

Vorlesung Lernen und Gedächtnis
Modul A2

Instrumentelles Konditionieren II

Prof. Dr. Thomas Goschke

Überblick

Thorndikes Gesetz des Effekts

Skinner's Forschung zum operanten Konditionieren

Was wird beim instrumentellen Konditionieren gelernt?

Shaping und Verhaltenssequenzen

Verstärker und Verstärkungspläne

Rolle von Kontiguität und Kontingenz

Abbau unerwünschter Verhaltensweisen und Bestrafung

Flucht- und Vermeidungstraining

Stimuluskontrolle: Generalisierung und Diskrimination

Artspezifische Reaktionsprogramme und angeborene Lernbereitschaften

Verstärkung und Verhaltenskontrolle

Skinner's operationale Definition:

- Verstärker = jeder Reiz, der die Auftretenswahrscheinlichkeit eines Verhaltens verändert (erhöht oder verringert)

Vielfalt spezifischer Verstärker

- *Primäre Verstärker*
 - Futter, Wasser, Sex etc.
- *Materielle Verstärker*
 - Spielsachen, Geld etc.
- *Soziale Verstärker*
 - Zuwendung, Anerkennung, Lob etc.
- *Aktivitätsbezogene Verstärker*
 - Angenehme Tätigkeiten
- *Informative Verstärker*
 - Informationen, Erkenntnisse

Arten von Verstärkern

	Verhaltenskonsequenz	
	Angenehmer Reiz	Unangenehmer Reiz
Auf die Reaktion folgt ein Reiz	Positive Verstärkung (Belohnung) ↑	Bestrafung 1. Art ↓
Reaktion eliminiert oder vermeidet Reiz	Bestrafung 2. Art (Omission training) ↓	Negative Verstärkung (Flucht / Vermeidung) ↑

	Verhaltenskonsequenz	
	Angenehmer Reiz	Unangenehmer Reiz
Auf die Reaktion folgt ein Reiz	Futter, Lob, Geld	Schmerz, Tadel
Reaktion eliminiert oder vermeidet Reiz	Futter, Lob, Geld wird entzogen	Schmerz wird beendet Tadel wird nicht erteilt

(Die Pfeile geben an, ob die Reaktionsrate zu- oder abnimmt)

Primäre und sekundäre Verstärker

Primäre Verstärker: Angeboren, wirken ohne Lernerfahrung (z.B. Futter, Schmerz)

Sekundäre Verstärker: Neutrale Reize, die durch Paarung mit primären Verstärkern zu Verstärkern werden (z.B. Ton, der mit Futtergabe assoziiert wird; Geld; etc.)

Beispielexperiment

- Phase 1

$S^D \rightarrow R \rightarrow \text{Futter}$

$S^N \rightarrow R \rightarrow \text{Kein Futter}$

- Phase 2

$S^N \rightarrow R \rightarrow S^D$

- Ergebnis:

Tier lernt, in Anwesenheit von S^N Reaktion R auszuführen $\rightarrow S^D$ wirkt als sekundärer Verstärker

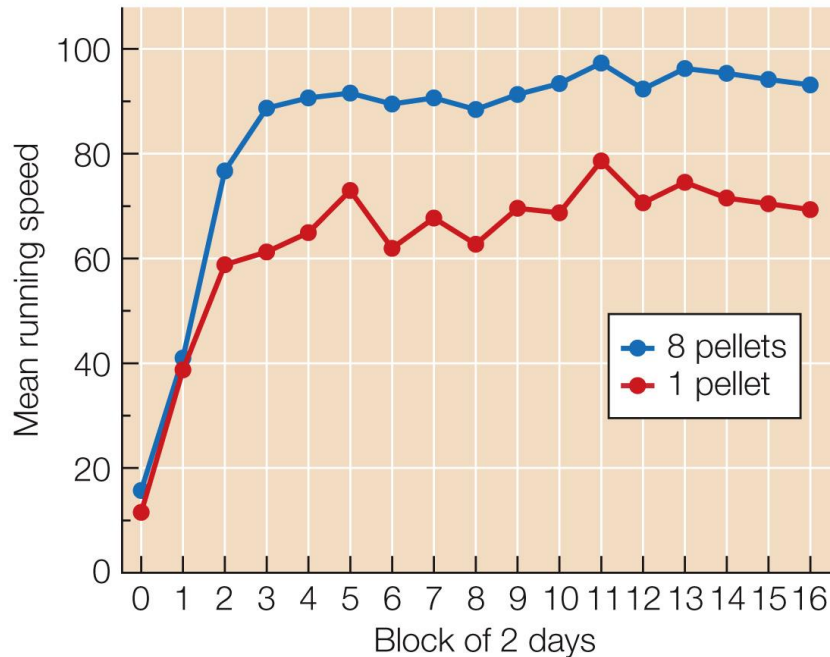
Bedeutung sekundärer Verstärker

- Chaining (z.B. bei Tierdressur)

- Überbrücken von längeren Zeitintervallen zwischen SD und primärem Verstärker

Effekte der Größe oder Intensität von Verstärkern

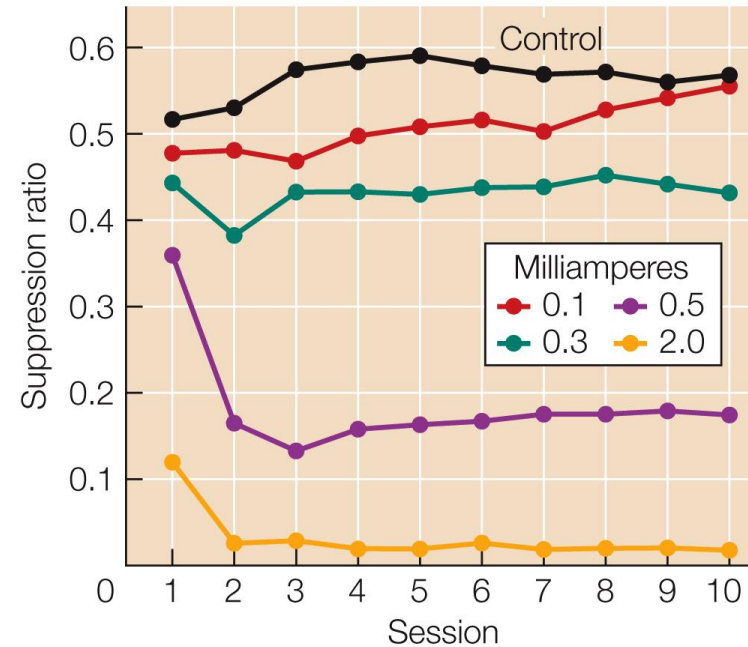
(A) Reward



LEARNING AND BEHAVIOR 2e, Figure 2.17 (Part 1)
© 2016 Sinauer Associates, Inc.

(A) Reward learning in rats in a runway (Bower, 1961)

(B) Punishment

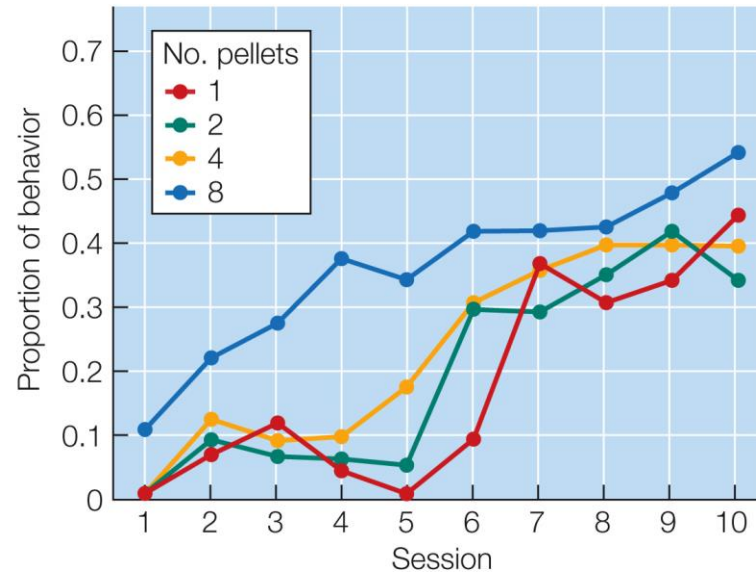


LEARNING AND BEHAVIOR 2e, Figure 2.17 (Part 2)
© 2016 Sinauer Associates, Inc.

(B) Punishment of lever pressing in rats with electric shock (Camp et al., 1967)

Größere / intensivere Konsequenzen fördern auch beim klassischen Konditionieren das Lernen

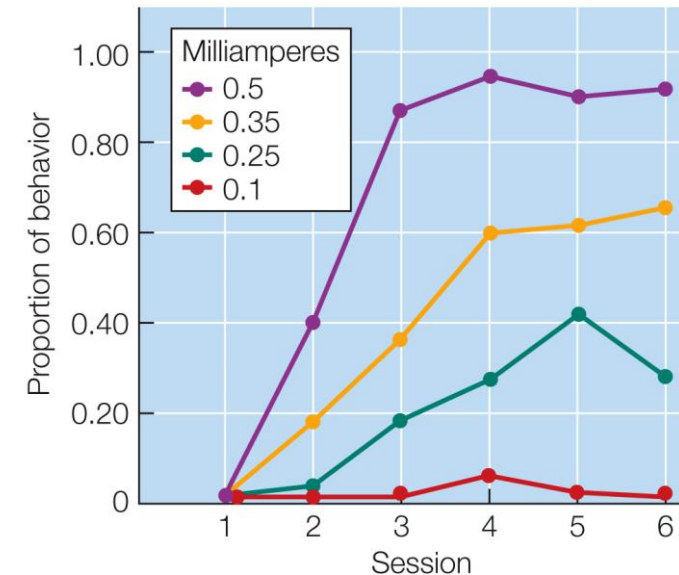
(A) Food Outcome



LEARNING AND BEHAVIOR 2e, Figure 2.18 (Part 1)
© 2016 Sinauer Associates, Inc.

(A) Classical conditioning with a food O. When an auditory cue signals food, rats show an excited “head jerk” response (Morris & Bouton, 2006).

(B) Shock Outcome

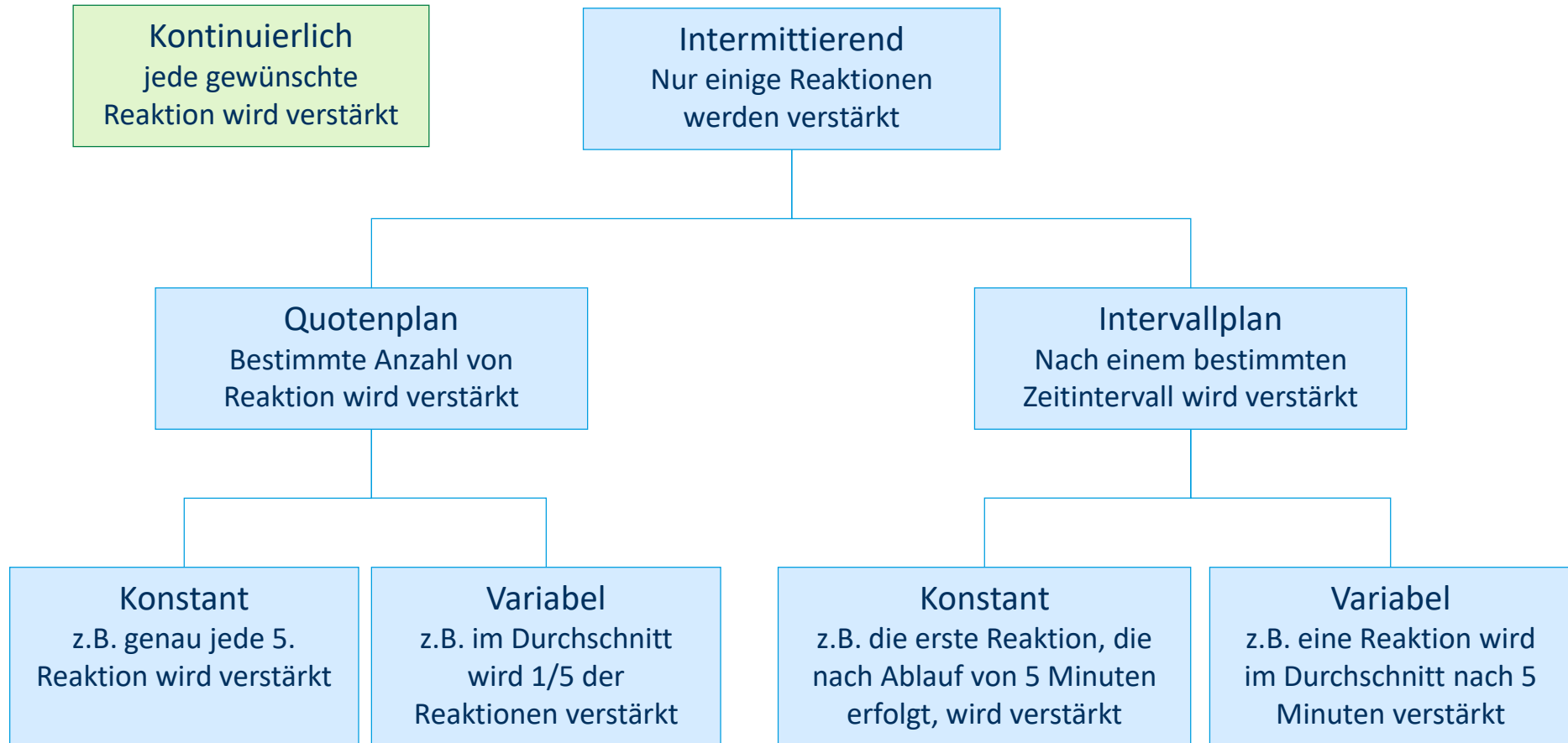


LEARNING AND BEHAVIOR 2e, Figure 2.18 (Part 2)
© 2016 Sinauer Associates, Inc.

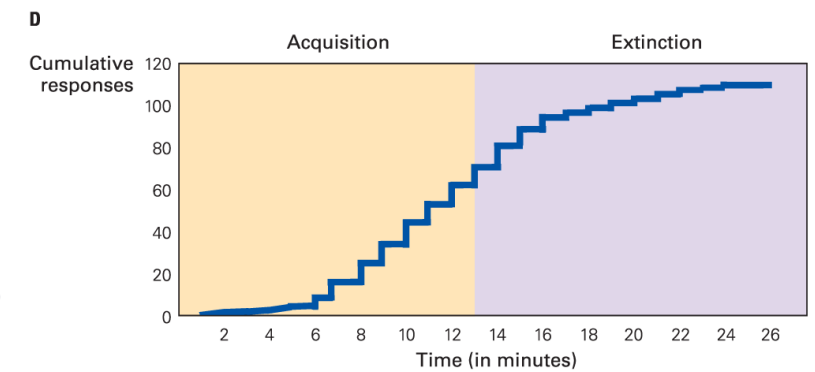
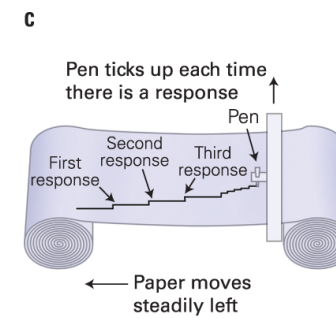
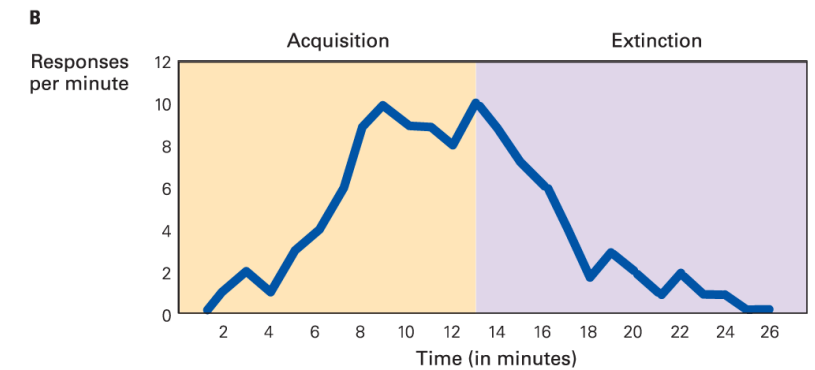
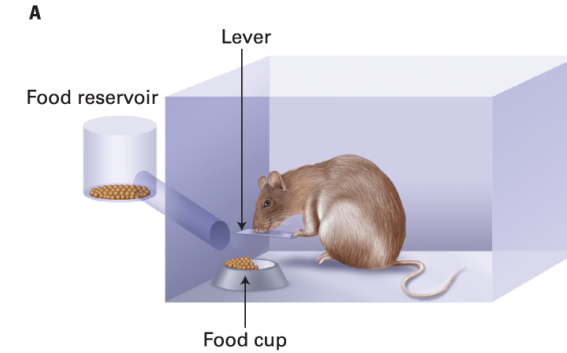
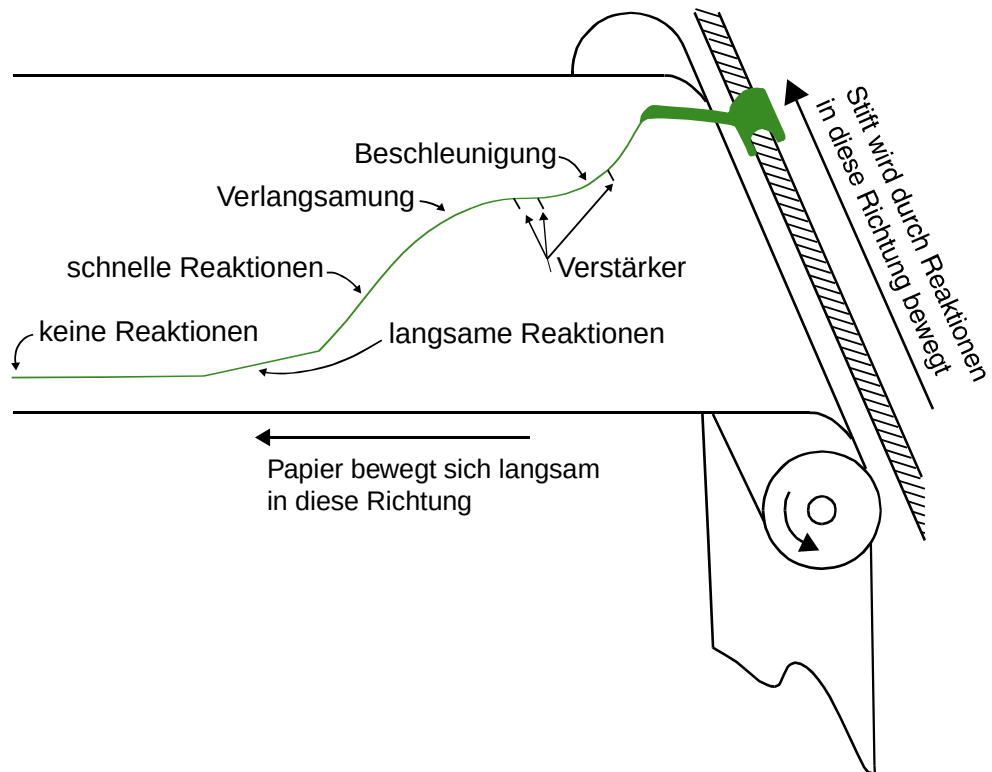
(B) Classical conditioning with an electric footshock O. Rats freeze in the presence of a cue signaling an electric shock (Holland, 1979)

Verstärkungspläne

Verstärkungspläne legen fest, wie häufig bzw. nach welchen Zeitintervallen Reaktionen zu Verstärkern führen

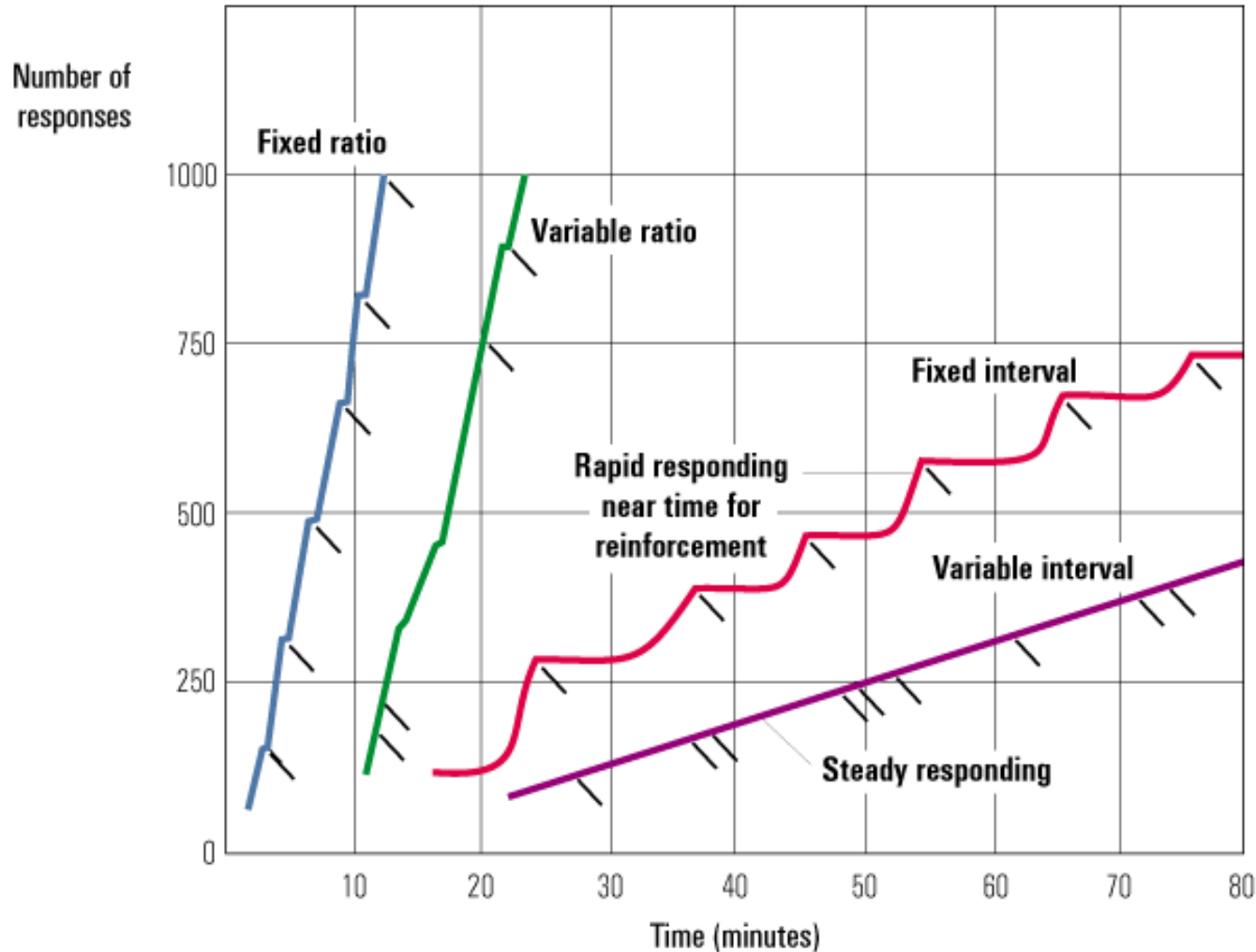


Kumulative Verhaltensaufzeichnung



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

Auswirkungen von Verstärkungsplänen



FR: konstante hohe Reaktionsrate bis zur Verstärkergabe; kurze Pause nach jedem Verstärker

FI: nach Verstärker zunächst keine Reaktionen; Zunahme der Reaktionsrate gegen Ende des Intervalls

VR: Konstante hohe Reaktionsrate ohne Nachverstärkungspause

VI: Konstante Reaktionsrate

Auswirkungen von Verstärkungsplänen auf die Löschung

Kontinuierliche Verstärkung:

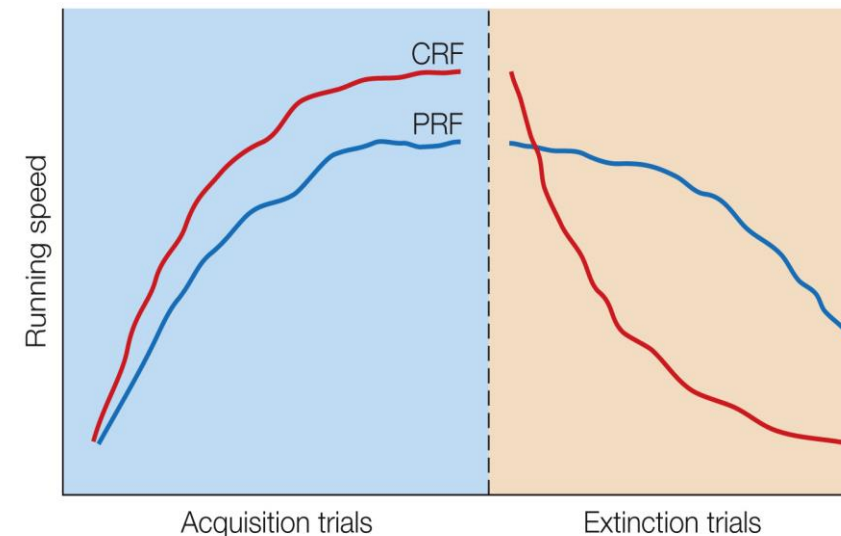
- Verhalten wird schnell erworben, wird aber auch schnell wieder gelöscht, wenn Verstärker ausbleibt

Intermittierende Verstärkung

- Höhere Lösungsresistenz
- Bsp.: Bettelnder Hund am Tisch, der bei jedem 10. Versuch ein Stück Wurst bekommt
- Bsp.: Getränkeautomat vs. Spielautomat

Experiment

- CRF: Kontinuierlich verstärkte Gruppe erhält Belohnung in 100% der Durchgänge
- PRF: intermittierend (partiell) verstärkte Gruppe erhält Belohnung in 50% der Durchgänge
- Während Extinktion zeigt PRF-Gruppe höhere Persistenz = höhere Lösungsresistenz



LEARNING AND BEHAVIOR 2e, Figure 9.13
© 2016 Sinauer Associates, Inc.

Erklärungen für den Effekt intermittierender Verstärkung (partial reinforcement effect)

Generalisierungserklärung

- *Kontinuierliche Verstärkung:*
 - Lebewesen hat nie Durchgänge ohne Verstärkung erlebt → Lern- und Löschphase sind sehr verschieden
- *Intermittierende Verstärkung:*
 - In Lernphase gibt es viele Durchgänge ohne Verstärker → Lern- und Löschphase sind ähnlich

Kognitive Erklärung

Kontinuierliche Verstärkung

- Lebewesen lernt, dass jede Reaktion zu einer Verstärkung führt und Ausbleiben des Verstärkers führt zu großer Verletzung der Erwartung → geringe Persistenz

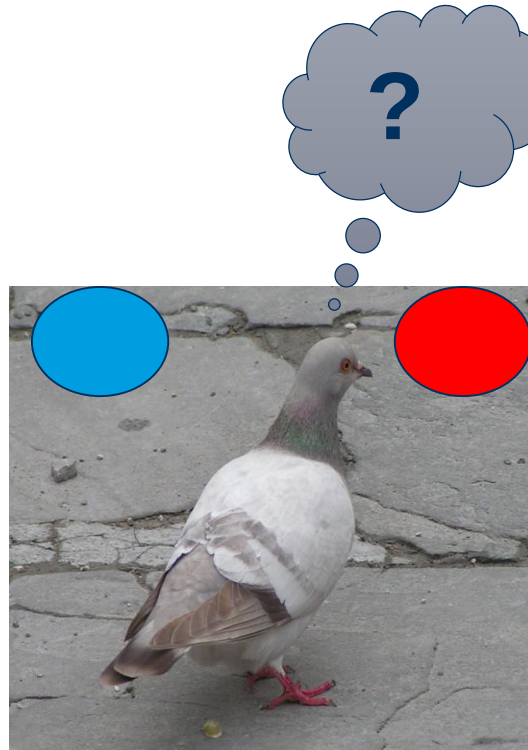
Intermittierende Verstärkung

- Lebewesen lernt, dass viele Reaktionen zur Erlangung des Verstärkers notwendig sein können und bildet entsprechende Erwartung aus → hohe Persistenz

Wahlverhalten: Zwei simultane Verstärkungspläne

VI 60 Plan

Verstärkt wird die erste Reaktion, die nach einem Intervall von durchschnittlich 60 Sekunden gezeigt wird



VI 120 Plan

Verstärkt wird die erste Reaktion, die nach einem Intervall von durchschnittlich 120 Sekunden gezeigt wird

Herrnsteins (1961) Matching Law

Optimales Verhalten = Maximierung des erhaltenen Futters durch optimale Verteilung der Reaktionen auf die beiden Alternativen

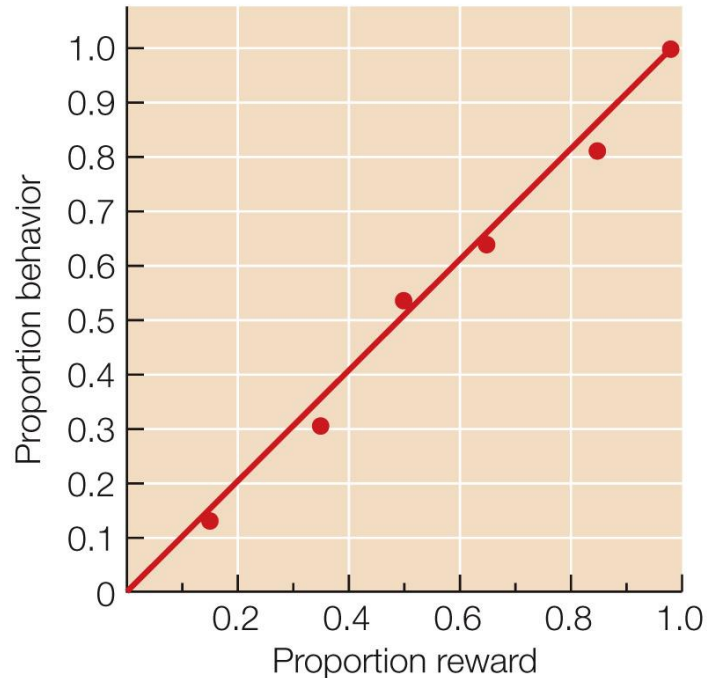


$$\frac{\text{Reaktionsrate A}}{\text{Reaktionsrate B}} = \frac{\text{Verstärkerrate A}}{\text{Verstärkerrate B}}$$

Herrnsteins (1961) Matching Law

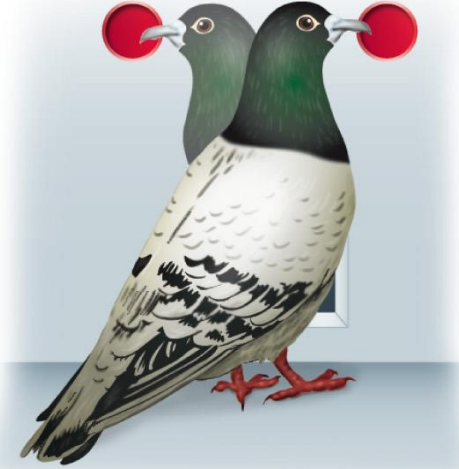
- Die Proportion einer Reaktionsalternative entspricht der Proportion der erhaltenen Belohnungen (Herrnstein, 1971)
- Die relative Reaktionsrate entspricht der relativen Rate von Verstärkungen

(A)



LEARNING AND BEHAVIOR 2e, Figure 7.9 (Part 1)
© 2016 Sinauer Associates, Inc.

Schedule 1 Schedule 2



$$\frac{R1}{R1+R2} = \frac{V1}{V1+V2}$$

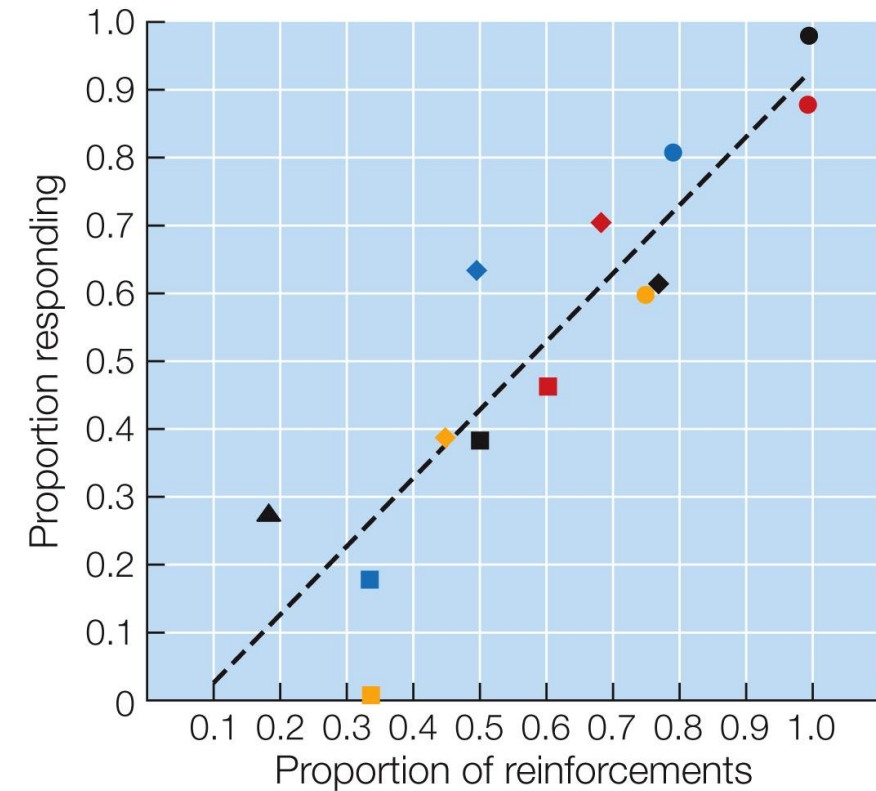
$$\frac{R1}{R2} = \frac{V1}{V2}$$

R = Anzahl Reaktionen
V = Anzahl Verstärker

Matching law bei Menschen

Conger & Killeen (1974)

- College Studenten unterhielten sich mit zwei Versuchsleitern
- Äußerungen der Studenten wurden durch positive Rückmeldungen verstärkt (z.B. “Das ist ein guter Punkt”)
- Verstärkung folgte zwei verschiedenen VI-Plänen
- A.V.: Anteil von Zeit, den die Studenten sich mit dem einen und dem anderen Versuchsleiter unterhielten



LEARNING AND BEHAVIOR 2e, Figure 7.10
© 2016 Sinauer Associates, Inc.

Premack-Prinzip: Verhalten als Verstärker

David Premack (1959, 1961)



Prinzip:

- Verstärkung findet statt, wenn ein instrumentelles Verhalten den Zugang zu einem stärker präferierten Verhalten ermöglicht

Experiment

- Ratten erhielten freien Zugang zu Trinkwasser und einem Hamsterrad
- Ratten verbrachten ca. 5 x soviel Zeit mit Rennen als mit Trinken
- Danach erhielten Ratten nur dann Zugang zum Laufrad, wenn sie bestimmte Wassermenge getrunken hatten
- Ratten tranken doppelt so lange wie zu Beginn
 - Ratten lernten zu trinken, um Zugang zum Laufrad zu erhalten
 - Verhalten (Rennen) wurde zum Verstärker für anderes Verhalten (Trinken)

Analoge Effekte bei Menschen

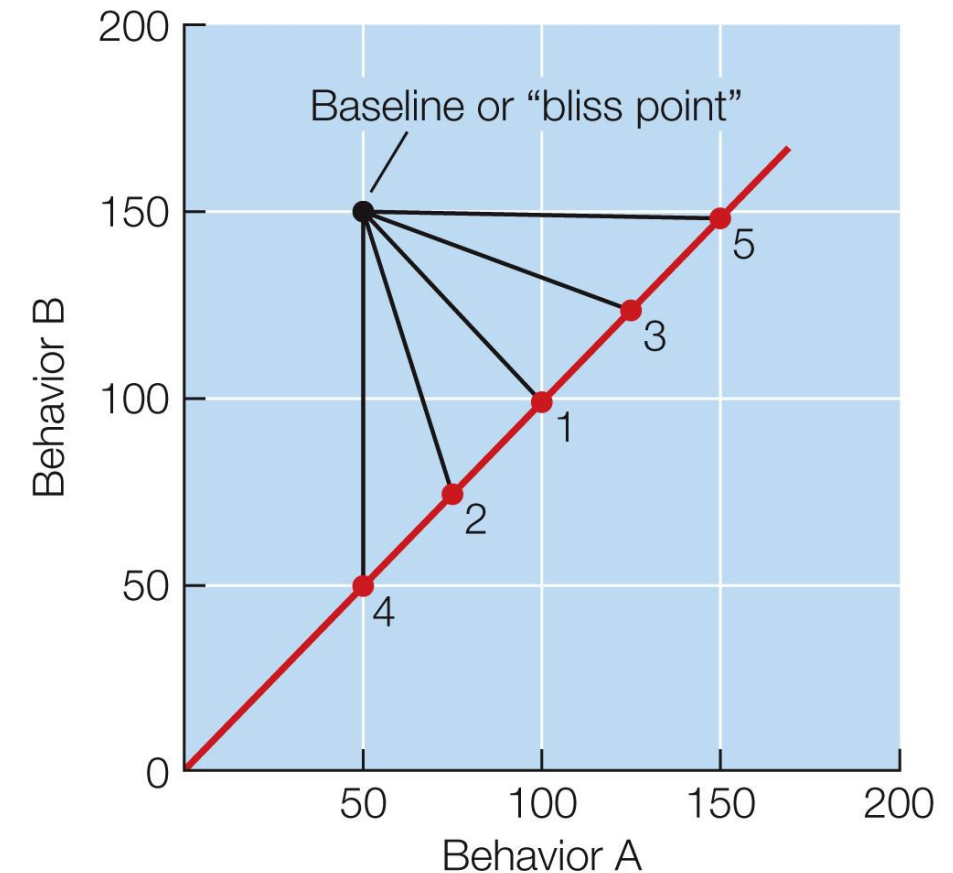
- Kinder, die größere Präferenz für Flipper spielen als für Süßigkeiten essen zeigten
- Wurde Zugang zum Flipper von Menge konsumierter Süßigkeiten abhängig gemacht, aßen die Kinder mehr Süßigkeiten

Problem

- Wie kann man relative Präferenzen für unterschiedliches Verhalten messen?

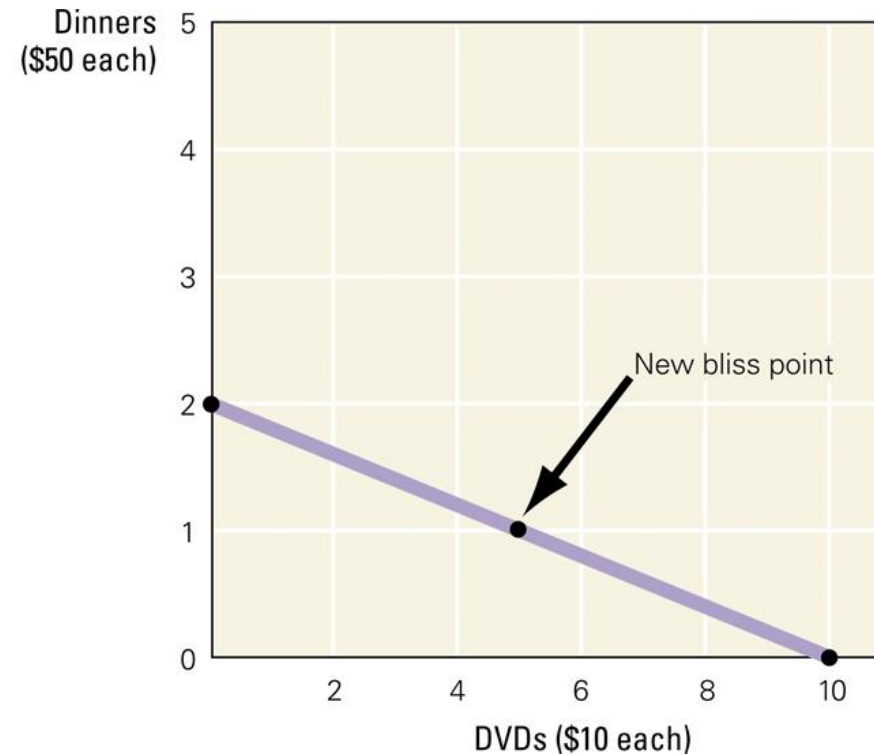
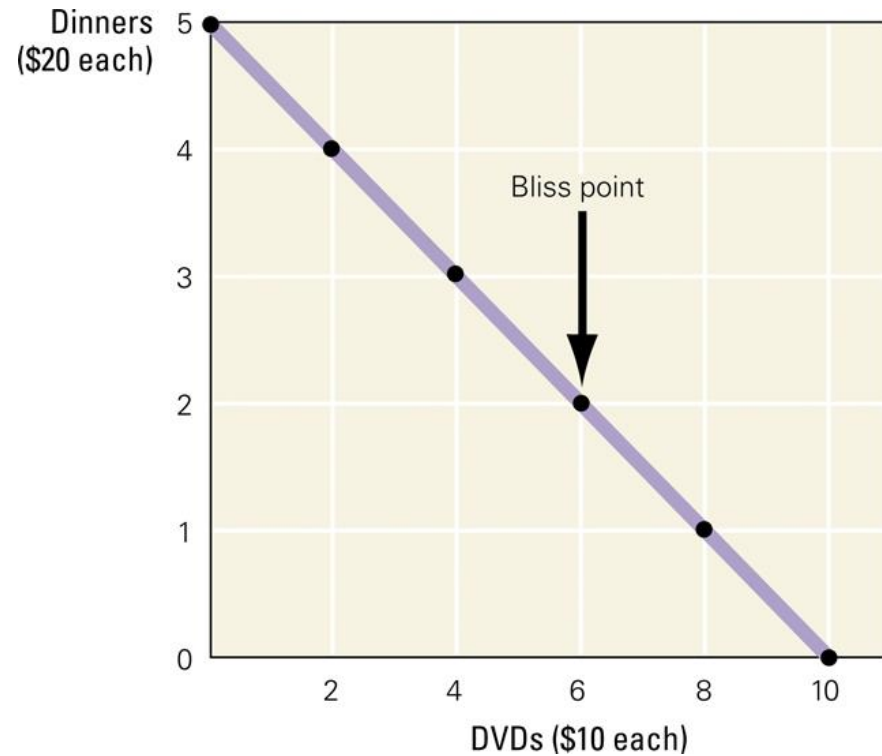
Verhaltensökonomie

- Lebewesen sind bestrebt, Zeit und Ressourcen so auf alternative Verhaltensweisen zu verteilen, dass der subjektive Nutzen maximiert wird
- “Bliss point”: optimales Verhältnis alternativer Verhaltensweisen bei freier Entscheidung
- Wenn eine bestimmte Anzahl von A-Reaktionen notwendig für eine B-Reaktion sind, liegen die möglichen relativen Reaktionshäufigkeiten auf einer Linie
- Lebewesen führen die Anzahl von Reaktionen aus, die die Distanz zum Bliss Point minimieren
- (Allison, 1989)



LEARNING AND BEHAVIOR 2e, Figure 7.19
© 2016 Sinauer Associates, Inc.

Verhaltensökonomie

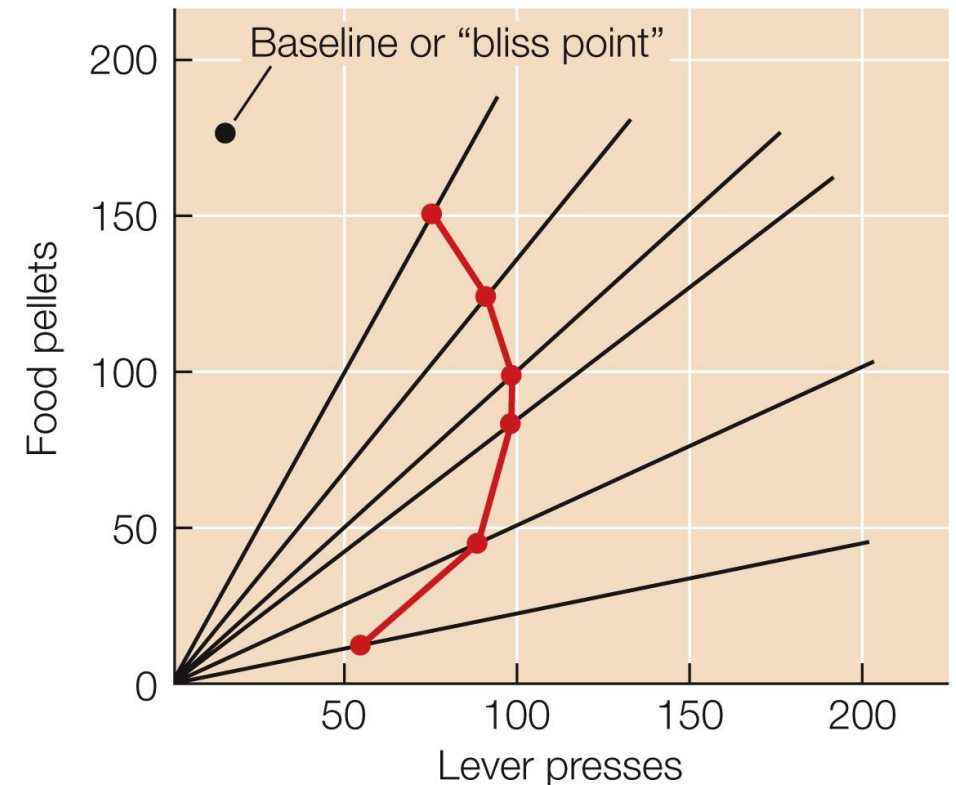


Verhaltensökonomie

Die Zahl von Reaktionen, die die Distanz zum Bliss Point minimiert, hängt vom Verstärkungsplan ab

Jede Linie repräsentiert einen Quotenplan, in dem eine unterschiedliche Zahl von A-Reaktionen notwendig ist, um eine B-Reaktion ausführen zu können

Die roten Punkte sind die Punkte für jeden Verstärkungsplan, die am nächsten zum Bliss Point liegen (nach Allison, 1989)



LEARNING AND BEHAVIOR 2e, Figure 7.20
© 2016 Sinauer Associates, Inc.

Überblick

Thorndikes Gesetz des Effekts

Skinner's Forschung zum operanten Konditionieren

Was wird beim instrumentellen Konditionieren gelernt?

Shaping und Verhaltenssequenzen

Verstärker und Verstärkungspläne

Rolle von Kontiguität und Kontingen

Abbau unerwünschter Verhaltensweisen und Bestrafung

Flucht- und Vermeidungstraining

Stimuluskontrolle: Generalisierung und Diskrimination

Artspezifische Reaktionsprogramme und angeborene Lernbereitschaften

Zeitliche Kontiguität: Effekt des Zeitabstands zwischen Reaktion und Konsequenz

Dickinson et al. (1992):
Kürzere Abstände zwischen
Verhalten (Hebel drücken) und
Verstärkung bei Ratten →
effektivere Konditionierung

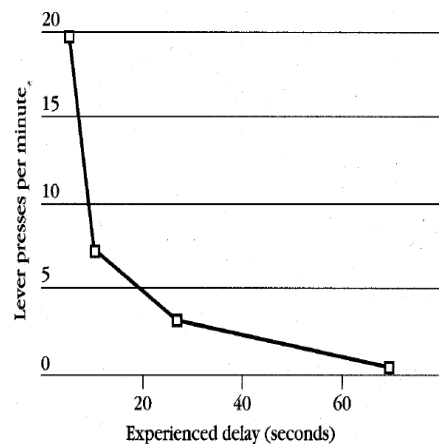
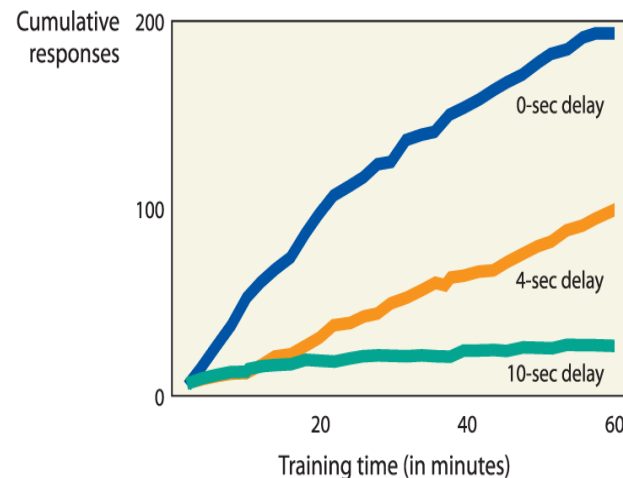


Figure 5.11 Effects of delay of reinforcement on acquisition of lever pressing in rats. (From "Free-Operant Acquisition with Delayed Reinforcement," by A. Dickinson, A. Watt, and W. J. H. Griffiths, 1992, *The Quarterly Journal of Experimental Psychology*, 45B, pp. 241–258. Copyright © 1992 by The Experimental Psychology Society. Reprinted by permission.)

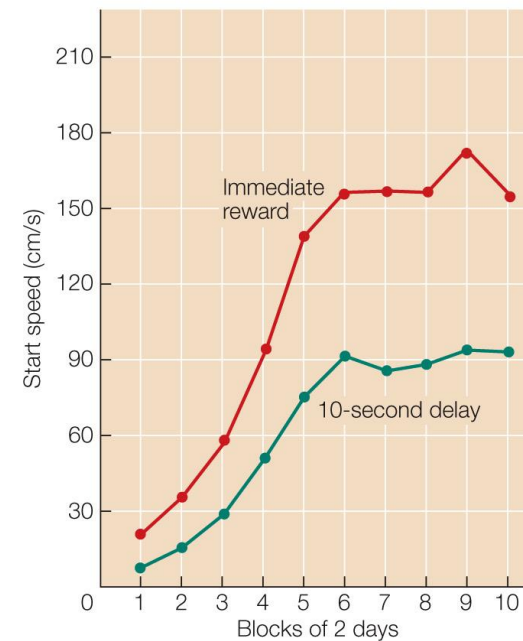
Schlinger & Blakey (1994):
Ratten erhielten nach
Drücken eines Hebels nach 0,
4 oder 10 Sekunden Futter als
Belohnung



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

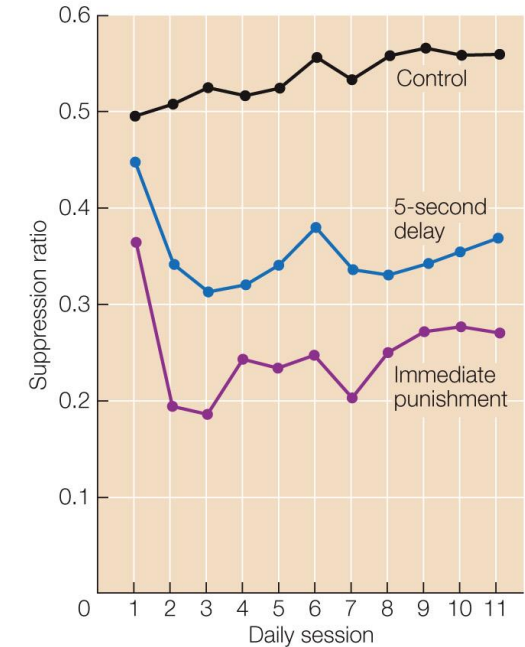
Belohnung und Bestrafung sind am effektivsten
wenn sie unmittelbar auf die Reaktion folgen
(Capaldi, 1978; Church, 1969)

(A) Reward



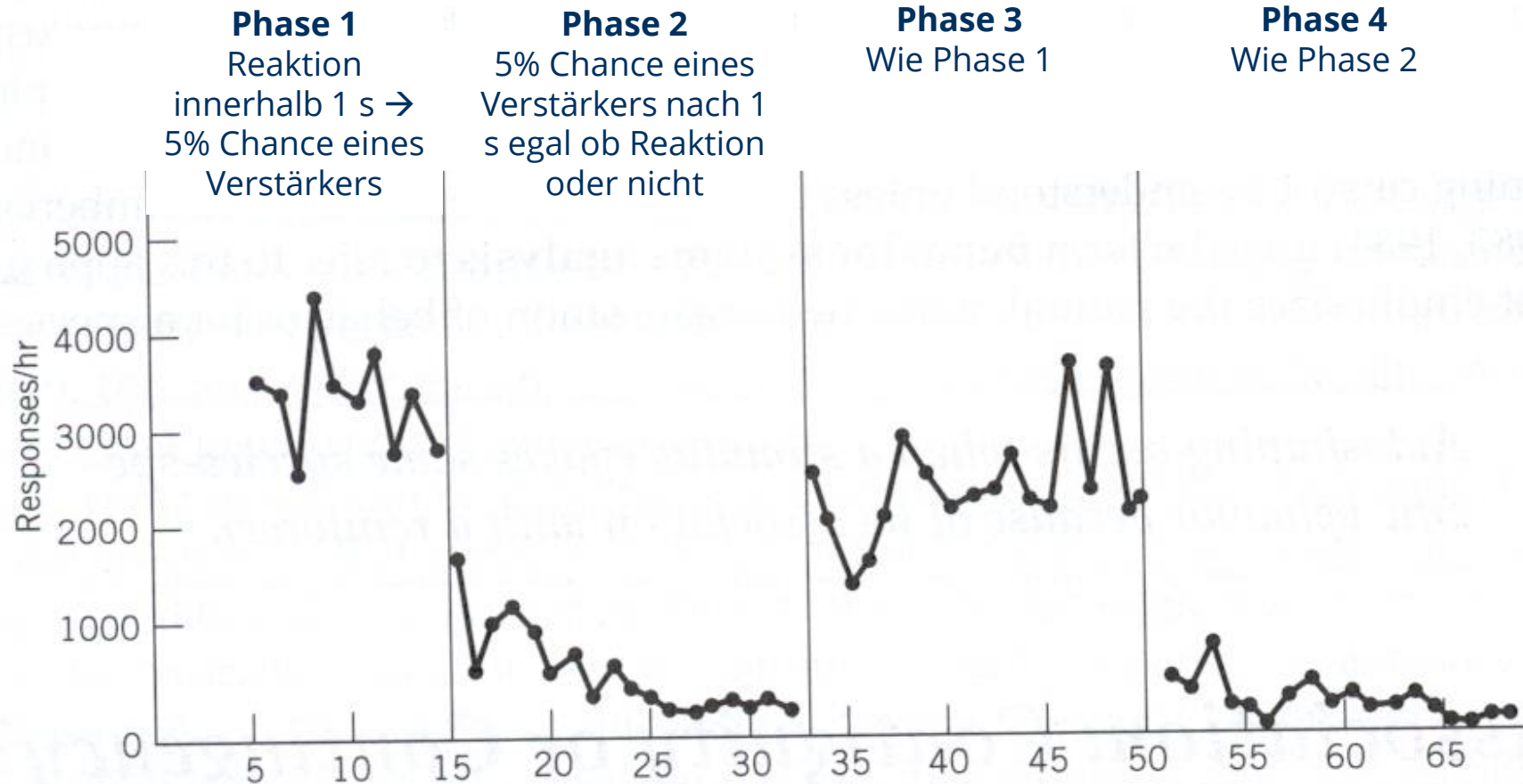
LEARNING AND BEHAVIOR 2e, Figure 2.15
© 2016 Sinauer Associates, Inc.

(B) Punishment



Bedeutung der Kontingenz

Hammon et al. (1980): Ratten wurden trainiert, Hebel zu drücken, um Futter zu erhalten



Instrumentelles Konditionieren als Inferenz von Reaktions-Konsequenz-Kontingenzen

Wasserman (1990)

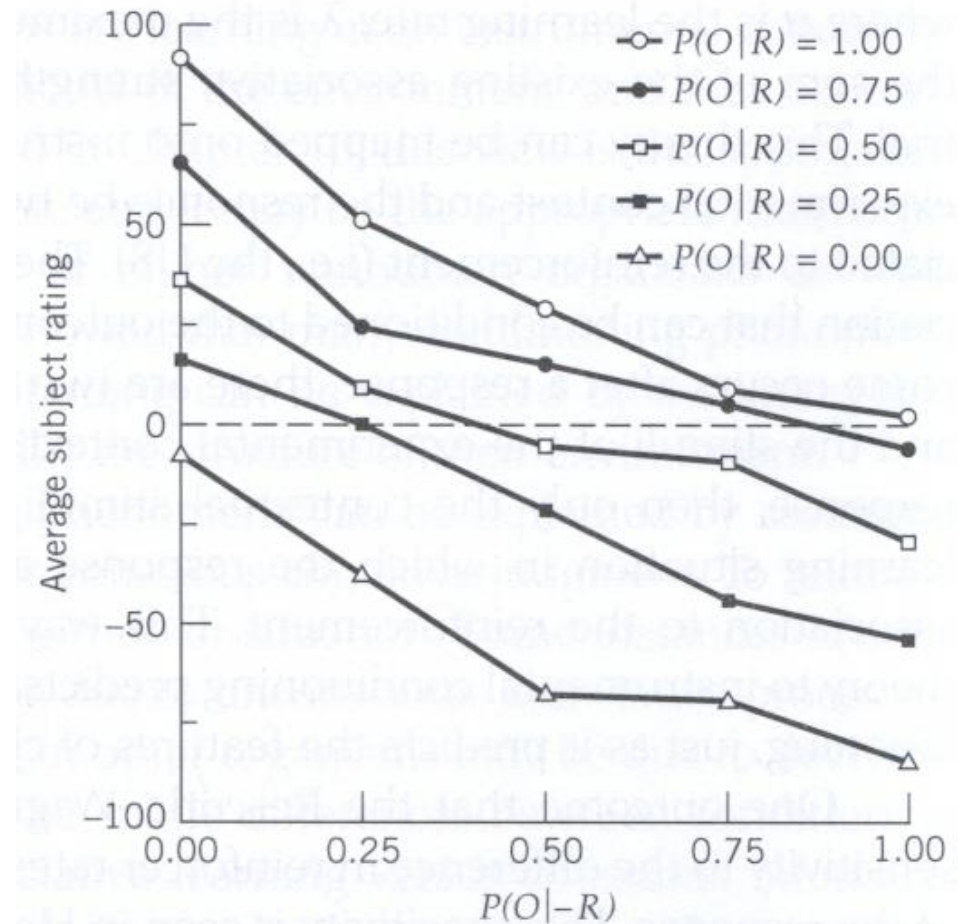
Vpn konnten Taste drücken, was in einem Teil der Durchgänge vom Aufleuchten eines Lichts gefolgt wurde

UV1: $p(\text{Licht} | \text{Taste})$: 0; 0.25; 0.5; 0.75; 1.0

UV2: $p(\text{Licht} | \text{keine Taste})$: 0; 0.25; 0.5; 0.75; 1.0

(vgl. Rescorla-Wagner-Theorie!)

(O = Outcome; R = Response)



Schlussfolgerungen

Organismen lernen kausale Regularitäten, die es ermöglichen, Effekte des eigenen Verhalten zu antizipieren

Dies ermöglicht es Organismen, sich adaptiv („rational“) zu verhalten = Verhalten zu selektieren, dass positive Konsequenzen hat und negative Konsequenzen vermeidet

Adaptives Verhalten muss nicht auf bewusster Überlegung oder Einsicht in die relevanten Kontingenzen beruhen

Relativ einfache assoziative Mechanismen können Verhalten erzeugen, das rational und zielgerichtet erscheint

Nicht-kontingente Verstärkung und „abergläubisches“ Verhalten

Skinner (1948):

- Tauben erhielten alle 15 s Futterkörner unabhängig vom Verhalten
- Einige Tiere entwickelten bizarre „Rituale“ (z.B. mehrmals gegen den Uhrzeigersinn drehen) und verhielten sich, als ob sie „glaubten“, durch ihr Verhalten die Verstärkergabe kontrollieren zu können

Skinner's Erklärung:

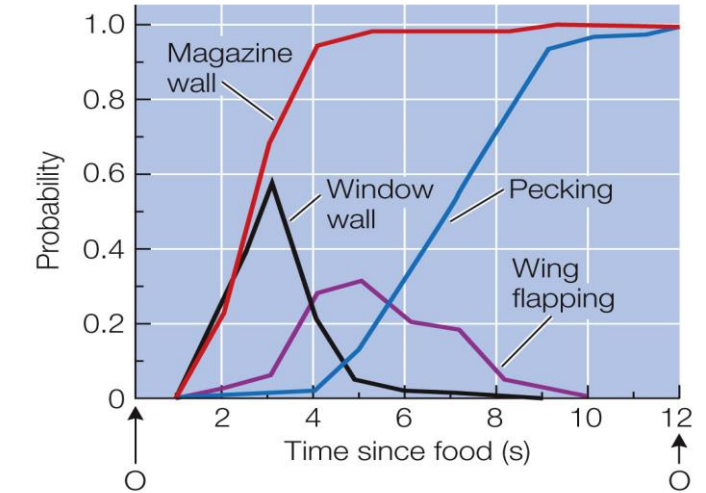
- Verstärker erhöht die Wahrscheinlichkeit jedes Verhaltens, dass zufällig unmittelbar vor der Verstärkergabe gezeigt wurde
- → höhere Wahrscheinlichkeit, dass dieses Verhalten (zufälligerweise) erneut vor der Verstärkergabe auftritt → weitere Verstärkung
- → Im Lauf der Zeit können sich komplexe Verhaltensmuster bilden (analog zu natürlicher Selektion)

Abergläubisches Verhalten: Folgestudien und Reinterpretation

Staddon & Simmelhag (1971)

- Tauben erhielten alle 12 s Futter
- Kurz nach der Futtergabe (O = outcome), bewegten sich Tauben in Richtung des Fensters der Skinnerbox und schlugen mit den Flügeln, aber diese Reaktionen nahmen mit der Zeit ab (*interim behaviors*).
- Gegen Ende des Ende, wenn wieder Futter erwartet wurde, bewegten die Tauben sich in Richtung des Futterspenders und zeigten Pickverhalten (*terminal behaviors*)

(A) Pigeon

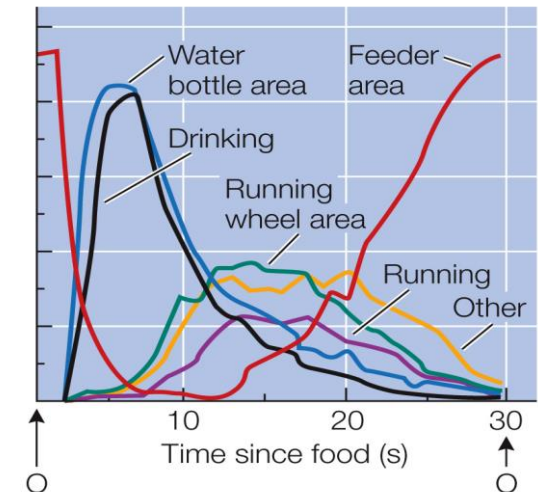


LEARNING AND BEHAVIOR 2e, Figure 10.13 (Part 1)
© 2016 Sinauer Associates, Inc.

Staddon & Ayres (1975)

- Ratten erhielten alle 30 s eine kleine Menge Futter
- Tiere zeigten analoges Muster von Verhaltensweise kurz nach der Futtergabe (*interim behaviors*) und Annäherung an den Futterspender gegen Ende des Zeitintervalls und vor der nächsten Futtergabe (*terminal behavior*)

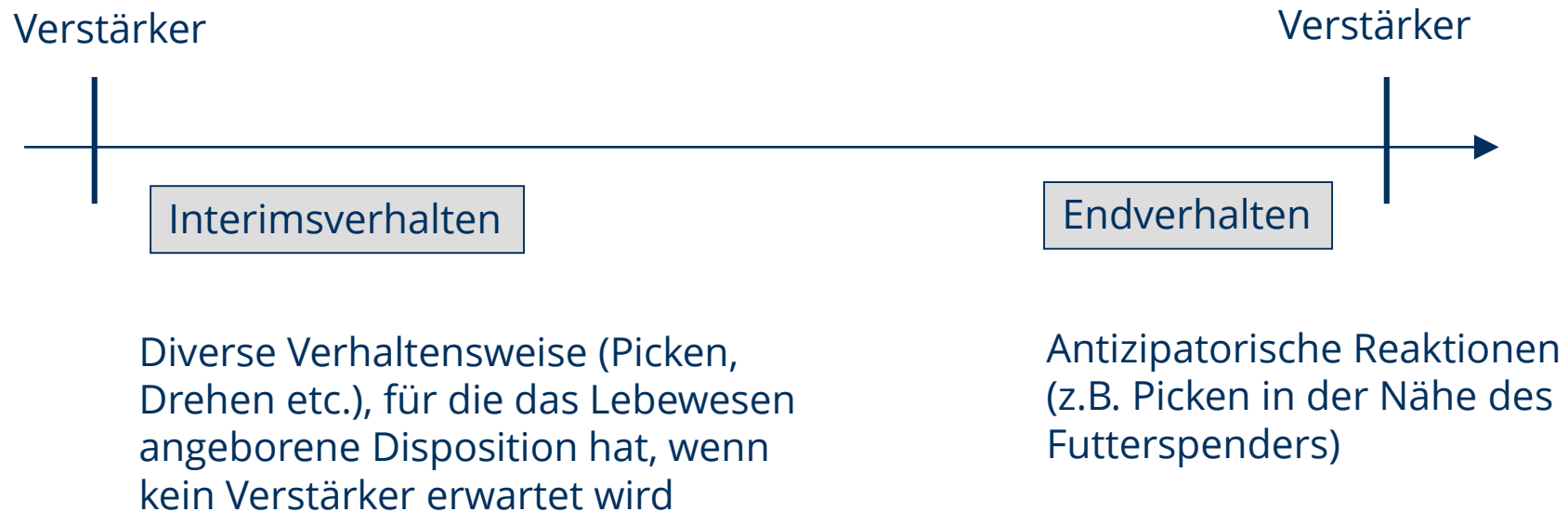
(B) Rat



LEARNING AND BEHAVIOR 2e, Figure 10.13 (Part 2)
© 2016 Sinauer Associates, Inc.

Abergläubisches Verhalten: Folgestudien und Reinterpretation

Staddon & Simmelhag (1971)



- Verhalten ist weniger „abergläubisch“ als ursprünglich angenommen, sondern scheint Anpassung an zeitliche Struktur des Verstärkungsplans zu sein
- Kontiguität allein reicht meist nicht aus, um Verhalten zu verstärken, sondern Organismen lernen Kontingenzen zwischen Verhalten und Verstärkern

Überblick

Thorndikes Gesetz des Effekts

Skinner's Forschung zum operanten Konditionieren

Was wird beim instrumentellen Konditionieren gelernt?

Shaping und Verhaltenssequenzen

Verstärker und Verstärkungspläne

Rolle von Kontiguität und Kontingenz

Abbau unerwünschter Verhaltensweisen und Bestrafung

Flucht- und Vermeidungstraining

Stimuluskontrolle: Generalisierung und Diskrimination

Artspezifische Reaktionsprogramme und angeborene Lernbereitschaften

Abbau unerwünschter Verhaltensweisen

Löschung: Entzug der das Verhalten aufrechterhaltenden Verstärker

Bestrafung des unerwünschten Verhaltens

Alternativverhalten aufbauen, das das unerwünschte Verhalten ersetzen soll

Löschung

Schritt 1: Bedingungsanalyse

- Die das unerwünschte Verhalten aufrechterhaltenden Bedingungen (Verstärker, diskriminative Hinweisreize) identifizieren

Schritt 2: Löschung

- Entzug eben dieser Verstärker

Beispiel:

- Ein Kind erhält durch ständiges Stören mehr Aufmerksamkeit
- Entzug der Aufmerksamkeit → Löschung des Verhaltens

Wichtig: Löschung ist besonders effektiv, wenn gleichzeitig ein Alternativverhalten positiv verstärkt wird

- Bsp.: Das Kind erhält Aufmerksamkeit für erwünschtes Verhalten

Bestrafung und Verhaltensalternativen

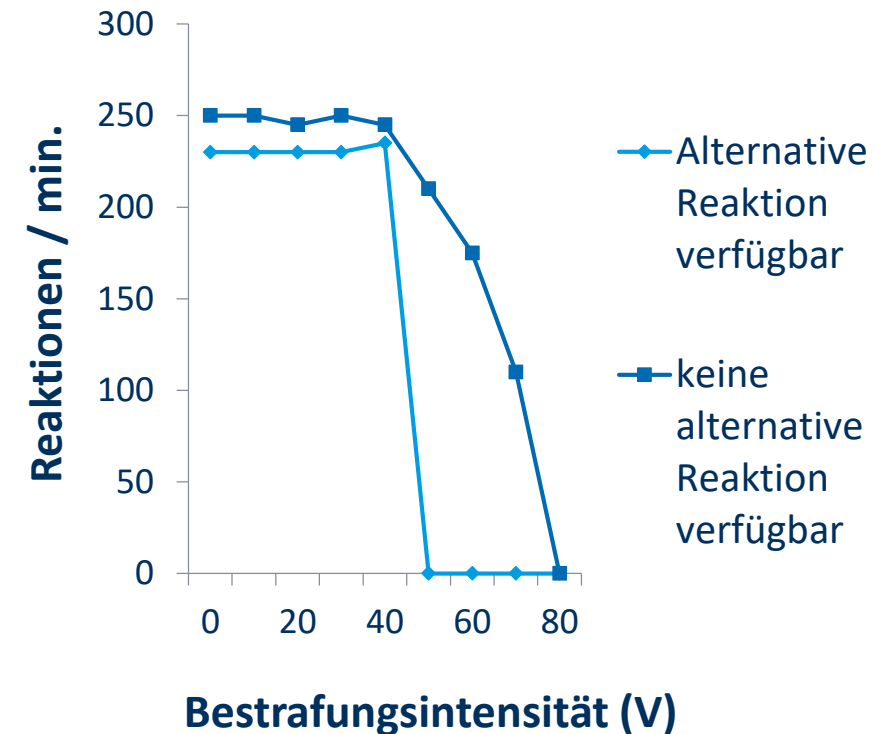
Bestrafung kann sehr wirksam sein:

- Eine einzige Lernerfahrung kann Verhalten eliminieren (z.B. heiße Herdplatte)

Unterdrückung eines Verhaltens durch Bestrafung ist effektiver, wenn alternative Verhaltensmöglichkeiten bestehen (insbesondere, wenn diese belohnt werden)

Anzrin & Holz (1966):

- Phase 1: Tauben lernten, auf einen Schalter zu picken, um Futter zu erhalten
- Phase 2: Auf Picken folgte Schock
- Gruppe A: Taube konnte auf anderen Schalter picken
- Gruppe B: Nur ein Schalter
- Ergebnis: Tauben in Gruppe A zeigten stärkere Unterdrückung des Pickens auf den ursprünglichen Schalter



Mögliche negative Nebeneffekte von Bestrafung

- Bestrafung kann unter bestimmten Bedingungen zur Unterdrückung von jeglichem Verhalten führen
- Bestrafung unterdrückt oft lediglich Verhalten (Performanz); aber hohe Wahrscheinlichkeit, dass unerwünschtes Verhalten wieder gezeigt wird, sobald keine Bestrafung mehr erwartet wird
- Diskriminative Hinweisreize, die Bestrafung signalisieren, können zu Täuschungsverhalten führen (Bsp.: Radarfalle)
- Beziehung zwischen Bestrafendem und Bestraftem leidet (kann Ärger und Aggression erzeugen)
- Falsche Vorbildfunktion: Kinder lernen, dass Bestrafung (Aggression) ein angemessenes Mittel zur Konfliktlösung ist
- Wenn Bestrafung unvermeidlich ist, sollte stets erwünschtes Alternativverhalten ermöglicht und verstärkt werden