

2 – 4 – 6

Wasons 2-4-6 Aufgabe

- Experimentator präsentiert der Vp eine Zahlensequenz (z.B. „2 4 6“), die einer gedachten Regel entspricht.
- Die Aufgabe der Vp ist es, eine Hypothese über die Regel zu generieren und diese dann zu testen, indem sie neue Zahlensequenzen produziert und sich vom Experimentator Rückmeldung geben lässt.
- Die Regeln sind bei dieser Aufgabe jeweils sehr allgemein gehalten (z.B. „Aufsteigende Zahlen“) und die erste Sequenz ist so gewählt, dass Vpn dazu tendieren, über-restriktive Regeln zu generieren (z.B. „Aufsteigende gerade Zahlen“)
- Vpn wurden instruiert, ihre Hypothese mitzuteilen, wenn sie sich relativ sicher sind, dass sie richtig ist.
- Konsequenz: Vpn, die eine Verifikationsstrategie anwenden (also Zahlensequenzen produzieren, die der Hypothese entsprechen), können ihre anfängliche über-restriktive Hypothese nie falsifizieren.

4-6-8 120 122 124 24 26 28 5002 5004 5006

2-4-6 Aufgabe

Anmerkung:

Die Regeln bei 2-4-6-Aufgaben sind Bi-Implikationen

z.B. „Wenn und nur wenn die Zahlensequenz aufsteigend ist, ist sie eine Instanz der gesuchten Kategorie“

P = Sequenz aufsteigend

Q = Sequenz ist Instanz der gesuchten Kategorie

$P \leftrightarrow Q$

Wasons 2-4-6 Aufgabe

- Generell zeigt sich in diesem Paradigma ein stark ausgeprägter *Verification Bias*. D.h. Vpn generieren Zahlensequenzen, die mit ihrer Hypothese in Einklang stehen, und bleiben deshalb bei ihren ursprünglichen über-restriktiven Hypothesen kleben.
- Einige wenige Vpn veränderten zwischendurch spontan ihre Hypothese.
- Insgesamt lösten jeweils nur etwa 20% der Vpn die Aufgabe. Diese 20 % gehörten allesamt zu der Gruppe der Vpn, die spontan Hypothesenvariationen vornahmen.
- Falsifikationsstrategien (Produzieren von Zahlensequenzen, die der Regel nicht entsprechen) wurden nicht verwendet.

Wasons 2-4-6 Aufgabe: Variationen

- Raten von gedachten Kategorien bei vorgegebenen Instanzen (z.B. „Mikrophon“ für „Unbelebte Objekte“) → *Verification Bias*
- Instruktion, alle Regeln zu finden, die korrekt sein könnten → *Verification Bias*
- Monetärer Abzug pro falsch angesagter Regel → Vpn sind etwas zögerlicher, aber immer noch *Verification Bias*

Positives vs. Negatives Testen

- Klayman & Ha (1987) schlagen vor, statt von Verifikations- und Falsifikationsstrategien von „positivem“ und „negativem Testen“ zu sprechen.
- Die positive Teststrategie besteht darin, Umstände zu untersuchen, in denen das Auftreten einer eines Ereignisses erwartet wird (laut Hypothese), um zu überprüfen, ob es tatsächlich eintritt / eingetreten ist.
- Die negative Teststrategie besteht darin, Umstände zu untersuchen, in denen das Auftreten einer Eigenschaft oder eines Ereignisses nicht erwartet wird (laut Hypothese), um zu überprüfen, ob es tatsächlich nicht eintritt / eingetreten ist.
- In beiden Fällen kann es zu falsifizierender Information kommen. Die Begriffe „Falsifikationsstrategie“ bzw. „Verifikationsstrategien“ ist demnach irreführend.
- Die Ergebnisse der 2-4-6 Aufgaben können nur bedingt als Evidenz für Verifikationsstrategien gewertet werden, wohl aber als Evidenz für positives Testen.

Positives vs. Negatives Testen

- Beispiel: Meteorologen haben Hypothese darüber, wann Tornados auftreten (nämlich unter Wetterbedingung X)
- Positive Teststrategie: Prüfen ob in Wettersituationen von Typ X ein Tornado auftritt, bzw. ob Wettersituationen, in denen ein Tornado aufgetreten ist, vom Typ X sind.
- Negative Teststrategie: Prüfen ob in Wettersituationen von Typ Not-X auch wirklich kein Tornado auftritt, bzw. ob Wettersituationen, in denen kein Tornado aufgetreten ist, auch wirklich nicht vom Typ X sind.
- Welche Strategie eher falsifizierende Information zutage bringt, hängt von der Häufigkeit der betrachteten Ereignisse ab.
- Bei seltenen Ereignissen ist positive Teststrategie günstiger!
- Verhalten bei der 2-4-6 Aufgabe ist nicht irrational.

Positives vs. Negatives Testen

- Anmerkung: Die beiden Teststrategien funktionieren nur bei Bi-Implikationen

$$P \leftrightarrow Q$$

Bi-Implikation

$$P \rightarrow Q$$

$$Q \rightarrow P$$

$$\text{not-}P \rightarrow \text{not-}Q$$

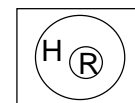
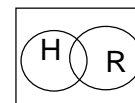
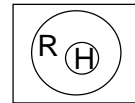
$$\text{not-}Q \rightarrow \text{not-}P$$

- Der positive Test überprüft „ $P \rightarrow Q$ “ und „ $Q \rightarrow P$ “
- Der negative Test überprüft „ $\text{not-}P \rightarrow \text{not-}Q$ “ und „ $\text{not-}Q \rightarrow \text{not-}P$ “

Positives vs. Negatives Testen: Klayman & Ha (1989)

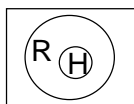
Vpn bekamen „2-4-6“-Aufgaben mit drei unterschiedlichen Regeltypen:

- *Embedded*: Anfängliche Hypothese ist in der tatsächlichen Regel enthalten (anfängliche Hypothese ist zu spezifisch).
Z.B. 2-4-6 / „aufsteigende Zahlen“
- *Overlapped*: Anfängliche Hypothese überschneidet sich mit Regel.
Z.B. 2-4-6 / Regel: „einstellige Zahlen“
- *Surrounded*: Regel ist in der anfänglichen Hypothese enthalten (anfängliche Hypothese ist zu generell).
Z.B. 2-4-6 / Regel: Aufsteigende Zahlen, die mit 2,4 und 6 enden.

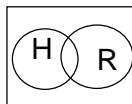


Positives vs. Negatives Testen: Klayman & Ha (1989)

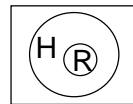
E:



O:



S:



Bei jedem Durchgang gab die Vpn ihre momentane Hypothese an, und generierte eine neue Zahlensequenz. →

Generell sehr viel positives Testen (67 %)

Alternative Hypothesen wurden in 22% der Fälle getestet (vor allem nachdem eine Hypothese viel positive Evidenz erhalten hatte).

Erfolgreiche Vpn wendeten mehr alternatives Testen an, als andere Vpn.

Bei Regeln vom Typ S (*Surrounding*) wurde wenig alternatives Testen verwendet, da hier positives Testen (wie erwartet) zu falsifizierender Information führte.

Single vs. Dual Rules

Regel-Entdeckungs-Leistungen bei der 2-4-6 Aufgabe können deutlich verbessert werden, wenn die Vpn angehalten werden, zwei Regeln zu finden, nämlich

- 1) die Regel, die die „Zielmenge“ beschreibt (z.B. „alle aufsteigenden Zahlenfolgen“) und
- 2) die Regel, die die „Komplementärmenge“ beschreibt (z.B. „alle nicht aufsteigenden Zahlenfolgen“)

→ Doppelt so viele Regel-Entdeckungen wie bei der klassischen 2-4-6 Aufgabe.

Dieses Ergebnis ist nicht so erstaunlich, da eine positive Teststrategie der Komplementärmengen-Hypothese einer negativen Teststrategie der Zielmengen-Hypothese gleichkommt (Wharton et al., 1993).

Single vs. Dual Rules

Beispiel: 2-4-6 Regel: Aufsteigende Zahlensequenz

Anfängliche Hypothese bezüglich Zielmenge: Aufsteigende Zahlen mit Abstand 2

Anfängliche Hypothese bezüglich Komplementärmenge: Zahlen, die nicht mit Abstand 2 aufsteigen

Vp schlägt als positiven Test für Komplementärmengen-Hypothese vor: 5-8-11

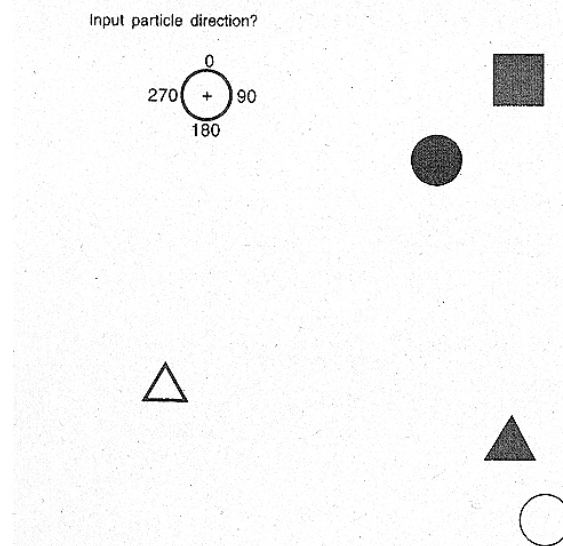
Experimentator klassifiziert diese Sequenz als zugehörig zur Zielmenge.

Sowohl die anfängliche Zielmengen-Hypothese als auch die anfängliche Komplementärmengen-Hypothese ist widerlegt.

Simulierte Forschungswelten: Mynatt et al., 1977

- Untersuchung des *Verification Bias* in komplexeren Umgebungen.
- Vpn saßen vor computer-generierten Displays mit Figuren unterschiedlicher Form und Helligkeit.
- Vpn konnten ein Geschoss abfeuern und dieses Geschoss wurde von einigen Figuren gestoppt von anderen nicht.
- Es galt herauszufinden, welche Figuren das Geschoss stoppen würden, und welche nicht.
- Anfänglich wurden die Vpn mit einer Anordnung trainiert, die eine falsche anfängliche Hypothese nahe legen sollte.
- Später konnten die Vpn dann aus verschiedenen Anordnungen auswählen, von denen einige geeignet waren, die anfängliche Hypothese zu widerlegen, andere nicht.

Versuchs- anordnung von Mynatt et al., 1977



Simulierte Forschungswelten: Mynatt et al., 1977

- Vpn wählten bevorzugt die Anordnungen, die vornehmlich eine Verifikation aber keine Falsifikation ihrer anfänglichen Hypothese erlaubte (→ *Verification Bias*)
- Wurde allerdings falsifizierende Information aufgedeckt, so benutzten die Vpn diese Information, um ihre Hypothesen zu revidieren.
- In einer späteren Untersuchung wurde gezeigt, dass falsifizierende Information in 33% der Fälle ignoriert wurde (Mynatt et al., 1978)

Simulierte Forschungswelten: Dunbar, 1993

- Sehr wirklichkeitsnahe Simulation Wissenschaftlicher Forschungssituation
- Vpn sollten mithilfe eines simulierten molekulargenetischen Labors eine Erklärung entwickeln, wie bestimmte Gene von anderen Genen kontrolliert werden (für diese Erklärung gab es 1961 einen Nobelpreis)
- Die tatsächliche Erklärung basiert darauf, dass zwei Gene ein anderes Gen, welches für die Produktion eines Enzyms zuständig ist, so lange hemmen bis dieses Enzym benötigt wird.
- Die Vpn wurden in Grundprinzipien der Genetik eingeführt und es wurden ihnen Beispiele für die Kontrolle von Genen durch andere Gene gegeben. Diese Beispiele basierten jedoch auf „Aktivierung“ nicht auf „Hemmung“ (d.h. die Vpn wurden hierdurch irregeleitet).

Simulierte Forschungswelten: Dunbar, 1993

- 4 von 20 Vpn lösten die Aufgabe
- Alle Vpn begannen mit einer Aktivationshypothese, unterschieden sich aber darin, wie sie auf Daten reagierten, die nicht zu ihrer Hypothese passten.
- Eine Untergruppe von Vpn versuchten daraufhin verstärkt Daten zu finden, die mit ihrer Hypothese in Einklang stehen würden.
- Eine andere Untergruppe von Vpn versuchten, Experimente zu machen, die die Abweichung zwischen Daten und Hypothese erklären helfen könnten.
- Die vier Vpn, die die Aufgabe schlussendlich lösten, gehörten zu der zweiten Gruppe.

Simulierte Forschungswelten: Dunbar, 1993

- Hypothese: Wenn die anfängliche Hypothese ein gewisses Maß an Bestätigung erhalten hat, so sind Vpn eher bereit, sich widersprüchlichen Datenpunkten zuzuwenden und Alternativhypothesen zu testen.
- Nachfolgeuntersuchung: Gleiches Paradigma, aber diesmal bestand die tatsächliche Erklärung aus Aktivierung und Hemmung.
- Ein Mechanismus (von insgesamt drei) entsprach also der anfänglichen Hypothese.
- Doppelt so viele Vpn lösten die Aufgabe.

Simulierte Forschungswelten: Dunbar, 1993

- Eine detaillierte Untersuchung der verbalen Protokolle ergab, dass die Hypothesen hierarchisch organisiert sind.
- Alle Vpn starten mit einer generellen Aktivationshypothese.
- Wenn diese Hypothese für das erste Gen falsifiziert wird, dann wird sie für das zweite Gen postuliert usw.
- Studie 1: Wenn die Aktivationshypothese für jedes Gen falsifiziert wurde, dann setzt sich ein Teil der Vpn das Ziel, eine Alternativhypothese zu finden, die die Daten erklären kann. Die anderen Vpn halten starr an ihrer Hypothese fest und versuchen weiter konsistente Daten zu finden.
- Studie 2: Nachdem die Aktivationshypothese für ein bestimmtes Gen akzeptiert wurde, wenden sich die Vpn den Daten zu, die nicht mit Aktivation erklärt werden können und versucht hierfür Alternativhypothesen zu generieren.

Tatsächliche Forschungssituationen

Untersuchungen, die sich mit dem Verhalten von Forschern in tatsächlichen Forschungssituationen beschäftigten, brachten ähnliche Ergebnisse zu tage, wie die Laborstudien zu diesem Thema.

Häufig wird nicht nach falsifizierender Information gesucht, sondern nach bestätigender Information.

Auch wird eine Theorie angesichts falsifizierender Information selten verworfen, sondern es wird versucht, Erklärungen für die abweichenden Datenmuster zu finden, die erlauben, die Theorie zu bewahren.

Beispiel: Vertreter der Newtonschen Theorie verwarfen die Theorie nicht, trotz offensichtlichen Fehlverhaltens einzelner Planeten (Uranus). Zur Erklärung des Fehlverhaltens wurde postuliert, dass es weitere bislang unentdeckte Planeten geben müsse, die für das Verhalten verantwortlich seien. Tatsächlich ist später ein solcher Planet entdeckt worden (Neptun). Andere Unstimmigkeiten wurden nie gelöst, aber die Theorie wurde dennoch nicht verworfen.

Tatsächliche Forschungssituationen

In einer Fragebogenuntersuchung von Mitroff (1974) mit 43 angesehenen Geologen, die am Apollo-Raumfahrtsprogramm teilnahmen ergab ähnliche Ergebnisse:

Forscher sahen ihr Forschungsziel nicht darin, ihre Theorien zu widerlegen, sondern eher darin bestätigende Information zu finden.

Die Theorien, die die Forscher versuchten zu falsifizieren waren die Theorien rivalisierender Forscher

Aber: Solange es rivalisierende Forscher mit unterschiedlichen Theorien gibt, regelt sich das Problem von selbst → dann ist nämlich die *scientific community* als Ganzes nicht voreingenommen (sondern nur einzelne Forscher – außerdem ist es ja unerheblich, wer bestehende Theorien zu widerlegen sucht.)

Fazit

Beim induktiven Schließen

- besteht eine Tendenz, Quellen potentiell falsifizierender Information zu ignorieren.
- besteht eine Tendenz, positiv statt negativ zu testen (wenn es sich bei den Regeln um Bi-Implikationen handelt).
- wird falsifizierende Information nur teilweise dazu genutzt, Theorien zu revidieren oder zu verwerfen.

Die Präferenz für positive Tests hat vermutlich damit zu tun, dass

- Positive Tests einfacher durchführbar sind (man testet Instanz von P und nicht eine der vielen Möglichkeiten von not-P)
- Positive Tests nur dann nicht zum Erfolg führen, wenn anfänglich eine Hypothese generiert wurde, die zu restriktiv ist.