



Carl Friedrich Gethmann

## 3.1 Einleitung

Spricht man KI als Instrument wissenschaftlicher Arbeit an, wird zugleich unterstellt, dass ein menschlicher Akteur dieses Instrument (vermeintlich) zweckgerichtet einsetzt. Mit welchem Erfolg und unter welchen Erfolgskriterien der instrumentelle Einsatz auch immer erfolgt, grundsätzlich ist der menschliche Akteur im epistemischen Prozess als solcher nicht grundsätzlich gefährdet, wenn sich seine Rolle auch ändern kann. Die Metapher von digitalen Agenten verlegt dagegen die Akteursrolle in das Instrument selbst, das begrifflich genau in diesem Moment aufhört, eines zu sein. Aus diesem Grunde ist auch der Begriff der „Autonomie“ in Bezug auf technische Artefakte allenfalls metaphorisch zu verwenden. Nur ein Akteur kann genau genommen sich selbst die Regeln des Handelns auferlegen. Ein technisches Gerät dagegen, dem die Regeln seiner Operationen vom Konstrukteur, Software-Entwickler u. a. vorgegeben werden, ist grundsätzlich „heteronom“ bestimmt, auch wenn es im Rahmen der vorgegebenen Regeln weiter „lernt“, seinen zweckgerichteten Einsatz zu optimieren. Ein digitaler Agent wäre demgegenüber der Souverän, der dank seiner künstlichen Intelligenz die Mittel bestimmte. Dieses Bild nimmt die Vorstellung in Anspruch, dass der Prozess des Generierens von Wissen prinzipiell ein regelbestimmtes Verfahren ist, also einer Methodologie folgt, dass ein solches Verfahren somit auch von einem technischen Artefakt übernommen und ohne die menschlichen Schwächen (Müdigkeit, Lustlosigkeit, Ehrgeiz, Geldgier u. a.) sogar mit größerer Perfektion ausgeführt werden kann. Sollte dieses Bild von einem regelgeleiteten Verfahren wissenschaftlicher Erkenntnisproduktion jedoch unzutreffend sein, wäre auch der Gedanke der Übernahme eines solchen Verfahrens durch einen „Erkenntnisroboter“ irreführend.

---

C. F. Gethmann (✉)  
Universität Siegen, Siegen, Deutschland  
E-Mail: [Cfg-ea@t-online.de](mailto:Cfg-ea@t-online.de)

Eine kohärente Rekonstruktion des Wissens- und Handlungsbegriffs kommt zu dem Ergebnis, dass der Mensch – trotz Entwicklungen von KI, Big Data und Machine Learning – der zentrale Akteur der forschenden Wissensarbeit bleibt. Als Handelnder ist der Mensch in der Genese wissenschaftlichen Wissens sowie in der method(olog)ischen Geltungsausweisung nicht (vollständig) substituierbar.

Zur Fundierung dieser These wird zunächst das Verhältnis von KI und menschlicher Wissensbildung untersucht (2). Der entscheidende Unterschied wird darin gesehen, dass künstlichen Agenten im Unterschied zum Menschen keine Handlungsurheberschaft zugeschrieben werden kann (3). Diese Unterscheidung ist Anlass, die weiterhin gültige Bedeutung des Bacon-Projekts der neuzeitlichen Wissenschaft angesichts künstlicher Intelligenz zu untersuchen. (4). Die Unterscheidung zwischen deterministischen und probabilistischen Voraussagen (5) sowie zwischen methodischen Verfahren und heuristischer Urteilskraft (6) sollen die Nicht-Ersetzbarkeit menschlicher Intelligenz weiter illustrieren.

## 3.2 KI und menschliche Wissensbildung

Von „künstlicher“ Intelligenz wird gesprochen, wenn Computer-Anwendungen in technischen Artefakten (Geräten, devices) Merkmale menschlicher Intelligenz aufweisen. In diesem Zusammenhang wird seit den siebziger Jahren des 20. Jahrhunderts kontrovers über die Frage diskutiert – in Anlehnung an einen bekannten Buchtitel von H.L. Dreyfus (1979, 1992) – was Computer können und nicht können bzw. demnächst können und nicht können werden. Um diese Frage zu beantworten, muss zunächst geklärt werden, von welchen impliziten Annahmen hinsichtlich der menschlichen Intelligenz dabei Gebrauch gemacht wird. Dazu wird jedoch in den Wissenschaften keine einfache Antwort angeboten.

### 3.2.1 Intelligenz

In der Psychometrie wird „Intelligenz“ zumeist als Fähigkeit definiert, zu lernen, sich an neue Umstände anzupassen oder komplexe Informationen zu verarbeiten. In der Konstruktion von Intelligenztests, die sich auf diese Definitionen stützen, werden Aufgaben gestellt, die verschiedene mentale Operationen erfordern. Der Wechsler-Intelligenztest (1958), auf dem die bekanntesten modernen Intelligenztests aufbauen, besteht aus zwei Teilbereichen: einem verbalen und einem sprachfreien Test. Der verbale Test enthält Unteraufgaben für Information, Verständnis, Wortschatz, Gedächtnisspanne für Zahlen, arithmetisches Denkvermögen und Erkennen verbaler Ähnlichkeiten. Im sprachfreien Teil werden verschiedene Aufgaben zur Mustererkennung und -verarbeitung gestellt, z. B. sind auseinandergeschnittene Teile zu einer vollständigen Figur zusammenzufügen, Bilder zu ergänzen

oder zu einer sinnvollen Sequenz zusammenzufügen oder Muster aus Klötzchen sowie Zahlen mit Symbolen zu kombinieren. In der Validierung von Intelligenztests steht zumeist deren prognostische Leistung im Zentrum des Interesses: Eine hohe Validität wird hierbei durch eine starke Korrelation von Testwerten und dem pragmatischen Erfolg (z. B. dem beruflichen Erfolg nach einem Eignungstest) angezeigt.

Insbesondere die Frage zur Struktur der Intelligenz, speziell danach, ob Intelligenz als eine einheitliche Fähigkeit betrachtet wird oder viele voneinander unabhängige Fähigkeiten umfasst, ist bis heute umstritten. Mit der Entwicklung der Kognitionspsychologie in den 1970er-Jahren rückte zunehmend die Analyse der Prozesse, die nötig sind, um die Aufgaben der Intelligenztests zu lösen, in den Vordergrund der wissenschaftlichen Untersuchung. Umfangreiche Forschungen zur Bearbeitung informationsverarbeitender Aufgaben, wie Informationskodierung und geteiltes Hören führten zu der Annahme, dass (verbale) Intelligenz durch die Fähigkeit zur Auswahl und Benutzung von Informationsverarbeitungsmethoden bestimmt wird (Hunt et al. 1975). Die Kernthemen der kognitionspsychologischen Forschung jedoch, nämlich Denken, Problemlösung und Entscheidungsfindung, die außerhalb der Disziplin der Psychologie häufig mit dem Begriff Intelligenz assoziiert werden und auf die häufig in der Entwicklung künstlicher Intelligenz Bezug genommen wird, fallen in der Psychologie nicht unter diesen Begriff.

In Intelligenztests wird auf ein Ensemble menschlicher Leistungen Bezug genommen, nämlich sprachliche Kompetenz, elementare Rechenfähigkeiten, räumliches Vorstellungsvermögen, logisches Schlussfolgern oder Gedächtnisleistungen. In jüngerer Zeit werden solche Leistungen durch soziale und emotive Beziehungen, kreative Gestaltung und komplexe Problemlösungen ergänzt. Spätestens diese Ergänzungsversuche werfen die grundsätzliche Frage auf, wie die Übertragung auf technische Artefakte zu verstehen ist. Bei einigen Merkmalen wie beispielsweise der elementaren Rechenfähigkeit, dem Schlussfolgern oder Gedächtnisleistungen scheint naheliegend, dass menschliche Fähigkeiten univok auf technische Artefakte übertragen werden können. Auch diesbezüglich sind schon kritische Fragen zu stellen, wie beispielsweise, ob die menschliche Erinnerungsfähigkeit in gleicher Weise eine Gedächtnisleistung ist wie die Aktivierung eines technischen Speichers. Einerseits sind die quantitativen Leistungen technischer Speicher der menschlichen Erinnerungsfähigkeit um Größenordnungen überlegen. Andererseits sortiert der Mensch seine Gedächtnisleistungen beispielsweise nach der jeweils kontextuell bestimmten Bedeutung, während ein technischer Speicher unterschiedslos je nach den technischen Vorgaben Daten aufnimmt oder nicht. Bei Intelligenzleistungen mit emotiven und kreativen Qualitäten verstärkt sich der Verdacht, dass es sich hierbei um anthropomorphe und somit äquivalente Übertragungen handelt. Die Klärung solcher Vergleichbarkeitsprobleme hängt somit wesentlich von den Kriterien ab, durch die man eine spezifisch menschliche Intelligenzleistung bestimmt sieht. Grosso modo wird man die Verwendung des Ausdrucks „Intelligenz“ in der Wortverbindung „künstliche Intelligenz“ eher als eine Metapher einordnen müssen, deren Beschreibungs- und Erklärungsfunktion genauerer Aufklärung bedarf.

### 3.2.2 Denken können

Auch ohne eine genauere Klärung des Begriffs der Intelligenz ist offenkundig, dass er das Verhältnis zwischen menschlichen und maschinellen Leistungsmerkmalen vor allem bezüglich kognitiver Fähigkeiten hervorhebt. Entsprechend war in der ersten Phase der Diskussion um KI ab den siebziger Jahren des 20. Jahrhunderts die Frage, „ob Computer denken können“ leitend. Für die Fragerichtung ist die kontroverse Diskussion um die Interpretation des von A.M. Turing 1950 (Turing 1950) vorgeschlagenen Turing-Test paradigmatisch.<sup>1</sup> Nach Turing kann die Frage, ob technische Artefakte „denken“ können, dadurch entschieden werden, dass ein Fragesteller (für ihn verdeckten) Menschen und Geräten beliebige Fragen stellt. Wenn in einer größeren Zahl von Durchgängen mit wechselnden Fragestellern und wechselnden Menschen bzw. Geräten die Antworten zu einem hinreichend großen Anteil (z. B. 50 %) nicht eindeutig Mensch bzw. Gerät zugeordnet werden können, gibt es nach Turing keinen Grund, technischen Artefakten weniger Denkvermögen zuzuschreiben als Menschen. Für die Beurteilung der Angemessenheit dieses Tests hängt offenkundig alles davon ab, nach welchen Kriterien in überprüfbarer Weise einer Entität das Vermögen „Denken“ zugeschrieben wird. Computerwissenschaftler/-wissenschaftlerinnen unterstellen dabei im Anschluss an Turing ein Verständnis von „menschlicher“ Intelligenz, relativ zu dem sie die Leistungsfähigkeit von „künstlichen“ Computern interpretieren. Beispielsweise wird das menschliche Gedächtnis nach dem Operationsraum von technischen Speichern „modelliert“. Für dieses von den Naturwissenschaften vom Menschen transferierte Verständnis von menschlicher Intelligenz sind drei *anthropologisch folgenreiche Annahmen* kennzeichnend:

1. *Messbarkeit*: Unter Intelligenz wird eine Eigenschaft verstanden, die vielen Sorten von Entitäten graduell zugeschrieben wird. Menschen wie technische Artefakte können somit mehr oder weniger intelligent sein, je nach ihrer messbaren Leistung. Qualitative Gewichtungen lassen sich auf quantitative zurückführen, oder aber sie sind bedeutungslos.
2. *Leiblosigkeit*: Die Eigenschaft der Intelligenz ist nicht an eine leibliche bzw. organische Realisierung gebunden. Die physische (organische oder anorganische) Beschaffenheit intelligenter Entitäten ist für die Frage, ob sie denken können, irrelevant. Damit wird das Zuspreehen von Intelligenz von der physischen Realisierung unabhängig gemacht.
3. *Nicht-Individualität*: Zustände, auf die mentale Termini referieren, sind durch äußere Reize und die Reaktionen auf sie zu erklären. Gleiche mentale Zustände müssen durch gleiche Ursachen festgelegt sein. Wenn Kognitionen intelligenter Entitäten affektive und emotive Varianz bei gleichen Reiz-Reaktions-Bedingungen aufweisen, sind diese als zu behebbende technische Störungen einzuordnen.

---

<sup>1</sup>Vgl. die kritische Darstellung bei K. Mainzer (1995), S. 113 f. und passim.

Diese Unterstellungen stammen aus identitätstheoretischen und funktionalistischen Konzeptionen der Philosophie des Geistes, gemäß der jede menschliche kognitive Leistung im Prinzip funktional äquivalent durch ein technisches Artefakt darzustellen und nachzustellen ist. Viele KI-Forscher/-Forscherinnen übernehmen somit von den Naturwissenschaften vom Menschen ein funktionalistisches Verständnis menschlicher Fähigkeiten, für die (ohne weitere Rechtfertigung) das „Denken“ für symptomatisch gehalten wird (vgl. die kritische Untersuchung bei Carrier und Mittelstraß 1989). Die Interpretation und Identifikation der Leistung des „Denkens“ übernehmen viele Neurowissenschaftler/-wissenschaftlerinnen von Beschreibungen der „künstlich“ hergestellten Funktionsweisen eines Computers, anscheinend ohne, dass ihnen der damit verbundene Explikationszirkel Probleme bereitet.

Die philosophische Kritik an der Vorstellung des „denkenden“ Computers setzt an den Präsuppositionen des Funktionalismus an. Sie bezieht sich dabei auf eine lange Geschichte eines oft mehr oder weniger explizit vertretenen anthropologischen Naturalismus. Der Ausdruck „Naturalismus“ kennzeichnet nicht die kognitiven Bemühungen von Naturwissenschaftlern/-wissenschaftlerinnen als solche, sondern aus den Naturwissenschaften vom Menschen abgeleitete, aber über sie hinausgehende Deutungsansprüche hinsichtlich des Selbstverständnisses des Menschen und seiner gesellschaftlichen Selbstorganisation (exemplarisch: im Zusammenhang mit dem Strafrecht und bezüglich Erziehungsmaßnahmen). Mit der Erfolgsgeschichte der Naturwissenschaften generell und der Naturwissenschaften vom Menschen speziell ist der Naturalismus in der Neuzeit – abhängig von der Entwicklung der Naturwissenschaften selbst – in immer neuen Varianten entwickelt worden. Als besonders einflussreich sind zu nennen:

- Mechanismus (La Mettrie, d’Holbach);
- Physikalismus (Laplace);
- Evolutionismus (Darwin, Haeckel);
- Empirismus (Schlick, Carnap)
- Soziobiologismus (Wilson);
- Neurowissenschaften (Singer, Roth).

Dabei geht es in einer auffälligen Unschärfe der Begriffe mal um die „natürliche“ Erklärung kognitiver, mal operativer Fähigkeiten des Menschen. In allen Fällen hatte der Naturalismus eine „wissenschaftliche“ Weltanschauung vom Menschen zum Ziel, die eine weitgehende *Transformation philosophischer Fragestellungen* in solche der Wissenschaften ermöglichen soll (vgl. ausführlicher Gethmann 2016).

Die Kritik am anthropologischen Funktionalismus, die besonders prominent von H.L. Dreyfus<sup>2</sup> und J. Searle (Searle 1984) entwickelt worden ist, versucht zu zeigen, dass die anthropologischen Präsuppositionen des Funktionalismus grundsätzlich

---

<sup>2</sup>Die von Hubert M. Dreyfus im Anschluß an M. Heideggers Analyse des In-der-Welt-seins formulierte Kritik an den anthropologischen Präsuppositionen der Vertreter der KI bleibt unbeachtlich der Fortschritte in der KI-Forschung gültig (Dreyfus 1979).

unangemessen sind, um als hinreichende Beschreibungsinstrumente für das Handeln menschlicher Akteure eingesetzt werden zu können:

1. *Messbarkeit*: Im Unterschied zu technischen Artefakten bestimmen menschliche Akteure ihre Lebensumstände in einem mehr oder weniger großen Umfang „selbst“, und zwar nach Maßgabe subjektiver qualitativer Kriterien. Diese lassen sich nicht (restlos) adäquat quantitativ darstellen. Die Wichtigkeit einer Freundschaft, die Verzweiflung über eine Erkrankung oder die Freude an einer Mahlzeit lassen sich in der mentalen Binnensphäre nicht quantitativ messen. Emotionen können zwar andere „anstecken“, aber nicht ohne weiteres mit Anspruch auf Geltung verbindlich gemacht werden.<sup>3</sup>
2. *Leiblosigkeit*: Die Eigenschaft der Intelligenz menschlicher Akteure ist immer an eine leibliche Realisierung gebunden. Einer der Hauptmängel des Funktionalismus ist die mangelnde Unterscheidung von Leib und Körper. Der Mensch *hat* nicht einen Leib, sondern er *ist* Leib. Deswegen sind beispielsweise Angriffe gegen seinen Leib (etwa durch Folter) nicht Angriffe gegen Gegenstände seines Eigentums, sondern Angriffe auf ihn selbst. Wesentliches Merkmal der Leiblichkeit ist die organische Struktur: Leib ist Leben und Leben lässt sich nicht (vollständig) mit physikalischen Beschreibungsinstrumenten erfassen.<sup>4</sup>
3. *Nicht-Individualität*: Wenn ein technisches Artefakt auf eine Ursache reagiert, ist zu erwarten, dass jedes baugleiche ungestörte technische Artefakt in gleicher Weise reagiert. Baugleichheit ist in Bezug auf lebende Organismen jedoch eine uneinschlägige Kategorie. Selbst genetische Zwillinge sind nicht in dem Sinne gleich, dass sie auf gleiche Stimuli auf gleiche Weise reagieren. Erst recht determiniert die genetische Gleichheit nicht die Urheberschaft gleicher Handlungen. Das menschliche Individuum reagiert nicht einfach auf Symbole, sondern verwendet und kontrolliert sie. Verhält sich ein Individuum relativ stabil gegenüber seiner Umwelt, entwickelt es auf diese Weise seine Identität. Identität als eine Handlung prägende Bedingung ist somit etwas völlig anderes als eine Handlung konditionierende Baugleichheit.<sup>5</sup>

<sup>3</sup>Dazu ist die in der Tradition der philosophischen Anthropologie entwickelte Konzeption von Emotionen und Affekten heranzuziehen, vgl. Scheler 1923.

<sup>4</sup>Auf den begrifflichen Unterschied von „Leib“ und „Körper“ hat wohl zuerst Scheler 1916, S. 397–402 hingewiesen. Vgl. ferner Plessner 1928, S. 367; Plessner 1941, S. 238 u. ö.; Hengstenberg 1957, S. 88–101; Scherer 1976, S. 157–173. Zur Bedeutung der Leiblichkeit im Zusammenhang mit der Debatte um das Gehirn vgl. auch Fuchs (2013) bes. S. 33–40, 95–110. An die Philosophische Anthropologie schließt sich Dreyfus 1979, S. 235–255 an.

<sup>5</sup>Der Unterschied zwischen physischer Baugleichheit und sozialer Identität ist vor allem im symbolischen Interaktionismus herausgearbeitet worden (Mead 1934).

### 3.2.3 Zusammenfassung

Die Vorstellung der vollständigen Substituierbarkeit menschlicher kognitiver Leistungen macht von anthropologischen Präsuppositionen Gebrauch, die einer kritischen Betrachtung nicht standhalten. Dabei ist nicht ausgeschlossen, dass kognitive Teilfunktionen wie beispielsweise Gedächtnisleistungen von technischen Artefakten nicht nur ersetzt, sondern sogar hinsichtlich geringerer Störanfälligkeit, höherer Leistungsfähigkeit, tieferer Vernetzung, u. a. übertroffen werden können. Der technische Erfolg, der zweifellos besteht, erzeugt dabei die problematische Suggestion, dass die anthropomorphe Interpretation der technischen Artefakte unproblematisch ist.

## 3.3 Künstliche Agenten und menschliche Handlungsurheberschaft

Die Diskussion um die KI hat sich etwa seit der Jahrtausendwende unmerklich von der Konzentration auf die möglicherweise kognitiven Fähigkeiten abgewandt und zunehmend auf die operativen Fähigkeiten konzentriert. Nicht die Frage, in welcher Weise, wenn überhaupt, Computer „denken“ können, sondern in welchem Sinne Computer (z. B. als humanoide Roboter) „handeln“ können, steht inzwischen im Vordergrund. Diese Verschiebung der Aufmerksamkeit geht Hand in Hand mit der technischen Entwicklung, die unter intelligenten Systemen nicht nur solche mit hoher Informationsverarbeitungs-Kapazität versteht, sondern Quasi-Akteure, die menschliches Handeln nicht nur nachahmen und übertreffen, sondern zunehmend „autonom“ ersetzen. Damit stellt sich die Frage, in welchem Sinne solche „Agenten“ überhaupt handeln und in welchem Umfang und in welcher Weise solche „autonomen“ Systeme in der Folge für ihr Handeln verantwortlich gemacht werden können. Entsprechend fordert diese Frage nicht nur – wie früher – die Epistemologie, sondern auch die Ethik heraus. Im Rahmen der Jurisprudenz hat sich eine breite Debatte um die Zuordnung der Handlungsurheberschaft zu „autonomen“ Systemen als e-persons und beispielsweise Fragen der Haftung ergeben (z. B. Hilgendorf 2015). Allerdings wäre die Diskussionslage zu einfach beschrieben, wenn man sie als Übergang vom Denken (Wissen) zum Handeln beschriebe; kognitive und operative Fähigkeiten und Leistungen lassen sich technisch nicht trennen, so dass die Frage, ob Computer denken können, mit der Frage, ob Computer handeln können, in aufzuklärender Weise zusammenhängt; dieser enge Zusammenhang wird sprachlich schon dadurch suggeriert, dass die Ausdrücke, mit denen kognitive Prozesse beschrieben werden wie beispielsweise „denken“ grammatisch „Tu-Wörter“ sind. Die pragmatische Wende in der Diskussion muss also so rekonstruiert werden, dass die epistemischen Kategorien in sie eingebettet werden können.

### 3.3.1 *Handlungsurheberschaft und Zurechenbarkeit*

Die in Abschn. 3.2.2 angesprochenen funktionalistischen Frage-Transformationen unterstellen eine Entität, deren Handlungen und die mit ihnen verbundenen Formen des Wissens und Sprechens grundsätzlich als Wirkungen von Ursachen zu interpretieren sind (kausalistische [ätiologische] Handlungstheorien). Dem steht die Selbst-Erfahrung des Akteurs entgegen, (gegebenenfalls bedingter, aber nicht verursachter) Urheber seiner eigenen Handlungen zu sein; Handlungen sind dementsprechend nicht Wirkungen von Ursachen, sondern Zweck-Realisierungsversuche (finalistische [teleologische] Handlungstheorien). Die Pointe finalistischer Handlungstheorien liegt darin, die Handlungserfahrung des Akteurs bezüglich seiner eigenen Handlungen, die „Vollzugsperspektive“, von der Handlungsbeschreibung eines Außenstehenden, der „Berichtsperspektive“, zu unterscheiden. Diese Unterscheidung ist grundlegend, weil die Naturwissenschaften vom Menschen ihre Gegenstände in der Berichtsperspektive (mehr oder weniger angemessen) beschreiben, methodologisch dagegen für die Vollzugsperspektive nicht zuständig sind (Gethmann 2010).

Die Selbstzuschreibungen in der Vollzugsperspektive (Askriptionen) lassen sich nicht restlos auf die Fremdzuschreibungen in der Berichtsperspektive (Deskriptionen) reduzieren. Das zeigt sich vor allem darin, dass der Ich-Autor einer (Rede-) Handlung nicht ohne semantischen Verlust auf den Akteur im Rahmen einer Handlungsbeschreibung reduziert werden kann. Während die Äußerung „Ich verspreche dir, dich morgen zu besuchen“ unter sprachlichen Normalbedingungen als Vollzug eines Versprechens bestimmbar ist, ist die Äußerung „Jemand verspricht dir, dich morgen zu besuchen“ jedenfalls kein Versprechensvollzug. Unter naheliegenden sprachlichen Randannahmen könnte es der Bericht über einen Versprechensvollzug sein (oder aber eine Beschreibung, Erzählung, ...). Die für das Abgeben eines Versprechens konstitutive Selbstverpflichtung ist für einen Bericht über ein Versprechen nicht konstitutiv. Während der Vollzug des Versprechens nur dann gelingt, wenn der Autor des Versprechens sich daran macht, zu gegebener Zeit die versprochene Handlung auszuführen, hängt das Gelingen des Vollzugs des Berichts über ein Versprechen davon ab, dass der Bericht z. B. verlässlich u. a. ist. Die Verwechslung des Vollzugs-Ich mit einem Bericht über einen Berichtsgegenstand in der Berichtsperspektive ist als Fehler des Vollzugswiderspruchs (*contradictio exercita*) anzusprechen.

Die Nicht-Reduzierbarkeit von Askriptionen auf Deskriptionen ist die anthropologisch-apriorische Basis dafür, dass Nicht-Menschen (das gilt für Tiere wie für technische Geräte in gleichem Maße) nicht durch Handlungsurheberschaft ausgezeichnet sind und ihnen somit raumzeitliche Episoden, die sie ursächlich auslösen, nicht als Handlungen zuzurechnen sind. Daher wird man niemals einem technischen Artefakt wie einem Roboter moralische Vorwürfe beispielsweise wegen eines gebrochenen Versprechens machen und deswegen auch keine moralischen Sanktionen in Form von Tadel, Mißachtung oder Strafe im Sinne sozialer Desintegration verhängen. Ein Mensch, der eine falsche Wegbeschreibung abgibt, wird

getadelt, bestraft und/oder aufgeklärt, ein Navigationsgerät bei gleicher Fehlleistung wird reklamiert, repariert oder ersetzt. Entsprechend ist auch unvorstellbar, von technischen Artefakten zu verlangen, in Rechtfertigungsdiskurse als Proponenten einzutreten. Entsprechend wird man niemals noch so „intelligente“ Roboter vor Gericht stellen und strafrechtlich sanktionieren.

Diese Rekonstruktionen lassen sich zu der Definition zusammenfassen, dass eine Intelligenzleistung genau dann als *menschliche* anzusprechen ist, wenn die Zuordnung zu einem Wesen mit Handlungsurheberschaft pragmatisch angemessen ist. KI-Leistungen können unter Gesichtspunkten angemessener Beschreibung menschlichen Intelligenzleistungen ähnlich sein, die technischen Artefakte, von denen KI-Leistungen ausgehen, können jedoch nicht in angemessener Weise als Handlungsurheber angesprochen werden. Der von John. R. Searle (1984, bes. Ch. 2) am Beispiels des Chinesischen Zimmers illustrierte Unterschied zwischen einer syntaktischen Prozedur und dem Beherrschen einer Sprache folgt aus dem Kriterium der Handlungsurheberschaft: nur ein Wesen, das (Rede-)Handlungen hervorbringt, kann diesen eine Bedeutung verleihen bzw. eine Bedeutung verstehen (vgl. Mainzer 1995, S. 653 ff.).

### 3.3.2 Kausale Geschlossenheit

Der Begriff der Handlungsurheberschaft und die damit verbundene finalistische (intentionalistische) Handlungserklärung folgt der Analyse von Kant, wonach die Vorstellung der Verursachung primär in der Erfahrung des Akteurs als Verursacher der Handlungsfolgen liegt („Kausalität aus Freiheit“; vgl. Kant 1956, S. A538/B566 ff.). Im Anschluss an Kant wurde dieser Gedanke durch die Agent-Causality-Konzeption des menschlichen Handelns weiterentwickelt (Chisholm 1964). Gelegentlich wird eingewandt, dass die Vorstellung eines Handlungsurhebers als unverursachter Verursacher das *Grundpostulat von der kausalen Geschlossenheit* des physikalischen Weltbildes verletze. In diesem Zusammenhang muß man sich allerdings vor Augen halten, dass das Kausalitätsprinzip („Jedes Ereignis ist als Wirkung einer Ursache zu interpretieren.“) – im Unterschied zu den Kausalgesetzen – nicht das Ergebnis physikalischer Beobachtung und Theoriebildung ist, sondern ein „regulative(r) Grundsatz“ (Falkenburg 2012, S. 27), eine grundlegende methodologische Präsupposition der Naturwissenschaften. Das Kausalitätsprinzip hat einen ähnlichen apriorischen Status wie das Prinzip der Homogenität des Raumes oder der Isotropie der Zeit. Somit steht die Vorstellung des „unverursachten Verursachers“ möglicherweise nicht in Einklang mit den methodologischen Grundlagen der Physik, was noch zu prüfen wäre, sie widerspricht jedoch nicht naturwissenschaftlichen Erkenntnissen (Falkenburg 2012, bes. S. 267–326) Dass menschliche Handlungen etwas „echt Neues“ in die Welt bringen, schließt im übrigen logisch nicht aus, dass sie in anderer Hinsicht verursacht sind. So wird schon von Aristoteles das Phänomen des Zufalls erklärt: dass ein Ereignis zufällig ist, bezieht sich auf eine bestimmte Kausalreihe, nicht auf alle Kausalreihen (Aristoteles 1936; *Phys* II 4–9 (197 b)).

### 3.3.3 Merkmale wissenschaftlicher Intelligenz

Die Frage nach den Folgen der Einführung künstlicher Intelligenz für Wissenschaftler-Arbeitsplätze spezialisiert die Frage nach dem Unterschied zwischen künstlicher und menschlicher Intelligenz auf die Sub-Spezies (menschlicher) Wissenschaftler/Wissenschaftlerinnen. Prima facie dürfte unbestritten sein, dass sich die Tätigkeit von Wissenschaftlern/Wissenschaftlerinnen von anderen Tätigkeiten unterscheidet, nämlich dadurch, dass Wissenschaftler/Wissenschaftlerinnen als spontane Handlungsurheber von Wissen verstanden werden. Ein Gerät der künstlichen Intelligenz kann aber a priori kein Handlungsurheber sein, sondern immer nur ein „Anwender“ von etwas, was durch menschliche Intelligenz vorgegeben wurde.<sup>6</sup> Die Explikation des Begriffes der künstlichen Intelligenz muss sich somit auf ein nicht-naturalistisches (finalistisches, intentionalistisches) Verständnis kognitiver Handlungen stützen.

Für die Beurteilung von Systemen Künstlicher Intelligenz in der Forschung bedeutet das näherhin, eine nicht-naturalistische Rekonstruktion der Tätigkeit des Wissenschaftlers zu entwickeln. Als Definitions-Skizze für die Explikation des Wissensbegriffs wird hier ein begründungstheoretischer Ansatz gewählt:

*X weiß, dass  $p$  := Für alle  $Y$ , es gibt ein  $p$  :  $X$  kann  $p$  gegenüber  $Y$  begründen*

Dieser Ansatz weicht von der oft verwendeten Definition ab, die Wissen durch Überzeugt-sein von einem Sachverhalts  $p$  und dem der-Fall-sein von  $p$  charakterisiert (Lenzen 1980, S. 53 ff.):

*X weiß\*, dass  $p$  :=  $\ddot{U}p$  und  $p$ .*

Dieser „nicht-epistemische“ Ansatz ist durch zwei kaum auflösbare Probleme belastet. Einmal ist „Überzeugt-sein“ in wohl keinem Kontext semantisch „einfacher“ als „Wissen“; die Definition steht unter einem Obscurum per obscurius-Verdacht. Ferner ist unklar, wie ein performativ oder modal uneingebundenes „ $p$ “ zu verstehen ist. Die Kommentare legen nahe, implizit den Modaloperator der Möglichkeit oder Wahrscheinlichkeit unterstellen (s. die Diskussion Lenzen 1980, S. 61 ff.).<sup>7</sup> Vor allem aber besteht ein Explikationszirkel dadurch, dass man die alethischen Modaloperatoren nicht ohne direkten Rekurs auf „Wissen“ explizieren kann. Schließlich besteht das Problem, dass ein modal oder performativ uneingebettetes „ $p$ “ zwar formal ein Satzradikal ist, das sich jeder Einbettung „andienen“ könnte, aber konnotativ einen präsupponierten epistemologischen Realismus suggeriert. Es wäre jedoch ein Verstoß gegen pragmatische Definitionsprinzipien,

<sup>6</sup>Vergl. Humm et al. in diesem Band, Kap. 2.

<sup>7</sup>Eine solche Deutung würde aber die Probleme der Semantik der Modallogik wie die De dicto-Problematik in die Explikation des Wissensbegriffs einschleusen, was jedenfalls das Obscurum per obscurius-Problem verschärfte.

durch eine Wortgebrauchsregel bereits eine Position in einer philosophischen Großdebatte (Realismus vs. Anti-Realismus) sozusagen unter der Hand zu entscheiden.

Auf der Basis des begründungspragmatischen Ansatzes lassen sich definieren:

*X erkennt  $p$  := X ist dabei, Wissen zu erwerben / herzustellen*

*X meint, dass  $p$  :=  $p$  ist „Kandidat“ des Wissens, dass  $p$*

*Y zweifelt, dass  $p$  := Y fordert X auf,  $p$  zu begründen*

Die gemeinsprachliche Verwendung des Ausdrucks „Wissen“ erweckt in Verbindung mit einem weitverbreiteten Vulgär-Cartesianismus den Eindruck, als sei Wissen ein privater innerer Vorgang, der gelegentlich von seinem „Besitzer“ „ausgedrückt“ wird.<sup>8</sup> Demgegenüber verwendet die obige Definitionsskizze den Ausdruck resultativ, d. h. als ein Ergebnis eines sozialen Prozesses, nämlich des Begründens. Begründen ist eine regelgeleitete Sequenz von Redehandlungen, die mit einem konstativen performativen Modus beginnt, für den das „Behaupten“ hier exemplarisch eingesetzt wird. Eine regelgeleitete Sequenz von Redehandlungen heie „Diskurs“. Diskurse lassen sich wie alle sprachlichen Phänomene empirisch beschreiben und erklären und fallen somit in den Gegenstandsbereich der empirischen Sprachwissenschaften. Um *interlingual* die Regeln korrekter diskursiver Verfahren zu rekonstruieren (Logik, Topik), muss man demgegenüber auf die Instrumente einer formalen Pragmatik von Redehandlung und Redehandlungssequenzen zurückgreifen (Gethmann 1979).

### 3.3.4 Behaupten und Begründen

Eine Redehandlungssequenz soll „Diskurs“<sup>9</sup> heißen, wenn gilt:

- (i) Die vorkommenden Redehandlungen lassen sich nach ihrem spezifischen „Ort im Diskurs“ klassifizieren, und zwar können sie sein:
  1. Eröffnungshandlungen (initiative Redehandlungen), wie beispielsweise das Behaupten oder Auffordern;
  2. Fortsetzungshandlungen (re-aktive Redehandlungen), wie beispielsweise das Zweifeln;
  3. Beendigungshandlungen (resultative Redehandlungen), wie beispielsweise das Zustimmung oder Abweisen.

<sup>8</sup>Zur generellen Kritik am nach-cartesischen Mentalismus s. Gethmann und Sander (2002).

<sup>9</sup>Mit „Diskurs“ wird hier der in der Erlanger Schule rekonstruierte „Dialog“ bezeichnet. Der Ausdruck „Dialog“ erscheint aus zwei Gründen verbal verfehlt: Einmal heißt  $\delta\iota\alpha$  nicht „zwei“ (vielmehr  $\delta\upsilon\omicron$ ), sondern „durch“ (per); ferner wird mit „Dialog“ oft eine eher emphatische Form von Verständigung gemeint, worauf es hier nicht ankommen kann.

- (ii) Eröffnungshandlungen sind solche, deren Zweck die Diskursbeendigungshandlung der Zustimmung ist. Sie können auch „geltungsbeanspruchende“ Redehandlungen heißen, sofern ihr Anspruch die Zustimmungsfähigkeit („Zustimmbarkeit“, „Akzeptabilität“) ist.
- (iii) Die Teilnehmer an Diskursen („Parteien“) nehmen genau zwei Rollen wahr:
  - erste Rolle: Diskurseröffnung und diese stützende („fundierende“): Proponent;
  - zweite Rolle: Diskursfortsetzung oder -beendigung: Opponent.
 Dabei können die Parteien Individuen oder Kollektive sein. Ferner können Parteien auch beide Rollen in Personalidentität wahrnehmen.
- (iv) Die Abfolge von Redehandlungen zwischen den Rollenträgern erfolgt gemäß impliziten Korrektheitsregeln, die konstitutiv in die Gelingensbedingungen der Einzelhandlungen eingehen.
  - „Ich behaupte, dass p“
  - kann dann paraphrasiert werden durch:
  - „Ich bin gegenüber einem Adressaten bereit, im Zweifelsfall eine Begründungshandlung zu vollziehen.“

Diskurse, deren Eröffnungshandlungen zum Typ der Konstativa gehören, sollen „Begründungsdiskurse“ heißen.

### 3.3.5 *Wahrheit als Wohlbegründetheit*

Die vorstehende Explikation des Begriffs des wissenschaftlichen Wissens bindet diesen an die Fähigkeit eines zweckorientierten Akteurs, Behauptungen im Zweifelsfall begründen zu können. Wissen kann somit nur ein Akteur ausbilden, der wenigstens über die Handlungskompetenzen des Behauptens, Bezweifeln, Aufforderns, Zustimmens bzw. Abweisens verfügt. Ein solcher Akteur verfügt über nicht messbare propositionaler Einstellungen, Leiblichkeit und Individualität und kann daher den Turingtest nicht bestehen. Daraus ergeben sich folgende Konsequenzen:<sup>10</sup>

#### (a) *Wahrheit*

Das Zusprechen des Beurteilungsprädikators „wahr“ zu einem propositionalen Gehalt ist nur zulässig, wenn ein diskursives, für jedermann nachvollziehbares Verfahren angegeben werden kann, das diesen propositionalen Gehalt als zustimmungsfähig auszeichnet. Das Hauptargument für einen solchen, wesentlich auf *Verfahren* bezogenen Wahrheitsbegriff („Prozeduralismus“) besagt in Anlehnung an M. Dummett (1978), dass das Wort „Wahrheit“ auf nichts referieren kann, was das Erkenntnisvermögen seiner Benutzer prinzipiell transzendiert; deshalb muss das Zu- und Absprechen von Wahrheit an Verfahren der Verifikation bzw. Falsifikation zurückgebunden werden. Solange eine Aussage weder als begründbar noch unbegründbar ausgewiesen ist, kann man deshalb nicht davon ausgehen, dass sie kraft ihrer Dar-

<sup>10</sup>Die folgenden Abschnitte sind eine Zusammenfassung von: Gethmann (2014).

stellungsfunktion von einer erkenntnisunabhängigen Realität „wahr gemacht“ wird. Die Bedeutung einer Aussage kann folglich nicht durch Angabe ihrer *Wahrheitsbedingungen* im Sinne der formalen Semantik erklärt werden, sondern nur durch Angabe ihrer *Behauptbarkeitsbedingungen*, d. h. der Umstände, unter denen ein Sprecher eine Aussage (korrekt) behaupten darf. Wäre eine Aussage allein deshalb wahr, weil der durch sie dargestellte Sachverhalt besteht, dann könnte die Feststellung ihrer Wahrheit das Erkenntnisvermögen des Sprechers unter Umständen prinzipiell übersteigen. Dies hätte zur Folge, dass sich die Kenntnis der Bedeutung dieser Aussage nicht im Handeln manifestieren kann, weil es keine Möglichkeit gibt, über das Zutreffen der Wahrheitsbedingungen zu befinden. Unentscheidbare Aussagen sind weder wahr noch falsch, weil die Wahrheit einer Aussage nicht über ihre prinzipielle Verifizierbarkeit hinausgehen kann. In dieser Fassung ist die Problemstellung nicht spezifisch auf mathematische oder andere analytische Wahrheiten beschränkt, sondern sie gilt für beliebige Kontexte; sie betrifft in gleichem Maße lebensweltliche Erfahrung, empirische Laborwissenschaften, exakte Formalwissenschaften und historisch-hermeneutische Kulturwissenschaften. In allen diesen Kontexten ist die epistemologische Grundfrage zu beantworten, wie das Verhältnis zwischen Wahrheit (als Eigenschaft einer mentalen oder lingualen Entität wie Urteil oder Behauptung) und dem Verfahren zu denken ist, das zur Feststellung dieser Eigenschaft führt.<sup>11</sup>

Die Bindung des Wahrheitsbegriffs an menschliche Verfahrenssouveränität gibt Anlass, auf den grundlegenden Unterschied von *Algorithmik und Heuristik* hinzuweisen. Mit der Entwicklung der KI ist die Vorstellung verbunden, dass die Wahrheitsfindung grundsätzlich durch Algorithmen übernommen werden kann und somit wenigstens langfristig (nach weiterer technischer Perfektionierung) Forscher/Forscherinnen „aus Fleisch und Blut“ durch KI-Maschinen abgelöst werden können. Dieser Vorstellung steht allerdings schon das von Gödel in den 30.er Jahren des 20. Jahrhunderts aufgestellte Unvollständigkeitstheorem im Wege, demgemäß jedes formale System bei der algorithmischen Erzeugung von ableitbaren Sätzen wahre Gesetze auslöst; dazu gehören insbesondere meta-sprachliche Sätze, beispielsweise solche über die Widerspruchsfreiheit des Systems. Aus philosophischer Sicht zeigen Gödels Theoreme die Grenzen des algorithmisch Erzeugbaren (Mainzer 1995, S. 100 f., 789 f.; Nida-Rümelin und Weidenfeld 2018, S. 111 f.) Spezifisch menschliche Intelligenz übersteigt deswegen jedes algorithmische Verfahren, weil heuristische Verfahren die algorithmischen nicht nur bei Bestimmung der Eigenschaften des formalen Systems übersteigen, sondern ihr wesentlich vorausgehen muss, indem im vorhinein der Zweck des Algorithmus festgelegt werden muss,

---

<sup>11</sup> Gegen diese Explikation des Wahrheitsbegriffs wird eingewandt, sie unterscheide nicht zwischen der Wahrheit einer mentalen oder lingualen Entität und ihrer epistemischen Qualifikation; das Wahrscheinlich sei vom Fürwahrgehaltenwerden strikt zu unterscheiden. Diesem Einwand ist entgegenzuhalten, dass die Abstraktion der Wahrheit vom Verfahren des Bewahrheitens zu einem sinnlosen Ergebnis führt: es kann nicht mehr erklärt werden, was es bedeutet, eine Proposition *p* für wahr zu halten, wenn man keine Vorstellung hat, wie man die Behauptung von *p* diskursiv begründen kann. – Zu dem hier vertretenen „Anti-realismus“ vgl. weiterführend: Wright 1992; Tennant 1997; Siegart 1997; s. a. den Diskussionsband Forum für Philosophie 1992.

damit auch seine Gelingensbedingungen und damit seine Qualitätsstandards. Jeder Anwendung von Algorithmen geht eine Heuristik voraus.

Am Ablaufschema eines Begründungsverfahrens lässt sich die Unvermeidbarkeit heuristischer, d. h. regelüberwachter Suchverfahren leicht illustrieren. Angenommen der Opponent äußert einen Zweifel an dem vom Proponenten behaupteten Sachverhalt. ein solcher Zweifel „folgt“ aus der Behauptung keineswegs. Der Opponent könnte auch sofort zustimmen, zum Beispiel, weil er bereits eine Begründung kennt oder an einer solchen kein Interesse hat. Die Behauptung könnte ihm relativ zu seinem bisherigen Überzeugungs-Ensemble als schwer integrierbar erscheinen. Er könnte auch überzeugt von der Behauptung sein, aber neugierig, worauf der Proponent sich im nächsten „Zug“ stützt. Der Proponent auf der anderen Seite muss „suchen“, welche propositionalen Gehalte zur Überzeugungswelt des Opponenten gehören *und* durch welche Regel er den Übergang zwischen den Prämissen und seiner Anfangsbehauptung herstellt. Insgesamt ist leicht zu sehen, dass das Begründen ein Suchprozess, ist, der sich grundsätzlich vom Ablauf eines algorithmischen Verfahrens unterscheidet. Was für Begründungen generell gilt, gilt auch für den Sonderfall des *Beweisens*, d. h. des Begründens im Rahmen einer formal konstituierten Sprache. Von formalen Sonderfällen abgesehen muss der Beweis für ein Theorem gesucht/gefunden werden, selbst dann, wenn die Beweisbarkeit des Theorems aus formalen Gründen a priori (im Rahmen der konstituierten Sprache) feststeht.

### **(b) Realität**

Die Behauptung der Existenz ist an den effektiven Existenznachweis gebunden. Verallgemeinert heißt das: *Realität* ist immer durch Verfahren vollzogene Realität. Es gibt keine schlechthin verfahrenstranszendente Realität. Versteht man unter „Realität“ das Ensemble der Tatsachen, d. h. *wahrer* Sachverhalte, dann hängt die Frage, was „real“ ist, von den Verfahren ab, die eine sprachliche Entität als „wahr“ auszeichnen. Eine solche Position ist erfahrungsgemäß dem Missverständnis ausgesetzt, die Realität werde als bloßes Produkt subjektiver Vollzüge angesehen. Es wird jedoch nicht behauptet, Tatsachen seien *Resultate* des Wahrmachens (dies wäre klassischer „Idealismus“ oder „Irrealismus“), sondern nur, dass es Tatsachen *nicht ohne* Verfahren des Wahrmachens „gibt“: Verfahren sind notwendige, nicht hinreichende Bedingung für das Zusprechen von Realität. Die Bindung des Wahrheitsbegriffs an den Begriff der Begründbarkeit (oder einen ähnlichen) setzt ein vorgängiges Verständnis der Fähigkeiten eines handelnden Wesens voraus. *Irrelative Wahrheit wird nicht durch Transzendenz (einer „Realität“), sondern durch Invarianz (eines wahr machenden Verfahrens) garantiert.* Die Qualifikationen der Begründbarkeit oder Unbegründbarkeit hängen von Verfahrensstrukturen („Strategien“), nicht vom jeweiligen raumzeitlichen Vollzug („Partie“) ab.<sup>12</sup>

<sup>12</sup>Die Unterscheidung von Strategie und Partie entspricht derjenigen von Handlungsschema und -vorkommnis (bzw. type und token); vgl. Lorenz 1976.

**(c) Subjektivität**

Damit sind hinsichtlich des Verfahrensverständnisses sowohl ein Moment der Faktizität als auch ein solches der Invarianz zu unterscheiden. Die Faktizität liegt darin, dass es einen Handlungsrahmen geben muss, relativ zu dem Verfahrensvorschriften formuliert werden, deren Ausführungen Handlungen sind. Solche faktischen Rahmenbedingungen sind z. B. Materialeigenschaften derjenigen Körper, aus denen Messgeräte hergestellt werden, mit deren Hilfe wiederum Messverfahren durchgeführt werden; oder die Kompetenzen derjenigen Akteure, die Verfahren durchführen, z. B. sprachliche elementare Fähigkeiten und Fähigkeiten der Körperbearbeitung. Diese setzen einen leiblich verfassten Akteur voraus, der in eine faktische Lebenswelt mit sozialen und natürlichen Momenten eingebettet ist, die (kontingent) sind.

Der Bezug auf einen Akteur, der die Kompetenz besitzt, Verfahren durchzuführen, traditionell gesprochen: das erkennende *Subjekt*, ist für das Zusprechen von Wahrheit und das Erkennen von Realität somit wesentlich. Das Verfügungkönnen über ein Verfahren ist vor allem deswegen als Wahrheitskriterium geeignet, weil es nicht nur um das private und raum-zeitlich einmalige Für-wahr-Halten, sondern um die prinzipielle Nach-Vollziehbarkeit für jedermann geht. Der Verfahrensbegriff ordnet die Wahrheitsqualifikation nicht nur einer faktisch-allgemeinen, sondern einer prinzipiell öffentlichen Sphäre zu. Dies bedeutet aber, dass das Verfügungkönnen über ein Verfahren nicht adäquat allein mit mentalen Termini wie Vorstellen, Denken usw. erfolgen kann. Ein Proponent, der äußert: „Ich kann mir ein Verfahren denken, aber ich kann es nicht mitteilen“, verstößt gegen eine wesentliche Gelin- gensbedingung des Wahrheitsdiskurses. Der Wahrheitsdiskurs verläuft nicht auf einer Hinterbühne, deren Spiel gelegentlich auch auf der Vorderbühne „ausgedrückt“ wird, sondern das Wahrheitsverfahren ist wesentlich ein öffentlicher Vorgang (vgl. Gethmann und Sander 2002).

Dieses „Manifestationsprinzip“ (Dummett 1993, S. 37) besagt also, dass eine Verdoppelung der Wirklichkeit in eine Denk- und eine Sprachsphäre überflüssig ist, weil mentale Prozeduren wiederum mit lingualen, also öffentlichen Kategorien rekonstruiert werden müssten. Um nämlich sichergehen zu können, dass zwei Sprecher Dasselbe „meinen“, muss sich die Bedeutung der verwendeten Ausdrücke im öffentlichen Sprachgebrauch niederschlagen können. Der über öffentlich zugängliche Verfahren verfügende bzw. diese kontrollierende und nachvollziehende menschliche Akteur muss als einer angenommen werden, der Handlungen gemäß Verfahren als Zweckrealisierungsversuche vollziehen kann.

Eine Verständigung über Handlungen ist generell nur möglich, wenn Akteure nicht nur über Handlungsvorkommnisse (tokens), sondern auch über Handlungsschemata (types) kontrolliert reden können. Über ein Handlungsschema reden heißt, wissen, wann *dasselbe noch einmal* getan wird (Kamlah und Lorenzen 1973, Kap. II.2; Lorenz 1976, S. 258). Je nach Handlungstyp gibt es Kennzeichen dafür, ob jemand diesen Handlungstyp beherrscht oder nicht. Ein Verfahren des Begründens muss wenigstens folgende Anforderungen erfüllen, was auf Seiten des Akteurs Fähigkeiten nicht-trivialer Art unterstellt:

- (i) Das Verfahren muss *regelmäßig* sein, d. h. es müssen wenigstens implizite Regeln rekonstruierbar sein, die für jedermann die Wiederholbarkeit sichern. Damit wird für jeden Akteur die Fähigkeit unterstellt, eine kanonische Handlungsfolge einzuhalten (regulative Kompetenz).
- (ii) Das Verfahren muss *lückenlos* sein, d. h. es darf beim schrittweisen Durchlaufen keine Schritte geben, die nicht durch die Regeln erlaubt sind. Dies unterstellt für jeden Akteur die Fähigkeit, Handlungsfolgen als Folgen zu bilden und wahrzunehmen (konsekutive Kompetenz).
- (iii) Das Verfahren muss *zirkelfrei* sein. Dies unterstellt die Fähigkeit, Wiederholungen in einer Handlungsfolge zu erkennen und zu vermeiden (repetitive Kompetenz).

Das Subjekt, auf das der Verfahrensbegriff verweist, muss also mit spezifischen, durchaus anspruchsvollen Fähigkeiten wie regulativer, konsekutiver und repetitiver Kompetenz ausgestattet gedacht werden.

Der Gedanke der Bindung des „Wahrseins“ an das Vorführendkönnen eines Verfahrens ist als Versuch zu verstehen, das Problem der Unkontrollierbarkeit von prä-tendierten Wahrheitsintuitionen zu vermeiden. Ein endlicher Akteur muss sich die Wahrheit einer sprachlichen Entität gemäß dem regelgerechten, lückenlosen und zirkelfreien Durchlaufen eines Verfahrens – diskursiv, nicht-intuitiv – im Prinzip jederzeit reproduzierbar vor Augen führen können.

### 3.3.6 Zusammenfassung

Technische Artefakte, die mit sog. künstlicher „Intelligenz“ ausgestattet sind, können grundsätzlich nicht als Akteure in Begründungsdiskursen auftreten, weil ihnen die pragmatischen Merkmale der Handlungsurheberschaft und Zurechenbarkeit nicht zukommen. Wie immer in unterschiedlichen kognitiven Kontexten Handlungen in Begründungsdiskursen wie Behaupten, Bezweifeln, Zustimmung u. a. propositional bestimmt sind, technische Artefakte können diese Handlungen zwar unterstützen, aber nicht selbst ausführen.

## 3.4 Das wissenschaftliche Bacon-Projekt und die KI

Der schärfste Epochenbruch, den die Wissenschaftshistoriografie bisher beschrieben hat, ist der Übergang von der antik-mittelalterlichen Auffassung der Wissenschaft zur früh-neuzeitlichen. Dieser Übergang ist besonderer Prägnanz in der programmatischen Schrift *Novum Organon* des Francis Bacon beschrieben worden, die diese Transformation zu methodologischen Forderungen zusammenfasst; deshalb spricht man auch vom Bacon-Projekt (Schäfer 1993; Gethmann 2003). Als markante Schwellenphänomene beschreibt Bacon eine zweifache Transformation der

neuen Wissenschaft gegenüber der alten. Es geht *einmal* (a) um eine Transformation der Erkenntnisstile, die man mit der Wendung „Kontemplation versus Intervention“ zusammenfassen kann: Die Kontemplation der Natur wird abgelöst durch die Intervention in die Natur. Es geht *zum anderen* (b) um eine Transformation der Wissensformen des poetischen und praktischen Wissens, die zueinander in Beziehung gesetzt werden. Beide Unterscheidungen sind im übrigen nicht neu, neu ist aber das Verhältnis zwischen ihnen, wie Bacon es bestimmt.

### 3.4.1 *Kontemplation versus Intervention*

Nach antiker und mittelalterlicher Auffassung kann sich Wissen nur aus einer kontemplativen, also in einer der inneren Schau gewidmeten Einstellung gegenüber der Natur ergeben. Die Begründung hat Platon prägnant formuliert: Wissen kann es nur vom Allgemeinen und Unveränderlichen geben, die Welt um uns herum ist aber partikulär und veränderlich, also können wir nur durch Kontemplation des Allgemeinen zum Wissen gelangen. Demgegenüber verdankt sich neuzeitliche Wissenschaft, deren Wurzeln freilich schon im späten Mittelalter liegen, der Überzeugung, dass nur durch Intervention in Naturabläufe kausale oder konditionale Beziehungen erkennbar werden. Das Paradigma dieser Transformation des Erkenntnisstils ist das Experiment, das – entgegen einem verbreiteten Oudit über die *Naturwissenschaften* – nicht in der gegebenen Natur, sondern an einem kulturell hergestellten Ort menschlicher Arbeit – dem *Labor* – stattfindet. Dabei transformiert sich durch Bacon das Natur- wie das Technikverständnis. Natur ist nicht mehr nur die gegebene oder phänomenale Natur, sondern das entscheidende Merkmal der Natur ist das Naturgesetz, dessen Struktur mathematisch darstellbar ist; Technik ist nicht mehr, wie bei Aristoteles, das Wider-Natürliche, sondern Technik ist das, was naturgesetzlich möglich ist. Zwischen einem kontemplativ erworbenen Wissen und der Verfügungsmacht (*potestas*) über die Natur gibt es in der vor-neuzeitlichen Wissenschaft keinen Konnex, wie Bacon kritisiert. Und somit sieht man bis ins späte Mittelalter auch so gut wie keinen systematischen Zusammenhang zwischen Wissen und Naturbeherrschung.

Erst durch ein aufgrund von Intervention in Naturabläufe gebildetes Wissen wird die Allianz von Wissen und Ingenieurskunst, die ein Markenzeichen der Neuzeit ist, möglich. Verfügungsmacht tritt im Verhältnis zum Wissen an zwei Orten auf, *einmal* als Folge des Wissens. Wer der Natur durch geschicktes Isolieren von Kausalfaktoren auf die Schliche kommt, kann die Abläufe im Prinzip auch verändern, er gewinnt Macht über die Natur. *Zum anderen* aber erreicht der Mensch Verfügungsmacht über die Natur durch Intervention in sie nur dadurch, dass er bereits über ein bestimmtes Know-how der Materialbearbeitung und der Herstellung von Messgeräten verfügt (Janich 2015).

Dass dieser Umstand häufig übersehen wird, zeigt sich zum Beispiel in der üblichen Bezugsetzung zwischen Ingenieurskunst und Wissenschaft. Dass wissenschaftlich fundierte Ingenieurskunst eine Folge von Wissen ist, ist bekannt, und dass

Ingenieurskunst insoweit angewandte Wissenschaft ist, mag dahinstehen. Aber die Ingenieurskunst geht auch dem Wissen voraus. Ohne eine lebensweltlich verfügbare elementare Körperbearbeitung und die Fertigkeit im Umgang mit elementaren Geräten beispielsweise der Körperbearbeitung (bis hin zu Messgeräten) gibt es keine Labore, keine Experimente, also keine Interventionen. In diesem Sinne ist Wissenschaft auch angewandte Ingenieurskunst, und deswegen schätzen die Wissenschaftler/Wissenschaftlerinnen die Erfinder. Bacon ist wohl der erste, der die Rolle der Erfinder für die Wissenschaft mit besonderer Hochschätzung würdigt.

Wenn das Know-how eine Folge bestimmten Wissens ist, aber auch möglicherweise Bedingung eines nächsten Wissens, dann kann man sich vorstellen, dass eine Art Zusammenhang, eine Kette von Wissensformen entsteht, jeweils verbunden über bestimmte Fähigkeiten im Geräteumgang und in der Naturintervention, ein kumulativer Zusammenhang, der sich zu einem „Fort-schritt“ verkettet. Und in diesem Sinn ist Fortschritt ein spezifisches Merkmal neuzeitlicher Wissenschaft. Nicht, dass es in Antike und Mittelalter nicht gelegentlich etwas Neues gegeben hätte, aber auf dem Hintergrund des interventionistischen Erkenntnisstils läßt sich ein Handlungswissenskontinuum vorstellen, das den Wissenserwerb planbar macht. „Wissen ist Macht“ heißt auch: Fortschritt ist planbar.

### 3.4.2 *Poietisches (technisches) und praktisches Wissen*

Die zweite Unterscheidung, die wesentlich in das Diktum von Bacon eingeht, ist die Unterscheidung zwischen dem *poietischen*<sup>13</sup> und dem *praktischen* Wissen. Neu ist auch hier ihre Verhältnisbestimmung. Aristoteles bestimmt definitorisch, dass das poietische Wissen, seinen Zweck außerhalb des erkennenden Akteurs hat, etwa im Zusammenhang mit der Herstellung eines Artefakts, während das praktische Wissen seinen Zweck im Akteur selbst oder in seinen Ko-Akteuren hat. Praktisches Wissen ist das Wissen, das man braucht, um einen Freund zu beraten, um eine Gesellschaft zu organisieren, etwa durch Ausbildung von Jurisprudenz, oder das man braucht, um die Heilkunst auszuüben. Bacons Diktum verbindet nun diese beiden Aspekte auf eine eigentümliche Weise. Das poietische, das technisch-naturwissenschaftliche Wissen, das durch Intervention in Naturabläufe gewonnen wird, dient einem praktischen Zweck, der im erkennenden Akteur liegt. Etwas verallgemeinert, Wissenschaft dient der Befreiung des Menschen von natürlichen und sozialen Zwängen. Eine solche Welt – befreit von natürlichen und sozialen Zwängen – hat Bacon in seinem utopischen Reiseroman „Atlantis“, einer der früh-neuzeitlichen Utopien, vorgestellt. Neuzeitliche Wissenschaft ist nach Bacon gerade nicht zweckfrei, sondern an einen allgemeinen humanen Zweck gebunden. Sie ermöglicht im gelingen-

---

<sup>13</sup>\*Gr. ποίησις (herstellendes Handeln) gegenüber πράξις (zwischenmenschliches Handeln). Statt „poietisch“ sagt man heute „technisch“, aber τέχνη (ars, Lehre/Wissenschaft) ist bei Aristoteles beides; die Unterscheidung zwischen Wissenschaft und Technik ist also eigentlich windschief (at odds).

den und günstigen Falle, das Verfügungswissen zu erlangen, das gebraucht wird, um die Befreiung von natürlichen und sozialen Zwängen zu bewirken. Zusammenfassend kann man vom *praktischen Zweck* neuzeitlicher Wissenschaft sprechen.

### 3.4.3 *Zweck und Nutzen wissenschaftlichen Wissens*

Die Rede von der *Zweckorientierung* der Wissenschaft muss allerdings gegen Missverständnisse gesichert werden, da damit nicht selten Vorstellungen von einer Indienstnahme der Wissenschaft für partikuläre ökonomische Interessen (wessen auch immer) verbunden werden. Vor allem diese Partikularisierung des Zweckbegriffs und die damit verbundene Partikularisierung des Interessenbegriffs sind maßgeblich dafür, dass viele Wissenschaftler/Wissenschaftlerinnen für die Beschreibung ihres Selbstverständnisses auf der „Zweckfreiheit“ bestehen. Zwar ist die Debatte über einer Gesellschaftstheorie und Geschichtsmetaphysik hinweggegangen, die die Wissenschaften für partikuläre Zwecke (durch Einnahme des sog. Klassenstandpunkts) instrumentalisiert. An die Stelle dieses Theoriesyndroms ist jedoch eine verbreitete Ökonomisierung der Wissenschaft im Rahmen der Arena des Marktes getreten. Demgegenüber ist daran festzuhalten, dass Wissenschaft zwar nicht utilitären und partikulären Nutzenerwägungen unterworfen ist, gleichwohl aber einem trans-utilitären Zweck dient; „Zweck“ – im Sinne einer universellen und humanen Zwecksetzung – ist von (ökonomischem) „Nutzen“ zu unterscheiden.

### 3.4.4 *Kausalität und Korrelation*

„Big Data“ und ihre Verarbeitung durch KI setzt nach Meinung einiger Wissenschaftsphilosophen die Wissenschaften in die Lage, auf das schwierige Geschäft der Kausalerklärung zu verzichten und sie durch Korrelationsanalysen zu ersetzen. In diesem Sinne wird von einem „Ende der Theorie“ (Anderson 2008) gesprochen, nicht in dem Sinne, dass man mit dem Untergang der Theorie etwas Wertvolles verlöre, sondern in dem Sinne, dass eine lästige, eigentlich undurchführbare Aufgabenstellung, nämlich die Explikation eines adäquaten Kausalitätsverständnisses, endlich aufgegeben werden kann. Kausale Hypothesen werden unter den Rahmenbedingungen des Bacon-Projekts durch Experimente, d. h. durch Intervention in die Natur gefunden. Somit bedeutet das Ende der Theorie auch das Ende der Notwendigkeit des Experiments. An die Stelle des Experiments sollen durch Algorithmen gesteuerte Simulationen und Modellkonstruktionen treten. Damit würde die neuzeitliche Vorstellung von Wissenschaft durch KI abgelöst und somit ein neuer Epochenbruch herbeigeführt.

Mit der Losung vom „Ende der Theorie“ wird somit in erster Linie auf die Ablösung der regulativen Idee der deterministischen Kausalerklärung durch Korrelati-

onsanalysen in Form von Mustererkennung in großen Datenmengen abgehoben. Der interventionistische Erkenntnisstil zielt dagegen auf die Beherrschung der Natur durch die Aufdeckung von Ursache-Wirkungs-Verhältnissen ab, deren Erkenntnis die notwendige Bedingung dafür ist, dass man durch Handlungen planvoll in sie eingreifen kann. Somit hängt der Erfolg dieses Wissenschaftsparadigmas von der korrekten Explikation von Ursache-Wirkungs-Verhältnissen ab. Ausgangspunkt der wissenschaftsphilosophischen Bemühungen um die begriffliche Rekonstruktion von Ursache-Wirkungs-Verhältnissen ist die bahnbrechende und folgenreiche Feststellung Humes, dass man zwar die scheinende Sonne und den etwas später erwärmten Stein optisch und haptisch wahrnehmen kann, nicht aber deren zeitliche Abfolge und schon gar nicht die Ursache-Wirkungs-Beziehung zwischen den Ereignissen. Behauptungen über zeitliche Sukzessivität und Kausalität zwischen zwei oder mehreren Ereignissen haben keine Referenzbeziehung zu wahrnehmbaren Gegenständen bzw. Ereignissen, sondern verbinden diese vielmehr durch bereits investierte Kategorien der zeitlichen Abfolge bzw. der Verursachung. Zeitliche Abfolge und Ursache-Wirkungs-Verhältnisse gehören damit neben anderen Kategorien zu den semantischen Netzwerken, über die das erkennende Subjekt bereits verfügen muss, wenn es Ereignisse zueinander in Beziehung setzt.

Hinsichtlich des Verhältnisses von Korrelationsanalyse und Kausalerklärung interagiert die begriffliche Organisation und Interpretation interventionistischer („experimenteller“) Forschung mit der grundsätzlichen Interpretation des Wahrscheinlichkeitsbegriffs. Wer der Überzeugung ist, dass die Welt „an sich“ deterministisch organisiert ist, Wahrscheinlichkeitsaussagen sich daher lediglich (vorläufigen) Erkenntnisdefiziten verdanken („Gott würfelt nicht.“), wird die Relationen zwischen unterschiedlichen Ereignissen vollständig und disjunkt in kausale und korrelative einteilen; es handelt sich dabei um eine *kontradiktorische* Unterscheidung. Wer demgegenüber unterstellt, dass Relationen zwischen Ereignissen grundsätzlich probabilistisch organisiert sind, wird diese Relationen auf einem Kontinuum (zum Beispiel zwischen 0 und 1) anordnen, dessen Endpunkte sich *polar-konträr* gegenüberstehen. Die damit angedeutete Kontroverse ist eine innerhalb des interventionistischen Paradigmas.

Sowohl Deterministen wie Probabilisten unterstellen ein Verständnis von Kausalität, das ihre Feststellung bezüglich des Zusammenhangs (Korrelation) bestimmter Ereignisse begründet. Der Determinist will kausale Zusammenhänge von bloß zufälligen, der Probabilist starke (signifikante) von schwachen (kontingenten) Korrelationen unterscheiden. Das Kausalitätsprinzip spielt somit eine bezüglich der hier betrachteten Positionen invariante kriteriale Rolle, die grundsätzlich unvermeidbar ist. Für das wissenschaftliche Geschäft wäre daher ein klares Verständnis von Kausalität unverzichtbar. Bedauerlicherweise ist die Rekonstruktion dieses Verständnisses jedoch eine der wissenschaftstheoretischen Baustellen, deren prekärer Zustand sogar zu tiefer Skepsis gegenüber dem interventionistischen Erkenntnisstil führen könnte.

Die wichtigste Funktion des Kausalitätsprinzips ist die Abwehr sogenannter ursachenloser Ereignisse oder Ereignisse mit obskuren Ursachen (zu denen nach Hume auch wundersame göttliche Eingriffe gehören; s. o. Abschn. 3.3.2). Das Kau-

salitätsprinzip ist das grundlegende Rationalitätsprinzip der neuzeitlichen interventionistischen Wissenschaften; würde man für bestimmte Ereignisse das Kausalitätsprinzip außer Kraft setzen, würde man auch kausal unerklärte oder durch obskure Ursachen erklärte Ereignisse in den Wissenschaften zulassen. Unter der Bedingung der Gültigkeit des Kausalitätsprinzips sind Situationen, in denen man eine Ursache nicht benennen kann, grundsätzlich als Antrieb zu verstehen, die Ursachenforschung fortzusetzen, da ausgeschlossen ist, dass es keine oder nur obskure Ursachen gibt. Es ist offenkundig, dass die Ausdehnung des Kausalitätsprinzips auf im klinischen Kontext diagnostizierte Symptome von großer Bedeutung für das Verständnis der medizinischen Praxis (gegenüber manchen sogenannten alternativen Medizinverständnissen) ist.

Während das Kausalitätsprinzip als Grundlage neuzeitlicher interventionistischer Wissenschaft weithin unumstritten ist, beginnen die Probleme mit der Explikation von *Kausalgesetzen*. Kausalgesetze sind allgemeine Behauptungen über Ursache-Wirkungs-Verhältnisse zwischen Wirkungseignissen und Ursacheereignissen. Nachdem Hume konstatiert hat, dass Ursache-Wirkungs-Verhältnisse nicht Referenzobjekte der menschlichen Wahrnehmung oder menschlicher Wahrnehmungs-Behauptungen sein können, wird in der Wissenschaftsphilosophie kontrovers diskutiert, welche Mindestbedingungen erfüllt sein müssen, um zu gültigen Kausal-Behauptungen zu gelangen.

Der Ansatz von G.H. von Wright schließt sich unmittelbar an den interventionistischen Erkenntnisstil an und versucht, das Verhältnis von Ursache und Wirkung unter Einbeziehung *handlungstheoretischer Kategorien* zu erklären (von Wright 1971). Von Wright weist auf die alltägliche Handlungserfahrung hin, zu der es gehört, dass der Akteur sicher ist, durch bestimmte Handlungen bestimmte Wirkungen herbeiführen zu können. Dabei wird „Handeln“ grundsätzlich nicht als physikalisch beschreibbaren Episode betrachtet (dies würde ersichtlich zu Zirkelproblemen führen), sondern als Zweck-Realisierungsversuch. Betrachtet man eine Handlung als absichtliche (geplante) Herbeiführung einer Wirkung, wird dagegen ein Kausalitätsgesetz noch nicht vorausgesetzt, da die Wahrheit von Aussagen über den Zusammenhang von verursachender Handlung und bewirkter Zweck-Erreichung keine methodische Vorbedingung für die Möglichkeit des Handelns ist. Die Allgemeinheit von Naturgesetzen im Sinne regelmäßig wiederkehrender (kausaler) Verbindungen von Ereignissen wird durch ein aus Handlungskontexten verallgemeinertes Verfügungswissen gewonnen, das sich in experimentellen Zusammenhängen bewähren muß. Dabei ist zu berücksichtigen, dass Experimente ihrerseits technisch hergestellte Zweckkontexte sind, die letztlich auf erfolgreich ausgeführte Herstellungshandlungen zurückzuführen sind (Janich 1997). Falsche Kausalitätsbehauptungen beruhen dann auf unrichtig ausgeführten Handlungssequenzen und sind durch Fehler in diesen erklärbar. Der handlungstheoretische Ansatz gibt somit grundsätzlich Kriterien für die Unterscheidung von wahren und falschen Kausalbehauptungen an die Hand. Der handlungstheoretische Ansatz steht jedoch mit dem Kausalitätsprinzip als grundlegendem Rationalitätsprinzip interventionistischer Wissenschaft nicht in Einklang, wenn man das Kausalitätsprinzip als generelle Bedingung für wissenschaftliches Wissen betrachtet. Handlungen im Sinne eines „finalistischen“ (inten-

tionalistischen) Handlungsverständnisses sind danach selbst keine möglichen Gegenstände interventionistischer Wissenschaft. Entsprechend kommt man bei einem so verstandenen Instrumentalismus nicht darum herum, die Existenz unterschiedlicher Typen wissenschaftlicher Wissensbildung zu unterscheiden.

### **3.4.5 Die Bedeutung von Big Data für die Wissenschaft**

Am Anfang des Prozesses der Genese von Informationen stehen elektrotechnisch durch Folgen von 0 und 1 hergestellte Signale, die zu Symbolen codiert werden. Diese werden durch syntaktische Regeln zu Daten zusammengeführt. Solche Daten werden in einen Kontext gesetzt und bekommen dadurch Bedeutung. Erst durch diese Bedeutung erhält man eine Information. So bedeutet das Datum 39,8 mit der Maßeinheit Celsius im Kontext der Medizin Fieber. Informationen werden dann mit anderen Informationen zu Wissen verbunden, beispielsweise die Diagnose einer Krankheit, um auf dieser Grundlage Probleme zu lösen, wie zum Beispiel eine Therapie einzuleiten. Durch die moderne Entwicklung u. a. im Bereich der Speichertechnik geht es heute allerdings nicht mehr nur um strukturierte Daten, sondern vor allem um unstrukturierte Daten. Dadurch entsteht eine gewaltige, amorphe Datenmasse, die durch die herkömmlichen Datenbanktechnologien und Algorithmen nicht bewältigt werden kann. Neuartige Such-Algorithmen, die diese Datenmassen nach Datenkorrelationen und Datenmustern durchsuchen, führen schließlich die Ergebnisse zusammen, um daraus beispielsweise Trends, Profile von Produkten und Personen abzuleiten. Diese Datenmassen führen zu einer tiefgehenden Veränderung in der Wissenschaft.<sup>14</sup> Zu den herkömmlichen Datenkorrelationen und daraus entwickelten Hypothesen kommen Machine Learning Algorithmen; durch das dadurch mögliche In-silico-Experiment mit Computersimulation und der Verbindung mit dem herkömmlichen In-vitro-Experiment gibt es die Möglichkeit, zu neuem Wissen zu kommen. Beispielsweise lässt sich anhand der Genomsequenz eines patientenspezifischen HIV-Erregers die Resistenzwahrscheinlichkeit für bestimmte Wirkstoffe berechnen und daraus das spezifische Profil eines Patienten berechnen. Zusammen mit dem „Internet der Dinge“ können so in Großzentren der Medizin Daten der Patienten und der Ärzte mit der technischen Infrastruktur zusammenwachsen.

### **3.4.6 Zusammenfassung**

Algorithmen sind auf dem Hintergrund von Theorien und Gesetzen zu interpretieren und auch kritisch zu beurteilen. Korrelationen und Datenmuster ersetzen keine Erklärungen und Begründungen von Ursachen. Das heißt, diese Daten müssen nicht nur qualitativ evaluiert, sondern auch normativ beurteilt werden. Die Algorithmen

---

<sup>14</sup>Vergl. Schmidt in diesem Band, Kap. 4.

müssen auf ihre Zwecke und Ziele hin überprüft werden. Diese Überprüfung kann nur von einem Wesen vorgenommen werden, das in der Lage ist, Handlungen als Zweckrealisierungsversuche auszuführen, zu verstehen und zu überprüfen. In diesem Sinne bleibt auch durch die KI-induzierten Innovationen das Bacon-Projekt in seinen Grundstrukturen erhalten.

## 3.5 Deterministische und probabilistische Voraussagen

Kausalerklärungen haben den großen pragmatischen Vorteil, dass sie sichere Voraussagen (Prognosen) ermöglichen. Viele Wissenschaftsphilosophen sehen in der Prognosefähigkeit – und nicht in einer repräsentativen Abbildung der „realen“ Wirklichkeit – sogar die zentrale operative Bedeutung von interventionistischen „Theorien“ (van Fraassen 1980). Korrelationsanalysen auf der Basis großer Datenmengen gelangen demgegenüber grundsätzlich nicht zu derartigen deterministischen Prognosen. Ihre Zukunftsantizipationen, die unter dem Begriff der *predictive analytics* entwickelt werden, hängen von probabilistischen Präsuppositionen ab.

### 3.5.1 Erklären und Voraussagen

Deterministische Voraussagen in der Physik unterstellen ein deterministisch organisiertes Referenzobjekt, wie es beispielsweise für das Planetensystem angenommen wird. Das Referenzobjekt von Big Data ist jedoch grundsätzlich probabilistisch organisiert. Der Ersetzung von Kausalerklärungen durch Korrelationsanalysen entspricht – aufgrund der Struktur-Isomorphie von Erklärungen und Voraussagen – die Ersetzung von unkonditionierten Voraussagen durch konditionierte Voraussagen, wie sie bei Wettervorhersagen verwendet werden. Die Voraussage der nächsten Sonnenfinsternis ist danach strukturell etwas völlig anderes als die Voraussage des nächsten Regenschauers. Wettervorhersagen beruhen nicht auf Kausalerklärungen, sondern auf mehr oder weniger großen in der Vergangenheit gewonnenen Datenmengen, die ein zukünftiges Ereignis mehr oder weniger nahelegen. Durch Mustererkennung bezogen auf Big Data und neue Techniken der Datenanalyse wächst die Sicherheit der Vorhersage, nähert sich in günstigen Fällen der Sicherheit der deterministischen Voraussage an, fällt aber grundsätzlich nicht mit ihr zusammen. Wenn es um lapidare pragmatische Probleme geht, im Beispiel: die Frage, ob man den Regenschirm mitnimmt, mag es dann zwischen diesen Fällen keinen pragmatischen Unterschied geben. In anderen pragmatischen Kontexten kommt es aber gerade darauf an, ob man von einer Restunsicherheit ausgehen muß oder nicht. Ob man einen Regenschirm unnötigerweise mitgenommen hat, mag in den meisten Fällen belanglos sein. In medizinischen Kontexten, in denen es nicht selten um Leben und Tod geht, versucht man daher, Restunsicherheiten falsch-positiver oder falsch-negativer Voraussagen möglichst einzugrenzen.

### 3.5.2 *Fehlschlüsse: Stratifizierung*

Wird der Unterschied zwischen die deterministischen Voraussagen und probabilistisch konditionierten Voraussagen nicht beachtet, stellen sich typische Fehlschlüsse ein. Beispielsweise werden durch auf der Grundlage von Big Data erstellten Korrelationen – durch Einteilung einer statistischen Stichprobe in kleinere Gruppen – sog. „Strata“ – zugeordnet (etwa regional definierten Kohorten wie die Anwohner einer Straße, Alterskohorten, Kohorten, die durch diagnostische Merkmale definiert sind), die ohne jede soziale Erlebnisqualität sein können. Die Bildung solcher Kohorten und die auf ihrer Basis durch Algorithmen produzierten Voraussagen können für den Verwender durchaus nützlich und ihre durch die Rest-Unsicherheit erzeugte mögliche Fehlermarge mit Bezug auf seine Zwecke erträglich sein. Für das Individuum mag der Effekt einer solchen statistischen Stratifizierung (etwa durch Verweigerung eines Kredits) ein erheblicher Nachteil sein, für den Verwender der Daten, (beispielsweise eine Bank) in der Summe aufgrund der summierten Risikovermeidung ein Vorteil.

Besonders prekär ist diese strukturelle Situation allerdings im Gesundheitsbereich, wenn ein Individuum aufgrund diagnostizierter Merkmale einem „Stratum“ zugeordnet und aufgrund von Algorithmen eine therapeutische Strategie festgelegt wird. Man kann von einem statistischen Kollektivismus sprechen, gegenüber dem das Individuum geltend machen können muß, als solches betrachtet und behandelt zu werden. Das bedeutet, dass im Rahmen der Medizin als praktischer Wissenschaft eine korrelative Zuordnung eines Individuums zwar ein gutes diagnostisches Indiz sein kann, das wiederum eine gute praktische Heuristik rechtfertigt, aber grundsätzlich nicht die einzige Evidenzgrundlage für eine Behandlung sein darf. Epidemiologische Studien grenzen entsprechend denkbare prioritäre Diagnosen ein und legen auf dieser Grundlage gewisse prioritäre Therapien nahe, dürfen jedoch nie allein die Grundlage für eine Behandlung sein. Der individuelle Patient kann immer ein statistischer Ausreißer sein, mag die Wahrscheinlichkeit auch noch so klein sein. Daher ist die Gefahr des „epidemiologischen Fehlschlusses“ immer zu beachten. Eine genetische Disposition beispielsweise kann durchaus zu unterschiedlichen Wahrscheinlichkeitsverteilungen bei großen Kollektiven führen, aber diese determinieren nicht das Individuum. Sollte es eine Wahrscheinlichkeitskorrelationen zwischen der Disposition „Musikalität“ und der Inzidenz von Magenkarzinomen geben, besagt das für das vor dem Arzt stehende Individuum so gut wie nichts. Big-Data Vorhersagen tragen entsprechend wesentlich einen Rest an Ungewissheit in sich, die nicht auf Unzulänglichkeit der eingesetzten Daten zurückgeführt werden kann. Auch auf der Basis von Gesetzeswissen gibt es Ungewissheiten; diese sind jedoch jedenfalls prinzipiell zu überwinden. Ein Akteur, der sich Big-Data-Vorhersagen zunutze machen will, muss daher entscheiden, welche Kosten (im monetären oder nicht-monetären Sinn) er zu übernehmen bereit ist.

### 3.5.3 *Beispiel: evidenzbasierte Medizin*

Diese pragmatischen Probleme bezüglich aus Big Data gewonnener Voraussagen werden noch einmal um eine Dimension verschärft, wenn das Referenzobjekt, auf das sich Wahrscheinlichkeitsaussagen beziehen, durch diese Aussagen selbst beeinflussbar ist, wie das für alle Kontexte menschlichen Handelns, z. B. die Volkswirtschaft oder das Gesundheitssystem gilt.

Die aus epidemiologischen Studien gewonnenen konditionierten Voraussagen können bei passender öffentlicher Verbreitung kollektive Verhaltensweisen erheblich verändern, derart, dass die Datengrundlage nicht mehr verlässlich ist. Von weitreichender praktischer Bedeutung ist die Heranziehung großer Datenmengen für das Konzept der evidenzbasierten Medizin.<sup>15</sup> Diese stützt sich vorzugsweise auf Erfahrungen zur Indikation, Wirksamkeit und Sicherheit von Therapieverfahren, die in Form von Expertenbewertungen oder systematischen Reviews („Studien“) zur Verfügung gestellt werden. Evidenzbasierte Medizin in diesem Sinne ist ein mittels nachvollziehbarer Kriterien standardisierter Prozeß der methodisch-rationalen Informationsreduktion und der Bewertung des jeweiligen, in Studien dokumentierten Standes der Wissenschaft. Die Anwendung des gesicherten und aktuellen Wissens setzt Einvernehmen darüber voraus, was als aktuelles und gesichertes Wissen zu verstehen ist. Die traditionell übliche Formel ‚Stand des Wissens‘ ist offenbar als unzureichend anzusehen. Die evidenzbasierte Medizin präzisiert deshalb diesen Stand des Wissens durch die Forderung, dass dieser durch die wissenschaftliche Literatur dargestellt wird, wobei nur diejenigen Publikationen als wissenschaftlich gewertet werden, die definierte wissenschaftliche Kriterien der Korrelationsanalyse erfüllen. In diesem Zusammenhang ist auch die Forderung nach direkt-vergleichenden randomisierten Studien als Mittel zur Verzerrungsvermeidung bei korrelativen Analysen zu verstehen. Aus diesen Studien ergeben sich die Kriterien der Indikation für die Intervention. Durch sie wird der Anwendungsbereich für therapeutische Interventionen bestimmt, wie er beispielsweise in der amtlichen Zulassung von Arzneimitteln und Medizinprodukten festgelegt wird. Wesentliche Dimensionen für die Beurteilung der Wirkungen von Interventionen sind Eintrittswahrscheinlichkeit, Erheblichkeit und Behandelbarkeit, wobei die Eintrittswahrscheinlichkeit in diesem Zusammenhang von besonderem Interesse ist. Die Eintrittswahrscheinlichkeit betrifft alle Arten von gesundheitsbezogenen Ereignissen wie Symptome, Testergebnisse, Behandlungserfolge, Krankheitsresiduen, Behinderungen (Funktions- und Leistungseinschränkungen). Grundsätzlich unterstellt also die evidenzbasierte Medizin ein probabilistisches Kausalitätsverständnis, und sie stellt somit probabilistisch konditionierte Voraussagen zur Verfügung. Dieses bietet für das ärztliche Handeln zweifellos eine sicherere Grundlage als bloß narrativ präsentierte okkasionelle Erfahrung. Dennoch ist der Ansatz der evidenzbasierten Medizin pragmatisch schwächer als eine deterministische Kausalanalyse, sofern man unterstellt, dass ein Kausalgesetz epistemisch mehr bietet als jede Korrelationsbe-

---

<sup>15</sup>Vergl. Humm et al. in diesem Band, Abschn. 2.3.3.

ziehung. Die Gefahr von Stratifizierung und epidemiologischen Fehlschlüssen ist grundsätzlich gegeben. Diese Gefahren lassen es als systematisch unmöglich erscheinen, Diagnose und Therapie generell durch Big Data-gestützte predictive analysis zu ersetzen.

### **3.5.4 Zusammenfassung**

Was für die Diagnose und Therapie durch den den Einzelfall würdigenden Arzt gilt, gilt analog für das Judiz des Richters, das pädagogische Ermessen des Lehrers und die kreative Einschätzung eines wissenschaftlichen Problems durch den Forscher/die Forscherin. In allen Fällen ist Big Data unter günstigen Bedingungen ein Hilfsmittel, aber kein geeignetes Instrument der definitiven Lagebeurteilung und Entscheidung. Pragmatische und heuristische Faktoren wie Prüfung der Kohärenz mit anderen Evidenzquellen, Erfolgseinschätzungen u. a. spielen eine nicht zu vernachlässigende Rolle.

## **3.6 Methodische Verfahren und heuristische Urteilskraft**

Bis in das 20. Jahrhundert folgte die methodologische Grundvorstellung des Bacon-Projekts dem Verifikationismus, d. h. dem Versuch, wissenschaftliche Hypothesen durch Verfahren des „Wahrmachens“ auszuweisen. Allerdings war spätestens seit Hume („Humesches Problem“) geläufig, dass endlich viele Bestätigungsinstanzen, beispielsweise Experimente, aus logischen Gründen nicht einen generalisierenden Satz vom Typ „Alles Kupfer leitet Elektrizität.“ wahr machen können. Aus dieser methodologischen Verlegenheit hat Karl Raimund Popper (1934) durch den Vorschlag geführt, den Verifikationismus durch den Falsifikationismus zu ersetzen. Danach soll man nicht versuchen, Hypothesen durch Erfahrung zu beweisen, sondern sie versuchsweise durch Erfahrung zu widerlegen. Wenn eine Hypothese dem Widerlegungsversuch widersteht, dann hat sie sich vorläufig bewährt. Aufgrund der Anzahl der (unabhängigen) Widerlegungsversuche läßt sich der Grad der Bewährung einer Hypothese angeben. Poppers Falsifikationismus weist eine starke Affinität mit dem von C.S. Peirce eingeführten Fallibilismus auf; danach geht es bei den menschlichen Erkenntnisbemühungen nicht darum, „Gewißheit“ zu erlangen („Certismus“), sondern Irrtümer zu vermeiden, soweit das einem endlichen Wesen überhaupt möglich ist (Spinner 1974). Der Grundgedanke des Falsifikationismus hat sich in den empirischen (sich auf Experimente, Befunde oder Befragungen stützenden) Wissenschaften weitgehend faktisch etabliert. Man kann von einem falsifikationistischen Ethos im Sinne eines meistens nicht explizit in Frage gestellten common sense der Wissenschaftler/Wissenschaftlerinnen sprechen.

### 3.6.1 *Falsifizierbarkeit und Falsifikation*

Der Falsifikationismus bietet eine Antwort-Strategie für zwei unterschiedliche wissenschaftsphilosophische Probleme an. Einmal geht es um die Abgrenzung von wissenschaftlich sinnvollen gegenüber wissenschaftlich sinnlosen Aussagen. Zu den sinnlosen Aussagen gehören nicht die Aussagen, die aufgrund einfacher syntaktischer („Aua!“) oder semantischer („Alle Bulgaren sind Primzahlen.“) Defizite nicht falsifizierbar sind, sondern vor allem solche, die wegen grundsätzlich empirisch nicht überprüfbarer Unterstellungen der Falsifikation entzogen sind. Dies gilt insbesondere für in Anspruch genommene wissenschaftliche „Theorien“, die inhaltliche Falsifikationsverfahren durch eine Immunisierungsstrategie verbieten.<sup>16</sup> Falsifizierbarkeit ist also zum einen ein Abgrenzungskriterium. Darüber hinaus bildet das Verfahren der Falsifikation ein Wahrheitskriterium im Sinne eines Bewährungskriteriums. Um eine Theorie gleich welcher epistemischen Herkunft (von der bloßen Fantasie bis zu Annahmen, die sich aus anderen Theorien nahelegen) zu überprüfen, muss folgende Schrittfolge vollzogen werden:

- (i) Aus Theorie-Kandidaten werden mit Hilfe der Logik Hypothesen deduziert.
- (ii) Aus diesen Hypothesen werden Prognosen deduziert; diese Prognosen können empirisch (durch Experiment, ...) widerlegt werden. Gelingt diese Widerlegung nicht, gilt die Theorie in Abhängigkeit von Häufigkeit und Härte des Testverfahrens als mehr oder weniger gut bestätigt. Dabei wird die Reproduzierbarkeit unterstellt, denn ein empirischer Test, der nur ein einziges Mal durchgeführt werden kann, ist kein Verfahren. Falsifizierbarkeit setzt Reproduzierbarkeit voraus. Bei der Falsifikation geht es nicht um eine generelle Erkenntnissicherung (wie dies im Rahmen des Fallibilismus unterstellt wird). Primäre Erfahrungen, wie sie Menschen durch ihre fünf Sinne im Alltag „widerfahren“ (nicht: „machen“), sind ausdrücklich nicht angesprochen. Vielmehr geht es um sekundäre Erfahrungen, die wesentlich auf Verallgemeinerung im Sinne von Hypothesen, Modellen, Theorien beruhen. Bereits der Einsatz von Messgeräten im Rahmen von Experimenten, die der Widerlegung von Prognosen dienen sollen, beruht auf Verallgemeinerungen und führt immer zu sekundärer Erfahrung. Vom einfachen Thermometer bis zum komplexen Geiger-Zähler sind empirische Testverfahren von vornherein theorie-imprägniert. Modelle. Aus ihnen lassen sich mathematisch formulierte Szenarien bilden, die viele Einzelbeobachtungen aus ganz unterschiedlichen Erkenntnisquellen zusammenfassen, um zu Verallgemeinerungen zu kommen. Das Verfahren der Falsifikation muß daher bei einer sehr schnell erreichten Stufe von Komplexität in zunehmendem Maße Vereinfachungen und Idealisierungen im Rahmen von Modellierungen vornehmen. Quantität und Qualität von Modellierungen werden nicht durch das

---

<sup>16</sup>Popper nennt insbesondere die Marxistische Gesellschaftstheorie (Wer diese in Frage stellt, gilt als Klassenfeind.), die Freudsche Psychoanalyse (Wer diese in Frage stellt, gilt als besonders schwer erkrankt) und die Christliche Metaphysik (wer diese in Frage stellt, gilt als Ungläubiger.) (Popper 1962). Zu Immunisierungsstrategien vgl. Albert (1968).

Referenzobjekt der Theorien festgelegt, sondern hängen wesentlich vom Zweck ab, der mit dem Geltungsanspruch der Theorie verbunden ist.

### ***3.6.2 Grenzen und Schwierigkeiten des Verfahrens der Falsifikation***

Ein großes Feld der wissenschaftsphilosophischen Diskussion des 20. Jahrhunderts besteht in der Auseinandersetzung mit Grenzen und Schwierigkeiten des Verfahrens der Falsifikation und Versuchen ihrer Überwindung. Eine bedeutende Grenze liegt darin, dass es sich ausschließlich auf empirisches Wissen bezieht, wie es beispielsweise durch Experimente bestätigt wird. Die für diese Form von Wissen konstitutiven Formalwissenschaften (vor allem Logik und Mathematik) fallen dabei völlig heraus. Zwar gibt es für diese Disziplinen kein Humesches Problem. Abgesehen davon, dass bei weitem nicht für alle interessanten Kalküle schon Widerspruchsfreiheitsbeweise geführt worden sind, gibt es bezüglich des Einsatzes formaler Instrumente aber durchaus ein Wahlproblem. Ob beispielsweise für die Modellierung von Wolkenbildung die Monte Carlo Methode der Wahrscheinlichkeitstheorie angemessen ist oder nicht, ist eine Frage der pragmatischen Adäquatheit, für deren Entscheidung es keine durch eine Methodologie vorgegebene Prozedur gibt. Ferner kann das Verfahren der Falsifikation nicht auf solche Wissensformen bezogen werden, die konstitutiv dafür sind, dass das Verfahren überhaupt sinnvoll ist. Es geht um grundlegende methodologische Postulate wie die der Konsistenz. Bezüglich dieser Wissensformen gibt es keine empirische Kontrolle, vielmehr bedarf es begrifflicher Überlegungen bezüglich der Angemessenheit und Reichweite dieser Postulate. Dabei spielt auch der Anwendungserfolg dieser Postulate eine Rolle, wobei die Festlegung der Maßstäbe des Erfolgs durchaus problematisch ist. Aber auch im Raum des empirischen Wissens gibt es erhebliche Stolpersteine. In den Naturwissenschaften gibt es keineswegs nur generelle Aussagen vom Typ der Gesetzesaussagen. Es gibt durchaus singuläre Phänomene und entsprechende Existenzbehauptungen (vom Typ „Es gibt einen 80. Mond des Jupiter.“ „Dies ist das Schlüsselbein des Australopithecus.“), die nicht nach dem Standardverfahren der Falsifikation auf ihre Bewährtheit hin überprüft werden können. Auch einmalige Groß-Phänomene („Das Universum nach dem Urknall.“ „Das Weltklima.“) haben die ontologische Eigenheit, dass es gewissermaßen keinen Außenraum gibt, von dem aus falsifizierende Phänomene vorgebracht werden können. Erhebliche methodologische Komplikationen werfen Hypothesen über ferne Zeiten und Räume auf. Da die einfache Falsifikation durch Abwarten („Das Klima im Jahr 3000 ...“) oder durch Fernreisen („Der Exoplanet in 1000 Lichtjahren Entfernung ...“) als Teststrategie entfällt, müssen indirekte Ersatzstrategien herangezogen werden, bei denen Hypothesen eingesetzt werden müssen, die ihrerseits einem Falsifikationsverfahren zu unterwerfen wären. In anderer Weise sind komplizierte indirekte Testverfahren einzusetzen, wenn es sich um Wahrscheinlichkeitsaussagen handelt, die sich auf ein empirisches Referen-

zobjekt beziehen. Wahrscheinlichkeitsaussagen enthalten gewissermaßen die Falsifikationsinstanzen schon in sich. Während der Satz „Alle Schwäne sind weiß.“ leicht zu falsifizieren wäre, wirft der Satz „Wahrscheinlich sind alle Schwäne weiß“ erhebliche Probleme auf.

### 3.6.3 *Unterbestimmtheit wissenschaftlicher Theorien*

Sowohl verifikationistische als auch falsifikationistische Methodologien unterstellen zunächst, daß zwischen Instanzen der Bestätigung und der mehr oder weniger bestätigten Theorie eine Relation besteht der Art, daß die Theorie um so mehr bestätigt ist, je mehr Instanzen für sie bzw. je weniger Instanzen gegen sie sprechen. Diese einfache Bestätigungsrelation ist schon dadurch in Frage gestellt, dass die Qualifizierung der Bestätigung einer Theorie auch von der absoluten Zahl der Verifikations- bzw. Falsifikationsverfahren abhängt, mehr aber noch von der Nicht-Trivialität der Bestätigungsinstanzen, beispielsweise der Experimente. Beide Gesichtspunkte stehen der intuitiv naheliegenden Vermutung im Wege, den Bestätigungsgrad einer Theorie einfach durch die Zahl der Verifikations- bzw. Falsifikationsinstanzen bestimmt sein zu lassen. Noch gravierender ist der Umstand, dass die logische Beziehung zwischen Bestätigungsinstanzen und Theorien in beiden Richtungen uneindeutig ist. Ein Experiment kann viele Theorien bestätigen und eine Theorie kann von unterschiedlichen Bestätigungsinstanzen bestätigt werden. Die Uneindeutigkeit wird noch einmal durch die zuerst von C.G. Hempel (1946) formulierte Paradoxie der Bestätigung (Rabenparadoxie) verschärft. Demgemäß bestätigt die Beobachtung eines weißen Schwans mit der Standardlogik<sup>17</sup> den Satz „Alle Raben sind schwarz.“ bestätigt.

Eine Folgerung aus der Hempelschen Paradoxie ist das Unbestimmtheitstheorem (Duhem-Quine These). Danach kann eine Theorie durch sehr verschiedene (sogar widersprüchliche oder disparate) Mengen von Evidenzen bestätigt werden. Vice versa gilt, dass eine konsistente Menge von Evidenzen viele Theorien bestätigen kann; man erhält damit ein Nachfolgeproblem des Humeschen Induktionsproblems, das man als Problem des semantischen Überschusses einer Theorie bezeichnen kann: Endlich viele Evidenzen reichen nicht aus, um einen generellen Satz (z. B. ein Naturgesetz) zu bestätigen. Anders formuliert: Eine Theorie weist immer einen (nicht bestätigten) Überschuß über das empirische Wissen aus. Aus beiden Problemen folgt zusammenfassend das Mehrdeutigkeitsproblem: Das Verhältnis der Mengen primärer Erfahrung zu wissenschaftlichen Theorien ist mehr-mehr-deutig. Primäre Erfahrung (z. B. durch ein Experiment erzeugt) bestätigt viele Theorien; eine Theorie wird durch verschiedene (sogar widersprüchliche) Erfahrungen bestätigt. Es bedarf also materieller Kriterien, die beispielsweise – mit Blick auf die Raben-Paradoxie – Schwäne grundsätzlich nicht als Bestätigungsinstanzen zulassen. Da-

<sup>17</sup>Wegen  $\Lambda x (Fx \rightarrow Gx) \leftrightarrow \Lambda x (\neg Gx \rightarrow \neg Fx)$  [kl. Kontraposition] gilt Best  $[(\neg Fa \wedge \neg Ga), \Lambda x (Fx \rightarrow Gx)]$ . Vgl. zum Logikwahlproblem in diesem Zusammenhang Gethmann (1980).

mit entsteht jedoch das gegenläufige Problem, dass man jede Theorie durch die Auswahl von Bestätigungsinstanzen verifizieren bzw. falsifizieren kann, indem man die falsifizierenden Instanzen als inhaltlich nicht einschlägig ausschließt oder als einschlägig einschließt. Die Unterbestimmtheit läßt sich auch schon am einfachen Verfahrensschema der Falsifikation erkennen: auch wenn die Auswahl der möglicherweise falsifizierenden Instanzen aus den Hypothesen logisch deduziert wird, wird dadurch keine inhaltliche Auswahl dieser Instanzen vorgegeben. Erst recht besteht keine eindeutige logische Beziehung zwischen der angeblich falsifizierenden Instanz und der Hypothese, die falsifiziert wird. In dieser Richtung handelt es sich um ein Schlußverfahren vom Typ der „Abduktion“. Der Wissenschaftler, der für die zur Debatte stehende Hypothese das Verfahren der Falsifikation durchlaufen will, muß also im Vorhinein unter Inkaufnahme möglicher Zirkel im Interesse der epistemischen Kontrolle pragmatische Kriterien heranziehen. Zu diesen gehören: wechselseitige Kohärenz mit anderen Theorien, Analogien zu anderen Wissensbereichen, der Blick auf andere Meinungen der Wissenschaftlergemeinschaft und andere Verfahren kluger Abwägung. Auf die Rolle der Heuristik im Sinne regelüberwachter Suchverfahren wurde bereits hingewiesen (siehe Abschn. 3.3.5 dieses Kapitels).

### ***3.6.4 Theoriendynamik und die epistemische Funktion von Wissenschaftlergemeinschaften***

Die dargestellten Probleme, das Verfahren der Falsifikation eindeutig zu rekonstruieren, haben seit der Mitte des 20. Jahrhunderts zu erheblichen Zweifeln an der Vorstellung geführt, der faktische Forschungsprozess entwickle sich entlang der Befolgung eines normativen Sets von methodischen Regeln. Statt dessen wurde der Blick auf historisch-kontingente Prozesse der Wissensbildung in den Wissenschaften gelenkt, insbesondere die sozialen Interaktions- und Kommunikationsverfahren innerhalb der Wissenschaftlergemeinschaften in Wechselwirkung mit wissenschaftsexternen Einflussfaktoren. Besonders einflussreich war dabei zunächst die Konzeption der diachronen Wissenschaftsphilosophie, die der amerikanischen Wissenschaftshistoriker und -philosoph Th. S. Kuhn in mehreren Anläufen entwickelt hat (Kuhn 1970, 2000). In Auseinandersetzung mit der Vorstellung, dass die Wissenschaftsentwicklung kumulativ-kontinuierlich, nach wissenschaftsinternen (methodologischen) Rationalitätskriterien verlaufe, beschreibt Kuhn einen diskontinuierlichen, auch von wissenschaftsexternen (sozialen) Faktoren bestimmten Wandel in der Wissenschaftsgeschichte. Der zentrale Begriff für die Erklärung des Phänomens des diskontinuierlichen Wandels ist der des Paradigmas, womit der Sachverhalt bezeichnet werden soll, dass einige anerkannte Vorbilder konkreter wissenschaftlicher Praxis die Muster liefern, gemäß denen die Wissenschaftlergemeinschaften ihre methodischen und sozialen Entscheidungen fällen. Was in einer Disziplin ‚rational‘ heißt, hängt danach jeweils von einem solchen Paradigma ab. Kuhn unter-

scheidet näherhin drei Phasen der Wissenschaftsentwicklung: (i) In der ‚vor-normalen (vor-paradigmatischen) Phase‘ hat sich noch kein verbindliches Paradigma durchgesetzt, vielmehr ist ein Pluralismus von konkurrierenden Ansätzen festzustellen; die so genannten Tatsachen werden im Rahmen der verschiedenen Ansätze unterschiedlich gewichtet. Es gibt keine einheitlichen Methodologien und keinen festumrissenen Wissenskanon. Demgegenüber ist (ii) in der ‚normalen Phase‘ ein wissenschaftliches Paradigma für die Definition einer Disziplin nach innen und außen leitend. Es besteht eine einheitliche Methodologie und eine allgemein von den Mitgliedern der Wissenschaftlergemeinschaft anerkannte Kommunikationsstruktur (z. B. wissenschaftliche Zeitschriften); das bereits errungene Wissen ist kanonisiert, z. B. in anerkannten Lehrbüchern niedergelegt. (iii) Die ‚revolutionäre Phase‘ besteht in einem sich mehr oder weniger schnell vollziehenden Paradigmenwechsel.

Für den Übergang von einer Phase zur nächsten gibt Kuhn soziale Indikatoren an, die diesen Prozess häufig begleiten. Zum Beispiel, ist der Übergang von der vor-normalen (vor-paradigmatischen) zur normalen Phase (i) an der Auflösung der Schulverbände, dem Ausschluss der das Paradigma nicht akzeptierenden Wissenschaftler/Wissenschaftlerinnen erkennbar (diese werden etwa als ‚Philosophen‘ aus der Disziplin fortgelobt), (ii) an der Entwicklung von Fachzeitschriften und Fachvereinigungen, (iii) an der Aufnahme der Disziplin in die akademischen Lehrpläne, (iv) an der Änderung der Publikationsformen von der Monografie (die alles von Anfang an darstellt) zum Fachartikel (der den Stand der Forschung als bekannt und anerkannt unterstellt), wodurch die wissenschaftlichen Publikationen einen esoterischen Charakter bekommen (das ‚gebildete Publikum‘ wird von der Diskussion ausgeschlossen). Die Tätigkeit des normalen Wissenschaftlers gleicht nach Kuhn (im Gegensatz zum Ideal des kritischen Wissenschaftlers bei K. R. Popper) der eines routinierten Aufräumens und Verdeutlichens; der paradigmatische Bestand des Wissens wird nicht angetastet, Erfolg hängt von der Anerkennung des Paradigmas ab. So leistet der normale Wissenschaftler z. B. im empirischen Bereich Präzisionsarbeit bezüglich der vom Paradigma als aufschlussreich bestimmten Fakten, bezüglich der gemäß der Theorie prognostizierten Fakten, bezüglich der Größen (Konstanten, Gesetze) der paradigmatischen Theorie selbst; er weitet die Anwendungsbereiche der anerkannten Theorien aus und bemüht sich um mathematische Formulierung und Verbesserung bestehender Formulierungen der Theorie (puzzle solving).

Im Anschluss an Kuhns Kritik an Popper hat es in der Wissenschaftsphilosophie des 20. Jahrhunderts eine breite Diskussion über das Verhältnis methodologischer Regeln zu kontingenten sozialen Einflüssen mit zahlreichen Lösungsversuchen gegeben (Radnitzky und Andersson 1980). Vor allem Imre Lakatos hat einen einflussreichen Vermittlungsversuch zwischen Popper und Kuhn entwickelt (Lakatos 1970). Im Gegensatz zu Kuhns Diskontinuitätsthese ist der Prozess der Wissenschaftsentwicklung nach Lakatos durch eine Kontinuität sich ablösender Theorien gekennzeichnet. Das Kontinuum entwickelt sich aus einer Menge methodologischer Regeln (‚Forschungsprogramm‘), die teilweise Verbote (‚negative Heuristik‘), teilweise Gebote (‚positive Heuristik‘) sind. Forschungsprogramme sind durch ei-

nen methodologischen ‚harten Kern‘ zu charakterisieren, der durch Hilfshypothesen gegen Überprüfung geschützt werden muss.

Die Diskussion über die Rolle einer wissenschaftlichen Methodologie und die Regeln interner Wissenschaftskommunikation in Wechselwirkung mit externen Einflussfaktoren, wie sie durch gesellschaftliche Erwartungen gebildet werden, die sich wiederum zu Instrumenten der Forschungsförderung verdichten können, spielt offenkundig eine epistemische und nicht nur eine pragmatische Rolle bei der Rekonstruktion wissenschaftlicher Geltungsansprüche und ihrer Einlösung. Das Gegenstandsfeld der Wissenschaftsphilosophie konvergiert demzufolge offenkundig mit dem der Wissenschaftshistoriografie und der Wissenschaftssoziologie; dies macht Versuche plausibel, die Forschungsperspektiven dieser drei Disziplinen zu einem einheitlichen Konzept der Wissenschaftsforschung/Wissenschaftswissenschaft zu vereinen. Im Rahmen eines solchen Programms sind die Beziehungen zwischen sozialen und kognitiven Strukturen und Vorgängen zu problematisieren. Die schematischen Deutungen der Wissenschaftsentwicklung, wonach diese entweder durch die Geltung ahistorischer Standards nur als intern induzierter Prozeß oder als ein Reflex nur externer sozialer Faktoren verstanden wird, müssen nach P. Weingart (1972) durch eine „intervenierende Variable“ des organisatorischen Aufbaus der Wissenschaft, ihrer Strukturen, Regeln und prozessualen Mechanismen miteinander verknüpft werden. Durch diese findet einerseits eine Vermittlung von sozialen Bedingungen und Einflüssen in wissenschaftliche, d. h. kognitive, Prozesse statt, andererseits werden ihre Resultate über die spezifische Organisation der Wissenschaft in die Gesellschaft vermittelt und wirken auf diese. Der Bezug der sozialen und kognitiven Strukturen aufeinander im Rahmen einer einheitlichen Konzeption ergibt sich nur, wenn die Institution der Wissenschaft (als intervenierende Variable) auf ihre normativen Grundlagen hin untersucht und mit den normativen Ansprüchen, die Wissenschaftler/Wissenschaftlerinnen kognitiv als Gründe erfahren, verglichen wird (Gethmann 1981).

### ***3.6.5 Zusammenfassung: Die Unverzichtbarkeit menschlicher Erkenntnissubjekte***

Der Falsifikationismus ist eine interne methodologische Wende innerhalb des Bacon-Projekts. Schon Popper hatte zunächst versucht, das Falsifikationsverfahren durch eine kanonische Prozedur anzugeben, die es sogar erlauben sollte, ein numerisches Maß für den Grad der Bestätigung einer Hypothese zu ermitteln. Ein solches Verfahren ließe sich grundsätzlich durch einen Bestätigungs-Algorithmus auf der Basis von Big Data verfeinern. Die weitere wissenschaftsphilosophische Debatte über den Begriff der Bestätigung, insbesondere das Problem der Unbestimmtheit, hat jedoch gezeigt, dass Bestätigungsversuche von materiellen Hintergrundannahme, Risiko-Abwägungen und heuristischen Fähigkeiten abhängen, die sich nicht in ein kanonisches Verfahren überführen lassen, sondern von faktisch-kontingenten

sozialen Faktoren abhängen. Die Vorstellung, die Wahrheitsprüfung von wissenschaftlichen Hypothesen letztlich Computern zu übertragen, hat sich daher als unhaltbar erwiesen. Erfahrung, Einfallsreichtum, Fingerspitzengefühl, Übersicht und kommunikative Kompetenz sind konstitutive Fähigkeiten, die Falsifikationsverfahren anleiten.

### 3.7 Fazit

Die Vorstellung der vollständigen Substituierbarkeit menschlicher kognitiver Leistungen macht von anthropologischen Präsuppositionen Gebrauch, die einer kritischen Betrachtung nicht standhalten. Dabei ist nicht ausgeschlossen, dass kognitive Teilfunktionen wie beispielsweise Gedächtnisleistungen von technischen Artefakten nicht nur ersetzt, sondern sogar hinsichtlich geringerer Störanfälligkeit, höherer Leistungsfähigkeit, tieferer Vernetzung, u. a. übertroffen werden können. Technische Artefakte, die mit sog. künstlicher „Intelligenz“ ausgestattet sind, können jedoch grundsätzlich nicht als Akteure in wissenschaftlichen Begründungsdiskursen auftreten, weil ihnen die pragmatischen Merkmale der Handlungsurheberschaft und Zurechenbarkeit nicht zukommen. Hinsichtlich praktischer Aufgabenstellungen ist Big Data unter günstigen Bedingungen ein Hilfsmittel, aber kein geeignetes Instrument der definitiven epistemischen Leistungsbeurteilung und Entscheidung. Pragmatische und heuristische Faktoren wie Prüfung der Kohärenz mit anderen Evidenzquellen, Erfolgseinschätzungen u. a. spielen eine nicht zu vernachlässigende Rolle. Die wissenschaftsphilosophische Debatte über den Begriff der Bestätigung, insbesondere das Problem der Unbestimmtheit, hat gezeigt, dass Bestätigungsversuche von materiellen Hintergrundannahme, Risiko-Abwägungen und heuristischen Fähigkeiten abhängen, die sich nicht in ein kanonisches Verfahren überführen lassen, sondern von faktisch-kontingenten sozialen Faktoren abhängen. Erfahrung, Einfallsreichtum, Fingerspitzengefühl, Übersicht und kommunikative Kompetenz sind konstitutive Fähigkeiten, die Falsifikationsverfahren anleiten müssen, aber nicht durch sie ersetzt werden können.

## Literatur

- Albert H (1968) Traktat über kritische Vernunft, 1. Aufl. J.C.B. Mohr, Tübingen
- Anderson Ch (2008) The end of theory: the data deluge makes the scientific method obsolete. WIRED. [http://archive.wired.com/science/discoveries/magazine/16-07/pb\\_theory](http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory). Zugegriffen am 11.12.2020
- Aristoteles (1936) In: Ross WD (Hrsg) Aristotle's physics. A revised text. Clarendon, Oxford
- Carrier M, Mittelstraß J (1989) Geist – Gehirn – Verhalten. de Gruyter, Berlin
- Chisholm RM (1964) Human freedom and the self. The Lindley lecture. University of Kansas Press, Lawrence

- Dreyfus HL (1979) *What computers can't do. A critique of artificial reason.* Cambridge University Press, Cambridge
- Dreyfus HL (1992) *What computers still can't do: a critique of artificial reason.* New Edition. MIT Press, Cambridge.
- Dummett M (1978) Truth. In: Dummett M (Hrsg) *Truth and other enigmas.* Harvard University Press, Cambridge, S 1–24
- Dummett M (1993) What is a theory of meaning (II). In: Dummett M (Hrsg) *The seas of language.* Clarendon Press, Oxford, S 34–93
- Falkenburg B (2012) *Mythos Determinismus.* Springer, Heidelberg
- Forum für Philosophie (1992) Ed. *Realismus und Antirealismus.* Suhrkamp, Frankfurt am Main
- Fuchs T (2013) *Das Gehirn – ein Beziehungsorgan, 4E. Aufl.* Kohlhammer, Stuttgart
- Gethmann CF (1979) *Protologik. Untersuchungen zur formalen Pragmatik von Begründungsdiskursen.* Suhrkamp, Frankfurt am Main
- Gethmann CF (1980) Die Logik der Wissenschaftstheorie. In: Gethmann CF (Hrsg) *Theorie des wissenschaftlichen Argumentierens.* Suhrkamp, Frankfurt am Main, S 15–42
- Gethmann CF (1981) *Wissenschaftsforschung? Zur philosophischen Kritik der nachkuhnischen Reflexionswissenschaften.* In: Janich P (Hrsg) *Wissenschaftstheorie und Wissenschaftsforschung.* Beck, München, S 9–38
- Gethmann CF (2003) Wissen als Macht. *Wissenschaftsphilosophische Überlegungen.* In: Emmermann R et al (Hrsg) *An den Fronten der Forschung: Kosmos – Erde – Leben.* Hirzel, Stuttgart, S 238–245
- Gethmann CF (2010) Menschsein – Menschbleiben. Zur Grammatik askriptiver Äußerungsmodi. In: Rösen J (Hrsg) *Perspektiven der Humanität. Menschsein im Diskurs der Disziplinen.* Transcript, Bielefeld, S 41–58
- Gethmann CF (2014) Wahrheit und Beweisbarkeit. Die philosophische Bedeutung von Intuitionismus und Konstruktivismus in der Logik. In: Mittelstraß J, von Bülow C (Hrsg) *Dialogische Logik. mentis, Münster, S 13–28*
- Gethmann CF (2016) What remains of the fundamentum inconcussum in light of the modern sciences of humans? *J Gen Philos Sci* 47:1–20
- Gethmann CF, Sander T (2002) Anti-Mentalismus. In: Gutmann M et al (Hrsg) *Kultur – Handlung – Wissenschaft.* Velbrück, Weilerswist, S 91–108
- Hempel CG (1946) A note on paradoxes of confirmation. *Mind* 55:79–82
- Hengstenberg HE (1957) *Philosophische Anthropologie.* Kohlhammer, Stuttgart
- Hilgendorf E (2015) Recht und autonome Maschinen – ein Problemaufriß. In: Hilgendorf E, Hötitzsch S (Hrsg) *Das Recht vor den Herausforderungen der modernen Technik (4).* Nomos, Baden-Baden, S 11–40
- Hunt E, Lunneborg C, Lewis J (1975) What does it mean to be high verbal? *Cogn Psychol* 7:194–227
- Janich P (1997) *Das Maß der Dinge. Protophysik von Raum, Zeit und Materia.* Suhrkamp, Frankfurt am Main
- Janich P (2015) *Handwerk und Mundwerk. Über das Herstellen von Wissen.* Beck, München
- Kamlah W, Lorenzen P (1973) *Logische Propädeutik. Vorschule des vernünftigen Redens.* Bibliographisches Institut, Mannheim
- Kant I (1956) *Kritik der reinen Vernunft.* Meiner, Hamburg
- Kuhn TS (1970) *The structure of scientific revolutions, 2. Aufl.* University of Chicago Press, Chicago
- Kuhn TS (2000) The road since „structure“. In: Conant J, Haugeland J (Hrsg) *Philosophical essays.* University of Chicago Press, Chicago
- Lakatos I (1970) Falsification and the methodology of scientific research programs. In: Lakatos I, Musgrave A (Hrsg) *Criticism and the growth of knowledge.* Cambridge University Press, Cambridge, S 91–196
- Lenzen W (1980) *Glauben, Wissen und Wahrscheinlichkeit. Systeme der epistemischen Logik.* Springer, Wien
- Lorenz K (1976) Sprachtheorie als Teil einer Handlungstheorie. In: Wunderlich D (Hrsg) *Wissenschaftstheorie der Linguistik.* Athenäum, Kronberg, S 250–280

- Mainzer K (1995) Computer – Neue Flügel des Geistes? de Gruyter, Berlin
- Mead GH (1934) In: Morris CW (Hrsg) Mind, self and society from the standpoint of a social behaviorist. Chicago University Press, Chicago
- Nida-Rümelin J, Weidenfeld N (2018) Digitaler Humanismus: Eine Ethik für das Zeitalter der Künstlichen Intelligenz. Piper, München
- Plessner H (1928) Die Stufen des Organischen und der Mensch. de Gruyter, Berlin. (Wiederabdr. Gesammelte Schriften IV (2003). Suhrkamp, Frankfurt am Main)
- Plessner H (1941) Lachen und Weinen. Eine Untersuchung nach den Grenzen menschlichen Verhaltens. Van Loghum Slaterus, Arnheim. (Wiederabdr. Gesammelte Schriften VII (2003). Suhrkamp, Frankfurt am Main)
- Popper KR (1934) Logik der Forschung. Julius Springer, Wien
- Popper KR (1962) The open society and its enemies. Routledge, London
- Radnitzky G, Andersson G (Hrsg) (1980) Fortschritt und Rationalität der Wissenschaft. J.C.B. Mohr (Paul Siebeck), Tübingen
- Schäfer L (1993) Das Bacon-Projekt. Von der Erkenntnis, Nutzung und Schonung der Natur. Suhrkamp, Frankfurt am Main
- Scheler M (1916) Der Formalismus in der Ethik und die materiale Wertethik. Niemeyer, Halle an der Saale
- Scheler M (1923) Wesen und Formen der Sympathie, 2. Aufl. F. Cohen, Bonn
- Scherer G (1976) Strukturen des Menschen. Ludgerus, Essen
- Searle J (1984) Minds, brains and science. Cambridge University Press, Cambridge
- Siegmund G (1997) Vorfragen zur Wahrheit. Ein Traktat über kognitive Sprachen. R. Oldenbourg, München
- Spinner HF (1974) Pluralismus als Erkenntnismodell. Suhrkamp, Frankfurt
- Tennant N (1997) The taming of the true. Clarendon, Oxford
- Turing AM (1950) Computing machinery and intelligence. Mind NS 59:433–460
- Van Fraassen BC (1980) The scientific image. Clarendon, Oxford
- Von Wright GH (1971) Erklären und Verstehen. Fischer-Athenäum, Frankfurt am Main
- Weingart P (1972) Einführung. Wissenschaftsforschung und wissenschaftssoziologische Analyse. In: Weingart P (Hrsg) Wissenschaftssoziologie I. Fischer Athenäum, Frankfurt am Main, S 11–42
- Wright C (1992) Truth and Objectivity. Harvard University Press, Cambridge

**Open Access** Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

