



ÖFFENTLICHE UNTERRICHTSPLANUNG

Ist da jemand?

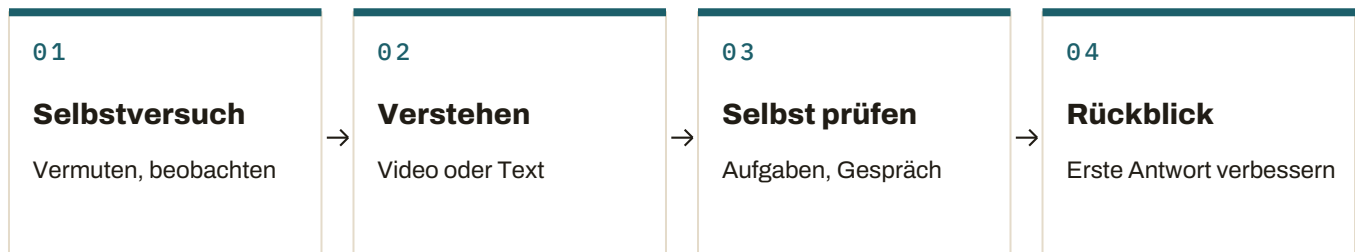
Bewusstsein bei Mensch, Tier und Maschine

Etappe 2 • Unterrichtsplan

Woher wissen Sie, dass jemand anderes etwas erlebt?

KLASSE 10–12 • 90 MINUTEN • IPAD

FASSUNG 1.2 • STAND 24. SEPTEMBER 2026



Für die Durchführung: Ablauf für 90 Minuten, Hinweise zu Lesefassung und Video, Tafelbild, Hinweise zum Selbstversuch »Die Dreiecke« mit seinem Ablauf und Differenzierung.

Zum Vergleichen: Lösungen mit Erläuterungen zu allen acht Aufgaben und zum Mill-Text, zwei Schülertexte zur Kalibrierung und Antworten auf Fragen, mit denen zu rechnen ist.

Etappe 2 ist mit Selbstversuch, Lesefassung und Aufgaben online; das Erklärvideo folgt nach der Audioaufnahme.

[Zur Etappe auf hoerstrup.com](https://www.hoerstrup.com)



Woher wissen Sie, dass jemand anderes etwas erlebt?

KURS »IST DA JEMAND? BEWUSSTSEIN BEI MENSCH, TIER UND MASCHINE« •

PHILOSOPHIE • KLASSEN 10–12

FASSUNG 1.2 VOM 24. SEPTEMBER 2026 • 90 MINUTEN • IPAD

Die Etappe beginnt mit einer kleinen Animation: Zwei Dreiecke und ein Kreis bewegen sich in und um ein Rechteck, und fast alle sehen darin eine Geschichte von Bedrängen, Hilfe, Flucht und Wut, obwohl die Formen nur einem Drehbuch aus Koordinaten folgen. Aus der Frage, woher die Geschichte kam, entwickelt die Stunde die Unterscheidung von Zuschreiben und Haben und dann eine unbequeme Frage: Worauf stützt sich unsere Gewissheit, dass die Person neben uns etwas erlebt? Zwei klassische Antworten der Philosophie, der Analogieschluss nach John Stuart Mill und der Schluss auf die beste Erklärung, werden auf ihre Stärke und ihre Grenze geprüft. Ein Fall aus der Medizin zeigt, was ein Indiz leisten kann und was nicht: eine Patientin im Wachkoma, die sich im Hirnscanner vorstellte, Tennis zu spielen. Am Ende steht der Turing-Test, den ein Sprachmodell in einer 2026 veröffentlichten Studie bestanden hat, und die Frage, was er über das Erleben zeigt.

Die Etappenseite enthält den vollständigen Lernweg. Dieser Plan ergänzt ihn für die Lehrkraft: Ablauf, Tafelbild, Hinweise zum Selbstversuch, Lösungen mit Erläuterungen, Differenzierung und Antworten auf Fragen, mit denen zu rechnen ist.

1. Auf einen Blick

Lernziel. Die Lernenden können das Problem des Fremdpsychischen in einem Satz formulieren, den Analogieschluss und den Schluss auf die beste Erklärung als zwei Antworten darauf prüfen, Zuschreiben von Haben unterscheiden und an Befunden aus dem Wachkoma und am Turing-Test sagen, was ein Indiz für Erleben zeigt und was nicht.

Kernbegriffe. Zuschreiben (ein Vorgang im Betrachter, der Bewegungen, Gesichter und Sätze als Zeichen von Absichten und Gefühlen deutet, schnell und fast von selbst) und Haben (eine Tatsache über das Wesen selbst: Es ist für das Wesen irgendwie, etwa wütend zu sein); das Problem des Fremdpsychischen (die Frage, woher ich weiß, dass andere etwas erleben, wenn ich ihr Erleben nie selbst wahrnehme, sondern nur ihren Körper, ihr Verhalten und ihre Sätze; nicht der Zweifel, ob es andere Menschen überhaupt gibt); Analogieschluss (aus einer Ähnlichkeit in Hinsichten, die man beobachten kann, auf eine Ähnlichkeit in einer Hinsicht, die man nicht beobachten kann; für das Fremdpsychische klassisch bei Mill 1865); Schluss auf die beste Erklärung (vom ganzen beobachteten Muster auf die Annahme, die es am besten erklärt) und



Konkurrenzerklärung (eine andere Annahme, die dasselbe Muster erklären könnte); Indiz und Beweis (ein Indiz ist ein Umstand, der für eine Annahme spricht, ohne sie sicher zu machen).

Drei Entscheidungen prägen die Etappe, und sie sollten im Unterricht erhalten bleiben. Die Etappe weckt keinen Zweifel an den Mitmenschen: Dass die Person nebenan etwas erlebt, glauben die Lernenden zu Recht, und gefragt wird nur, worauf sich diese Gewissheit stützt, damit man in den unklaren Fällen (Wachkoma, Tiere, Maschinen) begründet urteilen kann. Sie lässt die beiden Schlüsse nicht gegeneinander gewinnen, sondern stellt sie mit Stärke und Grenze nebeneinander; manche Philosophen halten beide für nötig (Melnik 1994; eine Übersicht über beide Antworten gibt Avramides 2023). Auch eine dritte Auffassung, nach der wir fremde Gefühle wahrnehmen, statt auf sie zu schließen, gilt in der Etappe nicht als Denkfehler; die Lesefassung nennt sie im Kasten »Zum Nachlesen«, die Rückmeldung in Aufgabe 2 nennt Scheler und Wittgenstein, und Abschnitt 7 gibt eine Antwort für das Gespräch. Und die Etappe bleibt in der Maschinenfrage offen: Gezeigt wird, dass die Sätze eines Sprachmodells wenig beweisen, nicht, dass es nichts erlebt.

Material. iPads mit der Etappenseite, möglichst dieselben Geräte und derselbe Browser wie in Etappe 1 (siehe Vorbereitung); Tafel; Kursnotizen (Dokument »Bewusstsein«, neuer Abschnitt »Etappe 2«); optional der Zahlenstrahl der Glaubenstafel aus Etappe 1 oder sein Foto und Klebepunkte in einer zweiten Farbe.

Vorbereitung.

1. Den Selbstversuch einmal selbst durchführen, mit Beschreibung und Nachmessen etwa zehn Minuten; die Animation allein dauert 62 Sekunden und braucht keinen Ton. Den Ablauf in Abschnitt 4 vorher lesen und die Nachmessung einmal genau ansehen, damit man im Blitzlicht auf die Bahnen und den Ausschnitt aus dem Drehbuch verweisen kann.
2. Die Glaubenstafel läuft in dieser Etappe zum zweiten Mal. Die Werte der ersten Runde liegen nur im Browser des Geräts, auf dem die Lernenden Etappe 1 bearbeitet haben, und nur dort erscheinen sie als Marken über den Reglern. Wer Gerät oder Browser gewechselt oder die Daten gelöscht hat, bekommt von der Seite einen Hinweis, stellt die Werte neu ein und vergleicht mit der Erinnerung oder dem Foto des Zahlenstrahls. Die Regler beginnen auch in der zweiten Runde bei »offen« und nicht beim alten Wert. Wer bei einer Einschätzung bleiben will, zieht den Regler deshalb genau auf die Marke E1, wie es die Anleitung sagt; liegt er einige Punkte daneben, zählt die Tafel das als Verschiebung.
3. Prüfen, ob die Schülerzugänge der Klasse freigeschaltet sind und die Einwilligung zur KI-Rückmeldung vorliegt. Ohne sie funktioniert Aufgabe 6 als Selbstprüfung an den vier Kriterien.
4. Der Fall im Scanner handelt von Menschen mit schweren Hirnverletzungen, und in Aufgabe 4 spricht ein Angehöriger. Wenn bekannt ist, dass jemand in der Klasse einen solchen Fall in der Familie erlebt, lohnt ein kurzes Gespräch vorher; die Etappe verlangt von niemandem, darüber zu sprechen.



2. Ablauf

ZEIT	SCHRITT	LERNENDE	LEHRKRAFT
0–5	Zum Ankommen: Aufgaben A1 und A2	öffnen die Etappe und lösen die beiden kurzen Aufgaben zu Etappe 1, zuerst ohne nachzuschlagen	beginnt ohne Einleitung; klärt nur dann kurz im Plenum, wenn viele an derselben Stelle scheitern (Abschnitt 5)
5–17	Selbstversuch und erste Antwort	vermuten, sehen die Animation, beschreiben sie in drei Sätzen, entscheiden beim Nachmessen über ihre Wörter, schreiben eine erste Antwort; danach kurzer Vergleich mit dem Nachbarn	sagt vorher nur: »Schauen Sie einfach zu, es gibt nichts falsch zu machen.« Nennt weder Heider und Simmel noch das Wort »Zuschreiben«; auch die Anleitung über dem Versuch verrät nichts. Achtet darauf, dass jede Person ihre Beschreibung schreibt, bevor sie die des Nachbarn liest.
17–20	Blitzlicht	Handzeichen: Wer hat eine Geschichte gesehen, wer nur Formen? Zwei, drei Beschreibungen werden freiwillig vorgelesen. Wie viele Wörter haben Sie gezählt, und bei welchem Wort fiel die Entscheidung schwer?	notiert die Verben aus den vorgelesenen Beschreibungen unkommentiert am Rand der Tafel (typisch: »jagt«, »flieht«, »hilft«, »will«, »ist wütend«) und fragt nur: Steht das im Drehbuch?
20–40	Den Gedanken verstehen	Lesefassung oder, sobald es da ist, Video (siehe unten), dabei Notizen in zwei Spalten »Analogieschluss« und »Schluss auf die beste Erklärung« und je ein Satz zum Scanner und zum Turing-Test; danach Mills Text, zu dem sie die drei Glieder seiner Reihe und eine Schwäche seines Schlusses notieren	achtet darauf, dass die Notizen beim Lesen entstehen und nicht erst am Ende; wer langsam liest, geht nach »Tennis im Scanner« zum Mill-Text über und liest den Rest zu Hause; wer früher fertig ist, beginnt mit Aufgabe 1
40–50	Aufgaben 1 und 2	bearbeiten die beiden Aufgaben und vergleichen danach kurz mit dem Nachbarn ihre Notizen zu Mills Reihe und ihre beiden Spalten	bei gehäuften Fehlern an einer Aufgabe kurze Klärung im Plenum (Abschnitt 5)
50–55	Sicherung	entwickeln mit der Lehrkraft den oberen Teil des Tafelbilds und übertragen ihn in die Kursnotizen	Tafelbild bis zu den beiden Spalten (Abschnitt 3); kehrt zu den Randverben zurück: Was davon ist Zuschreiben, und was müsste hinzukommen, damit daraus ein Grund für Haben wird?
55–65	Aufgaben 3 und 4	bearbeiten die beiden Aufgaben	wie bei den Aufgaben 1 und 2



ZEIT	SCHRITT	LERNENDE	LEHRKRAFT
65–73	Aufgabe 5: Tennis oder Haus	Gespräch zu zweit: A beantwortet drei Ja-Nein-Fragen von B nur in der Vorstellung, B rät; danach die Notiz »Von außen konnte ich ...«	achtet darauf, dass die Fragen harmlos bleiben
73–77	Bericht und unterer Teil des Tafelbilds	zwei Paare berichten, wie viele Antworten B getroffen hat und woran B sich orientiert hat	fragt die Klasse: Wenn ein Scanner die Antworten sicher zeigte, was wüssten wir dann über A und was nicht? Ergänzt dabei die Prüffrage und die Zeilen zu Indiz und Beweis.
77–80	Aufgabe 6 beginnen	lesen Material und Frage und notieren den Aufbau ihres Textes in vier Stichworten, einem für jedes Kriterium	gibt den Rest als Hausaufgabe: erste Fassung schreiben, prüfen lassen, Rückmeldung lesen und überarbeiten, bis zur nächsten Stunde
80–90	Rückblick	lesen ihre erste Antwort und verbessern sie in einem Satz, der einen der beiden Schlüsse nennt; stellen die Glaubenstafel neu ein und begründen die größte Verschiebung; beantworten die Reflexionsfrage	optional: Zahlenstrahl aus Etappe 1; jede Person klebt in der zweiten Farbe je einen Punkt für »Chatbot« und »Oktopus«, wer mag, auf einem dritten Strahl für »Mensch im Wachkoma«. Kurz fragen, wer einen Wert verschoben hat und aus welchem Grund.

Die Zeiten folgen der Etappenseite, mit drei Zusätzen. Die Seite rechnet mit 15 Minuten für den ersten Schritt, davon etwa 5 für die beiden Aufgaben zum Ankommen und 10 für den Selbstversuch mit dem Vergleich, dazu mit 20 Minuten für die Erklärung, zusammen 20 für die Aufgaben 1–4, 8 für Aufgabe 5, 12 für Aufgabe 6 und 10 für den Rückblick, insgesamt also mit 85 Minuten. Im Unterricht kommen das Blitzlicht (3 Minuten), die Sicherung (5 bis 7) und der Bericht nach Aufgabe 5 mit dem unteren Teil des Tafelbilds (etwa 4) hinzu. Außerdem braucht Aufgabe 6 mit Prüfung, Rückmeldung und Überarbeitung eher 15 Minuten als 12; so viel hat die Schreibaufgabe in Etappe 1 für 100 bis 220 Wörter, und hier sind es 120 bis 240 Wörter im Anforderungsbereich III. Alles zusammen ergäbe gut 100 Minuten. Der Plan gleicht das so aus: Der Selbstversuch bekommt zwei Minuten mehr als auf der Seite, das Blitzlicht füllt die übrigen drei der fünf freien Minuten, und Aufgabe 6 wird in der Stunde nur begonnen; die erste Fassung, die Rückmeldung und die Überarbeitung folgen bis zur nächsten Stunde. Das trifft die Aufgabe, die am wenigsten auf die Klasse angewiesen ist, und die Rückmeldung zu den vier Kriterien kommt auch zu Hause. Die Sicherung steht nach Aufgabe 2, weil diese Aufgabe den Mill-Text prüft und weil bis dahin auch die Lernenden aufgeholt haben, die lesen statt schauen.

Video und Lesefassung. Das Erklärvideo zu dieser Etappe folgt nach der Audioaufnahme und wird nach dem Sprechskript etwa neun Minuten dauern; bis dahin ist die Lesefassung der Zugang für alle. Sie hat rund 2.800 Wörter, etwas mehr als die der ersten Etappe (rund 2.400); dazu kommen etwa 170 Wörter in den



Abbildungen und gut 300 Wörter Mill-Text mit Einleitung. Bei 150 bis 200 Wörtern pro Minute dauert allein das Lesen 17 bis 22 Minuten, und dazu kommen sechs Abbildungen, drei davon zum Durchklicken, zwei Denkpausen und die Notizen. In 20 Minuten schaffen das nur schnelle Leserinnen und Leser. Wer langsamer liest, liest deshalb in der Stunde bis zum Ende von »Tennis im Scanner« (rund 2.100 Wörter) und dann den Mill-Text; die letzten beiden Abschnitte, »Das Imitationsspiel« und »Was wir wissen und woher« (zusammen knapp 700 Wörter), folgen zu Hause vor der ersten Fassung von Aufgabe 6. Die Aufgaben 1–5 brauchen diese Abschnitte nicht; die Konkurrenzerklärung für Sprachmodelle, auf die Aufgabe 3 zurückgreift, steht schon im Abschnitt über den Schluss auf die beste Erklärung. Die Zeile zum Turing-Test im Tafelbild ist für diese Lernenden eine Vorschau auf Aufgabe 6. Der Quellenauftrag verlangt nur die drei Glieder von Mills Reihe und eine Schwäche; die Standardform baut erst Aufgabe 2 mit Karten. Wer die Lesefassung als Zugang braucht, kann sie auch vorab als Vorbereitung bekommen, aber erst nach dem Selbstversuch, denn ihr erster Absatz nimmt dessen Auflösung vorweg.

Wenn die Zeit knapp wird, entfällt zuerst der Zahlenstrahl im Rückblick. Danach wird Aufgabe 3 zur Hausaufgabe, denn sie gibt zu jeder Einordnung eine Rückmeldung, und nach Aufgabe 5 berichtet nur ein Paar. Der Selbstversuch mit dem Blitzlicht, die Sicherung und das Gespräch in Aufgabe 5 sollten im Unterricht bleiben. Bleibt Zeit übrig, etwa weil viele schon das Video sehen können, schreiben die Lernenden die erste Fassung von Aufgabe 6 schon in der Stunde.

Als Blended-Learning-Variante lässt sich die Etappe teilen: Selbstversuch und Blitzlicht am Ende der vorigen Stunde (etwa 15 Minuten), Lesefassung oder Video und der Mill-Text zu Hause, in der Stunde dann die Aufgaben zum Ankommen, die Aufgaben 1–5 mit Sicherung und Bericht, Aufgabe 6 bis zur gelesenen Rückmeldung und der Rückblick. So passt auch die Schreibaufgabe ganz in die Stunde. Der Selbstversuch gehört aber in jedem Fall in den Unterricht und vor die Erklärung, weil der gemeinsame Moment, in dem die Klasse ihre eigenen Geschichten neben den nackten Bahnen sieht, die Frage erst zur Frage der Klasse macht.



3. Tafelbild

Woher wissen Sie, dass jemand anderes etwas erlebt?

ZUSCHREIBEN

ein Vorgang im Betrachter

»Ich sehe, dass es wütend ist.«

HABEN

eine Tatsache über das Wesen selbst

»Es ist wütend.«

Zuschreiben ist nicht Haben. Die Dreiecke folgen einem Drehbuch.

Das Problem des Fremdpsychischen: Mein Erleben kenne ich von innen.

Bei anderen sehe ich Körper und Verhalten und höre Sätze.

Woher weiß ich, dass sie etwas erleben?

ANALOGIESCHLUSS

(Mill 1865)

so geht er Sie ist wie ich, also
erlebt sie wie ich.

stark, weil die Ähnlichkeit groß ist
(Körper, Gehirn,
Verhalten, Sprache)

Grenze ein einziger Fall (ich);
schwächer, je unähnlicher

SCHLUSS AUF DIE BESTE ERKLÄRUNG

Nichts erklärt das ganze Muster
so gut wie ein Erleben.

sich die Konkurrenz prüfen lässt

nur so stark, wie die Erklärungen
ohne Erleben schwach sind

Prüffrage: Gibt es eine ebenso gute Erklärung ohne Erleben?

die Person neben mir keine in Sicht

die Dreiecke das Drehbuch

ein Sprachmodell das Training

starker Schluss

kein Grund für Erleben

die Sätze entscheiden es nicht

Indiz, nicht Beweis

Scan mit Muster auf die Bitte hin

Scan ohne Reaktion

Turing-Test bestanden

starkes Indiz: versteht und will
beweist nicht, dass niemand da ist
Gespräch nicht zu unterscheiden;
für Erleben ein schwaches Indiz,
weil Training und Rolle es auch
erklären

Den oberen Teil bis zu den beiden Spalten entwickelt die Klasse in der Sicherung nach Aufgabe 2; die Prüffrage und die Zeilen zu Indiz und Beweis ergänzt die Lehrkraft beim Bericht nach Aufgabe 5. In der letzten Zeile steht mit Absicht »schwaches Indiz« und nicht »sagt nichts«: Aus einem schwachen Indiz ein wertloses zu machen, wäre der skeptische Kurzschluss (F-U2.9), und die Rubrik von Aufgabe 6 erwartet genau diese Gewichtung.



4. Hinweise zum Selbstversuch

Was geschieht. Die Animation ist eine eigene nach dem Vorbild eines berühmten Versuchs von Fritz Heider und Marianne Simmel (1944), kein Nachbau ihres Films. Ein großes Dreieck, ein kleines Dreieck und ein Kreis bewegen sich 62 Sekunden lang in und um ein Rechteck auf einer hellen Fläche. Ein Stück der rechten Seite des Rechtecks ist beweglich wie eine Tür; es steht lange offen, schließt sich gegen Ende und schlägt danach dreimal kurz nach innen aus. Es gibt keinen Ton, keinen Text und keine Farben, die Rollen verraten: Die Formen unterscheiden sich nur durch Form und Größe. Keine Form reagiert auf eine andere. Jede folgt einem Drehbuch aus Zeitpunkten, Orten und Winkeln, und die Bilder dazwischen rechnet die Seite aus. Die Bewegungen sind so entworfen, dass die meisten Menschen darin eine Geschichte sehen, und zwar in diesen Takten:

SEKUNDE	TAKT	WAS GESCHIEHT
0–8	Zuhause	Kleines Dreieck und Kreis stehen im Rechteck, die bewegliche Seite ist offen. Das kleine Dreieck geht voran hinaus, der Kreis folgt; draußen bleiben beide nebeneinander.
8–11	Auftritt	Das große Dreieck kommt vom oberen Rand herein, hält an und richtet die Spitze auf den Kreis. Das kleine Dreieck dreht sich zu ihm.
11–18,6	Bedrängen	Das große Dreieck stößt dreimal auf den Kreis zu (bei 13, 14,5 und 16 s), ohne ihn zu berühren. Der Kreis weicht bis in die Ecke rechts unten zurück und zittert dort, das kleine Dreieck weicht aus.
18–25,6	Ablenken	Das kleine Dreieck nähert sich dem großen von hinten. Das große dreht sich um (20,5 s) und folgt ihm die rechte Seite hinauf, die Spitze auf das kleine gerichtet. Oben biegt das kleine nach links ab, das große bleibt stehen (23,7 s). Unterdessen geht der Kreis zurück zum Rechteck.
25–27,3	Flucht	Das große Dreieck dreht sich zum Kreis. Der Kreis flieht durch die offene Seite ins Rechteck, das große folgt ihm hinein.
27,3–32,1	Verfolgung im Haus	Der Kreis läuft unten entlang und die linke Wand hinauf, das große Dreieck hinterher. In der Ecke oben links wird der Kreis gestellt und zittert.
32,1–36,3	Entkommen	Der Kreis täuscht nach unten an, läuft oben am großen Dreieck vorbei und durch die noch offene Seite hinaus (35,5 s).
36–42,6	Trennung	Das kleine Dreieck kommt dem Kreis entgegen. Als das große herauskommt, laufen die beiden auseinander, der Kreis nach rechts unten, das kleine Dreieck nach rechts oben. Das große bleibt zwischen ihnen stehen und dreht sich hin und her.



SEKUNDE	TAKT	WAS GESCHIEHT
42,6–51,4	Wut an der Tür	Das große Dreieck schlägt von oben auf die offene Seite (45,85 s), und sie schwingt zu. Dann stößt es dreimal von außen dagegen (48,55, 49,65 und 50,75 s), und die Seite schlägt jedes Mal nach innen aus.
49,6–55,8	Wiedersehen	Kreis und kleines Dreieck finden am rechten Rand zusammen (52,3 s) und rücken gemeinsam ein Stück nach unten.
52,3–62	Abwenden	Das große Dreieck dreht sich zu den beiden, schaut, wendet sich langsam ab und kommt ab 58,5 s zur Ruhe.

Die Takte überlappen, weil mehrere Formen zugleich unterwegs sind. Auffällig ist, dass auch diese Tabelle kaum ohne Zuschreibungen auskommt: »flieht«, »täuscht an«, »Wut«. Im Drehbuch steht keines dieser Wörter. Das lässt sich im Blitzlicht nutzen: Die Tabelle in der Nachmessung zeigt das große Dreieck zwischen 12,1 und 16 Sekunden, also genau beim Bedrängen, und dort steht der Vorstoß nur als Zahlen: Bei 13, 14,5 und 16 Sekunden wachsen x und y deutlich, dazwischen gehen sie ein wenig zurück, und der Winkel dreht sich langsam mit. Ein Standbild aus der Geschichte bietet die erste Abbildung der Lesefassung, »Geschichte und Geometrie«. Sie zeigt den Augenblick bei 19,6 Sekunden, in dem der Kreis in der Ecke steht, das große Dreieck vor ihm und das kleine dicht dahinter. In der ersten Ansicht stehen daneben gestrichelte Fenster mit den Wörtern »bedrängt«, »hat Angst« und »hilft«, in der zweiten nur die Bahnen der Mittelpunkte von Sekunde 12 bis 19,6 und eine echte Zeile aus dem Drehbuch.

Die vier Schritte.

1. *Vermuten*: »Werden Sie gleich mehr sehen als bewegte Formen?« mit drei Antworten (nur bewegte Formen, eine kleine Geschichte, weiß nicht).
2. *Beobachten und beschreiben*: Die Animation startet nie von selbst, sondern erst mit »Abspielen«; sie lässt sich beenden und noch einmal abspielen. Ein beendeter Durchgang zählt erst ab Sekunde 53, wenn die Wut an der Tür vorbei ist und Kreis und kleines Dreieck wieder zusammengefunden haben; wer früher beendet, spielt neu ab. Der Knopf »Ohne Durchgang weiter« steht über der Bühne, nicht neben »Durchgang beenden«, und verschwindet, sobald ein Durchgang gezählt hat. Danach beschreiben die Lernenden in drei Sätzen, was passiert ist. Weiter geht es, sobald die Beschreibung zwölf Wörter hat, also etwa drei kurze Sätze.
3. *Nachmessen*: Unter »Was der Bildschirm gezeigt hat« stehen die drei Bahnen einzeln als Linien auf der Fläche, mit Start, Ende und Zeitmarken alle zehn Sekunden. Die Bahn des großen Dreiecks beginnt am oberen Rand, wo es bei Sekunde 8,9 ins Bild kommt; die Legende nennt den Ring deshalb »Start oder Eintritt ins Bild«. Bei Sekunde 60 fehlt die Marke, weil alle drei Formen dann schon an ihrem Ende stehen. Kleines Dreieck und Kreis haben zwischen Sekunde 40 und 50 an derselben Stelle gewartet; dort steht eine gemeinsame Marke »40–50«, die die Legende erklärt. Wo eine Zahl etwas weiter von ihrer Marke entfernt steht, führt eine kurze Linie zu ihr. Dazu kommt der Ausschnitt aus dem Drehbuch als Tabelle



(Zeit, x, y, Winkel). Unter »Ihre Wörter« steht die eigene Beschreibung Wort für Wort. Die Seite schlägt Wörter vor, die den Formen Absichten oder Gefühle zuschreiben, rahmt sie gestrichelt und zählt sie vorläufig; durch Antippen entscheidet die Person selbst, was zählt. Das Ergebnis lautet etwa »In Ihrer Beschreibung schreiben sieben Wörter den Formen Absichten oder Gefühle zu«, und daneben steht die eigene Vermutung.

4. *Deuten*: eine erste Antwort auf die Frage, was das für die Person neben Ihnen heißt. Sie erscheint im Rückblick wieder, zusammen mit der Vermutung.

Was die Nachmessung zeigt und was nicht. Die Bahnen zeigen, was der Bildschirm tatsächlich gezeigt hat; sie stammen aus denselben Daten, nach denen sich die Formen bewegt haben. Die Wortzählung ist dagegen keine Messung an der Person, sondern ein Spiegel ihrer eigenen Sätze, und sie wird mit niemandem verglichen. Die Vorschläge beruhen auf einer schlichten Liste von Wortanfängen (wollen, fliehen, jagen, helfen, Wut, Angst, Freude und ähnliche). Sie übersehen manches und schlagen gelegentlich ein Wort vor, das nichts zuschreibt. Das ist gewollt, denn die Entscheidung liegt bei der Person, und die schwierigen Fälle sind die besten Gesprächsanlässe: Ist »folgt« oder »stößt« schon eine Zuschreibung, »jagt« oder »will« dagegen sicher? Die Grenze verläuft dort, wo ein Wort mehr sagt als Ort, Richtung und Tempo.

Wenn jemand keine Geschichte sieht. Das kommt vor und ist kein Fehler. Auch bei Heider und Simmel beschrieb eine von 34 Versuchspersonen den Film fast ganz in geometrischen Begriffen, und selbst sie nannte das große Dreieck an einer Stelle »er«. Wer nur Formen gesehen hat, ist im Blitzlicht ein guter Zeuge: Wie schwer war es, bei der Geometrie zu bleiben? Manche kennen den Versuch auch schon, etwa aus Videos im Netz. Das schadet wenig, denn solche Deutungen geschehen schnell, von selbst und lassen sich kaum abstellen, auch wenn man weiß, dass es nur Formen sind (Scholl & Tremoulet 2000).

Die Studie von 1944. Heider und Simmel zeigten 34 Studentinnen des Smith College einen Trickfilm von etwa zweieinhalb Minuten, zweimal hintereinander, mit dem Auftrag, aufzuschreiben, was geschieht. Bis auf die eine erzählten alle von Lebewesen, fast immer von Menschen, zweimal von Vögeln, und 19 erzählten eine zusammenhängende Geschichte. Eine zweite Gruppe sollte die Formen ausdrücklich als Personen deuten; 97 Prozent nannten das große Dreieck angriffslustig, streitsüchtig oder tyrannisch, und eine Denkpause der Lesefassung greift diese Einigkeit auf. Rückwärts abgespielt sah eine dritte Gruppe wieder Handlungen, aber ihre Geschichten gingen weit auseinander.

Zwei Einschränkungen. Die Lesefassung lässt beide weg, für Nachfragen sind sie gut. Freie Beschreibungen mischen Wahrnehmung und nachträgliche Deutung; dass wir etwa eine Jagd wirklich sehen, zeigen deshalb erst Versuche, in denen man nur angibt, ob man eine Jagd erkennt (Gao, Newman & Scholl 2009). Dort wurde die Jagd nur gut erkannt, wenn der Verfolger höchstens etwa 30 Grad vom direkten Kurs abwich; bei allen Verfolgungen der Seite zeigt die Spitze des großen Dreiecks höchstens 19 Grad am Ziel vorbei. Und dieselben Forscher fragen, ob Versuchspersonen mit ihren Zuschreibungen recht haben, und antworten: am Ende nein, denn es sind animierte Formen auf einem Bildschirm, die ihre Urheber eigens so gebaut haben, dass sie belebt wirken.



Anders als im Film von Heider und Simmel. Der Film ist nicht gemeinfrei; die Seite zeigt deshalb eine eigene Animation und keine Bilder aus ihm. Auch die Geschichte ist eine andere. Bei Heider und Simmel geht das große Dreieck zu Beginn ins Haus und schließt die Tür, kommt heraus und kämpft mit dem kleinen, jagt den Kreis im Haus, bis das kleine Dreieck dem Kreis die Tür öffnet, jagt die beiden danach zweimal ums Haus und stößt am Ende gegen die Wände, bis sie brechen. Auf der Seite gehört das Rechteck dem kleinen Dreieck und dem Kreis, das große dringt von außen bei ihnen ein, die beiden Dreiecke berühren sich nie, der Kreis entkommt aus eigener Kraft, die Wut gilt der Tür, und die Wände bleiben stehen.

Sicherheit und Daten. Die Animation zeigt ruhig bewegte Formen ohne Flackern, ohne Blitze und ohne Ton. Sie startet nie von selbst. Wer im System reduzierte Bewegung eingestellt hat, bekommt einen Hinweis, und wer Bewegung auf Bildschirmen als unangenehm empfindet, wählt »Ohne Durchgang weiter« und prüft dann die Wörter einer erfundenen Beispielbeschreibung, in der die Seite zehn Wörter vorschlägt. Vermutung, Beschreibung, Auswahl der Wörter und erste Antwort bleiben im Browser dieses Geräts, auch nach einem Neuladen; übertragen wird nichts.

5. Lösungen und Erläuterungen

Aufgabe A1 · Licht in die Augen (Auswahl, aus Etappe 1)

Ein Mann hat nach zwei Schlaganfällen die Sehrinde verloren und sagt, er sehe kein Licht. Leuchtet ihm eine Ärztin in die Augen, werden seine Pupillen eng. Ein Medizinstudent notiert: »Die Pupillen reagieren auf das Licht, also sieht er es.«

Richtig: Die Reaktion zeigt, dass Auge und Gehirn das Licht verarbeiten. Ob der Mann es erlebt, zeigt sie nicht, und nach seinem Bericht sieht er es nicht. Der Pupillenreflex läuft vom Auge über den Hirnstamm zurück zur Iris und braucht die Sehrinde nicht; eine erhaltene Pupillenreaktion gehört deshalb zum Bild der Rindenblindheit.

Typische Fehler. Wer meint, wer auf Licht reagiere, sehe es auch, setzt Verhalten mit Erleben gleich (F-U2.2); die Straßenlaterne mit Dämmerungsschalter hilft. Wer meint, weil der Mann nichts sieht, verarbeite sein Gehirn das Licht auch nicht, macht denselben Fehler in umgekehrter Richtung (F-U2.6), denn der Hirnstamm ist Teil des Gehirns. Wer sagt, das könne niemand herausfinden, zieht den skeptischen Kurzschluss (F-U2.9): Der Bericht des Mannes und das Wissen, welche Teile seines Gehirns zerstört sind, sind prüfbare Gründe. Die Aufgabe bereitet nebenbei den Fall im Scanner vor, denn auch dort zeigt die Messung zunächst nur Verarbeitung.

Aufgabe A2 · Wach, aber nicht bei Bewusstsein (Auswahl, aus Etappe 1)

Eine Ärztin sagt über einen Patienten im Wachkoma: »Er ist wach, aber nicht bei Bewusstsein.« Widerspricht sie sich?



Richtig: Nein. »Bei Bewusstsein« heißt hier wie beim Notarzt aus Etappe 1 wach und ansprechbar. Wach ist der Patient, denn er öffnet die Augen und hat einen Schlaf-Wach-Rhythmus; ansprechbar ist er nicht, denn er reagiert auf keine Ansprache und keine Aufforderung. Im Wachkoma fällt auseinander, was sonst zusammengeht. Ob er etwas erlebt, sagt der Satz der Ärztin streng genommen nicht, und dass Wachsein und Erleben zweierlei sind, zeigte in Etappe 1 der Traum: Wer träumt, ist nicht wach und erlebt trotzdem etwas. Für das Wachkoma schlugen Fachleute 2010 den Namen »Syndrom reaktionsloser Wachheit« vor (englisch unresponsive wakefulness syndrome; Laureys u. a. 2010). Er beschreibt genau, was man am Bett beobachtet, und lässt offen, was erlebt wird.

Typische Fehler. »Wach« und »bei Bewusstsein« gleichsetzen (F-B4.5); aus offenen Augen auf Sehen schließen (F-U2.2), wobei der Schlafwandler hilft; aus »nicht bei Bewusstsein« folgern, dass der Patient sicher nichts erlebt (F-U2.3). Die Aufgabe greift die Einteilung aus der Zuordnungsaufgabe von Etappe 1 auf, in der der Notarzt für »wach und ansprechbar« stand; wer dort richtig lag, findet hier den Anschluss. Die Rückmeldung zum letzten Distraktor erwähnt schon die Befunde aus dem Scanner. Hier nicht vertiefen, denn die Erklärung kommt ausführlich darauf zurück.

Der Mill-Text (Quelle)

Auftrag: in den Kursnotizen die drei Glieder von Mills Reihe notieren und in einem Satz eine Schwäche seines Schlusses nennen, zusammen mit dem Glied oder der Annahme, an der sie ansetzt. Die Standardform baut erst Aufgabe 2 mit Karten.

Der Auszug ist gekürzt, die Auslassungen sind markiert. Er reicht bis zu Mills eigener Unterscheidung: Zwingend ist für ihn nur, dass es bei anderen überhaupt ein mittleres Glied gibt, das entweder dasselbe ist wie bei ihm oder ein anderes; er müsse sie also »entweder für lebendig halten oder für Automaten«. Dass dieses Glied von derselben Art ist wie seines, nimmt er an, weil der Fall seinem eigenen »in allen anderen Hinsichten ähnlich ist« (im Original »by supposing the link to be of the same nature«, S. 208).

Erwartung. Die drei Glieder sind: am Anfang Veränderungen des eigenen Körpers, also etwas, das ihm geschieht, etwa ein Stoß; in der Mitte Gefühle, bei Mill alles, was jemand erlebt; am Ende äußeres Verhalten, zu dem auch das Sprechen gehört. Bei anderen bezeugen Mill seine Sinne nur das erste und das letzte Glied. Als Schwäche kommen vier Antworten in Frage, und alle sind gut.

1. Die Reihe und ihren Zusammenhang kennt Mill nur aus einem einzigen Fall, seinem eigenen; so nennt es auch die Lesefassung. Für die Sicherung passt Wittgensteins Frage, wie man »den einen Fall in so unverantwortlicher Weise verallgemeinern« könne (*Philosophische Untersuchungen* § 293); sie steht nicht in der Lesefassung, und Wittgenstein zielt mit ihr auf die Bedeutung des Wortes »Schmerz«, nicht unmittelbar auf Mills Schluss.
2. Die Annahme, das mittlere Glied sei bei anderen von derselben Art, trägt nur so weit wie die Ähnlichkeit, und bei Wesen, die uns unähnlich sind, wird der Schluss schwach.
3. Mill behauptet, das erste Glied könne das letzte »ohne« das mittlere nicht hervorbringen. Woher weiß er das? Der Fall TN aus Etappe 1 legt nahe, dass Verhalten wie das Ausweichen auch ohne das passende



Erleben vorkommt, und der Pupillenreflex aus Aufgabe A1 ist eine Reaktion auf Licht, die ganz ohne Sehen abläuft.

4. Mill trennt selbst, was ihn die Erfahrung zu schließen zwingt, von dem, was er annimmt. Zwingend ist nur ein mittleres Glied überhaupt; dass es ein Gefühl wie seines ist, bleibt eine Annahme, und bei einem Automaten wäre es etwas anderes.

Die Antworten 3 und 4 sind anspruchsvoll und verdienen, in der Sicherung vorgelesen zu werden. Wer statt der Glieder schon Prämissen und Konklusion notiert, hat nichts falsch gemacht und ist für Aufgabe 2 gut vorbereitet.

Aufgabe 1 · Das wütende Dreieck (Urteil und Grund)

Nach dem Selbstversuch sagt Lea: »Das große Dreieck war wütend, als es die Tür zugeschlagen hat und dagegen gestoßen ist. Das habe ich doch gesehen.«

Richtig: Aus Leas Eindruck folgt nicht, dass das Dreieck wütend war. Gesehen hat sie Bewegungen, die sie als Wut deutet, und das geschieht fast von selbst. Ihre Zuschreibung zeigt etwas über ihr Sehen, nicht über das Dreieck, und das Drehbuch erklärt jede Bewegung ohne Wut.

Erläuterung. Lea meint den Takt »Wut an der Tür« (42,6 bis 51,4 Sekunden, Abschnitt 4): Das große Dreieck schlägt von oben auf die offene Seite, sie schwingt zu, und dann stößt es dreimal von außen dagegen. Im Drehbuch stehen dafür nur Orte und Winkel des Dreiecks und der Winkel der Seite. Die Lesefassung nimmt das Dreieck bei Mills zweiter Prämisse wieder auf: Es ähnelt einem wütenden Menschen nur in seinen Bewegungen, und ein Körper wie unserer fehlt ihm, also trägt der Analogieschluss hier nicht bis zur Wut.

Typische Fehler. Die Aufgabe stellt die richtige Antwort zwischen zwei Fehler. Der naive lautet, was alle übereinstimmend sehen, sei auch da (F-U1.3); in der Aufgabe heißt er »Fast alle Zuschauer sehen an der Tür dieselbe Wut«. Übereinstimmung stützt eine Annahme aber nur, wenn sie sich am besten dadurch erklärt, dass es das Gesehene gibt, etwa weil viele auf verschiedenen Wegen dasselbe finden. Hier erklärt sie sich schon dadurch, dass alle dieselbe Animation mit derselben menschlichen Art zu sehen betrachtet haben. Die Überkorrektur lautet, wenn schon Dreiecken Wut angesehen werde, sei auch jede Wut bei Menschen bloß hineingelesen (F-U1.4). Sie ist die Falle, die der Selbstversuch selbst stellt, denn wer eben erlebt hat, wie leicht Zuschreibung danebengeht, hält sie schnell überall für wertlos. Bei Menschen gibt es aber Gründe, die das Dreieck nicht bietet: einen Körper und ein Gehirn wie das eigene, ein ganzes Muster aus Gesicht, Stimme und Handlungen und keine gleich gute Erklärung ohne Erleben. Wer schließlich sagt, über Wut lasse sich grundsätzlich nichts wissen (F-U2.9), übersieht, dass man bei den Dreiecken sogar sehr genau weiß, woher ihre Bewegungen kommen. Leas Eindruck selbst ist kein Denkfehler, er ist echt und kaum abzustellen. Der Fehler liegt erst im Schritt vom Eindruck zum Dreieck.



Aufgabe 2 · Mills Schluss (Standardform)

Richtig: Auf die nummerierten Plätze gehören drei Prämissen, in beliebiger Reihenfolge: »Bei mir selbst folgt auf eine Einwirkung auf meinen Körper ein Gefühl und auf das Gefühl ein Verhalten«, »Bei mir selbst bringt die Einwirkung das Verhalten nur über das Gefühl hervor« und »Andere Menschen haben Körper wie ich, und bei ihnen folgt auf dieselben Einwirkungen ebenso regelmäßig dasselbe Verhalten«. Unter den Strich gehört »Also ist das mittlere Glied bei ihnen wahrscheinlich dasselbe wie bei mir: Sie haben Gefühle wie ich.« Liegen bleiben vier Karten:

KARTE	WARUM SIE NICHT DAZUGEHÖRT	CODE
»Ich sehe anderen Menschen ihre Gefühle direkt an.«	Genau das bestreitet Mill: Seine Sinne zeigen ihm bei anderen nur das erste und das letzte Glied, deshalb braucht er einen Schluss. Die Rückmeldung nennt Scheler und Wittgenstein, die dagegenhalten, und fügt an, dass derselbe Eindruck bei den Dreiecken entsteht, wo kein Gefühl ist.	F-U1.3
»Andere Menschen sagen mir, dass sie etwas fühlen, und das genügt als Beleg.«	Sprechen gehört bei Mill zum letzten Glied, zu dem Verhalten, das sein Schluss erklären will.	F-U2.11
»Wesen, deren Körper anders gebaut ist als meiner, fühlen nichts.«	Mill nennt den Körper zwar die Vorbedingung von Gefühlen, weiß das aber nur aus seinem eigenen Fall; die Karte macht daraus eine Regel für alle anders gebauten Wesen. Der Analogieschluss selbst sagt über unähnliche Wesen nichts, weder ja noch nein.	F-U1.7
»Also ist bewiesen, dass andere Menschen Gefühle haben wie ich.«	Zwingend ist für Mill nur, dass es bei anderen überhaupt ein mittleres Glied gibt; dass es von derselben Art ist wie seines, nimmt er an.	F-U1.5

Erläuterung. Neben den vier Karten fängt die Aufgabe zwei Verwechslungen ab: die Konklusion als Prämisse und die dritte Prämisse als Konklusion. Beide Male geht die Stützrichtung verloren (F-A3.8). Hier hilft die Frage des ersten Hinweises: Was weiß Mill über sich, was beobachtet er bei anderen, und was kann er nicht beobachten? Erschlossen wird nur das Letzte. Das »wahrscheinlich« in der Konklusion ist durch Mill gedeckt. Zwingend nennt er nur den Schluss auf ein mittleres Glied; dass es dasselbe ist wie bei ihm, nennt er eine Annahme (»by supposing«, S. 208), und der Auszug reicht bis zu dieser Stelle. Wer im Text »zwingt« liest und deshalb zur Karte »bewiesen« greift, bekommt in der Rückmeldung genau diese Unterscheidung. Die Karte »Ich sehe anderen Menschen ihre Gefühle direkt an« ist mit F-U1.3 verbunden. Der Code ist so gefasst, dass nicht die Wahrnehmungsauffassung als Fehler zählt, sondern erst der Schritt, mit dem ein Eindruck die Frage schon entscheiden soll (Abschnitt 7). Die Karte zu unähnlichen Körpern ist die diagnostisch wichtigste. Sie kehrt bei Tieren und Maschinen wieder, und dort muss sitzen, dass aus dem Schweigen des Analogieschlusses kein Nein folgt.



Aufgabe 3 · Ähnlichkeit, Erklärung oder Eindruck? (Einordnen)

Fünf Menschen begründen, dass in einem anderen Wesen etwas vorgeht, ein Schmerz, ein Gefühl oder ein Verstehen. Ist die Begründung jeweils ein Analogieschluss, ein Schluss auf die beste Erklärung oder eine bloße Zuschreibung, die gar keinen Schluss zieht?

BEGRÜNDUNG (GEKÜRZT)	EINORDNUNG
Die Katze hat Nerven und ein Gehirn wie ich und schreit auf, so wie ich aufschreien würde; also spürt sie wohl Schmerz.	Analogieschluss
Der Freund an der Herdplatte zuckte zurück, kühlte und schonte die Hand und sagte, es brenne noch; Schmerz erklärt das alles, und eine ebenso gute Erklärung ohne Schmerz fehlt.	Schluss auf die beste Erklärung
Der Hund schaut so schuldbewusst, das sieht man ihm doch an.	bloße Zuschreibung
Die Patientin zeigte 30 Sekunden lang das erbetene Muster und hörte auf, als man sie bat, sich auszuruhen; ein Reflex auf ein Wort dauert nur Sekunden.	Schluss auf die beste Erklärung
Das Sprachmodell schreibt über Freude und Trauer wie ich, also fühlt es wohl auch wie ich.	Analogieschluss, und zwar ein schwacher

Erläuterung. Die Erkennungszeichen nennt der zweite Hinweis: Ein Analogieschluss nennt Ähnlichkeiten mit dem eigenen Fall (»wie ich«), ein Schluss auf die beste Erklärung nennt ein Muster und fragt, ob eine andere Erklärung es ebenso gut erklärt, und eine bloße Zuschreibung nennt keinen Grund außer dem Eindruck. Die Rückmeldungen tragen drei Codes. Wer einen begründeten Schluss als bloße Zuschreibung einordnet, hält jede Zuschreibung für wertlos (F-U1.4); wer einen Schluss auf die beste Erklärung sieht, wo keine Konkurrenz geprüft wird, überschätzt ihn (F-U2.12); und wer den Blick des Hundes für einen Analogieschluss hält, nimmt den Eindruck schon als Beleg (F-U1.3). Zwei Stellen lohnen ein Gespräch. Wer den Satz zum Sprachmodell als bloße Zuschreibung einordnet, hält ihn vermutlich für zu schwach, um ein Schluss zu sein. Er nennt aber einen Grund, der sich prüfen lässt, nämlich eine Ähnlichkeit mit dem eigenen Fall; die Einordnung fragt nach der Art des Grundes, seine Stärke ist eine zweite Frage. Diese Verwechslung hat eine eigene Rückmeldung (F-U1.4). Der Satz zur Patientin schließt auf Verstehen und Folgen, noch nicht auf Erleben; deshalb fragt die Aufgabe nach dem, was in einem Wesen vorgeht, und nicht nur nach Erleben. Der Schritt zum Erleben wäre ein zweiter Schluss derselben Art, und wer das bemerkt, hat die Lesefassung genau gelesen. Zum Hund: In einem Versuch von Alexandra Horowitz (2009) mit 14 Hunden hing der »schuldbewusste« Blick nicht davon ab, ob ein Hund den verbotenen Leckerbissen gefressen hatte, sondern davon, ob sein Besitzer schimpfte. Die Rückmeldung der Aufgabe nennt den Versuch.



Aufgabe 4 · Zwei Befunde aus dem Scanner (Urteil und Grund)

Zwei Menschen reagieren am Bett auf keine Aufforderung. Im Scanner zeigt Frau A beim Vorstellen von Tennis und Haus dasselbe Muster wie gesunde Menschen, Herr B zeigt nichts. Ein Angehöriger sagt: »Der Scanner hat es bewiesen: Frau A erlebt etwas, und Herr B erlebt nichts.«

Richtig: Der Satz ist für keinen der beiden gedeckt. Bei Frau A ist das Muster ein starkes Indiz, dass sie versteht und mitmacht, und dass sie dabei etwas erlebt, ist gut begründet, aber nicht bewiesen. Bei Herrn B beweist das leere Ergebnis nichts, denn auch viele, die am Bett Aufforderungen befolgen, zeigen in der Messung keine Reaktion.

Erläuterung. Der Angehörige sagt »bewiesen« und nicht »wissen«, und darauf kommt es an. Die Etappe versteht Wissen so, wie es fast überall gemeint ist, als gut begründete Überzeugung und nicht als Beweis (Lesefassung, »Indiz, nicht Beweis«). Wer so denkt, könnte bei Frau A vertreten, man wisse jetzt aus guten Gründen, dass sie etwas erlebt. Bewiesen ist es trotzdem nicht, und genau das behauptet der Angehörige. Bei Frau A stehen zwei Schlüsse hintereinander. Vom Muster auf Verstehen und Wollen schließt man, weil keine andere Erklärung so gut passt: Das Muster kommt genau auf die Bitte hin, dauert so lange wie verlangt, liegt in den Arealen der jeweiligen Vorstellung und gleicht dem gesunder Menschen. Vom Verstehen und Wollen auf Erleben schließt man ein zweites Mal, und dieser Schritt stützt sich wieder auf beides, auf die Erklärungskraft und auf den eigenen Fall. Der Scanner misst dabei, wo mehr sauerstoffreiches Blut ankommt, ein Zeichen für die Arbeit von Nervenzellen und nicht das Erleben selbst. Bei Herrn B hilft die Logik eines Tests, der viele übersieht: Ein positives Ergebnis sagt viel, ein negatives wenig. In der Studie von Bodien u. a. (2024) zeigten von 112 Menschen, die am Bett Aufforderungen befolgten, nur 43 eine Reaktion im Scan oder im EEG; stumm blieb die Messung also bei rund sechs von zehn. Die Auswertung war absichtlich gegen falsche positive Befunde ausgelegt, und die Autoren vermuten, dass die Aufgaben im Scanner mehr verlangen als ein Befolgen am Bett. Untersucht wurde keine Zufallsauswahl, sondern eine Gelegenheitsstichprobe aus sechs Zentren. Wie unsicher schon die Diagnose am Bett ist, zeigt eine Zahl, die Monti u. a. (2010) aus älteren Studien anführen: Etwa 40 Prozent der Diagnosen in dieser Gruppe von Patienten gelten als Fehldiagnosen.

Typische Fehler. Wer den Scan für einen Blick ins Erleben hält, stimmt dem Satz für Frau A zu (F-U2.13). Wer ein leeres Ergebnis für einen Beweis hält, stimmt ihm für Herrn B zu (F-U2.14). Wer beides zugleich glaubt, wählt »Für beide«, vermutlich die häufigste naive Antwort; ihre Rückmeldung benennt beide Fehler und trägt den Code F-U2.14. Wer sagt, kein Gerät messe Erleben, also sage keiner der Befunde etwas (F-U2.9), hat wie bei Jonas in Etappe 1 das richtige Urteil aus dem falschen Grund. Über Frau A weiß man nach dem Scan viel mehr als vorher, und ihr Befund ist ein guter Grund, mit ihr zu sprechen wie mit jemandem, der versteht.

Aufgabe 5 · Tennis oder Haus (Gespräch)

Erwartung. Von außen kann B kaum mehr als raten. Bei drei Ja-Nein-Fragen sind im Mittel anderthalb Treffer zu erwarten, und alle drei trifft man durch Raten in einem von acht Fällen. Trifft B mehr, lohnt die Frage, woran B sich orientiert hat: an kleinen Bewegungen, am Gesicht, am Atem? Dann hat B Verhalten gedeutet,



also genau das getan, was bei der Patientin nicht möglich war. Ein Scanner könnte die Antworten erkennen, wenn die beiden Vorstellungen in verschiedenen Arealen verlässlich verschiedene Muster erzeugen. Dann wüsste B, dass A die Frage versteht, sich erinnert und eine Antwort wählt. Wie es für A ist, sich den Aufschlag vorzustellen, zeigte auch der Scanner nicht; dass A dabei etwas erlebt, bliebe die beste Erklärung.

Eine gute Notiz lautet etwa: »Von außen konnte ich nur raten, und einmal habe ich mich nach einem Zucken gerichtet. Mit einem Scanner könnte man sehen, ob die Areale für Bewegung oder für Orientierung arbeiten, und so die Antwort ablesen. Aber auch dann wüsste man nur, dass A versteht und antwortet, nicht, wie es für A ist, sich das vorzustellen.«

Die Aufgabe hält sich an das Original: In der 2010 veröffentlichten Studie wurde vor jeder Frage gesagt, welche Vorstellung Ja und welche Nein bedeutet, und in den Beispielen, die der Aufsatz nennt, stand Tennis für Ja und Haus für Nein; so hält es auch die Aufgabe. Die Lehrkraft achtet darauf, dass die Fragen harmlos bleiben; die Aufgabe nennt als Beispiel »Haben Sie heute schon Kaffee getrunken?«. Der Bericht zweier Paare mit der Frage an die Klasse folgt als eigener Schritt (Abschnitt 2).

Aufgabe 6 · Die Maschinenfrage (Schreibaufgabe mit Rubrik)

Auf der Etappenseite trägt die Aufgabe den Titel »Die Maschinenfrage: Hat es den Test bestanden?«.

In zwei Studien von Cameron Jones und Benjamin Bergen, deren Plan vorab festgelegt war, schrieben Versuchspersonen fünf Minuten lang gleichzeitig mit einem Menschen und einem Programm und sollten danach sagen, wer der Mensch war. Das Sprachmodell GPT-4.5 mit der Anweisung, eine bestimmte menschliche Rolle zu spielen, wurde in 73 Prozent der Fälle für den Menschen gehalten, öfter als der echte Mensch, ohne diese Anweisung in 36 Prozent; ELIZA, ein einfaches Regelprogramm aus den 1960er-Jahren, kam auf 23 Prozent (Jones & Bergen 2026). Nele: »73 Prozent! Dann denkt und fühlt es wie wir.« Tim: »ELIZA wurde auch in 23 Prozent der Fälle für einen Menschen gehalten, und ELIZA fühlt sicher nichts. Also fühlt GPT-4.5 auch nichts.«

Die Frage lässt offen, welche Art von Schluss Nele zieht, und Tims Schluss führt die Lesefassung nicht vor: Sie sagt nur allgemein, dass eine Erklärung ohne Erleben nicht zeigt, dass nichts erlebt wird. Den Vergleich mit ELIZA müssen die Lernenden selbst prüfen, und daran hängt der Anspruch der Aufgabe.

Kriterien (je 2 Punkte, formativ): Was der Test misst · Neles Schluss prüfen · Die Erklärung ohne Erleben · Tims Schluss prüfen.

Erwartungshorizont. Gemessen hat der Test, ob Befragte in einem fünfminütigen schriftlichen Gespräch ein Programm von einem Menschen unterscheiden können; die 73 Prozent sagen, wie oft sie das Programm für den Menschen hielten, und beschreiben also Urteile über ein Verhalten. Ob das Programm etwas erlebt, misst der Test nicht. Nele schließt aus einer Ähnlichkeit im Schreiben auf eine Ähnlichkeit im Denken und Fühlen, also mit einem Analogieschluss. Er ist schwach, weil er sich auf eine einzige Ähnlichkeit stützt: In Körper, Aufbau und Entstehung ist das Programm einem Menschen ganz unähnlich, während sich derselbe Schluss bei einem Menschen auf Körper, Gehirn und Verhalten zugleich stützt. Vor allem gibt es eine Erklärung für die 73 Prozent, die ohne Erleben auskommt. Das Modell hat an riesigen Mengen menschlicher



Texte gelernt, wie Menschen schreiben, und hatte die Anweisung, eine menschliche Rolle zu spielen; ohne diese Anweisung waren es nur 36 Prozent. Zwischen den beiden Fassungen hat sich nur die Anweisung geändert, Modell und Training blieben dieselben, und das passt genau zur Konkurrenzerklärung. Weil das Ergebnis auch ohne Erleben zu erwarten wäre, ist es nur ein schwaches Indiz für Erleben. Tim schließt ebenfalls aus einer einzigen Ähnlichkeit im Gespräch, nur in die Gegenrichtung und mit ELIZA statt mit einem Menschen. ELIZA zeigt, dass ein Programm ohne Erleben im schriftlichen Gespräch für einen Menschen gehalten werden kann; das Gespräch entscheidet die Frage also nicht. Dass GPT-4.5 nichts erlebt, zeigt ELIZA aber nicht. Die beiden Programme unterscheiden sich in Aufbau und Leistung stark (23 gegen 73 Prozent), und dieselbe Ähnlichkeit besteht auch mit dem echten Menschen im Test, der ebenfalls für einen Menschen gehalten wurde und etwas erlebt. Für seine Behauptung bräuchte Tim einen eigenen Grund, etwa dass Erleben an lebende Körper gebunden ist; ein solcher Grund wäre umstritten und stammt nicht aus dem Test.

Hintergrund. Für die Konkurrenzerklärung stehen in der Lesefassung das Training an menschlichen Texten, das die Lernenden aus Etappe 1 kennen, und das Nachahmungsargument (Schwitzgebel & Pober 2024). Butlin u. a. (2026) fassen dieses Gaming-Problem der Sache nach als Schluss auf die beste Erklärung: Ein Anzeichen gilt bei ihnen als »ausgetrickst«, wenn es sich besser dadurch erklären lässt, dass es ein System bewusst erscheinen lässt, als dadurch, dass das System bewusst ist. Den Ausdruck »Schluss auf die beste Erklärung« verwenden sie dabei nicht; ihre Bestimmung ist vergleichend (»better explained«, Box 3).

Die Rubrik ist in beide Richtungen offen. Wer den Test als guten Test für Denken oder Sprachgebrauch verteidigt, erfüllt das erste Kriterium, sofern er Erleben davon trennt. Wer Neles Reaktion als Schluss auf die beste Erklärung liest (Denken und Fühlen erklärten die 73 Prozent am besten) und sie daran misst, ob eine gleich gute Erklärung ohne Erleben bereitsteht, erfüllt das zweite ebenso wie ein Text, der der sprachlichen Ähnlichkeit Gewicht gibt, etwa weil Sprache viel über Denken verrät, sofern er die Unterschiede nennt und keinen Beweis daraus macht. Wer Tims Ergebnis mit einem eigenen Grund für wahrscheinlich hält, erfüllt das vierte, sofern er sagt, dass dieser Grund nicht aus dem Test oder dem Vergleich mit ELIZA stammt. Ebenso erfüllt es ein Text, der mit Sparsamkeit argumentiert (was sich ohne Erleben erklären lässt, erklärt man ohne Erleben) und Tims Ergebnis deshalb für die sparsamere Annahme hält, sofern er sagt, dass daraus keine Gewissheit folgt, weil Erleben da sein könnte, ohne die Sätze zu erklären. Die Etappe nennt Einfachheit selbst als Maßstab der besten Erklärung, und wer ihn folgerichtig anwendet, soll dafür nicht abgewertet werden. Überzogen ist erst, aus dem Test in die eine oder die andere Richtung eine Gewissheit über das Erleben zu machen.

Typische Fehlerbilder. Der Test gilt als Bewusstseinstest, und Neles Schluss wird übernommen (F-U2.15). Die Ähnlichkeit im Schreiben gilt als ausreichend (F-U1.6). Aus dem Vergleich mit ELIZA oder aus der Konkurrenzerklärung wird geschlossen, dass das Programm nichts erlebt, und Tims Schluss wird übernommen (F-U2.16). Die Frage wird zur Intelligenzfrage umgedeutet (F-U2.8). Der Text endet bei »Das kann man nie wissen« und prüft keine der beiden Reaktionen (F-U2.9). Mit Tims Umkehrung ist gerade bei guten Texten zu rechnen: Wer Neles Schluss eben mit der Konkurrenzerklärung zerlegt hat, hält den Vergleich mit ELIZA leicht für die Bestätigung. Eine zweite typische Schwäche: Der Text kritisiert Tim nur über den



Unterschied zwischen ELIZA und GPT-4.5, ohne zu sagen, was der Vergleich zeigt, oder er folgert aus seiner Kritik, das Programm fühle doch etwas (Kriterium 4 teilweise).

Zwei Schülertexte zur Kalibrierung (erfunden, typisch):

Text A (alle Kriterien erfüllt): »Gemessen hat der Test nur, ob die Befragten nach fünf Minuten Chat sagen konnten, wer der Mensch ist. Die 73 Prozent sind also Urteile über ein Verhalten, das Schreiben. Ob das Programm dabei etwas erlebt, hat niemand gemessen. Nele schließt: Es schreibt wie wir, also denkt und fühlt es wie wir. Das ist ein Analogieschluss, aber ein schwacher, denn die Ähnlichkeit gibt es nur beim Schreiben. Einen Körper und ein Gehirn aus Nervenzellen hat das Programm nicht, und bei einem Menschen stützt sich derselbe Schluss auf all das zugleich. Außerdem lassen sich die 73 Prozent ohne Erleben erklären: Das Modell hat an riesigen Mengen menschlicher Texte gelernt, wie Menschen schreiben, und sollte eine menschliche Rolle spielen. Ohne diese Anweisung waren es nur 36 Prozent. Das Ergebnis ist deshalb ein schwaches Indiz. Tim macht es sich aber zu leicht. ELIZA zeigt nur, dass man im Chat für einen Menschen gehalten werden kann, ohne etwas zu erleben. Das Gespräch entscheidet die Frage also nicht. Dass GPT-4.5 nichts fühlt, folgt daraus nicht, denn Tim schließt genau wie Nele aus einer einzigen Ähnlichkeit, nur mit ELIZA statt mit uns. Auch der echte Mensch im Test wurde für einen Menschen gehalten und fühlt trotzdem etwas. Wer behaupten will, dass das Programm nichts fühlt, braucht einen anderen Grund, etwa dass Erleben an einen lebendigen Körper gebunden ist, und der stammt nicht aus dem Test.«

Text B (teilweise erfüllt, ein Kriterium nicht): »Der Turing-Test prüft, ob ein Computer so gut chatten kann wie ein Mensch. GPT-4.5 hat ihn bestanden, weil die Befragten es in 73 Prozent der Fälle für den Menschen gehalten haben, sogar öfter als den echten Menschen. Nele meint, es denke und fühle deshalb wie wir. Das ist falsch, denn eine Maschine ist eben kein Mensch, auch wenn sie so schreibt. Außerdem hat das Programm an sehr vielen Texten von Menschen gelernt, wie Menschen schreiben, und es hatte die Anweisung, eine menschliche Rolle zu spielen. Ohne diese Anweisung haben es nur 36 Prozent für den Menschen gehalten. Es macht also nur nach. Deshalb hat Tim recht: Auch ELIZA wurde für einen Menschen gehalten, obwohl es nur ein einfaches Programm ist, das sicher nichts fühlt. Dann fühlt GPT-4.5 auch nichts, denn es macht im Grunde dasselbe wie ELIZA, nur besser.«

Rückmeldung in der Sache: Der Test ist richtig beschrieben, aber nicht vom Erleben getrennt (Kriterium 1 teilweise). Die Kritik an Nele sagt nur, eine Maschine sei kein Mensch, ohne die fehlenden Ähnlichkeiten zu nennen (Kriterium 2 teilweise). Die Konkurrenzerklärung ist gut gefunden und mit den 36 Prozent belegt, dient aber als Beweis für Tim statt als Einwand gegen Nele (Kriterium 3 teilweise). Tims Schluss wird übernommen und mit »dasselbe, nur besser« noch verstärkt, ohne zu prüfen, was der Vergleich mit ELIZA zeigt (F-U2.16, Kriterium 4 nicht erfüllt). Nächster Schritt: Nennen Sie, welche Ähnlichkeiten dem Programm fehlen, und prüfen Sie, was ELIZA wirklich zeigt: dass man ohne Erleben für einen Menschen gehalten werden kann, oder dass GPT-4.5 nichts erlebt?



Rückblick und Glaubenstafel

Eine gute verbesserte Antwort nennt einen der beiden Schlüsse beim Namen, etwa: »Ich sehe es nicht, ich schließe es: Sie ist mir in Körper, Gehirn und Verhalten ähnlich, und nichts erklärt ihr ganzes Verhalten so gut wie ein Erleben.« Auf der Glaubenstafel gibt es keine richtigen Werte, wohl aber gute und schwache Begründungen für eine Verschiebung. Beim Menschen im Wachkoma begründen die Befunde aus dem Scanner eine Verschiebung nach oben oder hin zu »offen«. Beim Chatbot ist eine Verschiebung nach unten gut begründet, wenn der alte Wert auf den Sätzen des Chatbots beruhte, denn diese Sätze haben in der Etappe an Gewicht verloren. Ein Wert von 0 mit der Begründung »er macht nur nach« wäre dagegen derselbe Schritt zu weit wie Tims Schluss in Aufgabe 6 (F-U2.16). Für alle Kandidaten gilt, was das Kurskonzept festhält: Wer 0 oder 100 einstellt, muss erklären, woher er diese Gewissheit nimmt.

6. Differenzierung

Unterstützung. Die Lesefassung in drei Stücken lesen lassen, mit einer kurzen Notiz nach jedem Stück: bis zum Ende von »Die Person neben Ihnen«, dann die beiden Schlüsse mit »Indiz, nicht Beweis«, dann »Tennis im Scanner« bis zum Schluss, wobei die letzten beiden Abschnitte auch zu Hause gelesen werden können (Abschnitt 2). Für die Notizen eine vorbereitete Tabelle mit den Zeilen »so geht er«, »stark, weil« und »Grenze« geben, wie im Tafelbild. Beim Mill-Text helfen drei Leitfragen: Was weiß Mill über sich selbst? Was sieht er bei anderen? Was schließt er daraus, und was nimmt er nur an? Bei Aufgabe 6 die Formulierungshilfen nutzen. Aufgabe 5 in der Dreiergruppe mit festen Rollen: A antwortet, B rät, C beobachtet A und protokolliert, woran B sich orientiert.

Vertiefung für Klasse 11 und 12.

1. *Lebendig oder Automaten.* Am Ende des Auszugs stellt Mill selbst eine Alternative auf: Das mittlere Glied sei bei anderen entweder dasselbe wie bei ihm oder ein anderes, und er müsse sie »entweder für lebendig halten oder für Automaten«. Welche Glieder seiner Reihe fehlen bei einem Sprachmodell, und was folgt daraus für Mills Schluss? Gute Antworten sehen, dass schon das erste Glied fehlt, ein Körper wie der eigene, und dass vom letzten nur die Sätze ähnlich sind. Mills Schluss reicht deshalb kaum bis zum Sprachmodell. Sehr gute Antworten bemerken, dass Mills Wort »Vorbedingung« hier sogar ein Nein nahelegt: kein Körper, keine Gefühle. Dieses Nein ruht aber wieder auf dem einen Fall, den Mill kennt, seinem eigenen, und der Analogieschluss selbst trägt es nicht; er schweigt über unähnliche Wesen (F-U1.7). Dass das Sprachmodell ein Automat in Mills Sinn ist, folgt daraus also nicht, nur dass dieser Schluss hier nicht weiterhilft.
2. *Der Streit um die Patientin.* Nachev und Husain (2007) wandten dreierlei ein: Aus einem Aktivierungsmuster dürfe man nur dann auf eine Tätigkeit schließen, wenn dasselbe Muster ohne sie nicht vorkommt; die Wörter könnten die Areale automatisch angestoßen haben; und wo es keine Wahl gebe, könne man nicht von einer Entscheidung sprechen. Welcher Einwand trifft die Ja-Nein-Antworten von Monti u. a. (2010) noch, welcher nicht mehr? Gute Antworten erkennen, dass das neutrale Startsignal



»answer« den Einwand der Automatik entkräftet und die Wahl zwischen zwei Vorstellungen den dritten, während der erste in abgeschwächter Form bleibt und ein einziger Patient wenig verallgemeinern lässt. Auch eine gewählte Antwort zeigt Verstehen und Wollen; dass dabei etwas erlebt wird, bleibt ein Schluss.

3. *Turing im Original.* Turings Antwort auf Jefferson steht in »Computing Machinery and Intelligence« (Mind 59, 1950, S. 445–447) und ist gemeinfrei. Lesen Sie die Stelle auf Englisch: Wofür argumentiert Turing, wenn er die »polite convention that everyone thinks« ins Spiel bringt, und warum ist das ein Argument für Gleichbehandlung und kein Beweis? Gute Antworten bemerken, dass Turing den Solipsismus selbst »vielleicht den folgerichtigsten Standpunkt« nennt und die Rätsel des Bewusstseins ausdrücklich zurückstellt; er fragt nach Denken, nicht nach Erleben. Nebenbei lässt sich zeigen, dass Turing sein Spiel zuerst mit einem Mann und einer Frau beschreibt und die heute übliche Fassung, Maschine gegen Mensch, auf S. 442 selbst nennt.

7. Fragen, die Schülerinnen und Schüler stellen

»**Heißt das, ich kann nicht wissen, ob meine Freunde etwas fühlen?**« Nein. Man weiß es in dem Sinn, in dem man fast alles weiß: aus guten Gründen und nicht durch einen Blick nach innen. Bei der Person nebenan stützen ein starker Analogieschluss und eine Erklärung ohne ernsthafte Konkurrenz dieselbe Annahme. Die Etappe fragt nicht, ob wir den Mitmenschen trauen dürfen, sondern worauf sich unsere Gewissheit stützt, damit wir in den schwierigen Fällen sehen, welche Gründe dort fehlen. Den Standpunkt, man könne nur vom eigenen Erleben wissen, nennt man Solipsismus; Turing hielt ihn für vielleicht den folgerichtigsten, aber für einen, der die Verständigung schwer macht.

»**Und wenn alle anderen nur so tun?**« Ausdenken lässt sich eine solche Erklärung immer. Die Lesefassung nennt einen vollkommenen Automaten, der alles tut, was ein Mensch mit Schmerzen tut, ohne dabei etwas zu erleben. Eine Schauspielerin, an die viele bei »so tun« denken, wäre kein gutes Beispiel, denn wer spielt, will täuschen, plant und beobachtet, erlebt also etwas, nur keinen Schmerz; eine Konkurrenzklärung im Sinn der Prüffrage muss ganz ohne Erleben auskommen. Robert Pargetter (1984) prüft drei solche Konkurrenzklärungen: dass ein Gott die anderen wie Marionetten lenkt, dass nur bei mir das Verhalten aus Erleben hervorgeht, und dass die anderen zwar etwas erleben, aber etwas ganz anderes als ich. Alle drei erklären dasselbe Verhalten schlechter als die einfache Annahme, dass andere erleben wie ich; man schließt darauf, wie Physiker auf Elektronen schließen, als beste Erklärung für alles, was sie beobachten. Andrew Melnyk (1994) wendet ein, dass dieser Schluss ganz ohne den eigenen Fall nicht auskommt, und hält deshalb beide Antworten für nötig (beides nach Avramides 2023).

»**Aber ich sehe doch, dass sie Schmerzen hat!**« Wer so fragt, vertritt eine ernsthafte Position, und die Etappe widerlegt sie nicht. Max Scheler (1913) meinte, wir seien mit der Freude eines anderen in seinem Lachen und mit seiner Scham in seinem Erröten unmittelbar bekannt. Ludwig Wittgenstein schrieb, wir sähen nicht erst die Gesichtsverziehungen und schlössen dann wie ein Arzt bei der Diagnose auf Freude oder Trauer (*Zettel* § 225). In der heutigen Phänomenologie vertreten etwa Dan Zahavi und Shaun Gallagher diese Auffassung (Überblick bei Avramides 2023). Zwei Dinge lassen sich im Gespräch trotzdem sagen. Auch eine



Wahrnehmung kann täuschen, und die Dreiecke zeigen, dass derselbe Eindruck dort entsteht, wo kein Gefühl ist. Und in den schwierigen Fällen hilft das Sehen wenig: Im Wachkoma zeigt das Gesicht nichts, beim Oktopus lesen wir es nicht sicher, und ein Sprachmodell hat keines. Dort braucht man Gründe, und dafür sind die beiden Schlüsse da. Der Code F-U1.3 zählt deshalb nicht die Wahrnehmungsauffassung als Fehler, sondern erst den Schritt, mit dem ein Eindruck die Frage schon entscheiden soll.

»**Wessen Wut ist es dann, die ich beim Dreieck sehe?**« Die Bewegungen hat der Autor der Animation so entworfen, dass sie wütend aussehen, und die Zuschauer sehen darin Wut. Wütend ist dabei niemand, auch der Autor nicht, denn er hat einen Anschein gebaut wie ein Zeichner von Trickfilmen. Die Absicht in der Geschichte steckt im Entwurf und im Blick der Zuschauer, nicht in den Formen.

»**Aber mein Gehirn ist doch auch ein Drehbuch.**« Die Frage trifft das Kriterium der Etappe an seiner empfindlichsten Stelle, denn auch eine vollständige Erklärung durch Nervenzellen käme scheinbar ohne das Wort »Erleben« aus. Eine solche Erklärung ist aber nur dann eine Erklärung ohne Erleben, wenn Erleben etwas anderes ist als diese Vorgänge, und genau das ist die Frage der Etappen 4 und 5. Beim Dreieck ist der Fall einfacher: Das Drehbuch erklärt jede Bewegung, und an den Formen selbst verarbeitet nichts etwas.

»**Hätte die Patientin nicht einfach auf das Wort »Tennis« reagieren können?**« Genau das wandten Kritiker ein (Nachev & Husain 2007; Greenberg 2007). Owen und sein Team antworteten (Owen u. a. 2007), eine automatische Reaktion auf ein Wort dauere wenige Sekunden und liege in Sprachregionen; bei ihr hielt die Aktivität in den Arealen der Vorstellung die vollen 30 Sekunden an, bis man sie bat, sich auszuruhen. Die Gegenprobe mit bloßen Sätzen wie »The man played tennis« machte sein Team allerdings nur an einer einzigen gesunden Person. Die Lesefassung lässt diese Erwiderung weg und geht gleich zu den Ja-Nein-Antworten aus der 2010 veröffentlichten Studie, die überzeugender sind: Das Startsignal war jedes Mal dasselbe neutrale Wort, und der Patient musste wählen.

»**Kann ein Hirnscanner Gedanken lesen?**« Nur in sehr engen Grenzen. In den Studien von 2006 und 2010 unterschied man zwei vorher vereinbarte Vorstellungen, die in ganz verschiedenen Arealen arbeiten; beliebige Gedanken las niemand. Neuere Verfahren können aus Scannerdaten den ungefähren Sinn gehörter oder vorgestellter Sätze erschließen, brauchen dafür aber viele Stunden Messung mit derselben Person und ihre Mitarbeit (Tang u. a. 2023). Und in jedem Fall misst der Scanner die Durchblutung als Zeichen der Arbeit von Nervenzellen, nicht das Erleben.

»**Ist der Turing-Test nicht längst bestanden? Und fünf Minuten sind doch kurz.**« In den beiden Studien von Jones und Bergen hielten die Befragten GPT-4.5 mit Rollenanweisung in 73 Prozent der Fälle für den Menschen, das einfache Programm ELIZA aus den 1960er-Jahren immerhin in 23 Prozent. Die fünf Minuten stammen von Turing selbst, aus seiner Vorhersage für das Jahr 2000. In einer dritten, ebenfalls vorab festgelegten Studie mit 15 Minuten wurde GPT-5 mit Rollenanweisung in 59 Prozent der Fälle für den Menschen gehalten. Gemessen wird in allen Fassungen die Menschenähnlichkeit im Gespräch; die Befragten achteten nach Auskunft der Autoren vor allem auf Stil und Umgangston, kaum auf Wissen. Über Erleben sagt der Test wenig, denn Training und Rollenanweisung erklären das Ergebnis auch ohne Erleben, und Turing fragte ohnehin nach Denken.



»Gibt es den echten Film von Heider und Simmel?« Ja, ein Trickfilm von etwa zweieinhalb Minuten, für den Pappfiguren auf einer beleuchteten Milchglasplatte Bild für Bild verschoben wurden. Er ist im Netz leicht zu finden, aber nicht gemeinfrei; die Seite zeigt deshalb eine eigene Animation mit einer anderen Geschichte (Abschnitt 4, »Anders als im Film von Heider und Simmel«). Wer den Film nach der Etappe zeigt, kann vergleichen lassen, welche Geschichte die Klasse darin sieht und welche Rollen sie den Formen dort gibt.

»Die Seite ist mit einer KI geschrieben. Kann sie über KI unparteiisch sein?« Die Lesefassung sagt es selbst: Die Seite ist mit Hilfe des Sprachmodells Claude entstanden, und dessen Hersteller Anthropic ist in der Frage, ob solche Systeme etwas erleben, selbst Partei. Die Sätze verantwortet die Lehrkraft. Die Frage ist ein guter Anlass, das Werkzeug der Etappe auf die Seite selbst anzuwenden: Woran erkennt man, ob eine Darstellung über KI einseitig ist? Etwa daran, ob sie Gründe in beide Richtungen prüft, wie es Aufgabe 6 verlangt.

8. Anschluss

Die dritte Etappe dreht die Frage um. Bisher ging es darum, wie wir von außen auf das Erleben anderer schließen. Nun lautet die Frage: Angenommen, wir wüssten alles, was sich über ein Gehirn von außen herausfinden lässt, wüssten wir dann auch, wie es sich anfühlt? Ihr Selbstversuch ist ein Nachbild, eine Farbe, die nicht auf dem Bildschirm ist; ihr Gedankenexperiment ist die Wissenschaftlerin Mary, die alles über Farben weiß und nie eine gesehen hat (Jackson 1982). Die Glaubenstafel kommt erst in der siebten Etappe wieder.

9. Quellen

- Akova, U., Yoo, H. & Launico, M. V. (2025): Neuroanatomy, Pupillary Light Reflexes and Pathway. StatPearls, Fassung vom 15. September 2025.
- Avramides, A. (2023): Other Minds. Stanford Encyclopedia of Philosophy, Fassung vom 19. Dezember 2023. plato.stanford.edu/entries/other-minds
- Bodien, Y. G. u. a. (2024): Cognitive motor dissociation in disorders of consciousness. New England Journal of Medicine 391(7), 598–608. DOI [10.1056/NEJMoa2400645](https://doi.org/10.1056/NEJMoa2400645).
- Butlin, P., Long, R., Bayne, T., Bengio, Y., Birch, J., Chalmers, D. u. a. (2026): Identifying indicators of consciousness in AI systems. Trends in Cognitive Sciences 30(6), 488–501. DOI [10.1016/j.tics.2025.10.011](https://doi.org/10.1016/j.tics.2025.10.011).
- Gao, T., Newman, G. E. & Scholl, B. J. (2009): The psychophysics of chasing: A case study in the perception of animacy. Cognitive Psychology 59(2), 154–179.
- Greenberg, D. L. (2007): Comment on »Detecting awareness in the vegetative state«. Science 315(5816), 1221b.
- Heider, F. & Simmel, M. (1944): An experimental study of apparent behavior. The American Journal of Psychology 57(2), 243–259.
- Horowitz, A. (2009): Disambiguating the »guilty look«: Salient prompts to a familiar dog behaviour. Behavioural Processes 81(3), 447–452.
- Jackson, F. (1982): Epiphenomenal qualia. The Philosophical Quarterly 32(127), 127–136.



- Jones, C. R. & Bergen, B. K. (2026): Large language models pass a standard three-party Turing test. Proceedings of the National Academy of Sciences 123(21), e2524472123. DOI [10.1073/pnas.2524472123](https://doi.org/10.1073/pnas.2524472123).
- Laureys, S. u. a. (2010): Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. BMC Medicine 8, 68.
- Melnyk, A. (1994): Inference to the best explanation and other minds. Australasian Journal of Philosophy 72(4), 482–491.
- Mill, J. S. (1865): An Examination of Sir William Hamilton's Philosophy. London: Longman, Green, Longman, Roberts & Green, Kapitel 12, S. 208. archive.org/details/examinationofsir00mill
- Monti, M. M. u. a. (2010): Willful modulation of brain activity in disorders of consciousness. New England Journal of Medicine 362(7), 579–589.
- Nachev, P. & Husain, M. (2007): Comment on »Detecting awareness in the vegetative state«. Science 315(5816), 1221a.
- Owen, A. M. u. a. (2006): Detecting awareness in the vegetative state. Science 313(5792), 1402.
- Owen, A. M. u. a. (2007): Response to comments on »Detecting awareness in the vegetative state«. Science 315(5816), 1221c.
- Pargetter, R. (1984): The scientific inference to other minds. Australasian Journal of Philosophy 62(2), 158–163.
- Scheler, M. (1913): Zur Phänomenologie und Theorie der Sympathiegefühle und von Liebe und Hass. Halle: Niemeyer; 2., erweiterte Auflage 1923 unter dem Titel »Wesen und Formen der Sympathie«.
- Scholl, B. J. & Tremoulet, P. D. (2000): Perceptual causality and animacy. Trends in Cognitive Sciences 4(8), 299–309.
- Schwitzgebel, E. & Pober, J. (2024): The Copernican argument for alien consciousness; the mimicry argument against robot consciousness. [arXiv:2412.00008](https://arxiv.org/abs/2412.00008).
- Tang, J., LeBel, A., Jain, S. & Huth, A. G. (2023): Semantic reconstruction of continuous language from non-invasive brain recordings. Nature Neuroscience 26, 858–866.
- Turing, A. M. (1950): Computing Machinery and Intelligence. Mind 59(236), 433–460.
- Wittgenstein, L. (1953): Philosophische Untersuchungen, § 293.
- Wittgenstein, L. (1967): Zettel. Hg. von G. E. M. Anscombe und G. H. von Wright. Oxford: Blackwell, § 225.