

**Der Einfluss von Kognition und Stresserleben auf die  
Entscheidungsfindung und deren Rolle für Alkoholabhängigkeit**

Kumulative Dissertation zur Erlangung der Würde des Dr. phil.  
Vorgelegt an der Humanwissenschaftlichen Fakultät der Universität Potsdam  
Strukturbereich Kognitionswissenschaften  
Sozial- und Präventivmedizin

08.06.2016

von  
Eva Friedel

**Gutachter**

1. Gutachter: Prof. Dr. Dr. Michael Rapp
2. Gutachter: Prof. Dr. Werner Strik

Online veröffentlicht auf dem  
Publikationsserver der Universität Potsdam:  
URN urn:nbn:de:kobv:517-opus4-91609  
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-91609>

## **Danksagung**

Mein außerordentlicher Dank gilt Prof. Dr. Dr. Michael Rapp für die Betreuung dieser Arbeit und den uneingeschränkten Rückhalt bei der Überwindung methodischer, inhaltlicher und vieler weiterer Herausforderungen während der gemeinsamen Zeit. Sein methodisches Wissen und die Kenntnisse zum Thema Lernen und Kognition waren wegweisend und bilden die theoretische und praktische Grundlage der durchgeführten Studien.

Mein besonderer Dank gilt außerdem Prof. Dr. Dr. Andreas Heinz und Dr. Florian Schlagenhaut für die intensive Förderung und Unterstützung meiner wissenschaftlichen und klinischen Tätigkeit an der Charité. Als Mentoren standen Sie mir zu jeder Zeit zur Seite und halfen, die Schwierigkeiten und Herausforderungen der letzten Jahre anzunehmen und zu meistern.

Ich möchte mich bei Prof. Dr. Dr. Henrik Walter für seine Hilfestellung und Förderung bedanken, sowie bei der gesamten Arbeitsgruppe Lernen und Kognition und Emotional Neuroscience, insbesondere bei meinen Kollegen Stefan Paul Koch, Claudia Gählsdorf, Miriam Sebold, Maria Garbusow, Lorenz Deserno und Claudia Hägele.

Abschließend gilt mein Dank meiner Familie und meinem Freund, ohne sie wären mein Weg und diese Arbeit nicht möglich gewesen.

Diese Arbeit wurde auf Basis der Richtlinien für die Anfertigung einer kumulativen Dissertation der Humanwissenschaftlichen Fakultät der Universität Potsdam Stand 24.6.2014 verfasst und orientiert sich an den Vorgaben des Leitfadens zur Gestaltung schriftlicher Arbeiten vom 22.02.2010 des Departments für Sport- und Gesundheitswissenschaften.

## Abkürzungsverzeichnis

AAT	Approach Avoidance Task
fMRT	funktionelle Magnetresonanztomographie
GR	Glukokortikoidrezeptor
HPA-Achse	Hypothalamus-Hypophysen-Nebennierenrinden-Achse
IQ	Intelligenzquotient
mPFC	medialer Präfrontalkortex/medial prefrontal cortex
PE	Prädiktionsfehler/Prediction Error
PET	Positron Emissions Tomographie
PIT	Pavlovian-to-Instrumental-Transfer
SARSA	State-Action-Response-State-Action
VS	Ventrales Striatum
VTA	Ventral Tegmentale Area/Ventrales Tegmentum

## Inhalt

<b>ZUSAMMENFASSUNG</b>	<b>1</b>
1.1 HINTERGRUND	1
1.2 ZIELSTELLUNG	4
1.3 METHODEN	5
1.4 ERGEBNISSE	6
1.5 DISKUSSION	8
<b>THEORETISCHER HINTERGRUND</b>	<b>11</b>
1.6 DIE PSYCHOLOGISCHEN UND NEUROBIOLOGISCHEN GRUNDLAGEN VON BELOHNUNGSABHÄNGIGEM LERNEN	11
1.7 DER EINFLUSS VON STRESS UND KOGNITION AUF BELOHNUNGSABHÄNGIGES LERNEN UND SEINE NEURONALEN KORRELATE	13
1.8 DER EINFLUSS VON LERNEN UND STRESS AUF DIE ENTSTEHUNG UND AUFRECHTERHALTUNG DER ALKOHOLABHÄNGIGKEIT	15
<b>DIE STUDIEN</b>	<b>18</b>
1.9 STUDIE 1: DEVALUATION UND SEQUENTIELLE ENTSCHEIDUNGSFINDUNG: DIE VERBINDUNG ZWISCHEN ZIELORIENTIERTEM UND MODELLBASIERTEM VERHALTEN	18
1.10 STUDIE 2: DER ZUSAMMENHANG ZWISCHEN CHRONISCHEM STRESSERLEBEN, DEM NEURONALEN PRÄDIKTIONSFEHLERSIGNAL UND INTELLIGENZ	23
1.11 STUDIE 3: WIE WIRD SUCHTVERHALTEN „ERLERNT“? RISIKOFAKTOREN UND LERNMECHANISMEN DIE ZUR ENTSTEHUNG UND AUFRECHTERHALTUNG DER ALKOHOLABHÄNGIGKEIT BEITRAGEN	26
1.12 DISKUSSION UND AUSBLICK	28
<b>LITERATURVERZEICHNIS</b>	<b>33</b>
<b>SELBSTSTÄNDIGKEITSERKLÄRUNG</b>	<b>40</b>
<b>ERKLÄRUNG ZUM PROMOTIONSVERFAHREN</b>	<b>40</b>
<b>EINGESCHLOSSENE PUBLIKATIONEN</b>	<b>41</b>

## Zusammenfassung

### 1.1 Hintergrund

Das Lernen aus den Konsequenzen eines Verhaltens und die darauf folgende Anpassung zukünftiger Verhaltensmuster sind ein wesentlicher Bestandteil unserer alltäglichen (Über-)Lebensfähigkeit. In seinem „Law of Effect“ schlägt Thorndike (1911) vor, dass positive und negative Konsequenzen von Verhalten einen unterschiedlichen Effekt haben: Die Belohnung eines spezifischen Verhaltens erhöht die Wahrscheinlichkeit seines erneuten Ausübens, während die Bestrafung eines Verhaltens dessen Auftretenswahrscheinlichkeit senkt. Somit kann erwünschtes Verhalten entweder durch Anerkennung und Belohnung verstärkt werden (Annäherungslernen), oder aber unerwünschtes Verhalten durch dessen Bestrafung minimiert werden (Vermeidungslernen). Beide Arten der Verhaltensmodifikation spielen eine wesentliche Rolle beim instrumentellen Lernen (*reinforcement learning*). Instrumentelles Lernen beschreibt den Prozess, der einen Organismus lehrt, Belohnung zu maximieren und Bestrafung zu minimieren (Sutton 1998). In den Neurowissenschaften besteht ein zunehmendes Interesse daran, die neuronalen Korrelate des instrumentellen Lernens zu identifizieren, um Klarheit darüber zu gewinnen, wie neuronale Netzwerke Feedback verarbeiten und für zukünftige Entscheidungen nutzen.

In der experimentellen Psychologie wurden entsprechend verschiedene Versuchsabläufe entwickelt, um belohnungsabhängiges Lernen und Entscheidungsfindungsprozesse zu erforschen. Es haben sich historisch und konzeptuell unterschiedlich gewachsene Bereiche (die theoretische Psychologie und die komputationalen Neurowissenschaften) mit jeweils eigenen Begriffen etabliert: Das zielgerichtete vs. das modellbasierte und das habituelle vs. das modellfreie Verhalten. In den kognitiven Neurowissenschaften gibt es eine anhaltende Diskussion darüber, ob es sich bei diesen Konzepten um dichotome Konstrukte handelt oder aber um Pole eines kontinuierlichen Spektrums, welche sich zwischen dem reflexiven und reflektiven Lern- und Entscheidungsverhalten bewegen (Dolan and Dayan 2013). Eine Grundannahme für diese Unterscheidung ist, dass für jede Entscheidungsmöglichkeit verschiedene zugrunde liegende Prozessklassen existieren, welche den Prozess dominieren und ein Resultat (die Handlung) bewirken. Reflektives Entscheidungsverhalten basiert dabei auf dem „vorausschauenden Abwägen“ zukünftiger Handlungsoptionen und ihrer Konsequenzen, während die reflexive Verhaltenskontrolle eher zurückliegende Erfahrungen mit entsprechend positiven oder negativen Handlungskonsequenzen nutzt.

Die Operationalisierung dieser intuitiv eingängigen Prozessunterscheidung wurde auf viele verschiedene Arten realisiert. Zur Erforschung des zielgerichteten vs. habituellen Verhaltens wurden vorrangig Devaluationsparadigmen (Valentin, Dickinson et al. 2007, Tricomi, Balleine et al. 2009) verwendet. Während der Devaluation wird ein belohnender Reiz solange zur Verfügung gestellt, bis eine Sättigung eintritt und er nicht mehr als belohnend

empfundener wird. Habituelles (reflexives) Verhalten führt weiter zur Wahl des devaluierten Stimulus (weil dieser in der Vergangenheit belohnt wurde), während zielgerichtetes (reflektives) Verhalten diesen meidet. In den komputationalen Neurowissenschaften wurde z.T. parallel und z.T. als Reaktion auf die Befunde der klassischen Psychologie Verhalten mit mathematischen Modellen beschrieben und u.a. auf der Basis dynamischer Programmierungstechniken (Bellman 1957) und statistischer Modelle (Sutton 1998) prädiziert (eine hervorragende Übersicht zu diesem Thema geben Dolan and Dayan 2013). Hierzu wurden sequentielle Aufgabenstrukturen wie der *decision-tree* (Huys, Eshel et al. 2012) und die *two-step* Aufgabe (Daw, Gershman et al. 2011) entwickelt. Diese Aufgaben ermöglichen eine Quantifizierung der Anteile von modellfreiem und modellbasiertem Verhalten. Der direkte Zusammenhang beider Operationalisierungen wurde jedoch bislang nicht untersucht und an einer gemeinsamen Population geprüft, obwohl die Begrifflichkeiten (modellfrei ~ habituell / modellbasiert ~ zielgerichtet) teilweise synonym verwendet wurden. Ziel der **Studie 1** war es, diesen Zusammenhang erstmals an einer Population empirisch zu prüfen.

Zwei wichtige Faktoren, die Einfluss auf das eingangs beschriebene Lern- und Entscheidungsverhalten nehmen, sind die individuelle kognitive Kapazität und Stresserleben. Stress beeinflusst kontextabhängiges Lernen (Schwabe, Bohringer et al. 2009) und fördert habituelles Verhalten (Dias-Ferreira, Sousa et al. 2009, Schwabe and Wolf 2009). Die Forschung fokussierte bislang jedoch maßgeblich auf den Effekt von akutem Stress (Schwabe and Wolf 2009, Otto, Raio et al. 2013, Buckert, Schwieren et al. 2014). Im Hinblick auf chronisches Stresserleben deuten erste Ergebnisse auf ähnliche Effekte mit Verstärkung habitueller Verhaltenskontrolle hin (Soares, Sampaio et al. 2012, Radenbach, Reiter et al. 2015). Stresserleben nimmt Einfluss auf kognitive Prozesse wie die Verarbeitungsgeschwindigkeit und deren Güte (Baumeister, Twenge et al. 2002) in Abhängigkeit von seiner Dauer, Valenz und Intensität (Lupien, McEwen et al. 2009). Gleichzeitig moderieren kognitive Variablen, in welchem Ausmaß Stresserleben Lernprozesse und Entscheidungsverhalten beeinflussen kann. Otto und Kollegen (2013) haben z.B. gezeigt, dass die Balance zwischen modellfreiem und modellbasiertem Verhalten nach Stressexposition durch die individuelle Verarbeitungskapazität (eine wesentliche Komponente der fluiden Intelligenz) moderiert wird und Probanden insbesondere dann eine Zunahme der modellfreien (habituellen) Komponente nach akutem Stresserleben zeigten, wenn sie eine geringere Verarbeitungskapazität besaßen. Über welche Mechanismen genau Stress sich auf Lernprozesse und die Entscheidungsfindung bei belohnungsassoziierten Reizen auswirkt (Schwabe and Wolf 2012, Otto, Raio et al. 2013) und welche kognitiven Variablen diesen Zusammenhang moderieren können ist unklar. Wesentliche Aspekte von Kognition (wie logisches Denken, Merkfähigkeit und erlerntes Wissen) wurden bislang nur unzureichend oder gar nicht berücksichtigt (Otto, Raio et al. 2013, Robinson, Overstreet et



al. 2013, Radenbach, Reiter et al. 2015), obwohl gerade das Verständnis der Aufgabenstruktur und die kognitive Repräsentation der möglichen Konsequenzen als eine wichtige Voraussetzung der modellbasierten Entscheidungsfindung identifiziert wurden (Gillan, Otto et al. 2015). Das Konstrukt der Intelligenz ist eine Möglichkeit, kognitive Kapazität umfassend abzubilden und soll in **Studie 2** dazu genutzt werden, den Zusammenhang zwischen chronischem Stresserleben und belohnungsabhängigem Lernen besser zu verstehen.

Wie belohnungsassoziierte Reize verarbeitet werden und welche Faktoren darauf Einfluss nehmen ist eine wichtige Komponente bei der Entstehung von Suchterkrankungen. Klassische und instrumentelle Konditionierungsprozesse, welche das „Erlernen“ einer Suchterkrankung begünstigen, konnten im Tiermodell ausführlich beschrieben werden (Flagel, Akil et al. 2009, Flagel, Robinson et al. 2010). Die Tendenz, suchtasoziierten Reizen einen hohen Anreizcharakter (im Sinne einer erhöhten Salienz) zuzuschreiben, scheint eng an das belohnungsassoziierte Lernen und seine dopaminerge Modulation im Ventralen Tegmentum (Ventral Tegmental Area, VTA) und Ventralen Striatum (VS) (wesentliche Bestandteile des sogenannten mesolimbischen Belohnungssystems) gekoppelt zu sein. Tiere, die mit einer erhöhten phasischen Dopaminausschüttung an der Synapse der VTA auf einen klassisch konditionierten Reiz reagierten, zeigten bei Darbietung eines suchtasoziierten Stimulus deutlich erhöhtes substanzsuchendes (instrumentelles) Verhalten (Flagel, Akil et al. 2009, Flagel, Robinson et al. 2010). Dieses „hijacking“ des Belohnungssystems, im Sinne einer Tendenz vermehrt auf drogenassoziierte Reize zu reagieren, kann allgemeine Lernprozesse und flexible Verhaltensanpassung auch auf nicht-drogenassoziierte Reize beeinträchtigen. In der Tat konnten aktuelle Studien am Menschen abweichende Lernprozesse als eine Komponente süchtigen Verhaltens replizieren: in einer Aufgabe zur sequentiellen Entscheidungsfindung (*two-step*) zeigten alkoholabhängige Patienten mehr modellfreies Verhalten im Vergleich zu den zielorientierten, modellbasierten Kontrollprobanden (Sebold, Deserno et al. 2014). Huys und Kollegen (Huys, Tobler et al. 2014) wendeten die eingangs beschriebenen mathematischen Modelle zur Erfassung von modellfreiem und modellbasierten Verhalten auf die oben beschriebenen tierexperimentellen Daten an, und konnten nachweisen, dass gerade modellfreies Verhalten eng an die phasische Dopaminausschüttung während der Salienzzuschreibung eines konditionierten Stimulus gekoppelt war und gleichzeitig einen Risikofaktor für die Entstehung und Aufrechterhaltung süchtigen Verhaltens darstellte. Modellfreie Lerner schreiben dem konditionierten Stimulus einen höheren Anreizcharakter ( $\sim$ Salienz) zu, verbunden mit einer stärkeren Modulation der dopaminergen Neurotransmission (höhere phasische Dopaminausschüttung) und reagieren entsprechend auch stärker auf konditionierte, drogenassoziierte Stimuli. Chronischer Alkoholkonsum kann in der Folge zu einer Entkopplung zwischen der dopaminergen Modulation und dem modellfreien neuronalen

Prädiktionsfehlersignal im VS führen, und so dysfunktionale Lernprozesse begünstigen (Deserno, Beck et al. 2015). Ein besseres Verständnis dieser Lernprozesse, ihrer moderierenden umwelt- und genassozierten Variablen und deren neurobiologisches Korrelat liefern zusammen wichtige Hinweise für die Ätiologie der Alkoholabhängigkeit und können helfen, wirksame Interventionen zu entwickeln. Aktuelle Daten und Erkenntnisse hierzu werden in **Studie 3** in einem Übersichtspapier dargestellt.

## 1.2 Zielstellung

Ziel der vorliegenden Arbeit ist es, belohnungsabhängiges (instrumentelles) Lernen und Entscheidungsfindungsprozesse auf Verhaltens- und neuronaler Ebene in Abhängigkeit von chronischem Stresserleben (erfasst über den *Lifetime Stress Inventory*, Holmes und Rahe 1962) und kognitiven Variablen (eingeteilt in eine fluide und eine kristalline Intelligenzkomponente) an gesunden Probanden zu untersuchen. Dabei steht zu Beginn die Sicherung der Konstruktvalidität zwischen den bislang oft synonym verwendeten Begriffen modellfrei ~ habituell, bzw. modellbasiert ~ zielgerichtet im Fokus. Darauf aufbauend soll dann insbesondere der differentielle und interaktionelle Einfluss von chronischem Stresserleben und kognitiven Variablen auf Entscheidungsprozesse (instrumentelles Lernen) und deren neuronales Korrelat im VS untersucht und dargestellt werden. Abschließend wird die Relevanz der untersuchten belohnungsabhängigen Lernprozesse für die Entstehung und Aufrechterhaltung der Alkoholabhängigkeit zusammen mit weiteren Einflussvariablen in einem Übersichtspapier diskutiert.

**Studie 1 (Friedel, Koch et al. 2014)** untersucht den Zusammenhang von zielgerichtetem (vs. habituellem) und modellbasiertem (vs. modellfreiem) Verhalten anhand unterschiedlicher Operationalisierungen (Devaluation vs. sequentielle Entscheidungsfindung) zur Prüfung der Konstruktvalidität.

**Studie 2 (Friedel, Schlagenhaut et al. 2015)** untersucht den Einfluss von chronischem Stresserleben auf den Zusammenhang zwischen dem neuronalen (modellfreien) Prädiktionsfehlersignal und Intelligenz (fluid und kristallin) bei gesunden Probanden mittels funktioneller Magnetresonanz Tomographie (fMRT) während einer sequentiellen Entscheidungsaufgabe (die *reversal*-Lernaufgabe).

**Studie 3 (Hagele, Friedel et al. 2014)** gibt eine Übersicht über aktuelle Risikofaktoren für die Entstehung und Aufrechterhaltung der Alkoholabhängigkeit mit einem Fokus auf den eingangs beschriebenen Lernmechanismen und damit assoziierten Gen X Umwelt Interaktionen.

### 1.3 Methoden

**Studie 1** Zur Erfassung von habituellem und zielgerichtetem Verhalten wurde das Devaluationsparadigma von Valentin und Kollegen (2007) adaptiert. Zur Erfassung von modellbasiertem und modellfreiem Verhalten wurde die von Daw und Kollegen (2011) etablierte *two-step*-Aufgabe verwendet. Alle 18 getesteten gesunden Probanden absolvierten beide Aufgaben, um ein gemeinsames bzw. überlappendes Konstrukt für die beiden Operationalisierungen von Lernen identifizieren zu können.

In der Devaluationsaufgabe lernten die Probanden einen (abstrakten) Stimulus auszuwählen, der probabilistisch (75/25) an den Erhalt einer Belohnung (einen primären natürlichen Verstärker: Schokoladenmilch oder Orangensaft) gekoppelt war. Nach der Lernphase wurde dieser Belohnungsreiz durch Sättigung (Konsum von Schokolade/Orangen bis zur subjektiven Angabe des Probanden, nicht mehr essen zu können, sowie der subjektiven Abwertung des Stimulus über eine visuelle Analogskala) devaluiert. In einer Testphase wurde nun überprüft, inwieweit der mit dem devaluierten Stimulus assoziierte Reiz weiterhin (habituell) ausgewählt oder aber (zielgerichtet) vermieden wurde. Der daraus errechnete Interaktionsscore (Wahl des belohnten vs. neutralen Stimulus vor und nach der Devaluation) diene als quantitatives Maß für das Ausmaß an zielgerichtetem vs. habituellem Verhalten.

In der *two-step* Aufgabe mussten die Probanden auf zwei Stufen eine Entscheidung zwischen zwei (abstrakten) Stimuli treffen, die nach der zweiten Auswahl probabilistisch (nach sog. *random-walks*) an eine Belohnung gekoppelt war (20 Cent). Außerdem war jedes Stimuluspaar der ersten Stufe probabilistisch (70/30) an ein Stimuluspaar der zweiten Stufe gekoppelt. Der modellbasierte Lerner berücksichtigte diese Transitionswahrscheinlichkeiten und wählte in Folge auch dann einen Stimulus erneut aus, wenn dieser nach der zweiten Stufe nicht belohnt wurde (weil er verstanden hat, dass es sich um einen seltenen Misserfolg handelt) bzw. wählte einen Stimulus auf erster Ebene trotz Belohnung nicht erneut aus, wenn es sich um einen seltenen Erfolg handelte. Der modellfreie Lerner hingegen ignorierte die Transitionswahrscheinlichkeiten zwischen Stufe eins und zwei und orientierte seine Auswahl ausschließlich am Gewinn nach der zweiten Stufe.

Der daraus errechnete Interaktionsscore (Wahl des belohnten vs. neutralen Stimulus in Abhängigkeit von der Transitionswahrscheinlichkeit) diene als quantitatives Maß für das Ausmaß an modellbasiertem vs. modellfreiem Verhalten. Zusätzlich wurden über mathematische Algorithmen der modellbasierte Parameter Omega ( $\bar{\omega}$ ) und der modellfreie Parameter Lamda ( $\lambda$ ) berechnet. Der Parameter  $\bar{\omega}$  gibt zusätzlich zu  $\lambda$  Auskunft über die Balance zwischen modellfreiem und modellbasiertem Verhalten.

Die Parameter beider Operationalisierungen wurden dann in einem letzten Schritt zur Prüfung des Zusammenhangs zwischen modellfreiem und habituellem sowie modellbasiertem und zielgerichtetem Verhalten miteinander korreliert.

**Studie 2** In der zweiten Studie wurde die von Park und Kollegen (2010) adaptierte *reversal*-Lernaufgabe an 16 gesunden Probanden durchgeführt. Es handelte sich bei der *reversal*-Lernaufgabe ebenfalls um eine probabilistische Lernaufgabe zur sequentiellen Entscheidungsfindung mit zwei Stufen: die Auswahl eines abstrakten Stimulus auf der ersten Stufe führte auf der zweiten Stufe mit einer Wahrscheinlichkeit von 80 vs. 20 % zu einer Belohnung (in diesem Fall kein monetärer Gewinn sondern ein lachendes vs. ein unglückliches Gesicht). Im Verlauf des Experimentes kam es mehrfach zu einem Wechsel der Stimulus-Belohnungs-Assoziation (sog. *switch*), d.h. ein zuvor belohnter Stimulus führte nicht länger zur erwarteten Belohnung und evozierte so einen Prädiktionsfehler. Die Berechnung dieses Prädiktionsfehlers erfolgte über ein Rescorla-Wagner Modell (Sutton 1998), der dann als Regressor in die Analyse der fMRT-Daten einging, um das neuronale Korrelat des Prädiktionsfehlersignals im VS erfassen zu können (O'Doherty, Hampton et al. 2007, Schlagenhaut, Rapp et al. 2013). Intelligenz wurde über eine Faktorenanalyse (Varimax Rotation mit Eigenwert größer 1.0) aus 10 neuropsychologischen Testverfahren (adaptiert aus der Berliner Altersstudie, Lindenberger and Baltes 1997) in eine kristalline und eine fluide Komponente eingeteilt. Chronisches Stresserleben wurde über die *Lifetime Stressor Checklist* von Holmes und Rahe (1967) als die gewichtete Summe von 42 als stressreich eingestuften Lebensereignissen in den letzten 12-24 Monaten erfasst.

Um den Zusammenhang zwischen Stress, Intelligenz und dem neuronalen Korrelat von Lernen im VS zu prüfen, wurde eine schrittweise Regression mit den Faktoren Stress, Prädiktionsfehlersignal und IQ durchgeführt. Um den Einfluss von Stresserleben auf den Zusammenhang und die Interaktion zwischen Intelligenz und dem neuronalen Korrelat von Lernen differenzierter betrachten zu können, wurde zusätzlich eine Median-Split-Analyse anhand der Anzahl stressreicher Lebensereignisse durchgeführt.

**Studie 3** Studie 3 stellt in einem Übersichtspapier den aktuellen Stand der Forschung zum Verständnis prädisponierender und risikoassoziierter Faktoren der Alkoholabhängigkeit, insbesondere im Hinblick auf belohnungsassoziertes Lernen, chronisches Stresserleben und deren Interaktion mit genetischen Vulnerabilitäts- und Resilienzfaktoren dar. Zusätzlich wurden Annahmen über deren neurobiologische Grundlagen anhand tierexperimenteller, bildgebender und mathematischer Modellierungsstudien dargestellt und in ein allgemeines Modell über den Pathomechanismus der Alkoholabhängigkeit eingebettet.

## 1.4 Ergebnisse

**Studie 1** Im Devaluationsparadigma zeigte sich, dass die Probanden während der Testphase lernten, den belohnenden Stimulus signifikant häufiger auszuwählen. Es fand eine selektive Devaluation bei allen eingeschlossenen Probanden statt, d.h. nach der Devaluation wurde die entsprechende Belohnung (Orangensaft oder Schokolade) selektiv abgewertet, während der nicht devaluierte Stimulus weiter als angenehm eingestuft und

weiterhin ausgewählt wurde. In der Extinktionsphase zeigte sich eine signifikante Zeit (pre/post Training) x Bedingung (devaluiert/valuiert/neutral) Interaktion. Das Gewicht dieses Interaktionsterms diente als Maß für die Stärke der Ausprägung von zielgerichtetem Verhalten, wie bereits von Valentin und Kollegen (2007) berichtet.

In der *two-step*-Aufgabe konnte sowohl ein signifikanter Haupteffekt für Belohnung (als Indikator für modellfreies Verhalten), als auch eine signifikante Interaktion von Belohnung und Transitionsverhalten, als Indikator für modellbasiertes Verhalten, nachgewiesen werden. Das Verhalten während der Aufgabe wurde also von beiden Formen der Entscheidungsfindung beeinflusst.

Es zeigte sich eine signifikante Korrelation der Interaktionsparameter aus beiden Lernaufgaben. Zusätzlich konnte ein signifikanter Zusammenhang zwischen dem Interaktionsparameter der Devaluation und dem Modellparameter  $\bar{U}$  für modellbasiertes Verhalten aus der *two-step*-Lernaufgabe nachgewiesen werden, während die Korrelation mit den Parametern für modellfreies Verhalten ( $\lambda$  und der Haupteffekt für Belohnung) nicht signifikant wurde. Modellfreies und habituelles Verhalten schienen also zumindest teilweise auf unterschiedlichen Mechanismen zu beruhen, während modellbasiertes und zielgerichtetes Verhalten einen deutlichen Zusammenhang zeigten.

**Studie 2** Ein Hauptergebnis der zweiten Studie ist die signifikante Aktivierung des VS als Korrelat des modellfreien Prädiktionsfehlersignals berechnet aus der *reversal*-Lernaufgabe über alle 16 Probanden. Dieses modellfreie Prädiktionsfehlersignal im VS korrelierte signifikant mit dem fluiden IQ und der Summe stressreicher Lebensereignisse. Fluiden IQ und Stresserleben zeigten keinen signifikanten Zusammenhang, während die Interaktion Stress x modellfreies Prädiktionsfehlersignal im VS signifikant war. Um diese Interaktion genauer zu analysieren, wurde eine Aufteilung der Probanden in viel und wenig Stresserleben vorgenommen (Median-Split der Summe stressreicher Lebensereignisse). Als Ergebnis konnte nachgewiesen werden, dass der Zusammenhang zwischen dem modellfreien Prädiktionsfehlersignal im VS und dem fluiden IQ dann besonders hoch war, wenn die Probanden viel Stresserleben in der Vergangenheit berichtet hatten. Im Gegensatz zum fluiden IQ zeigte sich auf neuronaler Ebene kein Einfluss des kristallinen IQ auf das modellfreie Prädiktionsfehlersignal und keine Interaktion mit chronischem Stresserleben.

**Studie 3** Die geschätzte Erbllichkeit der Alkoholabhängigkeit lag bei 50%, und war zwischen Männern und Frauen gleichverteilt. Das Risiko an einer Alkoholabhängigkeit zu erkranken, wurde in gleichem Ausmaß (um den Faktor 1.4) durch Umweltfaktoren (inklusive familiärem Umfeld, Peer Einfluss und Erziehung) und genetische Variablen bestimmt (Kendler, Ji et al. 2015). Insbesondere konnten Genotypen als Risikofaktoren klassifiziert werden, die eine Modulation der Alkoholverträglichkeit und Wirksamkeit bewirken und entsprechend den Anreizcharakter der Droge als konditionierten Stimulus moderieren konnten. Die Relevanz dopaminergener Gene scheint zusätzlich durch deren Einfluss auf die

emotionale Verarbeitung suchtassoziierter Stimuli im limbischen System (insbesondere der Amygdala) moderiert zu werden, mit einer entsprechend gehäuften Komorbidität von Angst- und Suchterkrankungen (Kienast, Hariri et al. 2008, Kienast, Schlagenhauf et al. 2013).

Die Ergebnisse der ausgewählten Studien deuten darauf hin, dass süchtiges Verhalten u.a. aus einer pathologischen Motivation gegenüber der Droge resultiert, die ihre Grundlage in einer dysfunktionalen dopaminergen Enkodierung suchtrelevanter Reize hat. Der einer Droge zugeschriebene hedonische Aspekt scheint dagegen eher durch das Opioid-System enkodiert zu werden (Berridge 2009, Berridge, Robinson et al. 2009). Zusätzlich konnten genetische Faktoren für klassische Konditionierungsprozesse als Grundlage süchtigen Verhaltens im Tierexperiment nachgewiesen werden (Flagel, Robinson et al. 2010). Es fand sich eine erhöhte Vulnerabilität für Suchterkrankungen bei solchen Tieren, die in einem klassischen Konditionierungsexperiment dem konditionierten Stimulus (CS) einen erhöhten Anreizcharakter zuschrieben und entsprechend auch mit einer erhöhten phasischen dopaminergen Neurotransmission innerhalb der VTA auf den CS reagierten. Wichtig war hierfür die Unterscheidung zwischen dem motivationalen Anreizcharakter (bei Erwartung) und dem hedonischen Belohnungsaspekt (bei Erhalt) der Droge.

Funktionelle Bildgebungsstudien am Menschen konnten diese Befunde zum Teil replizieren, und zeigten dass eine Kombination aus klassischem und instrumentellen Lernen (der sog. Pavlovian to Instrumental Transfer, PIT) die neuronale Enkodierung von suchtspezifischen Reizen im VS moderiert und die Rückfallrate von abhängigen Patienten z.T. präzisieren konnte (Garbusow, Schad et al. 2014, Garbusow, Sebold et al. 2014).

## **1.5 Diskussion**

Entsprechend der a priori Hypothese zeigten die Korrelationsanalysen zwischen Devaluation und sequentieller Entscheidungsfindung einen positiven Zusammenhang zwischen modellbasiertem und zielgerichtetem Verhalten. Dies liefert erste empirische Hinweise dafür, dass beiden Konstrukten tatsächlich ein gemeinsamer Mechanismus zugrunde liegt. Dieser Befund wurde in einer aktuellen Studie von Gillan und Kollegen (2015) an 185 gesunden Probanden repliziert, welche die Sensitivität gegenüber Devaluation nutzten, um modellbasiertes Verhalten vorherzusagen. In Übereinstimmung mit Studie 1, die keine signifikante Korrelation der modellfreien Parameter nachweisen konnte, fand sich auch in der Studie von Gillan et al. (2015) kein direkter Zusammenhang zwischen modellfreiem und habituellem Verhalten. Entweder sind diese beiden Komponenten der Entscheidungsfindung tatsächlich unabhängig voneinander (Dezfouli und Balleine 2013) oder aber die Operationalisierung ihrer Mechanismen ist bislang nicht ausreichend, um tatsächlich habituelle Verhaltensmuster zu evozieren (Tricomi, Balleine et al. 2009).

Wir fanden eine Kopplung des modellfreien Prädiktionsfehlersignals im VS mit fluider Intelligenz nur bei Probanden, die viel Stresserleben in den letzten beiden Jahren angaben. Es scheint, dass gerade diese Population (möglicherweise durch eine erhöhte Verarbeitungsgeschwindigkeit und parallele Prozessierungsmöglichkeiten) genau dann mehr Gebrauch von einem modellfreien Signal machte, wenn sich dieses als besonders effizient in der Vergangenheit erwiesen hat und ein Umlernen nicht erfolgsversprechend war. Möglicherweise bietet die modellfreie neuronale Enkodierung des Prädiktionsfehlers während der *reversal*-Aufgabe einen Vorteil, wenn in der Vergangenheit die Vorhersage von Belohnung nach kognitiven Karten (also modellbasiert) aufgrund stetig wechselnder Reiz-Belohnungsassoziationen (die Stress induzieren und gleichzeitig eine fehlerhafte Belohnungsvorhersage provozieren) nicht möglich war.

Als Ursache der Signalveränderung im VS nach Lernprozessen wurden Veränderungen an der dopaminergen Synapse im Bereich des Striatums und seiner Projektionen in den Präfrontalkortex bereits nachgewiesen (Yin, Knowlton et al. 2005, Yin, Ostlund et al. 2005, de Wit, Corlett et al. 2009, de Wit, Watson et al. 2012). Ähnliche Veränderungen im VS wurden auch als Folge von akutem und chronischem Stresserleben (Cao, Covington et al. 2010, Valenti, Lodge et al. 2011, Robinson, Overstreet et al. 2013) und in Abhängigkeit von kognitiven Variablen (Schlagenhauf, Rapp et al. 2013) beschrieben. Beide Faktoren, Stress und Lernen, evozieren also eine erhöhte phasische Dopaminausschüttung an der Synapse im VS. Eine Parallele die erklären könnte, warum gerade Probanden mit viel Stresserleben und einem hohen fluiden IQ die Signalverstärkung im VS während der *reversal*-Aufgabe zeigten.

Chronisches Stresserleben und eine damit verbundene veränderten Belohnungsverarbeitung konnte als Risikofaktor für erhöhten Alkoholkonsum identifiziert werden: in einer prospektiv angelegten Studie an Jugendlichen zeigte sich ein (kausaler) Zusammenhang zwischen chronischer Stressexposition, veränderter Belohnungsverarbeitung im medialen Präfrontalkortex (mPFC) und einem erhöhten Alkoholkonsum im Erwachsenenalter (Casement, Shaw et al. 2015). Chronisches Stresserleben scheint also über dysfunktionale Lernprozesse das Risiko eines riskanten Alkoholkonsums zu erhöhen.

Die Relevanz der dargestellten Lernmechanismen wurde sowohl in tierexperimentellen Studien, als auch in funktionellen Bildgebungsstudien am Menschen repliziert. Flagel und Kollegen (Flagel, Watson et al. 2008, Flagel, Akil et al. 2009, Flagel, Robinson et al. 2010, Flagel, Cameron et al. 2011) zeigten in einer Reihe von Experimenten an Ratten, dass die genetisch verankerte Prädisposition, dem konditionierten Stimulus (CS) eine hohe Salienz zuzuschreiben, mit einer erhöhten phasischen Dopamin Antwort auf den CS und gleichzeitig einem erhöhten Risiko für die Entstehung süchtigen Verhaltens einherging. Wichtig ist hier die Unterscheidung, dass nicht Lernen per se durch Dopamin enkodiert wurde, sondern der spezifische Wert, der einem erlernten Stimulus zugeschrieben wurde (seine Salienz). Dieser

Prozess konnte durch die Antagonisierung der phasischen Dopaminantwort selektiv unterdrückt werden (Flagel, Clark et al. 2011). Bildgebende Studien an alkoholabhängigen Patienten bestätigen die Relevanz der dargestellten Lernprozesse für süchtiges Verhalten: Patienten mit einer Alkoholabhängigkeit zeigten einen verstärkten Einfluss klassisch konditionierter Reize auf instrumentelles Lernen, der mit einer veränderten neuronalen Verarbeitung belohnungsrelevanter Reize im VS und höheren Rückfallraten assoziiert war (Garbusow, Schad et al. 2015).

Zusammenfassend zeigen die Ergebnisse der Studien 1-3, dass gerade das Zusammenspiel und die wechselseitige Beeinflussung von gewachsenen Lernstrukturen, zurückliegenden stressreichen Erfahrungen, kognitive Variablen und die damit assoziierten neurobiologischen Anpassungsprozessen wichtige Bausteine in einem besseren Verständnis süchtigen Verhaltens darstellen. Hieraus leiten sich wichtige Implikationen für die Prävention und Behandlung süchtigen Verhaltens ab.



## Theoretischer Hintergrund

### 1.6 Die psychologischen und neurobiologischen Grundlagen von belohnungsabhängigem Lernen

Lernvorgänge hängen von zahlreichen Faktoren ab: Wie passen wir unsere Entscheidungen an, wenn die wiederkehrende Wahl zwischen zwei belohnenden Alternativen getroffen werden soll? Welche Alternativen stufen wir in Abhängigkeit von bereits getroffenen Entscheidungen als belohnend ein? Welches Risiko gehen wir für eine erhoffte Belohnung ein und wie gehen wir mit einem Ausbleiben dieser um? Wichtige Einflussvariablen auf solche Lern- und Entscheidungsvorgänge sind unter anderem unsere früheren Erfahrungen mit diesen alternativen Zuständen, unser Wissen über die Wahrscheinlichkeit der erwarteten Belohnung jeder Alternative, die kognitive Fähigkeit, alle möglichen Alternativen gegeneinander abzubilden und nicht zuletzt auch der emotionale Zustand im Moment der Entscheidungsfindung. Veränderungen im Belohnungslernen haben entsprechend weit reichende Konsequenzen auf das tägliche Erleben, Leben und Handeln. Zugleich stellen sie eine wichtige Komponente in der Psychopathologie zahlreicher Erkrankungen, wie beispielsweise Suchterkrankungen (Heinz, Wrase et al. 2007, Heinz, Beck et al. 2009, Park, Kahnt et al. 2010), affektiven Erkrankungen (Abler, Greenhouse et al. 2008) oder Essstörungen (Simon, Skunde et al. 2014, Wierenga, Bischoff-Grethe et al. 2015) dar.

Im Fokus der vorliegenden Arbeit stehen zwei distinkte Lernsysteme, die als Grundlage des belohnungsassoziierten, instrumentellen Lernens (so genanntes *reinforcement learning*) gelten: das „reflexive“ Reiz-Reaktions-Lernen (Thorndike 1911, Dolan and Dayan 2013) und das zielorientierte, „reflektierte“ Lernen nach kognitiven Karten (Tolman 1948). Ersteres gründet auf dem Erlernen einer instrumentellen Kontingenz zwischen Verhalten, Reiz und Belohnung. Letzteres beruht zusätzlich auf dem Anreizwert, der einem Reiz zugeschrieben und durch die Motivation des Individuums maßgeblich mit beeinflusst wird (Balleine 2005). Die Evidenz für Reiz-Reaktions-Lernen kommt zu großen Teilen aus Devaluationsstudien (Valentin, Dickinson et al. 2007), während die Forschung zum reflektierten Lernen sich vor allem mit Aufgaben zur sequentiellen Entscheidungsfindung befasst (Daw, Gershman et al. 2011, Huys, Eshel et al. 2012).

In Studie 1 der folgenden Arbeit sollen diese beiden Systeme und ihre Operationalisierungen miteinander verglichen und in Studie 2 auf moderierende Variablen (insbesondere Stresserleben und Intelligenz) hin untersucht werden.

Ergänzend zur klassischen Konditionierung (Pawlow 1927) begründete Thorndike (1911) das Prinzip des operanten Konditionierens als Reiz-Reaktions-Lernen, mit dem wesentlichen Unterschied, dass hier der lernende Organismus erstmals Kontrolle über den Reiz und seine Reaktion hatte. Die Häufigkeit eines Verhaltens wird durch seine

angenehmen (appetitiven) oder unangenehmen (aversiven) Konsequenzen nachhaltig verändert. Während beim operanten Konditionieren spontanes Verhalten durch Belohnung/Bestrafung modifiziert wird, nutzt das instrumentelle Lernen Konsequenzen, um ein spezifisches zielorientiertes Verhalten zu modifizieren. Entsprechend wurden instrumentelle Lerner in die so genannten „zielorientierten“ und die „habituellen“ Lerner eingeteilt. Eine Operationalisierung dieser Lernform ist das Prinzip der Devaluation: Ein appetitiver Reiz wird solange appliziert, bis eine Sättigung eintritt und der Stimulus an Wert verliert. Der habituelle Lerner wählt trotz Devaluation häufiger den in der Vergangenheit belohnten Reiz. Der zielgerichtete Lerner kann sein Verhalten anpassen und meidet den devaluierten Reiz, da er nicht länger zu dem vormals angenehmen Ziel (Sättigung) führt. Die Grundlage bilden tierexperimentelle Studien (Dickinson 1985), welche dann in Humanstudien an Gesunden (Valentin, Dickinson et al. 2007, de Wit, Corlett et al. 2009, Tricomi, Balleine et al. 2009, Valentin and O'Doherty 2009) und Menschen mit psychischen Störungen (de Wit, Barker et al. 2011, Gillan, Papmeyer et al. 2011, Sjoerds, de Wit et al. 2013) repliziert wurden.

Im Vergleich zum Reiz-Reaktions-Lernen nutzen die Aufgaben zur sequentiellen Entscheidungsfindung eine sehr ähnliche Terminologie, welche jedoch historisch aus einem anderen Bereich (den komputationalen Neurowissenschaften im Vergleich zur experimentellen Psychologie) gewachsen ist. Sie verwenden die Begriffe modellfreies (~habituelles) im Gegensatz zu modellbasiertem (~zielgerichtetem) Verhalten (Dolan and Dayan 2013). Der modellfreie Lerner entscheidet sich in der Regel für den zuletzt belohnten Reiz, unabhängig von Wahrscheinlichkeitsprofilen und Aufgabenstrukturen, die ggf. dazu führen könnten, nicht belohnte Reize einer sofortigen Belohnung vorzuziehen (ähnlich wie der habituelle Lerner). Im Gegensatz dazu ist ein Verhalten modellbasiert, wenn die Entscheidung die Summe vorangegangener Erfahrungen und daraus entstandene internale Repräsentationen einer Aufgabenstruktur berücksichtigt. Somit ist das modellbasierte Verhalten in seiner Natur auch zielgerichtet. Ein rein modellfreier Lerner vernachlässigt die Aufgabenstruktur und wiederholt die zuletzt belohnte Verhaltenssequenz. Aufgaben zur sequentiellen Entscheidungsfindung wurden u.a. an Gesunden (Daw, Gershman et al. 2011, Wunderlich, Dayan et al. 2012, Smittenaar, FitzGerald et al. 2013) und Probanden mit Alkoholabhängigkeit (Sebold, Deserno et al. 2014) evaluiert.

Beiden Lerntheorien gehen davon aus, dass die Konsequenzen des Verhaltens zu einer Verhaltensänderung führen können. Eine Möglichkeit, diesen Lernprozess abzubilden, ist das Modell des so genannten Prädiktionsfehlersignals. Dieses Signal ist als neuronales Signal repräsentiert und basiert auf der Differenz zwischen der erwarteten und tatsächlich erhaltenen Belohnung. Das Modell postuliert nun, dass der Differenzwert als neuronales Signal verwandt werden kann, um zukünftiges Verhalten anzupassen. Wolfram Schultz konnte in tierexperimentellen Studien (Schultz, Dayan et al. 1997) erstmals einen direkten

Zusammenhang zwischen der Belohnungserwartung und der dopaminergen Neurotransmission in der VTA als neuronales Korrelat von belohnungsabhängigem Lernen nachweisen. Über Einzelzelleitungen an Affen wurde nachgewiesen, dass die Feuerrate dopaminergener Neurone zunächst bei Erhalt der Belohnung, in der Folge dann aber als Reaktion auf einen die Belohnung ankündigenden Reiz zunimmt und entsprechend bei Ausbleiben der Belohnung einen Abfall zeigt (Schultz, Dayan et al. 1997). Es findet also eine spezifische valenzabhängige Enkodierung der Belohnungserwartung an der dopaminergen Synapse in der VTA statt. Diese Ergebnisse wurden in der Folge häufig sowohl in tierexperimentellen Studien (Schultz, Dayan et al. 1997, Stuber, Klanker et al. 2008) als auch in Humanstudien (Frank, Seeberger et al. 2004, Volkow, Wang et al. 2006, Heinz, Wrase et al. 2007, Beck, Schlagenhauf et al. 2009) repliziert. Die Neurone der VTA projizieren ihr Signal ins VS und bilden wesentliche Bestandteile des sogenannten mesolimbischen Belohnungssystems, das bei der Verarbeitung von belohnungsassoziierten primären Verstärkern wie Essen, Genussmitteln und Sexualität, aber auch bei erlernten sekundären Verstärkern wie Geld, Drogen (Di Chiara and Imperato 1988) und zahlreichen anderen, insbesondere appetitiven Stimuli, aktiviert wird. Dysfunktionale Aktivierungsmuster im VS wurden mit verschiedenen psychischen Störungen, unter anderem der Alkoholabhängigkeit (Wrase, Schlagenhauf et al. 2007, Beck, Wustenberg et al. 2012, Kienast, Schlagenhauf et al. 2013), aber auch affektiven Störungen (Abler, Greenhouse et al. 2008) und der Anorexia Nervosa (Wierenga, Bischoff-Grethe et al. 2015) in Verbindung gebracht.

### **1.7 Der Einfluss von Stress und Kognition auf belohnungsabhängiges Lernen und seine neuronalen Korrelate**

Veränderungen der dopaminergen Neurotransmission in der VTA und dem VS nach akutem und chronischem Stresserleben (Cao, Covington et al. 2010, Valenti, Lodge et al. 2011, Robinson, Overstreet et al. 2013), zeigten ähnliche Muster wie sie bei belohnungsabhängigem Lernen stattfinden (einen Anstieg der dopaminergen Neurotransmission bei negativer Belohnungserwartung). Die Enkodierung von Lernen und Stresserleben scheint also zumindest auf molekularer Ebene in überlappenden Systemen stattzufinden.

Betrachtet man den Einfluss von Stresserleben auf Verhaltensebene, zeigt sich ein Anstieg des habituellen Verhaltens als Folge akuter Stressexposition (Schwabe and Wolf 2009, Schwabe, Tegenthoff et al. 2010, Schwabe and Wolf 2012, Otto, Raio et al. 2013). Probanden, die einem akuten Stressreiz ausgesetzt waren, zeigten keine Verhaltensanpassung nach Devaluation (Schwabe and Wolf 2009), wählten also weiterhin einen bereits devaluierten Stimulus aus. Zusätzlich besaßen die Probanden nach Stressexposition weniger explizites Wissen über die Stimulus-Belohnungs-Assoziation,

konnten also nicht angeben welcher Stimulus zu einer Belohnung führte. Der Einfluss von Stresserleben auf die Entscheidungsfindung scheint demnach zusätzlich über Prozesse der Gedächtniskonsolidierung moderiert zu sein. Stresserleben erleichtert prozedurales Lernen, während der Abruf aus dem deklarativen Gedächtnis erschwert wird (Schwabe, Wolf et al. 2010, Schwabe and Wolf 2012). Die Effekte von Stress auf die Entscheidungsfindung variieren jedoch in Abhängigkeit von der Dauer und Intensität der Stressexposition (Lupien, McEwen et al. 2009). Wenn Stress über einen längeren Zeitraum anhält, kann die chronische Exposition neben der akuten Verhaltensanpassung (*state*) auch zu langfristigen neurobiologischen Veränderungen und entsprechenden überdauernden Verhaltensmerkmalen (*trait*) führen (Lupien, McEwen et al. 2009). Humanstudien (Schwabe and Wolf 2009, Robinson, Overstreet et al. 2013) haben sich bislang fast ausschließlich auf die Folge von akutem Stress (z.B. die Applikation von Schmerzreizen), auf das instrumentelle Lernen und die Verarbeitung belohnungsassoziierter Reize konzentriert. Dabei wurden kognitive Variablen wie Arbeitsgedächtniskapazität und Verarbeitungsgeschwindigkeit nur unzureichend berücksichtigt. Dass die Berücksichtigung kognitiver Maße auf Verhaltensebene insbesondere nach Stressexposition jedoch hoch relevant ist, wurde bereits nachgewiesen (Schwabe and Wolf 2009, Schwabe, Hoffken et al. 2011, Schwabe and Wolf 2012, Otto, Raio et al. 2013). Otto und Kollegen (2013) berichten eine Abnahme des modellbasierten und eine Zunahme des modellfreien Verhaltens nach akutem Stress in einer sequentiellen Entscheidungsfindungsaufgabe (Daw, Gershman et al. 2011) ausschließlich für Probanden mit niedriger Arbeitsgedächtniskapazität. Eine hohe kognitive Kapazität hatte im Gegensatz dazu einen protektiven Einfluss auf die modellbasierte Komponente unter Stress. Radenbach et al. (2015) berichten ergänzend dazu, dass akuter Stress nur dann einen Einfluss auf die modellbasierte Komponente hatte, wenn die Probanden zusätzlich anhaltendes Stresserleben in den vorangegangenen 12 Monaten, also chronischen Stress, berichteten. Früheres Stresserleben erhöhte entsprechend die Vulnerabilität gegenüber akutem Stress und seinem Einfluss auf die Entscheidungsfindung.

Über welche Mechanismen genau sich Stress auf Lernprozesse und die Entscheidungsfindung insbesondere bei belohnungsassozierten Reizen auswirkt, wird kontrovers diskutiert (Schwabe and Wolf 2012, Otto, Raio et al. 2013). Wegweisende Studien hierzu deuten darauf hin, dass insbesondere die Sensitivität gegenüber Bestrafungen durch akuten Stress beeinflusst wird (Cavanagh, Frank et al. 2011, Robinson, Overstreet et al. 2013). Darüber hinaus werden Entscheidungen unter Stress oft als automatisierter oder auch habitueller (gewohnheitsmäßig) beschrieben (Schwabe and Wolf 2009, Otto, Raio et al. 2013). Habituelle Entscheidungsfindung ist mit einem geringeren kognitiven Aufwand verbunden und von Vorteil, wenn kognitive Ressourcen nicht abrufbar sind oder in der Vergangenheit der Abruf dieser Ressourcen zu fehlerhaften Entscheidungen

fürte (z.B. unter Zufallsentscheidungen oder bei sehr häufigen Wechseln von Reiz-Reaktions-Belohnungs-Wahrscheinlichkeiten).

Entsprechend ist eine weitere wesentliche Komponente beim (instrumentellen) Lernen neben dem Stresserleben auch die individuelle kognitive Kapazität. Das Konstrukt der Intelligenz ist eine Möglichkeit, diese Kapazität abzubilden. Dieses Konstrukt ist historisch gewachsen und unterliegt einem stetigen Wandel. Klassischerweise ist eine Einteilung in die fluide und kristalline Komponente der Intelligenz vorgenommen worden (Horn and Cattell 1966, Sternberg 2000). Die fluide Komponente der Intelligenz (fluider Intelligenz Quotient, IQ) beschreibt die Fähigkeit eines Individuums, logische Schlüsse zu ziehen, abstrakt zu denken, neue Stimuli zu interpretieren und schnelle flexible Verhaltensanpassung vorzunehmen. Auf neuronaler Ebene wurde dieser Aspekt der Kognition vor allem mit einer Aktivierung im dorsolateralen Präfrontalkortex (Goldman-Rakic 2000) und VS (Cools, Clark et al. 2002, Cools, Gibbs et al. 2008, Cooper and Knutson 2008) in Verbindung gebracht. Der kristalline Teil der Intelligenz hingegen wurde mehr mit Wissensakkumulation und der verbalen Intelligenz assoziiert (Schipolowski 2014) und nimmt im Gegensatz zur fluiden Komponente über die Lebensspanne nicht ab (Lindenberger and Baltes 1997).

### **1.8 Der Einfluss von Lernen und Stress auf die Entstehung und Aufrechterhaltung der Alkoholabhängigkeit**

Bisher bekannte ätiologische Faktoren, die im Zusammenhang mit der Entwicklung einer Alkoholabhängigkeit stehen, betreffen sowohl genetische (z.B. die genetisch bedingte geringere Sensitivität gegenüber der akuten Alkoholwirkung oder eine positive Familiengeschichte bezüglich Alkoholabhängigkeit) als auch umweltbedingten Faktoren (z. B. Stress, Trauma, dysfunktionale familiäre Strukturen und der Einfluss der Peers, Garbusow, Sebold et al. 2014). Insbesondere die Interaktion dieser gen- und umweltassoziierten Variablen spielt eine wesentliche Rolle bei der Entstehung süchtigen Verhaltens (Kendler, Ji et al. 2015). Ergebnisse aus klinischen und präklinische Studien (Koob and Bloom 1988, Nestler and Aghajanian 1997, Robinson and Kolb 2004, Kumar, Choi et al. 2005) deuten darauf hin, dass auf neuronaler Ebene Adaptationsprozesse stattfinden, die denen eines physiologischen Lernprozesses ähneln (von der Goltz and Kiefer 2009).

Wie genau diese Adaptationsprozesse im menschlichen Gehirn stattfinden, ist weiterhin Gegenstand der Forschung: Während im Tierexperiment über Einzelzelleitungen beim Affen die VTA und ihre kortikalen Projektionen als wesentliche Bestandteile der belohnungsverarbeitenden Strukturen identifiziert werden konnten (Schultz, Dayan et al. 1997), liefern funktionelle Bildgebungsstudien beim Menschen Hinweise darauf, dass Patienten mit einer Suchterkrankung Veränderungen in der Verarbeitung von

belohnungsassoziierten Reizen (Wrase, Schlagenhauf et al. 2007, Heinz, Beck et al. 2009) und damit einhergehende Veränderungen an der dopaminergen Synapse (Kienast, Hariri et al. 2008, Kienast, Siessmeier et al. 2008, Heinz, Beck et al. 2009, Kienast, Schlagenhauf et al. 2013) im Bereich zentraler Strukturen des mesolimbischen Belohnungssystems, dem VS und seinen Verbindungen zum Präfrontalcortex (Park, Kahnt et al. 2010) zeigen. Die dysfunktionale Regulation belohnungsverarbeitender Areale über den Präfrontalcortex zeigte dabei einen engen Zusammenhang mit Veränderungen der Lerngeschwindigkeit (Park, Kahnt et al. 2010).

Die Frage, ob es sich bei Veränderungen im Lernen und der Verarbeitung belohnungsassoziierten Reize um eine Prädisposition für, oder aber eine Folge der Erkrankung selbst handelt, bleibt damit noch unbeantwortet.

Tierexperimentelle Studien weisen darauf hin, dass die Verarbeitung konditionierter Reize, als eine zum Teil genetisch verankerte Prädisposition (Flagel, Robinson et al. 2010), das Risiko für die Entwicklung einer Suchterkrankung erhöht: Tiere, die ihre Aufmerksamkeit vermehrt einem konditionierten Stimulus zuwendeten (sog. *sign tracker*) und entsprechende Veränderungen an der dopaminergen Synapse als Folge dieses Lernprozesses zeigten, hatten ein erhöhtes Risiko für Suchtverhalten (Flagel, Akil et al. 2009). Im Vergleich dazu beobachteten Flagel und Kollegen bei sog. *goal tracker*, die zwar ebenfalls die Wertigkeit des ankündigenden Reizes (des CS) erlernten, diesem jedoch einen sehr viel geringeren Anreizwert zuschrieben (weniger aufsuchendes Verhalten und Hinwendung) und somit auch weniger süchtiges Verhalten. Interessanterweise war diese Art der Konditionierung zusätzlich mit einer verminderten Modulation an der dopaminergen Synapse auf den drogenankündigenden Reiz (CS) verbunden. Diese Beobachtungen stehen in Einklang mit der von Robinson und Berridge begründeten Theorie der „*Incentive Sensitization*“ (Robinson and Berridge 1993) als Erklärungsmodell für die Entstehung süchtigen Verhaltens. Dieses Konzept betont den motivationalen Aspekt, der durch den besonderen Anreizcharakter einer Droge konstituiert wird. Die Autoren schlagen vor, dass neuro-adaptive Prozesse als Folge wiederholten Drogenkonsums zu einer Hypersensitivität mesolimbischer Belohnungsstrukturen führen, welche in der Folge sensitiviert sind auf Drogen und drogenassoziierte Reize. Ein wesentlicher Unterschied zu früheren Konzeptualisierungen von süchtigem Verhalten ist, dass der hedonische Aspekt, das „*liking*“ einer Droge, nicht im Vordergrund steht, sondern ausschließlich ihr Anreizwert, das „*wanting*“. Süchtiges Verhalten resultiert demnach aus einer pathologischen Motivation, welche ihren Ursprung in einer fehlerhaften dopaminergen Modulation klassisch konditionierter Pavlovscher Reize hat, während hedonische Prozesse eher dem Opioidsystem (Berridge and Robinson 1998, Berridge 2009, Berridge, Robinson et al. 2009) zugeschrieben werden sollten.

Das Lernen erhöhten Alkoholkonsum moderieren, konnte in Humanstudien repliziert werden (Park, Kahnt et al. 2010, Sebold, Deserno et al. 2014): Patienten mit einer

Alkoholabhängigkeit waren weniger stark in der Lage, ihr Verhalten an wechselnde Kontingenzen anzupassen (beispielsweise von einem Prädiktionsfehler zu lernen, Park, Kahnt et al. 2010) und zeigten häufiger habituelles, modellfreies, im Gegensatz zum modellbasierten, zielgerichteten Verhalten (Sebold, Deserno et al. 2014). Modellfreies Lernen könnte erklären, warum die Devaluation der Droge, trotz bereits eingetretener Sättigung und massiver negativer Konsequenzen auf sozialer, finanzieller und beruflicher Ebene, bei alkoholabhängigen Patienten nur eingeschränkt möglich ist. Die Wertigkeit eines früher belohnenden Stimulus kann der aktuellen Situation nicht angepasst werden, das „Modell der Welt“ wird nicht aktualisiert, sondern ein automatisiertes Streben nach der Substanz fortgesetzt. Über mathematische Modelle konnte bereits gezeigt werden, dass ein direkter Zusammenhang zwischen modellfreiem Lernen, einer veränderten Salienzattribution auf klassisch konditionierte Reize und der damit assoziierten phasischen Dopaminmodulation existiert (Huys, Tobler et al. 2014), welche zusammen ein Vulnerabilitätsmodell für die Entstehung süchtigen Verhaltens liefern.

Ein weiterer Faktor, der in der industrialisierten Gesellschaft zunehmend an Bedeutung für die Pathologie und das Verständnis der Suchterkrankung gewinnt, ist die chronische Stresserfahrung. Ergebnisse deuten darauf hin, dass der Faktor Stress (gemessen sowohl als frühkindliche Stresserfahrung als auch als chronisches Stresserleben über die letzten Jahre) sowohl auf molekularer Ebene (über Veränderungen im Methylierungsgrad von Genen, sogenannte epigenetische Veränderungen) als auch auf kognitiver Ebene über Veränderungen in der Belohnungsverarbeitung, zu dysfunktionalem Lernen führt (Casement, Shaw et al. 2015). Casement und Kollegen. (2014) berichteten, dass anhaltende Stresserfahrungen über Veränderungen der neuronalen Modulation von Belohnungserwartung im mPFC zu einem problematischen Alkoholkonsum beitragen können: in einer prospektiven Studie an Jugendlichen (15.-18. Lebensjahr) fanden die Autoren einen positiven Zusammenhang zwischen chronischer Stressexposition während der Adoleszenz und einem erhöhten riskanten Alkoholkonsum im Erwachsenenalter. Die erhöhte Stressexposition war gleichzeitig auch mit einer verminderten Aktivierung des mPFC während der Gewinnerwartung in einer Lernaufgabe verbunden, die Höhe dieser Aktivierung konnte kritischen Alkoholkonsum im Erwachsenenalter mit einer erklärten Varianz von bis zu 33% vorhersagen - ein eindrucksvoller Hinweis auf den Zusammenhang zwischen Stresserfahrung, Belohnungslernen und riskantem Alkoholkonsum. Gleichzeitig kann Stresserleben über Methylierung von Genen, die mit der „Erblichkeit“ der Alkoholabhängigkeit assoziiert wurden, dazu führen dass das Risiko für die Entstehung und Aufrechterhaltung der Erkrankung erhöht wird (Harlaar and Hutchison 2013). Diese Veränderungen sind potentiell reversibel und damit ein möglicher Indikator therapeutischer Wirksamkeit (Philibert, Penaluna et al. 2014).

## Die Studien

### 1.9 Studie 1: Devaluation und sequentielle Entscheidungsfindung: Die Verbindung zwischen zielorientiertem und modellbasiertem Verhalten

Zur Erfassung von Entscheidungsverhalten im Sinne von zielorientiertem versus habituellem und modellfreiem versus modellbasiertem Vorgehen bei instrumentellem Lernen wurden unterschiedliche Operationalisierungen beschrieben. Ähnlich wie in tierexperimentellen Studien wurden hierzu unter anderem Devaluationsaufgaben verwendet. Eine neuere Entwicklung im Rahmen der komputationalen Ansätze zur Beschreibung und Vorhersage von Entscheidungsverhalten ist die Anwendung sequentieller Entscheidungsfindungsaufgaben zur Bestimmung von modellfreiem vs. modellbasiertem Verhalten. Eine anhaltende theoretische Diskussion beschäftigt sich mit der Frage, ob diese unterschiedlichen Operationalisierungen auf gleiche oder zumindest überlappende theoretische Konstrukte zurückgehen. Diese Verbindung wird zwar konstatiert, wurde bislang jedoch nie direkt in Humanstudien getestet. Zur Klärung dieser Frage haben wir 18 gesunde Männer (mit einer ausführlichen neuropsychologischen Testbatterie zum Ausschluss kognitiver Defizite, sowie dem Strukturiertes Klinischen Interview für DSM-IV (mit einer ausführlichen neuropsychologischen Testbatterie zum Ausschluss kognitiver Defizite, sowie dem Strukturiertes Klinischen Interview für DSM-IV, Wittchen 1997, zum Ausschluss von Achse I und II Störungen), untersucht. Die Probanden absolvierten sowohl eine Devaluationsaufgabe zur Erfassung von zielgerichtetem vs. habituellem Verhalten als auch eine sequentielle Entscheidungsfindungsaufgabe zur Erfassung von modellbasiertem und modellfreiem Verhalten. Fünf Probanden mussten ausgeschlossen werden, da sie entweder nicht ausreichend devaluierten (4 Probanden) oder das *two-step*-Experiment (1 Proband) vorzeitig abbrachen.

#### *Devaluationsparadigma*

Die gesunden 18 Probanden lernten den probabilistischen Zusammenhang zwischen einem abstrakten Stimulus und einem appetitiven Reiz (Orangensaft oder Schokoladenmilch, beide Reize mussten zuvor als besonders angenehm unter einer Auswahl natürlicher Verstärker eingestuft werden), der in 75 % der Fälle bei Auswahl des belohnenden Stimulus über einen Schlauch direkt in den Mund des Probanden appliziert wurde. Bei Auswahl eines neutralen Stimulus wurde mit derselben Wahrscheinlichkeit Wasser appliziert. Nach Ablauf der Lernphase wurde eine Devaluation des appetitiven Stimulus durch Sättigung herbeigeführt (die Probanden erhielten so lange entweder Orangen oder Schokoladenpudding, bis das subjektive Gefühl der Sättigung eingetreten und der entsprechende Reiz als weniger appetitiv eingestuft wurde). In der darauf folgenden Extinktionsphase wurden die Probanden erneut aufgefordert, zwischen den gelernten abstrakten Stimuli zu wählen, ohne dabei



jedoch eine entsprechende Belohnung zu erhalten. Um das Antwortverhalten während der Extinktionsphase aufrechtzuerhalten ohne gleichzeitig neues Lernen zu provozieren, folgte auf die Auswahl der neutralen Stimuli weiterhin die Applikation von Wasser über den Schlauch, während auf den zuvor belohnten Stimulus keine Applikation folgte. Vor Beginn der Lernphase, vor und nach der Devaluation und nach der Extinktion wurde jeweils über eine visuelle Analogskala die subjektive Wahrnehmung von jedem Reiz als angenehm bzw. unangenehm erfasst. Ein Reiz galt dann als devaluiert, wenn die Probanden eine subjektive Sättigung und Abwertung des vormals als positiv eingestuften Reizes angaben. Hat eine erfolgreiche Devaluation stattgefunden, dann wird der modellfreie Lerner weiterhin häufiger während der Extinktionsphase den devaluierten Stimulus auswählen, während der zielgerichtete Lerner diesen nach der Devaluation meidet. Das Ausmaß von habituellem vs. zielgerichtetem Verhalten wurde über eine 2 (Zeit: vor/nach Devaluation) x 2 (Wert des ausgewählten Stimulus: valuiert/devaluiert) Anova erfasst.

#### *Sequentielles Lernen über die two-step Lernaufgabe*

In der *two-step* Aufgabe trafen dieselben 18 Probanden (an einem zweiten Testtag) zwei aufeinander folgende Entscheidungen zwischen jeweils zwei abstrakten Reizen (geometrische Figuren), um nach der zweiten Entscheidung einen Geldgewinn zu erhalten. Dabei war das Erscheinen des Stimuluspaares der zweiten Ebene mit einer definierten Wahrscheinlichkeit an die Wahl des Stimulus der ersten Ebene geknüpft: In 70% führte ein Stimulus der ersten Ebene zu einem bestimmten Stimuluspaar der zweiten Ebene (sogenannte *common transition*). In 30% der Fälle führte die Auswahl zu dem anderen der beiden Stimuluspaare auf der zweiten Ebene (sogenannte *rare transition*). Nachdem auf zweiter Ebene eine Auswahl getroffen wurde, folgte ein Geldgewinn (20 Cent) oder kein Gewinn entsprechend sich langsam verändernder Wahrscheinlichkeiten (sog. *random walks*).

Die Rationale bei dieser Entscheidungsaufgabe ist, dass man in seltenen Fällen bei einer Entscheidung auf der ersten Ebene für eine bisher erfolgreiche Option, die auf der zweiten Ebene zu einem häufig belohnten Stimulus führte, überraschenderweise vor eine Wahlmöglichkeit gestellt wird, in der beide Optionen kaum belohnt werden. Wer diese komplexe Aufgabenstruktur überblickt und versteht, dass dies ein seltener Misserfolg einer ansonsten häufig belohnten Wahlmöglichkeit ist, wird die bisher erfolgversprechende Wahl wiederholen. Wesentlicher Indikator für ein modellbasiertes vs. modellfreies Verhalten bei *two-step* war daher die Auswahl des Stimulus auf der ersten Ebene: Führt die Wahl auf erster Ebene zu einer *rare transition* auf die zweite Ebene und in der Folge zu einem Ausbleiben der Belohnung, dann wird der modellbasierte Lerner diesen Stimulus auf erster Ebene dennoch wieder auswählen, da er in 70% der Fälle zu einem Stimuluspaar mit hoher Belohnungswahrscheinlichkeit (nämlich zur *common transition*) führt. Genauso wird er bei

Belohnung, die auf eine *rare transition* folgt, das Stimuluspaar auf erster Ebene wechseln, da er gelernt hat, dass dieser Stimulus mit 70% Wahrscheinlichkeit zu einem nicht belohnten Stimuluspaar führt. Der modellbasierte Lerner wird sein Verhalten also trotz Belohnung anpassen und diesen Reiz nicht erneut auswählen. Der modellfreie Lerner ignoriert die Transitionsmatrix und orientiert sein Verhalten ausschließlich an dem Erhalt oder Ausbleiben der Belohnung auf der zweiten Ebene.

Das Ziel von modellfreien und modellbasierten Algorithmen ist es, jedem der dargestellten Stimuli einen spezifischen Wert Q (quality) zuschreiben zu können, welche in der Aufgabe in drei Stimuluspaaren präsentiert werden (sA, sB, sC). Der Index a (action) beschreibt die Stimulusauswahl, die Indizes i stehen für die Anzahl der Stufen (i = 1 oder 2) und t (trial) steht für die Anzahl der Durchgänge (t=200). Der Algorithmus zur Berechnung der modellfreien Parameter war SARSA (State-action-response-state-action, bei der die stimulusbezogenen Handlungswerte die Handlungsplanung modellieren; Sutton 1998) und berechnete sich wie folgt:

$$[1] \quad Q_{MF_{s_{i,t+1}, a_{i,t+1}}} = Q_{MF_{s_{i,t}, a_{i,t}}} + \alpha (r_{i,t} - Q_{MF_{s_{i,t}, a_{i,t}}})$$

$$[2] \quad r_{i,t} = r_{i,t} + \alpha (Q_{MF_{s_{i+1,t}, a_{i+1,t}}} - Q_{MF_{s_{i,t}, a_{i,t}}})$$

Insbesondere gilt,  $r_{s_{i,t}} = 0$ , das bedeutet, dass es auf erster Ebene keine Belohnung gab.

Der Wert (Q) jedes Stimulus (s) orientierte sich ausschließlich an der Belohnung (r) auf der zweiten Ebene. Für jede Stufe i wurden unterschiedliche Lernraten ( $\alpha_i$ ) zugelassen. Der Algorithmus erlaubte eine Aktualisierung der Q-Werte der ersten Stufe durch den Prädiktionsfehler (d) auf zweiter Stufe, welcher durch  $\lambda$  gewichtet wurde.

$$[3] \quad Q_{MF_{s_{1,t+1}, a_{1,t+1}}} = Q_{MF_{s_{1,t}, a_{1,t}}} + \lambda (d_{1,t} - Q_{MF_{s_{1,t}, a_{1,t}}})$$

$\lambda$  berücksichtigt den Haupteffekt *reward* und gilt daher als Indikator für modellfreies Verhalten.  $\lambda$  kann aber nicht die Interaktion *reward* x *state* abbilden (wie sie für die Vorhersage eines modellbasierten Verhaltens notwendig wäre).

Stattdessen kann mit der Einführung der Transitionsmatrix in den modellbasierten Algorithmus diese Interaktion abgebildet und berücksichtigt werden. Der modellbasierte Algorithmus nutzt die Werte des modellfreien Algorithmus der zweiten Ebene, unter Berücksichtigung einer Transitionsmatrix und berechnet damit die Werte der ersten Stufe durch Multiplikation der besseren Option auf zweiter Stufe mit den Transitionswahrscheinlichkeiten:

$$[4] \quad Q_{MB_{s_A, a}} = P(s_B | s_A, a) \max_{a'} Q_{MF_{s_B, a'}} + P(s_C | s_A, a) \max_{a'} Q_{MF_{s_C, a'}}$$

$$[5] \quad Q_{MB_{s_{2,t}, a_{2,t}}} = Q_{MF_{s_{2,t}, a_{2,t}}}$$

Abschließend werden  $Q_{MF}$  und  $Q_{MB}$  in einem Hybridalgorithmus miteinander gekoppelt:

$$[6] \quad Q_{s_A, a} = \alpha Q_{MB_{s_A, a}} + (1 - \alpha) Q_{MF_{s_A, a}}$$

$$[7] Q_{S_2,a} = Q_{MB_{S_2,a}} = Q_{MF_{S_2,a}}$$

Der Parameter  $\omega$  stellt eine Gewichtung des relativen Einflusses von modellfreiem und modell-basiertem Lernen auf die Auswahl der Stimuli und ihrer Werte dar und war deshalb für die folgenden Auswertungen der wichtigste und interessanteste Parameter. Um die Wahrscheinlichkeit für eine Auswahl  $a$  auf Stufe  $i$  im trial  $t$  berechnen zu können, wurde folgender Softmax Algorithmus verwendet:

$$p(i, a, t) = \frac{\exp(\beta_i(Q_{S_{i,t},a'} + rep(a)))}{\sum_{a'} \exp(\beta_i(Q_{S_{i,t},a'} + rep(a'))))$$

$\beta$  kontrolliert in diesem Modell für die Zufälligkeit der Auswahl, welche auf beiden Stufen unterschiedlich berechnet wurde. Der zusätzliche Parameter  $\rho$  erfasst Perseverationen im Auswahlverhalten auf der ersten Ebene, und  $rep$  (*repetition*) ist eine Indikatorfunktion, die 1 entspricht, wenn auf erster Ebene die gleiche Auswahl im folgenden Durchgang erneut getroffen wird (Lau and Glimcher 2005, Daw, Gershman et al. 2011). Zusammenfassend hat der Algorithmus insgesamt 7 Parameter und kann auf 2 spezifische Fälle reduziert werden:  $\omega = 1$  (4 Parameter) und  $\omega = 0$  (5 Parameter). Um eine Normalverteilung innerhalb der Parameterschätzer zu erhalten, wurden die Parameter entweder durch eine logistische ( $\alpha, \lambda, \omega$ : Parameter zwischen 0-1) oder exponentielle ( $\beta$ : nicht begrenzt) Transformation angepasst. Die freien Parameter wurden zunächst mit Hilfe einer Maximum-Likelihood Methode geschätzt. Diese Werte dienten dann als Ausgangspunkt für ein iteratives *expectation-maximization* Verfahren, wobei letztlich die a posteriori Schätzer verwendet wurden. Für eine ausführliche Diskussion dieser Methode siehe Huys, Cools et al. (2011), sowie Huys, Eshel et al. (2012).

#### *Korrelationsanalyse von zielgerichtetem und modellbasiertem Verhalten*

In der selektiven Devaluationsaufgabe wurde das Ausmaß des modellbasierten Verhaltens über die Berechnung des Interaktionswertes '(valued stimulus pre devaluation – valued stimulus post devaluation) – (devalued stimulus pre devaluation - devalued stimulus post devaluation)' entsprechend der getroffenen Verhaltensauswahl ermittelt. Ein hoher Wert entspricht einem eher zielgerichtetem Verhalten, die Probanden haben hier häufiger den *valued* gegenüber dem *devalued* Stimulus nach der Devaluation gewählt. Ein ähnlicher Interaktionsterm für modellbasiertes Verhalten kann auch in der *two-step* Aufgabe über die Analyse der Wahrscheinlichkeit, einen zuvor gewählten Stimulus erneut zu wählen (*stay probability*), auf der ersten Stufe berechnet werden: '(rewarded common stimulus choice – rewarded rare stimulus choice) – (unrewarded common stimulus choice - unrewarded rare stimulus choice)'. Modellbasiertes Verhalten entspricht dem Verbleib (*stay*), nachdem eine Belohnung auf einen *common state* oder keine Belohnung auf einen *rare state* folgte. Des Weiteren wurde der bereits oben beschriebene Parameter  $\omega$  als Resultat der

mathematischen Modellierung für die Balance von modellfreiem und modellbasiertem Verhalten als Korrelationsparameter verwendet.

**Ergebnisse:** Im Devaluationsparadigma zeigte sich, dass die Probanden während der Testphase lernten, den mit einer höheren Wahrscheinlichkeit belohnten Stimulus in allen 3 Bedingungen signifikant häufiger auszuwählen (devaluiert:  $T=-3.50$ ,  $p=0.004$ , valuiert:  $T=-2.60$ ,  $p=0.022$ , neutral:  $T=-2.73$ ,  $p=0.017$ ). Instrumentelles Lernen wurde also erfolgreich induziert. Es fand eine selektive Devaluation bei allen eingeschlossenen Probanden statt, d.h. nach der Devaluation wurde die entsprechende Belohnung (Orangensaft oder Schokolade) selektiv abgewertet ( $T=2.67$ ,  $p=0.019$ ), während der nicht devaluierte Stimulus weiter als angenehm eingestuft und weiterhin ausgewählt wurde. In der Extinktionsphase zeigte sich eine signifikante Zeit (pre/post Training) x Bedingungs (devaluiert/valuiert/neutral) Interaktion ( $F=5.200$ ,  $p=0.04$ ). Die Höhe dieses Interaktionsterms diente als Maß für die Höhe des zielgerichteten Verhaltens, wie bereits von Valentin und Kollegen (2007) berichtet. In der *two-step*-Aufgabe zeigte sich ein signifikanter Haupteffekt für Belohnung ( $F=14.1$ ,  $p=0.002$ ) als Indikator für modellfreies Verhalten und eine signifikante Interaktion von Transitionsverhalten nach Belohnung auf der ersten Stufe in Abhängigkeit von der Bedingung ( $F=6.05$ ,  $p=0.026$ ), als Indikator für modellbasiertes Verhalten. Beide Komponenten, die modellbasierte und die modellfreie, hatten also einen signifikanten Einfluss auf das Verhalten.

13 Probanden (Mittleres Alter [in Jahren]=45, SD [Jahre]=9) wurden in die finale Analyse eingeschlossen (5 der vormals 18 Probanden mussten ausgeschlossen werden, da sie die Einschlusskriterien nicht erfüllten oder die Aufgabe vorzeitig abbrachen). Es zeigte sich eine signifikante Korrelation der Interaktionsparameter aus beiden Lernaufgaben (Spearman Rangkorrelation  $r = 0.708$ ,  $p<0.005$ ). Zusätzlich zeigte sich ein signifikanter Zusammenhang zwischen dem Interaktionsparameter der Devaluation und dem Modellparameter für modellbasiertes Verhalten ( $\omega$ ) aus der *two-step*-Lernaufgabe (Spearman Rangkorrelation  $r = 0.498$ ,  $p<0.005$ ). Die Korrelationen der Parameter für modellfreies Verhalten ( $\lambda$ ) und habituelles Verhalten (Interaktionsterm) wurden nicht signifikant ( $p>0.75$ ). Habituelles und modellfreies Lernen scheinen zumindest teilweise durch unterschiedliche Prozesse evoziert oder unterschiedlich operationalisiert zu sein.

### **1.10 Studie 2: Der Zusammenhang zwischen chronischem Stresserleben, dem neuronalen Prädiktionsfehlersignal und Intelligenz**

Fluide Intelligenz wird im Gegensatz zur kristallinen Intelligenz vor allem mit der Fähigkeit für schnelles Problemlösen in Verbindung gebracht. Sowohl der Einfluss von chronischem Stress als auch Lernen und seine neuronalen Korrelate (insbesondere Veränderungen der Aktivierung im VS auf das Prädiktionsfehlersignal) sind mit kognitiven Variablen wie der

fluiden Intelligenz assoziiert worden (Garrett, Grady et al. 2010, Otto, Raio et al. 2013, Schlagenhaut, Rapp et al. 2013). Hinweise über den Einfluss von Stress auf die dopaminerge Neurotransmission im mesolimbischen System (insbesondere der VTA und dem VS) ergeben sich aus tierexperimentellen Studien, welche Veränderungen an der Synapse nach akuter und chronischer Stressexposition berichten (Cao, Covington et al. 2010, Valenti, Lodge et al. 2011). Ein Effekt von akutem Stress auf das neuronale Prädiktionsfehlersignal im VS konnte entsprechend bereits in Humanstudien nachgewiesen werden (Robinson, Overstreet et al. 2013). Über die Effekte chronischer Stressexposition auf das neuronale Lernsignal und dessen Zusammenhang mit Intelligenz (fluide und kristalline) gab es bislang keine Evidenz. Zur Untersuchung dieser Frage wurde die 2. Studie konzipiert. Wir haben in dieser Arbeit die Interaktion zwischen einem adaptiven, modellfreien Lernsignal (unter Verwendung einer probabilistischen instrumentellen Lernaufgabe in Kombination mit fMRI), chronischem Stress und Intelligenz an einer Stichprobe von 16 gesunden Männern (Mittleres Alter [Jahre]=38.4, SD [Jahre]=11.9) untersucht. Chronisches Stresserleben wurde über die *Lifetime Stressor Checklist* von Holmes und Rahe (1967) erfasst. Hier geben die Probanden retrospektiv über einen Zeitraum von 24 Monaten an, welche von 42 als stressreich eingestuften Lebensereignissen auf sie zutreffen. Es werden Aspekte aus dem familiären Umfeld (Hochzeit, Tod eines nahen Angehörigen), der Ausbildung (Abschluss, Abbruch einer Ausbildung), sowie finanzielle Variablen (Schulden, Hauskauf) berücksichtigt. Die Summe dieser Ereignisse ging als moderierende Variable in die Analysen des Lernverhaltens ein.

Intelligenz wurde über eine Faktorenanalyse (Varimax Rotation mit Eigenwert größer 1.0) aus 10 neuropsychologischen Testverfahren (adaptiert aus der Berliner Altersstudie, Lindenberger and Baltes 1997) in eine kristalline und eine fluide Komponente eingeteilt.

### *Sequentielles Lernen über die reversal Lernaufgabe*

Das Prädiktionsfehlersignal wurde über die *reversal* Lernaufgabe (Park, Kahnt et al. 2010) evoziert. Ähnlich wie bei der *two-step* Lernaufgabe handelt es sich um eine zweistufige probabilistische instrumentelle Lernaufgabe über 200 Durchgänge. Die Belohnung war in diesem Fall ein lachendes vs. ein unglückliches Gesicht. Im Verlauf des Experimentes kam es mehrfach zu einem Wechsel der Belohnungswahrscheinlichkeit (80/20) nach der Auswahl des abstrakten Stimulus (geometrische Figur) auf der zweiten Stufe. Immer dann, wenn ein solcher Wechsel eines zuvor häufig belohnten Stimulus (80%) zu einem weniger belohnten (20%) stattfand, wurde der Prädiktionsfehler berechnet. Ein Wechsel der Belohnungswahrscheinlichkeit fand statt, wenn zwei Kriterien erfüllt waren: 1) mind. 10 Durchgänge wurden absolviert und 2) mindestens 70% korrekte Antworten im gesamten Block. Wenn nach 16 Durchgängen das Lernkriterium von 70% korrekten Antworten nicht erreicht wurde, fand automatisch ein Wechsel der Belohnungswahrscheinlichkeit statt. Die

Berechnung des Prädiktionsfehlers erfolgte über ein Rescorla-Wagner Modell (Sutton 1998). Das Modell postuliert, dass die Wahrscheinlichkeit ein Verhalten  $a$  (action) im trial  $t$  zu zeigen, proportional zum Erwartungswert  $Q_t(a)$  ist. Die Wahrscheinlichkeit  $p(a)$ , eine bestimmte Auswahl ( $a$ ) zu treffen, wurde wie bereits oben für *two-step* beschrieben über eine Softmax Funktion berechnet.

Die Erwartungswerte wurden wie folgt in Verhaltenswahrscheinlichkeiten umgerechnet:

$$p(a|Q_t) = \exp(Q_t(a)) / (\sum_{a'} \exp(Q_t(a'))).$$

Die Belohnungserwartung  $Q_t(a_{\text{chosen}})$  wurde schrittweise angepasst:

$$Q_t(a_{\text{chosen}}) = Q_{t-1}(a_{\text{chosen}}) + \varepsilon(R_t - Q_{t-1}(a_{\text{chosen}}))$$

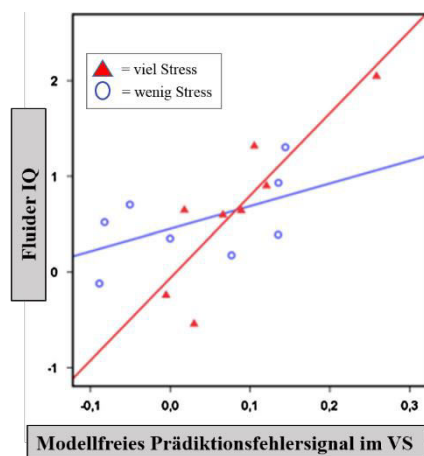
Hier stellt  $\varepsilon$  die Lernrate dar und  $R$  (Reward) die erhaltene Belohnung. Die Differenz  $R_t - Q_{t-1}(a)$  entspricht dem Prädiktionsfehler. Somit gilt: wenn  $R_t > Q_{t-1}(a_{\text{chosen}})$ , also die Belohnung ( $R$ ) zum Zeitpunkt  $t$  größer war, als der Wert ( $Q$ ) der gewählten Handlung ( $a_{\text{chosen}}$ ) zum Zeitpunkt  $t$ , dann erhöht sich der Wert von  $Q_t(a_{\text{chosen}})$  und erhält damit im folgenden Durchgang ( $t+1$ ) eine höhere Wahrscheinlichkeit ausgewählt zu werden. Die Variable  $R_t$  bestimmt den direkten Einfluss der Belohnung auf das Verhalten des Probanden und entspricht der sog. Belohnungssensitivität  $\beta$ . Die Variable  $\beta$  nimmt den Wert  $R_t = \beta_{\text{rew}}$  an, wenn eine Belohnung folgt und  $-\beta_{\text{pun}}$ , wenn eine Bestrafung folgt. Für jeden Proband wurde eine individuelle Lernrate  $\varepsilon$  und die Belohnungssensitivität für Belohnung und Bestrafung ( $\beta_{\text{rew}}$  and  $\beta_{\text{pun}}$ ) als maximum a posteriori estimates dieser Parameter mit Hilfe eines *gaussian prior* berechnet. Die freien Parameter ( $\beta$  und  $\varepsilon$ ) wurden wie bereits oben beschrieben mit Hilfe der *expectation-maximization* Methode bestimmt. Basierend auf diesen Parametern wurde für jeden einzelnen Probanden  $i$  eine individuelle Zeitreihe der PEs berechnet:

$$PE_t^i = R^i(t) - Q_{t-1}^i(a_t).$$

Diese Zeitreihe ging dann als Regressor in die fMRI Analysen ein, um das neuronale Prädiktionsfehlersignal im VS erfassen zu können (O'Doherty, Hampton et al. 2007, Schlagenhaut, Rapp et al. 2013).

**Ergebnisse:** Die Probanden berichteten im Mittel 16.31 (SD =4.59) als stressreich erlebte Ereignisse aus der Liste von 42 Ereignissen. Der mittlere fluide IQ lag bei 0.60 (SD = 0.64) nach der Varimax Rotation. Der mittlere kristalline IQ lag bei 32.31 (SD = 5.67). Im Gegensatz zum kristallinen IQ Wert zeigt der fluide IQ eine signifikante Abnahme mit zunehmendem Alter der Probanden (Pearson  $r = -0.61$ ,  $p=0.013$ ). Die Anzahl der als stressreich berichteten Lebensereignisse war weder mit dem IQ-Wert (Pearson  $r = 0.24$ ,  $p=0.376$ ) noch mit dem Alter (Pearson  $r = -0.11$ ,  $p=0.694$ ) assoziiert. Es zeigte sich eine signifikante Aktivierung des linken VS als Korrelat des modellfreien Prädiktionsfehlersignals berechnet aus der *reversal*-Lernaufgabe über alle 16 Probanden (left:  $T=3.38$ ,  $p_{\text{FWE}}$  korrigiert für  $_{\text{VS}} = 0.026$ ), sowie eine trendweise Aktivierung des rechten VS ( $T=2.47$ ,  $p_{\text{FWE}}$  korrigiert für  $_{\text{VS}} = 0.057$ ). Dieses modellfreie Prädiktionsfehlersignal im VS korrelierte signifikant (auch nach

Alterskorrektur) mit dem fluiden IQ (links:  $T=3.89$ ,  $p_{\text{FWE}} \text{ korrigiert für VS} = 0.012$ ; rechts:  $T= 3.71$ ,  $p_{\text{FWE}} \text{ korrigiert für VS} = 0.013$ ) und der Summe stressreicher Lebensereignisse (links:  $T=3.02$ ,  $p_{\text{FWE}} \text{ korrigiert für VS} = 0.047$ , rechts fand sich kein signifikanter Effekt  $p>0.2$ ). Die Interaktion Stress x modellfreies Prädiktionsfehlersignal im VS wurde ebenfalls signifikant ( $\beta=1.19$ ,  $T=2.53$ ,  $p<0.05$ ). Um diese Interaktion besser verstehen zu können, wurde eine Aufteilung der Probanden in viel und wenig Stresserleben vorgenommen (Mediansplit der Summe stressreicher Lebensereignisse). Die Ergebnisse zeigten, dass der Zusammenhang zwischen dem modellfreien Prädiktionsfehlersignal im VS und dem fluiden IQ dann besonders hoch war, wenn die Probanden viel Stresserleben in der Vergangenheit berichtet hatten (siehe **Abbildung 1**, Median-Split-Analyse,  $R^2 = 0.76$ ). Im Gegensatz zum fluiden IQ zeigte sich auf neuronaler Ebene kein Einfluss des kristallinen IQ auf das modellfreie Prädiktionsfehlersignal und keine Interaktion mit chronischem Stresserleben ( $p>0.05$ ). Es scheint also einen selektiven Einfluss von chronischer Stresserfahrung auf die Assoziation zwischen dem modellfreien Prädiktionsfehlersignal und der fluiden Intelligenzkomponente zu geben.



**Abb. 1:** Median-Split-Analyse: Probanden, die mehr Stresserleben (*Lifetime Stressor Checklist*) berichteten, zeigten einen stärkeren Zusammenhang zwischen dem modellfreien Prädiktionsfehlersignal im VS und dem fluiden IQ ( $R^2 = 0.76$ ) als Probanden mit wenig Stresserleben ( $R^2 = 0.20$ ). Der Interaktionsterm wurde signifikant ( $b = 1.19$ ,  $T = 2.53$ ,  $p<0.05$ ). Die *Outlier-Analyse* zeigte keine signifikanten Abweichungen ( $Z\text{-Wert} < 3$ ,  $\text{Cook's distance} < 1$ ,  $\text{max} = 0.218$ ,  $\text{SD} = 0.061$ ).

### 1.11 Studie 3: Wie wird Suchtverhalten „erlernt“? Risikofaktoren und Lernmechanismen die zur Entstehung und Aufrechterhaltung der Alkoholabhängigkeit beitragen

Die Entstehung von abhängigem Verhalten wurde mit zahlreichen Variablen in Verbindung gebracht, darunter insbesondere eine hohe genetische Prädisposition: die geschätzte „Erblichkeit“ der Alkoholabhängigkeit liegt bei 40-60% (Nieratschker 2013, Gelernter, Kranzler et al. 2014). Kendler und Kollegen (Kendler, Ji et al. 2015) konnten in einer umfassenden Adoptionsstudie an über 18.000 Probanden nachweisen, das genetische

Faktoren in gleichem Maße wie Umweltfaktoren (inklusive Erziehung, familiäres Umfeld, und frühkindliche Stresserfahrung) das Risiko einer Alkoholabhängigkeit erhöhen. Erste positive Erfahrungen mit Alkohol, die Erhöhung der subjektiv erlebten Selbstwirksamkeit und die Reduktion negativer Emotionen sind Variablen, die über Alkoholkonsum reguliert werden, und entsprechend als Risiko- und Resilienzfaktoren gleichermaßen gewertet werden müssen (Smith, Goldman et al. 1995, Chartier, Hesselbrock et al. 2010).

Stresserleben im Sinne von chronischem Stress scheint ähnlich wie frühkindliche und sogar pränatale Stresserfahrung einen erhöhten Alkoholkonsum zu evozieren (Huot, Thirvikraman et al. 2001, de Wit, Soderpalm et al. 2003, Cruz, Quadros et al. 2008), wobei der sedierende und angstlösende Effekt eine wesentliche Rolle zu spielen scheint. Gleichzeitig konnte in tierexperimentellen Studien nachgewiesen werden, dass bereits postnatale Stresserfahrung zu epigenetischen Veränderungen im Nachwuchs führt und entsprechende Veränderungen der Hypothalamus-Hypophysen-Nebennierenrinden-Achse (HPA)-Achse und des Glukokortikoid (GR)-Rezeptors bewirkt. Diese Prozesse führten in der Folge zu einer erniedrigten Resilienz gegenüber chronischem Stresserleben, waren im Gegensatz zu Gen-assoziierten Prozessen aber potentiell reversibel (Weaver, Cervoni et al. 2004).

Die geschätzte hohe Erbllichkeit der Alkoholabhängigkeit wird dementsprechend nicht allein durch die Weitergabe bestimmter Genotypen von biologischen Eltern auf ihre Kinder bedingt, sondern kann zusätzlich durch das oft von sozialem Stress geprägte Umfeld zu einer anhaltenden Veränderung im Ablesen des Erbgutes führen, welche die Basenabfolge jedoch unverändert lässt (Weaver, Cervoni et al. 2004). Ein oft beschriebener und vererbbarer Risikofaktor der Alkoholabhängigkeit ist z.B. die Verträglichkeit hoher Alkoholmengen, ohne die damit verbundenen aversiven Folgen der akuten Intoxikation. Eine gute Alkoholverträglichkeit konnte mit der Methylierung - und damit der epigenetischen Steuerung - in der Nähe des ALDH1A2 Gens in Verbindung gebracht werden (Harlaar, Bryan et al. 2014). Das ALDH1A2 Gen reguliert die Metabolisierung des Alkohols. Die Methylierung an einer anderen Stelle in der Nähe desselben Gens war mit dem Auftreten einer verminderten Kontrolle nach Alkoholkonsum verbunden, somit können umweltassoziierte Veränderungen in der Ausprägung eines des Erbgutes über epigenetische Mechanismen das Risiko für eine Abhängigkeitserkrankung verändern. Weitere Gene, die u.a. mit der Erbllichkeit der Alkoholabhängigkeit eng assoziiert waren, sind das ALDH-2 Gen (Maezawa, Yamauchi et al. 1995), das GABRA-1 Gencluster (Dick, Plunkett et al. 2006), Kalium-Kanal (KCNMA1) kodierende Gene (Davies, Pierce-Shimomura et al. 2003, Bettinger and Davies 2014), Serotonin-Transporter (5-HTTLPR) assoziierte Gene (Heinz, Schafer et al. 2003, Hu, Oroszi et al. 2005, Wrase, Reimold et al. 2006) und Dopamin assoziierte Gene (Kienast, Hariri et al. 2008, Kumakura, Gjedde et al. 2013).

Veränderungen im Lernen, die prä-morbid und als Folge der Suchterkrankung berichtet wurden, scheinen ebenfalls eine wichtige Rolle bei der Entstehung und Aufrechterhaltung



der Alkoholabhängigkeit zu spielen und waren sowohl mit genetischen als auch mit Umweltfaktoren assoziiert. Eine herausragende Funktion wurde in diesem Zusammenhang dem Dopamin zugeschrieben, welches spätestens seit der 1997 von Wolfram Schultz veröffentlichten Studie als *das* neuronale Substrat von Belohnungslernen und Belohnungserwartung galt (Schultz, Dayan et al. 1997) und gleichzeitig einen starken Zusammenhang mit süchtigem Verhalten (Heinz, Schafer et al. 2003) hat. Entsprechend postulierten Robinson und Berridge in ihrer Theorie der „*Incentive Sensitization*“ (Robinson and Berridge 1993), dass süchtiges Verhalten aus einer pathologischen Sensitivierung motivationaler Strukturen resultiert, welche als Zeichen der Suchterkrankung suchtrelevanten Reizen eine höhere Wertigkeit (Salienz) zuweisen, während neutrale Reize an Bedeutsamkeit verlieren. Diese dysfunktionale Salienzattribution war auf neuronaler Ebene mit einer veränderten dopaminergen Modulation im Bereich mesolimbischer Strukturen assoziiert. Studien an alkoholabhängigen Patienten konnten diese Befunde zum Teil replizieren und zeigten eine veränderte Aktivierung auf suchtspezifische Reize in der Amygdala (Kienast, Hariri et al. 2008) und im VS (Wrase, Schlagenhauf et al. 2007, Beck, Wustenberg et al. 2012, Kienast, Schlagenhauf et al. 2013). Die beschriebenen dysfunktionalen Aktivierungsmuster konnten auf entsprechende Veränderungen in der dopaminergen Neurotransmission zurückgeführt werden (Di Chiara and Imperato 1988, Robbins and Everitt 1999, Volkow, Wang et al. 2006).

Auf Rezeptorebene hat die verstärkte Dopaminfreisetzung bei chronischem Alkoholkonsum eine Verminderung der Dopamin D2 Rezeptoren im VS zur Folge (Heinz, Batra et al. 2012). Neben den D2-Rezeptoren befinden sich D3-Rezeptoren im VS, die durch Blockade (im Tiermodell) zu reduziertem Alkoholkonsum führten (Erritzoe, Tziortzi et al. 2014). Die Befunde zur Wirksamkeit der D3-Rezeptor Blockade sind jedoch widersprüchlich, da bei abstinenten Alkoholabhängigen gegenüber Gesunden D3-Rezeptoren vorwiegend im Hypothalamus - und nicht wie im Tiermodell beschrieben im Striatum - vermindert waren - so dass die klinische Relevanz dieses Befundes noch unklar ist (Erritzoe, Tziortzi et al. 2014).

Während sich dysfunktionale Lernprozesse auf neuronaler Ebene über eine Veränderung der dopaminergen Neurotransmission im VS am Menschen replizieren ließen, zeigte sich auf der Verhaltensebene eine Tendenz bei alkoholabhängigen Patienten, Entscheidungen eher modellfrei zu treffen (Sebold, Deserno et al. 2014). Zusätzlich konnte bei alkoholabhängigen Patienten eine Tendenz nachgewiesen werden, vermehrt auf klassisch konditionierte Reize während einer instrumentellen Lernaufgabe zu reagieren (sogenannter PIT Effekt). Diese Tendenz war mit einer veränderten Aktivierung im VS auf suchtspezifische Reize und einer erhöhten Rückfallrate assoziiert.

Über die mathematische Modellierung und Prädiktion von Verhalten als modellfrei und modellbasiert gelang es schließlich, einen direkten Zusammenhang zwischen modellfreiem Verhalten, der phasischen Dopaminantwort auf konditionierte Reize im Tierexperiment, und

süchtigem Verhalten herzustellen (Huys, Tobler et al. 2014). Sollte es möglich sein, in der Zukunft Variablen zu erfassen, welche vor einem Überwiegen von automatisiertem, habituellem Verhalten schützen, so wäre ein Training dieser Fähigkeiten in der Behandlung und Prävention von Sucht bedeutsam. Erste Erfolge zeigten sich durch die sogenannte Approach Avoidance Task (AAT), in der alkoholabhängige Patienten trainiert wurden, alkoholassoziierte Reize mehr von sich wegzuschieben und so eine automatisierte Annäherungstendenz überwunden wurde. Insgesamt konnte eine Rückfallreduktion von 13% nach einem Jahr erreicht werden (Eberl, Wiers et al. 2013).

Maladaptive Lernprozesse konnten in der Zusammenschau der dargestellten Studien gemeinsam (und in Interaktion) mit einer erhöhten Vulnerabilität durch genetische und epigenetische Faktoren sowie einer erhöhten Stressexposition als wesentliche Komponenten für die Entstehung und Aufrechterhaltung der Alkoholabhängigkeit identifiziert werden. In der Zukunft müssen prospektive randomisierte Studien folgen, um eine kausale Interpretation der Ergebnisse zu ermöglichen, da die aktuellen Erhebungen in der Mehrzahl aus korrelativen Berechnungen von Querschnittsuntersuchungen stammen.

## **1.12 Diskussion und Ausblick**

In der dargestellten Arbeit wurden drei gängige Aufgaben zur Erfassung von belohnungsabhängigem Lernen an gesunden Probanden auf verhaltens- und neuronaler Ebene erfasst. In einem ersten Schritt konnte gezeigt werden, dass die historisch und konzeptuell gewachsene Dichotomie des modellfreien vs. modellbasierten Lernens zu einem großen Teil dem Konzept des habituellen vs. zielgerichteten Lernens folgt. Zielgerichtetes Verhalten, wie es während der selektiven Devaluationsaufgabe erfasst wird, stellt dementsprechend eine von mehreren Facetten eines modellbasierten Verhaltens dar. Weitere Facetten werden beispielsweise über instrumentelle *reversal*-Lernaufgaben (Hampton, Bossaerts et al. 2006, Li and Daw 2011, Schlagenhaut, Rapp et al. 2013), Pavlovsche Konditionierungsprozesse (Huys, Eshel et al. 2012, Prevost, McNamee et al. 2013) und den sogenannten Pavlovian-to-Instrumental-Transfer (Garbusow, Schad et al. 2014) erfasst. Dabei hat neben dem Ausmaß der einzelnen Komponenten für sich genommen (modellfrei vs. modellbasiert) insbesondere auch deren Gleichgewicht und Zusammenspiel beim Lösen einer Aufgabe eine besondere Bedeutung. Vermutlich handelt es sich bei diesem Zusammenspiel um ein interindividuell relativ stabiles Merkmal (Doll, Simon et al. 2012, Dolan and Dayan 2013). Dennoch wurde die zeitliche Stabilität dieses Merkmals bislang nicht ausreichend nachgewiesen, wie es z.B. für kognitive Variablen wie das Arbeitsgedächtnis mehrfach repliziert werden konnte (Klein and Fiss 1999, Waters and Caplan 2003). Eine weitere Frage, die sich daraus ergibt ist, inwiefern beide Lernstrategien (modellfrei und modellbasiert) miteinander um Handlungsansätze konkurrieren (Daw, Niv et al. 2005). Wenn es um das Abwägen der Effizienz geht, so ist das modellbasierte System oft

kostenintensiver, da mit einem größeren kognitiven Aufwand und entsprechend mehr Verarbeitungskapazität verbunden. Lee und Kollegen (2014) sprechen daher von einer unabhängigen Instanz, dem Vermittler (*arbitrator*), der einen stetigen Vergleich beider Systeme durchführt, nach Abwägung der Effizienz ein System zuweist und dann entsprechende Handlungsoptionen, je nach Anforderung der Aufgabe und Kompetenz der zur Verfügung stehenden Systeme, initiiert. Die hier durchgeführte Aufgabe zur sequentiellen Entscheidungsfindung (*two-step*) und die selektive Devaluationsaufgabe lassen beide eine Trennung in modellfreies/modellbasiertes bzw. habituelles/zielgerichtetes Verhalten zu. Die Erfassung eines Vermittlers zwischen den Instanzen ist jedoch mit keiner der Aufgaben möglich. Einschränkend kommt hinzu, dass ein Zusammenhang lediglich für modellbasiertes und zielgerichtetes Lernen nachgewiesen werden konnte (über die Korrelation zwischen dem Parameter  $\omega$  und dem Interaktionsterm der Devaluationsaufgabe). Ein Zusammenhang zwischen dem Parameter für modellfreies Lernen ( $\lambda$ ) und einem niedrigen Interaktionsterm (als Indikator für habituelles Lernen) konnte nicht nachgewiesen werden. Diese fehlende Korrelation entspricht Studienergebnissen von Valentin und Kollegen (Valentin, Dickinson et al. 2007, 2009), die darauf hinweisen, dass auf neuronaler Ebene habituelles Lernen mit der selektiven Devaluationsaufgabe nicht in den zu erwartenden spezifischen Arealen (wie z.B. Putamen/VS) abgebildet werden konnte. Entsprechend konnten Gillan und Kollegen (2015) die Sensitivität gegenüber einer selektiven Devaluation mit modellbasiertem Verhalten vorhersagen, während sich kein Zusammenhang zwischen modellfreiem Verhalten und der Devaluationssensitivität fand. Es scheint also so zu sein, dass selektive Devaluationsaufgaben eher geeignet sind, zielgerichtetes Verhalten als eine Variante des modellbasierten Lernens abzubilden. Dieses steht im Gegensatz zum habituellem Lernen, welches besser mit Aufgaben erfasst werden kann, die einen Trainingseffekt, und sog. Überlernen (*overtraining*) provozieren können (Tricomi, Balleine et al. 2009). Um modellfreies Verhalten in der *two-step* Aufgabe erfassen zu können, wurden die Belohnungswahrscheinlichkeiten auf der zweiten Ebene langsam (nach sog. *random walks*) verändert. Hiermit sollte verhindert werden, dass die modellfreie Lernstrategie gegenüber dem kontinuierlichen Lernen und Abgleichen der Werte auf der zweiten Ebene sehr viel effizienter und damit vorteilhafter im Vergleich zum modellbasierten Lernen wird. Dieser Aspekt verhindert jedoch gleichzeitig, dass ein Überlernen durch immer wieder gleiche unveränderte Abläufe stattfindet, was eine wichtige Komponente des modellfreien Lernens darstellt (Tricomi, Balleine et al. 2009). Es ist also wichtig an dieser Stelle darauf hinzuweisen, dass habituelles Lernen als eine Facette des modellfreien Verhaltens nur in Anteilen durch die dargestellten Aufgaben erfasst werden konnte, während modellbasiertes und zielgerichtetes Verhalten enger miteinander verknüpft waren und besser durch die dargestellten Aufgaben erfasst werden konnten. Es besteht also weiterhin der Bedarf nach Operationalisierungen, die in der Lage sind, beide Systeme exakt voneinander zu trennen

und gleichzeitig auch eine übergeordnete, vermittelnde Instanz erfassen können, die zwischen den Systemen abwägt und ihre individuelle Ausprägung moderiert. Beispielsweise konnten Sjoerds und Kollegen (2013) in einer etwas veränderten Devaluationsaufgabe von de Wit (2009) tatsächlich nachweisen, dass Patienten mit einer Alkoholabhängigkeit im Vergleich zu gesunden Kontrollen mehr habituelles und weniger zielgerichtetes Verhalten zeigten. Insgesamt weisen die Ergebnisse darauf hin, dass Aufgaben zur sequentiellen Entscheidungsfindung besonders geeignet sind, modellfreies- und modellbasiertes Verhalten abzubilden, während habituelle Verhaltensmuster weniger genau dargestellt werden können.

Welchen Einfluss Intelligenz und chronisches Stresserleben auf die oben dargestellten Lernprozesse und ihre neuronale Repräsentation nehmen, wurde in Studie 2 dargestellt. Hier konnte gezeigt werden, dass ein grundlegendes neurobiologisches Korrelat von Lernen – das Prädiktionsfehlersignal im VS – und das Ausmaß von stressassoziierten Lebensereignissen in den letzten beiden Jahren des Probanden zusammen Einfluss auf die fluide Komponente der Intelligenz nahmen, auch wenn für deren Abnahme durch Alter korrigiert wurde. Es konnte nachgewiesen werden, dass der Zusammenhang zwischen Stress und dem Prädiktionsfehlersignal im VS (als Korrelat von modellfreiem Lernen) in die gleiche positive Richtung zeigt wie der Zusammenhang zwischen dem Prädiktionsfehlersignal und fluider Intelligenz. Dieser Zusammenhang zwischen einem modellfreien Lernsignal und einer kognitiven Fähigkeit, die vor allem mit logischem, vorausschauendem, also eher modell-basiertem Denken assoziiert ist, scheint auf den ersten Blick kontraintuitiv. Eine mögliche Interpretation ist das Vorliegen paralleler, gleichzeitig ablaufender Entscheidungsstrategien, welche in einem ständigen Austausch und in Konkurrenz zueinander stehen (Daw, Gershman et al. 2011). So mag es von Vorteil sein, kognitive Strategien in Richtung eines eher modellfreien Verhaltens auszurichten, wenn in der Vergangenheit viel Stress erlebt wurde und modellfreies Entscheiden unter Stress häufiger belohnt wurde. Diese Tendenz in der Entscheidungsfindung hin zu mehr modellfreiem Verhalten nach hoher Stressexposition, findet vermutlich vor allem bei Probanden mit einem hohen fluiden IQ statt, da hier ausreichend Kapazität für die parallele Bewertung und das permanente Abgleichen der konkurrierenden Systeme vorhanden ist. Die Zunahme des Prädiktionsfehlersignal nach (akutem) Stresserleben wurde von Robinson und Kollegen (2013) für eine ähnliche *reversal* Lernaufgabe an gesunden Probanden beschrieben. Im Gegensatz hierzu beschreiben Otto und Kollegen (2013) auf Verhaltensebene eine Tendenz von modellbasiertem hin zu modellfreiem Lernen unter Stress nur bei Probanden mit einer niedrigen Arbeitsgedächtniskapazität (eine wesentliche Komponente der fluiden Intelligenz). Eine erhöhte kognitive Kapazität scheint die Probanden hier vor der Abnahme der modellbasierten Komponente unter (akutem) Stress zu schützen, während das modellfreie System durch keine der Komponenten beeinflusst wurde. Die

abweichenden Ergebnisse weisen auf einen differentiellen Effekt von akutem und chronischem Stresserleben auf Lernprozesse sowohl auf neuronaler- als auch auf Verhaltensebene hin. Die Aktivierung im VS als neurobiologisches Korrelat von Lernen wurde durch Stresserfahrung und deren Interaktion mit kognitiven Strukturen maßgeblich beeinflusst: es zeigte sich nur dann ein Zusammenhang zwischen der fluiden Intelligenzkomponente und dem modellfreien Prädiktionsfehlersignal im VS, wenn die Probanden viel Stress in der Vergangenheit berichteten. Diese Aktivierungsunterschiede wurden mit Veränderungen an der dopaminergen Synapse in Verbindung gebracht (Schlagenhauf, Rapp et al. 2013), welche unter anderem durch Stresserleben evoziert werden können (Morgan, Grant et al. 2002, Cao, Covington et al. 2010, Valenti, Lodge et al. 2011). Das dopaminerge System im Bereich des VS scheint tatsächlich ein wichtiger Bestandteil für das Verständnis der Verbindung zwischen chronischem Stresserleben, der Intelligenz und dem Entscheidungsverhalten zu sein.

Für zahlreiche, insbesondere psychiatrische Erkrankungen, wurde eine dysfunktionale Belohnungsverarbeitung im VS bei den beschriebenen Lernaufgaben (Schlagenhauf, Juckel et al. 2008, Heinz, Beck et al. 2009, Beck, Wustenberg et al. 2012), gekoppelt an Veränderungen in der dopaminergen Neurotransmission (Deserno, Boehme et al. 2013) und Veränderungen in der Entscheidungsfindung (Deserno, Beck et al. 2015) berichtet. Ein Zuwachs im habituellen Lernen und eine Abnahme der modellbasierten Entscheidungsfindung wurden beispielsweise bei Zwangserkrankungen (Gillan, Morein-Zamir et al. 2014, Voon, Derbyshire et al. 2014, Gillan, Apergis-Schoute et al. 2015), und insbesondere bei Suchterkrankungen (Dickinson, Wood et al. 2002, Sjoerds, de Wit et al. 2013, Sebold, Deserno et al. 2014) beschrieben. Patienten mit Alkoholabhängigkeit zeigen vermehrt automatisierte, habituelle Verhaltensmuster, wie z.B. das Heranziehen suchtassoziierter Reize (Wiers, Stelzel et al. 2014). Eine Überwindung dieser automatisierten Tendenz durch Verhaltenstraining zeigte in einer ersten Studie von Eberl und Kollegen (2013) eine signifikante Reduktion der Rückfallrate. Auf neuronaler Ebene scheint chronischer Alkoholkonsum zu einer spezifischen Störung der dopaminergen Modulation des (vorwiegend modellfreien) Prädiktionsfehler-Lernsignals zu führen: Deserno und Kollegen (2015) fanden einen inversen Zusammenhang zwischen der Dopaminsynthesekapazität im VS (gemessen über PET), dem Prädiktionsfehlersignal im VS während der *reversal*-Lernaufgabe und dem Verlangen nach Alkohol bei abhängigen, aktuell abstinenten Patienten. Eine Moderatoranalyse zeigte, dass die Höhe dieses (negativen) Zusammenhangs (zwischen Dopaminsynthesekapazität und der neuronalen Enkodierung des Prädiktionsfehlersignals im VS) von der Höhe des chronischen Alkoholkonsums bestimmt wurde - während sich aufgabenunabhängig kein Unterschied in der Dopaminsynthesekapazität zwischen den Gruppen zeigte. Interessanterweise konnte zusätzlich ein enger Zusammenhang zwischen der dopaminergen Modulation des

Prädiktionsfehlersignals, chronischem Stresserleben im Jugendalter und der Entwicklung eines gefährlichen Alkoholkonsums im Erwachsenenalter (Casement, Shaw et al. 2014) nachgewiesen werden: Eine erhöhte Stressexposition im Jugendalter war mit einer Reduktion des neuronalen Lernsignals während der Belohnungserwartung im Erwachsenenalter assoziiert. Die Höhe der (fehlenden) Aktivierung im medialen Präfrontalkortex während der Gewinnerwartung konnte gefährlichen Alkoholkonsum im Erwachsenenalter vorhersagen.

Zusammenfassend kann festgestellt werden, dass ein hinreichend komplexes ätiologisches Konzept für die Entstehung der Alkoholabhängigkeit nur dann gelingen kann, wenn alle genannten Faktoren, insbesondere maladaptive Lernprozesse, genetische Prädispositionen, Stresserfahrungen und die individuelle kognitive Kapazität berücksichtigt und in ihrer Interaktion erfasst werden können.

Eine wesentliche Einschränkung der dargestellten Ergebnisse ergibt sich aus ihrer meist korrelativen Berechnung, wodurch kausale Interpretationen spekulativer Natur bleiben. Stress und Intelligenz scheinen jedoch einen wichtigen differentiellen Einfluss auf das Entscheidungsverhalten und dessen dopaminerge Modulation zu haben. Die zukünftige spezifische Erfassung und Operationalisierung von modellbasiertem vs. modellfreiem, sowie habituellem vs. zielgerichtetem Lernen an klinischen Populationen kann wichtige weitere Hinweise für die Entstehung insbesondere der Suchterkrankung liefern und ggf. sogar helfen, wirksamere Interventionen zu entwickeln. Zu diesem Zweck sollten Studien folgen, um die dargestellten Operationalisierungen weiter zu optimieren und an verschiedenen klinischen Populationen zu testen.

## Literaturverzeichnis

- Abler, B., I. Greenhouse, D. Ongur, H. Walter and S. Heckers (2008). "Abnormal reward system activation in mania." Neuropsychopharmacology **33**(9): 2217-2227.
- Balleine, B. W. (2005). "Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits." Physiol Behav **86**(5): 717-730.
- Baumeister, R. F., J. M. Twenge and C. K. Nuss (2002). "Effects of social exclusion on cognitive processes: anticipated aloneness reduces intelligent thought." J Pers Soc Psychol **83**(4): 817-827.
- Beck, A., F. Schlagenhauf, T. Wustenberg, J. Hein, T. Kienast, T. Kahnt, K. Schmack, C. Hagele, B. Knutson, A. Heinz and J. Wrase (2009). "Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics." Biol Psychiatry **66**(8): 734-742.
- Beck, A., T. Wustenberg, A. Genauk, J. Wrase, F. Schlagenhauf, M. N. Smolka, K. Mann and A. Heinz (2012). "Effect of brain structure, brain function, and brain connectivity on relapse in alcohol-dependent patients." Arch Gen Psychiatry **69**(8): 842-852.
- Bellman, R. (1957). Dynamic Programming, Princeton University Press.
- Berridge, K. C. (2009). "Wanting and Liking: Observations from the Neuroscience and Psychology Laboratory." Inquiry (Oslo) **52**(4): 378.
- Berridge, K. C. and T. E. Robinson (1998). "What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience?" Brain Res Brain Res Rev **28**(3): 309-369.
- Berridge, K. C., T. E. Robinson and J. W. Aldridge (2009). "Dissecting components of reward: 'liking', 'wanting', and learning." Curr Opin Pharmacol **9**(1): 65-73.
- Bettinger, J. C. and A. G. Davies (2014). "The role of the BK channel in ethanol response behaviors: evidence from model organism and human studies." Front Physiol **5**: 346.
- Buckert, M., C. Schwieren, B. M. Kudielka and C. J. Fiebach (2014). "Acute stress affects risk taking but not ambiguity aversion." Front Neurosci **8**: 82.
- Cao, J. L., H. E. Covington, 3rd, A. K. Friedman, M. B. Wilkinson, J. J. Walsh, D. C. Cooper, E. J. Nestler and M. H. Han (2010). "Mesolimbic dopamine neurons in the brain reward circuit mediate susceptibility to social defeat and antidepressant action." J Neurosci **30**(49): 16453-16458.
- Casement, M. D., D. S. Shaw, S. L. Sitnick, S. C. Musselman and E. E. Forbes (2014). "Life stress in adolescence predicts early adult reward-related brain function and alcohol dependence." Soc Cogn Affect Neurosci.
- Casement, M. D., D. S. Shaw, S. L. Sitnick, S. C. Musselman and E. E. Forbes (2015). "Life stress in adolescence predicts early adult reward-related brain function and alcohol dependence." Soc Cogn Affect Neurosci **10**(3): 416-423.
- Cavanagh, J. F., M. J. Frank and J. J. Allen (2011). "Social stress reactivity alters reward and punishment learning." Soc Cogn Affect Neurosci **6**(3): 311-320.
- Chartier, K. G., M. N. Hesselbrock and V. M. Hesselbrock (2010). "Development and vulnerability factors in adolescent alcohol use." Child Adolesc Psychiatr Clin N Am **19**(3): 493-504.
- Cools, R., L. Clark, A. M. Owen and T. W. Robbins (2002). "Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging." J Neurosci **22**(11): 4563-4567.
- Cools, R., S. E. Gibbs, A. Miyakawa, W. Jagust and M. D'Esposito (2008). "Working memory capacity predicts dopamine synthesis capacity in the human striatum." J Neurosci **28**(5): 1208-1212.
- Cooper, J. C. and B. Knutson (2008). "Valence and salience contribute to nucleus accumbens activation." Neuroimage **39**(1): 538-547.
- Cruz, F. C., I. M. Quadros, S. Planeta Cda and K. A. Miczek (2008). "Maternal separation stress in male mice: long-term increases in alcohol intake." Psychopharmacology (Berl) **201**(3): 459-468.
- Davies, A. G., J. T. Pierce-Shimomura, H. Kim, M. K. VanHoven, T. R. Thiele, A. Bonci, C. I. Bargmann and S. L. McIntire (2003). "A central role of the BK potassium channel in behavioral responses to ethanol in *C. elegans*." Cell **115**(6): 655-666.
- Daw, N. D., S. J. Gershman, B. Seymour, P. Dayan and R. J. Dolan (2011). "Model-based influences on humans' choices and striatal prediction errors." Neuron **69**(6): 1204-1215.

Daw, N. D., Y. Niv and P. Dayan (2005). "Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control." Nat Neurosci **8**(12): 1704-1711.

de Wit, H., A. H. Soderpalm, L. Nikolayev and E. Young (2003). "Effects of acute social stress on alcohol consumption in healthy subjects." Alcohol Clin Exp Res **27**(8): 1270-1277.

de Wit, S., R. A. Barker, A. D. Dickinson and R. Cools (2011). "Habitual versus goal-directed action control in Parkinson disease." J Cogn Neurosci **23**(5): 1218-1229.

de Wit, S., P. R. Corlett, M. R. Aitken, A. Dickinson and P. C. Fletcher (2009). "Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans." J Neurosci **29**(36): 11330-11338.

de Wit, S., P. Watson, H. A. Harsay, M. X. Cohen, I. van de Vijver and K. R. Ridderinkhof (2012). "Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control." J Neurosci **32**(35): 12066-12075.

Deserno, L., A. Beck, Q. J. Huys, R. C. Lorenz, R. Buchert, H. G. Buchholz, M. Plotkin, Y. Kumakara, P. Cumming, H. J. Heinze, A. A. Grace, M. A. Rapp, F. Schlagenhauf and A. Heinz (2015). "Chronic alcohol intake abolishes the relationship between dopamine synthesis capacity and learning signals in the ventral striatum." Eur J Neurosci **41**(4): 477-486.

Deserno, L., R. Boehme, A. Heinz and F. Schlagenhauf (2013). "Reinforcement learning and dopamine in schizophrenia: dimensions of symptoms or specific features of a disease group?" Front Psychiatry **4**: 172.

Dezfouli, A. and B. W. Balleine (2013). "Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized." PLoS Comput Biol **9**(12): e1003364.

Di Chiara, G. and A. Imperato (1988). "Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats." Proc Natl Acad Sci U S A **85**(14): 5274-5278.

Dias-Ferreira, E., J. C. Sousa, I. Melo, P. Morgado, A. R. Mesquita, J. J. Cerqueira, R. M. Costa and N. Sousa (2009). "Chronic stress causes frontostriatal reorganization and affects decision-making." Science **325**(5940): 621-625.

Dick, D. M., J. Plunkett, L. F. Wetherill, X. Xuei, A. Goate, V. Hesselbrock, M. Schuckit, R. Crowe, H. J. Edenberg and T. Foroud (2006). "Association between GABRA1 and drinking behaviors in the collaborative study on the genetics of alcoholism sample." Alcohol Clin Exp Res **30**(7): 1101-1110.

Dickinson, A., N. Wood and J. W. Smith (2002). "Alcohol seeking by rats: action or habit?" Q J Exp Psychol B **55**(4): 331-348.

Dolan, R. J. and P. Dayan (2013). "Goals and habits in the brain." Neuron **80**(2): 312-325.

Doll, B. B., D. A. Simon and N. D. Daw (2012). "The ubiquity of model-based reinforcement learning." Curr Opin Neurobiol **22**(6): 1075-1081.

Eberl, C., R. W. Wiers, S. Pawelczack, M. Rinck, E. S. Becker and J. Lindenmeyer (2013). "Approach bias modification in alcohol dependence: do clinical effects replicate and for whom does it work best?" Dev Cogn Neurosci **4**: 38-51.

Erritzoe, D., A. Tziortzi, D. Bargiela, A. Colasanti, G. E. Searle, R. N. Gunn, J. D. Beaver, A. Waldman, D. J. Nutt, M. Bani, E. Merlo-Pich, E. A. Rabiner and A. Lingford-Hughes (2014). "In vivo imaging of cerebral dopamine D3 receptors in alcoholism." Neuropsychopharmacology **39**(7): 1703-1712.

Flagel, S. B., H. Akil and T. E. Robinson (2009). "Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction." Neuropharmacology **56 Suppl 1**: 139-148.

Flagel, S. B., C. M. Cameron, K. N. Pickup, S. J. Watson, H. Akil and T. E. Robinson (2011). "A food predictive cue must be attributed with incentive salience for it to induce c-fos mRNA expression in cortico-striatal-thalamic brain regions." Neuroscience **196**: 80-96.

Flagel, S. B., J. J. Clark, T. E. Robinson, L. Mayo, A. Czuj, I. Willuhn, C. A. Akers, S. M. Clinton, P. E. Phillips and H. Akil (2011). "A selective role for dopamine in stimulus-reward learning." Nature **469**(7328): 53-57.

Flagel, S. B., T. E. Robinson, J. J. Clark, S. M. Clinton, S. J. Watson, P. Seeman, P. E. Phillips and H. Akil (2010). "An animal model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-related cues: implications for addiction." Neuropsychopharmacology **35**(2): 388-400.



Flagel, S. B., S. J. Watson, H. Akil and T. E. Robinson (2008). "Individual differences in the attribution of incentive salience to a reward-related cue: influence on cocaine sensitization." Behav Brain Res **186**(1): 48-56.

Frank, M. J., L. C. Seeberger and C. O'Reilly R (2004). "By carrot or by stick: cognitive reinforcement learning in parkinsonism." Science **306**(5703): 1940-1943.

Friedel, E., S. P. Koch, J. Wendt, A. Heinz, L. Deserno and F. Schlagenhauf (2014). "Devaluation and sequential decisions: linking goal-directed and model-based behavior." Front Hum Neurosci **8**: 587.

Friedel, E., F. Schlagenhauf, A. Beck, R. J. Dolan, Q. J. Huys, M. A. Rapp and A. Heinz (2015). "The effects of life stress and neural learning signals on fluid intelligence." Eur Arch Psychiatry Clin Neurosci **265**(1): 35-43.

Garbusow, M., D. J. Schad, M. Sebold, E. Friedel, N. Bernhardt, S. P. Koch, B. Steinacher, N. Kathmann, D. E. Geurts, C. Sommer, D. K. Muller, S. Nebe, S. Paul, H. U. Wittchen, U. S. Zimmermann, H. Walter, M. N. Smolka, P. Sterzer, M. A. Rapp, Q. J. Huys, F. Schlagenhauf and A. Heinz (2015). "Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence." Addict Biol.

Garbusow, M., D. J. Schad, C. Sommer, E. Junger, M. Sebold, E. Friedel, J. Wendt, N. Kathmann, F. Schlagenhauf, U. S. Zimmermann, A. Heinz, Q. J. Huys and M. A. Rapp (2014). "Pavlovian-to-instrumental transfer in alcohol dependence: a pilot study." Neuropsychobiology **70**(2): 111-121.

Garbusow, M., M. Sebold, A. Beck and A. Heinz (2014). "Too difficult to stop: mechanisms facilitating relapse in alcohol dependence." Neuropsychobiology **70**(2): 103-110.

Garrett, D. D., C. L. Grady and L. Hasher (2010). "Everyday memory compensation: the impact of cognitive reserve, subjective memory, and stress." Psychol Aging **25**(1): 74-83.

Gelernter, J., H. R. Kranzler, R. Sherva, L. Almasy, R. Koesterer, A. H. Smith, R. Anton, U. W. Preuss, M. Ridinger, D. Rujescu, N. Wodarz, P. Zill, H. Zhao and L. A. Farrer (2014). "Genome-wide association study of alcohol dependence: significant findings in African- and European-Americans including novel risk loci." Mol Psychiatry **19**(1): 41-49.

Gillan, C. M., A. M. Apergis-Schoute, S. Morein-Zamir, G. P. Urcelay, A. Sule, N. A. Fineberg, B. J. Sahakian and T. W. Robbins (2015). "Functional neuroimaging of avoidance habits in obsessive-compulsive disorder." Am J Psychiatry **172**(3): 284-293.

Gillan, C. M., S. Morein-Zamir, G. P. Urcelay, A. Sule, V. Voon, A. M. Apergis-Schoute, N. A. Fineberg, B. J. Sahakian and T. W. Robbins (2014). "Enhanced avoidance habits in obsessive-compulsive disorder." Biol Psychiatry **75**(8): 631-638.

Gillan, C. M., A. R. Otto, E. A. Phelps and N. D. Daw (2015). "Model-based learning protects against forming habits." Cogn Affect Behav Neurosci.

Gillan, C. M., M. Pappmeyer, S. Morein-Zamir, B. J. Sahakian, N. A. Fineberg, T. W. Robbins and S. de Wit (2011). "Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder." Am J Psychiatry **168**(7): 718-726.

Goldman-Rakic, P. (2000). "Localization of function all over again." Neuroimage **11**(5 Pt 1): 451-457.

Hagele, C., E. Friedel, T. Kienast and F. Kiefer (2014). "How do we 'learn' addiction? Risk factors and mechanisms getting addicted to alcohol." Neuropsychobiology **70**(2): 67-76.

Hampton, A. N., P. Bossaerts and J. P. O'Doherty (2006). "The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans." J Neurosci **26**(32): 8360-8367.

Harlaar, N., A. D. Bryan, R. E. Thayer, H. C. Karoly, N. Oien and K. E. Hutchison (2014). "Methylation of a CpG site near the ALDH1A2 gene is associated with loss of control over drinking and related phenotypes." Alcohol Clin Exp Res **38**(3): 713-721.

Harlaar, N. and K. E. Hutchison (2013). "Alcohol and the methylome: design and analysis considerations for research using human samples." Drug Alcohol Depend **133**(2): 305-316.

Heinz, A., A. Batra, N. Scherbaum and E. Gouzoulis-Mayfrank (2012). Neurobiologie der Abhängigkeit, Kohlhammer.

Heinz, A., A. Beck, S. M. Grusser, A. A. Grace and J. Wrase (2009). "Identifying the neural circuitry of alcohol craving and relapse vulnerability." Addict Biol **14**(1): 108-118.

Heinz, A., A. Beck, J. Wrase, J. Mohr, K. Obermayer, J. Gallinat and I. Puls (2009). "Neurotransmitter systems in alcohol dependence." Pharmacopsychiatry **42 Suppl 1**: S95-S101.

Heinz, A., M. Schafer, J. D. Higley, J. H. Krystal and D. Goldman (2003). "Neurobiological correlates of the disposition and maintenance of alcoholism." Pharmacopsychiatry **36 Suppl 3**: S255-258.

Heinz, A., J. Wrase, T. Kahnt, A. Beck, Z. Bromand, S. M. Grusser, T. Kienast, M. N. Smolka, H. Flor and K. Mann (2007). "Brain activation elicited by affectively positive stimuli is associated with a lower risk of relapse in detoxified alcoholic subjects." Alcohol Clin Exp Res **31(7)**: 1138-1147.

Holmes, T. H. R., R.H. (1967). "The Social Readjustment Rating Scale." J Psychosom Res **11 (2)**: 213-218.

Horn, J. L. and R. B. Cattell (1966). "Age differences in primary mental ability factors." J Gerontol **21(2)**: 210-220.

Hu, X., G. Oroszi, J. Chun, T. L. Smith, D. Goldman and M. A. Schuckit (2005). "An expanded evaluation of the relationship of four alleles to the level of response to alcohol and the alcoholism risk." Alcohol Clin Exp Res **29(1)**: 8-16.

Huot, R. L., K. V. Thirvikraman, M. J. Meaney and P. M. Plotsky (2001). "Development of adult ethanol preference and anxiety as a consequence of neonatal maternal separation in Long Evans rats and reversal with antidepressant treatment." Psychopharmacology (Berl) **158(4)**: 366-373.

Huys, Q. J., R. Cools, M. Golzer, E. Friedel, A. Heinz, R. J. Dolan and P. Dayan (2011). "Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding." PLoS Comput Biol **7(4)**: e1002028.

Huys, Q. J., N. Eshel, E. O'Nions, L. Sheridan, P. Dayan and J. P. Roiser (2012). "Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees." PLoS Comput Biol **8(3)**: e1002410.

Huys, Q. J., P. N. Tobler, G. Hasler and S. B. Flagel (2014). "The role of learning-related dopamine signals in addiction vulnerability." Prog Brain Res **211**: 31-77.

Kendler, K. S., J. Ji, A. C. Edwards, H. Ohlsson, J. Sundquist and K. Sundquist (2015). "An Extended Swedish National Adoption Study of Alcohol Use Disorder." JAMA Psychiatry.

Kienast, T., A. R. Hariri, F. Schlagenhauf, J. Wrase, P. Sterzer, H. G. Buchholz, M. N. Smolka, G. Grunder, P. Cumming, Y. Kumakura, P. Bartenstein, R. J. Dolan and A. Heinz (2008). "Dopamine in amygdala gates limbic processing of aversive stimuli in humans." Nat Neurosci **11(12)**: 1381-1382.

Kienast, T., F. Schlagenhauf, M. A. Rapp, J. Wrase, I. Daig, H. G. Buchholz, M. N. Smolka, G. Grunder, Y. Kumakura, P. Cumming, K. Charlet, P. Bartenstein, A. R. Hariri and A. Heinz (2013). "Dopamine-modulated aversive emotion processing fails in alcohol-dependent patients." Pharmacopsychiatry **46(4)**: 130-136.

Kienast, T., T. Siessmeier, J. Wrase, D. F. Braus, M. N. Smolka, H. G. Buchholz, M. Rapp, M. Schreckenberger, F. Rosch, P. Cumming, G. Gruender, K. Mann, P. Bartenstein and A. Heinz (2008). "Ratio of dopamine synthesis capacity to D2 receptor availability in ventral striatum correlates with central processing of affective stimuli." Eur J Nucl Med Mol Imaging **35(6)**: 1147-1158.

Klein, K. and W. H. Fiss (1999). "The reliability and stability of the Turner and Engle working memory task." Behav Res Methods Instrum Comput **31(3)**: 429-432.

Koob, G. F. and F. E. Bloom (1988). "Cellular and molecular mechanisms of drug dependence." Science **242(4879)**: 715-723.

Kumakura, Y., A. Gjedde, D. Caprioli, T. Kienast, A. Beck, M. Plotkin, F. Schlagenhauf, I. Vernaleken, G. Grunder, P. Bartenstein, A. Heinz and P. Cumming (2013). "Increased turnover of dopamine in caudate nucleus of detoxified alcoholic patients." PLoS One **8(9)**: e73903.

Kumar, A., K. H. Choi, W. Renthal, N. M. Tsankova, D. E. Theobald, H. T. Truong, S. J. Russo, Q. Laplant, T. S. Sasaki, K. N. Whistler, R. L. Neve, D. W. Self and E. J. Nestler (2005). "Chromatin remodeling is a key mechanism underlying cocaine-induced plasticity in striatum." Neuron **48(2)**: 303-314.

Lau, B. and P. W. Glimcher (2005). "Dynamic response-by-response models of matching behavior in rhesus monkeys." J Exp Anal Behav **84**(3): 555-579.

Lee, S. W., S. Shimojo and J. P. O'Doherty (2014). "Neural computations underlying arbitration between model-based and model-free learning." Neuron **81**(3): 687-699.

Li, J. and N. D. Daw (2011). "Signals in human striatum are appropriate for policy update rather than value prediction." J Neurosci **31**(14): 5504-5511.

Lindenberger, U. and P. B. Baltes (1997). "Intellectual functioning in old and very old age: cross-sectional results from the Berlin Aging Study." Psychol Aging **12**(3): 410-432.

Lupien, S. J., B. S. McEwen, M. R. Gunnar and C. Heim (2009). "Effects of stress throughout the lifespan on the brain, behaviour and cognition." Nat Rev Neurosci **10**(6): 434-445.

Maezawa, Y., M. Yamauchi, G. Toda, H. Suzuki and S. Sakurai (1995). "Alcohol-metabolizing enzyme polymorphisms and alcoholism in Japan." Alcohol Clin Exp Res **19**(4): 951-954.

Morgan, D., K. A. Grant, H. D. Gage, R. H. Mach, J. R. Kaplan, O. Prioleau, S. H. Nader, N. Buchheimer, R. L. Ehrenkaufer and M. A. Nader (2002). "Social dominance in monkeys: dopamine D2 receptors and cocaine self-administration." Nat Neurosci **5**(2): 169-174.

Nestler, E. J. and G. K. Aghajanian (1997). "Molecular and cellular basis of addiction." Science **278**(5335): 58-63.

Nieratschker, V., Batra, A., Fallgatter, A.J. (2013). "Genetics and epigenetics of alcohol dependence." Journal of Molecular Psychiatry 1:11.

O'Doherty, J. P., A. Hampton and H. Kim (2007). "Model-based fMRI and its application to reward learning and decision making." Ann N Y Acad Sci **1104**: 35-53.

Otto, A. R., C. M. Raio, A. Chiang, E. A. Phelps and N. D. Daw (2013). "Working-memory capacity protects model-based learning from stress." Proc Natl Acad Sci U S A **110**(52): 20941-20946.

Park, S. Q., T. Kahnt, A. Beck, M. X. Cohen, R. J. Dolan, J. Wrase and A. Heinz (2010). "Prefrontal cortex fails to learn from reward prediction errors in alcohol dependence." J Neurosci **30**(22): 7749-7753.

Pawlow, I. P. (1927). Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. Oxford, England, Oxford Univ. Press.

Philibert, R. A., B. Penaluna, T. White, S. Shires, T. Gunter, J. Liesveld, C. Erwin, N. Hollenbeck and T. Osborn (2014). "A pilot examination of the genome-wide DNA methylation signatures of subjects entering and exiting short-term alcohol dependence treatment programs." Epigenetics **9**(9): 1212-1219.

Prevost, C., D. McNamee, R. K. Jessup, P. Bossaerts and J. P. O'Doherty (2013). "Evidence for model-based computations in the human amygdala during Pavlovian conditioning." PLoS Comput Biol **9**(2): e1002918.

Radenbach, C., A. M. Reiter, V. Engert, Z. Sjoerds, A. Villringer, H. J. Heinze, L. Deserno and F. Schlagenhauf (2015). "The interaction of acute and chronic stress impairs model-based behavioral control." Psychoneuroendocrinology **53**: 268-280.

Robbins, T. W. and B. J. Everitt (1999). "Drug addiction: bad habits add up." Nature **398**(6728): 567-570.

Robinson, O. J., C. Overstreet, D. R. Charney, K. Vytal and C. Grillon (2013). "Stress increases aversive prediction error signal in the ventral striatum." Proc Natl Acad Sci U S A **110**(10): 4129-4133.

Robinson, T. E. and K. C. Berridge (1993). "The neural basis of drug craving: An incentive-sensitization theory of addiction." Brain Research Reviews **18**(3): 247-291.

Robinson, T. E. and B. Kolb (2004). "Structural plasticity associated with exposure to drugs of abuse." Neuropharmacology **47 Suppl 1**: 33-46.

Schipolowski, S. W., O.; Schroeders, U. (2014). "On the nature of crystallized intelligence: the relationship between verbal ability and factual knowledge." Intelligence **46**: 156-168.

Schlagenhauf, F., G. Juckel, M. Koslowski, T. Kahnt, B. Knutson, T. Dembler, T. Kienast, J. Gallinat, J. Wrase and A. Heinz (2008). "Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine." Psychopharmacology (Berl) **196**(4): 673-684.

Schlagenhauf, F., M. A. Rapp, Q. J. Huys, A. Beck, T. Wustenberg, L. Deserno, H. G. Buchholz, J. Kalbitzer, R. Buchert, M. Bauer, T. Kienast, P. Cumming, M. Plotkin, Y.

Kumakura, A. A. Grace, R. J. Dolan and A. Heinz (2013). "Ventral striatal prediction error signaling is associated with dopamine synthesis capacity and fluid intelligence." Hum Brain Mapp **34**(6): 1490-1499.

Schultz, W., P. Dayan and P. R. Montague (1997). "A neural substrate of prediction and reward." Science **275**(5306): 1593-1599.

Schwabe, L., A. Bohringer and O. T. Wolf (2009). "Stress disrupts context-dependent memory." Learn Mem **16**(2): 110-113.

Schwabe, L., O. Hoffken, M. Tegenthoff and O. T. Wolf (2011). "Preventing the stress-induced shift from goal-directed to habit action with a beta-adrenergic antagonist." J Neurosci **31**(47): 17317-17325.

Schwabe, L., M. Tegenthoff, O. Hoffken and O. T. Wolf (2010). "Concurrent glucocorticoid and noradrenergic activity shifts instrumental behavior from goal-directed to habitual control." J Neurosci **30**(24): 8190-8196.

Schwabe, L. and O. T. Wolf (2009). "Stress prompts habit behavior in humans." J Neurosci **29**(22): 7191-7198.

Schwabe, L. and O. T. Wolf (2012). "Stress modulates the engagement of multiple memory systems in classification learning." J Neurosci **32**(32): 11042-11049.

Schwabe, L., O. T. Wolf and M. S. Oitzl (2010). "Memory formation under stress: quantity and quality." Neurosci Biobehav Rev **34**(4): 584-591.

Sebold, M., L. Deserno, S. Nebe, D. J. Schad, M. Garbusow, C. Hagele, J. Keller, E. Junger, N. Kathmann, M. Smolka, M. A. Rapp, F. Schlagenhauf, A. Heinz and Q. J. Huys (2014). "Model-based and model-free decisions in alcohol dependence." Neuropsychobiology **70**(2): 122-131.

Simon, J. J., M. Skunde, M. Wu, K. Schnell, S. C. Herpertz, M. Bendszus, W. Herzog and H. C. Friederich (2014). "Neural dissociation of food- and money-related reward processing using an abstract incentive delay task." Soc Cogn Affect Neurosci.

Sjoerds, Z., S. de Wit, W. van den Brink, T. W. Robbins, A. T. Beekman, B. W. Penninx and D. J. Veltman (2013). "Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients." Transl Psychiatry **3**: e337.

Smith, G. T., M. S. Goldman, P. E. Greenbaum and B. A. Christiansen (1995). "Expectancy for social facilitation from drinking: the divergent paths of high-expectancy and low-expectancy adolescents." J Abnorm Psychol **104**(1): 32-40.

Smittenaar, P., T. H. FitzGerald, V. Romei, N. D. Wright and R. J. Dolan (2013). "Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans." Neuron **80**(4): 914-919.

Soares, J. M., A. Sampaio, L. M. Ferreira, N. C. Santos, F. Marques, J. A. Palha, J. J. Cerqueira and N. Sousa (2012). "Stress-induced changes in human decision-making are reversible." Transl Psychiatry **2**: e131.

Sternberg, R. J. (2000). "Cognition. The holy grail of general intelligence." Science **289**(5478): 399-401.

Stuber, G. D., M. Klanker, B. de Ridder, M. S. Bowers, R. N. Joosten, M. G. Feenstra and A. Bonci (2008). "Reward-predictive cues enhance excitatory synaptic strength onto midbrain dopamine neurons." Science **321**(5896): 1690-1692.

Sutton, R. B., A. (1998). Reinforcement Learning: An Introduction, A Bradford Book.

Thorndike, E. (1911). Animal Intelligence: Experimental Studies. New York, MacMillan.

Tolman, E. C. (1948). "Cognitive maps in rats and men." Psychological Review **55**: 189-208.

Tricomi, E., B. W. Balleine and J. P. O'Doherty (2009). "A specific role for posterior dorsolateral striatum in human habit learning." Eur J Neurosci **29**(11): 2225-2232.

Valenti, O., D. J. Lodge and A. A. Grace (2011). "Aversive stimuli alter ventral tegmental area dopamine neuron activity via a common action in the ventral hippocampus." J Neurosci **31**(11): 4280-4289.

Valentin, V. V., A. Dickinson and J. P. O'Doherty (2007). "Determining the neural substrates of goal-directed learning in the human brain." J Neurosci **27**(15): 4019-4026.

Valentin, V. V. and J. P. O'Doherty (2009). "Overlapping prediction errors in dorsal striatum during instrumental learning with juice and money reward in the human brain." J Neurophysiol **102**(6): 3384-3391.

Volkow, N. D., G. J. Wang, F. Telang, J. S. Fowler, J. Logan, A. R. Childress, M. Jayne, Y. Ma and C. Wong (2006). "Cocaine cues and dopamine in dorsal striatum: mechanism of craving in cocaine addiction." J Neurosci **26**(24): 6583-6588.

von der Goltz, C. and F. Kiefer (2009). "Learning and memory in the aetiopathogenesis of addiction: future implications for therapy?" Eur Arch Psychiatry Clin Neurosci **259 Suppl 2**: S183-187.

Voon, V., K. Derbyshire, C. Ruck, M. A. Irvine, Y. Worbe, J. Enander, L. R. Schreiber, C. Gillan, N. A. Fineberg, B. J. Sahakian, T. W. Robbins, N. A. Harrison, J. Wood, N. D. Daw, P. Dayan, J. E. Grant and E. T. Bullmore (2014). "Disorders of compulsivity: a common bias towards learning habits." Mol Psychiatry.

Waters, G. S. and D. Caplan (2003). "The reliability and stability of verbal working memory measures." Behav Res Methods Instrum Comput **35**(4): 550-564.

Weaver, I. C., N. Cervoni, F. A. Champagne, A. C. D'Alessio, S. Sharma, J. R. Seckl, S. Dymov, M. Szyf and M. J. Meaney (2004). "Epigenetic programming by maternal behavior." Nat Neurosci **7**(8): 847-854.

Wierenga, C. E., A. Bischoff-Grethe, A. J. Melrose, Z. Irvine, L. Torres, U. F. Bailer, A. Simmons, J. L. Fudge, S. M. McClure, A. Ely and W. H. Kaye (2015). "Hunger does not motivate reward in women remitted from anorexia nervosa." Biol Psychiatry **77**(7): 642-652.

Wiers, C. E., C. Stelzel, S. Q. Park, C. K. Gawron, V. U. Ludwig, S. Gutwinski, A. Heinz, J. Lindenmeyer, R. W. Wiers, H. Walter and F. BERPohl (2014). "Neural correlates of alcohol-approach bias in alcohol addiction: the spirit is willing but the flesh is weak for spirits." Neuropsychopharmacology **39**(3): 688-697.

Wittchen, H.-U., Zaudig, M., Fydrich, T. (1997). Strukturiertes Klinisches Interview für DSM-IV. SKID-I und SKID-II. Achse I: Psychische Störungen / Achse II: Persönlichkeitsstörungen. Göttingen, Hogrefe.

Wrase, J., M. Reimold, I. Puls, T. Kienast and A. Heinz (2006). "Serotonergic dysfunction: brain imaging and behavioral correlates." Cogn Affect Behav Neurosci **6**(1): 53-61.

Wrase, J., F. Schlagenhauf, T. Kienast, T. Wustenberg, F. BERPohl, T. Kahnt, A. Beck, A. Strohle, G. Juckel, B. Knutson and A. Heinz (2007). "Dysfunction of reward processing correlates with alcohol craving in detoxified alcoholics." Neuroimage **35**(2): 787-794.

Wunderlich, K., P. Dayan and R. J. Dolan (2012). "Mapping value based planning and extensively trained choice in the human brain." Nat Neurosci **15**(5): 786-791.

Yin, H. H., B. J. Knowlton and B. W. Balleine (2005). "Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning." Eur J Neurosci **22**(2): 505-512.

Yin, H. H., S. B. Ostlund, B. J. Knowlton and B. W. Balleine (2005). "The role of the dorsomedial striatum in instrumental conditioning." Eur J Neurosci **22**(2): 513-523.

## **Selbstständigkeitserklärung**

Ich erkläre, dass ich die Arbeit selbstständig ohne Hilfe Dritter verfasst habe und bei der Abfassung alle Regelungen guter wissenschaftlicher Standards eingehalten habe.

A handwritten signature in black ink, appearing to be 'Pre'.

Berlin den 13.06.2016

## **Erklärung zum Promotionsverfahren**

Ich erkläre, dass ich an keiner anderen Hochschule ein Promotionsverfahren eröffnet habe, sowie dass die Dissertation in der gegenwärtigen Fassung keiner anderen Hochschule zur Begutachtung vorliegt.

A handwritten signature in black ink, appearing to be 'Pre'.

Berlin den 13.06.2016

## Eingeschlossene Publikationen

1. Friedel, E.\*, Koch, S. P.\*, Wendt, J., Heinz, A., Deserno, L., Schlagenhauf, F. (2014). Devaluation and sequential decisions: linking goal-directed and model-based behavior. *Front Hum Neurosci* 8:587. \*equal contribution.
2. Friedel, E., Schlagenhauf, F., Beck, A., Dolan, R. J., Huys, Q. J.\*, Rapp, M. A.\*, Heinz, A.\* (2015). The effects of life stress and neural learning signals on fluid intelligence. *Eur Arch Psychiatry Clin Neurosci* 265(1):35-43. \*equal contribution.
3. Hägele, C., Friedel, E., Kienast, T., Kiefer, F. (2014) How do we 'learn' addiction? Risk factors and mechanisms getting addicted to alcohol. *Neuropsychobiology*: 70(2):67-76.