

Neurobiologie wirtschaftlicher Entscheidungen

Unterlagen zum Vortrag im Rahmen des 1. Fachverbandskongresses am 28. und 29. September 2017 des Fachverbandes Personenberatung und Personenbetreuung in der Wirtschaftskammer Österreich.

Thomas Münte

Klinik für Neurologie

Universität zu Lübeck

Thomas.muente@neuro.uni-luebeck.de

1. Was ist Motivation

Was ist Motivation? Die Definition von Wikipedia hält fest: „Motivation bedeutet „Triebkraft“ und bezeichnet das auf emotionaler und neuronaler Aktivität (Aktivierung) beruhende Streben des Menschen nach Zielen oder wünschenswerten Zielobjekten. Die Gesamtheit der Beweggründe (Motive), die zur Handlungsbereitschaft führen, nennt man Motivation. Die Umsetzung von Motiven in Handlungen nennt man Volition oder Umsetzungskompetenz.“

An diesem Wochenende beschäftigen wir uns speziell mit den neuronalen Grundlagen von Motivation: Dies impliziert zwei zentrale Aspekte: Wie werden Ziele umgesetzt in Wahrnehmen, Denken, Fühlen und Handeln, und warum werden bestimmte Ziele überhaupt verfolgt? Der erste Aspekt ist die kognitive Kontrolle, also die intentionalen Steuerungsprozesse des Verhaltens; der zweite Aspekt ist die Motivation, also die Gründe der Intentionen. Evolutionsbiologie und verschiedene ökonomische Theorien geben in Bezug auf die Motivation eine gemeinsame Antwort: Sie definieren das Streben nach größtmöglichem Nutzen als allgemeinste Motivation des menschlichen Verhaltens. Nutzen wiederum kann vom Organismus als Belohnung erlebt werden, und hier setzt die Neurobiologie an. Denn die vielleicht allgemeinste experimentelle Beobachtung über das Verhalten des Menschen (und vieler anderer Organismen) ist das Streben nach Belohnung und das Vermeiden von Bestrafung. Eine solche allgemeine Beobachtung impliziert, dass der Organismus in der Lage ist, die Varianten von Belohnung einheitlich, also mit einer gemeinsamen Kodierung, zu bewerten. Belohnung allerdings hat viele Dimensionen: Es gibt zahlreiche Formen von Belohnung (Befriedigung elementarer biologischer Bedürfnisse, soziale Anerkennung, Geld, abstrakter Erkenntnisgewinn), es gibt unterschiedliche Zeitskalen von Belohnung (kurzfristige und langfristige Belohnung), es gibt verschiedene soziale Aspekte von Belohnung (individuelle und oder auf das Gemeinwohl bezogene Belohnung), und die Neurobiologie ist daher mit folgender Fragen konfrontiert: Wie kann der Organismus so rasch und effizient, wie dies meistens der Fall ist, die verschiedenen Formen und Aspekte von Belohnung in Bezug auf die Ressourcen der Umgebung bewerten und schließlich sein Verhalten, d.i. sein Wahrnehmen, Denken und Handeln in diesem Sinne optimal steuern? Und umgekehrt: Wie kommen

Situationen zustande, in denen der Organismus trotz klarer Belohnungshierarchie sich aus scheinbar irrationalen Gründen suboptimal verhält?

Ein wesentliches Beispiel für solche „Entkopplungen“ von Verhalten von den Belohnungskontingenzen findet sich in Form von Gewohnheiten: Ich weiß, dass Rauchen schlecht für mich ist. Eigentlich schmeckt mir die Zigarette auch gar nicht. Trotzdem muss ich nach dem Essen unbedingt rauchen. Die Neurobiologie hat besondere Paradigmen für die Untersuchung von Gewohnheiten entwickelt, die zeigen, dass es hierbei zu einem „Shift“ in der neuronalen Verschaltung kommt, was dazu führt, dass das Verhalten unabhängig von den Belohnungskontingenzen führt.

Scheinbar irrationales Verhalten findet sich auch bei der „altruistischen Bestrafung“. Man verzichtet auf eine mögliche Belohnung, um ein jemanden zu „erziehen“ oder zu bestrafen, mit dem man (vermutlich) nie wieder etwas zu tun hat. Warum gehen wir solche Kosten ein? Welcher evolutionärer Vorteil steckt dahinter? Gibt es aber vielleicht Menschen, die diese altruistische Bestrafung nicht vollziehen? Was sind die neuronalen Ursachen dafür und sind diese Menschen vielleicht erfolgreicher als andere, weil sie sich die Kosten der Bestrafung ersparen?

Eine besondere Variante des motivgesteuerten Verhaltens ist das Reinforcement Learning (RL), das computertheoretisch formuliert werden kann und enge Verbindungen zur instrumentellen/ operanten Konditionierung aufweist. Fast alle Modelle von RL teilen bestimmte Basiseigenschaften. Danach berechnet das lernende System – in unserem Falle der Mensch – während der Durchführung eines Experimentes (aber natürlich auch in der „freien Wildbahn“) iterativ eine Schätzung des mittleren Wertes aller möglichen Aktionen, also die mittlere Belohnung, die mit eine jede Aktion assoziiert ist. Diese Schätzung dient zugleich als Prädiktion für die Belohnung (bzw. Bestrafung), die eine Aktion erzielen wird. Nach jeder abgeschlossenen Aktion wird die tatsächlich erhaltene Belohnung verglichen mit der erwarteten Belohnung. Die aus diesem Vergleich resultierende Differenz ist der sog. Prädiktionsfehler, der positiv ist, wenn die tatsächliche Belohnung größer ist als die erwartete. Dieser Fehler wird verwendet, um die Belohnungsprädiktion zu präzisieren und um das Verhalten in Bezug auf Belohnungsmaximierung anzupassen. Ein populäres Modell von RL ist das sog. Zeitdifferenz (Time difference, TD) Actor-Critic Modell, das (in verschiedenen Varianten) zwischen dem Actor- und dem Critic Signal unterscheidet. Das Critic-Signal codiert den Prädiktionsfehler; das Actor-Signal

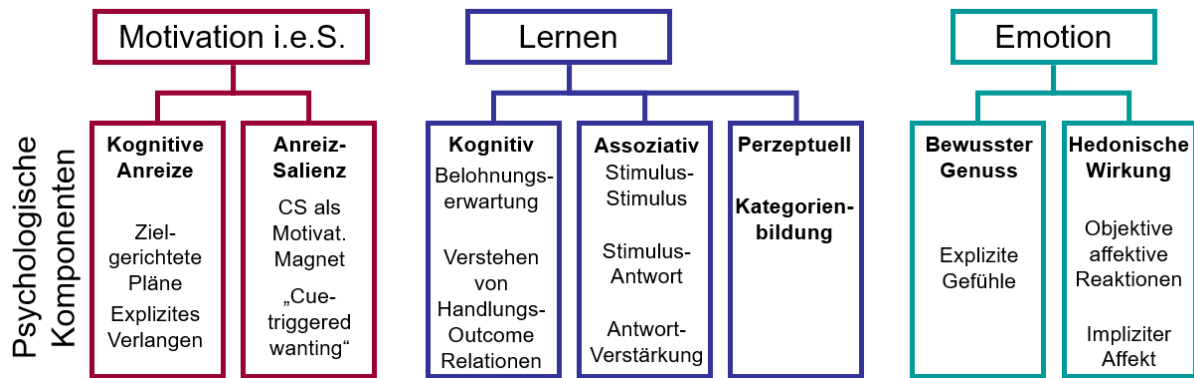
verwendet diesen Fehler für die adaptive Auswahl und Sequenzierung von Aktionen. Der TD Prädiktionsfehler schließt Informationen über zeitverzögerte Belohnungen als Folge der Aktionen der Vergangenheit ein. Daraus folgen spezifische Vorhersagen über den zeitlichen Verlauf einer Fehlerantwort:

Wenn in einem Konditionierungsparadigma eine Assoziation zwischen einem Stimulus und einer Belohnung gelernt werden soll, tritt vor dem Lernen dieser Assoziation ein positiver Fehler nur auf, wenn die Belohnung selbst präsentiert wird; während des Lernens verlagert sich die Antwort aber hin zum assoziierten Ereignis. Ferner löst ein Auslassen der Belohnung eine negative Fehlerantwort zum Zeitpunkt der erwarteten Belohnung aus. Wie der Prädiktionsfehler für die Selektion und das Erlernen von Aktionen verwendet wird, ist allerdings beim klassischen Lernen (d.h., die Belohnung ist unabhängig vom Verhalten) und beim instrumentellen Lernen (d.h., die Belohnung hängt ab vom Verhalten) verschieden. Ein wichtiger Unterschied beispielsweise betrifft Änderungen des motivationalen Zustandes und deren Konsequenzen für das Handeln.

Wenn Belohnungen so wichtig für unser Handeln sind stellt sich weitergehend die Frage, ob es nicht Sinn machen könnte, wenn der Erhalt oder Nichterhalt von Belohnungen bestimmt, ob wir glücklich sind oder nicht. Wenn dem so ist, wäre allerdings zu fragen, welches ein optimaler Algorithmus wäre, die Beziehung zwischen Belohnung und Glück zu determinieren. Ein Mensch der wunschlos glücklich ist, ist in Gefahr, den Anschluss zu verlieren, sich auszuruhen und von anderen überholt zu werden. Es sind in der letzten Zeit einige interessante Untersuchungen durchgeführt worden, die zeigen, dass offensichtlich in unserem Gehirn eine komplexe Beziehung von Belohnung und Glücklichkeit realisiert ist, die uns auch erklären kann, warum Geld alleine nicht glücklich macht.

2. Belohnung – Bedeutung für den Organismus

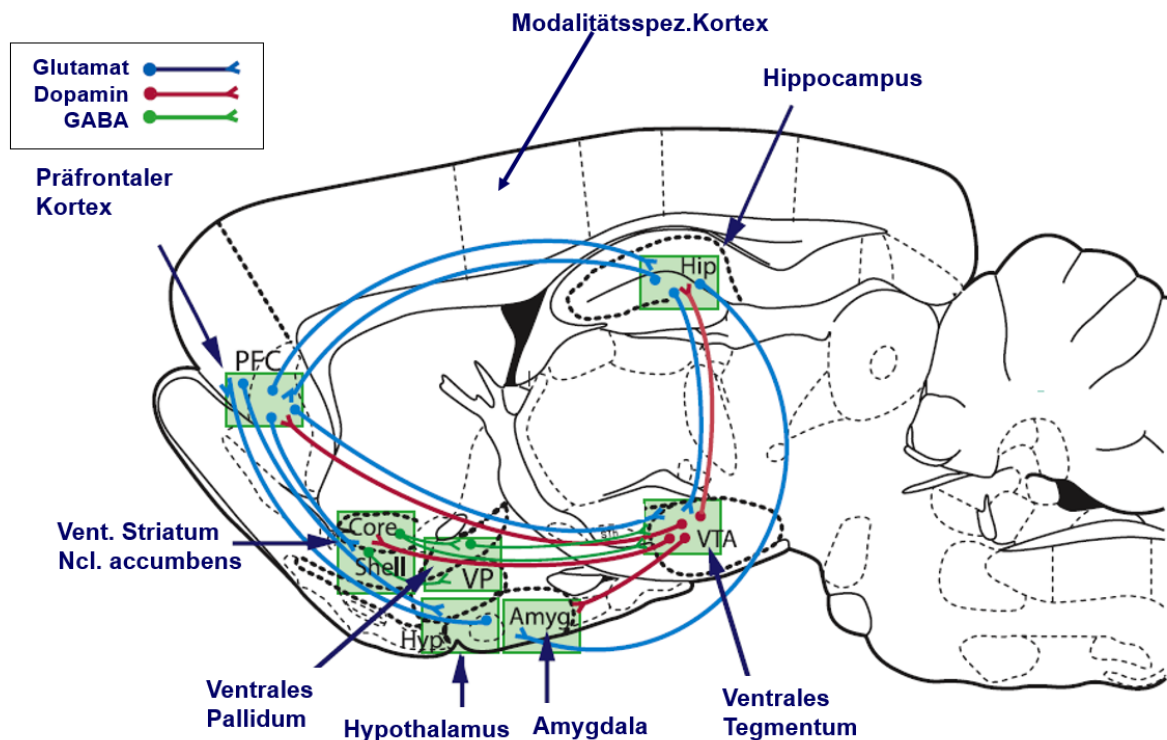
Motiviertes Verhalten ist für den Organismus überlebenswichtig. Man könnte dies auch als Trivialität betrachten, da man jegliches Verhalten als motiviert (also in irgendeiner Weise auf einen Zweck gerichtet betrachten könnte. Während meiner Zeit an der Universität Magdeburg leitete ich den dortigen DFG-Sonderforschungsbereich „Neurobiologie motivierten Verhaltens. Für diesen entwickelten wir folgendes Schema, welches verdeutlichen soll, dass motiviertes Verhalten verschiedene Aspekte umfasst.



Wir werden dieses Schema im Einzelnen durchgehen. Wichtig erscheint mir zunächst, dass Anreize / Belohnungen essentiell sind, und zwar zum einen in der Verhaltenssteuerung, was auf der linken (roten) Seite der Abbildung dargestellt ist. Hier geht es um das Erlangen von Belohnungen – ein Organismus (auch der Mensch) strebt also zu Belohnungen und vermeidet Bestrafungen. Ein essentieller Bereich für motiviertes Verhalten ist das Lernen. Hierzu haben Sie schon von Frau Dicke gehört. Wesentlich erscheint mir, dass Lernen in der Regel durch Anreize funktioniert. Wenn man einem Hund etwas beibringen will, kann man dies durch Belohnungen erreichen. Durch motivationale Prozesse sind verschiedene Arten von Lernen vermittelbar: Neben den Klassikern der klassischen und operanten Konditionierung ist in letzter Zeit auch gezeigt worden, dass perzeptuelles Lernen einer Steuerung durch motivationelle Prozesse unterliegt.

3. Belohnung – neurale Repräsentation

Welche Strukturen verarbeiten Belohnungen und sind somit maßgeblich für motiviertes Verhalten. Es hat sich aus zahlreichen Tierversuchen und in der Folge aufgrund von Untersuchungen beim Menschen mit funktionell bildgebenden Verfahren eine Reihe von Kernregionen herausgestellt, die ich in der folgenden Graphik für die Ratte einmal dargestellt habe:

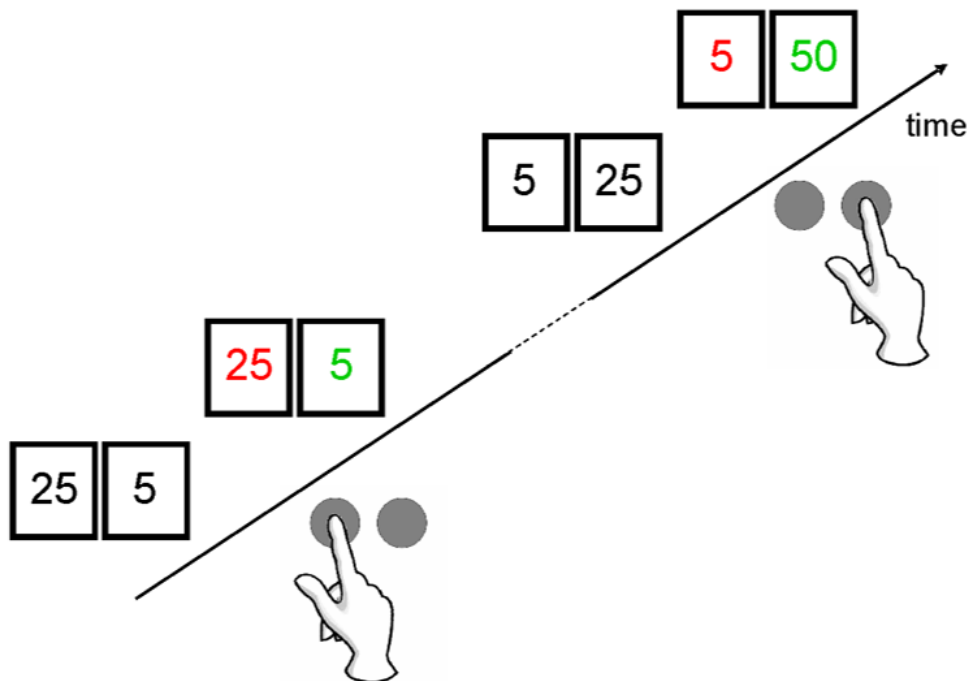


Ich möchte an dieser Stelle nicht auf Details eingehen, aber darauf hinweisen, dass sich hier zwei fundamentale Bauprinzipien unseres Gehirns ableiten lassen. Sie sehen, dass die Verbindungen zwischen den einzelnen Hirnstrukturen mehrere parallele Schleifensysteme bilden. Dies ist ein Bauprinzip, das auch in der Steuerung der Motorik eine Rolle spielt (wenn es in den Schleifen eine Fehlregulation gibt, resultiert eine Parkinsonsche Erkrankung) und dass sich hier auch findet. Ähnlich wie bei der Parkinson-Erkrankung kann eine Aktivitätsveränderung in diesen Schleifen zu einer Veränderung im motivierten Verhalten führen.

Ein zweites Bauprinzip, das Sie aus dieser Graphik ableiten können, ist die Wechselwirkung von kortikalen (in der Hirnrinde gelegenen) Strukturen und subkortikalen Strukturen. Keiner von uns weiß, ob die Ratte so etwas wie Bewusstsein hat. Für uns Menschen nehmen wir aber an, dass das Bewusstsein im Wesentlichen durch die in der Hirnrinde lokalisierten Neurone entsteht. Bewusste Entscheidungen und Reflektionen über Sinn und Zweck dieser Entscheidungen finden in der Hirnrinde statt. Die Hirnrinde wird aber in ihrer Aktivität kontrolliert und moduliert durch die tief im Inneren des Gehirns liegenden subkortikalen Strukturen.

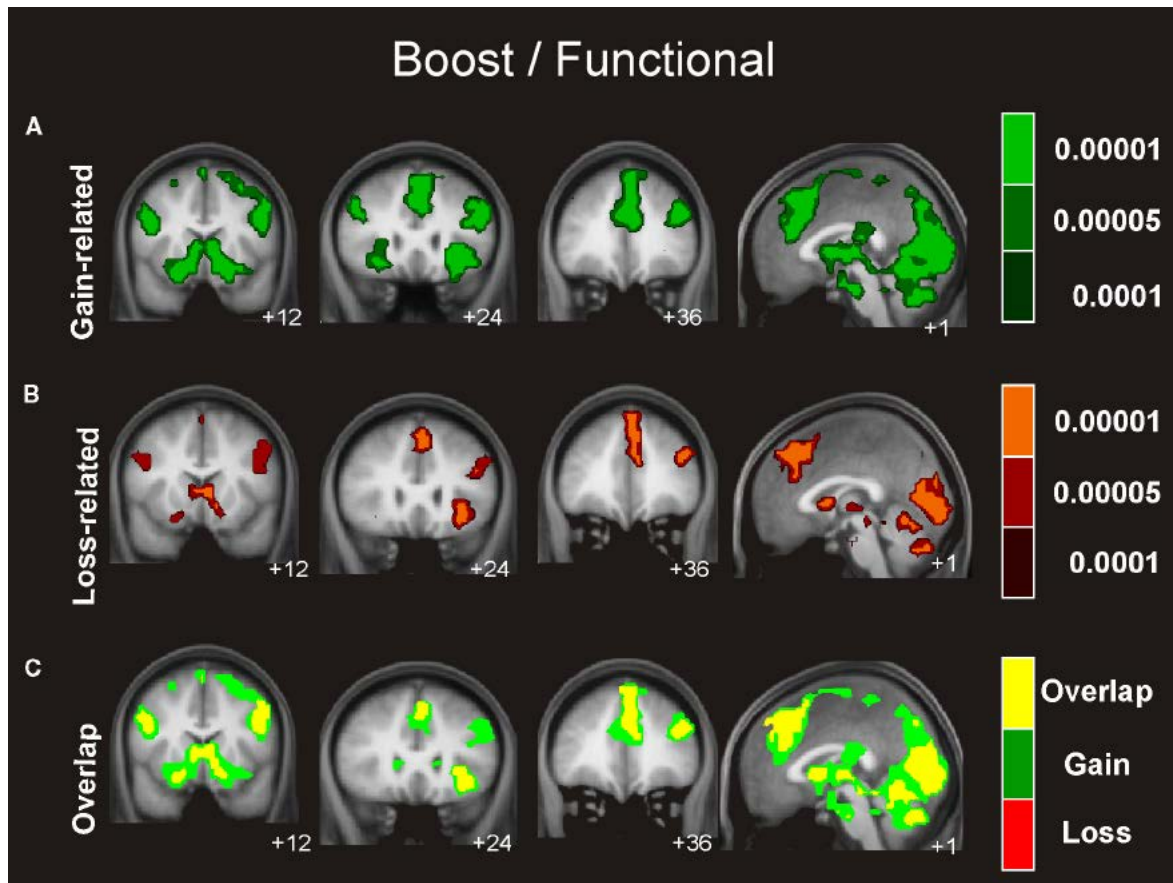
Schauen wir einmal an, wie dies beim Menschen aussieht. Wir, d.h. meine ehemaligen Mitarbeiter Estela Camara, Antoni Rodriguez-Fornells, Zheng Ye und ich, wollten vor einigen Jahren die Hirnstrukturen, die für die Verarbeitung von positiven (in diesem

Fall der Gewinn von Geld) und negativen (in diesem Fall der Verlust von Geld) Ereignissen zuständig sind. Hierzu führten wir ein einfaches Experiment durch, welches in der nächsten Abbildung illustriert ist:



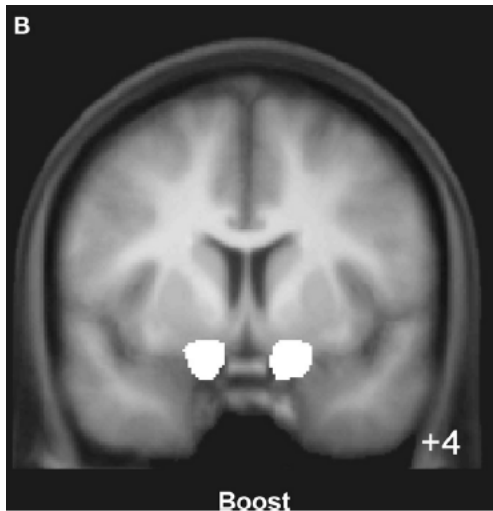
Die Probanden lagen in einem MRT-Scanner und sahen zunächst zwei schwarze Zahlen (eine höhere und eine niedrigere) und mussten durch Drücken eines Knopfes sich für eine der beiden Zahlen entscheiden. Im Beispiel oben entschied sich der Proband für die 25, also die höhere Zahl. Eine Sekunde nach der Entscheidung sprang die Farbe der Zahlen um: Rot bedeutete einen Verlust, grün einen Gewinn. Im ersten Versuch in der obigen Abbildung wurde die ausgewählte Zahl 25 rot; der Proband hatte nun 25 Eurocent verloren. Im zweiten Versuch der obigen Abbildung ist eine Variante dargestellt, die wir in unser Experiment einführten: Der Proband wählte in diesem Falle wieder die höhere Zahl 25 aus. Die Zahl wechselte jedoch nicht nur die Farbe sondern verdoppelte sich auch. Dies war ein seltenes Ereignis (10 % Wahrscheinlichkeit) und sollte die aus dem Tierexperiment bekannte Tatsache widerspiegeln, dass vor allem unerwartete Belohnungen zu einer Antwort des Belohnungssystems führen. Diese Versuchsdurchläufe mit vergrößertem Gewinn nannten wir „Boost“-Trials.

Wir untersuchten gesunde junge Probanden mit der funktionellen Kernspintomographie. Ich zeige Ihnen zunächst nur die Ergebnisse für die Boost-Trials (sowohl unerwartet hohe Gewinne als auch unerwartet hohe Verluste):

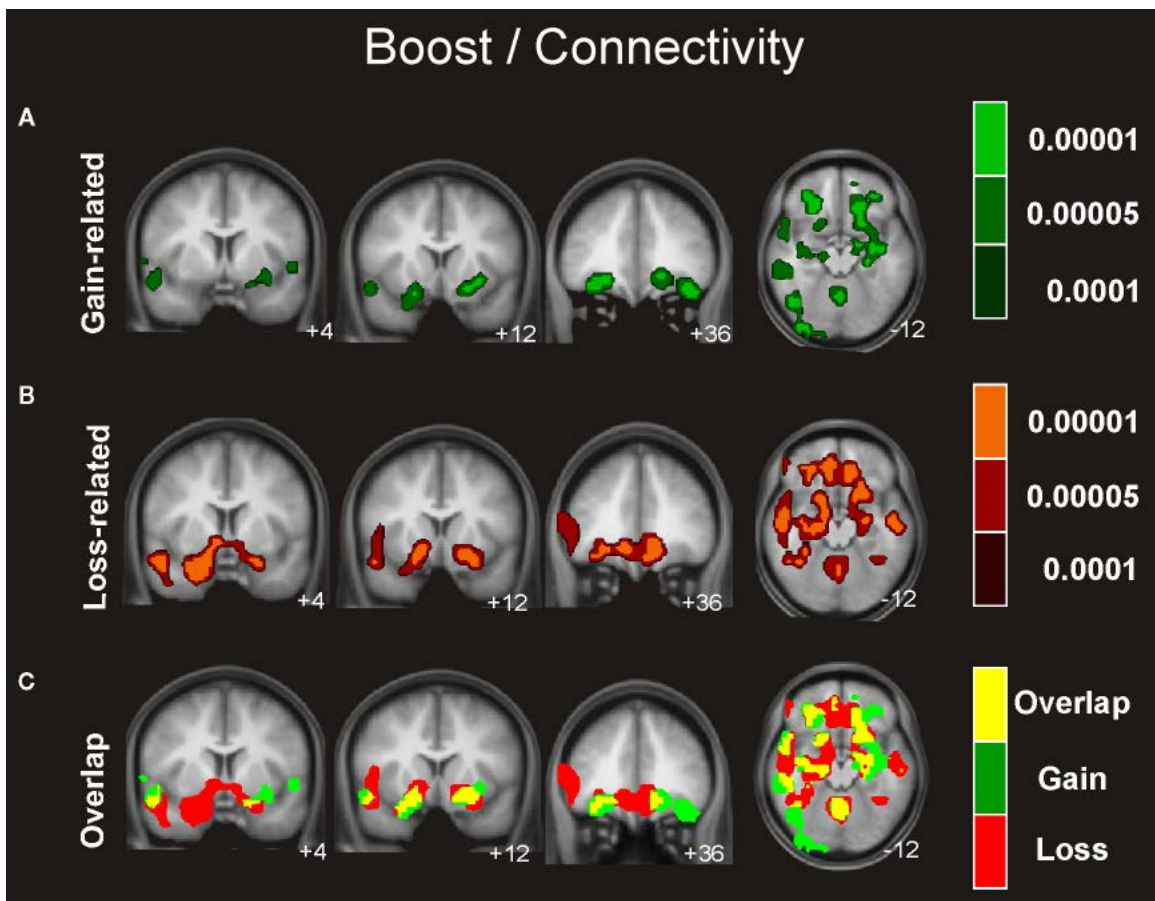


Sie sehen in der oberen Zeile Aktivierungen für die unerwarteten Gewinne, in der mittleren Zeile die Aktivierungen für die unerwartet hohen Verluste. In der unteren Zeile sind gelb die gemeinsam für Gewinne und Verluste aktivierten Areale in Gelb dargestellt. Es gibt also deutliche Überlappungen zwischen den für Gewinne und Verluste. Dies stellte sich in ähnlicher Weise auch für die normalen Gewinne und Verluste dar. In Übereinstimmung mit unseren Erwartungen fand sich eine wesentliche Region für Gewinne (weniger für Verluste) im ventralen Striatum, speziell im Nucleus accumbens.

Als nächsten Schritt der Analyse führten wir eine so genannte Konnektivätsuntersuchung durch. Hierzu platzierten wir im Nucleus accumbens eine „Region of Interest“ und korrelierten die Aktivität jeder einzelnen kleinen Volumeneinheit des Gehirns (wir sprechen von Voxeln) mit der Aktivität in dieser Region of Interest separat für die Verarbeitung von Gewinnen und Verlusten. Auf diese Weise kann man herausfinden, welche neuronalen Netzwerke mit dem Nucleus accumbens in einer aufgabenspezifischen Art und Weise zusammenhängen. Zunächst zeige ich Ihnen die Region of Interest im Nucleus accumbens:

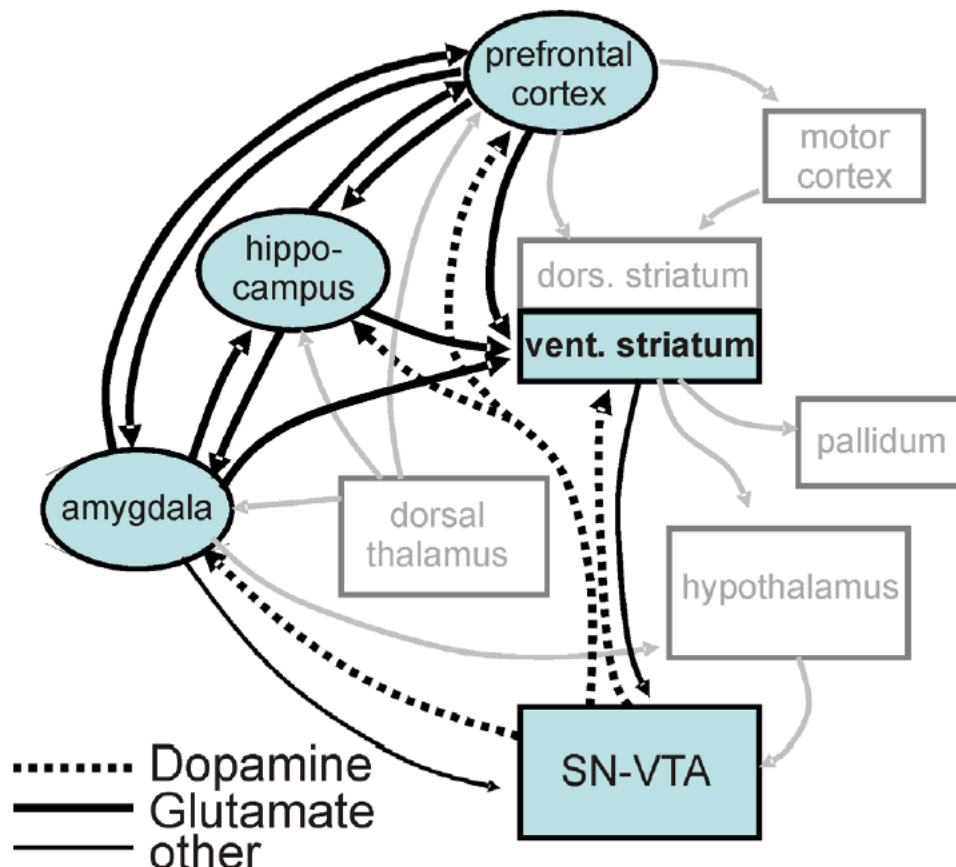


Das Ergebnis der Konnektivitätsanalyse sehen Sie in der nächsten Abbildung:



Die Konnektivitätsanalyse zeigt ein anderes Bild als die Standardanalyse. Es wird klar, dass insbesondere Verluste zu einer weitgespannten Konnektivität des Nucleus accumbens mit andere Hinstrukturen führen. Sehr wichtig ist hier der Mandelkern („Amygdala“), der Ihnen wahrscheinlich schon von Herrn Büchel nähergebracht worden ist. Dieser Mandelkern ist für die Verarbeitung von negativen Informationen zuständig. Wir haben die von uns gefundenen Aktivierungs- und

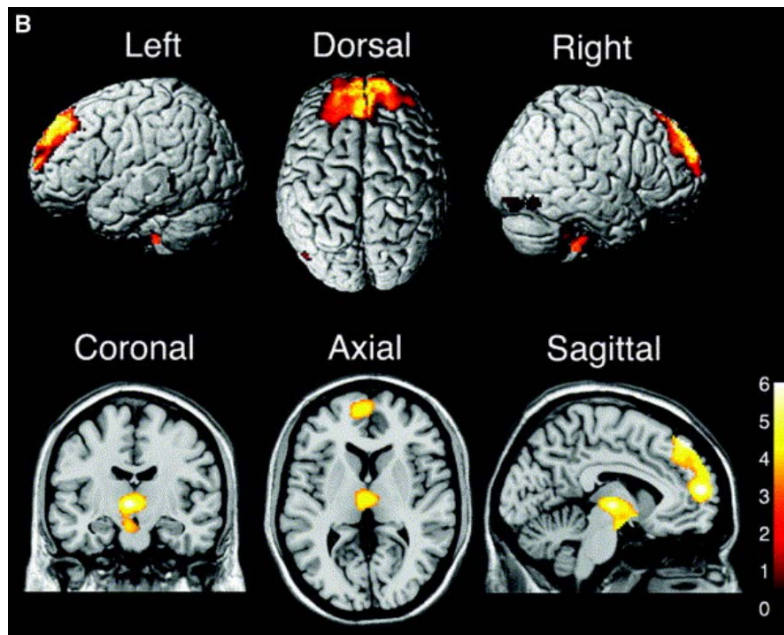
Konnektivitätsergebnisse in einem Schaltdiagramm dargestellt, welches Sie in der nächsten Abbildung sehen. Dabei sind die bläulich dargestellten Regionen diejenigen, die wir in unseren Versuchen gefunden haben. Die anderen Regionen sind aus tierexperimentellen Untersuchungen bekannt und dürften beim Menschen ebenfalls aktiv sein:



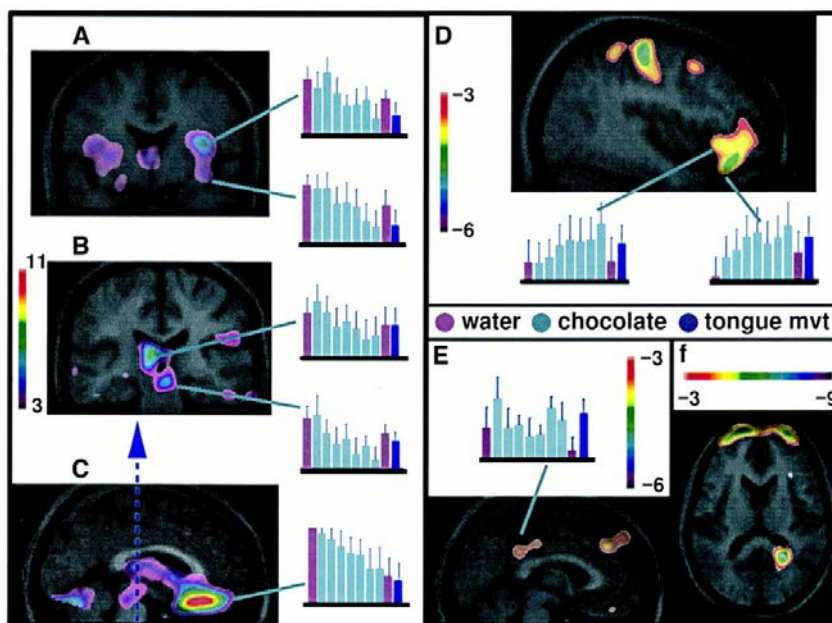
Denken wir noch einmal über Belohnungen nach, die unser Verhalten steuern: Neben den von uns (und vielen anderen Arbeitsgruppen) verwendeten Geldgewinnen und –verlusten gibt es eine Vielzahl von anderen Belohnungen. Als „primary rewards“ (primäre Belohnungen) können elementare Stimuli wie Nahrungsmittel, Getränke, Sex oder ähnliche betrachtet werden, wohingegen Geldgewinne eher als „secondary rewards“ (sekundäre Belohnungen) bezeichnet werden. Schließlich gibt es auch noch tertiäre Belohnungen, z.B. in Form von Lob oder anderen sozialen Reizen (z.B. ein lächelndes Gesicht).

Wie sieht es mit der Repräsentation von diesen verschiedenen Formen von Belohnungen aus?

In der nachfolgenden Abbildung aus einer Publikation von Killgore und Mitarbeitern (NeuroImage 2003) ist die Reaktion des Gehirns auf primäre Belohnungen in Form von Nahrungsmitteln (Bilder von Nahrungsmitteln, um genau zu sein) gezeigt:



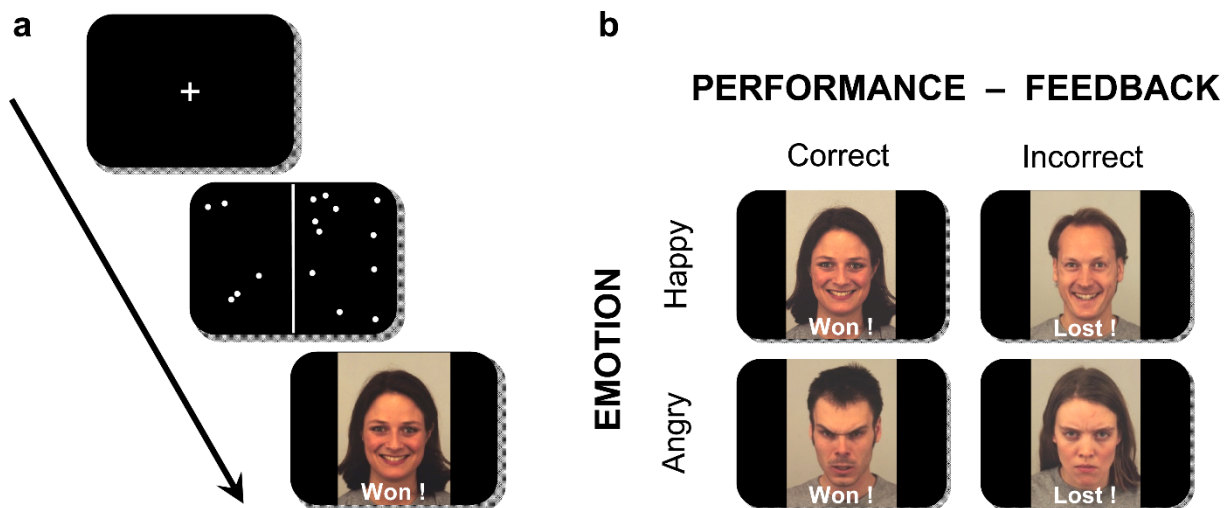
Es finden sich hier sehr ähnliche Hirnregionen, wie wir sie auch in unseren Untersuchungen zu sekundären Belohnungen gefunden haben. Ein anderes Beispiel ist ein Klassiker: Die Antwort des Gehirns auf Schokolade, die von der Arbeitsgruppe um Zatorre in Montreal untersucht wurde.



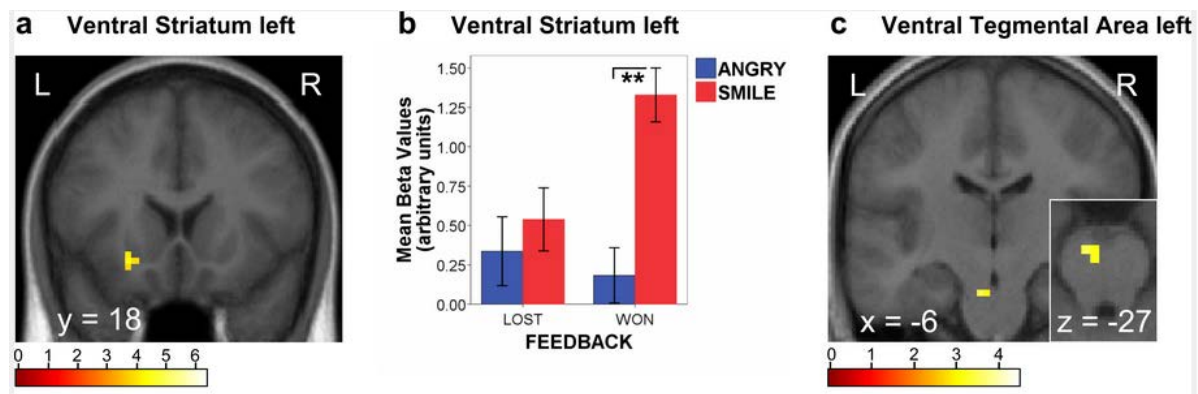
Die Abbildung wirkt etwas gewöhnungsbedürftig. Es handelt sich nämlich um eine relativ alte Studie, die statt der funktionellen Kernspintomographie die PET-Methode,

bei der dem Probanden radioaktiv markiertes Wasser gespritzt wird, um Blutflussunterschiede darstellen zu können. Das Isotop (Sauerstoff) zerfällt sehr rasch (Halbwertszeit 122 Sekunden), so dass man hier mehrfach nachgespritzt hat, während die Probanden mehr und mehr Schokolade zu sich nahmen, so dass diese am Ende nicht mehr belohnend war (die Probanden hatten sich überfressen). In der Abbildung B sehen Sie wieder das ventrale Striatum als Belohnungszentrum.

Wie steht es jetzt mit tertiären Belohnungen? Was passiert, wenn Sie jemandem ein Lächeln schenken? In einer Untersuchung der Universität Genf mussten Probanden entscheiden, ob die linke oder rechte Seite eines Stimulus mehr Punkte enthielt. Sie erhielten Feedback über die Richtigkeit ihrer Entscheidungen in Form von Gesichtern:



Die Probanden wurden darüber hinaus bezüglich ihrer Sozialfunktionen analysiert (genauer gesagt: Der Attachment Style der Probanden wurde erhoben). Es zeigten sich auf das Feedback robuste Aktivierungen von belohnungsrelevanten Arealen des Gehirns, nämlich der ventralen tegmentalen Area und des ventralen Striatums, wie Sie in der nächsten Abbildung sehen können:



In der Tat fand sich die Antwort im ventralen Striatum deutlich durch die Tatsache moduliert, dass zusätzlich zum verbalen Feedback die Exposition gegenüber dem lächelnden Gesicht zu einer erheblichen Aktivierung des Ventralen Striatums führte.

Wir können also als Zwischenergebnis festhalten, dass wir vergleichbare Aktivierungen des ventralen Striatums / Nucleus accumbens für primäre, sekundäre und tertiäre Belohnungen erfahren.

Es stellt sich hieran anschließend die Frage, wie der Wert von Dingen berechnet und repräsentiert wird? Oder anders gesagt: Wie vergleiche ich den Belohnungswert eines Lächelns mit dem Belohnungswert von Schokolade oder dem Belohnungswert eines 5-Euro Scheins? In unserer täglichen Umwelt müssen wir ständig Äpfel und Birnen vergleichen. Soll ich ein Sabbatjahr machen oder lieber Geld verdienen? Soll ich diese Schokolade kaufen oder das Geld lieber sparen? Um Entscheidungen zu steuern, müssen Belohnungen unterschiedlicher Art intern im Gehirn in ein vergleichbares Währungssystem umgesetzt werden. In vielfältigen Untersuchungen beim Menschen und am Tiermodell stellt sich heraus, dass wir offensichtlich frontale Hirnareale für die Bewertung benutzen, und zwar auf der Basis der berechneten Belohnungswerte. Der laterale orbitofrontale Cortex (IOFC) spielt eine zentrale Rolle beim Lernen der mit unterschiedlichen Gütern assoziierten Werte. Läsionen im IOFC führen dazu, dass man Stimuli keine präzisen Werte mehr beimessen kann. Die Repräsentationen von Werten in diesem Hirnareal erlauben die Steuerung von Entscheidungen. Interessanterweise (und notwendigerweise) zeigen die Neuronen in dieser Region eine so genannte Range-Adaptation. Dies bedeutet, dass die gleichen Neurone einmal die Entscheidung zwischen zwei relativ kleinen Belohnungen, die sich leicht in ihrer Größe unterscheiden, vermitteln und ein andermal die Entscheidung zwischen zwei großen Belohnungen. Das Problem dieser adaptiven Skalierung ist nicht trivial. Man kann dies aber bei seinen eigenen Entscheidungen auch beobachten: Einmal ist man nicht bereit 50 Cent auszugeben, weil man den Preis nicht angemessen findet, ein andermal – z.B. im Rahmen eines Eigenheimbaus – jongliert man mit Zehntausenden von Euro.

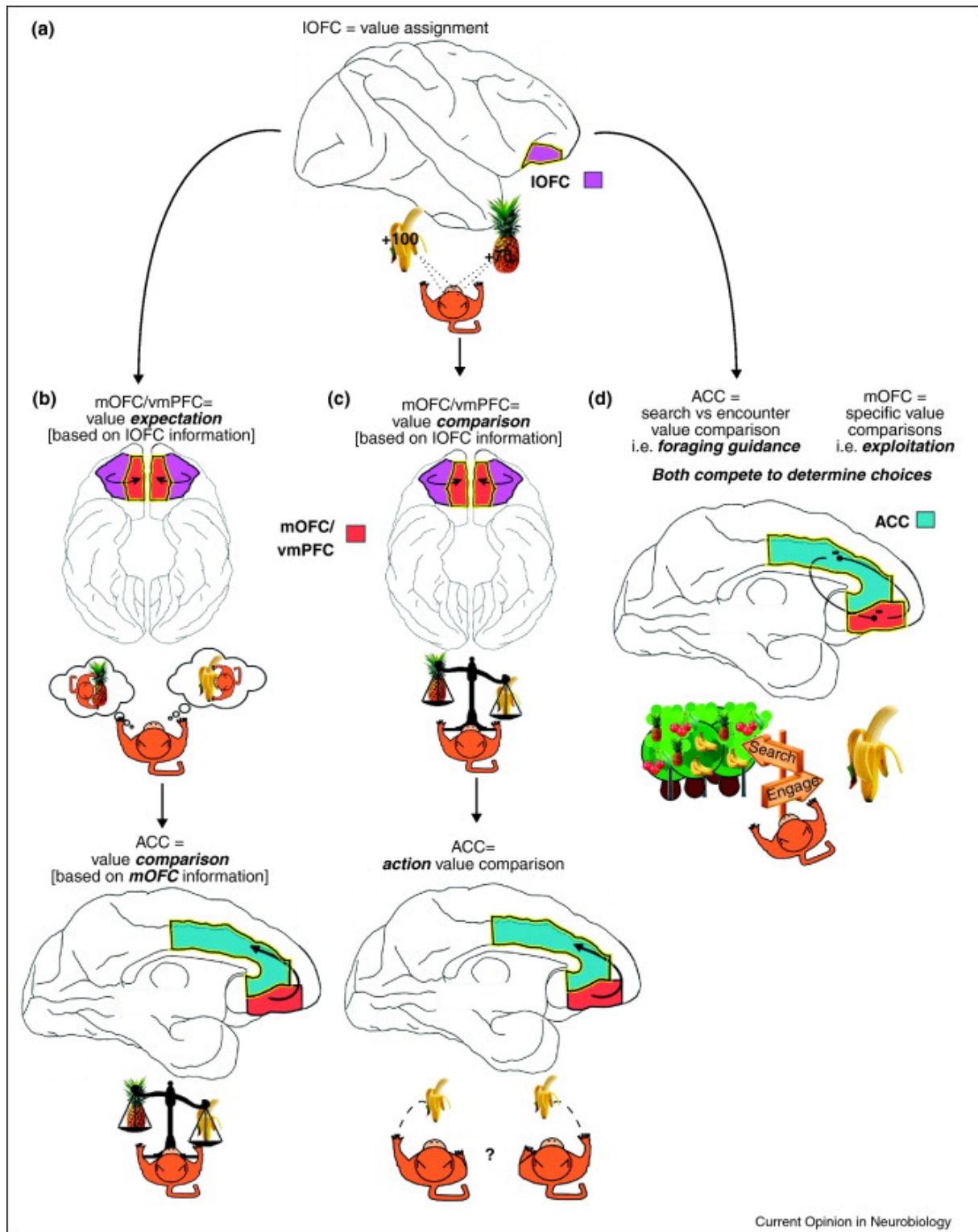
Der IOFC unterstützt die Aktualisierung von Wertrepräsentation auf der Basis von negativen und positiven Verhaltensergebnissen. Darüber hinaus ist der IOFC mit der Verbindung von Stimuli mit ihrem skalaren Belohnungswert befasst. Ferner kodiert er

Informationen über den Typ der Belohnung, der mit einer bestimmten Verhaltenswahl assoziiert sein wird.

Andere frontale Hirnstrukturen wie der ventromediale präfrontale Cortex und der mediale orbitofrontale Cortex (vmPFC/mOFC) sind eher mit der Umsetzung dieser Belohnungsrepräsentation in Verhalten befasst. Die Neuronen im vmPFC/mOFC kodieren zwar den Belohnungswert aber unabhängig von den damit assoziierten Stimuli (also eher den abstrakten Wert). In funktionellen Bildgebungsstudien ist die Aktivität in der vmPFC/mOFC Region proportional zu den Belohnungserwartungen zum Zeitpunkt der Entscheidung.

Es ist postuliert worden, dass die Wertrepräsentationen dann in einen Vergleichsprozess einmünden, der im dorsalen anterioren cingulären Cortex (ACC) bzw. im daran angrenzenden dorsomedialen präfrontalen Cortex (dmPFC) stattfindet. Somit werden die Werte zunächst berechnet und repräsentiert (vmPFC/mOFC) und dann in einen Vergleichsprozess eingefüttert (ACC/dmPFC).

Ein Grund für die Annahme, dass Wertvergleiche im ACC/dmPFC stattfinden, sind vielfältige Befunde, die zeigen, dass in dieser Region prominente Wert-Differenz-Signale messbar sind. Das ACC/dmPFC BOLD Signal nimmt zu, wenn die Differenz zwischen Werten potentieller Verhaltensoptionen abnimmt. Intuitiv könnte man sagen, dass es umso schwerer wird, zwei Werte miteinander zu vergleichen und sich zu entscheiden, je näher die Werte zusammenliegen. Somit bestimmt der Output des ACC/dmPFC Vergleichssystem die Aktivität des motorischen Systems und damit die Verhaltensantwort, die am Ende des Vergleichsprozesses steht. Diese Verhaltensantwort kann als das Endprodukt einer Serie von Prozessen der Wertrepräsentation und des Wertvergleichs angesehen werden, die im frontalen Cortex stattfinden.



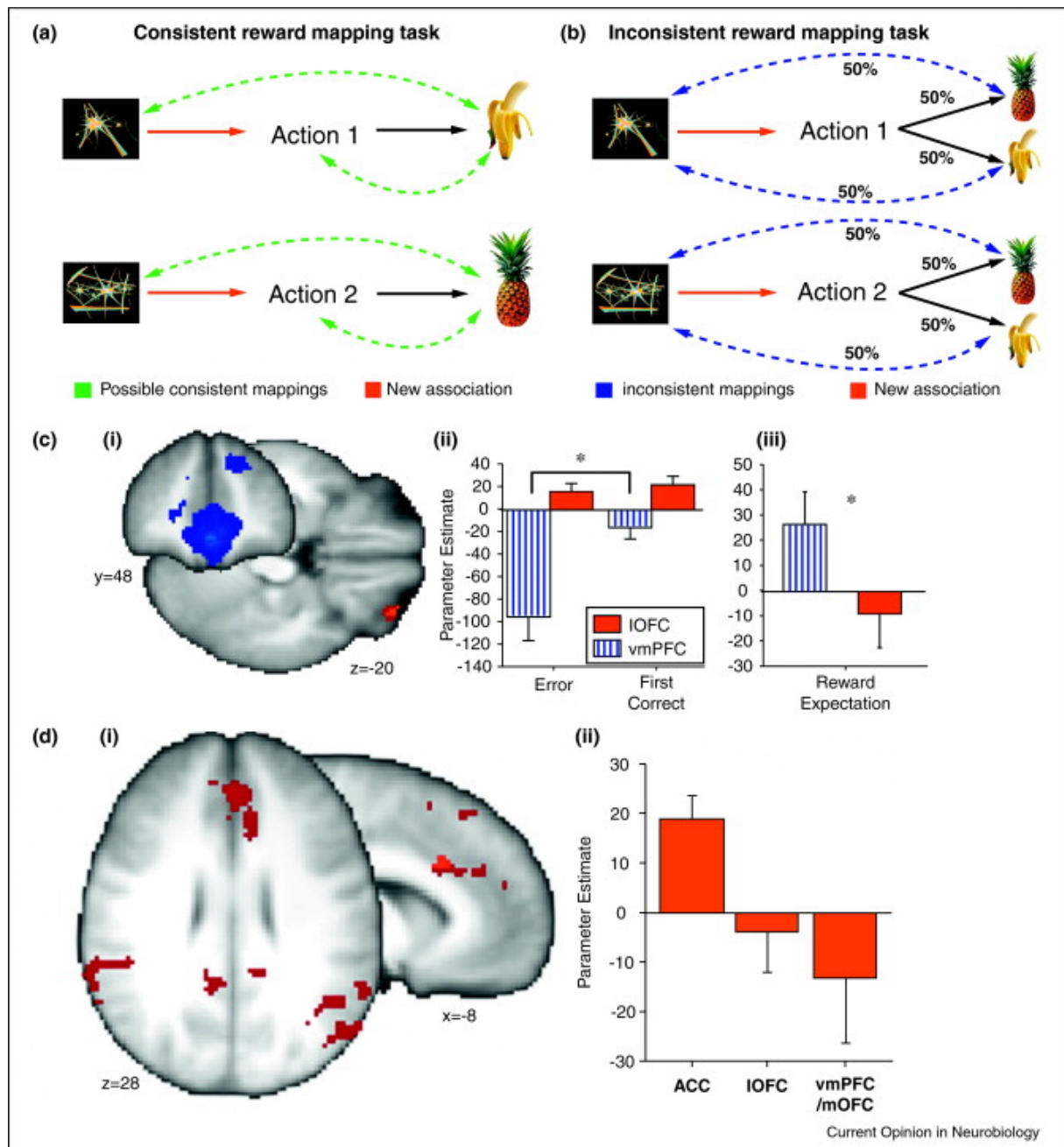
Die obenstehende Abbildung skizziert die Repräsentation von Werten in verschiedenen Teilen des frontalen Cortex. Werte und Informationen über den Belohnungstyp werden verschiedenen Optionen im IOFC zugeordnet (a). Sodann gibt es verschiedene Möglichkeiten, wie Entscheidungsmechanismen im frontalen Cortex stattfinden können. Belohnungserwartungen werden im vmPFC/mOFC repräsentiert und der eigentliche Wertvergleich erfolgt im ACC (b). Alternativ werden von einigen

Forschern unterschiedliche Mechanismen angenommen: Zunächst werden Entscheidungen getroffen, welche Belohnungen in den Fokus der Aufmerksamkeit gezogen werden sollen und damit verhaltensrelevant werden. Sodann werden in einem zweiten Schritt Entscheidungen über die Handlungen getroffen, die dazu führen sollen, die Belohnung zu erlangen (c). Die ganz rechts dargestellte Alternative besagt schließlich, dass es unterschiedliche Systeme für das Suchen von Belohnungen („foraging“) und das Entscheiden über Handlungen bezüglich der Belohnungen („exploitation“) gibt (d). Diese letzte Variante sieht vor, dass der ACC das Suchverhalten koordiniert, während der vmPFC, der nur bei Primaten eine nennenswerte Ausdehnung hat, besonders wichtig wird, wenn Entscheidungen gefällt werden müssen.

In der nächsten Abbildung geht es um die unterschiedlichen Rollen von IOFC, vmPFC/mOFC und ACC beim Lernen der Repräsentationen (und deren Modifikation) von spezifischen Belohnungen, die mit bestimmten Stimuli assoziiert sind einerseits und der Repräsentation von Belohnungswert unabhängig vom Stimulus- und Belohnungstyp andererseits und drittens um die Rolle dieser Hirnstrukturen bei der Verbindung von Belohnungen zu Handlungen. Im ersten Teil der Abbildung (a) findet sich eine schematische Darstellung einer Belohnungsaufgabe mit einer konsistenten Zuordnung von Stimulus und Belohnung. In einer solchen Aufgabe lernen Probanden, welche Aktion (1 oder 2) ausgeführt werden muss als Reaktion auf Stimulus 1 oder 2, um eine Belohnung zu erlangen. Eine korrekte Antwort im Sinne einer Aktion 1 auf Stimulus 1 wird stets durch die Belohnung 1 belohnt. Gleichfalls wird die korrekte Aktion 2 als Reaktion auf Stimulus 2 mit der Belohnung 2 belohnt. Dies bedeutet, dass Probanden die Antworten aufgrund von Stimulus (S) – Antwort (Response=R) Assoziationen (in der Abbildung als rote Linien dargestellt) oder über Assoziationen zwischen spezifischen Stimuli und spezifischen Belohnungs-Ergebnissen (Outcome=O; S-O Assoziationen) oder über die Assoziation zwischen Antworten und Outcomes (R-O Assoziationen) auswählen können (letztere dargestellt als grüne Linien).

In Abschnitt (b) der Abbildung wird zum Vergleich eine Belohnungsaufgabe mit einer inkonsistenten Zuordnung dargestellt. Wie in der Aufgabe mit konsistenter Zuordnung lernen die Probanden hier, welche Aktion (1 oder 2) auf einen bestimmten Stimulus 1 oder 2 ausgeführt werden muss. Im Unterschied zur konsistenten Aufgabe erfolgt nun aber auf die korrekte Ausführung von Aktion 1 auf Stimulus 1 entweder die Belohnung

1 oder die Belohnung 2. Eine korrekte Antwort 2 auf Stimulus 2 wird ebenfalls belohnt, aber entweder durch Belohnung 1 oder Belohnung 2. Dies bedeutet, dass in diesem Falle die Probanden die Antworten nur aufgrund von S–R Assoziationen (rote Linien) lernen können, da S–O oder R–O Assoziationen nicht gelernt werden können.



Im unteren Teil der Abbildung sind nun Befunde aus der funktionellen Bildgebung dargestellt. (ci) zeigt, dass die Aktivität des IOFC (rot) jegliche Form von Feedback widerspiegelt, also sowohl die Rückmeldung, dass eine Antwort nicht korrekt war als auch die Rückmeldung, dass eine Antwort korrekt war. Da die Aktivität sich zwischen konsistent und inkonsistent trainierten Probandengruppen unterschied, liegt es nahe anzunehmen, dass der IOFC Erwartungen über bestimmte Belohnungstypen erlernt

und nicht nur, dass eine Belohnung folgt. Im Unterschied hierzu reflektiert die Aktivität im vmPFC/mOFC lediglich den Belohnungswert. Die Aktivität ist für positive Outcomes größer als für weniger positive, gleichgültig, ob der jeweilige Outcome informativ für das Erlernen von S-O Assoziationen oder die Bildung von Belohnungserwartung ist (ci und cii).

Die Aktivität im ACC unterschied sich zwischen den beiden Gruppen (größer für die Gruppe mit der konsistenten S-O Beziehung) und wurde darüber hinaus moduliert durch die Wahrscheinlichkeit, eine korrekte Antwort zu geben (di).

4. Lernen und motiviertes Verhalten

Eine besondere Variante des motivgesteuerten Verhaltens ist das Reinforcement Learning (RL), das computertheoretisch formuliert werden kann und enge Verbindungen zur instrumentellen/ operanten Konditionierung aufweist. Fast alle Modelle von RL teilen bestimmte Basiseigenschaften. Danach berechnet das lernende System – in unserem Falle der Mensch – während der Durchführung eines Experimentes (aber natürlich auch in der „freien Wildbahn“) iterativ eine Schätzung des mittleren Wertes aller möglichen Aktionen, also die mittlere Belohnung, die mit eine jede Aktion assoziiert ist. Diese Schätzung dient zugleich als Prädiktion für die Belohnung (bzw. Bestrafung), die eine Aktion erzielen wird. Nach jeder abgeschlossenen Aktion wird die tatsächlich erhaltene Belohnung verglichen mit der erwarteten Belohnung. Die aus diesem Vergleich resultierende Differenz ist der sog. Prädiktionsfehler, der positiv ist, wenn die tatsächliche Belohnung größer ist als die erwartete. Dieser Fehler wird verwendet, um die Belohnungsprädiktion zu präzisieren und um das Verhalten in Bezug auf Belohnungsmaximierung anzupassen. Ein populäres Modell von RL ist das sog. Zeitdifferenz (Time difference, TD) Actor-Critic Modell, das (in verschiedenen Varianten) zwischen dem Actor- und dem Critic Signal unterscheidet. Das Critic-Signal codiert den Prädiktionsfehler; das Actor-Signal verwendet diesen Fehler für die adaptive Auswahl und Sequenzierung von Aktionen. Der TD Prädiktionsfehler schließt Informationen über zeitverzögerte Belohnungen als Folge der Aktionen der Vergangenheit ein. Daraus folgen spezifische Vorhersagen über den zeitlichen Verlauf einer Fehlerantwort:

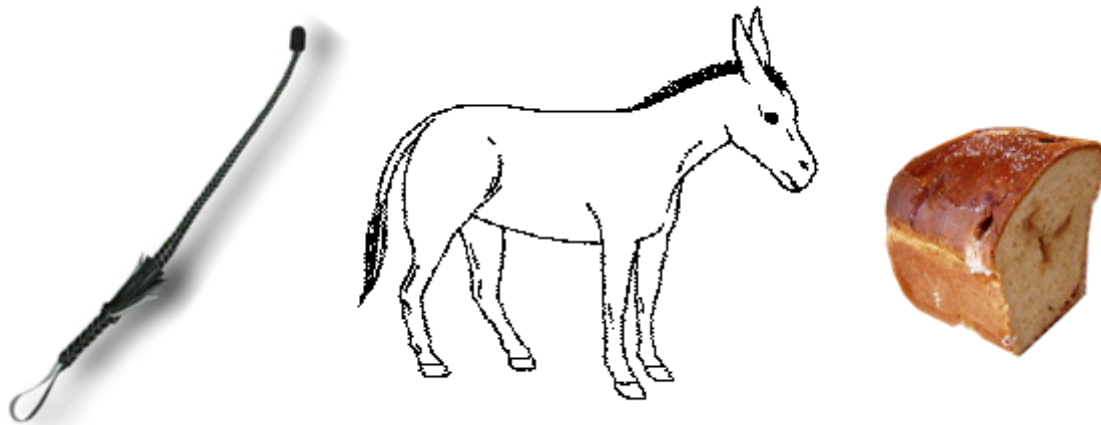
Wenn in einem Konditionierungsparadigma eine Assoziation zwischen einem Stimulus und einer Belohnung gelernt werden soll, tritt vor dem Lernen dieser Assoziation ein positiver Fehler nur auf, wenn die Belohnung selbst präsentiert wird; während des Lernens verlagert sich die Antwort aber hin zum assoziierten Ereignis. Ferner löst ein Auslassen der Belohnung eine negative Fehlerantwort zum Zeitpunkt der erwarteten Belohnung aus. Wie der Prädiktionsfehler für die Selektion und das Erlernen von Aktionen verwendet wird, ist allerdings beim klassischen Lernen (d.h., die Belohnung ist unabhängig vom Verhalten) und beim instrumentellen Lernen (d.h., die Belohnung hängt ab vom Verhalten) verschieden. Ein wichtiger Unterschied beispielsweise betrifft Änderungen des motivationalen Zustandes und deren Konsequenzen für das Handeln.

TD Modelle von RL lassen sich auf der Basis rein lerntheoretischer Überlegungen formulieren. Die Verbindung zur Neurobiologie beruht nun auf folgender grundlegender Beobachtung bei Vertebraten: Physische Antworten dopaminerger Mittelhirnneurone bei RL Aufgaben lassen sich interpretieren als Signale, die den TD Prädiktionsfehler kodieren und die in Bezug auf die verschiedenen Varianten von Belohnung ein einheitliche Metrik kodieren. Hierauf wird näher einzugehen sein. Diese Interpretation wurde beispielsweise in folgender Version eines Actor-Critic Models elaboriert: Zellen der ventralen tegmentalen Area (VTA) und der Substantia nigra pars compacta (SNc) kodieren denselben Prädiktionsfehler, aber verwenden ihn unterschiedlich. Aktivität der VTA Zellen ist assoziiert mit dem Critic Signal und kontrolliert Bewertungsprozesse im basolateralen Kern der Amygdala und im orbitofrontalen Kortex. Aktivität der SNc Zellen ist assoziiert mit dem Actor-Signal und kontrolliert das Erlernen von Aktionen in kompetitiven kortiko-striato-thalamo-kortikalen Schleifen. Diskutiert wird in diesem Zusammenhang, ob Dopamin den Prädiktionsfehler selber kodiert oder vielmehr die Bindung der hedonistischen Evaluation an ein Objekt oder eine Handlung („incentive salience“) vermittelt.

In jüngster Zeit konnten die tierexperimentellen Ergebnisse und Modellbildungen teilweise auf den Menschen übertragen werden. Denn es gelang mittels bildgebender und elektrophysiologischer Verfahren, distinkte räumlich-zeitliche neurale Aktivierungsmuster zu identifizieren, die mit der Verarbeitung von Belohnung, dem Vergleich von erwarteter und tatsächlichem Belohnung und der daran adaptierten Selektion und Sequenzierung von motorischer Aktion korrelieren.

5. Motivieren durch Strafe oder durch Belohnung

Eine fundamentale Frage, die sich nicht nur verzweifelten Eltern stellt (sondern auch im Justizvollzug, in betrieblichen Hierarchien und eigentlich überall), ist, ob man besser durch Belohnungen oder durch Bestrafungen dazu motiviert wird, bestimmte Handlungen zu vollführen oder zu entlassen.

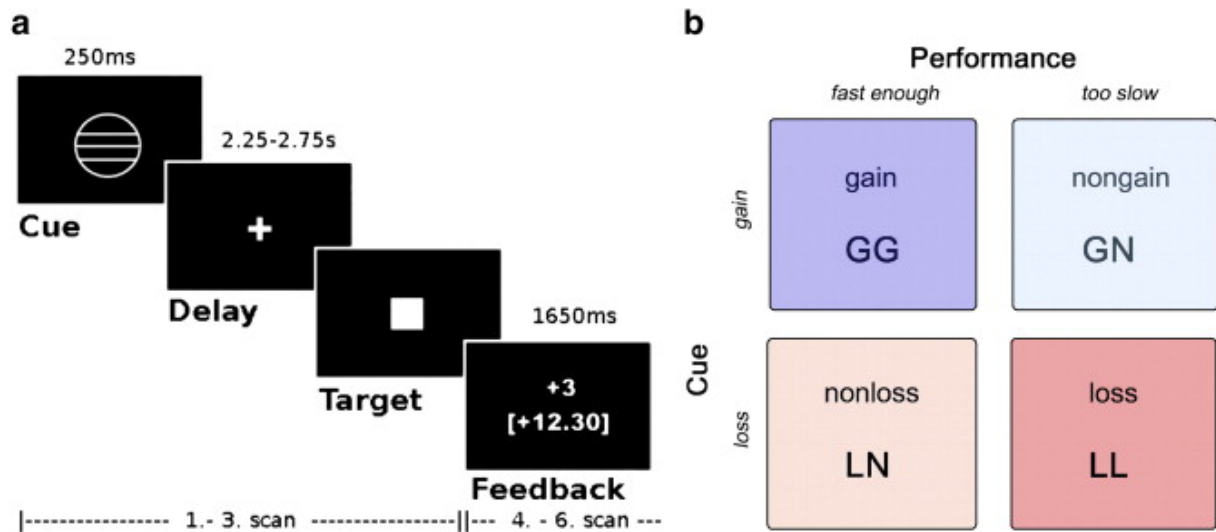


Den Esel in der obigen Abbildung können wir, wie das Sprichwort sagt, durch Zuckerbrot und Peitsche motivieren, sich von der Stelle zu bewegen. Im Englischen wird dies übrigens mit dem Idiom „By carrot or by stick“ in gleicher Weise gefasst.

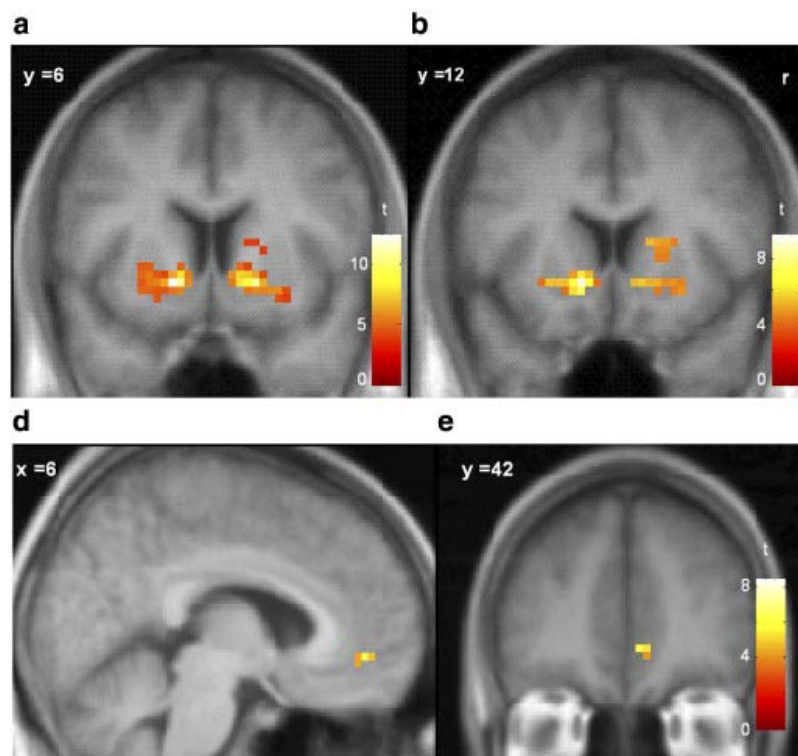
Unbestritten ist, dass man Menschen und Tiere sowohl mit Belohnungen dazu bringen kann, etwas zu tun, als auch mit Bestrafungen dazu motivieren kann, etwas nicht zu tun. Gibt es nun aber Unterschiede zwischen der neurobiologischen Verarbeitung von Bestrafungen und Belohnungen? Und: Ist das Weglassen einer Bestrafung gleichbedeutend mit dem Erhalt einer Belohnung?

Hierzu sind eine Reihe von Untersuchungen durchgeführt worden, die allerdings in ihrer Gesamtheit zu keinem einheitlichen Bild kommen. Ich möchte hier eine Studie erwähnen, die mir gut gefallen hat. Sie wurde in Berlin von Jana Wrase und Mitarbeitern durchgeführt und 2007 in der Zeitschrift *NeuroImage* veröffentlicht. Es wurden zwei verschiedene Versuchsbedingungen eingeführt. In der einen konnte der Proband bei ausreichend schneller Reaktionszeit einen Gewinn einfahren, war er nicht schnell genug, erhielt er nichts. In der anderen Bedingung erhielt er bei ausreichend

schneller Reaktionszeit nichts, bei zu langsamer Reaktionszeit wurde er allerdings bestraft durch einen finanziellen Verlust. Dies ist in der nächsten Abbildung illustriert:



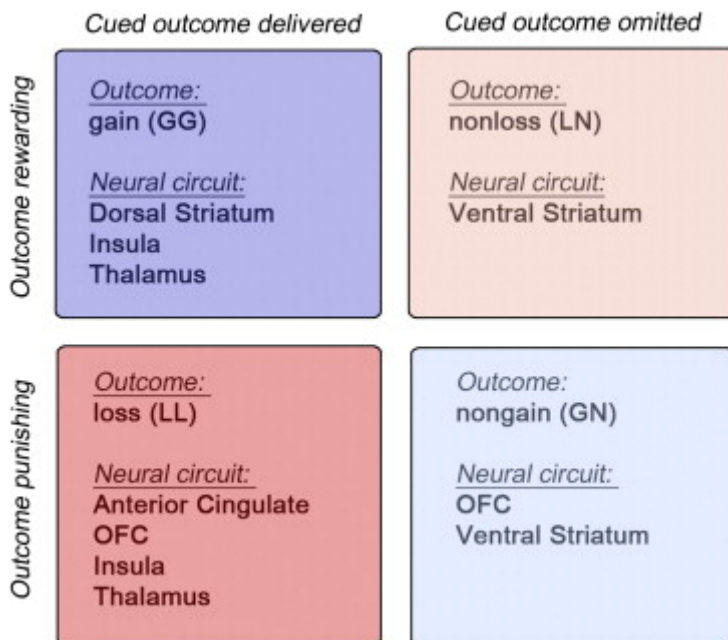
Es stellt sich die Frage, ob das Gehirn in beiden Bedingungen gleich reagiert. Ist also das Vermeiden einer Bestrafung gleichbedeutend mit dem Erhalt einer Belohnung. Zunächst schauten sich Wrase und Mitarbeiter die Aktivierung auf die Hinweisreize an:



Wie die Abbildung oben zeigt, führte die Antizipation eines potentiellen Geldgewinns zu einer Aktivierung des ventralen Striatums beidseits (a). Gleiches galt auch bei der Erwartung eines potentiellen finanziellen Verlusts (b). Es zeigte sich darüber hinaus

eine Aktivierung des medialen präfrontalen Cortex (MPFC) auf die Präsentation des Feedbacks für Geldgewinne im Vergleich zu keinem Gewinn (GG > GN) (d, e).

Es ergaben sich aus den weiteren Analysen dann folgende anatomische Überlegungen:



Die Verhaltensadjustierung in Richtung auf eine Maximierung von Belohnungen und Minimierung von Bestrafungen lässt sich, wie in der obenstehenden Abbildung gezeigt, bezüglich der Dimensionen „belohnendes“ vs. „bestrafendes“ Ergebnis (rewarding versus punishing outcomes) und erwartetes Ergebnis eingetreten vs. Nicht eingetreten (cued outcome delivered versus omitted) klassifizieren. Die Hirnregionen, die diesen Verhaltensanpassungen korrespondieren, sind ebenfalls in den jeweiligen Feldern gezeigt.

6. Delay Discounting

Was möchten sie lieber haben: 25 Euro sofort oder 50 Euro in 4 Wochen? Möchten Sie lieber diese Tafel Schokolade haben oder am Ende des Tages 2? Möchten Sie lieber ein Marshmallow haben oder in 5 Minuten 2?

Diese simplen Fragen weisen auf einen fundamentalen Mechanismus des menschlichen motivationalen Systems hin: Die Fähigkeit zum Belohnungsaufschub. Ein Extrembeispiel: Ein Waldbesitzer, der im Wald Eichen schlägt, bepflanzt die freigewordene Fläche mit neuen Bäumchen. Ein an sich völlig irrationales Verhalten

für ihn selbst, da erst seine Enkel in den Genuss kommen werden, die Bäume zu schlagen und möglicherweise damit Geld zu verdienen.

Wie entscheiden wir, ein Verhalten jetzt auszuführen oder vielleicht später? Solche Entscheidungen werden durch den Kontext beeinflusst, die Art des Verhaltens, die Wahrscheinlichkeit des Auftretens und natürlich auch durch Variablen, die in unserer Persönlichkeit begründet sind. Für Verhalten, die mit einer Belohnung einhergehen, wird unsere Entscheidungspräferenz oft bei der Entscheidung für „jetzt“ ausfallen, bei Verhalten, die mit einer Bestrafung einhergehen (oder die anstrengend sind), wird die Entscheidung häufig für „später“ fallen. Im Lauf der Evolution hat sich eine optimale Balance zwischen „jetzt“ und „später“ Entscheidungen herausgebildet, die dazu führt, dass wir sowohl unmittelbare Bedürfnisse stillen als auch unsere Langfristziele erreichen können. In diesem Zusammenhang ist das Konzept des Delay Discounting, das ich lieber nicht in die deutsche Sprache übersetzen möchte, von hoher Bedeutung.

Unter Delay Discounting (DD) verstehen wir einen Algorithmus der die Neigung eines Individuums beschreibt, eine unmittelbar verfügbare Belohnung zu erlangen oder eine unmittelbar unangenehme Folge zu vermeiden, selbst wenn später verfügbare Belohnungen größer sind (oder unmittelbare Bestrafungen kleiner). Letztlich geht es also um eine Abwertung einer in der Zukunft liegenden Belohnung (Abwertung = Discounting). Ein in Bezug auf DD oft benutzter Begriff ist der der Kurzsichtigkeit für die Zukunft (myopia of the future), womit man meint, dass die unmittelbaren Gewinne klar und deutlich sind, die in der Zukunft liegenden indessen offensichtlich vor dem mentalen Augen nicht oder nur unscharf sichtbar sind, so dass sie uns weniger Wert erscheinen.

Neben dem Delay Discounting, das durch die Eingangsfrage dieses Kapitels beschrieben wird (Möchtest du 25 Euro sofort oder 50 Euro in 4 Wochen?), unterscheiden wir auch noch andere Arten der Abwertung von Belohnungen (und Bestrafungen). Beim Probability Discounting (PD) erfolgt die Abwertung aufgrund von Wahrscheinlichkeiten: Möchtest du 25 Euro sicher oder eine Lotterie spielen mit 50 % Wahrscheinlichkeit um die Möglichkeiten 100 und 0 Euro? Offensichtlich ist die Lotterie mit einer durchschnittlichen Gewinnsumme von 50 Euro „mehr“ wert. Viele Probanden entscheiden sich dennoch für die sicheren 25 Euro. Ein Spatz in der Hand ist mehr wert als die Taube auf dem Dach. Es wird hier schon deutlich, dass diese im

Experiment leicht zu untersuchenden Entscheidungen etwas über die individuellen Eigenschaften einer Person aussagen können.

Beim Effort Discounting (ED) geht es um die Frage, wie viel kognitiver und / oder physischer Aufwand getrieben werden soll, um eine Belohnung zu bekommen. Jeder weiß, dass es eigentlich viele Monate dauert, einen Waschbrettbauch zu bekommen oder 10 kg abzunehmen. Aufgrund des damit verbundenen Aufwands fangen viele Personen erst gar nicht an, diese Ziele zu erstreben. Stattdessen kaufen sie Zeitschriften, die für „10 kg weniger in 5 Tagen“ oder mit „in 3 Wochen zum Waschbrettbauch“ aufmachen. Das Durchdenken dieser Beispiele macht schon deutlich, dass es bei der Beschäftigung mit DD, PD und ED auch um die Frage gehen muss, wie sich Menschen in ihrer Bewertung von frühen und späten Belohnungen (DD), von sicheren oder unsicheren Belohnungen (PD) oder von leicht erhältlichen und schwer verdienten Belohnungen (ED) unterscheiden.

Ich will mich hier zunächst auf DD konzentrieren aus den folgenden Gründen:

- DD hat eine hohe ökologische Validität
- DD hat eine hohe Augenscheinvalidität
- DD kann eine Reihe von ungesunden Verhaltensweisen (zu viel Essen, zu viel Trinken, Kaufsucht) erklären
- DD spielt bei schwierigen Persönlichkeitseigenschaften und psychiatrischen Erkrankungen eine Rolle
- Die neuronalen Grundlagen des DD sind gut untersucht und bereits gut verstanden

Wie lässt sich DD mathematisch fassen?

In der Ökonomie wurde ursprünglich davon ausgegangen, dass die Abwertung von in der Zukunft liegenden Gewinnen entsprechend einer exponentiellen Funktion erfolgt. Dies ist allerdings mit der empirischen Evidenz der letzten 30 Jahre nicht vereinbar. Man geht heute davon aus, dass die Entwertung entsprechend einer hyperbolischen oder quasi-hyperbolischen Funktion erfolgt. Am einfachsten lässt sich DD mit der von Mazur (1984) beschriebenen Funktion erfassen:

$$V = \frac{A}{1 + kD}$$

In dieser Gleichung repräsentiert V den Wert der in der Zukunft liegenden Belohnung A , die nach einer Verzögerung D verfügbar ist. K ist die so genannte "delay discount rate" oder Abwertungsrate. Von ihr hängt es ab, wie steil die Abwertungskurve ist.

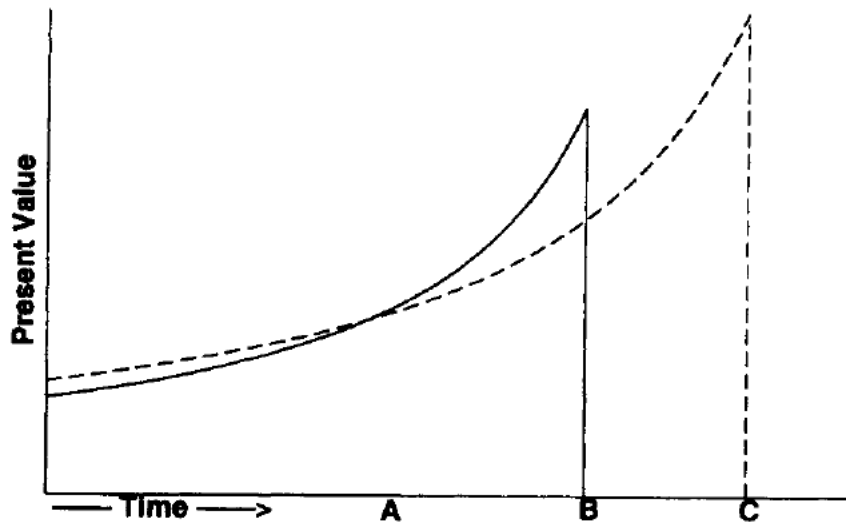


Abbildung: Hyperboles Discountingmodell

Man kann dies leicht überprüfen:

In einer kürzlich von uns gemeinsam mit Manfred Herrmann (Universität Bremen) und Stephan Miedl (Universität Salzburg) durchgeführten Studie mit pathologischen Spielern (und Normalpersonen) bestimmten wir den k -Wert für die Spieler mit $0,06$ und den der Normalpersonen mit $0,02$. Dies bedeutet, dass für die Spieler eine Summe von 100 Euro nach einer Verzögerung von 50 Tagen nur noch 25 Euro wert ist. Für Kontrollpersonen findet sich hingegen ein Wert von 50 Euro. Immerhin doppelt so viel! Gleichwohl findet in unserem Gehirn offensichtlich eine beispiellose Entwertung von in der Zukunft liegenden Gewinnen statt. Kann man dies ausnutzen? In der Tat: Nach der Arbeit sind Sie auf dem Weg nach Hause, der Sie eine halbe Stunde kosten wird. Sie gehen an einem Starbucks vorbei und leisten sich einen Kaffee für 3 Euro. In einer halben Stunde könnten Sie einen Kaffee zu Hause für einen Bruchteil dieses Preises genießen, schließlich kostet ein Pfund Kaffee „beim Aldi“ nur 5 Euro, ergibt aber viele Liter Kaffee.

Die Neurobiologie des DD

Im Folgenden gebe ich einige Hinweise zu den neurobiologischen Grundlagen des DD also der Entscheidungen für "jetzt" oder "später". Ein guter gedanklicher Ausgangspunkt, wie eine solche Entscheidungsfindung im Gehirn organisiert sein könnte, ist die Frage, ob eine sofortige Belohnung verfügbar ist. Vergleichen Sie einmal ihr Gefühl beim Beantworten der folgenden Fragen:

1. Was möchten Sie lieber: 25 Euro sofort oder 50 Euro in 4 Wochen?
2. Was möchten Sie lieber: 25 Euro in 2 Wochen oder 50 Euro in 6 Wochen?

Eigentlich sind die Fragen fast identisch und auch die Verzögerung zwischen den beiden Belohnungen ist gleich. Trotzdem ist es offensichtlich so, dass man intuitiv eher die 25 Euro in Frage 1, die sie sofort in Aussicht stellt, nehmen würde, als in Frage 2, die eine Wartezeit vorzieht.

Was könnte das für das Gehirn bedeuten?

Von einigen Autoren um Sam McClure ist vorgeschlagen worden, dass es im Gehirn zwei Systeme gibt: ein Beta-System wäre immer dann aktiv, wenn eine Entscheidung für eine sofortige Belohnung im Raum steht. Dieses Beta-System umfasst limbische und paralimbische Strukturen wie den ventromedialen präfrontalen Cortex (vmPFC), das ventrale Striatum (VS), den Hippocampus, die Amygdala und die Insula und ist für die Integration komplexer appetitiver und aversiver Informationen bei der Koordination von Verhalten zuständig. Dies ist vor allem immer dann notwendig, wenn sich die Gelegenheit zu einer sofortigen Belohnung ergibt.

Das zweite System, Delta-System genannt, besteht aus dem dorsolateralen präfrontalen Kortex (dlPFC), dem dorsalen anterioren cingulären Kortex (dACC) und posterioren parietalen kortikalen Arealen (PPC) und bewertet den subjektiven Wert von zukünftigen Ergebnissen von Aktionen. Es erlaubt die Simulation von Verhaltenskonsequenzen. Dieses System ist vor allem dann tätig, wenn sich ein Proband für eine verspätete Belohnung entscheidet.

Von Nora Volkow wurde darüber hinaus zusammengetragen, welche Botenstoffe im Gehirn für die „jetzt“ oder „später“ Entscheidungen zuständig sind. Interessanterweise dreht sich hier alles um den ihnen schon vielfach bekannten Botenstoff Dopamin (DA), wobei allerdings andere Botenstoffe für verschiedene mit dem jetzt / später Paradigma zusammenhängende Prozesse zuständig sein könnten, so z.B. vorschnelle Entscheidungen ohne Überlegung, Mangel an Durchhaltevermögen (insbesondere in

Aufgaben, die Energieaufwand bedeuten oder die unangenehm sind), Dringlichkeit (Tendenz auf starke Impulse präferentiell zu reagieren) oder Neuigkeitssuche (Englisch: Sensation Seeking).

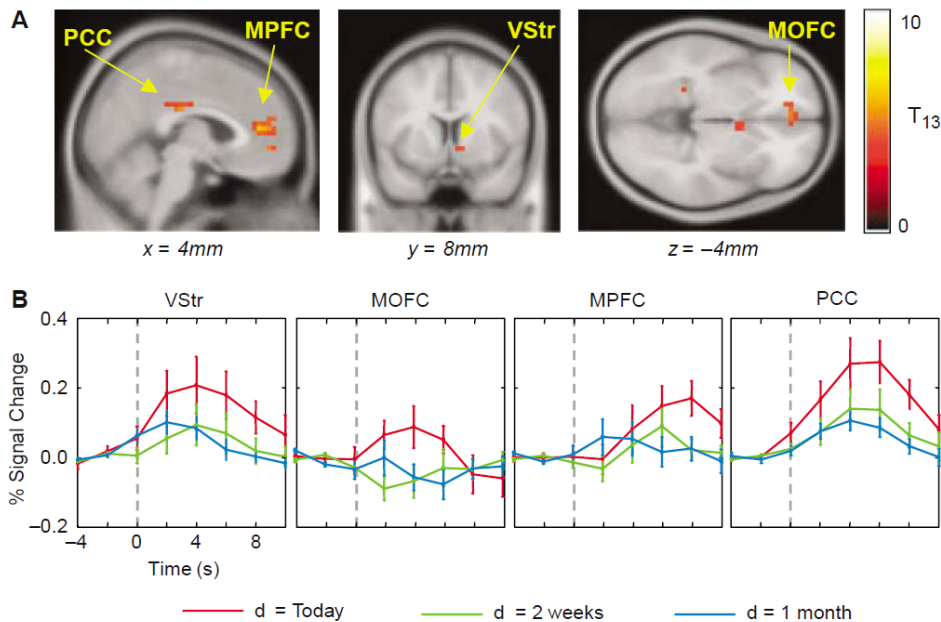


Abbildung oben (aus McClure et al., Science, 2004): Hirnregionen, die präferentiell für Entscheidungen aktiv sind, bei denen einen unmittelbar verfügbare Belohnung (today) zur Verfügung steht (Beta-System).

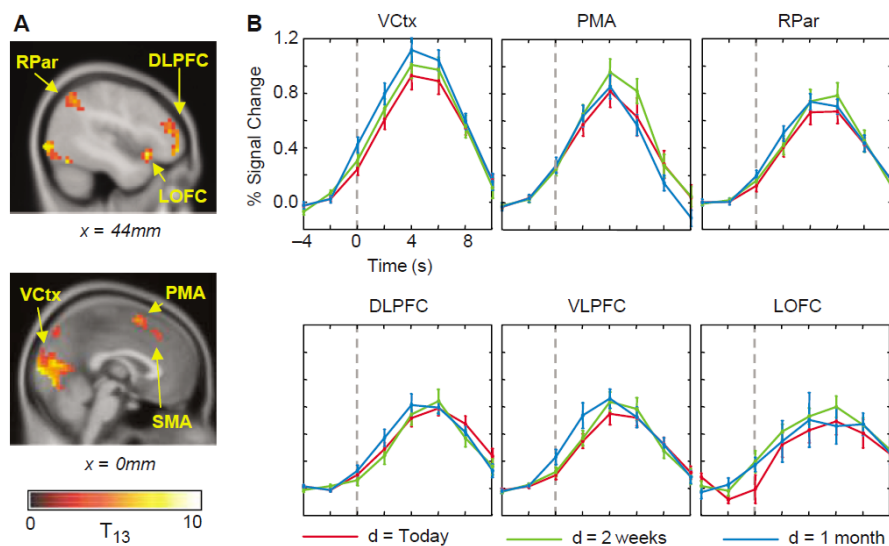


Abbildung oben (aus McClure et al., Science, 2004): Hirnregionen, die unabhängig davon aktiv sind, ob die Belohnung sofort oder erst verspätet zur Verfügung steht und die darum dem Delta-System entsprechen.

Wenn man auf Dopamin fokussiert, ergibt sich folgendes Bild: „später“ Prozesse erfordern eine tonische, dauerhafte Ausschüttung von DA, wohingegen „jetzt“ Prozesse durch schnelle, scharfe „bursts“ von DA Aktivität getrieben sind, die die Aufmerksamkeit auf einen Stimulus (eine Belohnung lenken). In der Tat konnte gezeigt werden, dass Stimuli, die zu einer schnellen, burst-artigen Ausschüttung von DA im Striatum führen, als belohnend empfunden werden und von den Probanden gewollt werden, wohingegen Stimuli, die langsame, tonische DA Ausschüttung auslösen, eher nicht sofort gewollt werden. Solche tonischen DA Ausschüttungen führen indessen dazu, dass die Fähigkeit, eine längerfristige (kognitive) Anstrengung zu erbringen, gesteigert und die Aufgabe als motivierend und interessant erlebt wird.

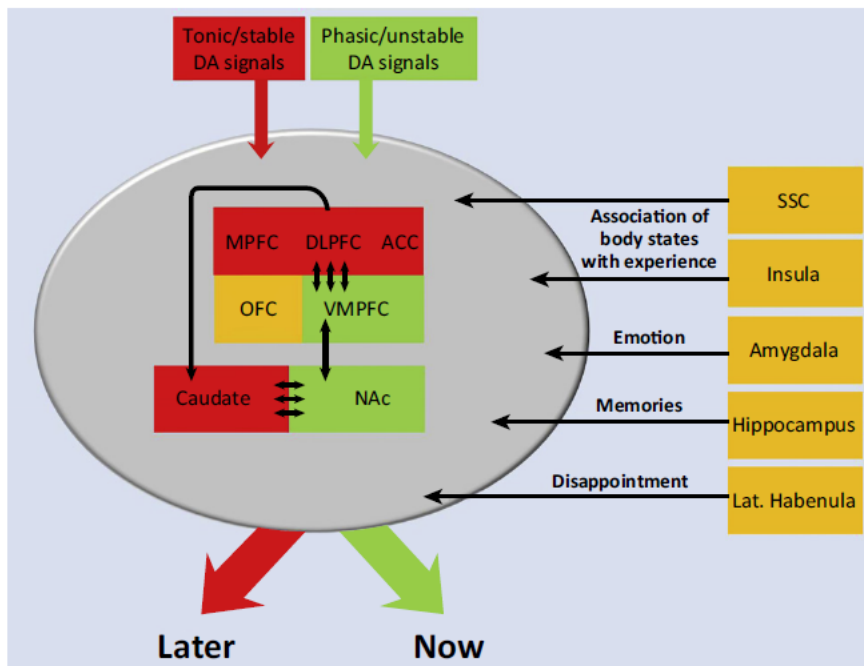


Abbildung oben (aus Volkow und Baler, 2015): Dargestellt sind die Komponenten des Hirnsystems, das auf unmittelbare Belohnungen anspringt (grün) und des Hirnsystems, welches auf spätere Belohnungen spezialisiert ist (rot). Bitte beachten Sie auch die gelb-orangen Hirnregionen, die zusätzliche Informationen beisteuern und die offensichtlich nicht dopaminerg sind.

Nicht nur der Modus der Ausschüttung („burst“-artig oder tonisch) sondern auch der anatomische Ort der Ausschüttung innerhalb des Striatums bestimmen die Funktion von DA: Im Nucleus caudatus (Schweifkern) moduliert DA die inhibitorische Kontrolle von kognitiven Prozessen, im ventralen Striatum (VS) werden impulsive

Entscheidungen beeinflusst und im Putamen, einem weiteren Teil der Basalganglien, moduliert DA Gewohnheiten und Routinehandlungen. Von den Basalganglien gehen Projektionen in die verschiedenen Teile des präfrontalen Kortex, so dass Schleifensysteme entstehen, die für verschiedene Aspekte motivierten Verhaltens zuständig sind.

Diese anatomischen und neurochemischen Überlegungen sind plausibel. Wir selber haben Untersuchungen gemacht, die für die Existenz des Beta- und Deltasystems sprechen.

Es soll hier allerdings nicht verschwiegen werden, dass es auch alternative Sichtweisen der neuroanatomischen Unterstützung von DD ("jetzt" versus "später") Prozessen gibt. So wird von einigen Forschern (zum Beispiel Paul Glimcher aus New York) angenommen, dass es ein einheitliches System gibt, welches den Wert einer jeden Belohnung berechnet, wobei unter anderem auch die Wartezeit, bis die Belohnung verfügbar wird, mitberücksichtigt wird.

7. Altruistische Bestrafung (Altruistic Punishment)

Bisher haben wir motivationale Prozesse aus einer egoistischen Perspektive betrachtet. Die Entscheidungen des einzelnen sind nach dieser Sichtweise geprägt durch Funktionen (und entsprechende neuroanatomische Regelkreise), die eine Handlungsoption und die damit assoziierte Belohnung bewerten, so dass die Entscheidung immer für diejenige Handlungsalternative ausfällt, die mit der größten Belohnung verbunden ist. Nimmt man eine evolutionäre Perspektive ein, so fällt es einem leicht, sich vorzustellen, dass Handlungen auch getrieben sein könnten durch Vorteile für die eigenen Nachkommen. So leuchtet es ein, dass es sinnvoll ist, meiner Tochter das Studium in Oxford zu finanzieren, damit dieser dann eine höhere Chance hat, ihre Gene (und das sind ja auch meine!) bestmöglich (d.h., in Kombination mit anderen kompetenten Genen) weiterzugeben.

Ernst Fehr aus Zürich hat jedoch darauf hingewiesen, dass menschliche Kooperation auch in vielen anderen Situationen auftritt und ein evolutionäres Rätsel darstellt. So kooperieren wir häufig mit genetisch nicht verwandten Fremden, oftmals in größeren Gruppen, mit Menschen, die wir nie wieder sehen werden, und in Situationen, in denen ein Reputationszuwachs minimal oder gar fehlend ist. Offensichtlich können derartige

Formen der Kooperation nicht mit verwandtschaftlichen Verhältnissen und evolutionären Überlegungen zur präferentiellen Förderung des eigenen genetischen Materials erklärt werden. Auch greifen Überlegungen der so genannten „Theorie des reziproken Altruismus“ (helf ich dir, so hilfst du mir) zu kurz.

Eine neuroökonomische Sichtweise von evolutionär wichtigen menschlichen Aktivitäten wie zum Beispiel die Jagd auf Großwild, das Teilen von Fleisch, die Erhaltung von Gemeinschaftseigentum und die Kriegführung definiert diese als Gemeinschaftsgüter. Ein jedes Gruppenmitglied profitiert von diesen Gütern, auch jene, die nichts in diese Güter investiert haben. Dies wirft laut Fehr die Frage auf, warum Menschen sich regelmäßig in Aktivitäten wie Kriegführung oder Großwildjagd stürzen, obwohl diese mit immensem persönlichen Aufwand und Kosten (Kriegführung kann ja bekanntlich das Leben kosten) verbunden sind. Eine Reihe von Theorien sind aufgestellt worden, um diese Verhaltensweisen zu erklären. Die Theorie der Verwandtschaftsselektion („kin selection“) nimmt an, dass auf diese Weise das Überleben von verwandten Individuen (und damit die Weitergabe von genetischem Material, welches dem eigenen sehr ähnlich ist) gesichert wird. Die Theorie der direkten Reziprozität stellt die selbstbezogenen Anreize zur Kooperation in bilateralen Langzeitinteraktionen heraus. Theorien der indirekten Reziprozität hingegen nehmen an, dass Kooperationen in größeren Gruppen dadurch entstehen können, wenn die kooperierenden Individuen eine Reputation aufbauen können, von der sie später „zehren“ können.

Warum aber Menschen auch in Situationen zusammenarbeiten, in denen genetisch nicht verwandte Personen nur einmalig interagieren und daher der Aufbau einer Reputation nicht erfolgen kann, war bisher ein Rätsel.

Das Konzept der (altruistischen) Bestrafung kann derartige Situationen erklären. Wenn Individuen, die versuchen andere auszunutzen, bestraft werden, kann sich Kooperation auszahlen. Beim Versuch einmalige Kooperation zwischen genetisch nicht verwandten Individuen durch Bestrafung von Schmarotzern / Ausnutzern zu erklären, stellt sich jedoch die Frage, wer die Kosten für diese Bestrafung tragen soll und warum er dieses tun sollte. Jedes Gruppenmitglied profitiert in der Tat davon, wenn Schmarotzertum unterbunden wird, aber niemand hat wirklich einen Anreiz, durch Einsatz eigener Mittel und Anstrengungen, die Bestrafung des Schmarotzers zu leisten. Laut Fehr stellt die Bestrafung von Schmarotzern ein Gemeingut zweiter

Ordnung dar. Das Problem von Gemeingütern zweiter Ordnung könne dann gelöst werden, wenn genügend Menschen eine Tendenz zur altruistischen Bestrafung haben. Solche Menschen sind motiviert, Schmarotzer zu bestrafen, obwohl die Bestrafung aufwändig ist und keinen direkten Vorteil für den Bestrafenden bringt.

Die Resultate von Fehr belegten, dass Schmarztum starke negative Gefühle auslösten. Darüber hinaus fanden sie Hinweise dafür, dass diese emotionalen Reaktionen die Bestrafung des Schmarotzers aus folgenden Gründen triggerten: Wenn emotionale Gründe für die Bestrafung eine Rolle spielen, sollte es so sein, dass die meisten Bestrafungen von solchen Probanden ausgeübt werden, die überdurchschnittlich viel gegeben haben, gegenüber solchen Probanden die unterdurchschnittlich viel gegeben haben. Dies war in den Untersuchungen von Fehr der Fall, da 74,2% aller Bestrafungen diesem Muster folgten. Zweitens sollte die Bestrafung zunehmen je weiter der Schmarotzer vom durchschnittlichen Investment der anderen Spieler entfernt ist. Auch dies war der Fall. Drittens war für die Schmarotzer die Bestrafungsandrohung sehr real, denn intuitiv wissen diese Personen, dass sie mit ihrem Verhalten starke negative Emotionen auslösen. In der Tat konnte Fehr beobachten, dass die Wegnahme der Bestrafungsoption aus dem Experiment sofort zu einer erheblichen Abnahme des Investments führte.

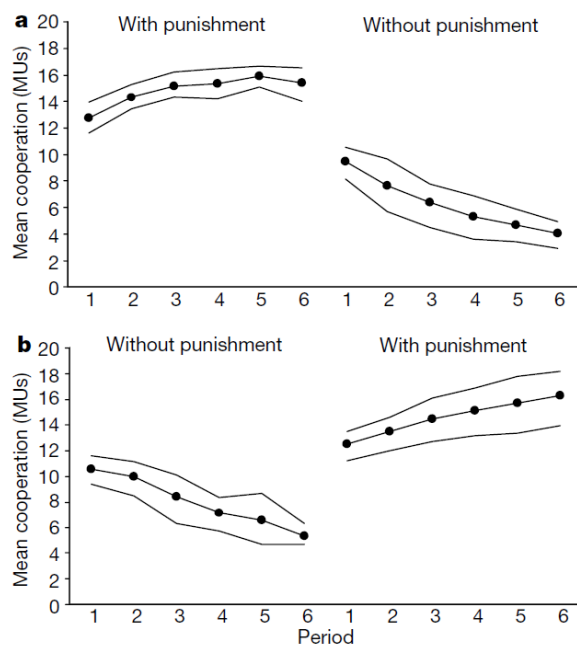


Abbildung oben: In Experiment (a) wurden zunächst 6 Blöcke gespielt, in denen die Bestrafung von Schmarotzern möglich war, danach wurden weitere 6 Blöcke gespielt, in den die Bestrafung nicht mehr möglich war. Dies führte zu einer Abnahme der

Kooperation. Umgekehrt war es in Experiment (b), in dem zunächst keine Bestrafung möglich war. In den zweiten 6 Blöcken wurde die Bestrafungsmöglichkeit eingeführt, was unmittelbar zur Folge hatte, dass die Kooperation der Probanden zunahm.

Zusammengefasst bedeutet dies, dass Emotionen ein herausragender Faktor bei der altruistischen Bestrafung sind. Dies hat, so Fehr, erhebliche Konsequenzen für die evolutionäre Sichtweise des menschlichen Verhaltens, welche klassischerweise die Förderung von verwandten Individuen und damit die bessere Möglichkeit eigenes genetisches Material zu verbreiten fokussiert hat. Fehr bleibt letztlich eine eigene Erklärung schuldig, fordert aber weiter Forschungen hierzu.

In einer weitere Arbeit, die Fehr zusammen mit Neurowissenschaftlern zur altruistischen Bestrafung durchgeführt hat (de Quervain et al, Science 2004), wurde das Kooperationsspiel im PET-Scanner gespielt.

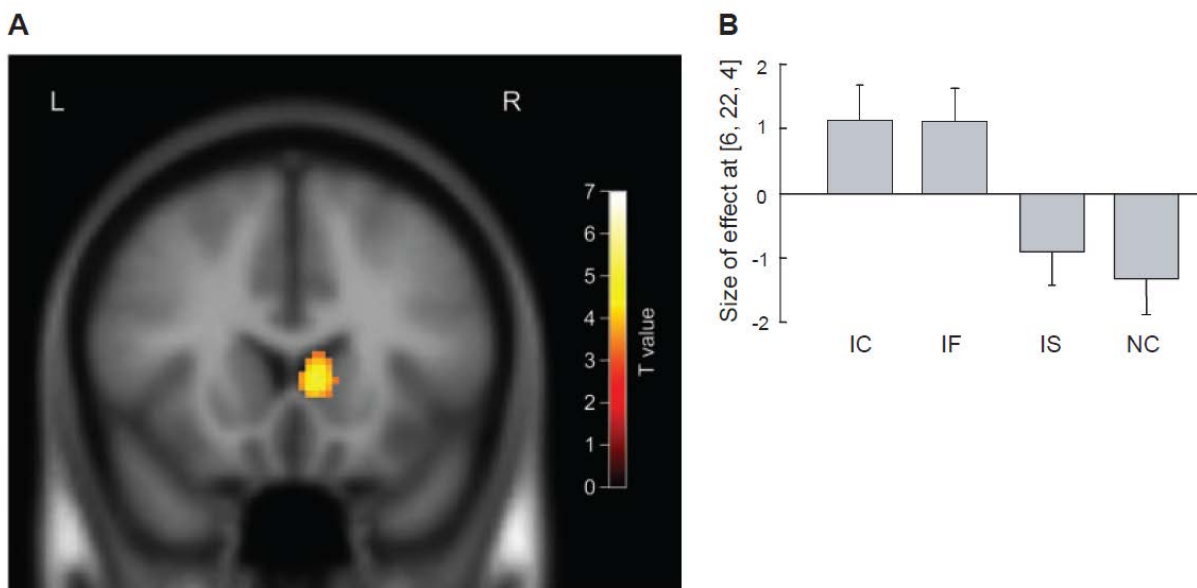


Abbildung oben aus de Quervain et al.: In (a) ist die Aktivierung des Nucleus caudatus dargestellt, die für Bedingungen auftrat, in den die Probanden die Neigung, ihren Mitspieler zu bestrafen, verspürten, und diese Neigung auch die Tat umsetzen konnten (Bedingungen IC und IF) im Vergleich zu Bedingungen, in denen sie entweder keine Neigung zur Bestrafung verspürten oder eine Bestrafung nicht erlaubt war (Bedingungen IS und NC). In (b) ist dies noch einmal in Form von Effektgrößen für die Aktivierung des Nucleus caudatus dargestellt.

8. Gewohnheiten

Es besteht ein grundsätzlicher Unterschied zwischen zielgerichtetem Verhalten und Gewohnheiten. Erstere sind getrieben durch unser Verlangen, ein Ziel zu erreichen. Dies kann bedeuten, dass ich dadurch ein Übel vermeide (z.B. erreiche ich als Schiffbrüchiger das rettende Ufer und vermeide so das Ertrinken) oder dass ich eine Belohnung erlange (z.B. erhalte ich eine Geldprämie, wenn ich eine bestimmte Leistung erfüllt habe). Gewohnheiten legen wir indessen auch an den Tag, wenn weder eine Bestrafung zu vermeiden noch eine Belohnung zu erlangen ist. Das Verhalten ist somit von Belohnungskontingenzen entkoppelt. Es macht sicher Sinn, einen Teil unserer täglichen Verhaltensweisen als Gewohnheiten zu realisieren. Dies befreit unser Gehirn von einer Menge Routineaufgaben. Nützliche Gewohnheiten sind z.B. das Zähneputzen, welches wir unseren Kindern durch Zureden, Loben, Versprechungen, Belohnungen und Bestrafungen beibringen, welches dann jedoch irgendwann von diesen (und von uns selbst) wie automatisch verrichtet wird. Andere Gewohnheiten (die Zigarette nach dem Essen, die Chips zur Sportschau, das übermäßige Trinken) sind indessen störend oder sogar gesundheitsschädlich. Dass Gewohnheiten schädlich sein können, kann man sehr gut an der Alkoholabhängigkeit illustrieren: Eine Substanzabhängigkeit ist einerseits assoziiert mit einem Verlangen und einer Erwartung von positiven Effekten (im Falle von Alkohol: Aufheiterung oder Beruhigung). Diese Belohnungen sprechen dafür, dass der Gebrauch von Substanzen mit intentionalem, zielgerichtetem Verhalten erklärt sein kann. Andererseits zeigt die Tatsache, dass viele Menschen nicht in der Lage sind, einen Alkoholkonsum zu beenden, selbst wenn sie dies „wollen“ und selbst wenn der Konsum bereits zu negativen Folgen (Verlust des Arbeitsplatzes) geführt hat, dass Alkoholkonsum eben nicht nur intentional getrieben sondern auch gewohnheitsmäßig erfolgt. Letztendlich sind sowohl belohnungsabhängige, intentionale Prozesse und belohnungs- und bestrafungsunabhängige Gewohnheiten bei der Entwicklung und der Aufrechterhaltung eines schädlichen Alkoholgebrauchs oder einer Alkoholabhängigkeit wichtig. Zur Erinnerung: Wir wollen hier zielgerichtetes motiviertes Verhalten als eine Aktion (oder eine instrumentelle Antwort) definieren, die durch ihre Assoziation mit einem erstrebenswerten Resultat (einer Belohnung / Vermeidung einer Bestrafung) definiert ist. Demgegenüber definieren wir Gewohnheiten als Verhaltensweisen, die unabhängig von ihrer Verhaltenskonsequenz

(positiv oder negativ) auftreten. Gewohnheiten werden oft durch einen „Trigger“- oder Hinweisreiz ausgelöst (die Bierreklame im Fernsehen). Im Falle von Suchterkrankungen geht man davon aus, dass die sinnvolle Maschinerie des Gewohnheitslernens unglücklicherweise genutzt wird, um eine durch suchtbezogene Hinweisreize getriebene von den Konsequenzen unabhängige Verhaltenskette aufzubauen. Dies gilt selbstverständlich nicht nur für stoffbezogene Süchte sondern auch für andere suchtbezogene Verhaltensweisen (Kaufsucht, Sexsucht, Spielsucht, Arbeitssucht).

Wie kann man Gewohnheitsbildung im Experiment untersuchen? Eine Grundvoraussetzung ist zunächst, dass die Transformation von einem zielgerichteten Verhalten zu einer von Belohnungen / Bestrafungen unabhängigen Verhaltensweise demonstriert wird. Bei zielgerichtetem Verhalten stellt man schnell fest, dass dieses ausbleibt bzw. rasch abnimmt, sobald die Belohnung ausbleibt, dies ist bei Belohnungen nicht der Fall.

Das erste bildgebende Experiment beim Menschen, das zu Belohnungsbildung durchgeführt worden ist, ist das von Tricomi und Mitarbeitern. Dieses soll im Folgenden kurz dargestellt werden.

In dieser bahnbrechenden Studie wurden normale Probanden während einer so genannten freien operanten Konditionierungsaufgabe, die vorsah, dass Antworten auf graphische Stimuli (Fraktale) nach einem variablen Intervall Schema belohnt wurden, und zwar entweder mit M&M Schokolinsen oder mit Fritos Chips, die unmittelbar nach der Scanningsitzung zu konsumieren waren. Ein Verstärkungsprotokoll mit variablen

Intervallen wurde gewählt, da dieses besonders gute Verhaltenseffekte aufweist.

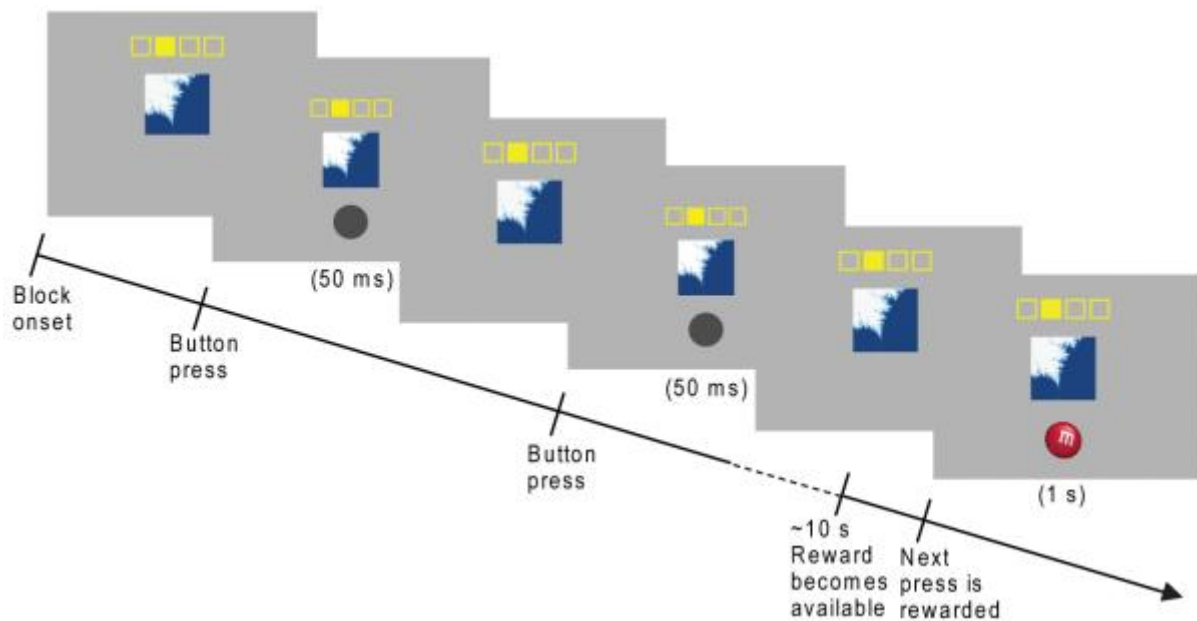


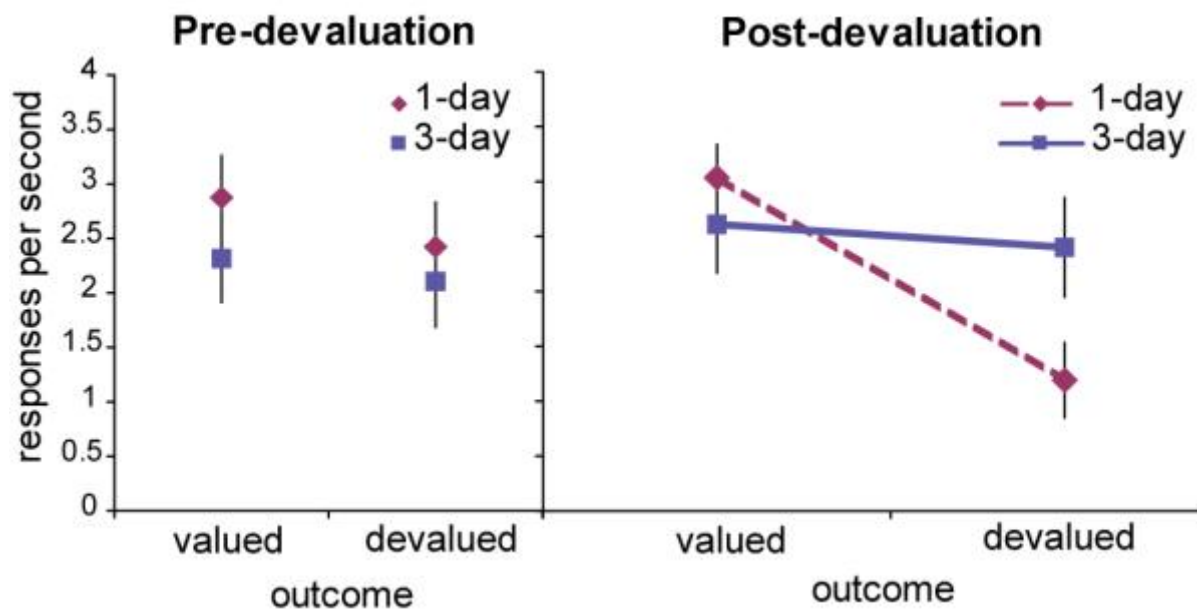
Illustration des Experiments vom Tricomi und Mitarbeitern. Ein fraktales Bild wurde während des gesamten Blocks gezeigt. Das ausgefüllte gelbe Quadrat zeigte an, welcher von 4 Knöpfen gedrückt werden sollte. Die Knopfdrücke konnten von den Probanden mit selbstgewählten Abständen ausgeführt werden. Nach einem nicht-belohnten Knopfdruck wurde ein grauer Kreis für 50 ms gezeigt. Eine Belohnung stand mit einer Wahrscheinlichkeit von 0,1 pro Sekunde zur Verfügung. Ein belohnter Knopfdruck zog die Präsentation eines Bildes einer M&M Schokolinse oder eines Frito Chips nach sich (Dauer 1 s), was die korrespondierende Belohnung, die nach dem Scan zur Verfügung stand, anzeigte.

Das Prinzip dieses Experimentes war also, dass eine Belohnung alle 10 Sekunden zur Verfügung stand (im Schnitt). Es gab nun 2 Gruppen von Probanden: Eine Gruppe von Probanden erhielt extensives Training über viele Versuchsblöcke, mehrere Tage lang, wohingegen die andere Gruppe kein solches Training erhielt.

Nun kommt der entscheidende Trick: Nach der letzten Trainingssitzung wurde einer der beiden Belohnungen (die M&Ms oder die Chips) entwertet. Dies wurde dadurch erreicht, dass die Probanden so lange die Belohnung konsumieren sollten, bis diese nicht mehr angenehm für sie war (sie sollten sich also „überfressen“). Welche der Belohnungen entwertet wurde, wurde jeweils zufällig festgelegt. Um die Effekte der Devaluierungsprozedur auf das Verhalten zu testen, wurden die Probanden nun wieder im Scanner untersucht für einen kurzen Test der Extinktion. Dieser Test erlaubt es, festzustellen, ob die Probanden eine Gewohnheit entwickelt hatten: Wenn das

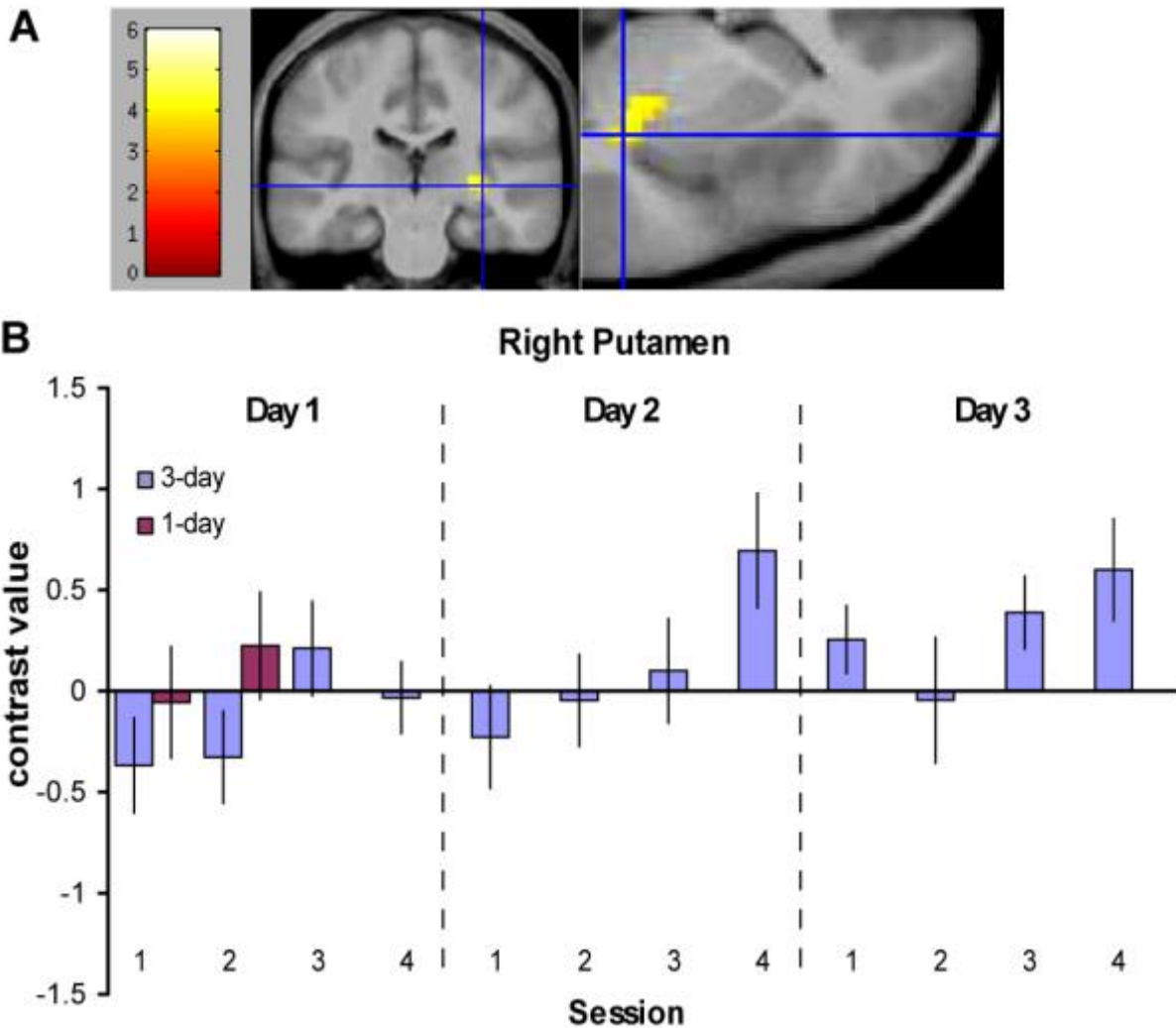
Verhalten noch im Sinne eines zielgerichteten Verhaltens (also abzielend auf den Erhalt einer Belohnung, die ja nun nichts mehr Wert war) ausgeprägt wäre, so würde man erwarten, dass die Probanden für das Fraktal, das die entwertete Belohnung anzeigt, weniger drücken, als für das Fraktal, das die noch wertvolle Belohnung anzeigt. Umgekehrt: Ist das Verhalten inzwischen zu einer Gewohnheit geworden, sollte sich keine Veränderung in der Frequenz der Knopfdrücke zeigen.

In der nächsten Abbildung ist das Resultat gezeigt:



Verhaltenseffekte: Im letzten Block vor der Devaluierung zeigten beide Gruppe, diejenige, die nur einen Tag trainiert wurde, und diejenige, die für drei Tage trainiert wurde, die gleichen Effekte: Für beide Belohnungen wurde mit einer gleich hohen Rate an Knopfdrücken geantwortet. Nach der Devaluierungsprozedure zeigte nur die 1-Tage Gruppe eine Abnahme der Knopfdrücke für die entwerteten Belohnungen. Bei der 3-Tage Gruppe hingegen war die Rate der Knopfdrücke genauso hoch wie für die nicht entwertete Belohnung. Es war also zu einer Ausbildung von Gewohnheiten gekommen.

In der nächsten Abbildung ist zunächst gezeigt, welche Region des Gehirns mit zunehmendem Training in ihrer Aktivität änderte.



Neuronale Korrelate der Gewohnheitsbildung: Es findet sich eine Zunahme der BOLD Antwort mit zunehmendem Training für die Aufgabenblöcke relativ zu so genannten Ruheblöcken in einer Struktur der Basalganglien dem rechten hinteren Putamen in der 3-Tage Gruppe (A). Verfolgt man die Entwicklung der Aktivität im rechten Putamen über die einzelnen Sitzungen an den verschiedenen experimentellen Tagen (B), so wird deutlich, dass die Aktivität in dieser Struktur über die Zeit zunimmt.

Was bedeutet dies für unser tägliches Verhalten?

Mit zunehmendem Training auf einer Aufgabe können die hierfür notwendigen Aktionen sich mehr und mehr von der Belohnung, die wir eigentlich für alle unsere Aktionen erstreben, abkoppeln, und zu einer belohnungs- und bestrafungsunabhängigen Gewohnheit werden. Dies ist manchmal gut: Ihr Hund setzt sich auf das Kommando „Sitz!“ hin, obwohl Sie ihm schon lange keine Belohnung mehr dafür geben. Andererseits fallen Ihnen bestimmt viele Situationen ein, wo Gewohnheiten unpassend, gesundheitsschädlich, nervtötend und kontraproduktiv

sind. Was die Studie von Frau Tricomi und Mitarbeitern zeigt, ist, dass es zu einem Shift vom ventralen Striatum, welches eher für zielgerichtetes (nicht gewohnheitsmäßiges) Handeln zuständig ist, zum dorsalen Striatum (speziell zum rechten hinteren Putamen) kommt, welches die Gewohnheiten unterstützt.

9. Neurobiologie des Momentanen Glücks

Das subjektive Wohlergehen von Menschen ist bedeutsam für unsere Gesellschaft. Obwohl „Glücklichsein“ durch äußere Lebensumstände wie z.B. die Demographie beeinflusst wird, wissen wir bisher nur wenig darüber, wie kleinere Ereignisse im täglichen Leben den momentanen Glückszustand modulieren. Könnte man nachweisen, dass es auf die Mischung und den Kontrast einzelner Ereignisse ankommt, wäre es möglich hier auch gezielt einzugreifen und die motivationale Kraft des Glücklichseins auszunutzen.

Eine kürzlich erschienene Studie von Rutledge und Mitarbeitern (2012) greift diese Idee auf. Diese Autoren setzten computationale Modelle sowie die funktionelle Kernspintomographie ein, um die emotionale Reaktivität von Probanden, operationalisiert als „momentanes Glücklichsein“, in einer probabilistischen Belohnungsaufgabe zu untersuchen. Als Hauptergebnis sei vorweggenommen, dass das momentane Glücklichsein **nicht** abhängig war von den in dieser Belohnungsaufgabe erworbenen Geldbeträgen, sondern durch eine Kombination von Belohnungserwartung und dem Prädiktionsfehler, der aus diesen Erwartungen resultierte, erklärt werden konnte.

Rutledge und Mitarbeiter präsentieren eine „Glücksformel“, die es ihnen erlaubt, zu berechnen, wie glücklich eine Person aufgrund einer Abfolge von Spielergebnissen ist. Zu dieser Glücksformel kommen wir gleich noch einmal zurück. Subjektives Glück ist offensichtlich abhängig von einer Reihe von Kontextvariablen und auch der individuellen Belohnungshistorie. Ein einfaches Beispiel: Stellen Sie sich vor, Sie wären Fan des FC Bayern München. Sollte ihr Verein in den nächsten 5 Spielen dreimal gewinnen, einmal unentschieden spielen und einmal verlieren, wäre ihr subjektives Glücksgefühl am Boden. Für mich als Fan von Hannover 96 wäre eine solche Serie indessen der Quell unerwarteter und langanhaltender Freude.

Diese unmittelbar einleuchtende subjektive Erfahrung haben Ruttledge und Mitarbeiter in eine experimentelle Situation umgesetzt, die zeigt, dass das momentane Glücklichein in einer probabilistischen Belohnungsaufgabe nicht von der Summe der in dieser Aufgabe gewonnen Punkte (oder einer gewonnenen Geldsumme) abhängt, sondern von einer Kombination aus Belohnungserwartung und Prädiktionsfehlern, die aus diesen Erwartungen resultieren.

Die Förderung des Glücklicheins wird als eine zentrale Aufgabe von Gesellschaften und der Politik angesehen. Der Philosoph Bentham hat vorgeschlagen, dass das Ziel sei „dass das größte Glück für die größte Zahl von Menschen ist das Maß dafür, was richtig und was falsch ist.“ Nach dieser Devise wird in vielen Nationen Politik gemacht und bewertet. Auf der anderen Seite ist Glück schwierig zu definieren, Es besteht offensichtlich keine unmittelbare Beziehung zum Wohlstand einer Person, so dass man meinen könnte, dass es keine Beziehung zwischen Belohnung und Glücklichein gibt. Aus dem, was wir bisher besprochen haben, ergibt sich aber, dass die Beziehung zwischen Belohnung und Glücklichein möglicherweise komplexer sein dürfte, als ein bloßer linearer Zusammenhang zwischen Belohnungsgröße und Glücksmenge. Auch intuitiv stellen wir fest, dass es häufig auf den Kontrast zwischen Erwartung und Realität ausmacht (Ein Lob von jemandem, von dem wir dieses nicht erwartet haben, freut uns ungleich mehr als ein Lob von jemandem, der gerne positiv verstärkt.).

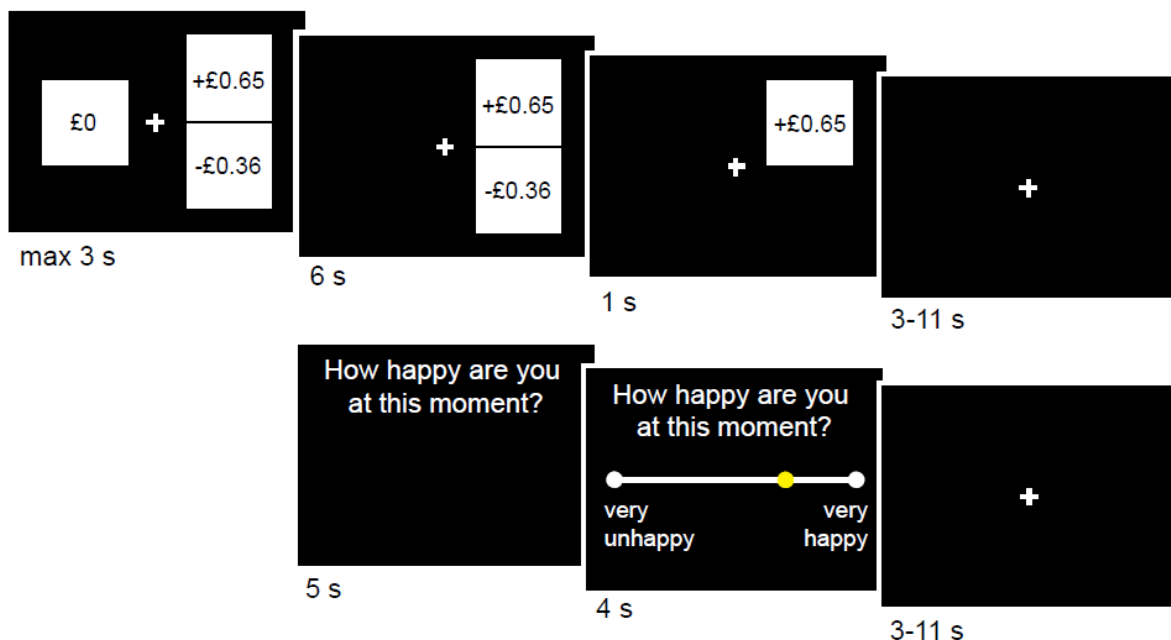
Wie kann man aber Glück berechnen und vorhersagen. Robb Ruttledge, mit dem ich seit kurzem zusammenarbeite, hat eine Studie hierzu durchgeführt, die großes Aufsehen erregt hat. Um die Sache einfacher zu machen, benutzte er quantifizierbare Belohnungen (=Geld), von denen wir wissen, dass sie affektive und motivationale Antworten produzieren. Er ging davon aus, dass für die Etablierung einer Beziehung zwischen Belohnung und Glücklichein die reliable Messung des subjektiven Glücks unbedingte Voraussetzung ist. Hierzu wandte er die Methode des „Experience Sampling“ an, mit der man den Zustand von Probanden abfragen kann, während sie sich „in der freien Wildbahn“ bewegen. Wenn man Probanden immer wieder nach ihrem momentanen Glückszustand fragt, kann man diesen zu zuvor erlebten Ereignissen (einschließlich Belohnungen) in Beziehung setzen. Es handelt sich hierbei also um Momentaufnahmen des Glücklicheins. Dies unterscheidet sich von Maßen der globalen Lebenszufriedenheit, die über einen langen Zeitraum (möglicherweise das gesamte Leben) integrieren.

In seiner Studie bat Rutledge seine Probanden an einer probabilistischen Belohnungsaufgabe teilzunehmen, die von diesen forderte, sich zwischen einer sicheren und einer risikobehafteten Spieloption (Lotterie) zu entscheiden. Alle paar trials bat er die Probanden anzugeben: „Wie glücklich sind Sie gerade?“.

Es wurde in dieser Untersuchung häufig das momentane Glücklichein abgefragt, um die Hypothese zu überprüfen, dass es zu einer raschen und von der Belohnungsgeschichte abhängigen Änderung des Glückszustandes kommen kann.

Wenn wir die bereits dargestellten Ergebnisse zum Einfluss von Dopamin auf Motivation und Belohnungsverarbeitung Revue passieren lassen, stellt sich die Frage, ob Dopamin auch mit dem momentanen Glücklichein zu tun haben könnte. Insbesondere stellte sich Robb Rutledge die Frage, ob die Aktivität der dopaminergen Neurone im Mittelhirn, die den Belohnungserwartungsfehler kodieren, auch mit dem momentanen Glücklichein zusammen hängen könnten.

Wie ging Rutledge nun vor? Zunächst führte er bei 26 Probanden eine Verhaltensstudie vor, bei der die Probanden sich jeweils zwischen einer Lotterie und einem sicheren Gewinn entscheiden mussten.



In der Abbildung entschied sich der Proband für die Lotterie, die mit 50 % Wahrscheinlichkeit einen Gewinn von 0,65 Pfund und einen Verlust von 0,36 Pfund

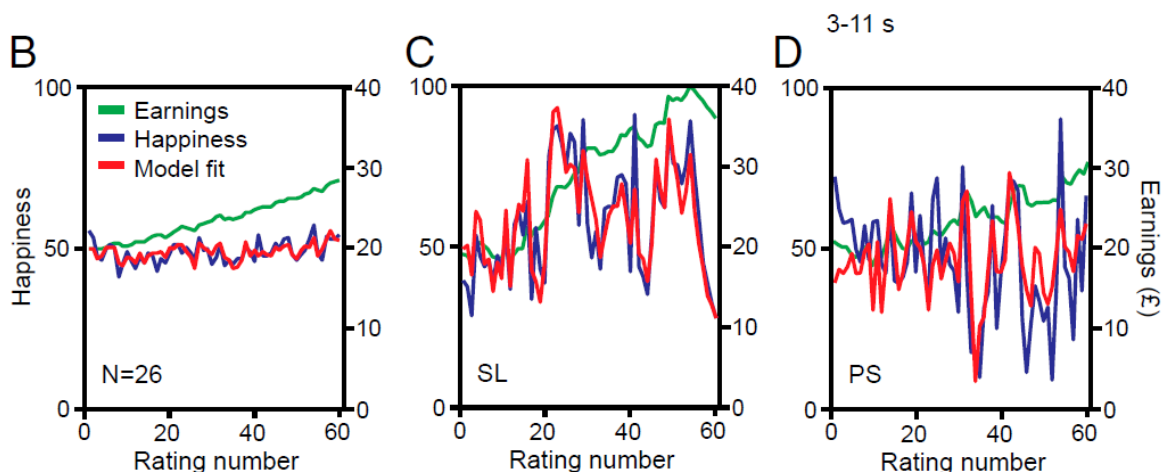
brachte. Wie aus dem dritten Bild ersichtlich ergab die Lotterie dann einen Gewinn. Danach erfolgte eine Abfrage.

Anhand der Ergebnisse dieses Versuchs wurde eine Glücksformel angepasst, die im Folgenden dargestellt ist:

$$\text{Happiness}(t) = w_0 + w_1 \sum_{j=1}^t \gamma^{t-j} CR_j + w_2 \sum_{j=1}^t \gamma^{t-j} EV_j + w_3 \sum_{j=1}^t \gamma^{t-j} RPE_j,$$

Hierbei ist CR der Betrag der sicheren Belohnung (Certain Reward), EV der Erwartungswert (Expected Value) und RPE der Belohnungsvorhersagefehler (Reward Prediction Error), also die Differenz zwischen dem erwarteten Gewinn und dem eingetretenen Gewinn. „ t “ entspricht der Versuchsnummer, w_0 ist eine Konstante und die anderen Gewichte w_1 , w_2 und w_3 bilden den Einfluss der verschiedenen Ereignistypen ab. „ γ “ kann zwischen 0 und 1 schwanken und ist ein Vergessensfaktor, der Ereignisse in kürzer zurückliegenden Versuchen stärker berücksichtigt als länger zurückliegende Versuche.

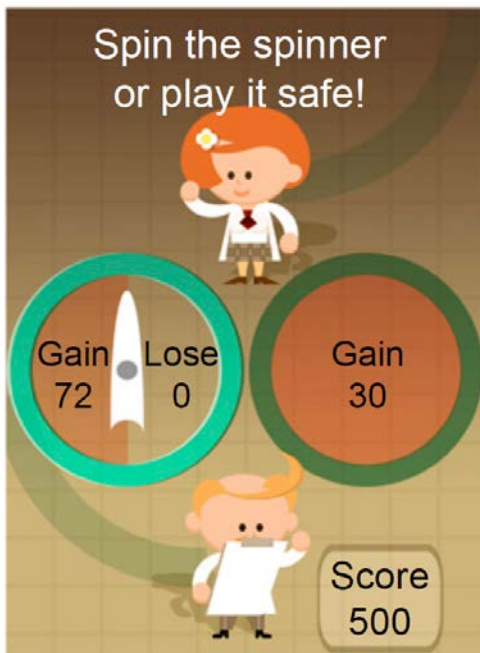
Es ergab sich bei dieser Untersuchung, dass durch dieses Modell das momentane Glück der Probanden sehr gut durch das Modell (die Glücksformel) vorhergesagt werden konnte, wie aus der nächsten Abbildung hervorgeht:



In der linken Abbildung (B) ist das Mittel aus 26 Probanden über insgesamt 60 Spiele dargestellt. In C und D finden sich die Ergebnisse von zwei individuellen Versuchspersonen. Ganz offensichtlich schwankt das momentane Glück im

individuellen Probanden sehr deutlich, kann aber durch die Glücksformel hervorragend vorhergesagt werden.

Nach einer Zwischenstudie mit 200 Probanden unternahm Rutledge eine weitere Untersuchung mit einem Smartphone-basierten Pendant des Experimentes, wie es in der nachfolgenden Abbildung dargestellt ist, an mehr als 18000 Probanden, wobei über 200000 Einzelmessungen des momentanen Glücks vorgenommen wurden



Auch bei dieser Untersuchung zeigte sich, dass die Glücksformel das durch die Belohnungshistorie im Experiment hervorgerufene Glück hervorragend abbildete.

Besonders wichtig ist, dass das Modell (enttäuschte und übererfüllte) Erwartungen mit berücksichtigt. Berechnet man ein Modell, welches nur auf den Belohnungen beruht, so kann dieses das momentane Glück nicht vorhersagen. Mit anderen Worten: Das Ausmaß der erhaltenen Belohnungen hat mit dem Glück nur wenig zu tun. Der Volksmund weiß das: Geld allein macht nicht glücklich.

Eine weitere Vorhersage dieses Modells ist, dass das endgültige Glück eines Probanden mehr durch die letzten Ereignisse als durch die ersten Ereignisse in einer Sequenz bestimmt wird. Hierzu ist der „Vergessensfaktor“ in die Glücksformel eingeführt worden und empirisch bestimmt worden.

Überlegen Sie sich beispielsweise, welchen Effekt der Erhalt von 1 Euro im aktuellen Versuchsdurchgang im Vergleich zum Erhalt von 1 Euro vor 5 Versuchsdurchgängen auf das momentane Glück haben sollte. Setzt man, wie empirisch bestimmt, einen

Vergessensfaktor von 0,61 an (Mittel aus 18000 Versuchspersonen), hätte der vor 5 Durchgängen erhaltene Gewinn nur 8 % so viel Einfluss auf das Glück wie der aktuelle Versuch. Der Vergessensfaktor schwankte bei den Untersuchungen von Rutledge etwa zwischen 0,4 und 0,8: Das bedeutet, dass der Einfluss eines 5 Durchgänge zurückliegenden Gewinns maximal 33 %, minimal nur 1 % des Einflusses des aktuellen Durchgangs hat.

Das am Ende des Versuchs gemessene Glück hängt also im Wesentlichen von den letzten Versuchsdurchgängen ab. In der Tat kann man das endgültige Glück eines Versuchsteilnehmers aus den letzten 10 Versuchsdurchgängen akkurat berechnen (wenn man die individuellen Parameter nimmt, die in den ersten 140 Versuchsdurchgängen bestimmt worden sind).

Überlegen Sie einmal zwischendurch:

Was hätte diese extrem kurzfristige Berechnung des momentanen Glücks im täglichen Leben für Auswirkungen?

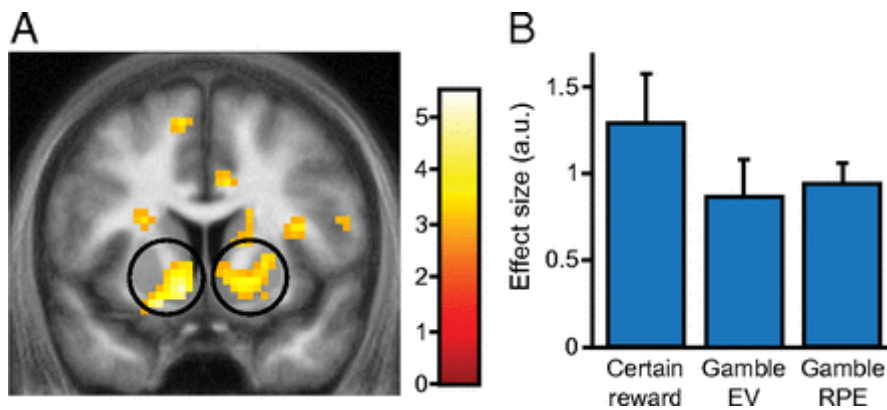
Wie könnte man (z.B. als Vorgesetzter) die hohe Vergessensrate von Mitarbeitern ausnutzen? Wie sollte man sie ausnutzen?

Wofür könnte diese kurzfristige Berechnung gut sein?

Ist die kurzfristige Berechnung möglicherweise ein phylogenetisches Erbe, welches in der heutigen Zeit eher Nachteile bringt?

In einem weiteren Schritt wurde ein Experiment mit der funktionellen Kernspintomographie durchgeführt. Hierbei wurde die so genannte BOLD Antwort im MRT korreliert mit den unmittelbar danach erhobenen Glücksratings. Es zeigte sich, dass die Aktivität im ventralen Striatum signifikant mit den zukünftigen Glücksratings korreliert war.

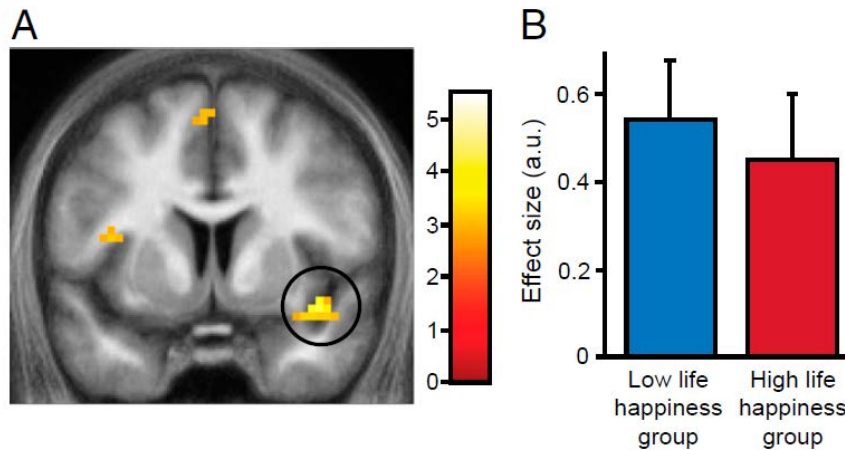
Dies ist schön im linken Teil der folgenden Abbildung zu sehen:



In einem weiteren Schritt wurde getestet, inwieweit Modellparameter die Aktivität im ventralen Striatum bestimmen. Dies ist im rechten Teil der Abbildung dargestellt. Es zeigt sich, dass die Aktivität des ventralen Striatums durch die einzelnen Teile der Glücksformel in einer Weise beeinflusst werden, die suggeriert, dass das ventrale Striatum in der Tat ein Computer zur Berechnung des momentanen Glücks ist. Wurde indessen versucht, den Einfluss von vergangenen Ereignissen auf die momentane Aktivität des ventralen Striatums zu berechnen, ergab sich kein signifikanter Effekt, was dafür spricht, dass im ventralen Striatum genau die Berechnungen durchgeführt werden, die das momentane Glück repräsentieren.

Wenn das Striatum der Computer ist, der Berechnungen zum momentanen Glück anstellt, stellt sich die Frage, welche Hirnstrukturen dann das momentane Glück repräsentieren? Hierzu bestimmte man die Hirnantwort zu dem Zeitpunkt, an dem die Probanden gefragt wurden, wie glücklich sie in diesem Moment gerade waren.

Interessanterweise war die BOLD-Aktivität im ventralen Striatum nicht mit der momentanen Glückseinschätzung korreliert. Vielmehr zeigte sich eine Korrelation der Aktivität in der rechten Inselregion mit der Glückseinschätzung. Dies ist in der nächsten Abbildung gezeigt.



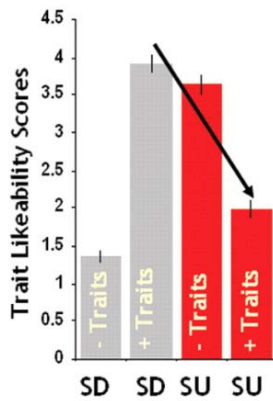
Der Befund einer Beziehung zwischen dem subjektiven Glücksgefühl und insulärer Hirnaktivität ist nicht wirklich überraschend. Andere Untersuchungen haben gezeigt, dass dieses Hirnareal interozeptive Funktionen besitzt, also Rückmeldungen aus dem Körper verarbeitet und in Handlungspräferenzen umsetzt sowie auch für die Bewusstwerdung von Emotionen wichtig ist. Es ist auch gezeigt worden, dass Personen, die sich insgesamt als glücklich erleben, in dieser Region ein größeres Volumen der grauen Substanz des Gehirns aufweisen. Bei solchen Beziehungen stellt sich immer die "Henne-Ei"-Frage: Sind Menschen glücklich, weil sie ein größeres Volumen der grauen Substanz in der vorderen Insel haben oder ist es umgekehrt: Ist durch das dauernde Glücklichkeit das Volumen der Insel angewachsen? Rutledge und Mitarbeiter haben sich daher gefragt, ob sich vielleicht die Antworten in der anterioren Insel zwischen Probanden mit einem größeren und geringerem globalem Glücksgefühl voneinander unterscheiden. Hierzu wurden die Probanden vor dem Experiment gefragt: "Wenn man alles zusammennimmt: Wie glücklich sind Sie generell mit Ihrem Leben?" Es zeigte sich für global glückliche und global weniger glückliche Probanden insgesamt eine vergleichbare Antwort der Insel in Bezug auf das momentane Glück. Man kann daraus schließen, dass es offensichtlich psychologische und neuronale Unterschiede zwischen momentanem und eher globalem Glücklichkeit gibt.

10. Belohnung durch Beobachtung des Glücks anderer Personen?

Seit gefühlt mehreren Jahrzehnten schauen sich Millionen von Menschen wöchentlich die Show „Wer wird Millionär“ an. Wer diese nicht mag, schaut Pilawa, „Schlag den Raab“ oder andere Formate an. Offensichtlich ist es so, dass wir Menschen eine

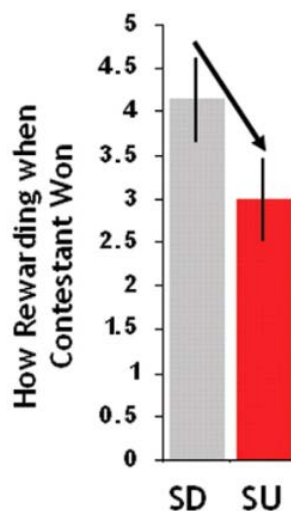
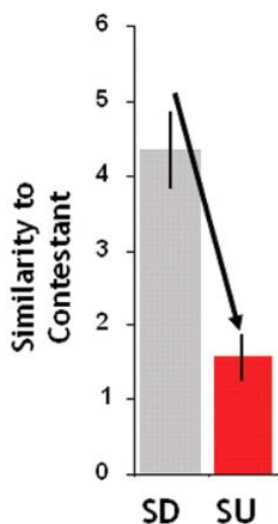
eingebaute prosoziale Tendenz haben, die uns dazu bringt, uns am Erfolg anderer zu freuen. Der Erfolg derartiger Quiz- oder Spielshows scheint eindrücklich zu belegen, dass dies so ist. Es bleibt aber – auch aus evolutionärer Sicht, unklar, warum wir das Gefühl einer Belohnung erfahren, obwohl wir keinerlei eigenen ökonomischen oder sonstigen Vorteil hiervon haben. Eine Hypothese in diesem Zusammenhang, die von einer Arbeitsgruppe um Dean Mobbs aus London verfolgt wurde, könnte sein, dass die Spieler in solchen Shows häufig große Ähnlichkeiten mit den Zuschauern haben, also den gleichen sozialen, kulturellen und ökonomischen Hintergrund haben, was dazu führen könnte, dass gruppen-motivierte Prozesse zu der Belohnungserfahrung führen können (Die alternative Hypothese, dass viele Menschen diese Shows schauen, um andere verlieren zu sehen, ist von Dean Mobbs bisher noch untersucht worden).

Sozial-kognitive Überlegungen legen nahe, dass wir, um die internen Zustände eines anderen zu simulieren, uns diesem anderen ähnlich fühlen müssen. In der Tat: Sich in die Haut eines anderen zu versetzen, erfordert eine gewisse Ähnlichkeit. Mobbs und Mitarbeiter überprüften nun zwei Hypothesen: Zum einen postulierten sie, dass das Betrachten eines sozial-kompatiblen Spielers die neuronalen Strukturen, die gewöhnlich mit Belohnungserfahrungen assoziiert sind, aktiviert. Zum anderen wurde postuliert, dass das Ausmaß dieser Aktivierungen mit der wahrgenommenen Ähnlichkeit zwischen der Versuchsperson und dem betrachteten Spieler zusammenhängt. Zu diesem Zweck adaptierten Mobbs und Mitarbeiter ein Game-Show-Format für die Präsentation in einem MR-Scanner. Die Probanden sahen zunächst Filme in denen zwei Schauspieler, die als Game-Show-Kandidaten auftraten, Fragen über persönliche, soziale und ethische Themen beantworteten. Einer der „Kandidaten“ spielte dabei einen sozial-erwünschten Charakter mit empathischen Eigenschaften (SD=socially desirable), wohingegen der andere sich eher unsozial präsentierte (SU=socially undesirable). In der Tat führte dies dazu, dass die Probanden den SD-„Kandidaten“ in einem Beliebtheits-Rating als sehr viel positiver einschätzten als den SU-„Kandidaten“ (siehe nächste Abbildung).



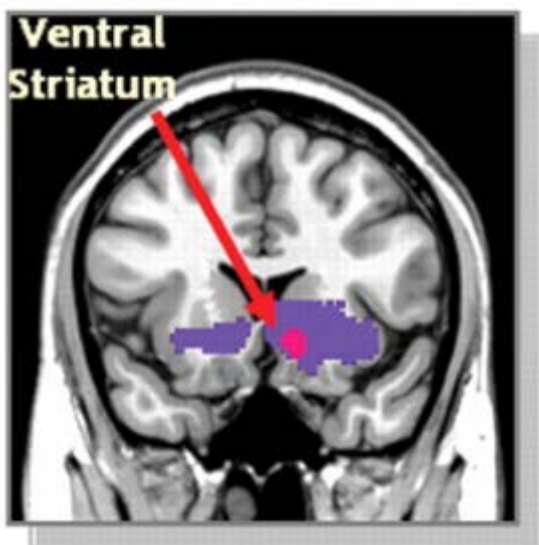
In der nächsten Phase des Experimentes wurden die Probanden in einem MR-Scanner untersucht. Sie betrachteten dabei die SD- und SU-„Kandidaten“, wie diese ein einfaches Spiel spielten: Sie mussten entscheiden, ob eine nicht sichtbare Spielkarte einen höheren oder niedrigeren Wert hat, als eine zweite, ebenfalls nicht sichtbare Karte. (Diese Spielzüge waren natürlich von den Experimentatoren von vornherein festgelegt, so dass eine definierte Anzahl von Gewinnen und Verlusten für die „Kandidaten“ resultierte). Eine korrekte Entscheidung eines „Kandidaten“ führte dazu, dass dieser £5 gewann. Nachdem die Probanden den „Kandidaten“ zugesehen hatten, durften sie das Spiel auch selbst spielen.

Was kam heraus? Die subjektiven Einschätzungen der Probanden nach dem Experiment zeigten, dass diese sich dem SD-„Kandidaten“ ähnlicher fühlten (nächste Abbildung links) und es belohnender empfanden, wenn der SD-„Kandidat“ gewann (nächste Abbildung rechts).

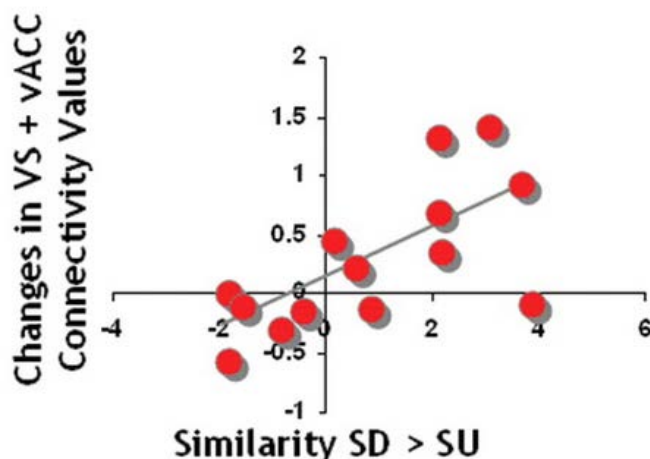


Ferner fanden sich Korrelationen zwischen den Ähnlichkeits-Ratings und den Einschätzungen, wie belohnend die Probanden es fanden, zu betrachten, wenn einer der „Kandidaten“ gewann.

Die Hirnaktivierungsdaten zeigten zunächst und nicht überraschend, dass selbst erlebte Belohnungen zu einer signifikanten Aktivierung des ventralen Striatum führten. Dies ist in der nächsten Abbildung in lila gezeigt. Ebenfalls kam es zu einer Aktivierung des ventralen Striatums, wenn man betrachtete, wie belohnend es empfunden wurde, den SD-„Kandidaten“ (im Vergleich zum SU-„Kandidaten“) gewinnen zu sehen. Dies ist in der Abbildung in pink dargestellt.



Eine so genannte psychophysiologische Interaktion zeigte darüber hinaus noch eine Modulation der Konnektivität zwischen dem ventralen Striatum und dem vorderen Teil des Gyrus cinguli als Funktion von (SD gewinnt versus SU gewinnt).



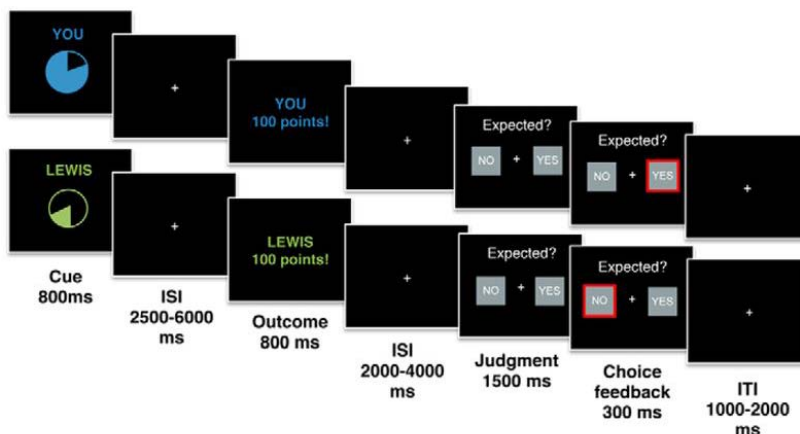
Was bedeutet dies im Zusammenhang mit unserem heutigen Thema Motivation? Bisher hatten sich Forscher darum bemüht, die neuronalen Repräsentationen der mentalen Zustände von anderen (Empathie bei Schmerz) in Gehirnen von Beobachtern darzustellen. Dabei hatte man zum Beispiel gefunden, dass die so genannte Schmerzmatrix auch dann aktiviert wird, wenn man beobachtet, dass jemand anderem Schmerzen zugefügt werden. Die Untersuchungsbefunde beim Beobachten und Erleiden von Schmerzen waren also vergleichbar. Die vorliegende Studie geht einen Schritt weiter und untersucht, wie wir reagieren, wenn jemandem etwas Gutes widerfährt. Offensichtlich schwingen wir auch da mit, allerdings mehr, wenn wir denjenigen, der die Belohnung erhält mögen (der SD-„Kandidat“). Wie die Untersucher weiter zeigten, spielt die wahrgenommene Ähnlichkeit des Kandidaten mit dem Probanden eine Rolle für die Aktivität im vorderen Teil des Gyrus cinguli, einer Region, die als wichtig für die Selbstrelevanz bekannt ist.

In einer im Oktober 2015 erschienenen Studie gehen Lockwood und Mitarbeiter noch einen Schritt weiter: Sie stellen die Frage, ob unsere Fähigkeit, auf Belohnungen, die anderen widerfahren, zu reagieren, mit dem Ausmaß an Empathiefähigkeit zusammenhängt. Sie gehen davon aus, dass sich soziale Situationen dadurch auszeichnen, dass Stimuli Belohnungen voraussagen, und zwar nicht nur für uns selbst, sondern auch für andere. Um effektiv kooperieren, wettstreiten und mitfühlen zu können, müssen wir in der Lage sein, Belohnungen für andere vorauszuberechnen. Allerdings ist bisher nur wenig darüber bekannt, wie Belohnungen für andere vorausgesagt werden und wie dies im Gehirn geschieht. Es stellt sich darüber hinaus die Frage, inwiefern individuelle Unterschiede in der Fähigkeit, sich in andere einzufühlen, hier eine Rolle spielen können.

Eine Struktur, die hier in besonderer Weise wichtig ist, ist der Gyrus cinguli, der Informationen über in der Zukunft liegende Belohnungen integriert und Vorhersagen, z.B. über die Wahrscheinlichkeit und die Größe einer Belohnung, trifft. Darüber hinaus ist diese Region auch aktiv, wenn soziale Informationen verarbeitet werden. Nach einem von Appss und Mitarbeitern vorgeschlagenen Modell, nach dem eine Region des Gyrus cinguli empfindlich auf Informationen über Belohnungen für andere Personen reagiert und darüber hinaus Berechnungen über die Wahrscheinlichkeit, dass eine andere Person eine solche Belohnung erhält, anstellt. In der Tat ist gezeigt worden, dass bei Affen in dieser Region Zellen vorhanden sind, die dann feuern, wenn ein anderer Affe eine Belohnung erhält. In einer anderen Studie an Affen wurden diese

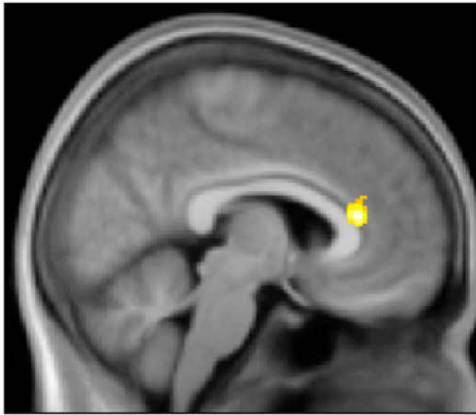
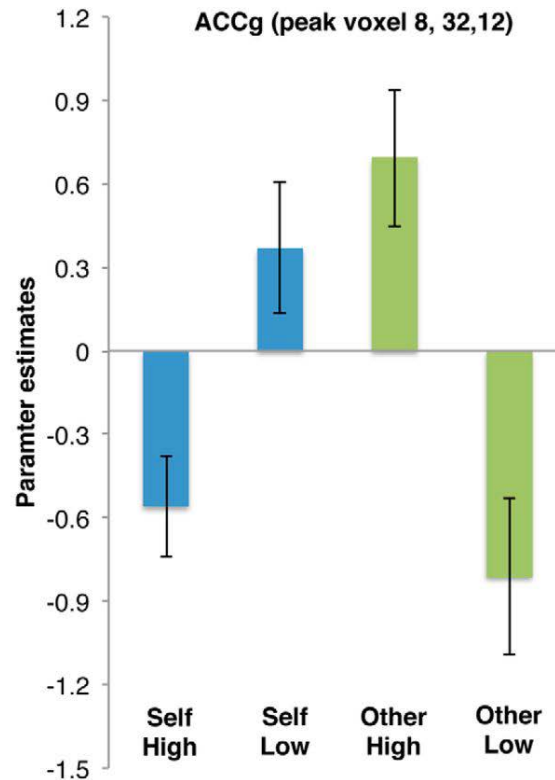
trainiert, an einem ökonomischen Spiel teilzunehmen. Die Aktivität von spezifischen Nervenzellen im Gyrus cinguli sagte voraus, welche Entscheidungen ein Artgenosse in diesem Spiel treffen würde. Beim Menschen ist gezeigt worden, dass Läsionen (z.B. im Rahmen von Schlaganfällen) des Gyrus cinguli den Wert, der sozialen Stimuli beigemessen wird, reduzieren. Demgegenüber bleibt die Verarbeitung von nichtsozialen Stimuli intakt. Schließlich ist auch gezeigt worden, dass die Aktivität dieser Region in der funktionellen Kernspintomographie den Netto-Wert von Belohnungen, die von anderen erhalten werden, reflektiert. Gleichfalls zeigte diese Region auch eine Sensitivität gegenüber Somit erscheint der Gyrus cinguli zentral für die Verarbeitung von sozialen Informationen einschließlich der Belohnungen, die andere Personen erhalten. Eine bisher noch nicht überprüfte Hypothese ist jedoch, ob der ACC auch beim Menschen Vorhersagen über die Wahrscheinlichkeit, mit der eine andere Person eine Belohnung erhält, anstellt. Da offensichtlich nach dem Modell von Apps der Gyrus cinguli auch besonders für die Verarbeitung von sozialen Informationen prädestiniert ist, ergibt sich die Frage, ob die individuelle Fähigkeit, Empathie mit anderen zu zeigen, die Antworten in dieser Hirnregion moduliert.

In dem von Lockwood und Mitarbeitern durchgeführten Experiment wurde folgender Versuchsaufbau durchgeführt:



Es war also so, dass entweder der Proband selbst, oder eine andere Person („Lewis“) um eine Belohnung, die mit einer bestimmten (durch einen Hinweisreiz angezeigten) Wahrscheinlichkeit eintreten konnte, spielten.

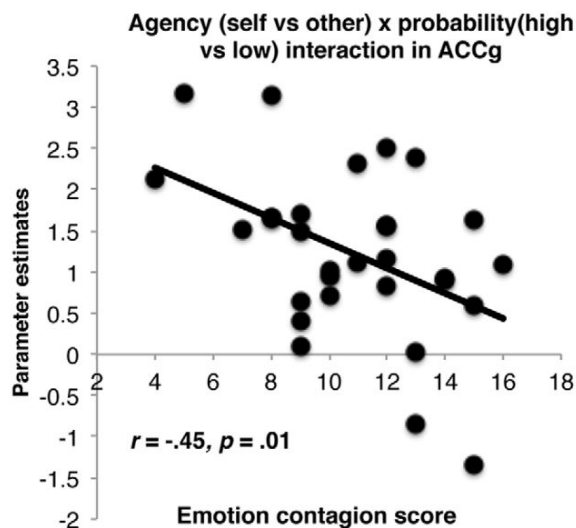
Die MRT-Aktivierungen sind in der nächsten Abbildung gezeigt:

A**B**

Auf der linken Seite sieht man den Interaktionseffekt zwischen dem Faktor „Agency“ (selbst oder andere Person) und der Belohnungswahrscheinlichkeit (high [80%] oder low [20%]) zum Zeitpunkt des Hinweisreizes.

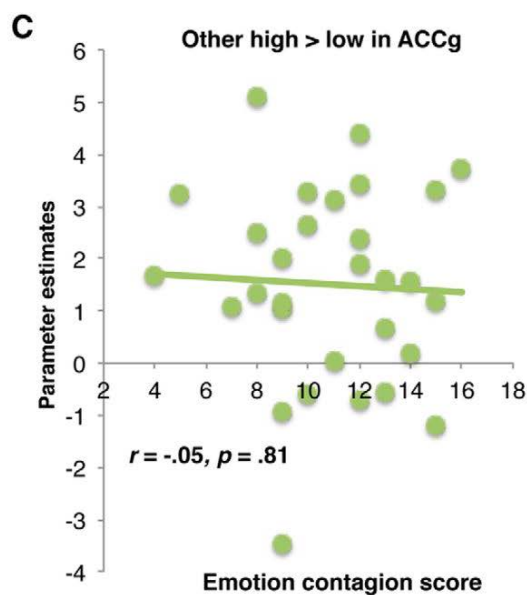
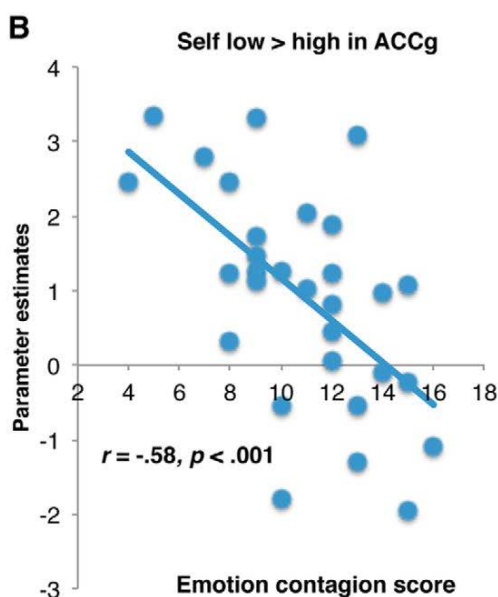
Es zeigt sich also, dass der Gyrus cinguli auf eigene und fremde Belohnungen in gewisser Weise spiegelbildlich reagiert, was zu der Interaktion führt, die sich in der Balkengraphik abbildet.

Die weitere Frage, die die Forscher um Lockwood beantworten wollten, war, ob die individuelle Empathiefähigkeit die Antwort des Gyrus cinguli modulieren kann. Bei den Probanden wurde diese mithilfe eines Fragebogens erfasst, der einen so genannten „emotional contagion“ Score lieferte, also so etwas wie die emotionale Ansteckbarkeit. Die Beziehung zwischen dem Interaktionseffekt im Gyrus cinguli und der individuellen emotionalen Ansteckbarkeit ist in der nächsten Abbildung dargestellt.



Wie man sieht, ergibt sich hier ein korrelativer Zusammenhang, der besagt, dass ein höherer emotionaler Ansteckbarkeitsscore mit einer schwächeren Interaktion im Gyrus cinguli einhergeht. Was könnte dies bedeuten?

Um diese Frage zu beantworten, wurden weitere Korrelationen berechnet, die in der nächsten Abbildung dargestellt sind:



Es zeigte sich in der Zusammenschau, dass in Individuen mit hohem emotionalen Ansteckbarkeitsscore der Gyrus cinguli Informationen über die relative Differenz zwischen hoch- und niedrigwahrscheinlichen Gewinnen für andere Personen widerspiegelte, wohingegen in Personen mit niedrigem Ansteckbarkeitsscore der Gyrus cinguli die Belohnungswahrscheinlichkeit für das Individuum selbst kodierte.

Eine wesentliche Erkenntnis aus diesem Experiment ist, dass die Rolle des Gyrus cinguli in der Vorhersage von Belohnungen für sich selbst und andere von der emotionalen Ansteckbarkeit, also von der Empathiefähigkeit für andere, abhängt. Die emotionale Ansteckbarkeit wird als wesentliche Voraussetzung für die Empathiefähigkeit gesehen. In diesem Zusammenhang ist es bedeutend, dass die emotionale Ansteckbarkeit auch mit den Reaktionszeiten für Entscheidungen, die die Belohnungen für andere betrafen, korrelierte in dem Sinne, dass die Probanden mit dem höchsten emotionalen Ansteckbarkeitsscore die schnellsten Reaktionszeiten aufwiesen.

Betrachtet man die menschliche Fähigkeit zur Empathie, wird oftmals eine affektive Komponente (also die Fähigkeit zu einer affektiven Resonanz mit dem Zustand eines anderen Menschen) von einer kognitiven Komponente (die Fähigkeit, die Intentionen und Gefühle eines anderen intellektuell zu verstehen) unterschieden. Die Regressionsanalysen, die von Lockwood und Mitarbeitern für die vorliegende Studie durchgeführt worden sind, sprechen dafür, dass es die affektive Komponente der Empathiefähigkeit ist, die mit der Vorhersage von Belohnungen für andere assoziiert ist.

Empathische Fähigkeiten sind essentiell für erfolgreiches Sozialverhalten. Umgekehrt liegen mangelnde Empathiefähigkeiten vielen psychiatrischen Störungen zugrunde, so z.B. der Psychopathie oder dem Autismus. In der Tat konnte gezeigt werden, dass der Gyrus cinguli bei Personen mit Autismus oder mit einer antisozialen Persönlichkeitsstörung atypisch aktiviert wird, vor allem in Situationen, die die Beurteilung des mentalen Zustandes einer anderen Person erfordern.

Die Studie von Lockwood legt nahe, dass eine Störung in der Funktionsfähigkeit des Gyrus cinguli den Sozialverhaltensstörungen zugrundeliegt und dass diese Störung in der unzureichenden Berechnung des Verhaltens anderer Menschen liegt. Der nächste Schritt wäre nun, diese Forschungen aus dem Labor in die Praxis zu holen. Kann man mit den hier skizzierten Methoden beispielsweise das Verhalten von Personen in realweltlichen Situationen vorhersagen?

Literatur

- Al-Khaled M, Heldmann M, Bolstorff I, Hagenah J, Münte TF. Intertemporal choice in Parkinson's disease and restless legs syndrome. *Parkinsonism Relat Disord*. 2015 Sep 25. pii: S1353-8020(15)00401-0.
- Buades-Rotger M, Brunnlieb C, Münte TF, Heldmann M, Krämer UM. Winning is not enough: ventral striatum connectivity during physical aggression. *Brain Imaging Behav*. 2015 Mar 11.
- Camara E, Krämer UM, Cunillera T, Marco-Pallarés J, Cucurell D, Nager W, Mestres-Missé A, Bauer P, Schüle R, Schöls L, Tempelmann C, Rodriguez-Fornells A, Münte TF. The effects of COMT (Val108/158Met) and DRD4 (SNP -521) dopamine genotypes on brain activations related to valence and magnitude of rewards. *Cereb Cortex*. 2010 Aug;20(8):1985-96.
- Camara E, Rodriguez-Fornells A, Münte TF. Functional connectivity of reward processing in the brain. *Front Hum Neurosci*. 2009 Jan 16;2:19. doi: 10.3389/neuro.09.019.2008.
- Camara E, Rodriguez-Fornells A, Münte TF. Microstructural brain differences predict functional hemodynamic responses in a reward processing task. *J Neurosci*. 2010 Aug 25;30(34):11398-402.
- Camara E, Rodriguez-Fornells A, Ye Z, Münte TF. Reward networks in the brain as captured by connectivity measures. *Front Neurosci*. 2009 Dec 15;3(3):350-62.
- de Quervain DJ, Fischbacher U, Treyer V, Schellhammer M, Schnyder U, Buck A, Fehr E. The neural basis of altruistic punishment. *Science*. 2004 Aug 27;305(5688):1254-8.
- Doñamayor N, Marco-Pallarés J, Heldmann M, Schoenfeld MA, Münte TF. Temporal dynamics of reward processing revealed by magnetoencephalography. *Hum Brain Mapp*. 2011 Dec;32(12):2228-40.
- Doñamayor N, Schoenfeld MA, Münte TF. Magneto- and electroencephalographic manifestations of reward anticipation and delivery. *Neuroimage*. 2012 Aug 1;62(1):17-29.
- Fehr E, Gächter S. Altruistic punishment in humans. *Nature*. 2002 Jan 10;415(6868):137-40.
- Flores A, Münte TF, Doñamayor N. Event-related EEG responses to anticipation and delivery of monetary and social reward. *Biol Psychol*. 2015 Jul;109:10-9.
- Heldmann M, Berding G, Voges J, Bogerts B, Galazky I, Müller U, Baillot G, Heinze HJ, Münte TF. Deep brain stimulation of nucleus accumbens region in alcoholism affects reward processing. *PLoS One*. 2012;7(5):e36572.
- Killgore WD, Young AD, Femia LA, Bogorodzki P, Rogowska J, Yurgelun-Todd DA. Cortical and limbic activation during viewing of high- versus low-calorie foods. *Neuroimage*. 2003 Aug;19(4):1381-94.
- Lockwood PL, Apps MA, Roiser JP, Viding E. Encoding of Vicarious Reward Prediction in Anterior Cingulate Cortex and Relationship with Trait Empathy. *J Neurosci*. 2015 Oct 7;35(40):13720-7.
- Marco-Pallarés J, Krämer UM, Strehl S, Schröder A, Münte TF. When decisions of others matter to me: an electrophysiological analysis. *BMC Neurosci*. 2010 Jul 29;11:86.
- Marco-Pallarés J, Mohammadi B, Samii A, Münte TF. Brain activations reflect individual discount rates in intertemporal choice. *Brain Res*. 2010 Mar 12;1320:123-9.
- Marco-Pallarés J, Münte TF, Rodríguez-Fornells A. The role of high-frequency oscillatory activity in reward processing and learning. *Neurosci Biobehav Rev*. 2015 Feb;49:1-7.
- McClure SM, Laibson DI, Loewenstein G, Cohen JD. Separate neural systems value immediate and delayed monetary rewards. *Science*. 2004 Oct 15;306(5695):503-7.
- Miedl SF, Wiswede D, Marco-Pallarés J, Ye Z, Fehr T, Herrmann M, Münte TF. The neural basis of impulsive discounting in pathological gamblers. *Brain Imaging Behav*. 2015 Feb 3. [Epub ahead of print]
- Milenkova M, Mohammadi B, Kollwe K, Schrader C, Fellbrich A, Wittfoth M, Dengler R, Münte TF. Intertemporal choice in Parkinson's disease. *Mov Disord*. 2011 Sep;26(11):2004-10.
- Mobbs D, Yu R, Meyer M, Passamonti L, Seymour B, Calder AJ, Schweizer S, Frith CD, Dalglish T. A key role for similarity in vicarious reward. *Science*. 2009 May 15;324(5929):900.
- Rushworth MF, Kolling N, Sallet J, Mars RB. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr Opin Neurobiol*. 2012 Dec;22(6):946-55.
- Rutledge RB, Skandali N, Dayan P, Dolan RJ. A computational and neural model of momentary subjective well-being. *Proc Natl Acad Sci U S A*. 2014 Aug 19;111(33):12252-7.

Rutledge RB, Skandali N, Dayan P, Dolan RJ. Dopaminergic Modulation of Decision Making and Subjective Well-Being. *J Neurosci*. 2015 Jul 8;35(27):9811-22.

Small DM, Zatorre RJ, Dagher A, Evans AC, Jones-Gotman M. Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain*. 2001 Sep;124(Pt 9):1720-33.

Tricomi E, Balleine BW, O'Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci*. 2009 Jun;29(11):2225-32.

Volkow ND, Baler RD. NOW vs LATER brain circuits: implications for obesity and addiction. *Trends Neurosci*. 2015 Jun;38(6):345-52.

Vrticka P, Andersson F, Grandjean D, Sander D, Vuilleumier P. Individual attachment style modulates human amygdala and striatum activation during social appraisal. *PLoS One*. 2008 Aug 6;3(8):e2868.

Wrase J, Kahnt T, Schlagenhauf F, Beck A, Cohen MX, Knutson B, Heinz A. Different neural systems adjust motor behavior in response to reward and punishment. *Neuroimage*. 2007 Jul 15;36(4):1253-62.

Ye Z, Hammer A, Camara E, Münte TF. Pramipexole modulates the neural network of reward anticipation. *Hum Brain Mapp*. 2011 May;32(5):800-11. doi: 10.1002/hbm.21067.