

Title:

Die Bedeutung des abduktiven Schließens in Erkenntnis- und Wissenschaftstheorie

Author:

Gerhard Schurz

IPS-PREPRINTS

Annual 1995 No. 2

Edited by Gerhard Schurz und Alexander Hieke

Vorveröffentlichungsreihe am **I**nstitut für **P**hilosophie der Universität **S**alzburg

Prepublication Series at the Department of Philosophy, University of Salzburg

Die Bedeutung des abduktiven Schließens in Erkenntnis- und Wissenschaftstheorie

Gerhard Schurz, Salzburg

1. Nichtkreative versus kreative Abduktion

Die empiristische Erkenntnistheorie versucht, rationales Schließen auf Deduktion und Induktion zu reduzieren.¹ Ausgehend vom unmittelbar Beobachtbaren soll allein durch Induktion und Deduktion das ganze Spektrum alltäglicher und wissenschaftlicher Erkenntnis rational begründet werden. Das Hauptproblem dabei ist, daß unsere Erkenntnis, gerade auch die naturwissenschaftliche, in hohem Maße Begriffe bzw. Konzepte enthält, die über das unmittelbar Beobachtbare hinausgehen, z.B. "Kraft", "Masse", "chemische Substanzart", usw. Weder durch Deduktion noch durch Induktion kann man aber auf Sätze schließen, die *neue Begriffe* in essentieller Weise enthalten (was weiter unten präzisiert wird).² Daher auch immer wieder der - letztlich zum Scheitern verurteilte - Versuch der Empiristen, alle wissenschaftlichen Begriffe definatorisch oder analytisch auf Beobachtungsbegriffe zurückzuführen.³

Analog geartet ist das erkenntnistheoretische *Realismusproblem*. Tatsächlich

¹ Je nachdem, wie eng oder weit man den Begriff "Empirist" faßt, waren es "alle" oder aber nur "alle radikalen" oder "naiven" (im Gegensatz zu den "gemäßigten" oder "kritischen") Empiristen. S. auch Fußnote 3.

² Die vorliegende Expositon der Abduktion beruht auf einer zumindest graduellen Unterscheidung zwischen Beobachtungsbegriffen und theoretischen Begriffen - eine Unterscheidung, die in der jüngeren Wissenschaftstheorie auch häufig angegriffen worden ist. Zur Neukonzeption und Verteidigung dieser Unterscheidung s. Schurz (1988, Kap. 3). Aber selbst, wenn man nicht von einer solchen Unterscheidung ausgehen möchte, kann man die Frage der ('kreativen') Abduktion als die Frage nach einer Schlußart betrachten, welche in die Konklusion neue generelle Terme (in essentieller Weise) einführt.

³ John Locke zufolge sind alle Ideen Komplexe einfacher, letztlich durch Beobachtung übermittelter Ideen (1690, 2. Buch). David Hume teilte (im wesentlichen) diese Ansicht. Sowohl der frühe Peirce wie der frühe Bridgeman und der frühe Carnap versuchten, theoretische Begriffe der Naturwissenschaft definatorisch auf Beobachtungsbegriffe zurückzuführen. Alle drei haben später ihre Auffassung revidiert - vgl. z.B. Peirce (1878a, Bd 5, § 403) versus (1905, Bd 5, § 435). Ein Meilenstein für die Einsicht in die Unmöglichkeit, theoretische Begriffe definatorisch auf Beobachtungsbegriffe zurückzuführen, ist Carnap (1936/37). Carnap will die Konsequenzen seiner Erkenntnisse hier noch nicht so recht zugestehen; voll zum Tragen kommen sie dann in Carnap (1956) und Hempel (1950).

fängt unsere Erkenntnis nicht mit Beobachtungssätzen im Sinne von Beobachtungsrealsätzen an, sondern mit Erscheinungssätzen bzw. *introspektiven Sätzen* – wie ich sie hier, ob ihrer essentiellen *Ich-Bezogenheit* wegen, nennen möchte. Der Schluß von "ich habe eine Maus-Erscheinung" auf "dort ist eine Maus" – der *Realschluß*, wie ich ihn im folgenden kurz nennen werde – führt ebenfalls neue Begriffe ein, sogar ein ganzes System von neuen Begriffen, das der *Realbegriffe*. Es scheint ganz unverständlich, wie man durch Deduktion und Induktion allein vom Scheinen zum Sein gelangen könnte. Viele Empiristen haben den Realismus daher als keiner rationalen Begründung fähig kurzerhand überbortet.

Verschiedene Erkenntnis- und Wissenschaftstheoretiker haben aus diesen Gründen nach Schluß- bzw. Argumentationsschemata gesucht, die über Deduktion und Induktion hinausgehen. Der meine Erachtens vielversprechenste diesbezügliche Ansatz ist Peirce' Schlußform der *Abduktion*. Peirce gibt aber zwei verschiedene Explikationen dieser Schlußform, eine *formelle* und eine *informelle*. Die formelle Explikation stellt die Abduktion der Deduktion und der Induktion – genauer, der *generalisierenden* Induktion⁴ – wie folgt gegenüber (s. Peirce 1868, Bd 5, § 270-274 und 1878b, Bd 2, § 619-625):

Beispiel: Alle Bären (F) halten Winterschlaf (G)

<i>Deduktion:</i>	<i>Induktion:</i>	<i>Abduktion:</i>
Gesetz: Alle F sind G	Fall: Dies und dies ... ist ein F	Gesetz: Alle F sind G
Fall: Dies ist ein F	Resultat: Dies und dies .. ist ein G	Resultat: Dies ist ein G
Resultat: Dies ist ein G	Gesetz: Alle F sind G	Fall: Dies ist ein F

Abduktion ist hier also der Schluß entgegen die Pfeilrichtung vom instanziierten Ge-

⁴ Innerhalb der Theorie induktiven Schließens kennt man mehrere Arten induktiver Schlüsse; neben dem induktiven Generalisierungsschluß (vom Sample auf die Gesamtheit) insbesondere noch den induktiven Voraussageschluß (vom Sample auf einen neuen Fall) und den 'inversen' induktiven Spezialisierungsschluß (von der Gesamtheit auf ein Sample oder ein Individuum). S. z.B. Carnap (1959), S. 81. Wir sprechen im folgenden auch von 'generalisierender', 'prognostizierender' versus 'spezialisierender' Induktion.

setzeskonsequens auf das Gesetzesantecedens, in kausaler Terminologie von der Wirkung auf die Ursache, bei *bekanntem* Gesetz. Man hat diese Schlußart auf Retrodiktion genannt.⁵ Peirce ist sich klar, daß diese Schlußschemata hinsichtlich ihrer Validität bzw. Sicherheit sehr verschieden zu beurteilen sind; er sagt: Deduktion verleiht der Konklusion Sicherheit, Induktion Wahrscheinlichkeit, Abduktion als schwächste Schlußart bloße Möglichkeit. Zunächst sei Abduktion Spekulation; die abduzierte Hypothese muß unter Anwendung von Deduktion und Induktion empirisch getestet werden (Peirce 1903, Bd 5, §§ 145, 171).

Informell charakterisiert Peirce Abduktion als das einzige logische Verfahren, daß eine neue Idee einführen kann (Peirce 1878b). "Jedes einzelne Stück wissenschaftlicher Theorie", so sagt er, "ist der Abduktion zu verdanken" (Peirce 1903, Bd 5, § 172). Die Abduktion ist es, mit der die Wissenschaftler theoretische Modelle entwickeln und neue theoretische Begriffe einführen. Ein einfaches und von Peirce immer wieder diskutiertes Beispiel ist das Ritzverhalten des Diamanten, der alle anderen Gegenstände ritzt, woraus wir abduktiv auf seine innere Eigenschaft oder Disposition der (maximalen) Härte schließen.⁶ Abduktion ist hier also der Schluß von beobachteten Regelmäßigkeiten auf eine nichtbeobachtbare zugrundeliegende innere Ursache. Abduktion ist nach Peirce auch jener Schluß, womit wir von unseren Sinneseindrücken auf eine objektive Realität (Realität "als Drittheit") schließen (1903, Bd 5, § 206-212).

Was den beiden Peirceschen Charakterisierungen gemeinsam ist: in beiden Fällen handelt es sich um den Schluß von der 'Wirkung' auf die 'Ursache'. Insofern handelt es sich hier um einen *Nachfahren* des traditionellen philosophischen *Prinzips*

⁵ Chisholm (1979, S. 101) spricht von "inverser Induktion".

⁶ Wohlgermerkt lehnt der spätere Peirce (im Gegensatz zum früheren) dezidiert ab, die "Härte" mit dem empirischen Ritzverhalten semantisch zu identifizieren, denn auch der Diamant, der sein Leben lang nie geritzt wurde, ist objektiv hart (Peirce 1905, Bd 5, § 453). Peirce hilft sich hier mit dem *kontrafaktischen Konditional* - die Härte des Diamanten besteht darin, daß er einem Ritzversuch widerstehen *würde*, wenn er geritzt werden *würde*; aber dies ist nur eine andere Umschreibung dafür, daß hier eine über die Beobachtung hinausgehende innere Disposition des Diamanten abduziert wird (s. Schurz 1991a, Kap. 4.1). Peirce ist hier sehr weitsichtig, insofern die Debatte um Dispositionsprädikate mit genau denselben Einsichten viele Jahrzehnte später wieder aufgenommen wurde (s. Tuomela 1978, ed.).

vom zureichenden Grunde.⁷ Peirce versteht das Abduktionsschema jedoch weder im Sinne des *Rationalismus* als Denknötwendigkeit, noch im Sinne der Transzendentalphilosophie als Bedingung der Möglichkeit von Erfahrung. Abduktion ist bei Peirce vielmehr als *heuristisches* Schema zu verstehen: grob gesprochen sollen wir nach zureichenden Gründen zu suchen; Sicherheit, daß wir sie immer finden, gibt es keine, aber nur so kommen wir weiter.⁸

Zwischen den beiden Peirceschen Charakterisierungen von Abduktion besteht aber eine offensichtliche Divergenz, ja ein Widerspruch. In der ersten, formellen Charakterisierung von Abduktion wird nämlich *kein* neuer Begriff in die Konklusion eingeführt, und kann es gar nicht werden, da der Antecedensbegriff auch im Gesetz enthalten sein muß, und dieses als *bekannt*, d.h. als Prämisse, angenommen wird. Umgekehrt, soll eine Abduktion wirklich einen *neuen* Begriff einführen, so müßte zugleich mit dem Antecedens bzw. dem Fall auch das Gesetz abduktiv erschlossen werden; die einzige Prämisse wäre dann das Konsequens bzw. Resultat. Ebendies ist beim Peirceschen Beispiel des Schlusses vom Ritzverhalten eines Gegenstandes auf seine Härte auch der Fall: vom regelmäßigen Ritzverhalten des Diamanten schließen wir 1. darauf, daß er die innere Eigenschaft der Härte besitzt und 2. daß jeder Gegenstand, der hart ist, dieses regelmäßige Ritzverhalten aufweist. Um den Unterschied terminologisch zum Ausdruck zu bringen, nenne ich Abduktion im ersten Sinn *nichtkreative* Abduktion, Abduktion im zweiten Sinn *kreative* Abduktion.

Nichtkreative Abduktion ist schon deshalb, weil sie keine neuen Begriffe einführt, weniger interessant und fundamental. Aber auch deshalb, weil man bei ihr plausibel argumentieren kann, daß sie sich auf eine Form des Wahrscheinlichkeitsschließens (also des im weiten Sinne induktiven Schließens) zurückführen läßt. Fumerton (1980) hat dies mit folgendem Beispiel zu zeigen versucht. An einem einsamen

⁷ Leibniz hat dieses Prinzip axiomatisch ausformuliert; es ist jedoch wesentlich älter, ohne daß ich einen Urheber nennen könnte – jedenfalls haben es sowohl Descartes, Berkeley und Locke als selbstverständliche Grundlage ihrer Argumentation benutzt, und Hume war der erste, der dieses Prinzip (ebenso wie das Induktionsprinzip) grundsätzlich in Frage stellte.

⁸ S. Peirce 1892, Bd 6, § 40f; 1902/3, Bd 2, § 113. Eine Einführung in Peirce' Gesamtsystem gibt Schurz (1991a).

Strand sehen wir menschliche Fußspuren. Aufgrund des bekannten Gesetzes "Wenn jemand am Strand geht, hinterläßt er Fußspuren" schließen wir darauf, daß hier jemand am Strand entlang gegangen ist. Ebensogut, so Fumerton, könnten wir aber auch darauf schließen, daß dort Jimmy Carter oder sagen wir Michael Jackson entlang gegangen ist; für die formale Korrektheit der Abduktion macht dies keine Unterschied. Warum würden wir aber diese zweite Abduktion als unsinnig ablehnen? Wohl darum, weil die Wahrscheinlichkeit, daß es gerade Jimmy Carter oder Michael Jackson und kein anderer war, der hier vorbei ging ist, sehr gering ist, während die Wahrscheinlichkeit, daß hier irgendein Mensch hier vorbei ging, und nicht etwa ein mit Menschenschuhen ausgestattetes zweibeiniges Tier, äußerst groß ist. Eine implizite Bedingung, um nichtkreative Abduktionen zu akzeptieren, ist daher, so Fumerton, daß die umgekehrte Wahrscheinlichkeit des Antecedens, gegeben das Konsequenz, hinreichend hoch ist. Wenn das aber so ist, so kann man jede nicht-kreative Abduktion durch einen gewöhnlichen induktiv-probabilistischen Spezialisierungsschluß ersetzen. Aus dem Schluß von " $\forall x(Fx \rightarrow Gx)$, Ga " auf " Fa " unter der impliziten Annahme " $p(F/G) = \text{hoch}$ " wird der hoch-wahrscheinliche Schluß von " $p(F/G) = \text{hoch}$, Ga " auf " Fa ".

Bei der kreativen Abduktion kann man einen solchen Reduktionsversuch sicher nicht anwenden. Weil hier ein *neuer* Begriff in die Konklusion eingeführt wird, kann es auch kein bekanntes Wahrscheinlichkeitsgesetz über ihn geben, das als implizite Prämisse benutzt wird. Kreative Abduktion ist also sicher irreduzibel - ein zweiter Grund, diese Form der Abduktion als erkenntnistheoretisch fundamental anzusehen.

Kreative Abduktion, generalisierende Induktion und Deduktion lassen sich in folgender Weise schön kontrastieren, nämlich im Hinblick auf die Frage, welche neuen Terme diese Schlußformen in die Konklusion in essentieller Weise einzuführen vermögen. *Die Deduktion vermag gar keine neuen (nichtlogischen) Terme essentiell einzuführen. Die generalisierende Induktion (wie auch die prognostizierende Induktion, s. Fn. 4) vermag neue singuläre Terme in die Konklusion essentiell einzuführen. Die kreative Abduktion schließlich- und nur sie - vermag neue generelle Terme (i.e. Begriffe) essentiell einzuführen.* Die logische Präzisie-

rung des Begriffs der "essentiellen Einführung" muß vorallem zwei Schwierigkeiten meistern. *Erstens* lassen sich auch durch deduktive Allspezialisierung $\forall xFx \vdash Fa$ *scheinbar* neue Individuenkonstanten gewinnen; aber eben nur scheinbar, weil der Allquantor in der Prämisse implizit über *alle* Individuen und daher auch über *a* spricht. Dem läßt sich am einfachsten dadurch Rechnung tragen, indem man annimmt, alle Individuen des zugrundeliegenden Bereichs hätten Namen und der Allquantor sei durch eine infinite Konjunktion ersetzt. *Zweitens* vermag die Deduktion neue generelle oder singuläre Terme in *irrelevanten* Weise in die Konklusion einzuführen; etwa $Fa \vdash Fa \vee Gb$. Für Prädikate und für die Deduktion läßt sich dem durch die anderswo entwickelte Theorie der relevanten Deduktion (Schurz 1991b) gut Rechnung tragen, welche die Irrelevanz eines Prädikats in der Konklusion mit seiner Ersetzbarkeit *salva validitate* identifiziert. Allerdings ist diese Theorie auf Induktion über Individuen(konstanten) nicht anwendbar.

Eine Lösungsmöglichkeit würde hier das Kriterium der *Eliminierbarkeit ohne Gehaltsverlust* liefern; es ist für Deduktionen mit dem Ersetzungskriterium äquivalent (Schurz 1996a), und hat den Vorteil, auch auf Induktion und Individuenkonstanten besser anwendbar zu sein. Dies hier auszuführen würde den Rahmen dieses Aufsatzes allerdings sprengen⁹, weshalb ich mich damit begnüge, die obigen Thesen für *atomare Konklusionen aus konsistenten Prämissen* zu begründen. Solche Konklusionen folgen nämlich, falls überhaupt, immer auch relevant. Angenommen, Quantoren seien durch infinite Konjunktionen bzw. Diskunktionen eliminiert, und auch die Identität sei auf übliche Weise eliminiert, so gilt aufgrund elementarer logischer Theoreme, daß wann immer $\text{Prem} \vdash Fa$ und der Atomsatz *Fa* kommt in *Prem* nicht vor (d.h. entweder *a* oder *F* oder beides kommt in *Prem* nicht vor), *Prem* inkonsistent

⁹ Induktives Schließen besteht eben gerade darin, daß auf *beliebige andere* Individuen verallgemeinert wird, ohne daß die damit gewonnenen Prognosen deshalb irrelevant wären. Der Unterschied zur Deduktion besteht genau darin, daß die Ersetzbarkeit von Individuenkonstanten im Falle induktiver Prognosen *nicht* eine Eliminierbarkeit ohne Gehaltsverlust impliziert, während dies im Fall irrelevanter Deduktionen immer so ist, denn irrelevante Deduktion beruht immer auf irrelevanter Abschwächung der Konklusion. Mit Eliminierbarkeit *salva validitate* ist dabei gemeint, daß ein- oder mehrere Instanzen einer Atomformel in der Konklusion entweder durch *Verum* oder durch *Falsum* ersetzbar (und somit über die üblichen logischen Operationen eliminierbar) sind, und der logische Gehalt der Konklusion dabei gleich bleibt oder größer wird.

sein muß. Der induktive Voraussageschluß $Fa_1, \dots, Fa_n \Rightarrow Fa_{n+1}$ vermag uns atomare Konklusionen mit neuen Individuenkonstanten (aber niemals mit neuen Prädikaten) zu liefern. Kreative Abduktion schließlich – deren logische Form wir noch gar nicht bestimmt haben – liefert uns atomare Konklusionen mit neuen Prädikaten, z.B. "a hat die Disposition der *Härte*". Die Forderung, der eingeführte Dispositionsterm muß *neu* sein, bedeutet, daß dieser Term im bisherigen epistemischen Hintergrundsystem noch keine (auch nicht partielle) semantische Interpretation besitzen darf. Diese Forderung bedeutet wohlgermerkt eben *nicht*, daß der neue Term *beliebig* sein darf und daher durch beliebiges ersetzbar ist, denn sobald man den neuen Term durch einen *alten* ersetzen würde, wäre die Abduktion inkorrekt – beispielsweise wäre es unsinnig, abduktiv aus dem Ritzverhalten eines Gegenstandes zu schließen, er hätte die Disposition der Wasserlöslichkeit, und jeder wasserlösliche Gegenstand zeigt das typische Ritzverhalten.

2. Probleme der kreativen Abduktion

Ich wende mich nunmehr der kreativen Abduktion zu. In der Terminologie der Erklärung gesprochen handelt es sich dabei um den Schluß vom *Explanandum* auf das gesamte *Explanans*, also dem Antecedens inklusive dem Gesetz. Dies entspricht weitgehend dem, was im anglosächsischen Sprachraum – leider zumeist ohne Bezugnahme auf Peirce – unter dem Titel *inference to the best explanation* diskutiert wurde, initialgezündet durch Harman's ebenso betitelten Aufsatz (1965). Sobald man dies als Schlußform präzisieren will, kommt man in die größten Schwierigkeiten. So wie wir es bisher diskutierten, müßte das Schlußschema der kreativen Abduktion eigentlich folgendes sein:

Ga _____ (daher:)	Beobachtetes Explanandum
Fa , und	Erschlossenes Explanans; F ist der neue Begriff
$\forall x(Fx \rightarrow Gx)$	

So ist dieses Schema freilich ganz untauglich. Ich nenne es das *spekulative Abduktionsschema*, denn man kann damit Märchen und Spekulationen jeglicher Art rechtfertigen. Wann immer man etwas beobachtet, erfinde man gemäß diesem Schema eine Geschichte, die das Beobachtete erklärt. Ein Waldureinwohner etwa sieht das erste Mal in seinem Leben einen Berg und schließt abduktiv, dies sei deshalb so, weil einst ein großer Waldriese sich dort eine Sitzgelegenheit baute. Oder ein Selig-Gottvertrauender schließt aus allem, was er erlebt, abduktiv, daß Gott es so wollte, und daß was immer Gott will, auch geschieht. Hinter dem spekulativen Abduktionsschema steckt der angeborenen Deutungs- und Erklärungswille der Menschen, sein "metaphysisches Bedürfnis". Die animistischen und spiritistischen Weltbilder der Frühkulturen benutzen durchgängig dieses Schema. Um ein wissenschaftliches bzw. ein aufgrund des Motivs der Wahrheitssuche allein rational gerechtfertigtes Schlußschema handelt es sich dabei freilich nicht.

Ein Weg, diese Schwierigkeit loszuwerden, ist von Harman seit Anbeginn beschritten worden -- indem man nämlich nur die *beste* Erklärung unter allen möglichen Erklärungen zuläßt. Auch Peirce hat dies, wie oben erläutert, bereits ähnlich gesehen.¹⁰ Die Schwierigkeit ist, die man lösen muß, ist, *erstens* zu präzisieren, was überhaupt als eine *Erklärung* zählt¹¹, und *zweitens* worin die Kriterien der *Güte* einer Erklärung bestehen. Im Falle des spekulativen Abduktionsschemas hilft die Forderung nach bester Erklärung jedenfalls *nichts*. Inwiefern sollte beispielsweise die Erklärung des Blühens einer Blume durch ihre innere Seele besser oder schlechter sein als die Erklärung durch den Willen Gottes? Wie wir weiter unten ausführen werden, sind beide Erklärungen, und alle Erklärungen dieser Art, gleich 'leer' bzw. unfundiert. In besonders auffallender Weise setzt sich die Erkenntnistheorie Mosers (1989) diesem Problem aus. Moser benutzt das inference-to-the-explanation Schema in äußerst unvorsichtiger Weise. Z.B. meint er, der Schluß vom Scheinen aufs Sein, in der simplen Form

¹⁰ Denn das abduktiv erschlossene Explanans ist prima facie bloß 'möglich' und muß sich durch Induktion und Deduktion erst bewähren.

¹¹ Eine Schwierigkeit, der sich viele inference-to-the-best-explanation Theoretiker erst gar nicht unterziehen; so schon Harman (1965) ebensowenig wie Thagard (1989, 1992) oder Moser (1989, S. 91ff).

(Explanandum:)	(Explanans:)
Mir erscheint, daß dort ein Baum ist	Daher ist dort ein Baum; und: normalerweise, wann immer dort ein F ist, scheint mir dort ein F zu sein

sei im Sinn der besten Erklärung zu rechtfertigen (Moser 1989, S. 107ff). Aber *warum* sollte obiges Explanans die beste Erklärung für meine Baum-Erscheinung sein? Z.B. könnte ich meine F-Erscheinung auch damit erklären, daß ich gerade einen Baum halluziniere. Moser wendet dagegen ein, daß diese Halluzinationserklärung eine unbeobachtbare Entität (mein Halluzinieren) einführe (1989, S.162f), aber auch die Realerklärung bzw. jede kreative Abduktion führt eine neue Entität ein, weshalb Mosers sich durch das ganze Buch ziehende Versuch, schlechte von guten Erklärungen dadurch abzugrenzen, daß erstere unbeobachtbare Entitäten ('gratuitous entities') einführen, untauglich ist. Oder aber, wir könnten das Wahrnehmungsbild des Baumes damit erklären, daß sich dort statt eines Baumes ein Cartesischer Dämon oder aber auch bloß das *Hologramm* eines Baumes befindet; auch für diese Fälle werden Mosers Entgegnungsversuche (S. 160-65) den Schwierigkeiten kaum gerecht.

Aus diesem Grunde haben andere Autoren, allen voran van Fraassen (1980, S. 19-40; 1989, ch. 6), das inference-to-the-best-explanation Schema radikal kritisiert. Warum *sollten* wir es überhaupt unsere empirischen Erscheinungen erklären – was bringen uns über die empirische Erscheinungsebene hinausgehende Erklärungen *überhaupt*, so würde van Fraassen zuallererst fragen. Andere Autoren, wie etwa Chisholm (1979) oder Pollock (1986), lehnen zwar den Realschluß nicht ab, wohl aber den Versuch, ihn durch Abduktion oder auch nur irgendwie zu begründen – es handle sich beim Realschluß einfach um eine grundlegende, in sich gerechtfertigte Schlußart.¹² Diese Lösung ist deshalb unbefriedigend, weil die bloße Einführung einer Schlußart als grundlegend noch keineswegs erklärt, *warum* die Anwendung

¹² Chisholm nennt es das *Prinzip der Evidenz* (1979, S. 115); Pollock (1986, S. 44) spricht von nichtinduktiven prima facie Gründen.

dieser Schlußart Erkenntnisse liefert, die *reliabel* sind, die m.a.W. hohe Wahrheitschancen haben oder zumindest einen anderen aufweisbaren Erkenntnisgewinn liefern. In diesem Sinn hat etwa Goldman (1986, S. 106) argumentiert, reliable kognitive Schlußprozeduren müßten die Tendenz besitzen, auf lange Sicht zu überwiegend wahren Erkenntnissen zu führen; und analog (nur in 'internalistischer' Variante¹³) hat Alston (1989, S. 102) argumentiert.

Um solchen und ähnlichen Schwierigkeiten zu entkommen hat man insbesondere zwei weitere und durchaus verwandte Kriterien für die Güte einer abduktiv erschlossenen Erklärung entwickelt: die Forderung der *Kohärenz*, und die Forderung der *Vereinheitlichung*.

Gemäß der ersten Idee ist eine Erklärung umso besser, je mehr sie die Kohärenz unseres Wissenssystems erhöht. Ein Wissenssystem ist dabei, grob gesprochen, umso kohärenter, je mehr seine Elemente (Sätze bzw. von solchen beschriebene Phänomene) durch Erklärungsrelationen untereinander verbunden sind. In jüngerer Zeit haben vorallem Lehrer (1974) und Thagard (1988, 1992) ihre Erkenntnis- bzw. Wissenschaftstheorie auf Kohärenz aufgebaut. Allerdings führt die Kohärenzidee in zwei charakteristische Schwierigkeiten. Denn warum sollte die Tatsache, daß ein Glaubenssystem intern kohärent ist, seine Wahrheitschancen erhöhen? Auch Märchen oder Spekulationen können intern sehr kohärent sein. Was der Kohärenzidee zu einem der empirischen Wissenschaft gerecht werdenden Evaluationskriterium fehlt, ist (vor allem) folgendes.

Erstens müssen zirkuläre Stützungen ausgeschlossen sein. Wenn ich z.B. die Tatsache, daß es stürmt, damit erkläre, daß Zeus zornig ist, umgekehrt daß Zeus zornig ist, damit erkläre, daß es stürmt, so habe ich zwar Kohärenz geschaffen, aber zirkuläre Kohärenz. Wenn man von der Kohärenz alle Zirkularitäten abzieht, gelangt man zum Begriff der *Vereinheitlichung*. In Schurz/Lambert (1994, S. 72) wurde dies so formuliert: *unification = coherence minus circularity*. Eine abduktiv erschlossene Erklärung ist also umso besser, je mehr sie unser Wissenssystem *vereinheitlicht*. Dabei ist ein Wissenssystem, grob gesprochen, umso mehr vereinheitlicht, je mehr

¹³ D.h. man müsse zumindest gute Gründe für die Annahme der Wahrheitstendenz besitzen.

elementare Tatsachen darin auf je weniger elementare Prinzipien plus elementare Tatsachen *argumentativ* zurückführbar sind. Die Idee der Vereinheitlichung hat wissenschaftstheoretische Tradition; sie ist bereits in Machs Ökonomieprinzip (Mach 1883, S. 465) oder in Whewells Begriff der consilience (Whewell 1847) enthalten, wurde später von Feigl (1970) und in jüngerer Zeit von Friedman (1974), Kitcher (1981) und Schurz/Lambert (1994) verteidigt und ausgebaut.

Auch gegen das Vereinheitlichungskriterium gibt es ein einschlägiges Argument, das kürzlich und sehr stringent von Humphreys (1993) ausformuliert wurde: warum sollte von mehreren konkurrierenden Theoriensystemen dasjenige, welches einheitlicher ist, auch eher wahr sein? Vielleicht ist die Welt eben *nicht* einheitlich. Um diesem Einwand zu entgehen, muß die Vereinheitlichung primär auf die *empirischen Tatsachen* bezogen werden, derart, daß von Vereinheitlichung eines Wissenssystems nur dann gesprochen werden kann, wenn am Ende der (nichtzirkulären) argumentativen Schlußkette die *Beobachtungssätze* vereinheitlicht werden. Dies hat zur Folge, daß eine größere Vereinheitlichung automatisch immer, sozusagen als Beiprodukt, eine größere oder zumindest nicht geringere empirische Bewährung impliziert. Eine derartige Vereinheitlichungstheorie haben Karel Lambert und ich entwickelt (Schurz/Lambert 1994; s. auch Schurz 1996b), auf deren Details hier nicht eingegangen werden kann.¹⁴

Kreative Abduktion, so lautet mein erster und grundlegender Vorschlag, ist unter die Randbedingung der *Vereinheitlichungsfunktion* zu stellen. Daraus ergibt sich sofort eine wichtige Konsequenz. Bezieht man kreative Abduktion auf *singuläre* Tatsachen, so fällt sie mit dem obigem spekulativen Abduktionsschema zusammen. Man kann jedoch zeigen, daß das spekulative Abduktionsschema niemals Vereinheitlichung leisten kann (Schurz/Lambert 1994, S. 82ff). Kreative Abduktion kann

¹⁴ Die Schwierigkeit des Vereinheitlichungskriteriums liegt weniger in der Grundidee als darin, eine funktionierende Theorie von Vereinheitlichung zu liefern. Die in Schurz/Lambert (1994) entwickelte Theorie enthält folgende Hauptbestandteile: (1) eine Theorie der *Wissensrepräsentation* in Form minimaler relevanter Wissens Elemente, (2) eine Klassifikation von *Argumenten* "im weiten Sinne", welche über traditionelle deduktive und induktive Argumentationsschemata weit hinausgeht, (3) eine Theorie der *Wissensrevision* und Wissensdynamik, und vorallem (4) ein *komparatives epistemisches Kosten-Nutzen-Konzept*, welches dem verwendeten Vereinheitlichungsbegriff zugrunde liegt. Die Theorie wird auf mehrere *Fallbeispiele* angewandt.

nur dann Vereinheitlichung erbringen, und macht daher nur dann wissenschaftlichen Sinn, wenn sie sich auf *empirische Regelmäßigkeiten* bezieht – auf empirische Gesetze, die durch übergeordnete Theorien erklärt werden. Hier und nur hier hat kreative Abduktion ihren Platz.

Auch andere Autoren haben dies (mit z.T. anderer Begründung) betont (Alston 1971, Friedman 1974), und auch das Peircesche Beispiel von der Härte des Diamanten legt es nahe, denn das Ritzverhalten, das durch die dispositionelle Erklärung explanativ vereinheitlicht werden soll, ist ja eine empirische Regelmäßigkeit. Daraus ergibt sich eine wichtige Konsequenz für das Verhältnis von *Abduktion* zur *Induktion*. Mehrere inference-to-the-best-explanation Theoretiker – etwa Harman (1965) und Moser (1989, S. 113) – versuchen nämlich, Induktion auf Abduktion zurückzuführen. Da aber Abduktion nur funktioniert, wenn sie auf empirische Regelmäßigkeiten angewandt wird, letztere aber nur durch Induktion gewonnen werden können, ist ein solcher Reduktionsversuch nicht möglich; Induktion muß *unabhängig* begründet werden (s. dazu Schurz 1988, Kap. 4).¹⁵

Kreative Abduktion besteht beim bisherigen Stande unserer Überlegungen also darin, empirische Regelmäßigkeiten durch übergeordnete Theorien so zu erklären, daß unser Wissenssystem dabei möglichst vereinheitlicht wird, wobei diese Art der Vereinheitlichung immer auch empirischer Bewährung impliziert. Dagegen läßt sich zuguterletzt immer noch folgendes einwenden. Die Leistungen eines solchen Vereinheitlichungsansatzes mögen ja unter anderem darin bestehen, eine über Poppers Bewährungstheorie hinausgehende, wesentlich ausgefeiltere Theorie der Theorieevaluation zu entwickeln. Doch es fragt sich, warum hier denn von einer *eigenen*

¹⁵ Neben diesem eher grundsätzlichen Argument gibt es freilich eine Reihe weiterer Einwände gegen den Versuch, Induktion als abduktive inference-to-the-best-explanation anzusehen. Wenn Moser argumentiert, die beste Erklärung dafür, daß alle bisher gekosteten Zitronen sauer seien, sei daß alle Zitronen sauer wären (1989, S. 113), so ist zuallererst anzumerken, daß eine bloße Allspezialisierung ($\forall xFx$, daher Fa) gemäß den Erklärungstheorien in der Hempel-Tradition überhaupt nicht als Erklärung zählt, da das singuläre Antecedens fehlt (Schurz 1988, Hrsg). Wenn man aber den Erklärungsbegriff so weit ausdehnen möchte, so ist zweitens sofort zu entgegnen, wieso denn nicht alternative Erklärungen, etwa die 'Goodman'-artige Erklärung "weil alle bisher gekosteten Zitronen sauer und die anderen süß sind", oder bloß die Selbsterklärung "weil alle bisher gekosteten Zitronen eben sauer waren", nicht ebenso gut sein sollten.

Schlußart gesprochen werden soll. Warum sollte man nicht besser bei der altvertrauten *Popperschen* Sichtweise verbleiben, derzufolge die Frage, *wie* wir erklärende Theorien generieren, gar nicht beantwortet zu werden braucht, weil sie nicht in den *Begründungs-*, sondern in den *Entdeckungszusammenhang* von Theorien fällt. Wie wir unsere Theorien entwickeln, ob durch ein Schema, im Traum oder durch Intuition, ist – wie Popper immer wieder betont hat (z.B. 1935, Kap. 1) – für die Wissenschaftstheorie belanglos: wichtig ist nur, daß wir unsere theoretischen Hypothesen hinterher einer strengen Prüfung unterziehen, worin sie sich als empirisch bewährt herausstellen. Kurz gesagt, existiert das Problem der kreativen Abduktion für den Popperianer gar nicht, denn es gehört in den Entdeckungszusammenhang.

In der Tat bedeutet mein nun folgender Versuch, ein abduktive Schlußschemata auszumachen, so etwas wie eine *Logik des Entdeckungszusammenhangs* zu entwickeln. Natürlich nicht als Verfahren, daß per se auch Begründungen liefert und damit separate Bewährungsversuche überflüssig macht, sondern als ein Verfahren, das uns *Heuristiken* angibt, wie wir methodisch neue Konzepte bilden können, und welches auch sagt, wann und warum sie die Vereinheitlichungsfunktion erfüllen. Wie wichtig etwas derartiges ist, wird unter anderem in der jüngeren Künstlichen Intelligenz-Forschung deutlich. Es erweist sich hier immer noch als ungemein schwierig, gewöhnliche 'intuitive' Denkleistungen des Common Sense algorithmisch zu implementieren. Das Problem beim Popperschen trial-and-error Verfahren ist die sogenannte *kombinatorische Explosion*. Wie es Peirce bereits einmal gesagt hat, gäbe es Millionen rein syntaktischer Möglichkeiten, zu gegebenen vorliegenden empirischen Regelmäßigkeiten irgendwelche Theorien zu erfinden, die diese Regelmäßigkeiten deduktiv implizieren – und dennoch haben die Wissenschaftler, so Peirce, die 'richtige' Theorie schon nach zwei oder drei Rateversuchen gefunden (1903, § 172). Peirce weiß keine wirkliche Erklärung hierfür und führt dies auf menschlichen *Abduktionsinstinkte* zurück, die sich im Laufe der Evolution entwickelt haben (1903, § 47, § 212). Es ist von daher ein spannendes Unternehmen, nach Schemata zu suchen, die unseren kreativen Abduktionen unterliegen. Im folgenden unternehme ich einen (ersten) solchen Versuch. Dabei möchte ich kreative

Abduktion keineswegs allein auf das von mir vorgeschlagene Schema reduzieren, es gibt noch andere kreative Abduktionsarten¹⁶; aber ich halte das von mir vorgeschlagene Schema für das elementarste.

3. Ein Schema der kreativen Abduktion

Das Schema hat zwei Komponenten:

Erstens, den (von mir sogenannten) *abduktive Dispositionsschluß*, den Schluß von einer *lokalen temporalen Regelmäßigkeit* auf eine Disposition.

Zweitens, die Vereinheitlichung, die dieser Schluß dann und nur dann leistet, wenn es sich um mehrere untereinander *korrelierte Regelmäßigkeiten* handelt, deren *gemeinsame Ursache* die abduktive erschlossene Disposition ist.

Nur wenn beides vorhanden ist, ist kreative Abduktion rational gerechtfertigt. Der antreibende Motor dieser Art kreativen Abduktion ist eine bestimmte Version des Prinzips des zureichenden Grundes. Wir betrachten zuerst den Vorgang in der Wissenschaftstheorie, und zeigen dann, wie der erkenntnistheoretische Realschluß ganz analog zu behandeln ist.

Gegeben eine temporale Regelmäßigkeit, die aber nur auf ein oder einige bestimmte Objekte zutrifft, eine sogenannte

¹⁶ Z.B. der *Analogie-Abduktionsschluß*, wo der neu eingeführte theoretische Term und das mit ihm verbundene Gesetz per Analogie aus einem anderen Bereich übertragen wird (Beispiel: statistische Theormodynamik – Newtonsche Ballistik). Solche Analogie-Abduktionen setzen die Existenz theoretischer Begriffe in anderen Bereichen bereits voraus und sind daher Abduktionen 'zweiter Stufe', während es uns um Abduktionen fundamentaler Art geht, worin theoretische Begriffe zuallererst eingeführt werden.

Lokale temporale Regelmäßigkeit:

$\forall t (Uat \rightarrow Eat)$ (Uxt – Umstände, 'Testbedingungen', Ext – beobachteter Effekt)

Beispiel: Dieses Objekt leitet, wann immer man eine Spannungsquelle anlegt, den Strom.

Frage: warum trifft diese Regelmäßigkeit nur auf dieses bzw. auf einige, aber nicht auf andere bzw. auf alle Objekte zu? Der Grund muß in einer inneren oder intrinsischen Eigenschaft liegen, die das Objekt a hat und die andere Objekte nicht haben. Hier wird somit folgende, als Heuristik anzusehende, Version vom Prinzip des zureichenden Grundes benutzt:

Temporale Regelmäßigkeiten, die an einigen Orten bzw. Individuen auftreten, haben ihre Ursache in intrinsischen Eigenschaften, die an diesen Orten bzw. an diesen Individuen instanziiert sind.

Wir vermuten in unserem Fall eine innere Eigenschaft bzw. Disposition als Ursache dieser Regelmäßigkeit und nennen sie *elektrische Leitfähigkeit*. Dies führt uns dazu, zwei theoretische Gesetze zu abduzieren. Das erste besagt, daß dies Disposition D die Ursache der Regelmäßigkeit ist:

$\forall x (Dx \rightarrow \forall t(Uxt \rightarrow Ext))$ äquivalent: $\forall xt: Uxt \rightarrow (Dx \rightarrow Ext)$

Das zweite theoretische Gesetz besagt, daß wir die Disposition der elektrischen Leitfähigkeit an eben dieser Regelmäßigkeit erkennen können. Dies können wir mithilfe der wahrheitsfunktionalen materialen Implikation nicht mehr so einfach wiedergeben. Die Formulierung $\forall x (\forall t(Uxt \rightarrow Ext) \rightarrow Dx)$ wäre *falsch*. Denn dann müßten alle Objekte, an die nie eine Spannungsquelle angelegt wurde, automatisch als elektrisch leitfähig betrachtet werden. Diese Schwierigkeit ist aus der Diskussion um die Definierbarkeit von Dispositionsprädikaten wohlbekannt (vgl. Tuomela 1978 oder Stegmüller 1970). Man kann vom empirischen Verhalten auf die zugrunde-

liegende Disposition nur dann schließen, wenn die Umstände bzw. Testbedingungen positiv realisiert sind. Wir nehmen einfachheitshalber an, es handle sich um eine Situation, wo wir wissen, daß Dispositionen entweder permanent auftritt oder nie: ein einziger Test genügt also, um herauszufinden, ob das Objekt leitfähig ist (so auch in Kaila's Schema; s. Stegmüller 1970, S. 222). Dann lautet das zweite theoretische Gesetz:

$$\forall x \forall t ((U_{xt} \wedge Ext) \rightarrow Dx) \quad \text{äquivalent: } \forall x (\exists t (U_{xt} \wedge Ext) \rightarrow Dx)$$

Die beiden theoretischen Gesetze lassen sich nun in der Form eines Carnapschen bilateralen Reduktionssatzes (s. Carnap 1936/37) zusammenfassen:

$$\forall xt (U_{xt} \rightarrow (Dx \leftrightarrow Ext))$$

Bei dieser Explikation des abduktiven Dispositionsschlusses habe ich bewußt vereinfacht. Folgende Komplikationen möchte ich nur am Rande erwähnen. *Erstens*, wenn wir annehmen, daß nicht schon ein einziger positiver Test genügt, um auf das Vorliegen der permanenten Disposition zu schließen, muß der obige bilaterale Reduktionssatz komplizierter aussehen, nämlich: $\forall x (\exists t (U_{xt} \wedge Ext) \rightarrow (\forall t (U_{xt} \rightarrow Ext) \leftrightarrow Dx)$. Aber ich glaube, zeigen zu können, daß sich dadurch an den folgenden Überlegungen nichts Prinzipielles ändert. *Zweitens* läßt sich der abduktive Dispositionsschluß nicht nur auf permanente, sondern auch auf nur temporal, also in einem gewissen Zeitintervall, auftretende Dispositionen anwenden -ein Beispiel wäre die Disposition, magnetisch zu sein, welches ein magnetisiertes Eisenstück eine Zeit lang hindurch besitzt. Auch hier treffen alle folgenden Betrachtungen grundsätzlich zu. *Drittens* besteht bekanntlich die Möglichkeit, Dispositionsbegriffe mithilfe der kontrafaktischen Implikation \Rightarrow zu definieren. Das kontrafaktische Konditional ist jedoch selbst als ein theoretischer Begriff anzusehen, welcher analoge Eigenschaften wie die oben eingeführte Disposition besitzt, weshalb sich ebenfalls nichts Grund-

sätzliches ändern würde.¹⁷

Zusammengefaßt ist hat der abduktive Dispositionsschluß folgende Form:

$\forall t(Uat \rightarrow Eat)$ (für einige aber nicht alle a), <i>daher</i> :	Explanadum
Da, und $\forall x \forall t(Uxt \rightarrow (Dx \leftrightarrow Ext))$	Explanans

Für sich genommen ist der abduktive Dispositionsschluß zu wenig, um als rational bzw. wissenschaftlich gerechtfertigt zu gelten, denn bei Anwendung auf nur *eine* Regelmäßigkeit kommt noch keine Wissensvereinheitlichung zustande (Schurz/Lambert 1994, S. 87). Dementsprechend läßt sich intuitiv einwenden, daß dieser Schluß keine wirklich neue Erkenntnis generiert, sondern die Dispositionsaussage bloß eine sprachliche *Abkürzung* dafür ist, daß eben eine entsprechende Regelmäßigkeit vorliegt. Hier kommt nun die zweite Bedingung zum Einsatz, derzufolge eine zu Recht eingeführte theoretische Entität die gemeinsame Ursache mehrerer korrelierter Regelmäßigkeiten sein muß. An unserem Beispiel: wir stellen nicht nur fest, daß einige Objekte immer dann, wenn wir eine Spannungsquelle anlegen, den Strom leiten, sondern auch, daß genau die Objekte, die diese Regelmäßigkeit aufweisen, zugleich eine Reihe anderer Regelmäßigkeiten aufweisen. Z.B. wenn man sie in heiße Umgebung bringt, leiten sie die Wärme (*Wärmeleitfähigkeit*), wenn man sie poliert und Bestrahlung aussetzt, glänzen sie (*Glanz*), wenn man sie starkem Druck aussetzt, dann biegen sie sich, ohne zu brechen oder zu reißen (*Biegsamkeit*), und weiteres mehr. Diese Korrelation von Regelmäßigkeiten kann nur so erklärt werden, daß eine tieferliegende intrinsische Qualität in diesen gewissen Substanzexemplaren vorhanden sein muß, die der gemeinsame Grund für alle diese Regelmäßigkeiten ist. Ansonsten wäre die Korrelation dieser Regelmäßigkeiten ein merkwürdiger Zufall. Wir abduzieren also eine noch tiefliegendere Disposition; in der Chemie nennt man sie *Metallcharakter*. Die angeführten Eigenschaften sind nämlich genau die charakteristischen Eigenschaften von Metallen. Im Gegensatz zu ein-

¹⁷ Z.B. interpoliert es, wie die Disposition, zwischen Konjunktion und materialer Implikation: $Uxt \wedge Ext \vdash Uxt \Rightarrow Fxt \vdash Uxt \rightarrow Fxt$.

fachen Dispositionsbegriffen, die nur einer Regelmäßigkeit entsprechen, liegt jetzt ein echter *theoretischer* Begriff vor – das Unterscheidungskriterium besteht darin, daß theoretische Begriffe mehrere Regelmäßigkeiten vereinen. (Ich spreche auch von Dispositionsbegriffen erster versus zweiter Stufe). Und jetzt leistet der abduktive Dispositionsschluß plötzlich eine ganz enorme Vereinheitlichung. Nehmen wir an, wir haben n verschiedene temporale Regelmäßigkeiten

$$\forall t(U_{iat_i} \rightarrow E_{iat_i}) \quad \text{für } 1 \leq i \leq n$$

die alle untereinander korreliert sind. Aus dieser Korrelation ergeben sich $n \cdot (n-1)$ Implikationsbeziehungen:

$$\forall x \forall t((U_{ixt} \wedge E_{jxt}) \rightarrow (U_{jxt} \rightarrow E_{jxt})) \quad \text{für } 1 \leq j \neq i \leq n$$

Zusammen mit den obigen n temporalen Regelmäßigkeiten sind dies $n \cdot (n-1) + n = n^2$ elementare empirische Gesetze. Von diesen abduzieren wir n bilaterale Reduktionssätze

$$\forall x \forall t(U_{ixt} \rightarrow (D_x \rightarrow E_{ixt})) \quad 1 \leq i \leq n$$

welche $2n$ elementaren theoretische Gesetze entsprechen - $2n$, weil die Äquivalenz (gemäß der erwähnten Theorie relevanter Wissensselemente; s. Fn. 14) in zwei elementare Implikationsgesetze zerfällt. Aus diesen $2n$ theoretischen Gesetzen können alle n^2 empirischen Gesetze deduktiv erklärt werden. Somit handelt es sich um eine Reduktion elementarer Gesetze von n^2 auf $2n$, bzw. allgemeiner gesprochen von quadratisch auf linear in der Anzahl der Umstands-Effekt-Paare, was eine beachtliche Vereinheitlichung darstellt.

Wie wir ausführten, geht in unsere zweite Bedingung ein Prinzip der *gemeinsamen Ursache* ein. Dieses Prinzip ist von seiner Idee her nicht neu. Reichenbach (1959, ch.3, ch.5) hat ein *common cause* Prinzip entwickelt, welches von Salmon

(1978; 1984, S.206-227) ausgebaut und von van Fraassen (1980, S.25-31) kritisiert wurde. Reichenbach's common cause Prinzip besagt, daß wann immer zwei Ereignistypen F_{xt} und G_{xt} gleichzeitig (statistisch oder im klassischen Fall gar deterministisch) korrelieren, diese Korrelation auf eine gemeinsame zeitlich davorliegende Ursache zurückzuführen sein muß. Reichenbach's Prinzip setzt offenbar die relativistische Kausalitätstheorie voraus, derzufolge Kausalbeziehungen sich mit endlicher Geschwindigkeit zeitlich vorwärts gerichtet ausbreiten. Andernfalls wäre ja auch die wesentlich einfachere Erklärung möglich, daß entweder F_{xt} die zeitlich simultane Ursache für G_{xt} ist, oder umgekehrt. Genau in diesem Sinn wurde Reichenbach's common cause Prinzip auch von van Fraassen kritisiert, unter Bezugnahme auf simultane Korrelationen ohne 'hidden variables' in der Quantenmechanik. In jedem Falle setzt Reichenbach's common cause Prinzip eine Menge Theorie und damit eine Menge abduzierter theoretischer Terme bereits voraus. Das ist der erste Unterschied zu dem hier vorgeschlagenen Prinzip: letzteres ist *elementar* und setzt keinerlei bereits vorhandenes theoretisches Vorwissen voraus (s. auch Fn. 16). Zweitens, und damit zusammenhängend, fehlt bei Reichenbach die Bedingung der Existenz mehrerer korrelierter Regelmäßigkeiten bzw. Korrelationen.

Ich zeige abschließend, wie der abduktive Dispositionsschluß in strukturell gleichartiger Weise den erkenntnistheoretischen Realschluß zu rechtfertigen vermag. Der Anfang unserer Erkenntnis sind introspektiven Sätze und introspektiven Begriffe, wie z.B. "Baumerscheinung". Beobachtungsrealbegriffe sind selbst Dispositionsbegriffe nullter Stufe, wie ich sie nenne. Wenn wir nun vom Erscheinen auf das Sein schließen, so liegen dem *introspektive Regelmäßigkeiten* zugrunde, z.B.: immer wenn ich dorthin sehe, habe ich eine dort lokalisierte Baumerscheinung. Wir beschränken uns hier wieder auf den Fall permanenter Dispositionen – d.h. der Baum ist immer dort. Introspektiv liegt also folgende Regelmäßigkeit vor:

$\forall t(S(i,s,t) \rightarrow F_S(i,s,t))$ $S(i,s,t)$ – ich sehe zur Zeit t auf den Ort s
 $F_S(i,s,t)$ – ich habe zur Zeit t eine am Ort s lokalisierte
 F-Erscheinung

In Worten, wannimmer ich auf den Ort s sehe, habe ich eine am Ort s lokalisierte F-Seherscheinung. Der Index S am F indiziert, daß es sich um eine visuelle Erscheinung handelt.¹⁸ Auch das Prädikat S ist als Erscheinungsprädikat aufzufassen: wannimmer ich zum Ort hinzusehen scheine. Auch sprechen wir auf dieser Stufe noch nicht von Individuen, sondern bloß von Orten oder Lokalisationen s .

Gemäß dem Schema des Dispositionsschlusses muß die Tatsache, daß diese Regelmäßigkeit nur für diesen und nicht für andere Orte gilt, den Grund in einer intrinsischen Eigenschaft oder Disposition haben, die an diesem Ort angesiedelt ist – diese Disposition nullter Stufe nennen wir Baum. Wir abduzieren also wie folgt:

$F(s)$, und: $\forall s \forall t(S(i,s,t) \rightarrow (F(x) \leftrightarrow F_S(i,s,t)))$

in Worten: am Ort s ist ein F , und (der bilaterale Reduktionssatz) wenn ich auf einen beliebigen Ort sehe, dann befindet sich dort ein F genau dann wenn ich eine F-Erscheinung habe. *Allein* leistet dieser Schluß noch keine Vereinheitlichung. Der springende Punkt ist erneut, daß *viele korrelierte* Regelmäßigkeiten vorliegen. Vor allem gibt es die die Korrelation von Sehsinn und Tastsinn. D.h. genau dann, wenn ich regelmäßig nach Hinsehen dort eine Baumerscheinung erblicke, kann ich auch regelmäßig durch Abtasten dort eine Baumerscheinung ertasten. Ich spreche hier von *intersensueller Korrelation*. Aber es gibt auch Unmengen von korrelierten Regelmäßigkeiten innerhalb eines Sinnes - sogenannte *intrasensuelle Korrelationen*. Je nachdem, von welcher Richtung ich beispielsweise schaue, sehe ich den Baum etwas anders, aber die verschiedenen Sehbilder sind durch geometrische Transformationen eindeutig korreliert. Es sind maximal vier Richtungen nötig (von

¹⁸ Man könnte dies auch anders formalisieren, z.B. in der modallogischen Form $S(i,Fst)$ mit S als 'Sehoperator' – worauf ich hier nicht näher eingehe.

vorn, von hinten, von rechts, von links), um aus den zweidimensionalen Objektbildern das dreidimensionale Objekt und damit sein Bild aus allen anderen möglichen Blickwinkeln eindeutig rekonstruieren zu können (was im geometrischen Zeichnen, in der Architektur und in der KI der Bildidentifizierung eine bedeutende Rolle spielt; der Common Sense erledigt dies sozusagen 'im Schlaf').

Allgemein gesprochen liegen nun n korrelierte introspektive Regelmäßigkeiten folgender Form vor. Dabei steht W_i für die Einsetzung eines Wahrnehmungssinnes in einer bestimmten Weise (z.B. Sehen von vorne, Ertasten von hinten, usw.).

$$\forall t(W_k(i,s,t) \rightarrow F_{W_k}(i,s,t)) \quad 1 \leq k \leq n.$$

Die Korrelation dieser n introspektiven Regelmäßigkeiten drückt sich in folgenden $n \cdot (n-1)$ introspektiven Implikationsgesetzen aus:

$$\forall x(W_k(i,s,t) \wedge F_{W_k}(i,s,t) \rightarrow (W_j(i,s,t) \rightarrow F_{W_j}(i,s,t)) \quad 1 \leq k \neq j \leq n.$$

Diese $n(n-1) + n = n^2$ introspektiven Beziehungen werden durch Abduktion erneut auf $2n$ elementaren Realgesetze (in der Form bilateraler Reduktionssätze) zurückgeführt:

$$\forall st(W_k(i,s,t) \rightarrow (F(s) \leftrightarrow F_{W_k}(i,s,t)) \quad 1 \leq k \leq n.$$

Auch bei meiner Explikation des abduktiven Realschlusses habe ich äußerst vereinfacht und kann folgende Komplikationen nur noch am Rande erwähnen. *Erstens* gibt es auf nächster Ebene erneute Korrelationen, denn verschiedene elementare Eigenschaften treten immer gebündelt auf. Diese Korrelationen veranlassen zur Konzeptualisierung von Individuen bzw. Gegenständen, welche diesen Eigenschaftsbündeln entsprechen. *Zweitens*, wenn durch elementare Realabduktion einmal die objektive Existenz *anderer* Menschen eingeführt wurde, so läßt sich in obigem Schema auch der *Ichparameter* "i" in Form eines Realsubjektparameters variieren,

und dadurch zwischen intersubjektiven *Beobachtungsmerkmalen* versus subjektiven *Geschmacksmerkmalen* unterscheiden. Das ist durchaus subtil, denn grob gesprochen ist Realabduktion *prima facie* auch bei Geschmacksmerkmalen gerechtfertigt, weil *ein* Grund, warum ich gerade dies regelmäßig als schön empfinde, und dies andere nicht, in den intrinsischen Eigenschaften des Objektes liegen muß. Der Unterschied zu Beobachtungsmerkmalen ist hier, daß im Geschmacksfall ein zweiter Grund auch in mir liegt – diesen kann ich aber erst herausfinden, wenn ich den Ichparameter variieren kann, also die intersubjektive Ebene betrete. Für Beobachtungsmerkmale ergeben sich dann *drittens* eine Unmenge weiterer Korrelationen, nämlich alle *intersubjektiven* Korrelationen.

Man kann in der hier vorgeschlagenen Rechtfertigung des erkenntnistheoretischen Realismus eine Einlösung des Quineschen Programmes einer *naturalisierten Erkenntnistheorie* erblicken (Quine 1975). Die Pointe unserer Rechtfertigung besteht ja darin, das dasselbe abduktive Vereinheitlichungsschema, welches die Einführung theoretischer Terme in der Wissenschaft zu modellieren vermag, auch die Einführung des Systems der Realbegriffe erklärt. Natürlich ist sich niemand bewußt, daß beim Schluß auf die Realität eine solche subtile Vereinheitlichung vor sich geht. Aber daß unbewußte oder angeborene kognitive Vorgänge sehr komplex sein können, ist kognitionspsychologisch unbestritten.

Ich glaube auch, daß die erläuterte Vereinheitlichungsfunktion auch ganz wesentlich ist für unser *Vertrauen in die Realität*. Um dies zu testen, können wir uns Szenarios vorstellen, in denen diese Korrelationen nicht mehr gegeben sind: wie schnell würde darin unser Vertrauen in die objektive Realität verschwinden! Was wäre etwa, wenn Seh- und Tastsinn nicht so schön korreliert wären? Diese Visionen sind uralt: es sind Gespenster- bzw. Geistesvisionen. Ein Gespenst ist, fast möchte man sagen 'per definitionem', etwas, daß man sehen kann, ohne es ertasten zu können: man kann durch das Gespenst hindurchlaufen, bzw. das Gespenst kann durch beliebige Gegenstände hindurchlaufen. Für das Raumerlernen des Kleinstkindes ist die Korrelation von Seh- und Tastsinn bekanntlich essentiell; in einer Gespensterwelt könnte es sich wahrscheinlich erst gar nicht orientieren lernen. Auch der Er-

wachsene nimmt in jenen seltenen Situationen, wo er seinem Sehbild nicht mehr vertraut, sofort seinen Tatssinn zuhilfe: er versucht, Tastkontakt mit der fraglichen Seherscheinung aufzunehmen; hat er sie ergriffen, so ist er sicher, daß es sich um ein wirkliches Objekt und keine Halluzination handelt. Gegen Solipsisten Berkeleyscher Prägung haben einige Philosophen spaßhalber eingewandt, man müsse ihnen den Prügel, der ihnen erscheint, bloß auf den Kopf schlagen, um sie vom Solipsismus abzubringen. Aber auch Gespensterweilten sind noch ziemlich wohlgeordnet. Stellen wir uns vor, auch die Korrelation der Sehbilder aus verschiedenen Richtungen sind nicht mehr vorhanden: aus der Teufelerscheinung wird, wenn wir uns leicht drehen, eine Engelserscheinung, und gehen wir 5 cm nach vor, eine Krokodilerscheinung. Man könnte dies beliebig fortsetzen, etwa indem man nun auch die Eigenschaftsbündel zerfließen läßt: hier eine konturlose Röte, dort eine Kontur ohne Inhalt, da ein tastbarer aber unsichtbarer Widerstand, usw. Mit derlei Phantasieanregungen sei dieser Aufsatz beschlossen.

Literatur

- Alston, W. P. (1971): "The Place of the Explanation of Particular Facts in Science", *Philosophy of Science* 38, 13-34.
- Alston, W.P. (1989): *Epistemic Justification*, Cornell Univ Press, Ithaca/London.
- Chisholm, R.M. (1979): *Erkenntnistheorie*, dtv, München (Orig. 1966).
- Carnap, R. (1936/7): "Testability and Meaning", *Philosophy of Science*, Part I: Vol. 3, 419 - 471; Part II: Vol. 4, 2 - 40. (Selbstständig erschienen: New Haven 1954).
- Carnap, R. (1956): "The Methodological Character of Theoretical Concepts", in: Feigl, H./Scriven, M. (ed.), *Minnesota Studies in the Philosophy of Science Vol. I*, Univ. of Minnesota Press, Minneapolis, S. 38-76.
- Carnap, R. (1959). *Induktive Logik und Wahrscheinlichkeit*, bearbeitet von W. Stegmüller, Springer, Wien.
- Feigl, H.: 1970, "The Orthodox View of Theories: Remarks in Defense as well as Critique", in: *Minnesota Studies in the Philosophy of Science*, Vol IV, University of Minnesota Press, Minneapolis.
- Friedman, M. (1974): "Explanation and Scientific Understanding", *Journal of Philosophy* 71, 5-19.
- Fumerton, R.A.(1980): "Induction and Reasoning to the Best Explanation", *Philosophy of Science* 47, 589-600.
- Goldman, A. (1986): *Epistemology and Cognition*, Harvard Univ. Press, Cambr./Mass.
- Harman, G.H. (1965): "The Inference to the best Explanation", *Philosophical Review* 74, 88-95.
- Hempel, C. G. (1950): "Problems and Changes in the Empiricist Criterion of

- Meaning", *Revue Internationale de Philosophie* 4, S. 41-63.
- Humphreys, P. (1993): "Greater Unification equals greater Understanding?", *Analysis* 53.3, 183-188.
- Kitcher, P. (1981): "Explanatory Unification", *Philosophy of Science* 48, 507-531.
- Lehrer, K. (1974): *Knowledge*, Clarendon Press, Oxford.
- Locke, J. (1690): *Über den menschlichen Verstand* (In vier Büchern), Übers. v. C. Winckler, 3. Aufl. Hamburg 1976.
- Mach, E. (1883): *Die Mechanik*, Wissenschaftliche Buchgesellschaft, Darmstadt (1973).
- Moser, P.K. (1989): *Knowledge and Evidence*, Cambridge Univ. Press.
- Pollock, J. (1986): *Contemporary Theories of Knowledge*, Rowman & Littlefield Maryland.
- Peirce, C. S. (1868): "Some Consequences of Four Incapacities", *Collected Papers* Bd 5, §§ 264-317.
- Peirce, C.S. (1878a): "How to Make Our Ideas Clear", *Collected Papers* Bd 5, §§ 388-410.
- Peirce, C.S. (1878b): "Deduction, Induction, and Hypothesis", *Collected Papers* Bd 2, §§ 619-644.
- Peirce, C.S. (1892): "The Doctrine of Necessity Examined", *Collected Papers* Bd 6, §§ 35-65,
- Peirce, C.S. (1902-3b): "Objective Logic" (from "Minute Logic"), *Collected Papers* Bd 2, §§ 111-118.
- Peirce, C.S. (1903): "Lectures on Pragmatism", *Collected Papers* Bd 5, §§ 14-212,
- Peirce, C.S. (1905): "Issues of Pragmaticism" *Collected Papers* Bd 5, §§ 8-463.
- Peirce, C.S.: *Collected Papers*, Bände 1-6 hrsg. von Hartshorne, C. und Weiss, P., 1931-35; Bände 7-8 hrsg. von Burks, A. W., 1958; Harvard Univ. Press, Cambridge/Mass.
- Popper, Karl R. (1935): *Logik der Forschung*, 10. Aufl., Tübingen, Mohr (1994).
- Quine, W.v.O. (1975): "Naturalisierte Erkenntnistheorie", in: ders., *Ontologische Relativität und andere Schriften*, reclam, Stuttgart, S. 97-126 (engl. Orig. 1969).
- Reichenbach, H. (1959): *Modern Philosophy of Science*, Routledge and Kegan Paul, London.
- Salmon, W. (1978): "Why ask 'Why?'" *Proc. Adr. Amer. Phil. Assoc.* 51, 683 -705.
- Salmon, W. (1984): *Scientific Explanation and the Causal Structure of the World*, Princeton Univ. Press, Princeton.
- Schurz, G. (1988): "Kontext, Erfahrung und Induktion: Antworten der pragmatischen Wissenschaftstheorie auf drei Herausforderungen", *Philosophia Naturalis* Band 25, Heft 3-4, 296-336.
- Schurz, G. (1988, Hrsg): *Erklären und Verstehen in der Wissenschaft*, Oldenbourg, Scientia Nova, München, 2. Auflage Paperback 1990.
- Schurz, G. (1991a): "Charles Sanders Peirce: Die pragmatische Theorie der Erkenntnis", in: Speck, J. (Hg.), *Grundprobleme der großen Philosophen*, Band Neuzeit V, Vandenhoeck und Ruprecht (UTB), Göttingen, 115 - 169.
- Schurz, G. (1991b): "Relevant Deduction", *Erkenntnis* 35, 391 - 437.
- Schurz, G. (1996a): "The Role of Relevance in Deductive Reasoning", *paper submitted*.
- Schurz, G. (1996b): "Unification, Understanding, and Scientific Progress", in: G. Meggle, P. Steinacker (Hrsg.), *Proceedings of Analyomen II*, de Gruyter, Berlin.

- Schurz, G. und Lambert, K. (1994): "Outline of a Theory of Scientific Understanding", *Synthese* 101, 65-120.
- Stegmüller, W. (1970): *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*, Band II, Studienausgabe Teil B, Springer, Berlin.
- Thagard, P. (1989): "Explanatory Coherence", *Behavioral and Brain Sciences* 12, 435-469.
- Thagard, P. (1992): *Conceptual Revolutions*, Princeton Univ. Press.
- Tuomela, R. (1978, Hrsg.): *Dispositions*, Reidel, Dordrecht.
- Van Fraassen, B. (1980): *The Scientific Image*, Clarendon Press, Oxford.
- Van Fraassen, B. (1989): *Laws and Symmetries*, Clarendon, Oxford.
- Whewell, W.: 1847, *The Philosophy of the Inductive Sciences*, 2nd edition, 2 Volumes, John W. Parker, London.