



Beispielbild

Lernen & Gedächtnis

Operante Konditionierung: Verstärkung

SoSe 2007

Einführung

Belohnung ist das einfachste Prinzip, um ein erwünschtes Verhalten häufiger auftreten zu lassen!

Oder ?



Lernen hat mittel- und langfristig Belohnungswert



Das Einstellen von Rauchen hat mittel- und langfristig Belohnungswert

Thorndike

Edward Lee Thorndike (1874-1949)

Erste experimentelle Studien zur Wirkung von Belohnung

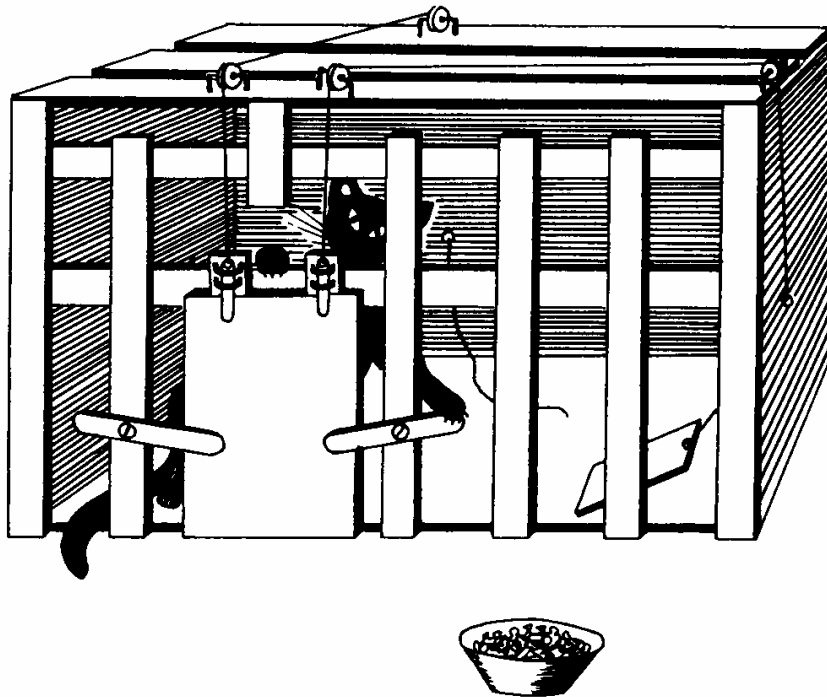


Thorndikes Ansatzpunkt war die Idee, dass andere Säuger scheinbar erstaunlich ‚intelligente‘ Leistungen erbringen können.

Kritik: Berichte über tierische Intelligenz sind anekdotisch, verfälscht und zeigen eher inzidentielles Verhalten als tatsächliche Kompetenz.

Thorndike

Gesetz der Wirkung (Law of Effect)



Puzzle Box: Käfig, der mit einem Hebel oder einer Schnur zu öffnen ist. Vor der Box steht eine Futterschale.

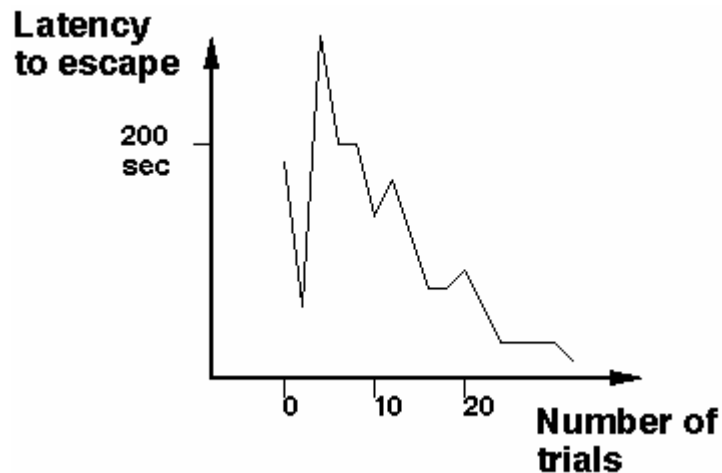
Eine Katze reagiert zunächst mit Kratzen an verschiedenen Teilen der Box – oder anderem ‚irrationalen‘ Verhalten.

Nach 5-10 Minuten wird per Zufall der Lösungsweg gefunden.

Aber was passiert im nächsten Trial?

Thorndike

Gesetz der Wirkung (Law of Effect)



Die Katze hat keine plötzliche Einsicht in den Lösungsweg, sondern reagiert wie beim ersten Trial.

Nur graduell wird die Latenz bis zur Befreiung aus dem Käfig kürzer.

Theorie: Die Belohnung außerhalb der Box führt dazu, dass die Katze langsam eine Assoziation zwischen den Cues in der Box und der operanten Reaktion (Hebeldruck) herstellt.

Gesetz der Wirkung (Law of Effect)

Law of Effect:

Eine Verbindung wird begünstigt, also „verstärkt“ (reinforced), wenn die Reaktion zu einer befriedigenden (lustbetonten) Nachwirkung führt. Sie wird abgeschwächt, wenn die Nachwirkung unbefriedigend (unlustbetont) ist.

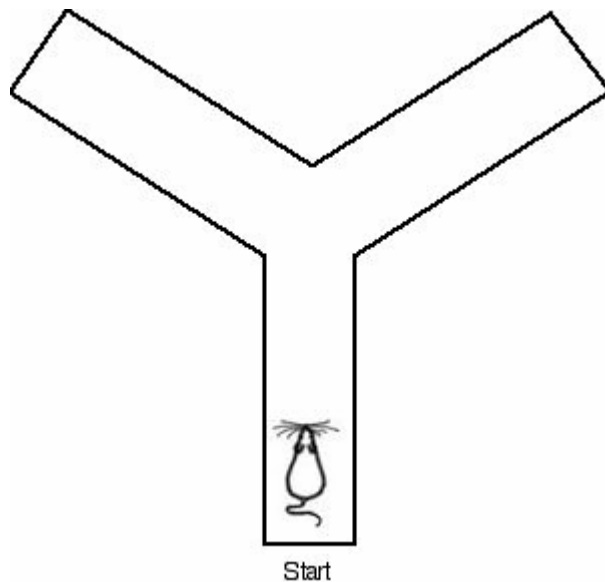
Problem 1: Was meint Thorndike mit ‚befriedigender‘ Nachwirkung? Ist dies nicht ein kognitives Konstrukt? Und macht es das Effektgesetz nicht zirkulär?

Problem 2: Was unterscheidet diesen Lernvorgang eigentlich von der klassischen Konditionierung?

Thorndike

Problem: Was ist ‚befriedigend‘?

Thorndike: Ein befriedigender Zustand besteht dann, wenn das Tier nichts unternimmt, um diesen Zustand zu vermeiden, jedoch vieles, um ihn zu erhalten.



Aber: Sheffield et al (1951) zeigten, dass selbst ein ‚unbefriedigender‘ Zustand (wiederholt unterbrochene Kopulation) eine Verhaltensreaktion häufiger und intensiver machen kann

Deshalb spricht man eher von Verstärkung als von Belohnung. Ein Verstärker erhöht die Frequenz des Verhaltens, wenn er nach einer bestimmten operanten Reaktion gezeigt wird.

Thorndike

Problem: Was unterscheidet Verstärkung und Konditionierung?

Klassische Konditionierung

Reiz (CS) – Reiz (US)

Präsentation des US hängt alleine von der Präsentation des CS ab.

Operante Konditionierung

Aktion (R) – Reiz (S)

Präsentation des Reizes hängt alleine von Durchführung eines Verhaltens ab.

Obwohl die Prozeduren unterschiedlich sind, können jedoch die Vermittlungsprozesse identisch sein:

1. Prinzip der Kontiguität oder der Kontingenz
2. Graduelle Verstärkung einer assoziativen Beziehung .

Typen von Verstärkern

Primäre Verstärker

Primäre Verstärker sind alle Reize, die von Geburt an effektiv sind. Sie benötigen kein spezielles Training. Dazu gehören...

...Futter

...Wasser

...Fortpflanzungsmöglichkeit

...Sensorische Verstärkung

Butler (1954):

Reaktion 1: Fenster zum Laborraum öffnet sich für 30 Sekunden.

Reaktion 2: Sichtblende geht herunter.

Für Affen ist die Reaktion 1 ein effektiver Verstärker.

Insgesamt kann jede visuelle Stimulation als Verstärker eingesetzt werden

Typen von Verstärkern

Primäre Verstärker



David Premack

Der Zugang zu einer Aktivität, die mit hoher Wahrscheinlichkeit ausgeübt wird, kann als Verstärker für eine Aktivität verwendet werden, die seltener ausgeübt wird. (Premack-Prinzip)

Kann effektiv eingesetzt werden – und wird intuitiv eingesetzt: Kinder können sehr effektiv zu diszipliniertem Verhalten (seltene Reaktion) geführt werden, wenn Ihnen kurze Phasen von undiszipliniertem Verhalten (häufige Reaktion) gestatte werden (Homme et al., 1963).

Typen von Verstärkern

Sekundäre Verstärker

Sekundäre Verstärker werden erst durch Erfahrung erworben.

Zu Ihnen gehört Geld, aber auch Lob.

Die sekundären Verstärker werden dadurch ‚erworben‘, indem sie häufig mit einem primären Verstärker aufgetreten sind.

Wolfe-Experiment (1936):

Affen erhalten nach Hebeldruck zunächst Trauben. Dann wird der Hebeldruck schwieriger – und die Trauben durch Spielmünzen ersetzt.

Die Affen drücken weiter den Hebel – und zeigen innerhalb der Gruppe ein (Menschen-) typisches Arbeitgeberverhalten.

Typen von Verstärkern

Soziale Verstärker

Soziale Verstärker können nicht genau von primären und sekundären Verstärkern abgegrenzt werden. Sie werden jedoch alleine aus dem Verhalten von Mitgliedern aus der gleichen Spezies gezogen.

Soziale Verstärker können eine angeborene Disposition sein, die erfahrungsabhängig moduliert wird. Zudem sind wir den Verstärkern ununterbrochen ausgesetzt (Aufmerksamkeit, Lob, Tadel.).

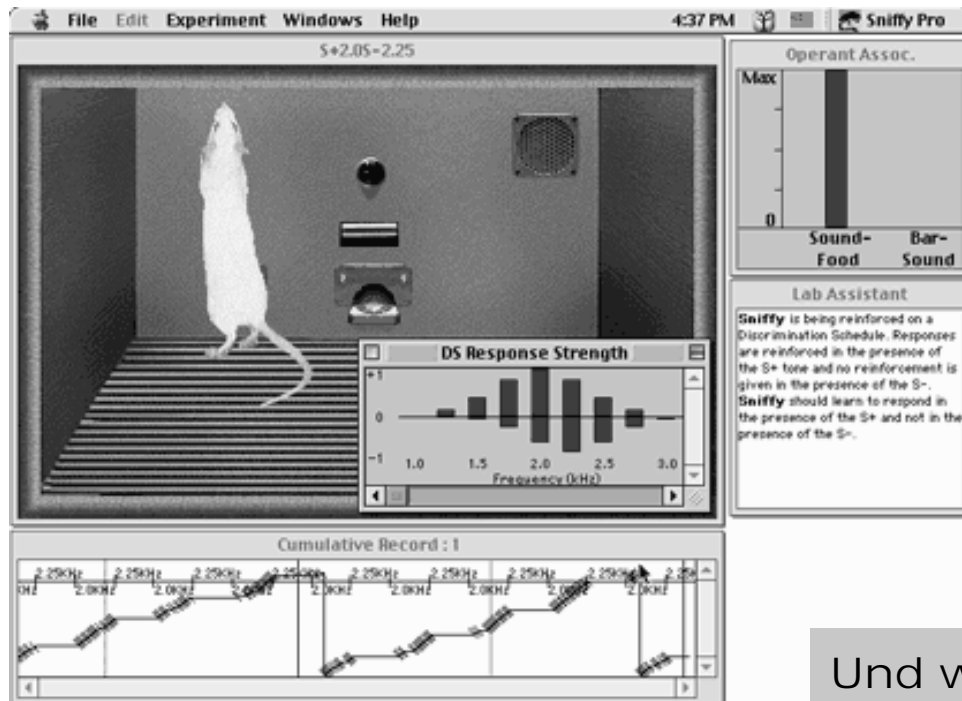
Experiment von Allen et al. (1964)

Was macht man mit einer 4-jährigen, die nicht mit anderen Kindern spielt – und eher den Kontakt von Erwachsenen sucht?

1. Verringerung der Zuwendung, wenn sie sich isoliert
2. Verstärkung der Zuwendung, wenn sie mit anderen Kindern spielt.

Resultat: Anstieg des Anteils des sozialen Spielens von 10% auf 60%. Zudem konnte die Verstärkung langsam reduziert werden.

Verstärkung



Wie bekomme ich Sniffy dazu, den Hebel zu drücken?

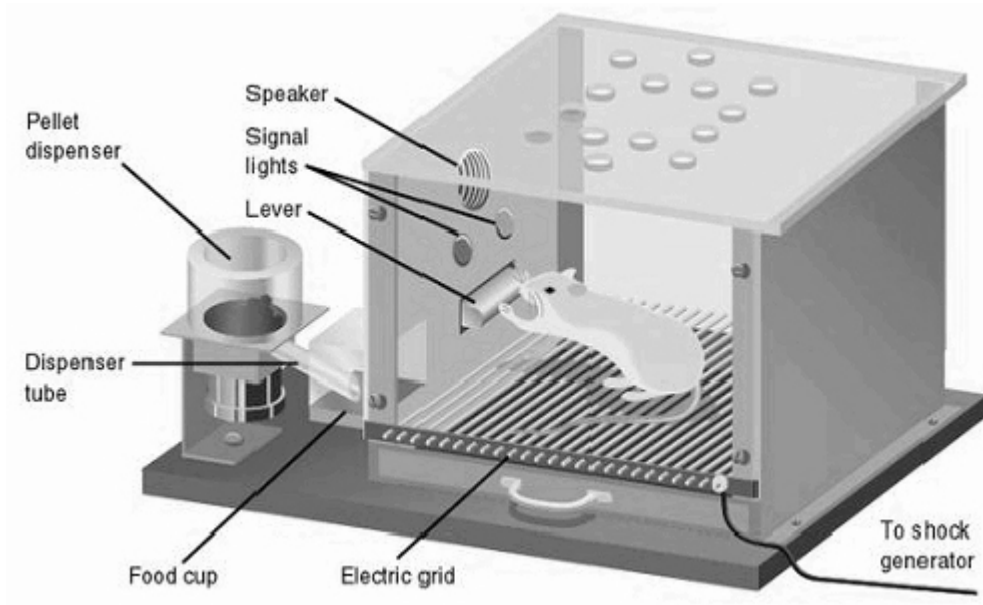
Was ist das ‚Magazin-Training‘?

Was ist ‚Shaping‘?

Und was ist eine kumulative Frequenz?

Verstärkung und Verzögerung

Wie wichtig ist die Kontiguität zwischen Reaktion und Reiz?



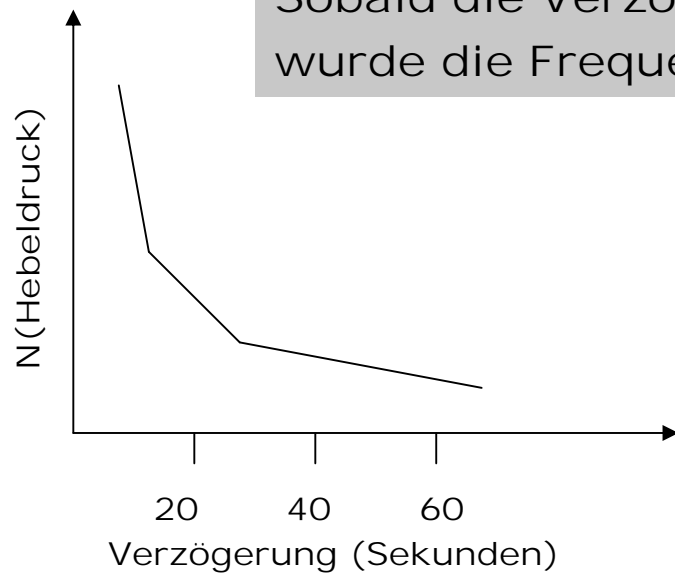
Wie tolerant ist die Sniffy, wenn sie nach ihrem Tastendruck nicht unmittelbar die Verstärkung erhält?

Und wie kann man möglicherweise ihre Toleranz steigern?

Verstärkung und Verzögerung

Welche zeitliche Verzögerung wird toleriert?

Sobald die Verzögerung auch nur einige Sekunden betrug, wurde die Frequenz der Reaktion drastisch reduziert.



Der Grund ist nicht das schlechte Gedächtnis für die Reize, sondern das Problem der Reaktions-Kontingenz. Der Organismus muss herausfinden, welche seiner Reaktionen den Reiz bewirkt hat.

Experimentator:

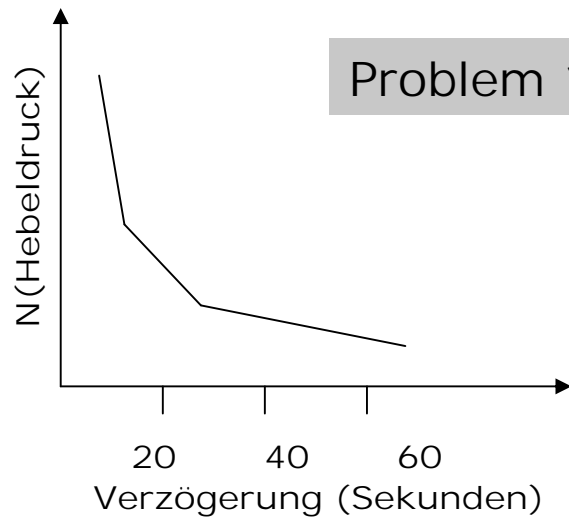
R(x) - Verstärker

Versuchstier:

R(1) – R(2) - R(x) – R(3) – R(4) -Verstärker

Verstärkung und Verzögerung

Wie kann Sniffy toleranter werden?



Problem 1

Problem 2

Versuchstier:

$R(1) - R(2) - R(x) - R(3)$
– $R(4)$ -Verstärker

Ein sekundärer Verstärker kann über die Zeitdauer helfen.

Die Kontingenz kann eindeutiger gemacht werden, wenn synchron mit dem erwünschten Verhalten ein weiteres äußeres Ereignis einsetzt, welches selber keinen Verstärkungscharakter haben muss

Verstärkung und Verzögerung

Konsequenzen für das menschliche Lernen

Sind Menschen toleranter für zeitliche Verzögerungen?

Shanks et al. (1989):

Tastendruck auf der Tastatur induziert einen visuellen Reiz auf dem Bildschirm

Kontingenz zwischen Reaktion und Reiz wird in 80% der Fälle richtig erkannt, wenn sie gleichzeitig erfolgt.

Kontingenz zwischen Reaktion und Reiz wird in 30% der Fälle richtig erkannt, wenn sie um 4s verzögert ist!

Verstärkung und Verzögerung

Konsequenzen für das menschliche Lernen

Verzögerung und Attraktivität der Verstärkung: Tierexperiment

Auf Reaktion 1 von Tauben folgt 2s Zugang zu Futter

Auf Reaktion 2 von Tauben folgt 4s Zugang zu Futter

In 95% der Fälle führen die Tauben Reaktion 1 aus

Verzögerung und Attraktivität der Verstärkung: Humanexperiment

Studenten werden nach Teilnahme am Versuch mit einer Zeitverzögerung entlohnt:

12\$ nach 6 Tagen oder

16\$ nach 12 Tagen

Studenten bevorzugen schnelle Belohnung

Verzögerung der Belohnung kann mit Hilfe der Sprache zwar vermittelt werden, beseitigt jedoch nicht die grundlegende Bevorzugung schneller Belohnung (inzentive Verstärkung)

Verstärkung und Verzögerung

Konsequenzen für das menschliche Lernen

Phillips-Programm (1968)

Jungen sollen lernen, ihre Hausaufgaben zu erledigen.

Verstärkung: Sie dürfen an den Wochenenden eine Stunde länger wach bleiben.

Problemfall: Tom

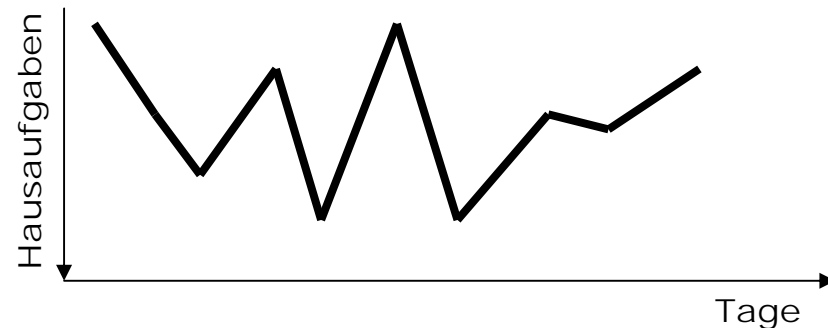


Problemfall: Tom

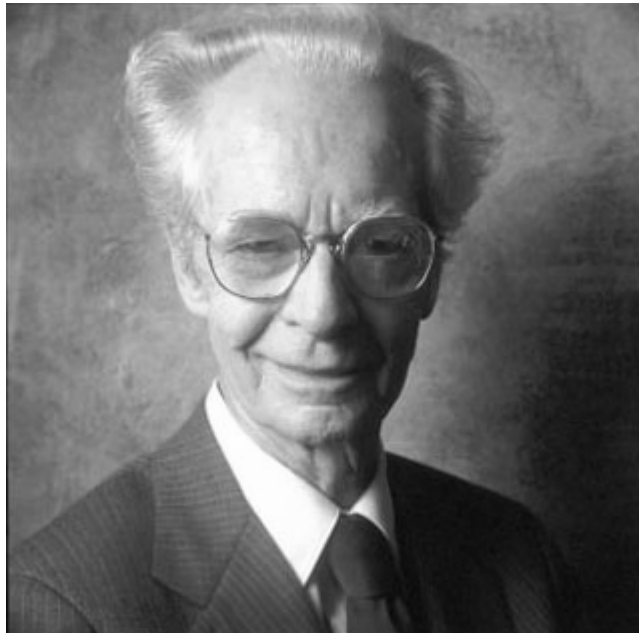
Ist die Verstärkung nicht attraktiv genug?

Ist nicht zeitnah genug?

Effekt von sofortiger Verstärkung:
Tom darf am gleichen Abend eine Stunde länger aufbleiben.



Verstärkungspläne



Problem für engagierte Behavioristen am Wochenende:
Die Futterpillen gehen langsam aus!

Konsequenz:
Nicht mehr jede Reaktion wird belohnt, sondern nur noch jede erste Reaktion nach einer Minute.

SNIFFY

Effekt:
Ratten reagieren initial weniger stark, pendeln sich dann aber auf ein stabiles Niveau ein.

Verstärkungspläne

Verstärkung jeder Reaktion:
CRF (Continuous Reinforcement Schedule)

	RATIO	INTERVALL
FIXED	FR	FI
VARIABLE	VR	VI

Verstärkungspläne



Ratio - Schedule
 Verstärkung beruht auf der Zahl von Reaktionen, die ausgeführt worden sind

Interval - Schedule
 Verstärkung beruht auf der Zeit, die seit der letzten Verstärkung vergangen ist. Wichtig: Reaktion muss trotzdem erfolgen

	RATIO	INTERVAL
FIXED	FR	FI
VARIABLE	VR	VI

Verstärkungspläne



Fixed - Schedule
 Das Intervall oder die Zahl der Reaktionen bleibt konstant

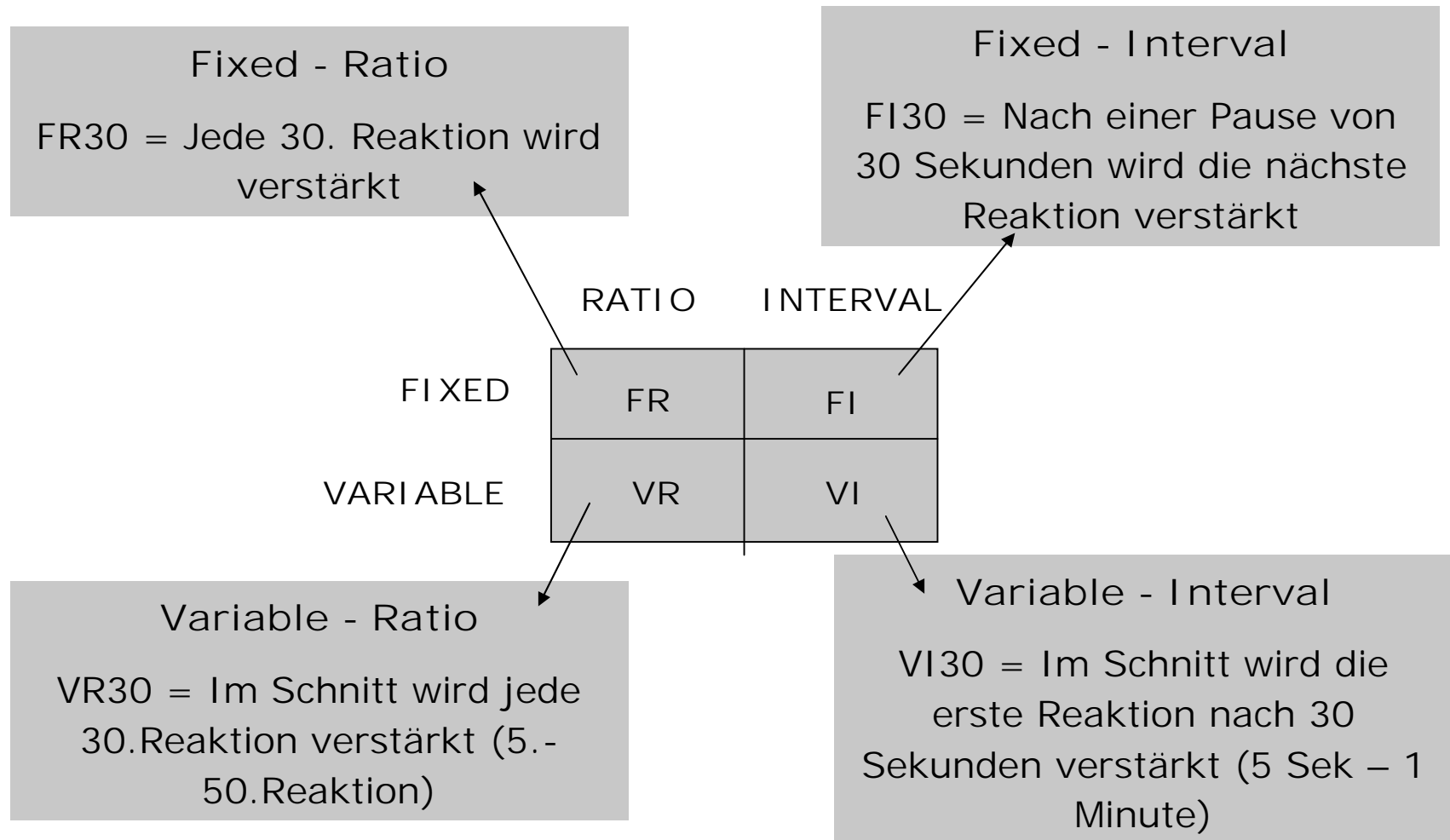
Variable - Schedule
 Das Intervall oder die Zahl der Reaktionen bewegen sich nur um einen Mittelwert.

FIXED

VARIABLE

RATIO	INTERVAL
FR	FI
VR	VI

Verstärkungspläne



Verstärkungspläne

Jeder Plan hat eine unterschiedliche Auswirkung auf das Verhalten

Aufnahme der kumulativen Antworten

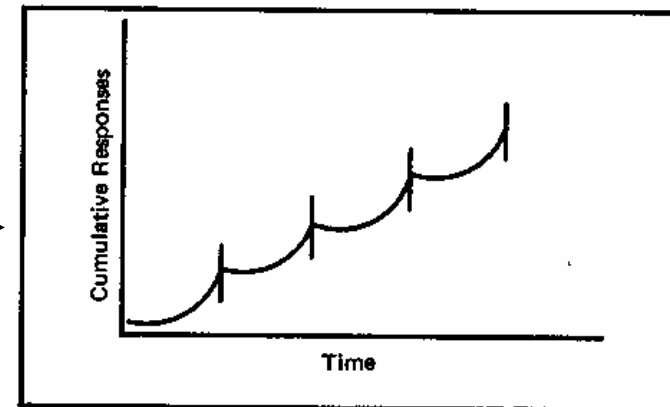
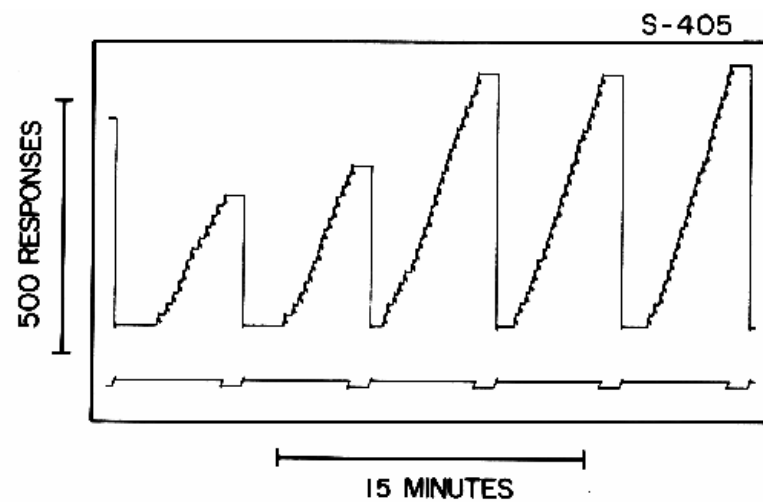


Figure 9. Cumulative response curve for a fixed interval schedule

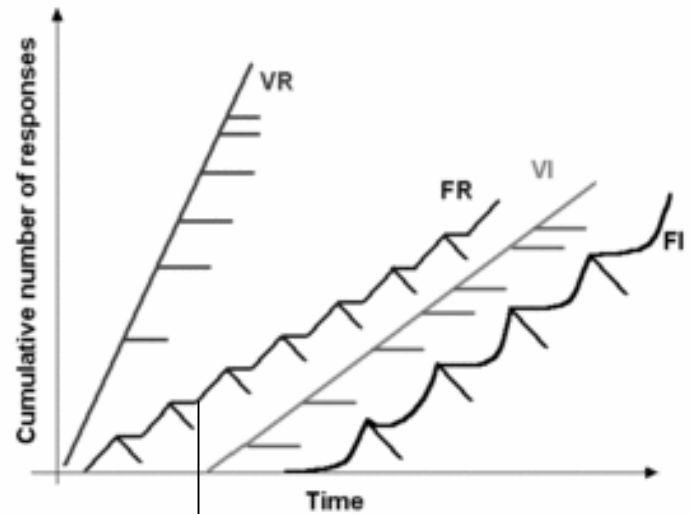
Linearer Anstieg = monotoner Anstieg

Schwingungen hängen von der Prädiktabilität der Verstärkung ab!

Verstärkungspläne



Gewünscht:
Konstante
Reaktion über
einen
Zeitbereich
hinweg

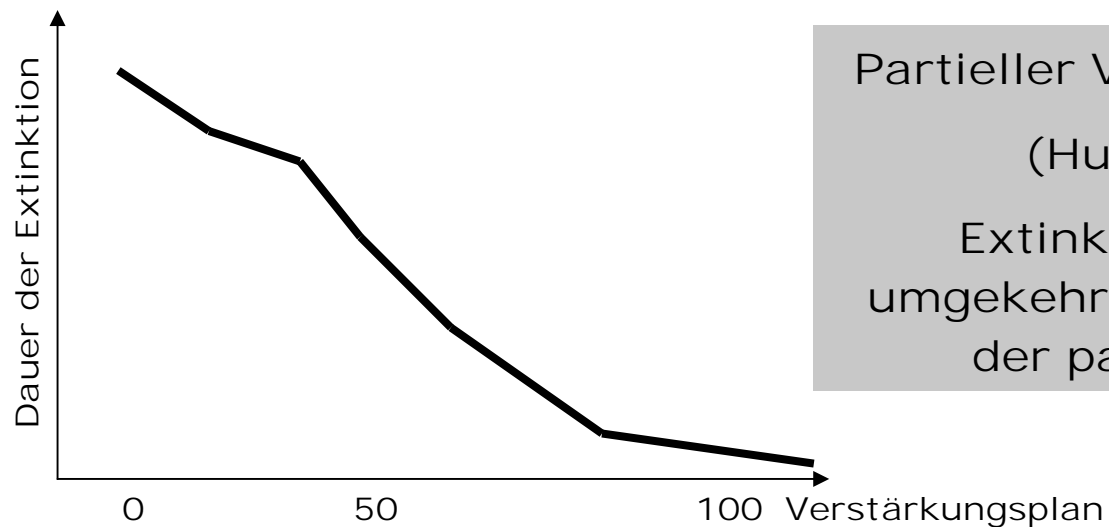


Ungewünscht:
Reaktion steigt
nur in
Erwartung
einer Reaktion
zyklisch an

Fixed Ration:
Konstanter Anstieg – bis auf die
Phasen unmittelbar nach der
Verstärkung (Run-and-Pause)

Verstärkungspläne

Aber: welcher Plan ist der Beste?



Partieller Verstärkungseffekt (PRE):
(Humphreys Paradox)
Extinktionsrate verhält sich
umgekehrt proportional zum Anteil
der partiellen Verstärkung

Ursache: Wie schwierig ist es für den Lernenden das Ende der Verstärkung des Verhaltens zu erkennen?

Regelmäßige Verstärkung: Effekt ist eindeutig

Unregelmäßige Verstärkung: Effekt ist nicht eindeutig

Verstärkungspläne

Aber: welcher Plan ist der Beste?

Vorteile kontinuierlicher Verstärkung (CRF)

1. Erwerb einer Reaktion geht sehr schnell

Nachteile kontinuierlicher Verstärkung (CRF)

1. Schnelle Extinktion
2. Hohe Kosten und Aufwand

Konsequenz

1. Phase 1: CRF-Plan
2. Phase 2: Übergang in einen variablen Verstärkungsplan (VR, VI)

VR-Plan, wenn viele Reaktionen gewünscht werden.
Gefahr der Frustration!

VI-Plan, wenn wenige Reaktionen auch einen konstanten Niveau gewünscht werden.

Verstärkungspläne

Anwendungsbeispiel von Kandel et al (1976)

Problem:

Gefängnisinsassen sollen Schulabschluss nachholen.

Für jede bestandene Prüfung (Grades) bekommen Sie eine fixen Prozentsatz an Punkten, der Ihnen Verstärker (Zigaretten, etc.) bringt.

Plan funktioniert, aber sehr langsam!

Lösung:

Abgestufter Verstärkungsplan:

Bestande Prüfung in 90 Tag bringt 120 Punkte.

Bestande Prüfung in 4 Tag bringt 900 Punkte.

Bestande Prüfung in 1 Tag bringt 4700 Punkte.

Effekt:

Ein Englischprogramm (4.6 Jahre), ein Leseprogramm (4.9 Jahre), ein Mathematikprogramm (6.6 Jahre) werden in Monaten bewältigt!

Der Lernende wies in diesem Fall einen IQ von 65 auf!