**PsyBSc 11: Lernen & Gedächtnis 5**

Literatur: Gluck, Mercado & Myers (2010). Lernen & Gedächtnis. Chapter 5.

Topic: Instrumentelles Konditionieren

Inhalt

[1. Behaviorale Prozesse 2](#_Toc94004164)

[1.1 Die „Entdeckung“ der Instrumentellen Konditionierung 2](#_Toc94004165)

[1.2 Klassische vs. Instrumentelle Konditionierung 2](#_Toc94004166)

[1.3 Akquisition und Extinktion 3](#_Toc94004167)

[1.4 Komponenten (S-R-C) 3](#_Toc94004168)

[1.4.1 Stimulus 3](#_Toc94004169)

[1.4.2 Reaktionen 4](#_Toc94004170)

[1.4.3 Konsequenzen 4](#_Toc94004171)

[1.5 Aufbau der S-R-C Assoziation 5](#_Toc94004172)

[1.5.1 Der Einfluss von Timing 6](#_Toc94004173)

[1.5.2 Unterschiedliche Verstärkungspläne 6](#_Toc94004174)

[1.6 Wahlverhalten 8](#_Toc94004175)

[1.6.1 Das Matching-Law von Herrnstein 8](#_Toc94004176)

[1.6.2 Das Premack-Prinzip 8](#_Toc94004177)

[2. Gehirnsubstrate 9](#_Toc94004178)

[2.1 Dorsales Striatum und Reiz-Reaktions-Lernen 9](#_Toc94004179)

[2.2 Der Orbitofrontale Cortex und Lernen, Konsequenzen vorherzusagen 10](#_Toc94004180)

[2.3 Das dopaminerge System und Bewertung der Konsequenz 10](#_Toc94004181)

[2.3.1 Ventral tegmental area (VTA) 10](#_Toc94004182)

[2.3.2 Dopamin und Verstärkung 11](#_Toc94004183)

[3. Anwendungen 13](#_Toc94004184)

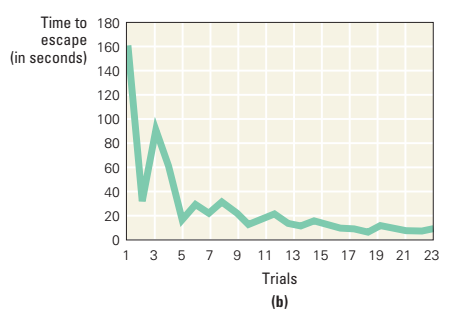
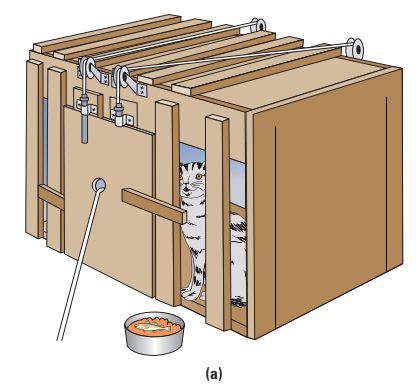
# 1. Behaviorale Prozesse

**Was ist instrumentelle Konditionierung?**

Der Prozess, bei dem ein Organismus lernt, bestimmte Verhaltensweisen zu tätigen oder zu vermeiden, um ein bestimmtes Ergebnis zu erzielen oder zu vermeiden. Wird seit Jahrtausenden angewandt in der Reinlichkeitserziehung von Kindern oder dem Dressieren von Tieren.

## Die „Entdeckung“ der Instrumentellen Konditionierung

**Edward Thorndike** war der erste, der die instrumentelle Konditionierung systematisch untersucht hat. Dazu nutzte er **Puzzleboxen**, in die er Katzen einsperrte.



Die Katze kann sich befreien, wenn sie eine korrekte Abfolge von Bewegungen ausführt. Wenn man die Katze erneut in die Box steckt, kann sie sich immer schneller befreien.

**Gesetz des Effekts**

= wenn auf die Reaktion des Tiers ein befriedigendes Ereignis folgt (z.B. aus Futterbox kommen oder Essen kriegen), so erhöht sich die Wahrscheinlichkeit, dass dieses Verhalten in der Zukunft erneut gezeigt wird. Ist die Konsequenz neutral oder negativ, nimmt die Wahrscheinlichkeit ab.

Discriminative Stimulus SD 🡪 Response R 🡪 Outcome O

SD Ein Reiz, der signalisiert, ob ein bestimmtes Verhalten zu einem bestimmten Ergebnis führt. Hilft dem Organismus, die Bedingungen zu erkennen, unter denen R zu O führt.

R Wenn das Outcome positiv ist, erhöht sich Wahrscheinlichkeit, dass R gezeigt wird. Dadurch verstärkt sich die SD 🡪 R – Verknüpfung

O Konsequenz aus dem Verhalten (entweder positiv oder negativ)

**Verstärkung**

= Erhöhung der Wahrscheinlichkeit eines Verhaltens durch Festlegen einer bestimmten Konsequenz für eine Reaktion

## Klassische vs. Instrumentelle Konditionierung

Was ist der Unterschied zwischen klassischer und instrumenteller Konditionierung?

a) Falls der Outcome unabhängig vom Verhalten ist 🡪 klassische Konditionierung

Lidschlusskonditionierung: Nach dem Ton (CS) folgt ein Luftstoß (US), sodass das Kaninchen lernt, einen Lidschluss (CR) beim Erklingen des Tons durchzuführen. Der Luftstoß wird unabhängig davon gegeben, ob CR erfolgt oder nicht.

b) Falls der Outcome abhängig vom Verhalten ist 🡪 instrumentelle Konditionierung

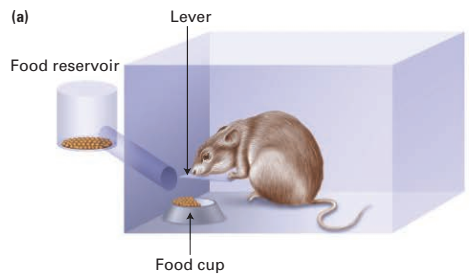
Bei der Puzzlebox (SD) muss die Katze eine Verhaltenssequenz durchführen (R), um zu entkommen und Futter zu kriegen (O). Wenn R nicht durchgeführt wird, dann findet O nicht statt.

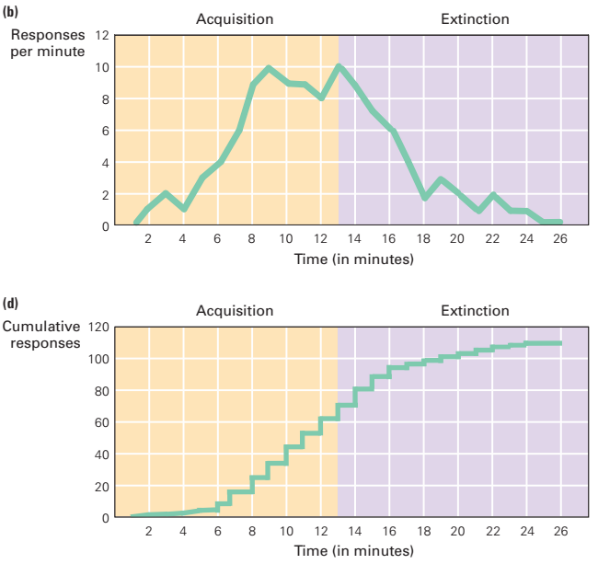
**Gemeinsamkeiten**: Verlauf der Lernkurve, Extinktion

Unterscheidung instrumentell (Reaktion, die man zeigt, um Konsequenz zu erreichen) vs. operant (Unterart von instrumentell, von Skinner geprägt mit sog. Skinner-Box: anstatt von diskreten Trials mit klarem Start/Ende, Versuche, wo Tiere selbst Trials starten)

## Akquisition und Extinktion

**Skinner Box**

* In der Skinner Box wird das Betätigen eines Hebels verstärkt durch die Gabe von Futter in einen Futterbecher
* Verstärkung findet nur statt, wenn ein Licht im Käfig angeschaltet ist



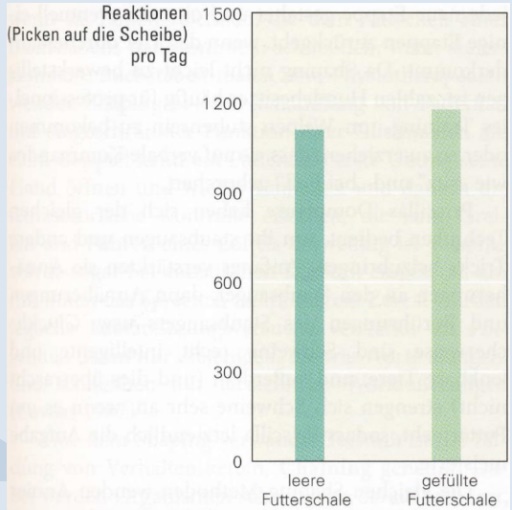
* Das Tier lernt mit der Zeit, dass O auf R folgt, wenn das Licht (SD) an ist: Verknüpfung von SD und R wird verstärkt
* Wenn das Licht ausgeschaltet ist, lernt das Tier, dass die Verknüpfung von R und O nicht stimmt. Und die Wahrscheinlichkeit, dass R gezeigt wird, nimmt ab.

## Komponenten (S-R-C)

### Stimulus

Es muss ermittelt werden, ob ein differenzierender Reiz vorliegt, der anzeigt, ob ein bestimmtes Verhalten nun belohnt/bestraft werden.

Stolz & Lott (1964): Wenn ein solcher Stimulus ermittelt wurde, zeigen Tiere auch dann noch die Reaktion, auch wenn es eigentlich gar nicht nötig wäre, weil auf dem Weg bspw. Futter ist. („Gewohnheitsfehler“)

**Protestantischer Ethikeffekt**

Hier wurde Tauben beigebracht, dass sie Futter erhalten, wenn sie auf eine leere Futterscheibe picken. Das machen sie auch dann weiterhin, wenn man eine gefüllte Futterschale neben sie stellt, weil die SD-R-Verknüpfung so stark automatisiert ist.

### 1.4.2 Reaktionen

Wichtig ist: es geht nicht um motorische Sequenz, sondern Erreichen eines Ziels. Es ist nicht wichtig, wo die Tochter anfängt, Zimmer aufzuräumen; Hauptsache, sie tut es!

**Shaping**

= Technik der instrumentellen Konditionierung, bei der sukzessive Annäherungen an die gewünschte Reaktion verstärkt werden. Beispiel:

Ratte erhält Futter in bestimmten Abteil des Labyrinths, wenn sie das gelernt hat, erhält sie nur Futter, wenn sie sich in der Nähe des Hebels befindet. Dann nur noch Futter, wenn Hebel berührt wird… bis die Ratte lernt, dass Hebel betätigt werden muss, um Futter zu erhalten.

**Chaining**

Technik der IK, bei der graduell beigebracht wird, komplexe Verhaltenssequenzen durchzuführen, die sich aus mehreren diskreten Schritten zusammensetzen. Beispiel:

Skinner trainierte eine Ratte: 1) an Schnur ziehen, was eine Murmel bringt; 2) Murmel greifen; 3) Murmel zu einem Rohr bringen; 4) Murmel in das Rohr fallen lassen

Die Sequenz wurde nicht auf einmal gelernt, sondern die Schritte wurden nacheinander zusammengefügt.

Beispiele: Instrumente lernen, Pilotentraining, etc.

### 1.4.3 Konsequenzen

= Verstärker (reinforcers) sind Konsequenzen eines Verhaltens, die dazu führen, dass sich die Wahrscheinlichkeit des Verhaltens in der Zukunft erhöht.

**Primäre Verstärker**

= Reize (z.B. Essen, Wasser, Sex, Schlaf), die einen inhärenten biologischen Wert für den Organismus haben und somit als Verstärker genutzt werden können.

**Sekundäre Verstärker**

= Reize (z.B. Klicken, Geld oder Token), das keinen inhärenten biologischen Wert hat, das jedoch mit primären Verstärkern verknüpft wurde oder Zugang zu ihnen ermöglicht.

**Negativer Kontrast**

= ein normalerweise akzeptabler Verstärker löst eine schwächere Reaktion aus, wenn ein bevorzugter Verstärker erwartet wird (Tinklepaugh, 1928).

* Studie von Kobre & Lipsitt (1972)
* Zuckerwasser ist bei menschlichen Säuglingen beliebter als normales Wasser
* Wenn in Sitzung 1 süßes Wasser gegeben wird und in Sitzung 2 normales Wasser, dann werden die Säuglinge weniger Saugen (R) zeigen

Beispiel: Wenn Kinder Fanta trinken, dann wird ihnen Wasser nicht gefallen. Wenn Kinder immer Wasser trinken, dann ist Wasser gut.

**Bestrafung**

= Konsequenzen, die dazu führen, dass sich die Wahrscheinlichkeit für ein Verhalten (R) reduziert.

Thorndike beobachtete jedoch, dass die Konditionierung von Verhalten mit Bestrafung nicht so gut funktioniert.



**Faktoren, die die Wirksamkeit von Bestrafung beeinflussen**

1. Bestrafungen führen zu mehr Variabilität im Verhalten

Bestrafung verringert die Auftretenswahrscheinlichkeit eines bestimmten Verhaltens, gibt aber nicht vor, durch was dieses Verhalten ersetzt werden soll. Der Organismus exploriert mögliche andere Verhaltensweisen.

2. Differenzierende Reize (SD) verleiten zum „Betrügen“

Beispiel: Für einen Raser ist die Anwesenheit eines Polizeiautos ein differenzierender Reiz, der ankündigt, dass Rasen jetzt bestraft werden würde. Der Fahrer lernt, nur dann zu rasen, wenn kein Polizist anwesend ist.

3. Konkurrierende Verstärkung kann Bestrafung untergraben

Beispiel: Obwohl beim Rasen eine hohe Geldstrafe riskiert wird, mag der Effekt der Bestrafung dadurch vermindert werden, dass das schnelle Fahren durch Spaß oder soziale Anerkennung verstärkt wird.

4. Intensität der ersten Bestrafung

Instrumentelle Konditionierung mit Bestrafung funktioniert am besten, wenn die erste Strafe möglichst hart ist. Ansonsten findet ein Habituieren an die Bestrafung an. Dies ist jedoch oftmals im Gegensatz zu unserem Gerechtigkeitsverständnis, wo man zunächst mit milden Strafen beginnt und diese dann erhärtet, wenn das Verhalten weiterhin auftritt.

Sicht von Frau Windmann: Verstärkung meist effektiver, weil Bestrafung aversiv sind (Stimmungen negativ), Belohnungen sind respektierend und wertschätzend. Bestrafungen aktivieren Angst/Vermeidungssystem (Amygdala), während Belohnungen Approach-System (PFC)

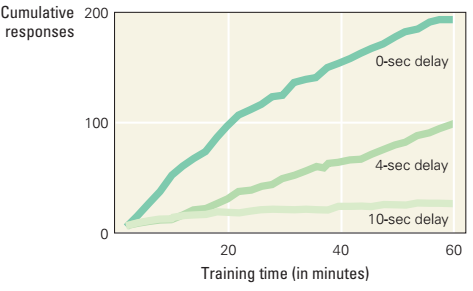
## Aufbau der S-R-C Assoziation

Nachdem wir die einzelnen Komponenten betrachtet haben, schauen wir jetzt, auf welche Weise sie miteinander verknüpft werden können.

**Verstärkungspläne**

= Regeln, die festlegen, wenn Konsequenzen in einem Experiment eintreten

### 1.5.1 Der Einfluss von Timing

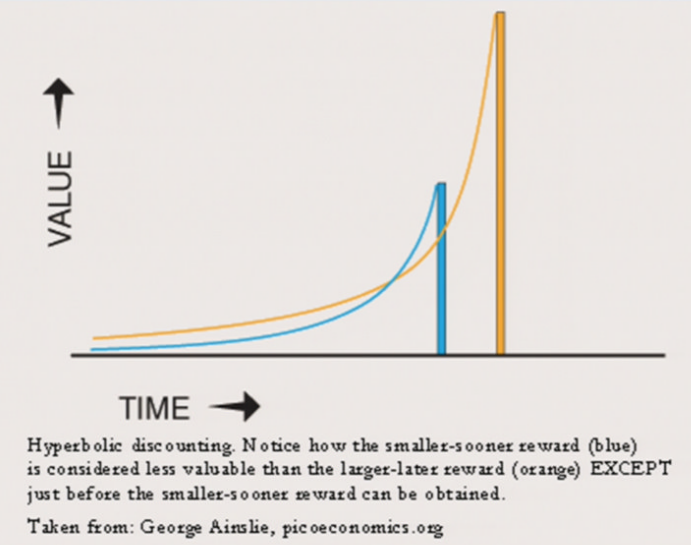
* Studie von Schlinger & Blakely (1994) mit Ratten, die einen Hebel betätigen müssen für Verstärkung durch Futter
* Konsequenzen ohne Verzögerung nach der Reaktion produzieren das schnellste Lernen
* Das gleiche Prinzip gilt für Bestrafungen, die auch am effektivsten sind, wenn sie unmittelbar nach dem unerwünschten Verhalten auftreten
* Warum führt zu späte Konsequenz zu Aberglaube? Wenn Tauben nicht-kontingent verstärkt wurden, zeigen dieses absurdes Verhalten wie im Kreisdrehen, in der Ecke picken, weil ihnen nicht klar ist, was verstärkt wird.

**Selbstkontrolle**

= Fähigkeit / Wille, auf einen kleinen unmittelbaren Verstärker zu verzichten (z.B. Eiscreme), um später einen größeren Verstärker zu erhalten (gesundes Körpergewicht). Weitere Beispiele: Rauchen aufgeben.

Dabei kann helfen die 1. **Stimuluskontrolle:** alle Süßigkeiten oder Zigaretten wegwerfen, sodass Einkaufen nötig wäre, um Drang nachzugeben. Oder 2. **Precommitment**: Lerngruppe beitreten oder Diät verkünden. Oder 3. **Attraktivität des langfristigen Ziels visualisieren**.

**Temporal Reward Discounting**



Je weiter entfernt ein Ziel erscheint, desto kleiner kommt der Wert vor.

### 1.5.2 Unterschiedliche Verstärkungspläne

Die Konsequenz muss nicht immer auf ein Verhalten folgen, damit Lernen stattfindet.

**Continuous reinforcement schedules**

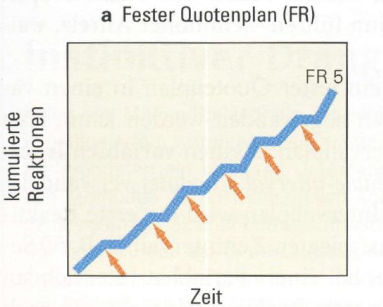
= Verstärkungspläne, bei denen bei jedem Verhalten eine Konsequenz erfolgt.

**Intermittierende (partielle) Verstärkungspläne**

= Verstärkungspläne, bei denen nur bei manchen Reaktionen eine Konsequenz erfolgt.

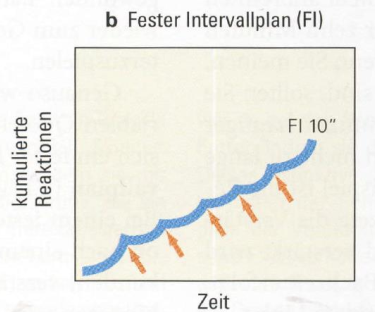
**Merke: intermittierende Verstärkung erhöht die Löschungsresistenz!**

**1. Fester Quotenplan (FR)**

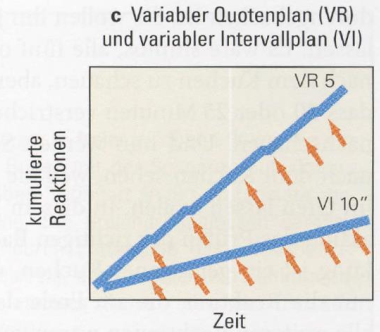
Hier muss eine bestimmte Anzahl an Reaktionen gezeigt werden, damit es zur Verstärkung kommt. FR 5 bedeutet, dass 5 Mal das Verhalten gezeigt werden muss, damit verstärkt wird.

* Ratten mit diesem Verstärkungsplan zeigen das Verhalten zuverlässig und pausieren dann kurz nach der Verstärkung (postreinforcement pause)
* Länge der Pause ist abhängig davon, wie viele Reaktionen für Verstärkung notwendig sind: ist länger bei einem FR 50 Plan
* Beispiel: Nach 10 Kaffees bei Coffee Bay, gibt es einen Kaffee gratis

**2. Fester Intervallplan (FI)**

Hier kann ein Verstärker nach einem fixen Zeitintervall erhalten werden. Bei einem Verstärkungsplan FI 10‘‘ wird eine Ratte für die erste Reaktion nach einem 10 Sekunden Intervall verstärkt.

* Man beobachtet, dass Ratten nach einer Verstärkung kurz pausieren/wenige Reaktionen zeigen und dann anschließend ihre Reaktionsrate steigt, wenn sich das Intervall dem Ende nähert, bis sie die Verstärkung erhalten
* Anderes Beispiel: Beim Nachsitzen schaut man am Ende der Stunde häufiger auf die Uhr als am Anfang

**3. Variabler Quotenplan (VR)**

Hier gibt es einen Verstärker bei einer durchschnittlichen Anzahl von Reaktionen. VR 5 bedeutet, dass eine Verstärkung gibt nach durchschnittlich 5 Reaktionen, ist aber auch schon früher möglich. Dieses Verstärkerplan reduziert dadurch deutlich die post-Verstärkung-Pause, die man bei einem FR Plan sieht.

Beispiel beim Menschen: Slotmaschine, die durchschnittlich nur bei jedem 10. Spiel ein Gewinn ausschüttet.

**4. Variabler Intervallplan (VI)**

Ähnlich wie fester Intervallplan, nur dass Verstärkung nach durchschnittlich 10 Sekunden möglich ist. Diese Pläne verursachen höhere Reaktionsraten als die FI-Pläne. Beispiel: Toms Freundin macht ein Auslandssemester und postet ungefähr 1 x am Tag einen Blogpost. Wenn dieser jeden Tag zur selben Zeit kommt, würde er nur einmal am Tag gucken. Falls er nicht weiß, wann er kommt, guckt er mehrmals am Tag.

Kurve zeichnen oder Beispiele geben können

## Wahlverhalten

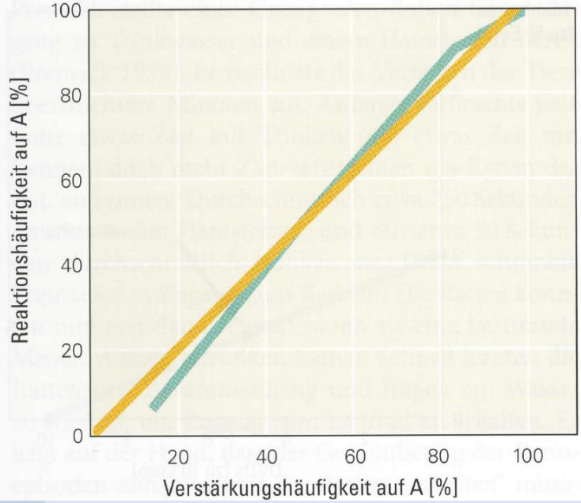
Wahlverhalten bei **konkurrierenden Verstärkungsplänen**

= Verstärkungsplan, bei dem ein Organismus verschiedene Reaktionen zeigen kann, die jeweils zu unterschiedlichen Konsequenzen führen. Ermöglichen es zu untersuchen, wie die Zeit und Anstrengung zwischen verschiedenen Optionen aufgeteilt werden.

### 1.6.1 Das Matching-Law von Herrnstein

Szenario: Taube befindet sich in einer Kammer mit zwei Tasten, Taste A und B. Betätigen der Taste A wird mit einem VI 1-m Plan verstärkt, Taste B mit einem VI 2-m Plan.

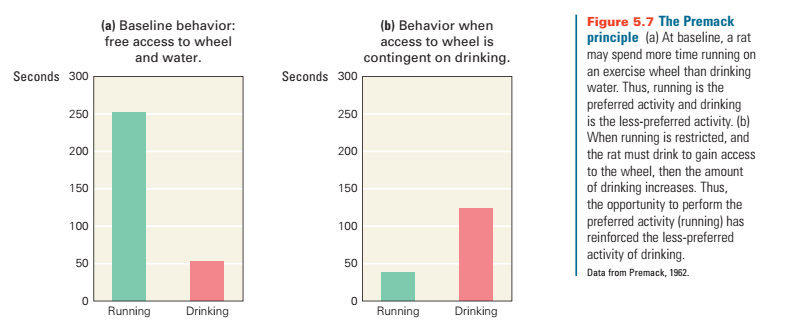
* Tauben werden 67% der Reaktionen bei A zeigen und 33% bei Taste B.

Ein Bild, das Text enthält.

Automatisch generierte Beschreibung

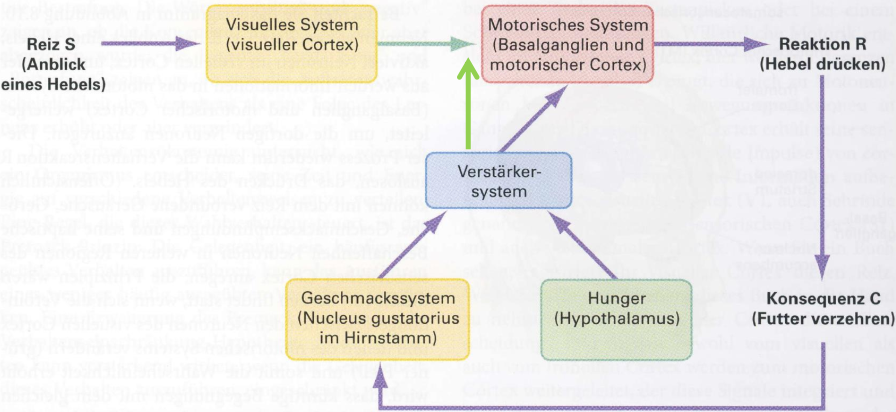
### 1.6.2 Das Premack-Prinzip

= Verhalten selbst kann ein Verstärker sein. Fernsehen oder im Hamsterradrennen machen einfach so Spaß. Wenn man dieses Verhalten einschränkt (z.B. Rennen), dann kann man das Rennen auch als verstärkende Konsequenz einsetzen.



Nicht wirklich verstanden.

# 2. Gehirnsubstrate



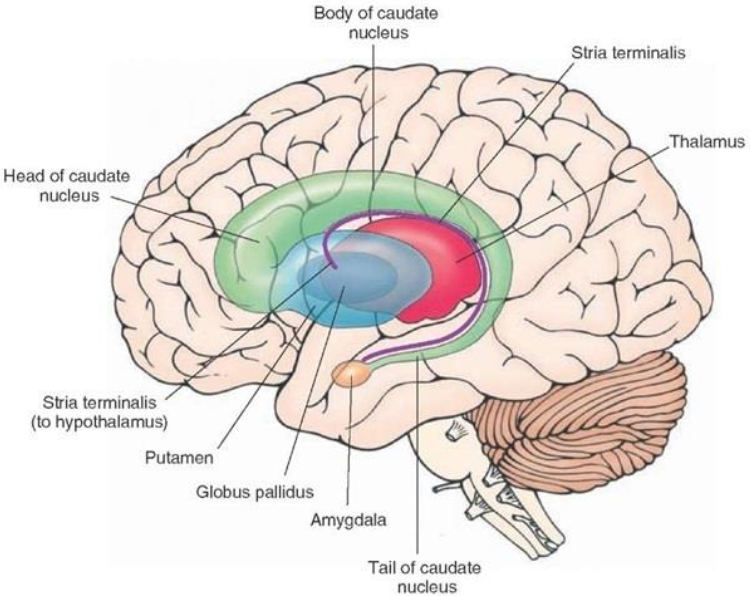
Es gibt einen diskriminierenden Reiz und eine Reaktion, die miteinander gekoppelt werden sollen. Diese Verbindung wird beim Instrumentellen Lernen durch ein Verstärkersystem gestärkt. Voraussetzung: es gibt eine Motivation (es besteht ein Bedürfnis, z.B. Hunger und es gibt etwas, das dieses Bedürfnis stillt, z.B. Futter).

Zwei Hirnarealen kommt eine Schlüsselfunktion zu:

1. Dorsales Striatum, das zum Lernen der SD 🡪 Verknüpfung sehr wichtig ist
2. Orbitofrontaler Cortex, der für das Erlenen erwarteter Konsequenzen wichtig ist
3. Weitere Hirnareale, die bewerten, ob die Konsequenzen verstärkend oder bestrafend sind

## 2.1 Dorsales Striatum und Reiz-Reaktions-Lernen

Anatomie

Informationen aus dem sensorischen Kortex können über die **Basalganglien** zum motorischen Kortex weitergeleiten werden.

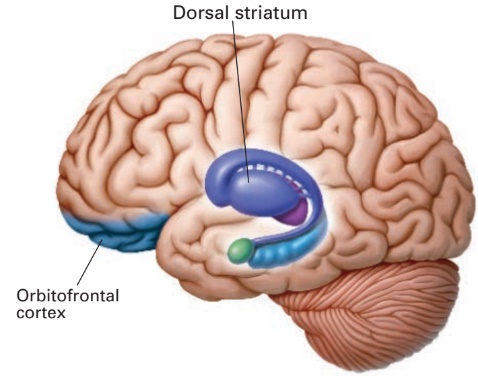
= subkortikale Kerngebiete

* Dorsales Striatum ist Teil der Basalganglien und kann in Putamen und Nucleus caudatus unterteilt werden
* Es erhält hoch verarbeitete Information aus den sensorischen Kortexarealen und projiziert zum Motorkortex, welcher eine behaviorale Reaktion erzeugt

Funktionen

* Dorsales Striatum wird benötigt, um SD – R – Assoziation aufgrund von Verstärkung oder Bestrafung zu erlenen
* Die erzeugten Assoziationen scheinen relativ automatisch abzulaufen

## 2.2 Der Orbitofrontale Cortex und Lernen, Konsequenzen vorherzusagen

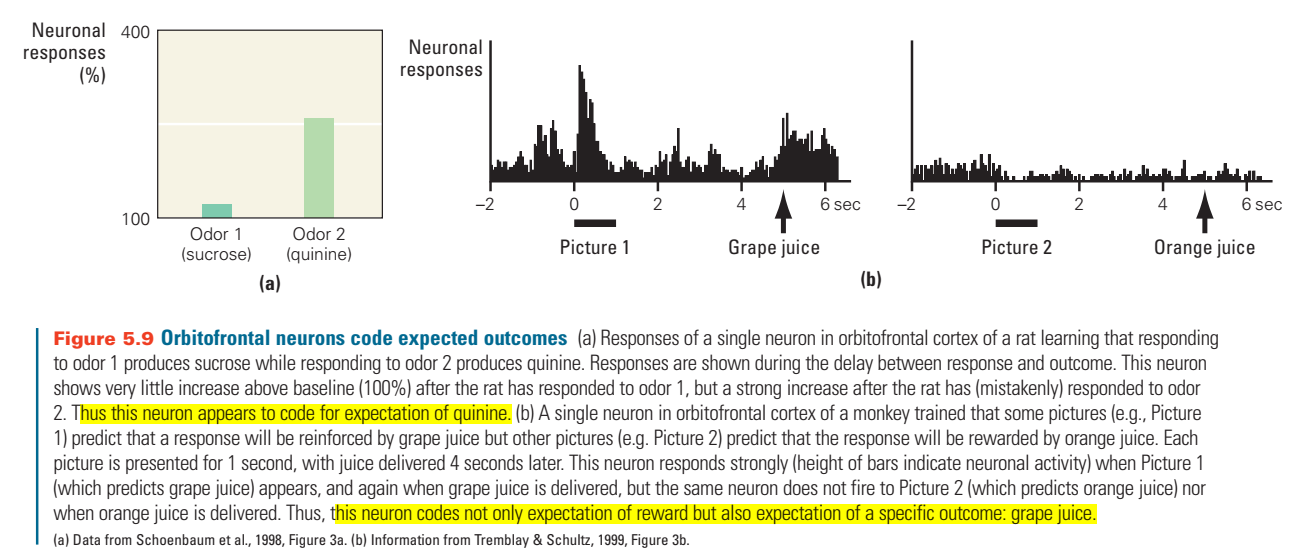
Anatomie

* Erhält Input aus allen sensorischen Modalitäten und Informationen aus viszeralen Empfindungen (wie Hunger und Durst)
* Integration verschiedener Informationstypen
* Outputs zum Striatum und dadurch Beteiligung bei der Koordination von motorischen Verhalten

Funktion

Der orbitofrontale Cortex ist involviert, wenn es darum geht, die Konsequenzen von Verhalten vorherzusagen. Er leistet somit außerdem einen Beitrag für zielgerichtetes Verhalten, indem er vorhergesagte Konsequenzen prädiziert.

Evidenz stammt aus neuronalen Aktivitätsaufnahmen:

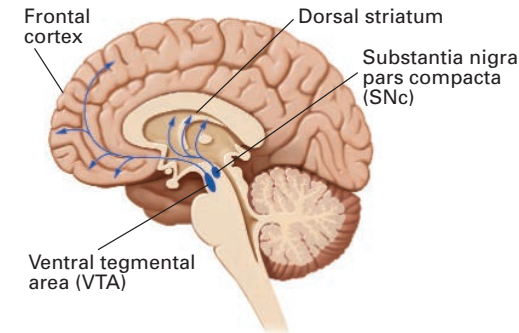
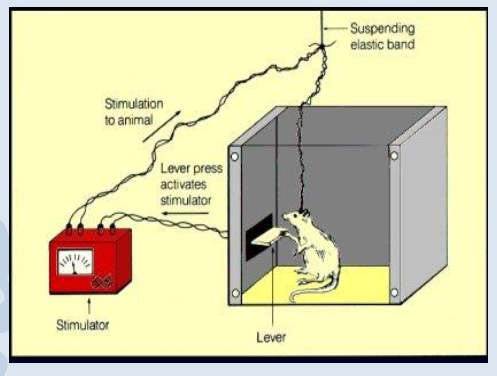


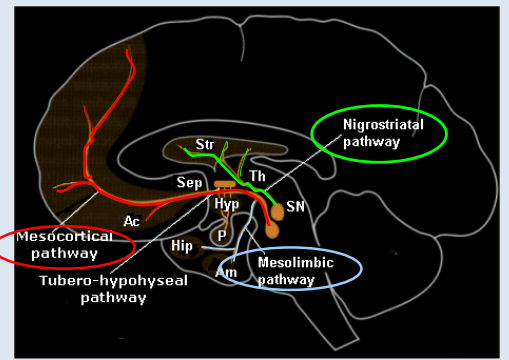
## 2.3 Das dopaminerge System und Bewertung der Konsequenz

Wie bestimmt das Hirn, ob eine Konsequenz ein Verstärker oder eine Bestrafung ist?

### 2.3.1 Ventral tegmental area (VTA)

* Olds (1954) fand das VTA als „Verstärkerzentrum“
* Ratten betätigen einen Hebel, der das VTA stimuliert, bis sie vor Erschöpfung zusammenbrechen
* VTA ist eine Region im Mittelhirn, das dopaminerge Neuronen besitzt, die in den Frontalkortex und andere Hirnareale projizieren

****



### 2.3.2 Dopamin und Verstärkung

Striatum und Dopamin wird aktiviert, wenn Belohnungsreize gezeigt oder antizipiert werden:

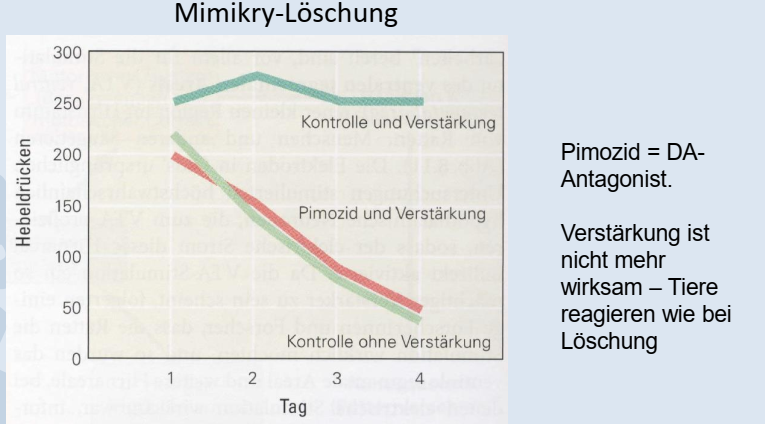
Bei Ratten: Futter, Sex, Drogen und sek. Verstärker

Bei Menschen: Saft, Kokain, Geld, Humor, Partner, schöne Menschen

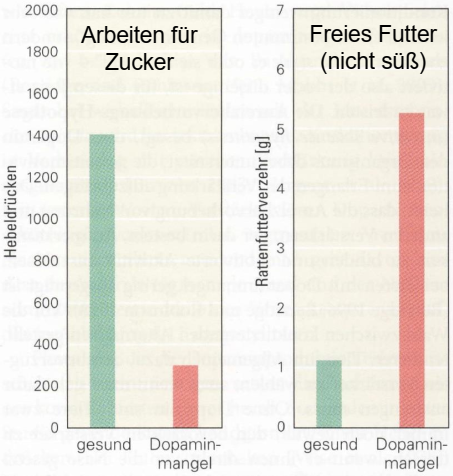
Was genau macht das Dopamin?

**1. Anhedonie Hypothese**

Pimozid-Gruppe verhält sich so, als würde sie keine Verstärkung erhalten (Mimokry-Löschung). Hypothese, dass das Dopamin die lustvolle Qualität des Verstärker vermittelt. Dopamin vermittele Hedonie, Freude am Verstärker.

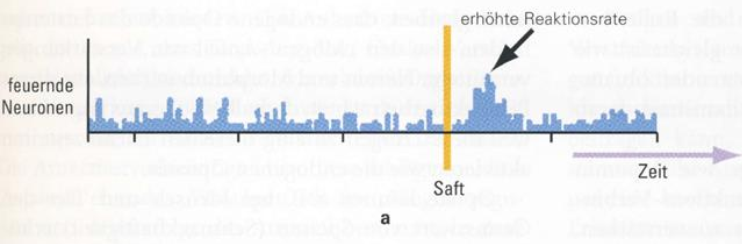


**Die Anreizhervorhebungshypothese** (Berridge & Robinson, 1998) besagt, dass Dopamin den Organismus dazu motiviert, für einen Verstärker zu „arbeiten“.

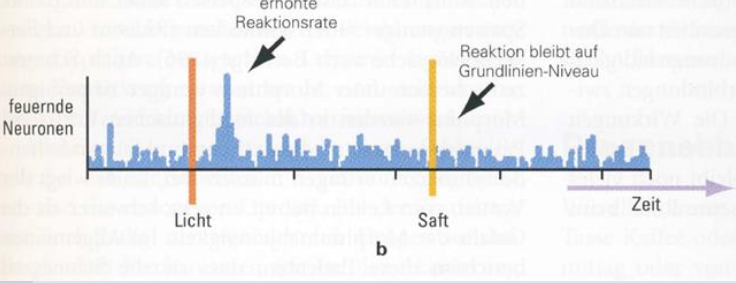
* Stimulation des dopaminergen Systems erhöht das Verlangen nach einem Stimulus
* Dopamin verstärkt die Assoziation zwischen Stimulus und Reaktion während der operanten Konditionierung
* Dieser Befund spricht gegen die Anhedonie-Hypothese, da der Anreiz/Verstärkungswert noch vermittelt wird. Sondern: Motivationaler Wert des Anreizes wird durch Dopamin vermittelt.

**Belohnungserwartungshypothese (Schultz)**

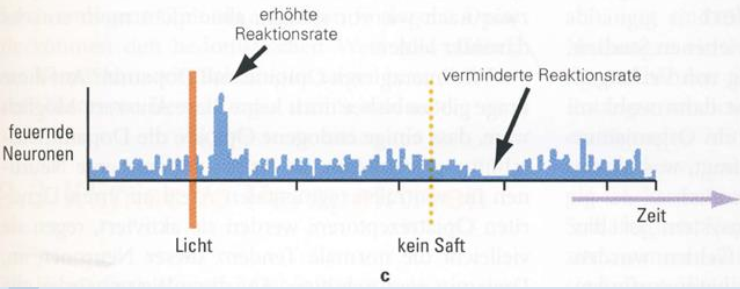
Ableitungen aus Neuronen aus dem Striatum



1. Affe erhält überraschend Saft und dadurch gibt es einen Aktivitätsburst.



1. Tiere wurden appetitiv klassisch konditioniert (CS: Licht; US: Saft). Dies ist ein Trial, indem der Affe gelernt hat, dass Licht Saft vorhersagt. Aktivitäts-Burst jetzt nach dem Licht und nicht nach dem Saft. Codierung für: das Licht sagt den Saft vorher.



1. Trial, bei dem Licht gezeigt wird, aber kein Saft gegeben wird. Nach der ausgelassenen Saftrate gibt es eine verminderte Reaktionsrate.

Im Sinne des Rescorla-Wagner-Modells: dopaminerge Neuronen codieren Erwartungen und Vorhersagefehler.

# 3. Anwendungen

z.B. Verhaltens- oder Substanzabhängigkeiten