoe een subject zich ‘dient’ te gedragen, terwijl het streeft naar optimaal gedrag. Op deze manier wordt er een aantal kernprocessen gedefinieerd



**Figuur 3** Niveaus van computationele psychiatrie. Klinische en non-klinische populaties worden getest aan de hand van speciaal ontworpen computertaken. Door het toepassen van computationele modellen op ruwe gedragsdata [72], bijvoorbeeld reactietijd of stimuluskeuze, kunnen cognitieve en neurocognitieve processen (bijvoorbeeld leersnelheid) worden afgeleid en gekwantificeerd als wiskundige parameters in een model. Het schatten van deze processen in de verschillende taken, in combinatie met andere factoren, zoals neurale activiteit gemeten met EEG en fMRI,

en gekwantificeerd aan de hand van meetbare subprocessen. Er kan bijvoorbeeld gedacht worden aan verstoringen in werkgeheugen, aandacht, leren van beloningen, verwerking van sociale informatie, emoties, enzovoorts. Op basis hiervan kan suboptimaal gedrag en afwijkende neurobiologie gedefinieerd en gedetecteerd worden, waardoor de modellen toepasbaar zijn binnen de psychiatrie. De toepassing van computationele modellen binnen psychiatrisch onderzoek betekent echter niet dat alle data opnieuw verzameld dienen te worden. Er bestaat al een grote hoeveelheid aan experimentele taakdata waarop computationele modellen toegepast kunnen worden.

Beloningsleren (*reinforcement learning*) [6] is een cognitief proces ten aanzien waarvan ons begrip grote stappen heeft gemaakt door computationele modellen en empirisch onderzoek te integreren. Deze vorm van leren is relevant bij het gros van de psychiatrische aandoeningen [31]. In de volgende sectie zullen wij daarom aan de hand van een voorbeeld uit het beloningsleren de benadering van de computationele psychiatrie concreter maken.

#### **Box 1. Voorspellingsfouten als voorbeeld van meetbare latente processen en hun neurale correlaten**

In dit artikel beargumenteren wij dat het meten van latente processen en hun koppeling aan neurobiologische mechanismen een van de belangrijkste voordelen is van de computationele benadering. Door gedrag met computationele modellen te benaderen, kan bepaald worden waar het

leveren een multidimensionaal computationeel profiel op. Op basis van deze individuele profielen kunnen patiënten geïdentificeerd en geclusterd worden, om zo subgroepen te definiëren in klinische en non-klinische populaties. Voor deze classificatie en clustering wordt gebruikgemaakt van methoden uit de computerwetenschappen, bijvoorbeeld de zogeheten *supervised/unsupervised learning* algoritmes. De afbeelding is geïnspireerd door figuur 1 uit [17]. Figuur aangepast met toestemming van de auteurs.

op het niveau van latente processen en neurale mechanismen fout gaat binnen psychiatrische aandoeningen. Ter illustratie van dergelijke meetbare latente processen bespreken wij hier het voorbeeld van de ‘voorspellingsfout’ (Engels: *prediction error*). Voorspellingsfouten staan centraal in een breed scala aan theorieën over aspecten van cognitie, van leertheorie tot aandacht en waarneming. De voorspellingsfout is simpelweg het verschil tussen een verwachte uitkomst (bijvoorbeeld de hoogte van een beloning, of het zien van een bepaalde stimulus), en de daadwerkelijke uitkomst. Volgens gangbare theorieën over beloningsleren wordt deze voorspellingsfout gebruikt om de verwachting van toekomstige beloning aan te passen, en het mogelijk te maken optimaal te kiezen tussen acties met verschillende verwachte uitkomsten. Als de uitkomst van een actie beter is dan verwacht, spreekt men van een positieve voorspellingsfout. De gemaakte actie stijgt dan in ‘waarde’, zodat de volgende keer de kans groter is dat deze actie weer uitgevoerd wordt. Is de daadwerkelijke uitkomst slechter dan verwacht, dan spreekt men van een negatieve voorspellingsfout, en daalt de actie in waarde. Is de uitkomst compleet volgens de verwachting (en is er dus geen voorspellingsfout), dan hoeft de verwachting niet meer aangepast te worden, en hoeft er dus niet meer geleerd te worden. De mate waarin een voorspellingsfout invloed heeft op de waarde van de actie, oftewel hoe snel er geleerd wordt, verschilt tussen individuen. Echter, de leersnelheid



kan niet direct geobserveerd worden aan de hand van bijvoorbeeld reactietijd of responspatroon, maar wel met behulp van een leermodel, dat inschat hoe elk individu informatie over de gemaakte responses en geobserveerde uitkomsten combineert.

De signalering van verschillende type voorspellingsfouten heeft een duidelijk aangetoonde neurale basis. Bijvoorbeeld, elektrofysiologische studies bij dieren hebben aangetoond dat de positieve voorspellingsfout (oftewel de uitkomst was beter dan verwacht) wordt gesignaleerd door een verhoging van het fasisch vuren van dopamineneuronen in het ventrale tegmentum [75]. Beeldvormende studies bij mensen, in combinatie met gedragstaken en computationele modellen, hebben aangetoond dat verstoring van dit signaal ook leidt tot verstoord beloningsleren [37, 61, 64].

Verstoord beloningsleren is geobserveerd in een scala aan psychiatrische aandoeningen, en wordt gekarakteriseerd door een aanhoudend inadequate voorspelling van uitkomsten. De oorzaak hiervan kan bijvoorbeeld liggen in verstoord signalering van de voorspellingsfout in de hersenen, of in het feit dat dit (onverstoord) signaal niet gebruikt wordt in het optimaliseren van gedrag. Dit kan leiden tot onaangepast, inflexibel gedrag. Met functionele beeldvormende technieken (fMRI) kan onderzocht worden in welke mate signalering van de voorspellingsfout verstoord is in het brein. Aan de hand van die informatie wordt dan duidelijk wat ten grondslag ligt aan verstoord beloningsleren in psychiatrische aandoeningen: een verstoord signaal in het brein, of verstoord gebruik van het signaal. Dit kan vervolgens een handvat bieden om in toekomstige interventies belangrijke verstoord onderliggende mechanismen te normaliseren, en zo waarneembaar verstoord gedrag te behandelen.

De voorspellingsfout is daarom een goed voorbeeld van een latent proces, dat niet direct extern waarneembaar is, maar wel met wiskundige algoritmes te schatten is op basis van geobserveerd gedrag, en direct gekoppeld kan worden aan neurale signalen. De voorspellingsfout vormt hierdoor ook de basis van de meest succesvolle modellen binnen het huidige computationeel psychiatrisch onderzoek.

## Beloningsleren als een van de bouwstenen van psychiatrie

Beloningsleren houdt zich bezig met hoe een subject (een persoon, een dier of, bij kunstmatige intelligentie, een robot) in complexe omgevingen leert de juiste beslissingen te nemen om beloningen te maximaliseren en verlies of negatieve uitkomsten te vermijden. Een veelbelovende toepassing van beloningsleren binnen computationele psychiatrie is het kwantificeren van de (verstoord) balans tussen automatische handelingen en gecontroleerd overwogen gedrag. De invloed op gedrag van aan de ene kant een ‘reflectief’ en aan de andere kant een ‘reflexief’ systeem is te herleiden naar de heel vroege psychologie, en de noties van *cognitive maps* [32] en de *law of effect* [33, 34], waarin gesteld wordt dat acties die gevolgd worden door beloning herhaald zullen worden, terwijl acties die leiden tot straf/verlies verminderd zullen optreden. Het eerste van de twee systemen is een doelgericht systeem, dat een model ‘bouwt’ van de structuur van de omgeving. Op basis van dat omgevingsmodel kunnen de consequenties van acties worden voorspeld. Dit systeem is flexibel en past zich snel aan een veranderende omgeving aan, maar vergt daardoor rekenkracht en aandacht. Dit systeem wordt aangeduid als een ‘modelgebaseerd’ systeem [35; 36], en lijkt voornamelijk te steunen op de prefrontale cortex [37–39]. Stel je bijvoorbeeld voor dat een collega je vertelt dat de kortste route naar je werk is afgesloten. Jouw modelgebaseerde systeem kan nu op basis van zijn kennis van de stad een andere route plannen. Als je je door dit systeem laat leiden, fiets je de volgende dag zonder problemen naar je werk. Het tweede, meer subcorticale [37, 38, 40], systeem is snel, automatisch, en betrokken bij gewoontegedrag. Dit systeem ‘ontlast’ de cortex, zodat het tegelijkertijd andere taken kan uitvoeren. Het past zich echter minder snel aan een veranderende omgeving aan. Het leert langzaam, en alleen van directe ervaringen, en maakt geen gebruik van kennis over de omgeving. Het wordt daarom aangeduid als ‘modelvrij’ [35, 36]. Als jij op je fiets naar je werk alvast in gedachten verzonken bent over je volgende onderzoek of patiënt, kan het zo zijn dat je alsnog bij de afgesloten straat uitkomt.

Onder normale omstandigheden is de samenwerking tussen deze twee systemen in balans. Het is echter aangetoond dat bijvoorbeeld onder stress [41–43] de balans wordt verstoord, en rigide gedrag gaat overheersen, wat kan leiden tot gewoontegedrag. Bij een aantal psychiatrische aandoeningen, voornamelijk binnen het

impulsiviteit-compulsiviteitspectrum (fig. 1) bestaat het vermoeden dat deze twee systemen in disbalans zijn. Zo is bijvoorbeeld vanuit dieronderzoek verslaving klassiek beschreven als een stoornis die wordt gekenmerkt door gewoontegedrag [44], en bouwt het bewijs hiervoor zich ook binnen humaan onderzoek op [45–48]. Recent gebruik van de computationele benadering bevestigt bij verschillende clusters (verslaving, dwangstoornissen, eetstoornissen, schizofrenie en depressie) [48–53] dat de balans tussen de twee systemen is aangedaan. Verschillende computationele modellen kunnen nog preciezer aantonen welke onderliggende processen en neurale mechanismen deze balans verstoren (voor een voorbeeld van zo een dergelijk latent proces dat gemeten kan worden, zie box 1). Zo wordt er bijvoorbeeld gedacht dat een verstoring in de balans kan liggen aan verstoord schakelen tussen de twee systemen [53, 54]. Binnen psychiatrische aandoeningen worden ook andere cognitieve processen die betrokken zijn bij beloningsleren met deze computationele aanpak verder ontrafeld [8]. Zo kunnen depressieve symptomen aan de ene kant voortkomen uit een afwijkende inschatting van de hoogte of waarde van een toekomstige beloning [9]. Dit kan bijvoorbeeld ontstaan door een overtuiging over hulpeloosheid, waarbij de patiënt denkt dat zijn gedrag geen invloed heeft op de hoogte van de beloning. Aan de andere kant kunnen taakirrelevante negatieve stimuli keuzes binnen het besluitvormingsproces beïnvloeden door een onvermogen om mentale representaties van aversieve gebeurtenissen te remmen [51]. Deze voorbeelden laten zien dat de computationele benadering een link kan aantonen tussen verstoorde latente cognitieve processen en psychiatrische aandoeningen. Echter, een beter begrip van de neurobiologische en neurocognitieve basis van deze stoornissen op zich is niet voldoende om te leiden tot een verbetering van diagnose en behandeling. In de laatste sectie willen wij bespreken welke concrete stappen er genomen moeten worden alvorens de computationele psychiatrie een concrete bijdrage kan leveren aan het welzijn van patiënten.

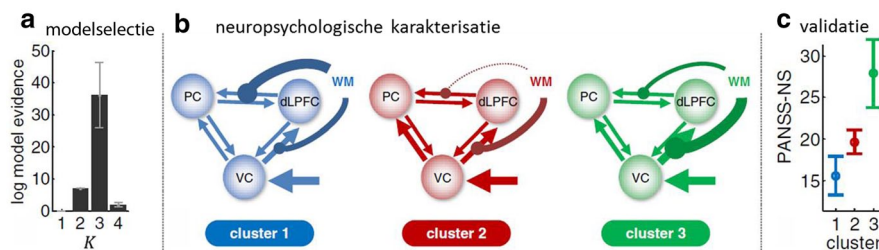
### From bench to bedside

Het doel van de computationele psychiatrie is een meer dimensionale classificatie van patiënten op basis van latente (cognitieve) processen, die cruciaal voorspellend is voor ziekteverloop en behandelings succes. Met andere woorden, de computationele psychiatrie bewijst pas zijn

nut op het moment dat ze daadwerkelijk tot *verandering van behandeltraject* leidt. Classificatie aan de hand van onderliggende mechanismen en endofenotypen met behulp van computationele modellen kan een dergelijke dimensionale indeling van psychiatrische aandoeningen bevorderen. Computationele modellen zijn namelijk expliciet niet beperkt tot de categorische diagnoses van de DSM en ICD. Dit idee sluit aan op de recentelijk door het National Institute of Mental Health ontwikkelde Research Domain Criteria (RDoC) [55], met als doel een nieuwe onderzoeksbenadering te ontwikkelen buiten het huidige categorische classificatiesysteem, door de kracht van moderne onderzoekstechnieken (genetica, neurowetenschappen, gedragswetenschappen) te gebruiken voor een transdiagnostische en biologische benadering van de psychiatrie.

Deze relatief nieuwe benaderingswijze van psychiatrische stoornissen is nog sterk in ontwikkeling, en heeft daarom enkele beperkingen. Om het uiteindelijke doel te bereiken om een kwantitatief profiel van verstoorde cognitieve processen te ontwikkelen dat voorspellend is voor ziekteverloop en differentiële behandelstrategie, zullen naar onze mening de volgende stappen doorlopen moeten worden.

*Stap 1. Ontwikkeling en verfijning van gedragstaken en normatieve wiskundige modellen.* Het theoretisch kader volstaat nog niet voor toepassing van de juiste modellen binnen een klinische omgeving. De ontwikkeling en verfijning van (bestaande) gedragstaken en wiskundige modellen voor het kwantificeren van relevante cognitieve en neurale processen is al wel in volle gang. Belangrijk hierbij is dat de ontwikkelde taken enerzijds door patiënten gemakkelijk uitvoerbaar zijn, maar anderzijds, tegelijkertijd, geschikt zijn voor het meten van relevante processen. Validatie van normatieve data in verschillende populaties is nodig, om natuurlijke variatie in processen (bijvoorbeeld op basis van leeftijd, geslacht of cultuur) in kaart te brengen. Theoretici en empirische onderzoekers zullen nauwer moeten samenwerken, teneinde nieuwe wiskundige modellen te ontwikkelen of bestaande modellen aan te passen voor het verrichten van onderzoek naar gezonde cognitieve processen [28, 30, 35, 37, 56–63]. Alleen dan kunnen afwijkende processen gedefinieerd worden bij verschillende psychiatrische aandoeningen. Een belangrijk component van deze stap is het vaststellen van de biologische basis van deze latente cognitieve processen. Fundamenteel onderzoek, zoals farmacologische en elektrofysiologische studies



**Figuur 4** Voorbeeld van patiëntclustering aan de hand van een computationeel (netwerk)model op basis van fMRI-data. In deze studie voerden 41 patiënten met schizofrenie een simpele werkgeheugentaak uit. De auteurs schatten vervolgens met gebruik van een neurobiologisch netwerkmodel (dynamic causal modeling) [73] de verbindingsterkte tussen de hersengebieden die betrokken zijn bij deze taak, namelijk de dorso-laterale prefrontale cortex (dLPFC), de pariëtale cortex (PC) en de visuele cortex (VC). a Op deze verbindingsterktes lieten de onderzoekers een clusteringsalgoritme los. Hun eerste bevinding was dat er sterk bewijs was voor drie ( $K=3$ ) subclusters (de veelgebruikte maat hiervoor is de log model evidence: de mate waarin een model

een goede representatie van de data is). b Deze drie clusters vertonen verschillende patronen in de verbindingsterkte in de betrokken gebieden tijdens deze werkgeheugentaak (WM), wat verbeeld is met de breedte van de pijlen. c Validatie van de drie clusters op basis van symptoomprofiel (gemeten met de Positive and Negative Syndrome Scale (PANSS) [74]) toonde aan dat de drie clusters duidelijk onderscheiden konden worden op basis van negatieve symptomen, en worden geoperationaliseerd door de negatieve symptomen-schaal van de PANSS. PANSS-NS Positive and Negative Symptom Scale - Negative Symptoms, K Aantal clusters per model. (Bron: [70]. Figuur gereproduceerd en vertaald met toestemming van de auteurs en uitgever (Elsevier)).

bij zowel mens als dier, zal hier een belangrijke bijdrage aan leveren.

*Stap 2. Externe (klinische) validatie.* Vervolgens is externe validatie nodig; er moet worden vastgesteld in hoeverre de met de modellen gemeten processen voorspellend zijn voor individuele verschillen binnen zowel gezonde als psychiatrische populaties. Deze stap is noodzakelijk om een duidelijker beeld te krijgen van de robuustheid en toepasbaarheid van de voor de psychiatrie meest relevante modellen. Er verschijnen steeds meer studies waarin specifieke modellen worden gevalideerd binnen kwetsbare en psychiatrische populaties [64–66]. Binnen neuropsychiatrische studies wordt gemeten in hoeverre patiënten kunnen worden geclassificeerd en de ernst van symptomen kunnen worden voorspeld [49, 50, 67–69]. Deze validatie wordt in eerste instantie gedaan op basis van bestaande DSM-5-definities, alvorens overgegaan wordt naar een meerdimensionale benadering op basis van latente processen en mechanismen.

*Stap 3. Dimensionale herdefinitie van psychiatrie.* Het uitgangspunt van de vorige stap wordt nog steeds gevormd door de bestaande DSM/ICD-criteria en dus de categorische indeling van aandoeningen. Op grond daarvan zullen de hier beschreven resultaten per definitie niet leiden tot een veranderende behandelingsstrategie. Daar is een meer dimensionale classificatie van patiënten voor nodig, op basis van computationele processen, die

cruciaal op betrouwbaarder wijze ziekteverloop en behandelingsucces voorspelt. De volgende noodzakelijke stap omvat daarom het in kaart brengen van mechanismen en latente processen bij overlappende psychiatrische aandoeningen, die de bestaande criteria overstijgen, wat zal leiden tot herclustering van patiënten. Onderzoekers aan de technische universiteit Zürich en het Max-Planck Instituut voor neurowetenschappen in Leipzig publiceerden recentelijk een illustratief voorbeeld van deze fase (fig. 4; [70]). Zij scanden schizofrene patiënten terwijl die een simpele werkgeheugentaak uitvoerden in de MRI-scanner. Op basis van deze data konden zij de sterkte van de hersenverbindingen schatten met behulp van een wiskundig generatief model van een netwerk van motorgebieden. Zij toonden aan dat de 41 patiënten in drie subclusters op te delen waren op basis van verschillen in verbindingsterkte. De externe validiteit van deze subclusters bleek uit verschillen in de ernst van negatieve symptomen in de verschillende clusters. Deze studie laat zien hoe met behulp van computationele modellen heterogene patiëntengroepen op basis van neurobiologische mechanismen (maar niet op basis van de DSM) opgesplitst kunnen worden in meer homogene subgroepen.

*Stap 4. Prospectieve studies naar behandelingsucces.* De laatste en belangrijkste stap is het aantonen van de differentiële voorspellende waarde (t.o.v. de DSM-5) van deze modelgedreven



dimensionale clusters. Er zijn prospectieve studies nodig om vast te stellen: 1) hoe cognitieve processen clusteren en 2) welke combinatie van aangedane processen voorspellend zijn voor behandelsucces en ziekteverloop. De resultaten van deze prospectieve behandelstudies kunnen inzicht geven in het klinisch verloop en de uitkomst van deze processen, en een belangrijke bijdrage leveren aan het testen van de klinische toepasbaarheid van computationele modellen binnen de psychiatrie.

Het moge duidelijk zijn dat dit ‘stappenplan’ geen lineair proces zal zijn, waarbij de ene fase voltooid moet worden voor er aan de andere begonnen kan worden; integendeel, er zal een continue terugkoppeling zijn naar de eerdere fasen. Zo zullen bijvoorbeeld modellen doorlopend verfijnd worden (stap 1) op basis van de prospectieve behandelstudies (stap 4). De toepassing van computationele modellen binnen de psychiatrie zal dus niet van de een op de andere dag plaatsvinden, maar het zal, zoals bij elke nieuwe benadering, een tijd duren voor deze toepassing volledig geïntegreerd kan worden, en op dit moment vooral een waardevolle toevoeging zijn aan bestaand neuropsychiatrisch onderzoek. Daarnaast heeft de beschreven computationele benadering van psychiatrie zich vooralsnog voornamelijk gericht op cognitieve processen, zoals besluitvormingstheorieën, beloningsleren, aandachtenwerkgeheugen, en deneurobiologische basis van deze processen in termen van neurale activiteit en neurofarmacologie. Dit komt echter meer doordat de computationele benadering van psychiatrie haar oorsprong vindt in de computationele cognitieve neurowetenschappen dan dat het een fundamentele beperking is. De computationele benadering zelf heeft geen theoretische beperkingen ten aanzien van de aard van de informatie die gemodelleerd wordt. Uiteindelijk zullen ook andere factoren geïntegreerd dienen te worden in de modellen van de computationele psychiatrie, bijvoorbeeld patiënt-omgevingsinteracties, alsook perifere oorzaken

van bijvoorbeeld endocriene of immunologische aard. Deze ontwikkelingen vereisen, in elk stadium van het proces, terugkoppeling met klinici, die van alle betrokkenen het best bekend zijn met de symptomatologie en fenomenologie van bepaalde psychiatrische aandoeningen, waardoor zij met de belangrijke hypothesevorming en -toetsing kunnen helpen.

## Conclusie

Hoewel het veld van computationele psychiatrie nog relatief jong is, wordt met computationele psychiatrie een nieuwe, elegante toevoeging geïntroduceerd binnen de neuropsychiatrie, met als doel latente processen te kwantificeren en te verbinden met neurobiologie en symptomatologie. Met deze computationele benadering kunnen dimensionale profielen van verstoorde cognitieve/neurale processen ontwikkeld worden, die voorspellend zijn voor ziekteverloop en differentiële behandelstrategie. De computationele psychiatrie zal een belangrijke bijdrage leveren aan het aangaan van de uitdagingen in de diagnostiek en behandeling binnen de geestelijke gezondheidszorg. Bij goede samenwerking tussen klinici, empirici en theoretici bestaat er een vruchtbare toekomst voor computationele psychiatrie, zowel binnen het wetenschappelijk onderzoek, als in de kliniek.

**Dankwoord** Graag willen wij een aantal mensen bedanken voor hun kritische blik en input op dit stuk. Drs. Esmée N. Arredondo, psychiater GGZ Noord-Holland Noord, Kliniek Westfriesland; drs. Dirk E.M. Geurts, wetenschappelijk onderzoeker en psychiater in opleiding, Radboud Universiteit Nijmegen en Maudsley Hospital Londen; dr. H.G. (Eric) Ruhe, psychiater en klinisch onderzoeker, stemmings- en angststoornissen, afdeling Psychiatrie Universitair Medisch Centrum Groningen/Rijksuniversiteit Groningen. Daarnaast danken wij voor toestemming voor hergebruik van figuren: prof. dr. Trevor W. Robbins en coauteurs voor fig. 1, dr. Thomas V. Wiecki en coauteurs voor delen uit fig. 3 en dr. Kay H. Brodersen en coauteurs voor fig. 4.

## Literatuur

1. Cressey D. Psychopharmacology in crisis – Researchers warn of “withdrawal of hope” as funding shrivels. *Nature*. 2011. doi:10.1038/news.2011.367. <http://www.nature.com/news/2011/110614/full/news.2011.367.html>. Geraadpleegd op: 21 aug. 2015.
2. Nutt D, Goodwin G. ECNP Summit on the future of CNS drug research in Europe 2011: report prepared for ECNP by David Nutt and Guy Goodwin. *Eur Neuropsychopharmacol*. 2011;21:495–9.
3. Hyman SE. Revolution Stalled. *Sci Transl Med*. 2012;4:155cm11.
4. Sahakian BJ, Malloch G, Kennard C. A UK strategy for mental health and wellbeing. *Lancet*. 2010; 375:1854–55.
5. Nutt DJ, Attridge J. CNS drug development in Europe—past progress and future challenges. *Neurobiol Dis*. 2014;61:6–20.
6. Sutton RS, Barto AG. Reinforcement learning: an introduction Cambridge. Cambridge: MIT Press; 1998.

7. Kishida KT, King-Casas B, Montague PR. Neuroeconomic approaches to mental disorders. *Neuron*. 2010;67:543–54.
8. Adams RA, Huys QJM, Roiser JP. Computational psychiatry: towards a mathematically informed understanding of mental illness. *J Neurol Neurosurg Psychiatry*. 2015. jnnp-2015-310737. doi:10.1136/jnnp-2015-310737.
9. Huys QJM. Computational Psychiatry. In: Jaeger D, Jung R, redactie. *Encyclopedia of computational neuroscience* New York. New York: Springer; 2013. pp 1–10.
10. Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci*. 2011;14:154–62.
11. Huys QJM, Moutoussis M, Williams J. Are computational models of any use to psychiatry? *Neural Netw*. 2011;24:544–51.
12. Montague PR, Dolan RJ, Friston KJ, Dayan P. Computational psychiatry. *Trends Cogn Sci*. 2012;16:72–80.
13. Friston KJ, Stephan KE, Montague R, Dolan RJ. Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry*. 2014;1:148–58.
14. Stephan KE, Mathys C. Computational approaches to psychiatry. *Curr Opin Neurobiol*. 2014;25:85–92.
15. Wang X-J, Krystal JH. Computational Psychiatry. *Neuron*. 2014;84:638–54.
16. Huys QJM, Guitart-Masip M, Dolan RJ, Dayan P. Decision-theoretic psychiatry. *Clin Psychol Sci*. 2015;3:400–21.
17. Wiecki TV, Poland J, Frank MJ. Model-based cognitive neuroscience approaches to computational psychiatry: clustering and classification. *Clin Psychol Sci*. 2015;3:378–99.
18. Bystritsky A, Nierenberg AA, Feusner JD, Rabinovich M. Computational non-linear dynamical psychiatry: a new methodological paradigm for diagnosis and course of illness. *J Psychiatr Res*. 2012;46:428–35.
19. Computational Psychiatry, CP13. 22–23 October 2013, Miami, Florida. <http://computationalpsychiatry.org/>. Geraadpleegd op 16 juli 2015.
20. Siddique H. World's first computational psychiatry centre opens in London. 2014. <http://www.theguardian.com/science/2014/apr/02/worlds-first-computational-psychiatry-centre-london>. Geraadpleegd op 16 juli 2015.
21. Max Planck UCL Centre for Computational Psychiatry and Ageing Research, homepage. <https://www.mps-ucl-centre.mpg.de/en>. Geraadpleegd op 16 juli 2015.
22. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: Dsm-5*, 5th edn. Arlington: American Psychiatric Publishing; 2013.
23. World Health Organization. *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*. Geneva: WHO; 1992.
24. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders, fourth edition*. Washington, DC: American Psychiatric Publishing; 2001.
25. Charney D, Barlow DH, Botteron K, Cohen JD, Goldman D, Gur R, et al. Neuroscience research agenda to guide development of a pathophysiologically based classification system. In: Kupfer DJ, First M, Regier DA, redactie. *A Res agenda DSM-V (Vol. 20)* Washington, DC: American Psychiatric Association; 2002. pp. 31–83.
26. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012;17:1174–9.
27. Forstmann BU, Wagemakers EJ, Eichele T, Brown S, Serences JT. Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? *Trends Cogn Sci*. 2011;6:272–9.
28. Den Ouden HEM, Kok P, Lange FP de. How prediction errors shape perception, attention, and motivation. *Front Psychol*. 2012;3:1–12.
29. Stephan KE, Baldeweg T, Friston KJ. Synaptic plasticity and dysfunction in schizophrenia. *Biol Psychiatry*. 2006;59:929–39.
30. Den Ouden HEM, Daw ND, Fernandez G, Elshout JA, Rijpkema M, Hoogman M, et al. Dissociable effects of dopamine and serotonin on reversal learning. *Neuron*. 2013;80:1090–100.
31. Maia TV. Reinforcement learning, conditioning, and the brain: successes and challenges. *Cogn Affect Behav Neurosci*. 2009;9:343–64.
32. Tolman EC. Cognitive maps in rats and men. *Psychol Rev*. 1948; 55:189–208.
33. Thorndike EL. *Animal intelligence: experimental studies*. New York: Macmillan; 1911.
34. Thorndike EL. *Animal intelligence: an experimental study of the associative processes in animals*. *Psychol Rev Monogr Suppl*. 1898;2:1–109.
35. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*. 2005;8:1704–11.
36. Huys QJM, Cruickshank A, Seriès P. Reward-based learning, model-based and model-free. In: Jaeger D, Jung R, redactie. *Encycl Comput Neurosci*. New York: Springer; 2014. pp. 1–10.
37. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011;69:1204–15.
38. Dolan RJ, Dayan P. Goals and habits in the brain. *Neuron*. 2013;80:312–25.
39. De Wit S, Corlett PR, Aitken MRF, Dickinson A, Fletcher PC. Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J Neurosci*. 2009;29:11330–8.
40. Tricomi E, Balleine BW, O'Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci*. 2009;29:2225–32.

41. Schwabe L, Wolf OT. Stress prompts habit behavior in humans. *J Neurosci*. 2009;29:7191–8.
42. Radenbach C, Reiter AMF, Engert V, Sjoerds Z, Villringer A, Heinze H-J, Schlagenhaut F. The interaction of acute and chronic stress impairs model-based behavioral control. *Psychoneuroendocrinology*. 2015;53:268–80.
43. Maier SU, Makwana AB, Hare TA. Acute stress impairs self-control in goal-directed choice by altering multiple functional connections within the brain's decision circuits. *Neuron*. 2015;87:621–31.
44. Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci*. 2005;8:1481–9.
45. Sjoerds Z, Wit S de, Van Den Brink W, Robbins TW, Beekman ATF, Penninx BWJH, Veltman DJ. Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Transl Psychiatry*. 2013;3:e337.
46. O'Tousa D, Grahame N. Habit formation: implications for alcoholism research. *Alcohol*. 2014;48:327–35.
47. Barker JM, Taylor JR. Habitual alcohol seeking: modeling the transition from casual drinking to addiction. *Neurosci Biobehav Rev*. 2014;47:281–94.
48. Sebold M, Deserno L, Nebe S, Schad DJ, Garbusow M, Hägele C, et al. Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology*. 2014;70:122–31.
49. Voon V, Derbyshire K, Rück C, Irvine MA, Worbe Y, Enander J, et al. Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry*. 2014;20:345–52.
50. Chen C, Takahashi T, Nakagawa S, Inoue T, Kusunagi I. Reinforcement learning in depression: a review of computational research. *Neurosci Biobehav Rev*. 2015;55:247–67.
51. Huys QJM, Daw ND, Dayan P. Depression: a decision-theoretic analysis. *Annu Rev Neurosci*. 2015;38:1–23.
52. Schlagenhaut F, Huys QJM, Deserno L, Rapp MA, Beck A, Heinze H-J, et al. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage*. 2014;89:171–80.
53. Gruner P, Anticevic A, Lee D, Pittenger C. Arbitration between action strategies in obsessive-compulsive disorder. *Neuroscientist*. 2015. doi:10.1177/1073858415568317.
54. Lee SW, Shimojo S, O'Doherty JP, Doherty JPO. Neural computations underlying arbitration between model-based and model-free learning. *Neuron*. 2014;81:687–99.
55. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;167:748–51.
56. Gläscher J, Daw ND, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010;66:585–95.
57. Gillan CM, Otto AR, Phelps EA, Daw ND. Model-based learning protects against forming habits. *Cogn Affect Behav Neurosci*. 2015;15:523–36.
58. Geurts DEM, Huys QJM, Ouden HEM den, Cools R. Serotonin and aversive Pavlovian control of instrumental behavior in humans. *J Neurosci*. 2013;33:18932–9.
59. Rutledge RB, Skandali N, Dayan P, Dolan RJ. Dopaminergic modulation of decision making and subjective well-being. *J Neurosci*. 2015;35:9811–22.
60. Kahnt T, Park SQ, Cohen MX, Beck A, Heinz A, Wrase J. Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. *J Cogn Neurosci*. 2009;21:1332–45.
61. Deserno L, Huys QJM, Boehme R, Buchert R, Heinze HJ, Grace AA, Dolan RJ, Heinz A, Schlagenhaut F. Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proc Natl Acad Sci USA*. 2015;5:1595–1600.
62. Schlagenhaut F, Rapp MA, Huys QJM, Beck A, Wüstenberg T, Deserno L, et al. Ventral striatal prediction error signaling is associated with dopamine synthesis capacity and fluid intelligence. *Hum Brain Mapp*. 2013;34:1490–9.
63. Schad DJ, Jünger E, Sebold M, Garbusow M, Bernhardt N, Javadi AH, et al. Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Front Psychol*. 2014;5:1450.
64. Deserno L, Wilbertz T, Reiter A, Horstmann A, Villringer A, Heinze H, Schlagenhaut F. Lateral prefrontal model-based signatures are reduced in healthy individuals with high trait impulsivity. *Transl Psychiatry*. 2015. (In Press).
65. Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci U S A*. 2013;110:20941–6.
66. Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci*. 2005;17:51–72.
67. Gillan CM, Morein-Zamir S, Urcelay GP, Sule A, Voon V, Apergis-Schoute AM, et al. Enhanced avoidance habits in obsessive-compulsive disorder. *Biol Psychiatry*. 2014;75:631–8.
68. Rolls ET, Loh M, Deco G, Winterer G. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nat Rev Neurosci*. 2008;9:696–709.

69. Chekroud AM. Unifying treatments for depression: an application of the free energy principle. *Front Psychol.* 2015;6:153.
70. Brodersen KH, Deserno L, Schlagenhaut F, Lin Z, Penny WD, Buhmann JM, Stephan KE. Dissecting psychiatric spectrum disorders by generative embedding. *NeuroImage Clin.* 2014;4:98–111.
71. Robbins TW, Gillan CM, Smith DG, Wit S de, Ersche KD. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn Sci.* 2012;16:81–91.
72. Daw ND. Trial-by-trial data analysis using computational models. In: Delgado MR, Phelps EA, Robbins TW. *Decision making, affect and learning, attention and performance XXIII.* New York: Oxford University Press; 2011. pp. 3–38.
73. Friston KJ, Harrison L, Penny W. Dynamic causal modelling. *Neuroimage.* 2003;19:1273–302.
74. Kay SR, Opler LA, Lindenmayer JP. The Positive and Negative Syndrome Scale (PANSS): rationale and standardisation. *Br J Psychiatry.* 1989;7:59–67.
75. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science.* 1997;275:1593–9.