

10 „**\$\$\$Overshadowing\$\$\$**“ beim kausalen Lernen

Michael Hildebrandt und Michael R. Waldmann

Kurzbeschreibung

Lernen wir beim Erwerb von Wissen über kausale Zusammenhänge allein assoziative Beziehungen zwischen den Lernereignissen (**\$\$\$Assoziationstheorien\$\$\$**) oder nutzen wir beim Lernen unser Wissen über kausale Eigenschaften von Ereignissen, wie etwa die Gerichtetheit kausaler Beziehungen (**\$\$\$Theorie der Kausalmodelle\$\$\$**)? Das Experiment prüft diese beiden konkurrierenden Positionen anhand einer klassischen Lernaufgabe aus der Assoziationspsychologie, dem „**\$\$\$Overshadowing\$\$\$**“-Paradigma.

Stichwörter: „**\$\$\$Overshadowing\$\$\$**“, **\$\$\$Assoziationstheorie\$\$\$**, **\$\$\$Theorie der Kausalmodelle\$\$\$**

10.1 Einleitung

Die Fähigkeit, Ereignisse vorhersagen zu können, gehört zu unseren wichtigsten kognitiven Kompetenzen, ohne die Organismen nicht überleben könnten. Weiß eine Maus etwa, dass der Schatten einer Katze Gefahr bedeutet, dass ein bestimmter Geruch auf Gift in der Nahrung hinweist oder dass sie durch eine bestimmte Aktion Zugang zu Futter erhält, dann heißt dies, dass sie irgendwann einmal Wissen über **\$\$\$statistische Beziehungen\$\$\$** zwischen den genannten Ereignissen erworben hat (vgl. Domjan, 1998). Die bisher genannten Beispielen sind vergleichsweise elementar. Es gibt aber auch Beispiele für komplexes statistisches Wissen, das wir im Alltag oder im Berufsleben erwerben. So lernen wir etwa, mit welchen Tasten man ein Videogerät einschalten kann, wie ein Computer funktioniert, warum wir bestimmte Krankheitssymptome bei uns beobachten oder, wenn wir psychologische Forschung betreiben, dass Aggression eine mögliche Folge von Frustration ist. Dies ist nur eine kleine Stichprobe von Ereignissen, die aufgrund einer kausalen Verbindung in statistischer Beziehung zueinander stehen.

Konkurrierende Lerntheorien

Die genannten Beispiele belegen, dass wir Wissen über **\$\$\$statistische Beziehungen\$\$\$** erwerben können, sie sagen aber nichts darüber aus, welche Lernprozesse beim Erwerb dieses Wissens eingesetzt werden. In diesem Kapitel sollen zwei konkurrierende Ansätze miteinander kontrastiert werden, **\$\$\$Assoziationstheorien \$\$\$** und die **\$\$\$Theorie der Kausalmodelle\$\$\$**:

\$\$\$Assoziations-theorien\$\$\$. Der dominierende Ansatz im Bereich Lernen nimmt an, dass wir *assoziative* Lernprozesse beim Erwerb von Wissen über \$\$\$statistische Beziehungen\$\$\$ zwischen Ereignissen einsetzen. \$\$\$Assoziations-theorien\$\$\$ gehen davon aus, dass es zwei Arten von Ereignissen gibt, \$\$\$Hinweisreize\$\$\$ bzw. \$\$\$Cue\$\$\$s, mit denen man Vorhersagen macht, und \$\$\$Kriteriumsreize\$\$\$ bzw. \$\$\$Outcome\$\$\$s, die man vorhersagen möchte. Solche \$\$\$Cue\$\$\$s können beispielsweise konditionierte Stimuli beim klassischen Konditionieren wie etwa ein Ton- oder ein Lichtreiz sein, während der vorherzusagende Kriteriumsreiz in diesem Fall ein unkonditionierter Reiz wie zum Beispiel ein Futterreiz oder ein Schockreiz sein könnte (vgl. Domjan, 1998; Gluck & Myers, 2001). Die \$\$\$Hinweisreize\$\$\$ können sich aber auch auf andere Ereignisse beziehen, wie etwa Krankheitssymptome, die man mit bestimmten Krankheitsdiagnosen verknüpfen lernt, oder Begriffsmerkmale wie „pelzig“, „klein“, „schwarz“ usw., die man mit Kategorienbezeichnungen (z.B. „Maus“) assoziieren lernt. Assoziative Beziehungen können \$\$\$statistische Beziehungen\$\$\$ zwischen allen Arten von Informationen ausdrücken, ungeachtet dessen, welche semantische Information die assoziierten Reize tragen.

\$\$\$Assoziations-theorien\$\$\$ gehen davon aus, dass Assoziationsgewichte zwischen \$\$\$Cue\$\$\$s und Kriteriumsinformationen erlernt werden, die das Ausmaß der statistischen Beziehung zwischen diesen Reizen ausdrücken. Ist beispielsweise ein Tonreiz ein guter Prädiktor von Futter, dann drückt sich dies durch ein starkes Assoziationsgewicht aus. Eine wichtige Funktion von assoziativen Netzwerken ist die Vorhersage von Kriteriumsinformationen. So kann es beispielsweise sein, dass ein Tier vorhersagen möchte, ob es Futter erwarten sollte, wenn es gerade einen Ton- und einen Lichtreiz erfahren hat. Liegt nur ein Reiz vor, wird einfach das entsprechende Assoziationsgewicht für die Vorhersage der Aktivierung des Outcomes herangezogen. So würde ein Gewicht von 1 etwa bedeuten, dass mit Sicherheit Futter erwartet wird ($\text{VorhergesagterOutcome}=1$), während Werte darunter (z.B. 0.5) höhere Grade von Unsicherheit ausdrücken. Liegen mehrere Reize vor, werden die entsprechenden Assoziationsgewichte einfach *summiert*. Dabei ist wichtig, dass nur diejenigen \$\$\$Cue\$\$\$s bei der Vorhersage des Outcomes verwendet werden, die beim jeweiligen Lerndurchgang („trial“) *anwesend* sind.

Wie werden nun die Assoziationsgewichte gelernt? Hierzu gibt es eine Reihe konkurrierender Ansätze (vgl. Domjan, 1998), im Augenblick dominiert aber eine Theorie, die den Lernprozess als Irrtumskorrektur-Lernen konzeptualisiert. Diese Theorie wurde unabhängig von Ingenieuren und Psychologen entwickelt und wird hier in vereinfachter Form dargestellt (nach Gluck & Myers, 2001).

Irrtumskorrektur-Lernen geht davon aus, dass das Ziel des Lernprozesses darin besteht, Fehler bei der Vorhersage zu minimieren. Dazu wird bei jedem einzelnen Lerndurchgang die durch die anwesenden \$\$\$Cue\$\$\$s vorhergesagte Aktivierung der \$\$\$Outcome\$\$\$s verglichen mit der Aktivierung, die aufgrund des Feedbacks im Versuch richtig gewesen wäre. Dieser Fehler wird als Differenz nach der Formel

$$\text{VorhergesagterOutcome} - \text{Outcome} = \text{GewünschterOutcome} - \text{Outcome} \quad (1)$$

berechnet. Dieser Vorhersagefehler ist nun der wesentliche Faktor, der das Ausmaß der Veränderung von Assoziationsgewichten bestimmt. Dabei wird bei jedem einzelnen Lerndurchgang der Vorhersagefehler (Outcome-Fehler) mit einem Lernparameter (Lernrate) multipliziert, der zwischen 0 und 1 liegen kann. Der resultierende Wert gibt dann an, um wie viel das alte Assoziationsgewicht verändert werden soll,

$$\text{NeuesGewicht} = \text{AltesGewicht} + (\text{Lernrate} \cdot \text{Outcome-Fehler}) \quad (2).$$

Angenommen, bei einem Lerndurchgang wird ein Tonreiz präsentiert, der mit einem Assoziationsgewicht von 0.4 mit dem Outcome verknüpft ist, der einen Futterreiz repräsentiert. Aufgrund der Aktivationsregel ist also der vorhergesagte Outcome des Futter-Ereignisses 0.4. Wird nun tatsächlich Futter dargeboten, dann bedeutet dies, dass eine perfekte Vorhersage dem Futter-Ereignis die Aktivierung von 1 (was Anwesenheit bedeutet) zugewiesen hätte. Der Vorhersagefehler (Outcome-Fehler) beträgt also 0.6 (Outcome-Fehler = 1 - 0.4). Nimmt man nun eine Lernrate von 0.2 an, dann bedeutet dies, dass das neue Gewicht der Summe von 0.4, dem alten Assoziationsgewicht, und dem Produkt von 0.2, der Lernrate, und 0.6, dem Vorhersagefehler, entsprechen würde. Das neue Gewicht würde also 0.52 betragen. Wie man sieht, nähert sich das Gewicht dem gewünschten Wert von 1 an, wobei der Lernzuwachs sowohl von der Lernrate als auch dem Ausmaß des Vorhersagefehlers abhängt.

Was bedeutet nun die Lernrate? Mit der Lernrate kann man zum Ausdruck bringen, dass nicht alle Lernereignisse gleich auffällig sind. So kann es sein, dass in einem Experiment ein schwacher Lichtreiz mit einem sehr lautem Tonreiz gepaart wird. Dies kann dazu führen, dass der Tonreiz stärker beachtet wird, so dass bei gleichartigen Lernereignissen dennoch die Beziehung zwischen Ton und Futter schneller gelernt wird als die zwischen Licht und Futter. Diesen Unterschieden kann man durch verschiedene Lernparameter (z.B. 0.4 für Ton, 0.2 für Licht) Rechnung tragen. Im Folgenden gehen wir aber davon aus, dass die verwendeten Reize im Experiment äquivalent sind, so dass man einen einheitlichen Wert für die Lernrate annehmen kann.

Bisher wurden die assoziativen Lernmechanismen an einem Beispiel zum klassischen Konditionieren (Ton-Futter-Assoziation) diskutiert, man kann aber auch Beispiele aus anderen Bereichen nehmen. Werden etwa Ursachen wie Nahrungszusätze mit Effekten wie etwa allergischen Erkrankungen assoziiert, dann würde man die Ursachen der Cue-Ebene und die Erkrankungen der Outcome-Ebene zuweisen. Man könnte aber auch eine Aufgabe wählen, bei der Effekte wie etwa Krankheitssymptome (z.B. Husten, Fieber) mit Ursachen wie etwa einer Virusinfektion assoziiert werden. In diesem Fall würden die Effekte der Cue-Ebene und die Ursache der Outcome-Ebene zugewiesen werden. Der erste Typ von Aufgabe (Ursache-Effekt-Lernen) wird *prädictives Lernen* genannt, der zweite (Effekt-Ursache-Lernen) *diagnostisches Lernen* (nach Waldmann & Holyoak, 1992). Es ist wichtig zu sehen, dass Assoziationstheorien keinen Unterschied machen zwischen diesen beiden

kausalen Lernformen. In beiden Aufgaben geht es um die Assoziation von $$$$Cue$$$$ s (Ursachen oder Effekte) mit $$$$Outcome$$$$ s (Effekte oder Ursachen).

$$$$Theorie der Kausalmodelle$$$$. Die Grundannahme dieser Theorie ist, dass Lernende in den meisten kausalen Lernsituationen Vorannahmen über den kausalen Status von Ereignissen haben und diese Vorannahmen beim Lernen nutzen (vgl. Waldmann, 1996, 1997; Waldmann & Holyoak, 1992). So gehen wir davon aus, dass Schalter in der Regel Ursachen des Aufleuchtens von Lämpchen oder das Blutbild ein Effekt eines Krankheitserregers ist, auch wenn wir ansonsten nichts über diese spezifischen Ereignisse wissen. Kausales Lernen ist dieser Theorie zufolge nicht ein Prozess, in dem beliebige $$$$Cue$$$$ s mit beliebigen $$$$Outcome$$$$ s assoziiert werden, sondern ein Vorgang, bei dem Lernereignisse in Verbindung gebracht werden mit hypothetischen Kausalmodellen. Abb. 10.1 zeigt zwei verschiedene Kausalmodelle, ein Gemeinsame Ursache-Modell, bei dem eine Ursache mehrere Effekte produziert, und ein Gemeinsamer Effekt-Modell, bei dem mehrere alternative Ursachen einen Effekt verursachen können. Ein Beispiel für ein Gemeinsame Ursache-Modell wäre etwa eine Erkrankung, bei der ein Virus mehrere Symptome wie Husten, Fieber und Erhöhung der Leukozyten verursacht. Ein Beispiel für ein Gemeinsamer Effekt-Modell könnte eine Erkrankung sein, die alternativ durch Bluttransfusion, ungeschützten Sexualverkehr oder durch kontaminierte Spritzen verursacht werden kann. Besteht die Lernaufgabe darin, die zwei Ereignisse unten (d.h. zwei Ursachen bzw. zwei Effekte) mit dem Ereignis oben (Effekt oder Ursache) zu assoziieren, dann würden $$$$Assoziationstheorien$$$$ keinen Unterschied zwischen den beiden Aufgaben machen (Shanks & Lopez, 1996). Im Gegensatz dazu sagt die $$$$Theorie der Kausalmodelle$$$$ vorher, dass Versuchsteilnehmer sensitiv sind für die Unterschiede zwischen den beiden Modellen, die durch die Umkehrung der $$$$Kausalrichtung$$$$ erzeugt werden.

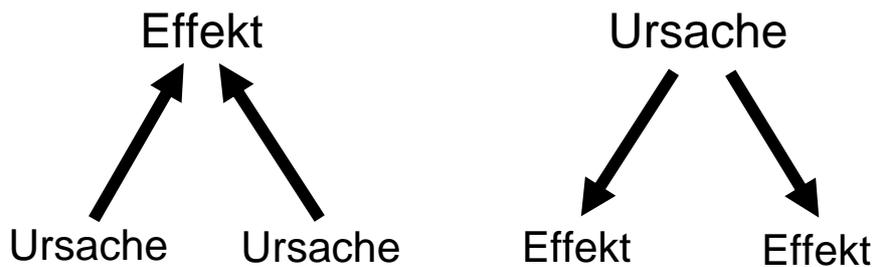


Abbildung 10.1 Beispiel für ein Gemeinsamer Effekt-Modell (links) und ein Gemeinsame Ursache-Modell (rechts).

Bei Kausalmodellen mit mehreren Ursachen wird der $$$$Theorie der Kausalmodelle$$$$ zufolge versucht, die statistische Beziehung zwischen jeder Ursache und dem Effekt abzuschätzen. Ein Maß für statistische Stärke ist das $$$$Kontingenzmaß$$$$ (ΔP , sprich: Delta-P), bei dem die Wahrscheinlichkeit, dass ein Effekt bei Anwesenheit der Ursache vorliegt ($P(E|U)$), verglichen wird mit der Wahrscheinlichkeit des Effekts in Abwesenheit der Ursache ($P(E|\sim U)$),

$$\Delta P = P(E|U) - P(E|\sim U) \quad (3).$$

Die beiden bedingten Wahrscheinlichkeiten werden auf der Basis der beobachteten relativen Häufigkeiten der Lernereignisse abgeschätzt. Das Kontingenzmaß ergibt einen positiven Wert, wenn die Ursache die Wahrscheinlichkeit des Effekts steigert, und einen negativen Wert, wenn sie sie senkt. Beträgt der Wert 1, dann handelt es sich um eine deterministische Beziehung, bei 0 liegt keine statistische Beziehung zwischen Ursache und Effekt vor.

Die Theorie der Kausalmodelle geht nun davon aus, dass Lernende versuchen, die Kontingenz in der Ursache-Effekt-Richtung abzuschätzen, gleichgültig, in welcher Reihenfolge die Lernereignisse dargeboten werden. Beim Gemeinsamen Ursache-Modell wird also die Kontingenz zwischen der gemeinsamen Ursache und jedem einzelnen Effekt getrennt geschätzt. Diese Kontingenz drückt das Ausmaß der Stärke der Kausalbeziehung aus. In manchen Aufgaben mag es allerdings gefordert werden, die *diagnostische* Beziehung zwischen Effekt und Ursache einzuschätzen (z.B. „Wie sicher ist es, dass eine Krankheit vorliegt, wenn Symptom A zu beobachten ist?“). Für diesen Fall geht die Theorie der Kausalmodelle davon aus, dass die Versuchsteilnehmer die Zielbeziehung vergleichen mit alternativen Ursachen des Effekts. Besteht die Aufgabe etwa darin, die diagnostische Beziehung von Fieber im Hinblick auf Grippe abzuschätzen, dann werden die Versuchsteilnehmer keine hohen Werte angeben, selbst wenn sie davon ausgehen, dass Grippe immer zu Fieber führt. Der Grund für die Absenkung der diagnostischen Urteile liegt daran, dass allgemein bekannt ist, dass es auch alternative Ursachen des beobachteten Symptoms Fieber gibt, so dass der diagnostische Schluss mit Unsicherheit behaftet ist.

Bei Gemeinsamer Effekt-Modellen werden Kontingenzen ebenfalls in der Ursache-Effekt-Richtung ermittelt. Hier ergibt sich aber das Problem einer potentiellen Konfundierung durch alternative Ursachen (vgl. auch den Beitrag von Waldmann in diesem Band). Beobachtet man beispielsweise eine starke positive Kontingenz zwischen Rauchen, einer Ursache, und Herzerkrankung, dem Effekt, dann muss man berücksichtigen, ob nicht alternative Ursachen, beispielsweise schlechte Ernährung, gleichzeitig vorliegen, die zu einer Scheinkorrelation zwischen Rauchen und Herzerkrankung führen können. Um diese mögliche Konfundierung zu kontrollieren, muss man den Ko-faktor (z.B. Ernährung) konstant halten, das heißt, man sollte Kontingenzen (Gleichung 3) berechnen für die Teilklasse der Beobachtungen, bei der die alternative Ursache (z.B. schlechte Ernährung) nach Möglichkeit abwesend ist. Eine Reihe von Studien hat gezeigt, dass Versuchsteilnehmer tatsächlich dazu neigen, Ko-faktoren konstant zu halten, wenn sie eine bestimmte Ursache-Effekt-Beziehung abschätzen (vgl. Waldmann, 1996, 1997).

10.2 Fragestellung und Hypothesen

Ziel des Experiments ist ein empirischer Test zwischen Assoziationstheorien des Kausallernens und der Theorie der

Kausalmodelle. Dabei steht die Frage im Mittelpunkt, ob Versuchsteilnehmer sensitiv sind für die Unterschiede von Kausalmodellen wie in Abb. 10.1, die im wesentlichen durch Unterschiede der kausalen Richtung zustande kommen, oder ob sie im Sinne von Assoziationstheorien Cues mit Outcomes ungeachtet der kausalen Rolle dieser Ereignisse assoziieren. Eine Möglichkeit für einen solchen Test bieten Aufgaben, bei denen die Probanden in den Vergleichsbedingungen gleiche Cues mit gleichen Outcomes assoziieren müssen, so dass Assoziationstheorien identisches Lernen vorhersagen. Der Unterschied zwischen den Bedingungen besteht lediglich darin, dass in eingangs gegebenen Instruktionen in einer Bedingung die Cues als Ursachen eines Effekts dargestellt werden (prädiktives Lernen eines Gemeinsamer Effekt-Modells), während in der zweiten Bedingung die gleichen Ereignisse als Effekte einer gemeinsamen Ursache charakterisiert werden (diagnostisches Lernen eines Gemeinsame Ursache-Modells).

Als Aufgabe sollen die Versuchsteilnehmer Informationen über das Verhalten von Schalter und Lämpchen einer artifiziellen Kiste erhalten (vgl. Waldmann, in press, Experiment 2). Auf der Vorderseite der Kiste finden sich drei Lämpchen, auf der Rückseite zwei (vgl. Abb. 10.2). In den beiden experimentellen Bedingungen erhalten alle Versuchsteilnehmer identische Informationen über die Lämpchen auf der Vorderseite (Cues) und müssen lernen, den Zustand der beiden Lämpchen auf der Rückseite (Outcomes) vorherzusagen. Da alle Versuchsteilnehmer gleiche Lerninformationen erhalten, sagen alle Assoziationstheorien gleichartiges Lernen vorher. Es sollen nun zwei Bedingungen verglichen werden, die sich lediglich dadurch unterscheiden, dass in den eingangs gegebenen Instruktionen die Lämpchen auf der Vorderseite entweder als Schalter, d.h. Ursachen, beschrieben werden und die Lämpchen auf der Rückseite als Effekte (prädiktives Lernen), oder die Lämpchen auf der Vorderseite werden als Effekte und die Lämpchen auf der Rückseite als Schalter, d.h. Ursachen, charakterisiert (diagnostisches Lernen). Zeigen sich im Lernen trotz identischer Lernereignisse charakteristische Unterschiede zwischen diesen beiden Bedingungen, dann spricht dies eher für die Theorie der Kausalmodelle als für Assoziationstheorien.

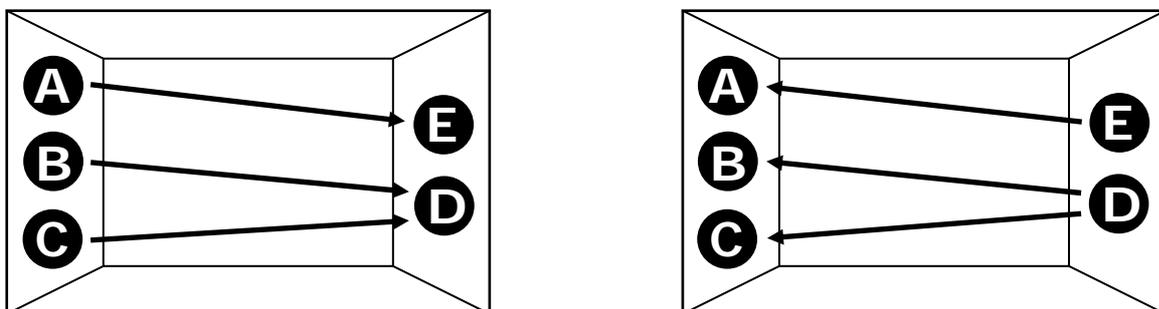


Abbildung 10.2: Realisierung der prädiktiven (links) vs. diagnostischen (rechts) Version eines „Overshadowing“-Paradigmas an einer Kiste

mit Lämpchen und Schaltern (nach Waldmann, in press, Experiment 2).

„*Overshadowing*“. Bisher wurde noch nichts über die Struktur der Lernaufgabe gesagt. In dem Experiment soll eines der zentralen Phänomene untersucht werden, die von aktuellen *Assoziationstheorien* vorhergesagt werden. Eine interessante Implikation von *Assoziationstheorien* ist, dass der Lernzuwachs und die maximale Assoziationsstärke einzelner *Cue*s nicht nur von der Anzahl der Paarungen mit dem *Outcome* abhängig ist, sondern auch davon, wie stark die Assoziationsstärke der *anderen Cue*s ist, die gleichzeitig vorliegen. Diese Eigenschaft des Wettbewerbs zwischen *Cue*s („*cue competition*“) ist eine Folge dessen, dass der Lernzuwachs für jedes einzelne Assoziationsgewicht eine Funktion des Fehlers bei der Vorhersage ist (Gleichung 1), wobei dieser Vorhersagefehler auf der Grundlage des Vergleichs der Summe der Assoziationsstärken *aller* anwesenden *Cue*s mit dem gewünschten *Outcome* berechnet wird.

Dies lässt sich am besten an Hand des Beispiels des „*Overshadowing*“ zeigen (vgl. auch Domjan, 1998). Bei diesem Paradigma gibt es zwei kritische Lernereignisse (vgl. Abb. 10.2): Zum einen wird ein einzelner *Cue* A (prädiktiver *Cue*) mit einem *Outcome* E gepaart. Daneben gibt es aber auch Lerndurchgänge, bei denen zwei weitere *Cue*s B und C (redundante *Cue*s) immer zusammen auftreten und mit einem anderen *Outcome* D gepaart werden. Man beachte, dass alle *Cue*s gleich häufig mit den ihnen zugeordneten *Outcome*s gepaart werden. Dennoch sagt die in Abschnitt 1 skizzierte *Assoziationstheorie* vorher, dass die Assoziationsstärken für die drei *Cue*s am Ende des Lernprozesses unterschiedlich hoch ausfallen sollten. Die Assoziationsstärke für *Cue* A sollte maximal (also 1) sein, weil dieser *Cue* alleine *Outcome* E vorhersagt und am Ende des Lernprozesses der gewünschte *Outcome* perfekt durch diesen einen *Cue* vorhergesagt wird. Dem gegenüber teilen sich die beiden anderen *Cue*s B und C die Vorhersage von *Outcome* D. Geht man von identischen *Lernrate*n aus, sollten sie sich also gemeinsam dem Zielwert annähern. Da die Vorhersage des *Outcome*s auf der Basis der *Summe* der Assoziationsstärken für B und C ermittelt wird, wird Lernen nur so lange stattfinden, bis diese Summe dem gewünschten Zielwert von 1 (für die Anwesenheit von *Outcome* D) entspricht. Das heißt, jeder *Cue* erhält nur die Hälfte des maximalen Assoziationsgewichts, was einem Wert von 0.5 für jeden *Cue* entspricht. Die beiden *Cue*s B und C, die individuell perfekte Prädiktoren von *Outcome* D wären, scheinen sich also gegenseitig zu überschatten. Zusammen sagen sie aber die Kriteriumsinformation perfekt vorher (*Outcome*-Fehler=0). Die Reduktion der Assoziationsstärke durch die Anwesenheit eines anderen redundanten *Cue*s, also der Unterschied zwischen 1 für den prädiktiven einzelnen *Cue* und 0.5 für jeden der beiden redundanten *Cue*s wird in der Literatur „*Overshadowing*“ genannt.

Während *Assoziationstheorien* einen „*Overshadowing*“-Effekt unabhängig davon, ob es sich um prädiktives oder *diagnostisches Lernen*

handelt, vorhersagen, macht die Theorie der Kausalmodelle unterschiedliche Vorhersagen für diese beiden Lernbedingungen. In der prädiktiven Lernbedingung repräsentieren die Cues Ursachen. Der einzelne prädiktive Cue kovariiert perfekt mit dem zugehörigen Effekt auf der Rückseite der Kiste mit einer Kontingenz (ΔP) von 1. Der Effekt tritt immer auf, wenn die Ursache vorhanden ist, aber nie, wenn sie abwesend ist. Anders sieht es bei den beiden redundanten Ursachen aus. Beide kovariieren zwar individuell perfekt mit dem gemeinsamen Effekt, bei multiplen Ursachen sollte man allerdings versuchen, Situationen für die Abschätzung der kausalen Wirksamkeit auszuwählen, in denen die alternative Ursache nach Möglichkeit abwesend ist, um Konfundierung zu vermeiden. Da beide Cues immer gemeinsam auftreten, gibt es eine solche Situation während der Lerndurchgänge nicht, so dass die Versuchsteilnehmer unsicher über den kausalen Status dieser Ursachen sein sollten. Es ist möglich, dass beide Cues Ursachen repräsentieren, es ist aber auch möglich, dass nur einer der beiden Cues kausal wirksam ist, oder dass die beiden Ursachen bei der Produktion des Effekts interagieren. Diese Unsicherheit sollte sich in reduzierten Einschätzungen der Stärke der Ursache-Effekt-Beziehung für die redundanten Cues relativ zum prädiktiven Cue widerspiegeln. In der prädiktiven Bedingung ist also ein Muster zu erwarten, das dem „Overshadowing“-Effekt entspricht, auch wenn von der Theorie der Kausalmodelle ein anderer Mechanismus als von Assoziationstheorien angenommen wird.

Anders sieht es allerdings in der diagnostischen Bedingung aus. Hier sollten die Probanden erkennen, dass eine Ursache (auf der Rückseite der Kiste) deterministisch ($\Delta P=1$) einen Effekt (prädiktiver Cue) und die andere Ursache deterministisch ($\Delta P=1$) jeden der beiden anderen redundanten Cues verursacht. Da keine alternativen Ursachen erwähnt werden, sollten die Versuchsteilnehmer dazu tendieren, jeden dieser Cues individuell als perfekten diagnostischen Indikator für die zugehörige Ursache anzusehen. Es sollte also zu keinen Unterschieden zwischen den Einschätzungen für den prädiktiven und die redundanten Cues kommen, was einer Abwesenheit eines „Overshadowing“-Effekts entspricht.

Der Vergleich der prädiktiven mit der diagnostischen Lernbedingung erlaubt es, die beiden konkurrierenden Theorien zu kontrastieren. Assoziationstheorien sagen bezüglich der Einschätzungen einen signifikanten Unterschied zwischen dem prädiktiven Cue und den redundanten Cues vorher („Overshadowing“-Effekt), wobei dieser Unterschied gleich hoch in den beiden Lernbedingungen ausgeprägt sein sollte (Haupteffekt). Die Theorie der Kausalmodelle hingegen sagt eine Interaktion zwischen den beiden Faktoren vorher. Während es in der prädiktiven Bedingung zu einem „Overshadowing“-Effekt kommen sollte, werden allgemein hohe Einschätzungen für den prädiktiven Cue und die redundanten Cues in der diagnostischen Bedingung erwartet, die sich nicht signifikant voneinander unterscheiden sollten.

10.3. Methode

10.3.1 Versuchsteilnehmer

Da die Situation in der prädiktiven Bedingung dem entspricht, was man in der Methodenlehre unter **\$\$\$Konfundierung\$\$\$** versteht, ist darauf zu achten, dass die Probanden nicht aktuell oder vor kurzem eine Methodenveranstaltung besucht haben, in der dieses Problem diskutiert wurde. Außerdem sollten die Probanden auch nicht bereits an ähnlichen Experimenten teilgenommen haben.

10.3.2 Versuchsmaterial

Der Versuch arbeitet mit der oben beschriebenen ‚artifiziellen Kiste‘. Im Versuchssteuerungsprogramm können die Häufigkeiten der einzelnen Ereignisse, d.h. die Kombinationen an- oder ausgeschalteter Lämpchen an der Vorder- und Rückseite, eingestellt werden (vgl. CD-ROM).

10.3.3 Versuchsablauf

Die experimentelle Manipulation des Faktors Lernbedingung (prädiktives vs. **\$\$\$diagnostisches Lernen\$\$\$**) wird durch die Eingangsinstruktion (siehe CD-ROM) realisiert. In der prädiktiven Bedingung erfahren die Probanden, dass sich an der Vorderseite der Kiste drei Schalter befinden, die zwei Lämpchen an der Rückseite anschalten. Über die Art der Verschaltung wird nichts genaueres gesagt. In beiden Gruppen wird erwähnt, dass auch die Schalter mit darüber liegenden Lämpchen ausgerüstet sind. Diese leuchten auf, wenn der jeweilige Schalter gedrückt wurde. Daher würde im Folgenden nur mehr von ‚Lämpchen‘ gesprochen (an der Vorder- oder Rückseite). Dieses Vorgehen ist notwendig, um in beiden Bedingungen absolut identisches Material in der Lernphase präsentieren zu können. Würde man die Lernphase zwischen den Bedingungen variieren, könnten Assoziationstheoretiker etwaige Unterschiede in den Daten auf diese Variation zurückführen.

Das Instruktionsverständnis ist von zentraler Bedeutung für den Erfolg des Experiments. Um dies zu überprüfen, fordert man die Probanden auf, die Instruktionen mit ihren eigenen Worten zusammenzufassen. Der Versuchsleiter muss beurteilen, ob der Proband den Zusammenhang und die **\$\$\$Kausalrichtung\$\$\$** hinreichend verstanden hat, was man an Äußerungen wie ‚verursacht‘ oder ‚schaltet an‘ erkennen kann. Das Experiment ist somit nicht ohne weiteres für den Einsatz als Web-Experiment geeignet, denn bei dieser Durchführungsart ist ja keine Versuchsleiterkontrolle gegeben.

Den Instruktionen folgt die Lernphase, die für beide Gruppen identisch ist. Die Probanden sehen zwölf Lerndurchgänge, in denen entweder kein Lämpchen, eines (prädiktiver **\$\$\$Cue\$\$\$**) oder zwei Lämpchen (redundante **\$\$\$Cue\$\$\$s**) auf der Vorderseite der Kiste an sind. Die Probanden sollen bei jedem Lerndurchgang

angeben, ob eines der beiden Lämpchen auf der Rückseite an ist oder beide Lämpchen aus sind. Je nach Bedingung entsprechen diese Lämpchen Ursachen oder Effekten der beobachteten Lämpchen auf der Vorderseite. Den Lernitems sieht man diese Unterschiede aber nicht an. Die kausale Bedeutung der Lämpchen wird aufgrund der eingangs gegebenen Instruktion von den Probanden an die Lernsituation herangetragen. Erfahrungsgemäß lernen die Probanden die Zusammenhänge sehr schnell.

Nach der Lernphase folgt die Testphase, in der die Versuchsteilnehmer jedes Lämpchen auf der Vorderseite einzeln sehen und die anderen Lämpchen abgedeckt sind. Die Probanden sollen nun für jedes Lämpchen einzeln auf einer Ratingskala, die von 0 („keine Vorhersage“) bis 10 („perfekte Vorhersage“) reicht, angeben, wie gut das jeweilige Lämpchen den Zustand jedes der beiden Lämpchen auf der Rückseite vorhersagt (vgl. Instruktion auf der CD-ROM).

10.3.4 Versuchsplan

Zusammenfassend handelt es sich also um ein zweifaktorielles Design mit zwei Ausprägungen der beiden Faktoren (2×2 -Design). Ein Faktor mit unabhängigen Versuchsgruppen bezieht sich auf die beiden Lernbedingungen „prädiktives“ und „diagnostisches Lernen“. Der zweite Faktor ist ein Messwiederholungsfaktor (bzw. „Within-subjects“-Faktor) „Overshadowing“ mit den Ausprägungen „prädiktiver Cue“ und „redundante Cue“. Als abhängige Variable dienen die Ratings der Probanden in der Testphase.

10.4. Auswertung

Abb. 10.3 zeigt beispielhaft die Ergebnisse von Experiment 2 aus Waldmann (in press). Die Werte für die redundanten Cues entsprechen jeweils dem Mittelwert der Ratings für die beiden Cues (in der Regel geben die Probanden für beide redundanten Cues identische Urteile ab).

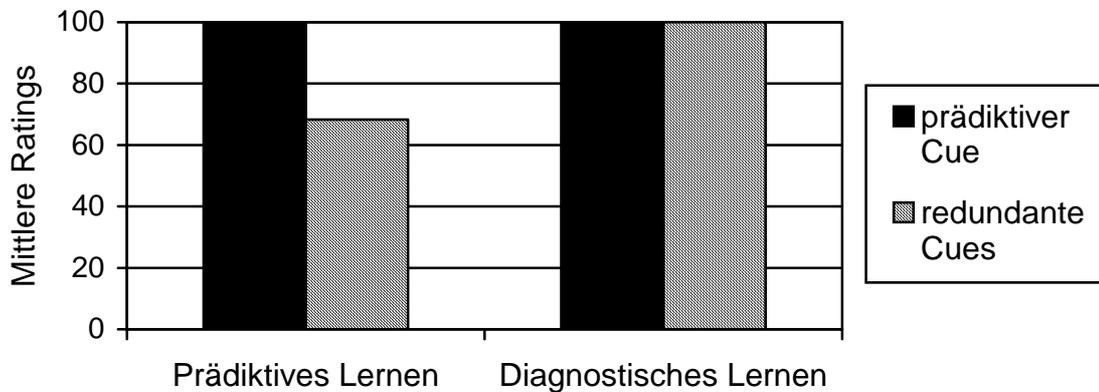


Abbildung 10.3: „*Overshadowing*“ beim prädiktiven vs. diagnostischen Lernen (nach Waldmann, in press, Experiment 2).

In der diagnostischen Bedingung, bei der die Lämpchen auf der Vorderseite als Effekte der Schalter auf der Rückseite dargestellt wurden, zeigte sich kein „*Overshadowing*“-Effekt: Sowohl der prädiktive *Cue* als auch die beiden redundanten *Cues* sagen nach Meinung der Probanden den jeweiligen *Outcome* auf der Rückseite perfekt vorher. In der prädiktiven Bedingung hingegen, bei der die Lämpchen auf der Vorderseite Ursachen repräsentieren, stellte sich für die redundanten *Cues* ein „*Overshadowing*“-Effekt ein: Viele Probanden sahen die redundanten *Cues* auf der Vorderseite nicht als perfekte Prädiktoren für den *Outcome* auf der Rückseite an, sondern beurteilten deren Vorhersagequalität geringer als die des prädiktiven *Cue*s. Eine 2×2 Varianzanalyse bestätigt diese Beobachtung durch eine hoch signifikante Interaktion zwischen den beiden Faktoren Lernbedingung und „*Overshadowing*“, $F(1, 22) = 22.2, p < .01$. Weitere geplante Vergleiche haben gezeigt, dass nur in der prädiktiven, aber nicht in der diagnostischen Bedingung ein signifikanter „*Overshadowing*“-Effekt zu beobachten war.

10.5. Fragen und Diskussionsthemen

Die Ergebnisse von Waldmann (in press) stützen die Annahmen der *Theorie der Kausalmodelle*, wonach kausales Lernen sich nicht auf den Erwerb von Assoziationen zwischen *Cues* und Kriterieninformationen reduzieren lässt, sondern Wissen über die kausale Struktur der Zusammenhänge den Lernprozess prägt. Sensitivität für den Aspekt der *Kausalrichtung* ließ sich auch bei vielen

anderen Fragestellungen nachweisen (für einen Überblick siehe Waldmann, 1996, 1997).

Anders als *Assoziationstheorien* beansprucht die *Theorie der Kausalmodelle* nicht, dass alle Arten von Lernen mit diesem Ansatz erklärbar sind. Sie ist vielmehr auf Aufgaben beschränkt, in denen es um den Erwerb von Wissen über kausale Beziehungen geht. Da solche Aufgaben prinzipiell auch mit anderen Lernmechanismen (z.B. assoziativen) bewältigt werden können, stellt sich die Frage, welche Randbedingungen darüber entscheiden, welchen Lernmodus die Probanden einsetzen. Denkbar wäre auch eine Mischung mehrerer Lernsysteme. Bislang ist nur sehr wenig über diese Frage bekannt. Eine mögliche Hypothese wäre, dass der Aufbau von Wissen über Kausalmodelle gebunden ist an Situationen, in denen die Probanden klare Intuitionen über die zugrundeliegende Kausalstruktur haben und die Aufgabe nicht allzu komplex ist. Insbesondere *diagnostisches Lernen* erfordert Prozesse, die die Informationsverarbeitungskapazität stark belasten können.

Weitere Fragen, die ebenfalls zukünftiger Forschung vorbehalten bleiben müssen, betreffen die Spezifikation der Prozesse, die beim Aufbau von Kausalwissen eingesetzt werden. Sind die gefundenen Effekte eher ein Resultat der Lernphase oder werden kausale Denkvorgänge erst im Anschluss an die Lernphase beim Abruf von Wissen eingesetzt? Mit welchen Lernmechanismen werden die für die Kausalmodelle relevanten Kontingenzen gelernt?

10.6 Extensionen

Mit der beschriebenen Versuchsanordnung läßt sich eine Reihe anderer klassischer Lernphänomene aus der Assoziationspsychologie (z.B. „Blocking“, „Overexpectation“) im Hinblick auf die Frage untersuchen, welche Theorie den Lernprozess am besten beschreibt (vgl. Domjan, 1998; Waldmann, 1996, 1997). Das Versuchssteuerungsprogramm läßt sich dazu einfach abwandeln.

Die artifizielle Kiste eignet sich besonders gut als Lernmaterial für die Untersuchung von Kausalmodellen. Es handelt sich um ein abgeschlossenes System, bei dem keine alternativen Ursachen oder Effekte außer den beobachteten auftreten. Die Probanden erstellen daher auch keine Hypothesen über mögliche weitere Einflüsse, die auf das System einwirken könnten. Die Rolle solcher Vorannahmen läßt sich gut untersuchen, indem man die Instruktionen und die Semantik der Lernereignisse variiert. So hat Waldmann (in press) in seinem Experiment 1 eine fast identische Versuchsanordnung eingesetzt, die Lernitems allerdings als Krankheiten und Substanzen im Blut dargestellt. Dabei wurden die Substanzen je nach Bedingung als Ursachen oder als Effekte einer Krankheit charakterisiert. Mit diesem Material zeigte sich zwar wiederum ein deutlicher Unterschied zwischen prädiktiver und diagnostischer Bedingung, allerdings auch ein kleiner „*Overshadowing*“-Effekt in der diagnostischen Bedingung. Waldmann (in press) erklärt dies damit, dass die Probanden weitere Annahmen über die kausale Struktur von Krankheiten an die Lernaufgabe herangetragen haben. Im Hinblick auf die in Abschnitt 10.5 skizzierten

Fragen wäre es auch interessant, mit verschiedenen Stufen der Komplexität der Aufgabe zu experimentieren. Das Versuchsprogramm lässt solche Variationen zu.

Literatur

- Domjan, M. (1998). *The principles of learning and behavior* (4th. ed.). Pacific Grove: Brooks/Cole.
Ausgezeichnetes Lehrbuch zum assoziativen Lernen.
- Gluck, M. & Myers, C. (2001). *Gateway to memory*. Cambridge: MIT Press.
Gute Einführung in das Irrtumskorrektur-Lernen und in die neuronalen Grundlagen des assoziativen Lernens.
- Shanks, D. R. & Lopez, F. J. (1996). Causal order does not affect \$\$\$cue\$\$\$ selection in human associative learning. *Memory & Cognition*, 24, 511-522.
Beispiel für die Position, dass Lernende den Aspekt der Kausalrichtung vollkommen vernachlässigen.
- Waldmann, M. (in press). Predictive versus diagnostic causal learning: Evidence from an \$\$\$overshadowing\$\$\$ paradigm. *Psychonomic Bulletin & Review*.
Basisartikel zum vorgestellten Experiment. Als PDF-Datei unter:
[www.psych.uni-goettingen.de/abt/1/waldmann/\\$\\$\\$Overshadowing\\$\\$\\$01.pdf](http://www.psych.uni-goettingen.de/abt/1/waldmann/$$$Overshadowing$$$01.pdf)
- Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 47-88). San Diego: Academic Press.
Überblick über Befunde, die die \$\$\$Theorie der Kausalmodelle\$\$\$ stützen.
- Waldmann, M. R. (1997). Wissen und Lernen. *Psychologische Rundschau*, 48, 84-100.
Einführung in die Theorie der Kausalmodelle.
- Waldmann, M. R. & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in \$\$\$cue\$\$\$ competition. *Journal of Experimental Psychology: General*, 121, 222-236.
Untersuchung der Sensitivität für die Kausalrichtung am Blockierungs-Paradigma.