

Lernen als Bildung von Reiz-Reaktions-Verbindungen

Joachim Hoffmann

2.1 Lernen bei Tieren als Modell für menschliches Lernen – 11

2.1.1 Der Behaviorismus: Lernen als Bildung von Reiz-Reaktions-Verbindungen bei Tier und Mensch – 11

2.2 Klassische Konditionierung – 12

2.2.1 Pavlov'scher bedingter Reflex – 12

2.2.2 Klassische Erklärung bedingter Reflexe – 13

2.2.3 Ausgewählte Eigenschaften bedingter Reflexe – 13

2.3 Instrumentelle Konditionierung – 16

2.3.1 Versuche von Thorndike – 16

2.3.2 Skinner-Box – 16

2.3.3 Effektgesetz („law of effect“) – 16

2.3.4 Ausgewählte Eigenschaften instrumentellen Lernens – 17

2.4 Diskriminationslernen – 19

2.4.1 Unterscheidung von verhaltensrelevanten und verhaltensirrelevanten Reizbedingungen – 19

2.4.2 Positives und negatives Patterning – 20

2.4.3 Bildung von Reizkategorien – 21

2.5 Die Selektivität der Bildung von S-R-Verbindungen – 22

2.5.1 Latente Hemmung der Ausbildung eines bedingten Reflexes – 22

2.5.2 Blockierung der Ausbildung eines bedingten Reflexes – 22

2.5.3 Erlernte Hilflosigkeit: Die Blockierung des Vermeidungslernens – 23

2.5.4 Preparedness: angeborene verhaltensgebundene Aufmerksamkeit – 24

2.6	Rescorla-Wagner-Modell elementaren S-R-Lernens – 25
2.6.1	Modellbeschreibung – 26
2.6.2	Modellerklärungen – 28
2.6.3	Bewertung des RWM – 28
2.7	Fazit – 29

Lernziele

- Warum können wir aus dem Studium der Lernprozesse bei Tieren etwas über elementares Lernen beim Menschen erfahren?
- Worin unterscheiden sich Lernsituationen und Lernprozesse bei der klassischen und instrumentellen Konditionierung?
- Inwiefern verbessern klassische und instrumentelle Konditionierung die Überlebens- und Fortpflanzungschancen von Organismen?
- Wie werden verhaltensrelevante von verhaltensirrelevanten Reizbedingungen unterschieden und ggf. in Kategorien zusammengefasst?
- Von welchen äußeren und inneren Bedingungen sind klassische und instrumentelle Konditionierungen abhängig?
- Auf welchen Grundannahmen beruht das Rescorla-Wagner-Modell klassischen Konditionierens, und wo liegen seine Grenzen?

Beispiel

Menschen, die die Bombenangriffe des 2. Weltkrieges erlebt haben, bekommen teilweise noch heute Herzrasen und Schweißausbrüche beim Klang einer Sirene. Gleichermäßen spürt jemand, der einmal einen heftigen elektrischen Schlag bekommen hat, noch nach Jahren ein Kribbeln in den Fingern, wenn er elektrische Drähte berührt, selbst wenn er sicher ist, dass die Sicherungen ausgeschraubt sind. Die Lernmechanismen, die diesen und ähnlichen Phänomenen zugrunde liegen, werden im Folgenden diskutiert.

2.1 Lernen bei Tieren als Modell für menschliches Lernen

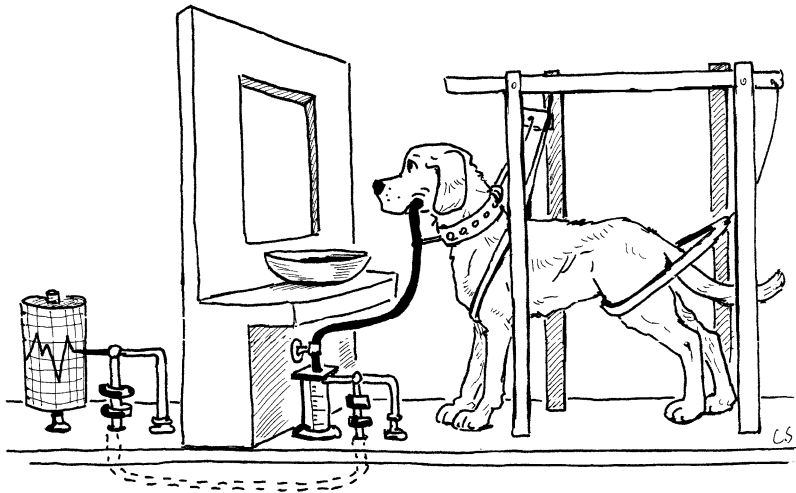
Der Schimpanse ist unter den Tieren der genetisch naheste Verwandte des Menschen. Das ist allgemein bekannt. Weniger bekannt ist, dass auch umgekehrt der Mensch der naheste Verwandte des Schimpansen ist. Nicht etwa Gorilla und Orang Utan sind seine Brüder und Schwestern, sondern wir Menschen, und die anderen Menschenaffen sind unsere gemeinsamen Cousins und Cousinen. Wir Menschen sind also lediglich besonders begabte Affen. Wie alle Tiere tragen wir das Erbe von Millionen Jahren Evolution

in uns. In der Evolution wird bewahrt, was „fit“ macht, wie es Charles Darwin genannt hat. „**Fitness**“ bezieht sich in diesem Zusammenhang auf alles, was zu Fortpflanzungsvorteilen führt. Fortpflanzungsvorteile werden u. a. durch anpassungsfähiges Verhalten gewonnen. Dies gilt vor allem, wenn sich die Lebensbedingungen ändern. Tiere, die dann ihr Jagdverhalten, ihren Schutz vor Angreifern oder die Aufzucht ihrer Nachkommen schnell und effektiv den neuen Bedingungen anpassen, haben bessere Chancen, Nachkommen zu zeugen und aufzuziehen, als weniger anpassungsfähige Tiere. Daraus ergibt sich: Lernmechanismen, die eine effektive Anpassung des Verhaltens an Umgebungsbedingungen ermöglichen, werden bevorzugt an die nächste Generation weitergegeben und bleiben damit in der Evolution erhalten. Wir dürfen deshalb davon ausgehen, dass sich auch die **Lernmechanismen**, mit denen wir geboren werden, schon lange vor uns herausgebildet haben.

2.1.1 Der Behaviorismus: Lernen als Bildung von Reiz-Reaktions-Verbindungen bei Tier und Mensch

Wie bereits in ► [Kap. 1](#) erwähnt war der **Behaviorismus** ein die Lernpsychologie beherrschendes Paradigma, das die Forschung zum Lernen auf Untersuchungen zur Entstehung und Veränderung von Reiz-Reaktions-Beziehungen reduzierte. Neben die Konzentration auf Reiz-Reaktions-Beziehungen trat die Überzeugung, dass die Gesetzmäßigkeiten ihrer Herausbildung universal sind und deshalb an Tieren gleichermaßen untersucht werden können wie an Menschen. Mehr noch, die Behavioristen waren der Überzeugung, dass die Gesetze der Bildung von Reiz-Reaktions-Verbindungen bei Tieren den Schlüssel für das Verständnis der geistigen Fähigkeiten des Menschen liefern. So schreibt etwa Edward L. Thorndike, einer der Pioniere der behavioristischen Bewegung:

- » The main purpose of the study of the animal mind is to learn the development of mental life down through the phylum, to trace in particular the origin of human faculty ...



■ Abb. 2.1 Versuchsanordnung von Pavlov zur Untersuchung des Speichelreflexes. © Claudia Styrsky

For the origin and development of human faculty we must look to [these] processes of association in lower animals (Thorndike 1898, S. 1).

Die Auffassung, dass sich Verhaltensanpassungen auf Assoziationen zwischen Reizen und Reaktionen zurückführen lassen, konnte sich auf zwei Beobachtungen berufen. Die eine Beobachtung wurde vom Physiologen Ivan Petrovitsch Pavlov aus St. Petersburg und die andere Beobachtung wurde von Edward Lee Thorndike aus Harvard berichtet.

2.2 Klassische Konditionierung

2.2.1 Pavlov'scher bedingter Reflex

Pavlov interessierte sich für die Physiologie der Verdauung und dabei insbesondere für die Anpassung der Speichelzusammensetzung an die aufgenommene Nahrung. Als Versuchstiere wählte er Hunde. Den Hunden wurde operativ eine Fistel gelegt, durch die der produzierte Speichel kontrolliert aufgefangen werden konnte (■ Abb. 2.1). Jede Regung der Speicheldrüsen war nun unmittelbar messbar, und der Einfluss verschiedener Bedingungen auf das Speicheln konnte gezielt untersucht werden.

Pavlov beobachtete, dass die Hunde nicht erst zu speicheln anfangen, wenn sie Futter aufnahmen, sondern bereits, wenn der Wärter den Käfig betrat. Dass Hunden (und nicht nur ihnen) auch schon vor dem Fressen das Wasser im Maul zusammenläuft, ist allgemein bekannt. Pavlov erkannte jedoch das Erstaunliche an diesem Vorgang: Als Physiologe war ihm verständlich, dass durch Reizung von Rezeptoren im Maul der Hunde Erregungen ausgelöst werden, die durch bestehende neuronale Verbindungen unmittelbar die Tätigkeit der Speicheldrüsen anregen. Das war ein fester, von keinen weiteren Bedingungen abhängiger neurophysiologischer Vorgang, ein **unbedingter Reflex**. Wie aber konnte ein visueller Reiz, wie das Erscheinen des Wärters, ebenfalls die Speicheldrüsen anregen, obwohl direkte Verbindungen zwischen den visuellen Rezeptoren und den Speicheldrüsen nicht anzunehmen waren?

Um dieser Frage nachzugehen, untersuchte Pavlov die Bedingungen, unter denen ein neutraler, nicht futterbezogener Reiz, wie etwa der Ton einer Glocke, zum Auslöser des Speichelflusses wird. Dies geschah insbesondere dann, wenn der Glockenton wiederholt kurz vor dem Futter dargeboten wurde: Neben die unbedingte Reaktion der Speicheldrüsen auf das Futter trat nach mehrmaligem gemeinsamen Auftreten von Glockenton und Futter die bedingte Speichelreaktion auf den Glockenton. Weiterhin konnte gezeigt werden, dass die **bedingte Reaktion**

2.2 · Klassische Konditionierung

auf den Glockenton wieder schwächer wurde und schließlich völlig ausblieb, wenn dem Tier wiederholt der Glockenton ohne Futter dargeboten wurde. Es handelt sich beim bedingten Reflex also um eine nicht dauerhafte, wieder auflösbare Verbindung zwischen einem ursprünglich neutralen Reiz und einer ursprünglich unbedingten reflektorischen Reaktion.

2.2.2 Klassische Erklärung bedingter Reflexe

Die **Abb. 2.2** zeigt schematisch die **Grundstruktur eines bedingten Reflexes**. Ausgangspunkt des Lernvorgangs ist ein unbedingter Reflex, d. h. eine feste unwillkürliche Verbindung zwischen einem unbedingten Reiz (US) und einer unbedingten Reaktion (UR). Beispiele für in Konditionierungsversuchen verwendete unbedingte Reflexe sind das Speicheln auf die Gabe von Futter (Speichelreflex), das Schließen der Augen als Reaktion auf einen Luftstoß (Lidschlussreflex), die Verengung der Pupille auf Lichteinfall (Pupillenreflex) oder Fluchtverhalten bzw. Verhaltensstarre als Reaktion auf eine elektrische Reizung (Vermeidungsreflex). In einer Lernphase wird der unbedingte Reiz (US), der die unbedingte Reaktion (UR) auslöst, mit einem zu konditionierenden Reiz (CS) dargeboten. Als zu konditionierende Reize werden akustische oder visuelle Reize verwendet, die die Aufmerksamkeit der Tiere auf sich ziehen. In einer anschließenden Testphase wird allein der CS dargeboten. Der bedingte Reflex gilt als ausgebildet, wenn auf den CS eine ähnliche Reaktion eintritt wie zuvor auf den US. Dies wird konditionierte Reaktion (CR) genannt. Ein ausgebildeter konditionierter Reflex (CS→CR) kann auch wieder gelöscht werden (Extinction). Um dies zu erreichen, wird der CS wiederholt ohne US dargeboten. In der Folge lässt die Stärke der CR nach, bis sie schließlich völlig ausbleibt.

Als Physiologe beschrieb Pavlov die Ursachen für die Verhaltensänderung in physiologischen Termini. Er nahm an, dass das gemeinsame Auftreten von bedingtem und unbedingtem Reiz dazu führt, dass sich der „Punkt“ des Zentralnervensystems (ZNS), wie er sich ausdrückte, der durch den unbedingten Reflex aktiviert wird, mit dem „Punkt“ des ZNS verbindet, der durch den neutralen Reiz aktiviert wird.

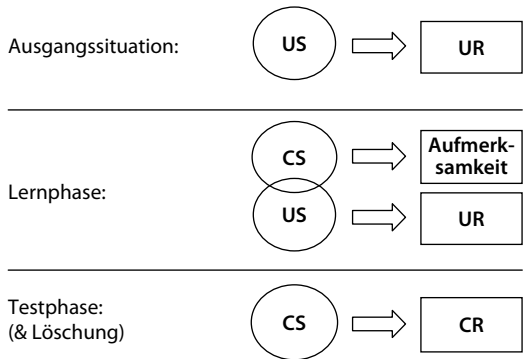


Abb. 2.2 Grundstruktur einer klassischen Konditionierung

In der Folge beginnt der neutrale Reiz die unbedingte Reaktion ebenfalls auszulösen. Die Verbindung wird umso stärker, je häufiger die beiden Aktivierungen gemeinsam auftreten, und sie schwächt sich ab, wenn die beiden „Punkte“ nicht gemeinsam aktiv sind. Damit waren die Grundgesetzmäßigkeiten des Aufbaus und der Löschung bedingter Reflexe beschrieben, in denen Pavlov die Grundlage für die „feinste Anpassung tierischen Verhaltens an seine Umwelt“ sah (Pavlov 1904).

Bezieht man diese Überlegungen auf die im ersten Kapitel besprochenen notwendig zu spezifizierenden Elemente einer Lerntheorie, dann wird angenommen, dass 1. (reflektorisches) Verhalten durch Reiz-Reaktions-Verbindungen determiniert wird. Der Lernvorgang besteht 2. darin, dass in bestehenden Reiz-Reaktions-Verbindungen der ursprüngliche Reiz durch einen anderen, neuen Reiz substituiert wird, wenn 3. die beiden Reize in einem raum-zeitlichen Zusammenhang (Kontiguität) wiederholt erlebt werden. Das ist die klassische Erklärung, die allerdings, wie wir sehen werden, unvollständig ist.

2.2.3 Ausgewählte Eigenschaften bedingter Reflexe

Generalisierung

Ist ein bedingter Reflex auf einen bestimmten CS hin erworben, wird die CR auch durch Reize ausgelöst, die dem verwendeten CS hinreichend ähnlich sind.

Ist z. B. eine Lidschlussreaktion auf einen Ton von 1000 Hertz (Hz) konditioniert worden, erfolgt der Lidschluss auch auf tiefere (etwa 800 Hz) oder höhere Töne (etwa 1200 Hz). Der bedingte Lidschlussreflex wird jedoch schwächer, je stärker der aktuell verwendete Ton vom konditionierten Ton abweicht (Generalisierungsgradient). Die Generalisierung bedingter Reflexe verweist darauf, dass nicht äußere Reize, sondern die von diesen Reizen ausgelösten neuronalen Aktivierungen konditioniert werden. In der Folge überträgt sich die Tendenz, die CR auszulösen, auf alle Reize, die eine hinreichend ähnliche Aktivierung im ZNS hervorrufen wie der CS. Die Generalisierung wird jedoch eingeschränkt, wenn dem CS ähnliche Reize ohne Zusammenhang zum US dargeboten werden. Die Tiere lernen dann zwischen den bekräftigten und den nichtbekräftigten Reizen zu unterscheiden (► Abschn. 2.4).

Der Einfluss der Zeitverhältnisse bei der Darbietung von CS und US

Im Hinblick auf die Zeitverhältnisse zwischen CS und US in der Lernphase lassen sich simultane, vorwärts gerichtete und rückwärts gerichtete Konditionierungen unterscheiden:

- Bei der **simultanen Konditionierung** werden der CS (z. B. ein Ton) und der US (z. B. Futter) gleichzeitig dargeboten.
- Bei der **vorwärts gerichteten Konditionierung** wird der CS vor dem US dargeboten.
- Bei der **rückwärts gerichteten Konditionierung** wird der CS nach dem US dargeboten.

Es wäre durchaus plausibel, anzunehmen, dass eine simultane Konditionierung zu den besten Resultaten führt, da hier der CS präsent ist, während der unbedingte Reflex abläuft, sodass er leicht mit ihm verbunden werden kann. Dies ist aber nicht der Fall. Es zeigt sich vielmehr, dass ein bedingter Reflex am schnellsten aufgebaut wird, wenn der CS kurz vor dem US dargeboten wird. Im Kontrast dazu werden rückwärts gerichtete Konditionierungen, wenn überhaupt, nur sehr schwer erworben.

Dies gibt einen Hinweis auf die Funktion bedingter Reflexe im Verhalten der Tiere: Der Erwerb bedingter Reflexe dient nicht dazu, unbedingtes

Verhalten auf neue Reize zu übertragen (Warum sollten Hunde auch lernen, auf Glockentöne zu speicheln?). Es wird vielmehr gelernt, dass es Reize gibt, die das Auftreten von verhaltenswichtigen (unbedingten) Reizen ankündigen, sodass sich die Tiere darauf vorbereiten können (z. B. die Hunde auf das kommende Futter, indem sie schon beim Ton der Glocke vorsorglich zu speicheln beginnen). Da eine solche Verhaltensvorbereitung nur Reize erlauben, die *vor* dem kritischen US auftreten, werden diese Reize besonders schnell konditioniert.

Konditionierte Hemmung

Reize können auch die Tendenz erwerben, eine konditionierte Reaktion zu unterdrücken. Im Gegensatz zu einem **aktivierenden CS** (CS+) spricht man von einem **hemmenden CS** (CS–) oder einer konditionierten Hemmung. Um eine konditionierte Hemmung aufzubauen, ist zunächst eine konditionierte Reaktion auszubilden. Nehmen wir z. B. an, dass Hunde gelernt haben, auf einen Ton (CS_T) verlässlich zu speicheln. In weiteren Versuchen wird nun der konditionierte Ton zusammen mit einem weiteren Reiz, sagen wir einem Lichtsignal, dargeboten (CS_T + Licht), jedoch ohne den Hunden nachfolgend Futter zu geben. Die Hunde erleben damit, dass auf Ton Futter folgt, aber auf Ton + Licht kein Futter folgt. Dieser Erfahrung entsprechend, erlischt die Speichelreaktion auf den kombinierten Ton-Licht-Reiz (Extinction), während die Hunde auf den Ton allein weiterhin speicheln. Die Hunde haben offensichtlich gelernt, dass auf den Lichtreiz kein Futter folgen wird, sodass der Lichtreiz die vom Ton angelegte Speichelproduktion vermutlich hemmt. Diese Vermutung wird in einem **Summationstest** bestätigt. Im Summationstest wird der Lichtreiz mit einem weiteren CS₂ zusammen dargeboten, der ebenfalls verlässlich eine Speichelreaktion auslöst. Wenn allein der Lichtreiz die Tendenz zu speicheln hemmt, dann sollte er auch die Speichelreaktion auf den neuen CS₂ unterdrücken, obwohl er nie zusammen mit diesem Reiz erlebt worden ist. Das kann gezeigt werden. Die aktivierende Wirkung des CS₂ und die hemmende Wirkung des Lichtes summieren sich, was bei etwa gleichstarken Tendenzen zu einem Ausbleiben der CR führt. Generell gilt: Reize, die verlässlich das Nichteintreten eines US vorhersagen, obwohl er

erwartet wird, erwerben die Tendenz, die jeweilige CR zu hemmen.

Bedingte Reflexe höherer Ordnung

Bei der Ausbildung eines bedingten Reflexes höherer Ordnung geht man ähnlich vor wie bei der Ausbildung konditionierter Hemmung. Zunächst wird ein erster Reiz verlässlich konditioniert (CS_1+) und dann mit einem weiteren Reiz kombiniert (CS_2), ohne dass nachfolgend der US dargeboten wird. Im Unterschied zur konditionierten Hemmung wird der neue Reiz jedoch nicht simultan mit dem ersten Reiz ($CS_1 + CS_2$), sondern kurz vor ihm dargeboten ($CS_2 \rightarrow CS_1$). Diese kleine Änderung führt zu einem vollständig anderen Lernresultat: Während bei simultaner Darbietung der neue Reiz zu einem hemmenden CS_2- wird, wird er bei vorausgehender Darbietung zu einem aktivierenden CS_2+ , d. h., er erwirbt die Tendenz, eine konditionierte Reaktion hervorzurufen. Würde man z. B. wiederholt kurz vor einem Ton, der bereits verlässlich eine Speichelreaktion auslöst, einen Lichtreiz darbieten, würden die Tiere beginnen, auch auf den Lichtreiz hin zu speicheln, obwohl sie den Lichtreiz nie in Zusammenhang mit der Gabe von Futter erlebt haben. Die Tendenz, auf den Ton zu speicheln, wird auf den Lichtreiz übertragen. Man spricht von einem **konditionierten Reflex zweiter Ordnung**. Die Übertragbarkeit der CR auf einen Reiz, der einem bereits erworbenen CS verlässlich vorausgeht, unterstreicht die **Vorhersagefunktion** bedingter Reflexe: Auch Reize, die einen US nur vermittelt ankündigen, erwerben die Fähigkeit, eine CR auszulösen.

Kontiguität und Kontingenz

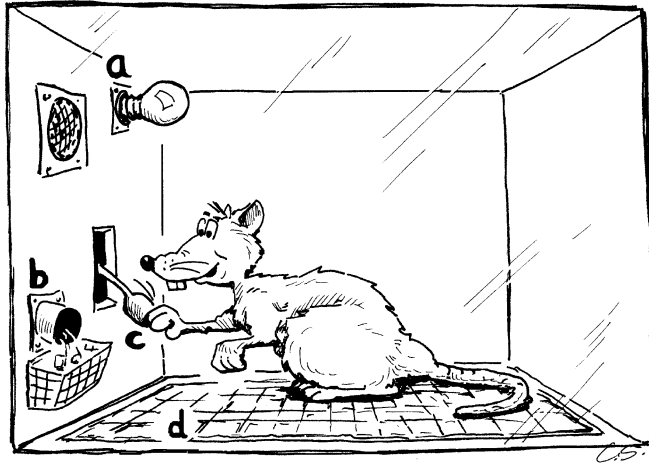
Die Existenz von Reflexen höherer Ordnung lässt bereits vermuten, dass es bei der Ausbildung bedingter Reflexe nicht notwendig darauf ankommt, dass CS und US in einem unmittelbar zeitlichen und räumlichen Zusammenhang auftreten (Kontiguität), sondern vielmehr darauf, dass ein Zusammenhang zwischen dem Auftreten des CS und dem Auftreten des US (Kontingenz) hinreichend verlässlich erlebt wird. Diese Vermutung wurde von Robert Rescorla (1968), einem der einflussreichsten Erforscher tierischen Lernens, experimentell überprüft:

Die Versuchstiere waren Ratten. Den Ratten wurde ein zwei Minuten anhaltender Ton als CS und ein leichter Elektroschlag als US dargeboten, auf den die Tiere eine unbedingte Vermeidungsreaktion zeigen. Allerdings wurde auch in den Pausen zwischen den Tönen der Schlag gelegentlich verabreicht. Rescorla variierte nun systematisch die Wahrscheinlichkeit, mit der der Schlag mit dem Ton ($p(US/CS)$) und die Wahrscheinlichkeit, mit der der Schlag in den Pausen zwischen den Tönen verabreicht wurde ($p(US/\sim CS)$). Im Ergebnis erwarben die Ratten nur dann eine konditionierte Vermeidungsreaktion allein auf den Ton, wenn die Wahrscheinlichkeit einen Schlag zu erhalten in Anwesenheit des Tones höher war als bei Abwesenheit des Tones. Wenn der Schlag mit gleicher Wahrscheinlichkeit auftrat, egal, ob ein Ton zu hören war oder nicht, wurde der Ton nicht konditioniert. Für die Ausbildung eines bedingten Reflexes ist danach nicht das gemeinsame Auftreten von CS und US, die **Kontiguität**, entscheidend, sondern ihr (statistischer) Zusammenhang, die **Kontingenz**. In anderen Worten: Ein bedingter Reflex wird nur dann ausgebildet, wenn der bedingte Reiz die Vorhersagbarkeit des US erhöht. Damit wird erneut die Funktion bedingter Reflexe als vorbereitende Verhaltensanpassung auf vorhersagbare Ereignisse bestätigt.

Definitionen

Kontiguität zwischen CS und US ist bestimmt durch die Wahrscheinlichkeit, mit der beide Reize in einem raum-zeitlichen Zusammenhang gemeinsam auftreten ($p(CS \& US)$).

Kontingenz wird durch den Anstieg der Wahrscheinlichkeit bestimmt, mit der der US unter der Bedingung eintritt, dass der CS eingetreten ist. Die stärkste Kontingenz besteht, wenn der US nur dann eintritt, wenn auch der CS dargeboten wird ($p(US/CS) = 1$ und $p(US/\sim CS) = 0$). Kein Zusammenhang besteht, wenn beide Wahrscheinlichkeiten gleich sind ($p(US/CS) = p(US/\sim CS)$). Ein negativer Zusammenhang besteht, wenn die Wahrscheinlichkeit für den US durch den CS gesenkt wird (z. B. $p(US/CS) = ,40$ und $p(US/\sim CS) = ,80$).



■ **Abb. 2.3** Eine typische Skinner-Box mit den Möglichkeiten, (a) einen Lichtreiz darzubieten, (b) Futterkugeln zu spenden, (c) eine Taste zu drücken und (d) einen leichten Stromschlag zu applizieren. © Claudia Styrsky

2.3 Instrumentelle Konditionierung

2.3.1 Versuche von Thorndike

Etwa zu der Zeit, als Pavlov in Petersburg seine Studien bei Hunden durchführte, untersuchte Edward Lee Thorndike an der Universität von Harvard das Verhalten von Katzen. Die Katzen befanden sich in einem Käfig, der sich durch das Herunterdrücken eines Pedals öffnen ließ. Außerhalb des Käfigs wurde verlockendes Futter angeboten. Die hungrigen Katzen versuchten alles Mögliche, um an das Futter zu kommen. Wenn sie dabei zufällig das Pedal herunterdrückten, öffnete sich der Käfig, und es wurde ihnen kurz erlaubt, zu fressen, bevor sie erneut in den Käfig gesperrt wurden. Thorndike beobachtete nun, dass die Katzen von Versuch zu Versuch immer schneller das Pedal zum Öffnen bedienten, bis sie es schließlich ohne jede Umschweife taten. Umgekehrt drückten die Katzen das Pedal zunehmend seltener und schließlich überhaupt nicht mehr, wenn sich der Käfig dadurch nicht mehr öffnen ließ.

2.3.2 Skinner-Box

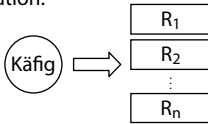
Burrhus F. Skinner wandelte das Vorgehen Thorndikes auf effektive Weise ab. Während bei Thorndike der Versuchsleiter nach jedem Versuch die

Katze wieder in den Käfig zurückbringen musste, entwickelte Skinner einen Experimentierkäfig – die sog. **Skinner-Box** –, in dem die Tiere die zu konditionierende Reaktion, im Beispiel das Drücken des Pedals, ohne jedes Einschreiten des Versuchsleiters beliebig oft wiederholen konnten (■ [Abb. 2.3](#)). Eine Skinner-Box ermöglicht typischerweise aber nicht nur eine, sondern verschiedene Verhaltensweisen, wie etwa das Drücken eines Hebels, das Zerren an einer Schnur, das Kratzen an einer Sperre oder das Picken auf eine Scheibe. Darüber hinaus ist die Box mit Vorrichtungen ausgestattet, mit denen Belohnungen wie Futterkugeln oder Wassertropfen, aber auch ein leichter elektrischer Schlag als Bestrafung verabreicht werden können. Schließlich gibt es Vorrichtungen für die Darbietung von visuellen und/oder akustischen Reizen.

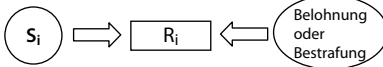
2.3.3 Effektgesetz („law of effect“)

Die ■ [Abb. 2.4](#) zeigt die Grundstruktur **instrumentellen Konditionierens**. Ausgangspunkt des Lernvorgangs ist die Ausführung frei gewählter Verhaltensweisen ($R_1 \dots R_n$) unter den gegebenen Reizbedingungen eines Experimentierkäfigs. Während der Lernphase wird ein bestimmtes experimentell ausgewähltes Verhalten (R_i) bei bestimmten Reizbedingungen (S_i) mehr oder weniger verlässlich

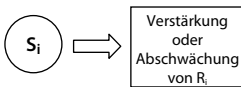
Ausgangssituation:



Lernphase:



Testphase:



■ **Abb. 2.4** Grundstruktur instrumentellen Konditionierens

belohnt bzw. bestraft. Instrumentelles Lernen manifestiert sich in der Testphase im Anstieg von Häufigkeit oder Intensität, mit der belohntes Verhalten bzw. in der Reduktion von Häufigkeit oder Intensität, mit der bestrafte Verhalten unter den jeweils experimentell gesetzten Reizbedingungen (S_i) ausgeführt wird.

Thorndike beschrieb als Psychologe und Verhaltensforscher die Ursachen der Verhaltensänderungen nicht in physiologischen Termini (wie Pavlov), sondern in den psychologischen Termini von Belohnung und Bestrafung: Nach seinen Überlegungen wird die Verbindung zwischen einer Reizsituation (S_i) und einem Verhalten (R_i) verstärkt („reinforcement“), wenn das Verhalten zu einer Belohnung führt. In der Folge tritt das Verhalten umso wahrscheinlicher in der jeweiligen Situation auf, je häufiger seine Ausführung in dieser Situation belohnt wurde. Umgekehrt wird die Verbindung zwischen einer Situation und einem Verhalten geschwächt, wenn das Verhalten in der entsprechenden Situation zu keiner Belohnung bzw. zu einer Bestrafung führt („punishment“). In der Folge tritt das Verhalten in der Situation zunehmend seltener auf. Das ist das sog. **Effektgesetz** („law of effect“) instrumentellen Konditionierens.

Bezieht man diese Überlegungen auf die zu fordernden Elemente einer Lerntheorie, dann wird (wie bei der klassischen Konditionierung) angenommen, dass (instrumentelles) Verhalten 1. durch Reiz-Reaktions-Verbindungen determiniert wird. Der Lernvorgang besteht 2. darin, dass die Stärke von Reiz-Reaktions-Verbindungen verändert wird, wobei 3. die Verbindungen durch nachfolgende Belohnung verstärkt und durch nachfolgende Bestrafung bzw.

das Ausbleiben von Belohnung geschwächt werden. Auch diese Erklärung ist, wie wir sehen werden, unvollständig.

2.3.4 Ausgewählte Eigenschaften instrumentellen Lernens

Verstärkungspläne

Wenn ein hungriges Tier nach einem bestimmten Verhalten im Experimentierkäfig stets Futter bekommt (kontinuierliche Verstärkung), dann überrascht es nicht, dass das Tier das entsprechende Verhalten in diesem Käfig immer häufiger zeigt. Was aber passiert, wenn das Verhalten nicht immer, sondern nur gelegentlich verstärkt wird (partielle oder intermittierende Verstärkung)? Dieser Frage ist man in zahlreichen Untersuchungen nachgegangen. Dabei wird zwischen Quoten- und Intervallverstärkung unterschieden, die jeweils nach einem festen oder nach einem variablen Plan realisiert sein können (■ [Tab. 2.1](#)):

- Bei einer **Quotenverstärkung** von z. B. 10 % wird entweder jedes zehnte Auftreten des kritischen Verhaltens (feste Quote) oder aber jedes Auftreten des kritischen Verhaltens mit einer Wahrscheinlichkeit von ,10 (variable Quote) verstärkt.
- Bei der **Intervallverstärkung** wird das kritische Verhalten nicht in Abhängigkeit von der Anzahl oder Häufigkeit seines Auftretens, sondern nur in bestimmten Zeitintervallen verstärkt. Bei einem festen 60-Sekunden-Intervall wird z. B. nach jeder Belohnung für 60 Sekunden jede weitere Belohnung ausgesetzt, unabhängig davon wie häufig das kritische Verhalten ausgeführt wird. Erst danach wird das nächste Auftreten des kritischen Verhaltens wieder belohnt. Bei variabler Intervallverstärkung variiert das Aussetzen der Belohnung zwischen kurzen und langen Intervallen, um aber im Mittel wieder einer vorgegebenen Zeitdauer, im Beispiel 60 Sekunden, zu entsprechen.

Tiere passen sich in der Regel schnell an solche partiellen Verstärkungspläne an: Bei festen Verstärkungsplänen treten typischerweise nach dem Erhalt

■ Tab. 2.1 Übersicht Verstärkungspläne

	Quote	Intervall
Fix	Beispielsweise wird jedes zehnte Verhalten bekräftigt.	Nach einer Bekräftigung werden weitere Bekräftigungen für z. B. 60 Sekunden ausgesetzt.
Variabel	Verhaltensakte werden mit einer Wahrscheinlichkeit von z. B. ,10 bekräftigt.	Die Bekräftigungspausen betragen im Mittel z. B. 60 Sekunden.

einer Belohnung sog. Nachverstärkungspausen auf, in denen die Tiere das belohnte Verhalten nicht zeigen. Nach dieser Pause zeigen die Tiere bei einer festen Quote das kritische Verhalten schnell hintereinander bis sie die nächste Belohnung erhalten, um sich dann erneut eine Pause ,zu gönnen‘. Bei einem festen Intervall steigt dagegen nach der Pause die Verhaltenshäufigkeit nur langsam an, um erst in der Nähe des nächstmöglichen Belohnungszeitpunktes ein Häufigkeitsmaximum zu erreichen – erhöhte Aktivität lohnt ja zuvor nicht, da Belohnungen ausgesetzt sind. Im Gegensatz zur festen Verstärkung sind bei variabler Verstärkung Nachverstärkungspausen kaum zu beobachten. Die Tiere wiederholen hier das kritische Verhalten eher kontinuierlich. Dies entspricht der Erfahrung, dass nach jeder Belohnung das kritische Verhalten sofort wieder belohnt werden kann, auch wenn die Wahrscheinlichkeit dafür gering sein mag.

Die partiellen Verstärkungspläne führen im Vergleich zur kontinuierlichen Verstärkung auch zu einer höheren **Löschungsresistenz** (► Beispiel). Dies erscheint auf den ersten Blick paradox. Nach dem „law of effect“ sollte kontinuierliche Verstärkung zu stärkeren S-R-Verbindungen führen als partielle Verstärkung, und stärkere Verbindungen sollten löschungsresistenter sein. Partielle Verstärkungspläne vermitteln aber auch die Erfahrung, dass auf eine Reihe unverstärkter Reaktionen immer wieder mal eine Belohnung folgen kann, sodass es sich durchaus lohnt, das Verhalten trotz ausbleibender Verstärkung beizubehalten. Partielle Verstärkungen halten gewissermaßen die Hoffnung auf eine Belohnung wach. Entscheidend für das Aufrechterhalten eines bestimmten Verhaltens ist auch hier wieder nicht die Häufigkeit, sondern die Erwartbarkeit von Belohnungen.

Beispiel

Die erhöhte Löschungsresistenz für Verhalten, das nur partiell verstärkt wird, sichert den Betreibern von Glücksspielautomaten hohe Renditen. Bei Glücksspielautomaten erfolgt die Bekräftigung durch einen Jackpot nach einer variablen Quote: Nach jeweils unterschiedlich langen Folgen von Spielen ohne Gewinn erfolgt eine Bekräftigung durch einen Jackpot. Die Erwartung eines Jackpot wird damit immer wieder neu genährt, sodass viele Spieler selbst einen defekten Glücksspielautomaten noch lange mit Münzen füttern würden, bis sie die Hoffnung auf einen Jackpot aufgeben. In einen defekten Getränkeautomaten wird dagegen wohl niemand lange Zeit Münzen einwerfen, bevor er aufgibt, da hier eine Belohnung (die Ausgabe des Getränkes) immer erwartet wird.

Konditionierte Verstärkung und die Ausbildung von Verhaltensfolgen

Zur instrumentellen Konditionierung von Verhaltensweisen werden zumeist Belohnungen eingesetzt, die elementare physiologische Bedürfnisse der Tiere unmittelbar befriedigen. In aller Regel erhalten hungrige Tiere Futter und durstige Tiere Wasser. Man spricht von **primären Verstärkern**. Als Verstärker können aber auch neutrale Reize dienen, wenn sie mit einem primären Verstärker assoziiert werden. Stellen Sie sich z. B. eine Gruppe von Ratten vor, denen zunächst wiederholt kurz vor der Fütterung ein Ton dargeboten wird. Anschließend wird in einer Skinner-Box das Drücken einer Taste von eben diesem Ton gefolgt. Im Ergebnis steigt die Rate des Tastendrückens an. Die Erklärung dafür liegt auf der Hand: Durch die Paarung mit dem Futter (US) wird der Ton klassisch konditioniert. Er kündigt

nun als CS+ eine bevorstehende Fütterung an. Und diese Erwartung einer zukünftigen Belohnung reicht auch hier wieder, um vorangehendes Verhalten (hier das Drücken der Taste) zu verstärken. Man spricht deshalb von einem konditionierten oder **sekundären Verstärker**.

Belohnungserwartungen spielen eine wichtige Rolle beim Erwerb von **Verhaltensfolgen**, die erst am Ende eine primäre Verstärkung erfahren. Im Zirkus kann man oft Dressuren bewundern, in denen Tiere eine Folge ungewöhnlicher Verhaltensweisen zeigen, z. B. einen Hund, der auf eine Leiter klettert, einen schmalen Steg entlang balanciert, an einer Schnur zerrt, um eine Tür zu öffnen, die den Weg zu einer Rutsche frei gibt, die er herunterrutscht, um schließlich belohnt zu werden. In dieser Verhaltensfolge bewegt sich der Hund vom Fuß der Leiter hinauf auf das Podest, über den Steg bis vor die Tür mit der Schnur, durch die Tür vor die Rutsche und schließlich zum Ende der Rutsche, wo die Belohnung erfolgt. Jede der Zwischenstationen bringt das Tier verlässlich näher an die Belohnung und kann so als konditionierter Verstärker für den vorangegangenen Verhaltensakt wirken. Gleichzeitig liefert die jeweils gegebene Situation den (An-)Reiz für die Ausführung des nächsten Verhaltensschrittes.

Kontiguität und Kontingenz

Bei der klassischen Konditionierung kommt es, wie wir gesehen haben (► Abschn. 2.2.3 „Kontiguität und Kontingenz“), nicht auf die Kontiguität, sondern auf die Kontingenz zwischen CS und US an: Ein bedingter Reflex wird nur dann ausgebildet, wenn das Auftreten des CS die Erwartbarkeit des US erhöht. Bei der instrumentellen Konditionierung tritt an die Stelle des Verhältnisses von CS und US das Verhältnis von Verhalten (R) und Belohnung. Um auch hier den Einfluss der Kontingenz zu untersuchen, geht man ähnlich vor wie bei der klassischen Konditionierung: Es wird ein **Zeitfenster** festgelegt. Wenn innerhalb des Zeitfensters das interessierende Verhalten ausgeführt wird, erfolgt eine Belohnung mit einer bestimmten Wahrscheinlichkeit ($p(B+/R)$). Das Tier erhält aber auch dann gelegentlich eine Belohnung, wenn innerhalb des Zeitfensters das kritische Verhalten nicht

ausgeführt wurde ($p(B+/\sim R)$). Durch die Variation dieser beiden Wahrscheinlichkeiten kann die Stärke des Zusammenhanges (die Kontingenz) zwischen dem Verhalten und dem Eintreten einer Belohnung beliebig justiert werden (► Definition in Abschn. 2.2.3 „Kontiguität und Kontingenz“). Im Ergebnis entsprechender Untersuchungen zeigt sich erneut, dass sich die Rate des jeweils beobachteten Verhaltens umso stärker erhöht, je stärker seine Ausführung die Wahrscheinlichkeit einer Belohnung erhöht, je größer die Differenz $p(B+/R) - p(B+/\sim R)$ also ausfällt. Wenn dagegen beide Wahrscheinlichkeiten gleich sind, wird das Verhalten nicht konditioniert. Wenn z. B. Ratten mit gleicher Wahrscheinlichkeit Futterkugeln erhalten, egal, ob sie eine Taste drücken oder nicht, erhöht sich die Rate des Tastendrückens kaum. Warum auch? Das Drücken der Taste verschafft den Tieren ja kein zusätzliches Futter.

2.4 Diskriminationslernen

2.4.1 Unterscheidung von verhaltensrelevanten und verhaltensirrelevanten Reizbedingungen

Klassische und instrumentelle Konditionierung führen dazu, dass verhaltensrelevante von verhaltensirrelevanten Umgebungsreizen unterschieden werden. Der Erwerb gezielter Unterscheidungen, ist schon früh untersucht worden. Bereits Pavlov berichtete von einem Experiment, in dem Hunde nach der Darbietung eines Kreises gefüttert wurden, während sie nach der Darbietung eines Quadrates kein Futter bekamen. Anfänglich speichelten die Hunde auf beide Reize, aber nach einigem Training speichelten sie nur noch bei der Darbietung des Kreises. Die Hunde hatten gelernt, dass allein der Kreis eine bevorstehende Fütterung ankündigt. Auch beim **instrumentellen Konditionieren** lernen Tiere Reizbedingungen, unter denen ein Verhalten belohnt wird, von Reizen zu unterscheiden, unter denen das gleiche Verhalten zu keiner Belohnung führt. Wenn z. B. bei grünem Licht das Drücken eines Hebels den Futterspender aktiviert, aber keine Futterkugeln gegeben werden, wenn ein rotes Licht leuchtet, dann drücken

die Ratten den Hebel nur noch bei „Grün“, aber nicht mehr bei „Rot“. Sie lernen auch, unter verschiedenen Reizen verschiedenes Verhalten zu zeigen. Wenn etwa bei grünem Licht nur das Drücken einer Taste und bei rotem Licht nur das Zerren an einer Schnur zur Belohnung führt, dann werden die Tiere nach einigen Versuchen bei „Grün“ bevorzugt die Taste drücken und bei „Rot“ bevorzugt an der Schnur zerren.

Dass Tiere zwischen Reizbedingungen zu unterscheiden lernen, unter denen ihr Verhalten zu unterschiedlichen Konsequenzen führt, ist wenig erstaunlich. In ihrer natürlichen Umwelt haben Tiere solche Unterscheidungen ständig zu treffen: Sie müssen zwischen Freund und Feind, zwischen eigenen und fremden Nachkommen, zwischen gefährlichen und sicheren Orten, zwischen erreichbarer und nicht-erreichbarer Beute, zwischen bekömmlicher und unbekömmlicher Nahrung usw. unterscheiden – und selbstverständlich tun sie das auch. Die Forschung zum Diskriminationslernen hat aber auch weniger selbstverständliche **Diskriminationsleistungen** offenbart. Zwei dieser Leistungen wollen wir im Folgenden behandeln.

2.4.2 Positives und negatives Patterning

Der Erfolg eines Verhaltens hängt oftmals nicht nur von einem, sondern von mehreren Situationsmerkmalen ab. Eine Amsel wird z. B. das Füttern ihrer hungrigen Küken aussetzen, wenn ein Bussard über dem Nest kreist. Das Verhalten wird hier durch die Kombination von zwei Merkmalen determiniert: vom Betteln der Küken und vom Bussard. Wie Tiere Verhaltensabhängigkeiten von solchen Reizkombinationen erwerben, ist in sog. „Patterning-Studien“ untersucht worden. Das vermutlich erste Experiment dieser Art wurde von Woodbury 1943 berichtet. Hunde wurden mit Futter belohnt, wenn sie mit der Schnauze eine Sperre lösten. Die Belohnung wurde aber nur dann gegeben, wenn während des Versuchs entweder nur ein hoher oder nur ein tiefer Ton zu hören war. Waren beide Töne zu hören, erhielten die Tiere keine Belohnung (negatives Patterning).

Anfänglich lösten die Hunde die Sperre, egal ob nur einer der beiden oder beide Töne zu hören waren. Nach etwa 300 Versuchen nahm jedoch die Verhaltensrate in den Durchgängen mit beiden Tönen rapide ab. Die Hunde hatten nun gelernt, dass es sich nicht lohnt, die Sperre zu lösen, wenn beide Töne zu hören sind.

Die ■ Abb. 2.5 stellt die Situation noch einmal formal dar. Es sind zwei Reize A und B gegeben. Eine Belohnung erfolgt nur dann, wenn einer der beiden Reize allein, aber nicht wenn beide Reize gemeinsam dargeboten werden (A+, B+, AB–; negatives Patterning). Beim positiven Patterning wird umgekehrt das Verhalten nur dann belohnt, wenn beide Reize, aber nicht, wenn nur einer der Reize dargeboten werden (A–, B–, AB+). Darüber hinaus gilt für beide Fälle, dass das Verhalten nie belohnt wird, wenn keiner der beiden Reize vorliegt.

Wie die Abbildung deutlich macht, führen positives und negatives Patterning bei gleich häufigen Merkmalskombinationen dazu, dass das Verhalten auf die beiden einzelnen Reize für sich genommen mit jeweils gleicher Häufigkeit belohnt und nicht belohnt wird. Nach dem „law of effect“ sollte jede Belohnung zu einer Stärkung und jedes Ausbleiben von Belohnung zu einer Schwächung der jeweiligen S-R-Beziehung führen. Stärkung und Schwächung sollten sich im Verlauf des Trainings also gegenseitig aufheben. Dementsprechend sollte konditioniertes Verhalten bei keinem der beiden Bekräftigungsmuster beobachtbar sein. Dies ist nicht der Fall. Es ist vielmehr wiederholt gezeigt worden, dass Tiere die Merkmalskombinationen zu unterscheiden lernen. Sie reagieren bei positivem Patterning dann nur noch auf die Kombination der beiden Merkmale und bei negativem Patterning nur noch auf die Einzelmerkmale. Dieses Ergebnis verweist darauf, dass sich das Lernen nicht nur auf einzelne Reize, sondern auch auf **Reizkombinationen** („stimulus compounds“) beziehen kann, die dann als eigenständige Einheiten mit Verhaltensweisen verbunden werden (positives Patterning) bzw. nicht verbunden werden (negatives Patterning). Mit anderen Worten: Reizkombinationen können, wenn sie Verhaltensrelevant sind, zu eigenständigen Reizmustern integriert werden.

Positives Patterning ($A \wedge B$) Negatives Patterning ($A \vee B$)

	A	$\neg A$
B	+	-
$\neg B$	-	-

	A	$\neg A$
B	-	+
$\neg B$	+	-

+ Belohnung

■ **Abb. 2.5** Positives und negatives Patterning von zwei Reizen (A und B)

2.4.3 Bildung von Reizkategorien

Wenn unter natürlichen Lebensbedingungen zwischen Freund und Feind, zwischen essbarer und giftiger Nahrung, zwischen guten und schlechten Fluchtwegen usw. zu entscheiden ist, dann finden sich in den jeweiligen Kategorien Beispiele unterschiedlichen Aussehens, auf die dennoch gleiches Verhalten gefordert ist. Um diese Anforderung im Experiment nachzustellen, wurden als verhaltenskritische Reize Bilder in den Experimentierkäfig projiziert. In jedem Versuchsdurchgang wurde ein neues Bild gezeigt, und in Abhängigkeit vom Bild wurde ein bestimmtes zu konditionierendes Verhalten, wenn es auftrat, belohnt oder nicht belohnt. In einer viel zitierten Untersuchung von Herrnstein et al. (1976) an Tauben wurden z. B. 80 Bilder von natürlichen Szenen verwendet. Auf 40 dieser Bilder waren Bäume zu sehen, während die anderen 40 Bilder keine Bäume zeigten. Das Picken der Tauben auf eine Plastikscheibe wurde nur dann mit Futterkörnern belohnt, wenn ein Bild mit Bäumen zu sehen war. Bei Bildern ohne Bäume wurde das Picken auf die Plastikscheibe nicht belohnt. Nach einigem Training pickten die Tauben nur noch bei Bildern mit Bäumen. Dieses Verhalten wurde auch auf Bilder übertragen, die im Training nie dargeboten wurden. Die Tauben hatten offensichtlich gelernt, Bilder mit Bäumen von Bildern ohne Bäume zu unterscheiden.

Dieses Experiment hat viele weitere Untersuchungen angeregt, in denen die unterschiedlichsten **kategorialen Unterscheidungen** fast immer mit Erfolg getestet wurden. So lernen Tiere nur auf Bilder von Menschen, von Katzen oder von

Blumen zu reagieren, so, wie sie es lernen, Autos von anderen Objekten, den Buchstaben A von anderen Buchstaben, Bilder von Monet von Bildern von Picasso oder moderne Musik von Barockmusik zu unterscheiden. Dabei wurden neben Tauben auch Affen, Chinchillas, Hühner, Wachteln, Papageien usw. untersucht. Diese überwältigende Evidenz für kategoriale Unterscheidungen bei verschiedenen Tierarten hat Herrnstein 1990 (S. 138) zu der Schlussfolgerung veranlasst, dass kategoriale Unterscheidungen auf jedem Niveau des Tierreiches gefunden werden können, wenn man nur kompetent danach sucht.

So beeindruckend die kategorialen Unterscheidungen sind, so schwierig ist es, zu bestimmen, worauf sie beruhen. Es ist verlockend, anzunehmen, dass die Unterscheidungen auf Konzepte verweisen, die denen menschlicher Konzeptbildungen entsprechen. Vielleicht ist es tatsächlich so, dass die Tiere abstrakte Konzepte wie „Baum“, „Mensch“, „Auto“ oder „Barockmusik“ bilden. Das muss aber nicht so sein. Es könnte auch sein, dass die Tiere lediglich eines oder mehrere Merkmale derjenigen Bilder abstrahieren, auf die sie eine Belohnung erleben, um dann das belohnte Verhalten immer dann zu zeigen, wenn auf einem Bild wenigstens eines dieser Merkmale zu sehen ist. Möglicherweise speichern die Tiere aber auch nur konkrete Erinnerungen an so viele belohnte Bilder wie möglich und reagieren auf ein neues Bild immer dann, wenn es einem dieser positiven Beispiele hinreichend ähnlich ist. Oder es wird ein Prototyp als Durchschnitt aller erlebten positiven Beispiele gespeichert, und die Reaktion auf ein aktuelles Bild hängt von seiner Ähnlichkeit zu diesem Prototyp ab. Untersuchungen haben für alle diese Alternativen Hinweise gefunden, was die Schlussfolgerung erlaubt, dass alle genannten Möglichkeiten kategorialer Unterscheidungen von den Tieren genutzt werden können (vgl. Güntürkün 1996, S. 105ff; Mazur 2006, S. 351ff; Pearce 1997, S. 118ff). Für menschliche Konzeptbildungen gilt übrigens das Gleiche, sodass die kategorialen Unterscheidungen bei Tieren durchaus als Vorläufer menschlicher Konzeptbildungen verstanden werden können (► [Abschn. 5.3](#)).

2.5 Die Selektivität der Bildung von S-R-Verbindungen

Nach der bisherigen Diskussion werden Reiz-Reaktions-Verbindungen zwangsläufig gebildet, wenn ein Reiz die Wahrscheinlichkeit für das Auftreten eines unconditionierten Reflexes (klassische Konditionierung) bzw. die Wahrscheinlichkeit für die Belohnung eines Verhaltens erhöht (instrumentelle Konditionierung). Lernen erscheint damit ausschließlich durch äußere Umstände determiniert. Wenn dem tatsächlich so wäre, müsste unter gleichen äußeren Bedingungen jeweils gleiches Lernen beobachtbar sein. Dies ist jedoch nicht der Fall.

2.5.1 Latente Hemmung der Ausbildung eines bedingten Reflexes

Wenn in einem Experimentierkäfig ein neuer Reiz dargeboten wird, dann unterbrechen die Tiere gewöhnlich ihr Verhalten, um sich dem Reiz zuzuwenden. Diese Hinwendung zu einem neuen Reiz nennt man **Orientierungsreaktion**. Unter den natürlichen Lebensbedingungen der Tiere sichert die Orientierungsreaktion, dass Veränderungen der Situation bemerkt werden. Die Tiere sind damit davor geschützt, das Erscheinen etwa eines Feindes, aber auch das Auftauchen einer lohnenden Beute zu übersehen. Wird ein Reiz jedoch wiederholt dargeboten, ohne dass er für das Tier verhaltensrelevant ist, erlischt die Orientierungsreaktion (Habituation). Die Tiere integrieren dann den Reiz in ihre gewohnte Umgebung und beachten ihn nicht weiter. Verwendet man anschließend einen so habituierten Reiz in einem Konditionierungsexperiment, dann zeigt sich, dass er umso schwerer konditioniert wird, je schwächer die Orientierungsreaktion zu ihm ausfällt. Dieses Phänomen hat man **latente Hemmung** genannt: Die Konditionierung eines Reizes wird durch seine Nichtbeachtung gehemmt. Es lässt sich damit schlussfolgern, dass ein Tier einen Reiz beachten muss, um ihn konditionieren zu können. Über Reize, die die Tiere nicht beachten (die keine Orientierungsreaktion bei ihnen auslösen), wird nichts gelernt (vgl. Pearce 1997, S. 65ff).

2.5.2 Blockierung der Ausbildung eines bedingten Reflexes

Die Entdeckung der **Blockierung** geht auf ein Experiment zurück, das 1969 von Kamin berichtet wurde. Das Experiment wurde an Ratten durchgeführt und gliederte sich in **drei Phasen**:

- In der ersten Phase wurde einer Experimentalgruppe von Versuchstieren wiederholt nach einem Ton ein leichter elektrischer Schlag versetzt, was dazu führte, dass allein der Ton bereits zu einer konditionierten Vermeidungsreaktion führte.
- In der zweiten Phase wurde die Konditionierung mit der Experimentalgruppe fortgeführt und mit einer Kontrollgruppe neu begonnen. Allerdings wurde in beiden Gruppen zusätzlich zum Ton ein Lichtreiz dargeboten.
- In der dritten Phase wurde schließlich der in der zweiten Phase verwendete Lichtreiz alleine dargeboten, um zu prüfen, ob er eine konditionierte Vermeidungsreaktion auslösen würde. Dies war nur in der Kontrollgruppe der Fall. Die Experimentalgruppe zeigte keine spezifische Reaktion auf den Lichtreiz.

Für beide Gruppen war der Lichtreiz neu, sodass er sicherlich von den Tieren beachtet wurde. Beide Gruppen erlebten den Lichtreiz auch als zuverlässigen Prädiktor des elektrischen Schlages. Warum kommt es dann nur in der Kontrollgruppe zu seiner Konditionierung?

Der Unterschied zwischen den Gruppen bestand lediglich in der Vorerfahrung hinsichtlich des mit dem Licht dargebotenen Tones. In der Experimentalgruppe war der Ton bereits konditioniert, während er für die Kontrollgruppe, so wie der Lichtreiz, neu war. Erinnern wir uns: Ein konditionierter Reiz löst die konditionierte Reaktion aus, weil er das Eintreten des unconditionierten Reizes (hier den elektrischen Schlag) verlässlich ankündigt. Für die Tiere der Experimentalgruppe kündigte in der zweiten Phase des Experimentes also bereits der Ton den elektrischen Schlag verlässlich an, die den Schlag damit erwarten konnten. Für die Tiere der Kontrollgruppe

kam der Schlag dagegen überraschend. Das Ergebnis deutet damit darauf hin, dass eine bereits bestehende Erwartung des US (hier der Schlag) die Konditionierung neu hinzutretender Prädiktoren blockiert. Die neuen Reize werden nicht mehr in die Verhaltenssteuerung einbezogen. Warum auch – das Tier stellt sich ja bereits auf die zu erwartenden Verhaltensanforderungen (hier die Vermeidung des Schlages) ein. In den Worten von Kamin:

- » ... perhaps for an increment in an associative connection to occur, it is necessary that the US instigate some 'mental work' on behalf of the animal. This mental work will occur only if the US is *unpredicted* – if it in some sense surprises the animal (Kamin 1969, S. 59).

2.5.3 Erlernte Hilflosigkeit: Die Blockierung des Vermeidungslernens

Tiere lernen für gewöhnlich schnell, wie sie vorhersehbaren Gefahren ausweichen können, wenn sie dazu die Gelegenheit haben. Wird etwa Hunden durch ein Licht ein leichter elektrischer Schlag angekündigt, dem sie durch das Überspringen einer Barriere in einen anderen Teil des Käfigs entfliehen können, dann genügen wenige Versuche, um dieses Fluchtverhalten stabil an den Lichtreiz zu binden: Die Tiere fliehen dann unmittelbar nach dem Licht in den anderen Teil des Käfigs und vermeiden damit den Schlag (z. B. Solomon und Wynne 1953).

Damit dies gelernt wird, müssen die Hunde dem Schlag zu entfliehen suchen, denn anders können sie nicht erfahren, dass sie der Sprung über die Barriere vor dem Schlag bewahrt. Diese gewissermaßen natürliche Flucht vor einem schmerzhaften Reiz kann, wie Martin E. P. Seligman von der Universität Pennsylvania gezeigt hat, blockiert werden (z. B. Seligman 1975; Seligman und Maier 1967). In einem entsprechenden Versuch wurde zwei Gruppen von Hunden zunächst eine Reihe von leichten Stromschlägen verabreicht. Die Hunde der einen Gruppe konnten durch das Drücken eines Hebels den Stromschlag beenden. Die Hunde der anderen Gruppe erhielten die gleiche Anzahl von Stromschlägen in gleicher Dauer, ohne

aber Einfluss darauf nehmen zu können. Anschließend wurden die Hunde beider Gruppen in die oben geschilderte Lernsituation mit dem zweigeteilten Käfig gebracht. Während alle Hunde der Hebelgruppe schnell lernten, den angekündigten Schlag durch einen Sprung über die Barriere zu vermeiden, lernten es zwei Drittel der anderen Gruppe nicht. Diese Hunde legten sich vielmehr hin und erduldeten winselnd den Stromschlag, ohne den Versuch zu machen, ihm zu entkommen. Seligman interpretierte dieses Verhalten als Ausdruck einer negativen Lernerfahrung, die zu einer, wie er es nannte, **erlernten Hilflosigkeit** führte: In der ersten Phase des Versuchs machten die Hunde die Erfahrung, dass sie, egal was sie taten, dem Schlag nicht entfliehen konnten. Dementsprechend bildeten sie die Erwartung aus, dass ihr Verhalten keinen Einfluss auf den schmerzhaften Reiz hat. Diese Erwartung führte dazu, dass die Tiere auch in anderen Situationen nicht mehr versuchten, schmerzhaften Reizen zu entkommen. Sie hatten in einer Situation erfahren, dass sie hilflos sind und übertrugen diese Erfahrung auf andere Situationen in denen sie sich durchaus hätten helfen können. Sie versuchten es nur nicht.

Instrumentelles Lernen setzt grundsätzlich eine allgemeine Verhaltensbereitschaft voraus: Nur wenn gegebene Verhaltensmöglichkeiten auch ausprobiert werden, kann ein Verhalten gefunden werden, das zur Belohnung bzw. zur Vermeidung von Bestrafung führt. Jede Erfahrung, die zu einer **Reduktion der allgemeinen Verhaltensbereitschaft** führt, bewirkt damit auch eine Verminderung der Gelegenheiten für instrumentelles Lernen und damit zu Hilflosigkeit im Sinne des Ausgeliefertseins an die jeweilige Situation. Das Phänomen der erlernten Hilflosigkeit verweist darauf, dass die Erfahrung, dass eigenes Verhalten keinen Einfluss auf Bestrafungen hat, zu einer solchen Reduktion von Verhaltensbereitschaft führen kann, was in der Folge die Chancen, neues Verhalten zu lernen, generell vermindert. Erlernte Hilflosigkeit als Folge des Erlebens von Situationen, in denen Bestrafungen oder auch nur aversive Reize nicht vermieden werden können, sondern erduldet werden müssen, ist bei vielen unterschiedlichen Arten experimentell gefunden worden. Auch bei Menschen ist das Entstehen erlernter Hilflosigkeit experimentell untersucht worden (► Exkurs).

Exkurs

Erlernte Hilflosigkeit beim Menschen

In einer der ersten Untersuchungen zur erlernten Hilflosigkeit beim Menschen wurden Versuchspersonen lautem Lärm ausgesetzt (Hiroto und Seligman 1975). Es wurde ihnen gesagt, dass es eine Möglichkeit gäbe, den Lärm zu beenden. Tatsächlich konnte jedoch nur ein Teil der Versuchspersonen den Lärm durch Drücken eines Knopfes beenden; die anderen Versuchspersonen hatten dagegen keinen Einfluss auf den Lärm, den sie, egal was sie taten, erdulden mussten. In einer anschließenden Testphase wurden alle Versuchspersonen erneut mit lauten Tönen konfrontiert, die durch ein Lichtsignal angekündigt wurden und nun durch das Drücken eines Hebels abstellbar waren. Die Versuchspersonen, die zuvor den Lärm kontrollieren konnten, lernten schnell, den Hebel gleich nach dem Lichtsignal zu bedienen und so die unangenehmen

Töne in etwa 90 % der Durchgänge zu vermeiden. Diejenigen Versuchspersonen aber, die in der Trainingsphase „hilfloos“ waren, zeigten eine deutlich beeinträchtigte Lernleistung: Sie bedienten den Hebel zur Vermeidung der Töne in weniger als durchschnittlich 50 % der Fälle. Interessant ist, dass bei weiteren Versuchen das Erdulden unkontrollierbaren Lärms zu einem Lerndefizit nicht nur bei der Vermeidung von Lärm, sondern auch bei der Lösung von Anagrammaufgaben führte. Dies deutet darauf hin, dass die Erfahrung der Unkontrollierbarkeit in einem Verhaltensbereich auf andere Verhaltensbereiche übertragen und damit zu einer generalisierten Hilflosigkeit führen kann. Es gilt wohl generell, dass uns unkontrollierbare Ereignisse frustrieren. Wenn unser Verhalten keinerlei Einfluss auf das Eintreten

unangenehmer, aber auch auf das Eintreten angenehmer Ereignisse hat, also auf Bestrafungen und Belohnungen, dann tendieren wir dazu, weitere Anstrengungen sowohl zur Vermeidung von Bestrafung als auch zur Erlangung von Belohnung zu unterlassen und schmälern damit unsere Chancen, entsprechendes Verhalten zu erlernen. Die Erfahrung, positive wie negative Ereignisse durch das eigene Verhalten kontrollieren zu können, ermutigt dagegen, auch in neuen Situationen die Wirkungen des eigenen Verhaltens auszuprobieren und erhöht damit die Chancen, Neues zu lernen. Eine wichtige Erziehungsmaxime lautet demnach: Lasst Kindern die Freiheit, unterschiedliches Verhalten auszuprobieren, und gebt verlässliche konsistente Rückmeldungen über Erfolge und Misserfolge des Verhaltens.

2.5.4 Preparedness: angeborene verhaltensgebundene Aufmerksamkeit

Die beiden vorangegangenen Abschnitte haben gezeigt, dass die Ausbildung von Reiz-Reaktions-Verbindungen nicht allein von den äußeren Verhältnissen abhängt, sondern dass die Tiere Verhaltensweisen generieren und „mentale Arbeit“, wie dies Kamin ausgedrückt hat, investieren müssen, damit Lernen stattfindet. Die mentale Arbeit bezieht sich vor allem auf die Ausrichtung der Aufmerksamkeit und dabei sowohl auf eine Beachtung der zu konditionierenden Reize als auch auf eine Beachtung der positiven wie negativen Verhaltenskonsequenzen. Weitere Beobachtungen haben nun gezeigt, dass die Generierung von Verhalten und die Ausrichtung der Aufmerksamkeit nicht unabhängig voneinander

sind sondern dass vielmehr die Ausrichtung der Aufmerksamkeit durch das aktuelle Verhalten wesentlich mitbestimmt wird. Tiere sind genetisch darauf vorbereitet („prepared“) in bestimmten Verhaltenskontexten auf bestimmte Reize zu achten. Eine frühe Demonstration dieses Zusammenhanges liefert ein **Konditionierungsexperiment** von Garcia und Koelling (1966).

Als unbedingte Reflexe wurden die Vermeidung eines leichten elektrischen Schlags bzw. die Vermeidung einer leichten Übelkeit verwendet. Die zu konditionierenden Reize (CS) waren der Geschmack von Zuckerwasser und ein visuell-akustischer Reiz (Licht mit Klickergeräusch). In der Konditionierungsphase tranken durstige Ratten vom Zuckerwasser, während gleichzeitig Licht mit Klickern dargeboten wurden. Beide CS wurden also gleichzeitig dargeboten. Kurz nach dem

Trinken wurde den Tieren entweder der elektrische Schlag versetzt oder es wurde eine Übelkeit provoziert. Dem zu vermeidenden Ereignis (US, Schlag oder Übelkeit) gingen also stets zwei unterschiedliche CS verlässlich voraus: Der süße Geschmack und der kombinierte Licht-Klicker-Reiz. In der anschließenden Testphase wurde geprüft, inwieweit die beiden CS im jeweiligen Verhaltenskontext konditioniert wurden, also inwieweit sie alleine eine Vermeidungsreaktion auslösten. Im Kontext eines zu vermeidenden elektrischen Schlages wurde nur der Licht-Klicker-Reiz, aber nicht der süße Geschmack des Wassers konditioniert. Im Kontext der zu vermeidenden Übelkeit wurde umgekehrt nur der süße Geschmack, aber nicht der Licht-Klicker-Reiz konditioniert.

Das Ergebnis entspricht der Lebenserfahrung der Ratten: Äußere Attacken wie ein Schlag gehen immer von Änderungen in der Umgebung aus, die sich zumeist durch visuelle und/oder akustische Reize ankündigen. Umgekehrt ist Übelkeit fast immer ein Resultat der Nahrungsaufnahme, die durch ihren Geschmack differenziert werden kann. Die Ratten sind dementsprechend darauf vorbereitet, zur Vermeidung von äußeren Attacken vor allem auf visuelle und/oder akustische Reize zu achten bzw. zur Vermeidung von Übelkeit vor allem auf Geschmacksreize zu achten, die dann jeweils auch bevorzugt konditioniert werden.

Vergleichbare Ergebnisse sind auch von anderen Tierarten berichtet worden: So werden z. B. bei Tauben Ton-Schlag-Verbindungen außerordentlich schnell, aber Farb-Schlag-Verbindungen überhaupt nicht konditioniert. Umgekehrt werden Ton-Futter-Verbindungen nur schwer, aber Farb-Futter-Verbindungen leicht konditioniert. Auch hier entsprechen die Ergebnisse den natürlichen Lebensumständen: Das Fluchtverhalten auf einen leichten elektrischen Schlag wird leicht auf Töne übertragen, weil sich attackierende Feinde in der Regel durch Geräusche ankündigen, und das Picken von Körnern wird leicht auf Farben übertragen, weil sich Körner etwa von ungenießbaren Kieselsteinen in ihrer Färbung unterscheiden (Güntürkün 1996, S. 96ff; Seligman 1970). Insgesamt verweisen die Beobachtungen darauf, dass die Tiere dazu genetisch disponiert

(„prepared“) sind, in elementaren Verhaltensbereichen wie Verteidigung, Flucht oder Nahrungsaufnahme ihre Aufmerksamkeit jeweils auf diejenigen Reize zu richten, die für den Verhaltenserfolg wichtig sind. Damit wird das Lernen von vornherein auf relevante Reiz-Reaktions-Zusammenhänge konzentriert. Man spricht auch von **Lerndispositionen** (Tinbergen 1952), die dafür sorgen, dass bestimmtes Verhalten an bestimmte Reize schnell, an andere Reize aber nur schwer gebunden wird.

Im Übrigen ist „preparedness“ auch beim instrumentellen Konditionieren insofern zu finden, als zur Vermeidung von Bestrafungen bzw. zur Erlangung von Belohnungen jeweils bestimmte Verhaltensweisen leicht, andere dagegen schwer konditioniert werden. So führte z. B. Bolles (1973, zit. nach Alcock 1996, S. 31) einen Versuch zum Vermeidungslernen an Ratten in einem Laufrad durch. Die Ratten erhielten wiederholt nach einem Ton einen leichten elektrischen Schlag. In einer Gruppe von Ratten konnte der Schlag vermieden werden, wenn die Ratten sofort nach dem Ton im Laufrad wendeten. Die Ratten einer zweiten Gruppe konnten den Schock vermeiden, wenn sie sofort nach dem Ton aufhörten zu laufen und sich aufrichteten. Nur das Wenden, aber nicht das Aufrichten wurde als Vermeidungsreaktion gelernt. Durch ihre genetische Verhaltensausrüstung sind Ratten offensichtlich darauf vorbereitet, einer zu erwartenden äußeren Attacke eher durch Wegrennen als durch Stillstehen zu entgehen.

2.6 Rescorla-Wagner-Modell elementaren S-R-Lernens

Das Rescorla-Wagner-Modell (RWM) ist der bis heute einflussreichste Versuch einer theoretischen Integration von Befunden zum **assoziativen Lernen** bei Tieren. Das RWM ist ursprünglich zur Erklärung der Bildung konditionierter Reflexe entworfen worden, kann aber ebenso auf das instrumentelle Konditionieren und das Diskriminationslernen angewendet werden. An die Stelle der Bildung von Assoziationen zwischen konditionierten und unkonditionierten Reizen (CS→US) tritt dann die Bildung

von Assoziationen zwischen Reizen und (bekräftigten) Reaktionen (S→R).

der Assoziationsstärken aller Reize im aktuellen Versuch ist.

Alle drei Annahmen lassen sich in der folgenden Formel zum Ausdruck bringen:

$$\Delta V_j = \alpha(\lambda_{US} - \Sigma V_j)$$

Legende:

2.6.1 Modellbeschreibung

Rescorla und Wagner (1972) gingen bei der Entwicklung des Modells davon aus, dass ein bedingter Reflex auf einer assoziativen Verbindung zwischen dem CS und dem US beruht. Das Modell sollte also Aufbau und Abbau von CS-US-Verbindungen im Resultat aufeinanderfolgender Konditionierungs- bzw. Löschungversuche erklären. Es wurden drei Annahmen zugrunde gelegt:

- Die Assoziationsstärke (V_j) zwischen einem CS_j und einem US steigt schrittweise (inkrementell) an, wenn beide Reize zusammen erlebt werden (Konditionierung), und sie sinkt inkrementell, wenn der CS ohne US dargeboten wird (Löschung).
- Die insgesamt erreichbare Assoziationsstärke zu einem bestimmten US ist auf einen maximalen Wert (λ_{US}^{\max}) begrenzt. Damit wird gewährleistet, dass die Assoziationsstärke zu einem US nicht endlos anwachsen kann.
- Werden ein CS_j und ein US zusammen erlebt, wird der Anstieg der Assoziationsstärke zwischen ihnen durch die Differenz zwischen der maximal erreichbaren Assoziationsstärke und der im aktuellen Versuch insgesamt gegebenen Assoziationsstärke ($\lambda_{US}^{\max} - \Sigma V_j$) bestimmt. Zu Beginn des Lernens, wenn die Differenz zur maximalen Assoziationsstärke noch beträchtlich ist, werden also große Lernfortschritte gemacht, die mit der Annäherung an die maximale Assoziationsstärke immer geringer werden. Wird der US nicht dargeboten (Löschungversuch), bestimmt sich die Reduktion der Assoziationsstärken nach der Differenz der aktuell insgesamt gegebenen Assoziationsstärke zum Wert Null ($0 - \Sigma V_j$). Die Nichtdarbietung des US führt also zu einer umso größeren Reduktion der Assoziationsstärken, je größer die Summe

- In der Formel bezeichnet ΔV_j die inkrementelle Veränderung der Assoziationsstärke zwischen einem CS_j und dem US in einem Konditionierungs- bzw. Löschungversuch.
- Die Variable α dient der Anpassung an aktuelle Bedingungen und wird zumeist genutzt, um die unterschiedliche Assoziierbarkeit verschiedener CS_j zum Ausdruck zu bringen.
- Durch ΣV_j wird die Summe der Assoziationsstärken aller im jeweiligen Versuch dargebotenen Reize bezeichnet, die zum Ausdruck bringt, mit welcher Wahrscheinlichkeit das Eintreten des US aktuell erwartet wird.
- Für λ_{US} wird die maximal erreichbare Assoziationsstärke λ_{US}^{\max} eingesetzt, wenn der US dargeboten wird, während bei Nichtdarbietung des US der Wert Null eingesetzt wird.

Die Differenz in der Klammer lässt sich somit als Abweichung der aktuellen Vorhersage (ΣV_j) vom tatsächlich eintretenden Ereignis (λ_{US}^{\max} bzw. 0) interpretieren. Da es nach der Formel nur dann zu Veränderungen der Assoziationsstärken kommt, wenn diese Differenz ungleich Null ist, lässt sich der Grundgedanke des Rescorla-Wagner-Modells in der Aussage zusammenfassen:

- » Organisms only learn, when events violate their expectations (Rescorla und Wagner 1972, S. 75).

Die Anwendung der Formel des RWM erlaubt präzise Vorhersagen des Lernverlaufs unter verschiedenen Bedingungen (► Exkurs).

Exkurs

Aufbau und Löschung eines bedingten Reflexes nach dem Rescorla-Wagner-Modell für nur einen und für zwei gleichzeitig dargebotene CS

Die **Abb. 2.6** beschreibt Anstieg und Abfall der Assoziationsstärke eines CS_1 bei einem α von 0,2 und einem λ_{\max} von 100. Im ersten Versuch beträgt V_1 Null, d. h. es besteht keine Assoziation zwischen dem CS_1 und dem US. Nach dem RWM wird durch eine Bekräftigung die Assoziationsstärke um $0,2 \times 100 = 20$ Punkte erhöht. Bei der nächsten Bekräftigung beträgt die Erhöhung $0,2 \times (100 - 20) = 16$ usw. Das heißt, die Assoziationsstärken wachsen mit der Annäherung an den

maximalen Wert immer langsamer. Wenn nach acht Versuchen der US nicht mehr dargeboten wird, sinkt die Assoziationsstärke entsprechend dem RWM zunächst schnell und nachfolgend immer langsamer, je näher sich die Assoziationsstärke dem Wert Null nähert. In der Abbildung wird zudem der Lernverlauf für einen CS_1 unter der Bedingung, dass er alleine konditioniert wird, mit dem Lernverlauf unter der Bedingung verglichen, dass ein zweiter Reiz

($CS_1 + CS_2$) gleichzeitig dargeboten wird. Man sieht, dass entsprechend dem RWM der Lernverlauf für CS_1 verzögert wird, wenn ein zweiter Prädiktor in Konkurrenz tritt (**► Abschn. 2.6.2 „Overshadowing“**). Bei der ersten Bekräftigung steigen die Assoziationsstärken noch um $20 (0,2 \times 100)$. Bei der nächsten Bekräftigung steigen die Assoziationsstärken aber nur noch um $0,2 \times (100 - 20) = 12$ usw.

0	$\alpha \times (\lambda - V)$	ΔV	US	V
1	$0,2 \times (100 - 0)$	20	+	20
2	$0,2 \times (100 - 20)$	16	+	36
3	$0,2 \times (100 - 36)$	12,8	+	48,8
4	$0,2 \times (100 - 48,8)$	10,24	+	59,04
5	$0,2 \times (100 - 59,04)$	8,19	+	67,23
6	$0,2 \times (100 - 67,23)$	6,55	+	73,78
7	$0,2 \times (100 - 73,78)$	5,24	+	79,02
8	$0,2 \times (100 - 79,02)$	4,19	+	83,22

9	$0,2 \times (0 - 83,22)$	-16,44	-	66,58
10	$0,2 \times (0 - 66,58)$	-13,32	-	53,26
11	$0,2 \times (0 - 53,26)$	-10,65	-	42,61
12	$0,2 \times (0 - 42,61)$	-8,52	-	34,09
13	$0,2 \times (0 - 34,09)$	-6,82	-	27,27
14	$0,2 \times (0 - 27,27)$	-5,45	-	21,82
15	$0,2 \times (0 - 21,82)$	-4,36	-	17,45
16	$0,2 \times (0 - 17,45)$	-3,49	-	13,96
17	$0,2 \times (0 - 13,96)$	-2,79	-	11,16

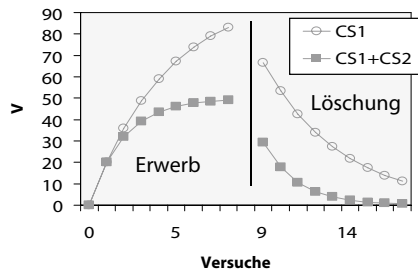


Abb. 2.6 Erwerb und Löschung eines CS nach dem RWM (für $\alpha=0,2$ und $\lambda_{\max}=100$). Die Tabellen zeigen die jeweilige inkrementelle Veränderung (ΔV) und die resultierenden Assoziationsstärken (V) für einen CS; links für acht aufeinanderfolgende Bekräftigungsversuche (US+) und rechts für neun nachfolgende Lösungsversuchen (US-). Die Grafik vergleicht den in den Tabellen dargestellten Verlauf der Assoziationsstärke für einen Reiz (CS_1) mit einer Situation, in der ein weiterer Reiz gleichzeitig konditioniert wird ($CS_1 + CS_2$, in den Tabellen nicht dargestellt)

2.6.2 Modellerklärungen

Overshadowing

Da die insgesamt erreichbare Assoziationsstärke zu einem US begrenzt ist (mehr als eine 100 %ige Vorhersage ist nicht möglich), müssen Reize, die gleichermaßen den US vorhersagen, die Assoziationsstärke untereinander aufteilen. Wenn also z. B. eine bevorstehende Fütterung durch einen Ton und ein Licht stets gemeinsam verlässlich angekündigt wird, dann sollte jeder der beiden Reize eine weniger starke Assoziation zur unbedingten Reaktion ausbilden, im Beispiel also eine weniger starke Speichelreaktion auslösen, als wenn er nur alleine konditioniert worden wäre (► Exkurs „Aufbau und Löschung eines bedingten Reflexes nach dem Rescorla-Wagner-Modell für nur einen und für zwei gleichzeitig dargebotene CS“). Pavlov hat dieses Phänomen des Overshadowings, wie es später genannt wurde, als erster beobachtet, und es ist nachfolgend immer wieder berichtet worden: Werden mehrere Reize gleichzeitig konditioniert, dann „**überschattet**“ die Assoziationsstärke jedes einzelnen Reizes die Assoziierbarkeit der anderen Reize.

Blockierung

Wenn ein erster Reiz CS_1 vollständig konditioniert wurde, dann entspricht seine Assoziationsstärke zum US der maximal erreichbaren Assoziationsstärke ($V_1 \cong \lambda^{\max}$). Wenn ein so bereits konditionierter Reiz nachfolgend zusammen mit einem zweiten Reiz CS_2 wiederholt mit dem US angeboten wird, dann ergibt nach dem RWM die Differenz zwischen der Summe der Assoziationsstärken beider Reize ($\Sigma V_j = V_1 + V_2$) und der maximal erreichbaren Assoziationsstärke λ^{\max} einen Wert von Null, sodass für den neu hinzukommenden Reiz CS_2 keine Assoziation aufgebaut wird. Ein bereits vollständig konditionierter Reiz **blockiert** so die Konditionierung jedes weiteren Reizes selbst dann, wenn dieser Reiz ein zuverlässiger Prädiktor des US ist.

Konditionierte Hemmung

Wird zusammen mit einem bereits konditioniertem Reiz CS_1 ein weiterer Reiz CS_2 immer dann präsentiert, wenn der US nicht dargeboten wird, erwirbt der

zusätzliche Reiz CS_2 die Tendenz, die konditionierte Reaktion zu hemmen. Konditionierte Hemmung ergibt sich nach dem RWM aus folgendem Zusammenhang: Für den bereits konditionierten CS_1 ist eine hohe Assoziationsstärke anzunehmen, während die Assoziationsstärke des neu hinzukommenden Reizes CS_2 nahe Null zu kalkulieren ist, da mit ihm noch keinerlei Konditionierungserfahrungen gemacht wurden. Bei Nichtdarbietung des US ergibt sich somit eine relativ hohe Differenz zwischen dem Wert Null (für λ^{\max}) und der Summe der aktuellen Assoziationsstärken und damit eine relativ hohe Reduktion der Assoziationsstärken beider Reize. Da die aktuelle Assoziationsstärke für den CS_2 nahe Null liegt, sinkt sie durch den Abzug unter Null und wird negativ. Die Darbietung eines Reizes mit negativer Assoziationsstärke senkt in Kombination mit beliebigen Reizen zwangsläufig die Summe der aktuellen Assoziationsstärken und muss damit die konditionierte Reaktion auch auf Reize hemmen, mit denen er noch nie gemeinsam dargeboten wurde – ein Effekt der, wie wir oben diskutiert haben, in Summationstests wiederholt bestätigt wurde.

2.6.3 Bewertung des RWM

Das RWM ist noch heute, mehr als 40 Jahre nach seiner Veröffentlichung, ein weithin akzeptiertes Modell assoziativer Lernprozesse (Miller et al. 1995). Seine Akzeptanz beruht vor allem darauf, dass einige wenige plausible Grundannahmen in einer einfachen Formel zum Ausdruck gebracht werden. Das ist nicht nur elegant, sondern es erlaubt vor allem, präzise und damit überprüfbare Voraussagen abzuleiten. Darüber hinaus wurden, wie wir im vorigen Abschnitt beispielhaft gezeigt haben, erstaunlich viele dieser Voraussagen durch experimentelle Untersuchungen bestätigt. Dies spricht dafür, dass das RWM grundlegende Eigenschaften assoziativer Lernmechanismen korrekt widerspiegelt: Dazu gehören vermutlich der inkrementelle Auf- und Abbau assoziativer Verbindungen, der asymptotische Verlauf des Lernfortschritts und schließlich die Feststellung, dass assoziative Bindungen zum gleichen US untereinander in Konkurrenz stehen.

Es gibt aber auch Beobachtungen, die das RWM nicht erklären kann. Dazu gehört das **Patterning**

(► Abschn. 2.4.2). Beim positiven wie beim negativen Patterning wird, wie wir gesehen haben, jeder der beiden zu konditionierenden Reize gleichhäufig bekräftigt und nicht bekräftigt. Das RWM sagt unter diesen Bedingungen voraus, dass die bei Bekräftigungen aufgebauten Assoziationsstärken bei Nichtbekräftigung wieder abgebaut werden und somit keine Konditionierung stattfindet. Die Tiere lernen aber, entweder nur auf beide Reize gemeinsam (positives Patterning) oder aber nur auf die Einzelreize zu reagieren (negatives Patterning). Um Patterning mit dem RWM erklären zu können, müsste man annehmen, dass neben den einzelnen Reizen auch Reizmuster unabhängige Assoziationen eingehen können. Bei zwei Reizen CS_1 und CS_2 würde dann auch deren gemeinsames Auftreten $CS_1 + CS_2$ als unabhängiger Reiz konditioniert bzw. nicht konditioniert werden können. Diese Zusatzannahme stößt allerdings schnell an Grenzen, denn die Zahl der möglichen Reizmuster steigt mit der Zahl der zu betrachtenden Einzelreize (n) nach der Formel 2^n exponentiell an (die Mächtigkeit des sog. Power Set). Bei der Vielzahl der unter natürlichen Bedingungen wirkenden Reize ist es äußerst unwahrscheinlich, dass alle möglichen Reizkombinationen unabhängig voneinander konditioniert werden.

Ein weiteres Beispiel, für das das RWM keine Erklärung liefert, ist die **latente Hemmung**: Die wiederholte Darbietung eines Reizes bereits vor dem ersten Konditionierungsversuch erschwert seine nachfolgende Konditionierung. Das lässt vermuten, dass der Reiz in die gewohnte Umgebung des Tieres integriert und demzufolge nicht mehr beachtet wird. Für eine solche Veränderung der Beachtung von Reizen bietet das RWM keinen Mechanismus an. Nach dem RWM sollte die Darbietung eines neuen Reizes allein überhaupt kein Lernen aktivieren: Ein neuer Reiz lässt per se nicht das Auftreten irgendeines unbedingten Reizes erwarten, und wenn dann auch kein unbedingter Reiz eintritt, wird die Erwartung bestätigt. Da nach dem RWM aber nur dann gelernt wird, wenn Erwartungen verletzt werden, sollte für einen einfach nur neu auftretenden Reiz nichts gelernt und seine spätere Konditionierbarkeit auch nicht beeinflusst werden.

Beide Beispiele beziehen sich auf die Wirkung zu konditionierender Reize und verweisen damit auf ein grundsätzliches Erklärungsdefizit des RWM: Das

Modell bietet keinen Mechanismus zur Bestimmung der in das Lernen einzubeziehenden Reize an. Nach dem RWM ist die grundsätzlich treibende Kraft des Lernens die Verbesserung der Vorhersage des US bzw. der Bekräftigung – gelernt wird nur, so die Kernaussage, wenn Vorhersagen verletzt werden. Der Grad der Sicherheit der Vorhersagen und damit die resultierenden Reaktionsstärken werden dabei vollständig durch die Summe der Assoziationsstärken aller aktuell gegebenen Reize bestimmt. Der Frage aber, wie die Reize bestimmt werden, von denen Vorhersage und Verhalten abhängig gemacht werden, wird nicht nur *nicht* nachgegangen, sondern sie wird gar nicht aufgeworfen.

Für Konditionierungsexperimente, in denen die Reize durch den Experimentator kontrolliert werden, ist dies in der Regel kein Problem. Unter natürlichen Bedingungen stehen jedoch stets unabzählbar viele Reize für eine Konditionierung zur Verfügung, und es ist durchaus ein Problem, aus dieser Vielfalt die Reize zu bestimmen, die in das Lernen einbezogen werden: Sind z. B. nur für Einzelreize oder auch für Reizmuster Assoziationsstärken zu verändern? Sind auch Reize, die zur gewohnten Umgebung gehören, in das Lernen einzubeziehen? Aufgrund welcher Merkmale werden Kategorien von Reizen gebildet? Auf diese und verwandte Fragen hat das RWM keine Antwort. Die Frage nach der Differenzierung verhaltensrelevanter Reize ist von grundsätzlicher Bedeutung. Wenn Lernen als Bildung von Reiz-Reaktions-Assoziationen und/oder von Assoziationen zwischen Reizen definiert wird, müssen die Reize, zu denen die Assoziationen aufgebaut werden, bestimmt werden können. Wenn sie nicht bestimmt werden können und wenn auch nicht die Faktoren bekannt sind, die auf ihre Auswahl Einfluss nehmen, kann Lernen nicht oder wenigstens nicht vollständig als Bildung von reizbezogenen Assoziationen beschrieben werden.

2.7 Fazit

Tierisches Verhalten passt sich auf dreierlei Weise an gegebene Umweltbedingungen an: Verhalten, das durch bestimmte Reize reflektorisch ausgelöst wird, kann erstens an neue Reize gebunden werden, wenn diese verlässliche Prädiktoren der ursprünglichen

Auslösebedingungen sind. Das ist klassisches Konditionieren. Zum zweiten werden neue Verhaltensweisen erworben, wenn diese verlässlich zu Bekräftigungen führen. Das ist instrumentelles Konditionieren. Ist das Erreichen der Bekräftigung von Situationsbedingungen abhängig, lernen die Tiere schließlich drittens zwischen den verhaltensrelevanten Situationen zu unterscheiden. Das ist Diskriminationslernen. Aus behavioristischer Sicht liegt allen drei Lernvorgängen die Bildung bzw. Modifikation von reizgebundenen Assoziationen zugrunde: CS-US-Assoziationen bei der klassischen Konditionierung und S-R-Assoziationen beim instrumentellen Konditionieren und beim Diskriminationslernen. Die reizgebundenen Assoziationen bilden sich jedoch nicht zwangsläufig. Ihre Ausbildung erfordert vielmehr eine aktive Beteiligung der Tiere: Eine aktive Exploration der Umgebung durch das Ausprobieren von möglichen Verhaltensweisen, eine aufmerksame Beachtung von möglicherweise verhaltensrelevanten Situationsbedingungen sowie eine kontinuierliche Aktualisierung von Erwartungen über das Eintreten von positiven wie negativen Bekräftigungen. Gelernt wird nur, wenn Erwartungen nicht bestätigt werden. Lernvorgänge werden danach wesentlich durch ein Streben nach Vorhersage von zu erwartenden Bekräftigungen getrieben. Allerdings geht es bei den behavioristischen Modellüberlegungen allein um eine Vorhersage des Grades von „Belohnung“ und/oder „Bestrafung“. Die konkreten Verhaltenskonsequenzen, also das, was mit dem Verhalten tatsächlich erreicht wird, spielt für das Lernen dagegen keine Rolle. Dies ist, wie wir im nächsten Kapitel zeigen werden, ein grundsätzliches Defizit aller behavioristischen S-R-Theorien.

? Kontrollfragen

1. Was sind Anliegen und grundlegende Annahmen des Behaviorismus?
2. Unter welchen Bedingungen wird ein Reiz konditioniert?
3. Welche Zeitverhältnisse zwischen CS und US sind für die Ausbildung eines bedingten Reflexes besonders günstig? Was sagt dies über die Funktion bedingter Reflexe aus?
4. Was beinhaltet das „law of effect“ von Thorndike?
5. Was sind sekundäre Verstärker?
6. Was versteht man unter „positivem Patterning“ und „negativem Patterning“?
7. Sind Tiere, wie z. B. Tauben, in der Lage, Konzepte zu bilden?
8. Unter welchen Bedingungen kommt es zu einer Blockierung der Ausbildung eines bedingten Reflexes für einen CS, der den US verlässlich prädiziert?
9. Was wird im Zusammenhang mit der Ausbildung bedingter Reflexe unter „preparedness“ verstanden?
10. Auf welcher Grundannahme beruht das Konditionierungsmodell von Rescorla und Wagner?

Weiterführende Literatur

- Kiesel, A. und Koch, I. (2012). *Lernen, Grundlagen der Lernpsychologie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mazur, J. E. (2006). *Lernen und Verhalten*. München: Pearson.
- Pearce, J. M. (1997). *Animal learning and cognition – an introduction*. Hove: Psychology Press.

Lern- und Gedächtnispsychologie

Hoffmann, J.; Engelkamp, J.

2017, XVI, 223 S. 56 Abb. Mit Online-Extras., Softcover

ISBN: 978-3-662-49067-9