

# ETHIK UND SOZIALWISSENSCHAFTEN

Streitforum für Erwägungskultur

EuS 1 (1990) Heft 1

## INHALT

**EDITORIAL** 5

### ERSTE DISKUSSIONSEINHEIT UND METAKRITIK

#### HAUPTARTIKEL

Hannelore Bublitz: *Theorie-Disziplin der Moderne* 7

#### KRITIK

- Rainer Greshoff: *Wissenschaftliche Aufklärung als Einfältigkeit?* 18  
Carol Hagemann-White: *Soll Theorie gesellschaftliche Urteilskraft sein?* 20  
Marlis Krüger: *Einige Anmerkungen zu "Theorie-Disziplin der Moderne" von Hannelore Bublitz* 22  
Marianne Krüll: *Die Rekursivität der Erkenntnis und ihre (möglichen) Folgen* 23  
Werner Loh: *Emanzipation durch Selbstentmündigung?* 25  
Bärbel Meurer: *Eine "neue Logik" oder "Können Hexen fliegen?"* 26  
Peter Rech: *Wider die formale Rationalität* 28  
Elvira Scheich: *Das abstrakte "Konkrete" und die zweite Natur der objektiven Wissenschaften* 29  
Hannelore Schröder: *Patriarchalismus-Theorie statt Rationalitätsdiskussionen* 31  
Johannes Weiß: *Anything goes - but where?* 33

#### REPLIK

Hannelore Bublitz: *Die Einfalt der sinnlichen Vernunft oder Wunschbild und Alptraum der begrifflichen Abstraktion* 34

#### METAKRITIK

Ilse Modelmog: *Hin und Wider* 40

### ZWEITE DISKUSSIONSEINHEIT

#### HAUPTARTIKEL

Hans Lenk und Matthias Maring: *Verantwortung und soziale Fallen* 49

#### KRITIK

- Dieter Bierlein: *Verantwortung in spieltheoretischer Sicht* 57  
Nina Hager: *Vernunft und Verantwortung* 59  
Bernd Halfar: *Das Prisoners' Dilemma als theoretische Falle* 61  
Winfried Hassemer: *Kollektive Verantwortung und moralische Fallen* 63  
Rainer Hegselmann: *Zur spieltheoretischen Rekonstruktion desasterträchtiger Situationen* 65

<b>Karl Homann:</b>	<i>Kollektive Probleme und individualethisches Paradigma</i>	67
<b>Franz-Xaver Kaufmann:</b>	<i>Leistet Verantwortung, was wir ihr zumuten?</i>	70
<b>Hartmut Kliemt:</b>	<i>Statische Sozialfallen und repetitive Spiele</i>	72
<b>Karl-Heinz Ladeur:</b>	<i>Ethik der Komplexität und gesellschaftliche Institutionen</i>	74
<b>Werner Loh:</b>	<i>Unverantwortbarer Fortschritt ohne Fortschritt der Verantwortung?</i>	77
<b>Heinz-Ulrich Nennen:</b>	<i>Kritik des technologischen Ökologismus</i>	80
<b>Otto Neumaier:</b>	<i>Zur theoretischen Vermeidung sozialer Fallen</i>	81
<b>Rainer Piepmeier:</b>	<i>Zur Pragmatik und Begründung des Begriffs Verantwortung</i>	84
<b>Friedrich Rapp:</b>	<i>Wahrnehmungsschwellen, Zielkonflikte und Institutionen</i>	87
<b>Günter Ropohl:</b>	<i>Ethik in der "individualistischen Falle"</i>	89
<b>Erwin K. Scheuch:</b>	<i>Ethische Probleme als Forschungsartefakte</i>	91
<b>Viktor Vanberg:</b>	<i>Die Grenzen von Verantwortung und die Bedeutung von Regeln</i>	93
<b>Hellmut Willke:</b>	<i>Von der Theorie der Verantwortung zur Verantwortung der Theorie</i>	95

**REPLIK**

<b>Hans Lenk und Matthias Maring:</b>	<i>Autoren in der Interdisziplinaritätsfalle?</i>	97
---------------------------------------	---	----

**DRITTE DISKUSSIONSEINHEIT, METAKRITIK UND BRIEF****HAUPTARTIKEL**

<b>Dieter Wandschneider:</b>	<i>Die Gödeltheoreme und das Problem Künstlicher Intelligenz</i>	107
------------------------------	--	-----

**KRITIK**

<b>Lutz-Michael Alisch:</b>	<i>Limitieren Grundagentheoreme das menschliche Denken?</i>	116
<b>Ansgar Beckermann:</b>	<i>Alles klar?</i>	119
<b>Ulrich Blau:</b>	<i>Zum Erweiterungsprozeß formaler Systeme</i>	123
<b>Lothar Eley:</b>	<i>Zwei kritische Anmerkungen zum semantischen Ansatz des Gödeltheorems</i>	123
<b>André Fuhrmann:</b>	<i>Über die Rolle von Selbstreferenz und Widerspruchsfreiheit im Beweis von Gödels Unvollständigkeitssatz</i>	125
<b>Sybille Krämer:</b>	<i>Das Scheitern der universalen Denkmaschine</i>	128
<b>Rosemarie Rheinwald:</b>	<i>Gödel zwischen Objekt- und Metaebene</i>	130
<b>Michael Roth:</b>	<i>Zur KI-Interpretation der Gödeltheoreme von Wandschneider</i>	133
<b>Pirmin Stekeler-Weithofer:</b>	<i>Die (Selbst)Begrenzung mathematischen Denkens durch die Identifikation von Beweis und Deduktion</i>	134
<b>Joachim Stolz:</b>	<i>Gödeltheoreme, semantische und syntaktische Selbstreferenz</i>	138
<b>Rainer Stuhlmann-Laeisz:</b>	<i>Semantik, Unbeweisbarkeit und Selbstreferenz</i>	140
<b>Holm Tetens:</b>	<i>Gödel und die semantische Selbstreferenz</i>	142
<b>Matthias Varga von Kibéd:</b>	<i>Gödel wußte sehr wohl, was er meinte</i>	145

**REPLIK**

<b>Dieter Wandschneider:</b>	<i>Gödelsche Selbstreferentialität und maschinelle Aspekte</i>	148
------------------------------	--	-----

**METAKRITIK**

<b>Manfred Wolff:</b>	<i>"Semantische Selbstreferentialität" - ein neuer Gödel-Mythos</i>	154
-----------------------	---	-----

**BRIEF**

<b>Dieter Wandschneider</b>	159
-----------------------------	-----

## VIERTE DISKUSSIONSEINHEIT UND METAKRITIK

### HAUPTARTIKEL

Franz M. Wuketits: *Moral - eine biologische oder biologistische Kategorie?* 161

### KRITIK

Kurt Bayertz: *Wie biologisch ist die Moral?* 168

Bettina Blanck: *Erwägungen und Fragen zu einer evolutionären Wurzelbestimmung des Moralischen* 170

Ingrid Craemer-Ruegenberg: *Evolutionäre Ethik: Bescheidenheit und problematische Zurückweisungen* 173

Rainer Greshoff: *Biologismus durch die Hintertür?* 174

Michael Hampe: *Die Abstraktheit der Suche nach den biologischen Wurzeln der Moral* 176

Marianne Krüll: *Warum eine "Evolutionäre Ethik"?* 178

Wolfgang Kuhlmann: *Evolutionäre und normative Ethik* 180

Werner Leinfellner: *Wie sozial ist die Ethik?* 182

Werner Loh: *Eine Selbstauflösung Evolutionärer Ethik in ihrer unaufgeklärten Geschichtlichkeit* 185

Wilhelm Lütterfelds: *Die Kontingenz des moralischen Norminhaltes und seine evolutionäre Erklärung* 186

Peter Meyer: *Ethik als Metatheorie sozialer Ordnung: Evolutionistische Überlegungen* 188

Annemarie Pieper: *Evolutionäre Ethik und philosophische Ethik: unvereinbare Gegensätze?* 190

Günter Tembrock: *Ethologie und Ethik* 192

Karl-Friedrich Wessel: *Biotische Determination - nur eine Voraussetzung der evolutionären Ethik* 194

### REPLIK

Franz M. Wuketits: *Moral als eine biologische Kategorie* 195

### METAKRITIK

Matthias Rath: *Moral zwischen Geltung und Genese* 200

## ANHANG

**BRIEF** 209

**EuS-PROGRAMM** 210

**EuS-STATUT** 211

**LISTE DER BEIRATSMITGLIEDER VON EuS** 212

**EuS-THEMENLISTE** 214

**RICHTLINIEN FÜR BEITRÄGE IN EuS** 216

**LISTE DER VERÖFFENTLICHUNGSVORHABEN FÜR EuS** 217

	H. Tetens	: Te
	M. Varga v. Kibéd	: Va
Replik von	D. Wandschneider	: War

((2)) Im folgenden möchte ich zunächst in I das Anliegen von D. Wandschneider referieren, dann in II die Hauptpunkte der Kritiken skizzieren, in III die Replik von D. Wandschneider kritisch beleuchten und schließlich in IV eine Zusammenfassung der mir wesentlich scheinenden Ergebnisse der schriftlichen Diskussion angeben.

Vorab möchte ich jedoch schon sagen, daß ich alle Kritiken mit großem Gewinn für mich gelesen habe. Einige von ihnen sind knapper und berühren nicht alle Punkte von Wa, andere sind sehr umfangreich. Manche zeigen ein hervorragendes didaktisches Geschick, die Gödel-Theoreme korrekt darzustellen. Schließlich gehen die Erläuterungen etwa von Ste weit über den gesteckten Rahmen hinaus. Es ist mir unmöglich, alle Kritiken en detail zu würdigen. Ich werde mich auf die Hauptpunkte beschränken und - die Autoren mögen mir dies verzeihen - die hierzu vorliegenden Stellungnahmen nur sehr summarisch referieren.

Es ist mir ein Bedürfnis, den Kritikern zu danken. Durch ihre Beiträge ist das von Wandschneider angesprochene Thema so substantiell behandelt, daß der Leser wirklich einen Gewinn davon hat.

## I

### Metakritik

#### "Semantische Selbstreferentialität" - ein neuer Gödel-Mythos

Manfred Wolff

((1)) Für die Bezugnahme auf die Texte will ich folgende Abkürzungen verwenden:

Der Originalartikel		: WA
Kritiken von	L.M. Alisch	: AL
	A. Beckermann	: Be
	U. Blau	: Bl
	L. Eley	: El
	A. Fuhrmann	: Fu
	S. Krämer	: Kr
	S. Rheinwald	: Rh
	M. Roth	: Ro
	P. Stekeler-Weithofer	: Ste
	J. Stolz	: Sto
	R. Stuhlmann-Laeisz	: Stu

((3)) Mit seinem Aufsatz möchte D. Wandschneider zum Problem der Intelligenz von Mensch und Maschine Stellung nehmen. Sein Anliegen präzisiert er wie folgt: (Wa((3))) "Es soll der Frage nachgegangen werden, ob für Computer prinzipielle Leistungsgrenzen bezüglich ihrer logischen Möglichkeiten anzunehmen sind oder umgekehrt: ob die dem Denken zugrunde liegende Logik möglicherweise nicht objektivierbar, d.h. nicht in effektive Computeroperationen übersetzbar ist."

In dieser Form hängt die Frage eng mit der Theorie der rekursiven Funktionen zusammen und damit, wie Wandschneider im gleichen Absatz bemerkt, auch mit Gödels sog. Unvollständigkeitstheorem. Insbesondere Lucas hat diesen Satz als Argumentationshilfe für seine These benutzt, die zitierte Frage sei zu bejahen. Im vorliegenden Artikel (Wa) soll dagegen gezeigt werden, daß Gödels Resultate (und seine für sie gegebenen Beweise) keine Argumente für eine solche Bejahung liefern.

((4)) Wa ist in drei Teile gegliedert: in I (Abschnitte((4)) bis ((8)) werden die Resultate vorgestellt, in II (Abschnitte ((9)) bis ((20))) werden die Beweise analysiert. Dabei wird als wesentliche Eigenschaft des von Gödel angegebenen Satzes G (der in der zum konsistenten und genügend reichhaltigen System S gehörigen Sprache ausdrückbar, aber in S nicht entscheidbar ist) dessen semantische Selbstreferentialität ange-

sehen. In III (Abschnitte ((21)) bis ((39)) werden dann die gewonnenen Einsichten zu einem Begründungsversuch für die Hauptthese des Artikels benutzt, die man so formulieren kann:

(HT): Gödels Abhandlung (GÖDEL 31) liefert keine ausreichenden Argumente dafür, daß "die Logik des Denkens und die der Maschinen von prinzipiell verschiedener Art sind."

((5)) In Wa wird diese Hauptthese verschieden und durchaus auch nicht immer logisch äquivalent dazu formuliert. So findet man z.B. in Wa ((35)): "Gödels Theoreme, soviel ist damit klar, können weder als Beleg für eine grundsätzliche Differenz der Logik des Denkens und der Maschine noch auch im Sinne einer prinzipiellen Überlegenheit des Denkens über die Maschine gedeutet werden". Wie man eine solche Folgerung aus Gödels Theoremen ziehen kann, ist von trivialer Einfachheit (vgl. ((24)) unten). Ich hatte jedoch den Eindruck gewonnen, Wandschneider versuche, durch eine sorgfältige Analyse des Gödelschen Beweises noch andere interessante Argumente für die Schlußfolgerung in (HT) zu gewinnen. Deshalb habe ich in (HT) statt "Theoreme" den Ausdruck "Abhandlung" gewählt.

((6)) Außer diesem Hauptgedanken enthält der Abschnitt III von Wa noch allgemeinere Überlegungen zum Problem des Unterschiedes zwischen Mensch und Maschine, eingeschlossen eine spekulative These über das menschliche Gehirn und seine Funktionsweise.

## II

((7)) In einer Grundtendenz sind sich alle Kritiker einig: Die in (Wa) entwickelte Argumentation taugt nicht dazu, um die Hauptthese (HT) zu stützen.

Dies sagt jedoch noch nicht, daß alle Kritiker die These ablehnen. Rh ((1)) - ((3)) stimmt mit Wandschneider überein: "Wie immer das genaue Verhältnis zwischen Mensch und Maschine einzuschätzen sein mag, aus Gödels Theorem folgt nicht die intellektuelle Überlegenheit des Menschen über die Maschine..." Rheinwald zeigt anhand von Literaturangaben, daß dies die heute vorherrschende Meinung ist. Sto ((5)) sieht keinen Zusammenhang zwischen Gödels Theorem und dem Mensch-Maschine - Problem. Damit wird also indirekt (HT) unterstützt. Auch Te ((9)) enthält eine Bestätigung der These. Va ((33)) äußert sich vorsichtig in gegenteiliger Richtung. Ähnlich ist die Argumentation in Kr.

((8)) Bevor ich auf den Hauptteil der kritischen Einwände zu Wandschneiders Argumentation eingehe, möchte ich noch betonen, daß sich eine Reihe von Kritikern mit den weitergehenden Bemerkungen Wandschneiders ((29)), ((35)ff) zum Problem der künstlichen Intelligenz auseinandersetzen. Alisch ist der Ansicht, Wandschneider sage hierzu sehr wenig. Da in Wa das allgemeine Problem der KI zwar ausführlich ange-

sprochen ist (obwohl es wenig mit Gödels Theoremen zu tun hat, und daher nicht im Zentrum von Wa steht), aber kaum präzisiert wurde, möchte ich hier zitieren, wie Alisch es ausdifferenziert (Al ((1)), ((2))):

- "(a) Ist das Beweisen von Theoremen der entscheidende Punkt menschlicher Intelligenz?  
 (b) Warum sollten Geist-Modelle in Betracht gezogen werden, deren Wissen nur in logisch kodierter Form vorliegt und deren mentale Operationen sich in Deduktionen über geschlossenen Informationsmengen erschöpfen?  
 (c) Sollten Turing-Maschinen oder abstrakte Computer Gegenstand des Vergleichs sein?  
 (d) Sollte KI zweckmäßigerweise nur als äquifinale Reduplikation der menschlichen Intelligenz konzipiert werden?  
 (e) Wenn sich durch Rekurrieren auf semantische Aspekte der Selbstreferentialität eine sinnvolle Vergleichsdimension eröffnet, sollte man dann bei den erschöpfenden Ergebnissen... aus der extensionalen Logik stehen bleiben oder den Vergleich nicht besser mit Mitteln der Modallogik... führen?"

((9)) Außer der schon erwähnten Kritik von Alisch an Wandschneiders Behandlung des allgemeinen Problems gibt es noch eine Reihe weiterer negativer Beurteilungen dieses Teils von Wa, die dann auch eine Rolle bei der Kritik des Hauptteils spielen. Der Hauptvorwurf hierbei ist, daß die verwendeten Begriffe nicht präzisiert worden sind. Damit ist deren Verwendung in Argumenten unmöglich. So bleiben die Begriffe "Logik des Denkens", "Logik der Maschine", sowie die Art der Maschine, mit der verglichen werden soll, ungeklärt (AL((15))). Ähnlich kritisch äußern sich Ste((1)), Stu ((5)), ((15)), Rh ((13)), Sto ((5)). Besonders wichtig scheint mir in diesem Zusammenhang der Hinweis von Stolz auf die Ansicht Wandschneiders über das Arbeiten des Gehirns (Wa((39))). Nach neueren biologischen Forschungsergebnissen muß man das Gehirn eher als assoziativen Speicher, nicht jedoch als programmierten Computer ansehen (vgl. Sto ((5)) und die Literaturhinweise in Sto, Anm. 9).

((10)) Ich komme nun zur Kritik der Argumentation, die Wandschneider für die Begründung von (HT) verwendet, d.h. zur Kritik am Hauptteil von Wa. Außer Roth, dessen gesamte Kritik von einem anderen philosophischen Ansatz her kommt, der also gar nicht die Absicht hat, sich mit der Wandschneiderschen Analyse der Gödel-Theoreme auseinanderzusetzen, kritisieren alle Autoren mehr oder weniger ausführlich und übereinstimmend die gleichen Punkte dieser Analyse. Sie läßt sich entsprechend dem Aufbau von Wa in zwei Teile gliedern:

- (1) Kritik an der adäquaten Wiedergabe der beiden Gödelschen Sätze und an der Analyse ihrer Beweise.
- (2) Kritik an der Anwendung (Wa III) dessen, was Wandschneider in Wa II herausgearbeitet hat.

((11)) Da die Kritiken im wesentlichen übereinstimmen, beziehe ich mich aus Gründen der eigenen Bequemlichkeit auf Fu, da dort die Gödelschen Resultate am ausführlichsten dargestellt sind. Alle Kritiken, die überhaupt Gödel, wenn

auch noch so kurz referieren, sind klar und eindeutig und tragen damit zur Entmythologisierung dessen bei, was außerhalb der Mathematik zum Teil als "Gödel-Theoreme" rezipiert worden war. Der Darstellung Fu ((6)) - ((10)) ist nichts hinzuzufügen.

((12)) Der erste kritische Punkt ist die von Wandschneider ausführlich behandelte "Selbstreferenz" des in Fu ((8)), Formel (DLG) auftretenden Satzes G. Fuhrmann präzisiert den Begriff in der im Zusammenhang mit Gödels Beweis seines 1. Theorems (des Unentscheidbarkeitstheorems) einzig sinnvollen und relevanten Weise: G ist logisch äquivalent zu der Aussage, daß G (in S!, Hinzufügung von mir) nicht ableitbar ist (vgl. Fu ((12)), Fuhrmann schreibt "beweisbar" statt "ableitbar", ich schließe mich hier der Terminologie von El ((6)) an, der auf Lorenzen zurückgreift. In diesem Zusammenhang vgl. auch Ste, wo der Unterschied zwischen (halbformalen) Beweisen und (syntaktisch formalem) Deduzieren pointiert herausgestellt wird).

((13)) Nach Ansicht der Kritiker ist schon die Deutung der Gödelisierung als Interpretation der Ausdrücke der Sprache über S "falsch". Zumindest ist der Gebrauch des Wortes "Interpretation" in der Semantik anders festgelegt, nämlich etwa so, wie dies in Stu ((9)) dargestellt ist. Wichtig dabei ist z.B., daß den logischen Konstanten Funktionen, deren Wertevorrat Wahrheitswerte bilden, zugeordnet werden. Bei der Gödelisierung werden ihnen jedoch (arithmetische) Konstanten zugeordnet. Somit wird Stu ((10)) zufolge keine ("neue" oder "Sekundär"-) Semantik im (formalen) System S etabliert. Ähnlich explizit argumentiert Tetens (Te ((7))).

((14)) Aus Fuhrmanns Darstellung wird nun auch klar, daß "der von Gödel konstruierte Ausdruck G" (Wa((7)) keinesfalls "prinzipiell", d.h. unabhängig von allen Beweismodalitäten unbeweisbar und gleichwohl wahr [ist]" (Wa ((7)). Die Frage der Unsinnigkeit dieser Aussage erörtern explizit u.a. Rheinwald (Rhe ((8)), und Varga v. Kibed (Va ((26))).

((15)) Wandschneiders These, daß die semantische Selbstreferentialität des von Gödel über das (DLG) (Fu ((8))) (gewissermaßen) konstruierten Satzes G entscheidend in den Beweis des Unentscheidbarkeitstheorems (Fu ((9))) eingeht, wird von allen Kritikern bezweifelt, die sich mit dieser Frage überhaupt auseinandersetzen (besonders einfach dargestellt in Te ((7))). Gewiß mag die alte Lügner-Antinomie heuristisch Pate gestanden haben, als Gödel seinen Beweis fand, jedoch konkret wird - so sind sich die Kritiker einig - nur die logische Äquivalenz eines (im übrigen arithmetischen) Satzes G mit der folgenden Aussage benutzt: Eine bestimmte zweistellige Relation, in der an zweiter Stelle die Gödelzahl von G steht, besteht für keine Zahl x an erster Stelle, die ihrerseits die Gödelnummer einer (formalen) Ableitung für G ist. (Dies ist eine sprachliche Formulierung des (DLG) von Fu ((8))).

((16)) Ich komme nun zur Kritik von Wandschneiders Argumentation für die Hauptthese (HT) (s. ((4)) oben). Wand-

schneiders Argumentation ist gerade derjenigen entgegengesetzt, die üblicherweise benutzt wird, um aus Gödels Abhandlung heraus die Begrenztheit von Maschinen gegenüber dem Menschen zu begründen. So jedenfalls hat Tetens (Te ((4))) Wandschneiders Schlußweise referiert.

Im Zusammenhang damit stehen einmal die von Wandschneider behauptete semantische Selbstreferentialität des von Gödel konstruierten Satzes G (vgl. Fu ((8)), DGL) und dann eng damit zusammenhängend die Meinung Wandschneiders, der Mensch verbleibe bei seinem Beweis innerhalb der Mittel von S, benutze also keine anderen Hilfsmittel als S. bzw. eine Maschine, die S realisiert, bieten.

Beide Behauptungen werden von allen Kritikern (die überhaupt darauf eingehen), als falsch dargestellt (Be ((2)) ff, Bl ((1)) ff, Rhe ((11)), Te ((8)), die übrigen Kritiken gehen implizit ebenfalls davon aus).

Damit ist nach Meinung der Kritiker Wandschneiders Argumentation nicht haltbar.

### III

((17)) Wandschneider gliedert seine Replik (War) in zwei Teile. In I geht er auf die Kritik an seiner Auffassung des "Gödelproblems" ein, in II behandelt er die möglichen Konsequenzen für Denkmaschinen.

((18)) Im ersten Teil setzt sich Wandschneider vor allem mit der Kritik an seiner Auffassung der semantischen Selbstreferentialität von G auseinander. Dabei scheint er den Begriff der Interpretation so abzuschwächen, daß tatsächlich als Selbstreferentialität nur noch die Aussage des Diagonallemmas (DLG in Fu ((8))) übrig bleibt. Dennoch bleibt er dabei, daß es sich um einen semantischen Aspekt handle.

In War ((5)) wird dann auch die These von Wa ((7)), G sei unabhängig von S prinzipiell unbeweisbar, zu klären versucht. Gemeint sein könnte meiner Meinung nach: sobald eine Gödelisierung möglich ist und "S | - DLG" gilt (S muß gewisse Schlußmöglichkeiten erlauben, s. SMORENSKY, S. 827), ist G unentscheidbar in S. So wenigstens wäre es richtig.

((19)) In War ((10)) wird der Versuch unternommen, die in Wa ((24)) ff, ((32)) ff vorgenommene Unterscheidung zwischen dem Satz G und dem durch G ausgedrückten Sachverhalt zu verdeutlichen, indem er G ganz als zu S gehörig betrachtet, aber eine Übersetzung von G in ein S umfassendes System einführt, die erlaubt, über G und damit über den durch G ausgedrückten Sachverhalt zu sprechen.

((20)) In War ((16)) wird der "Mangel des axiomatischen Beweisverfahrens" darauf zurückgeführt, daß Gödels Ausdruck G "durch seine spezifische, selbstreferentielle Konstruktion gewissermaßen 'außer Reichweite' der S-Axiome

ist''.

((21)) In Teil II seiner Replik reduziert Wandschneider seine Argumentation aus Wa ganz wesentlich. Hier stellt er nur noch fest (War ((20)), daß aus Gödels Theoremen natürlich nicht folgt, daß "Maschinen technische Realisierungen formaler Systeme" sind. Weiterhin kann nach War ((21)) nicht gefolgert werden, daß jenes Metasystem, von dem aus Gödels Beweis geführt wird, über eine grundsätzlich andere Logik verfügt.

Schließlich betont Wandschneider noch einmal (s. Wa 29), "daß die Festlegung der Maschine auf irgendwie eingeschränkte Operationsmöglichkeiten schon eine Vorentscheidung des Mensch-Maschine-Problems und damit eine *Petitio principii* wäre". Seine eigenen Überlegungen führen Wandschneider schließlich in War ((25)) dazu, daß eine Maschine, die Gödels Theoreme beweisen soll, über eine gewisse pragmatische Kompetenz verfügen müßte, "kurz: derartige Leistungen gehen eindeutig über den Bereich des rein Formalen hinaus. Aber gehen sie damit auch über das hinaus, was einer Maschine grundsätzlich möglich ist? Der Begriff des Algorithmus muß, scheint mir, auch auf die Möglichkeit semantischer und pragmatischer Maschinenstrukturen hin durchdacht werden."

#### IV

((22)) In diesem letzten Abschnitt möchte ich Wandschneiders Ausführungen (Wa und War) sowie die verschiedenen Kritiken noch einmal aus meiner Sicht beleuchten. Um Wiederholungen zu vermeiden und die Ausführungen nicht mit Detail-Erörterungen unnötig zu dehnen, konzentriere ich mich auf die Hauptpunkte: die Analyse der Gödel-Resultate und die Argumentation für die Begründung von (HT) (s. ((4)) oben). Es würde auch zu weit führen, die Kritiken im einzelnen zu würdigen, zumal sie in der Kritik an Wandschneiders Analyse der Gödel-Resultate im wesentlichen übereinstimmen. Die unterschiedlichen Ansichten zur Hauptthese sind in verschiedenen Maschinenbegriffen begründet, wie leicht zu sehen ist (vgl. a. unten).

((23)) Beckermann und Tetens haben sich die Mühe gemacht, Wandschneiders Argumentation zur Begründung von (HT) besonders explizit zu rekonstruieren. Alle Kritiker gingen - wie sicher auch unbefangene Leser von Wa - davon aus, daß Wandschneiders herausgearbeiteter Hauptpunkt der Selbstreferentialität des von Gödel konstruierten Satzes G auch in dieser Argumentation die entscheidende Rolle spielt - warum sonst die langen Ausführungen, so fragt man sich. Doch Wandschneider selbst ist da anderer Ansicht. In War ((20)) schreibt er: "Über den Charakter von Maschinen sagen diese (die Gödel-Theoreme, Anm. d. Verf.) nichts aus, so daß von daher auch nichts über prinzipielle Grenzen von Maschinen zu erschließen ist. Ein von mir offenbar provoziertes Mißverständnis besteht in der Annahme einiger Kritiker..., als solle das eben Gesagte aus dem semantischen bzw. selbstre-

ferentiellen Charakter von G erschlossen werden, was natürlich keinen Sinn gibt."

((24)) An dieser Stelle ist der Leser vielleicht enttäuscht. Denn die triviale Schlußfolgerung, daß aus Gödels Theoremen keine Aussage über den Charakter von Maschinen und damit auch (HT) folgt, hätte keiner so langen Abhandlung wie Wa bedurft, wie aus dem folgenden (diesem und dem nächsten Absatz) klar wird. Im gleichen Absatz (War ((20)) gibt Wandschneider nämlich das Argument für die Lucas'sche These, aus den Gödeltheoremen folge die Beschränktheit für Maschinen m.E. richtig wieder:

Prämisse: Maschinen sind technische Realisierungen formaler Systeme

Satz (Gödel): Es gibt grundsätzlich Grenzen formaler Systeme

Schlußfolgerung: Es gibt grundsätzliche Grenzen für Maschinen

(Diese Argumentation ist sehr grob und ungenau, man müßte "grundsätzliche Grenzen", "formale Systeme" etc. präzisieren. Aber die Argumentationslinie dürfte auch so verständlich sein). Dann fährt Wandschneider fort: "Das Unzutreffende dieser Argumentation wird aber sofort deutlich, wenn man sich vergegenwärtigt, daß die erste Prämisse (oben: Prämisse, Anm. d. Verf.) ... jedenfalls nicht aus Gödels Theoremen folgt." - absolut korrekt und absolut trivial.

((25)) Mit der in ((24)) angegebenen Argumentation wird nun klar, daß die Verwendung des Gödel'schen Satzes als Argument für die Beschränktheit von Maschinen von der Prämisse abhängt, d.h. von der Maschinen-Definition. Wenn ich grundsätzlich der Ansicht bin, es könnte Maschinen geben, die nicht technische Realisierungen formaler Systeme (die den Voraussetzungen des 1. Gödelschen Theorems genügen) sind, so kann ich trivialerweise auch nicht mit dem Gödel'schen Resultat folgern, sie seien beschränkt. Auch diese Erkenntnis hätte keiner so aufwendigen Abhandlung bedurft, es sei denn, man hätte über den Begriff der Maschine mehr als nur gängige Spekulationen zu sagen gehabt. Nebenbei sei angemerkt, daß die verschiedenen Einstellungen der Kritiker (vgl. ((7))) zum Mensch-Maschine - Problem von der Art abhängen, wie die obige Prämisse gesehen wird.

((26)) Nun hat eine erneute Analyse der Gödelschen Resultate ihre eigene Berechtigung, dargestellt zu werden, wenn sie neue Einsichten liefert. Zunächst ist freilich darauf hinzuweisen (und die meisten Kritiker haben dies getan), daß Gödels 1. Theorem zumindest in vielen interessanten Fällen ohne Rückgriff auf das DLG (Fu((8))) bewiesen werden kann (vgl. Al, Rhe Anm. 7). Das reduziert die Bedeutung des Satzes G im DLG und seiner Struktur für das betrachtete Problem.

((27)) Damit bleibt also die Analyse der Gödelschen Abhandlung als eigentlich bedeutsamer Kern von Wa. Die hieran geübte Kritik ist so einmütig und so substantiell, daß sich eine Zusammenfassung eigentlich erübrigt. In War ((10)) wird denn auch eine der wichtigsten Thesen von Wa ((23)), ((27)) als verfehlt zurückgezogen. Sie war jedoch als ein zentraler Hinweis dafür angesehen worden, daß sich die Logik des Denkens von der der Maschine womöglich doch nicht unterscheidet.

((28)) Die Art der Selbstreferentialität von G ist in Wa nicht deutlich geworden. Schauen wir uns noch einmal das DLG an (vgl. Fu ((8))).

Es gibt einen Satz G mit  $(DLG) S \vdash G \leftrightarrow \exists y \text{bew}(\ulcorner G \urcorner, y)$ .

Aus S ist also eine logische Äquivalenz zwischen G und der rechten Formel ableitbar. Wendet man nun die Umkehrabbildung der Gödelisierung auf die rechte Seite an, so erhält man im Metasystem  $G \leftrightarrow S \not\vdash G$ . Das heißt, im Metasystem erhält man, daß G logisch äquivalent zur Aussage ist: G ist aus S nicht (effektiv) ableitbar. Diese Art der "Selbstreferentialität" ist im Gödelschen Beweis in der Tat wesentlich. Sie ist aber keinerlei semantische Selbstreferentialität im üblichen Sinne, denn die eigentliche Bedeutung von G als arithmetischem Satz braucht man nicht zu kennen. Pointiert formuliert: logische Äquivalenz liefert noch keine semantische, sondern allenfalls eine formale Selbstreferentialität (Blau spricht von syntaktischer Selbstreferentialität).

((29)) Man könnte nun meinen, logische Äquivalenz bedeute trotzdem so etwas wie inhaltliche Äquivalenz oder anders ausgedrückt: Der in einem Satz A ausgedrückte Sachverhalt ist die Äquivalenzklasse aller zu A logisch äquivalenten Sätze. Eine solche Definition ist jedoch nicht a priori vorgegeben und sprach-psychologisch zumindest zweifelhaft. Wenden wir nämlich diese Betrachtung auf einen Satz G der in ((28)) genannten Art an, so erhalten wir einerseits, daß G inhaltlich äquivalent ist zu  $S \not\vdash G$ . Andererseits ist G (im Metasystem) auch logisch äquivalent zur Konsistenz von S (s. SMORENSKY, S. 829), d.h. G würde zur Konsistenz von S inhaltlich äquivalent sein. Setzt man "inhaltlich äquivalent" gleich mit "G sagt aus", so erhält man einerseits: "G sagt seine eigene Unableitbarkeit in S aus", andererseits: "G sagt die Konsistenz von S aus". Unterdrückt man nun noch die Referenz zu S und setzt Unableitbarkeit mit Unbeweisbarkeit gleich, so erhält man jene Unklarheit, die mir das Lesen von Wa und War so schwer macht.

Des ungeachtet kann man - wenn man will - meiner Meinung nach den Begriff "Sachverhalt" konsistent wie oben definieren. Die Unklarheiten in Wa und War beruhen ja nur auf deren gerade oben dargestelltem Schritt der Unterdrückung des Bezugs auf S. In der vorliegenden Diskussion würde ich um der Klarheit willen allerdings stets nur von logischer Äquivalenz reden, denn dann vermeidet man den von Wandschneider ungewöhnlich und nicht eindeutig benutzten Begriff "Selbstreferentialität".

((30)) Alles in War deutet darauf hin, daß Wandschneider dort die Selbstreferentialität im Sinne von ((29)) meint, d.h. die logische Äquivalenz. Damit stellt er sich aber im Gegensatz zur Originalarbeit, denn in Wa ((19)) ergibt sich, daß G (s. ((28))) semantisch unfundiert ist, ein Irrtum, auf den fast alle Kritiker aufmerksam machen und der durch ((29)) oben sofort erkennbar wird.

((31)) Auch die Unterscheidung zwischen "Satz" und "Sachverhalt" bleibt in Wa trotz War unklar. Damit zusammenhängt die Ausführung in Wa ((34)), daß der Ausdruck G, als rein formales Gebilde genommen, sehr wohl beweisbar sein (könnte), wenn er nicht außerdem selbstreferentiell wäre." Als "rein formales Gebilde" ist G jedoch logisch äquivalent zur Konsistenz von S. Wäre G aus S ableitbar, so auch die Konsistenzaussage - entgegen dem zweiten Theorem von Gödel.

((32)) Die Kritik an Wandschneiders Aufsatz und an seiner Replik sollte hier enden. Weniger sein Artikel als vielmehr die vielen dazu eingegangenen Kritiken stellen einen anregenden Beitrag zur Entmythologisierung dessen dar, was von Nichtmathematikern häufig unter Gödels Resultaten verstanden worden ist. Weiter wird durch die Kritiken auch das Problem des Vergleichs von Mensch und KI-Maschine schärfer umrissen, insbesondere auf folgendes hingewiesen: Notwendig ist vor aller Diskussion eine (weder in Wa noch War angegangene) Klärung der Begriffe "Logik des Denkens" und "(Logik der) Maschine". Erst dann können die wichtigen, u.a. in Al, Ro, Te, Va angerissenen Probleme der KI vernünftig diskutiert werden. Entsprechende Aufsätze sind in dieser Zeitschrift angekündigt. Man darf auf sie gespannt sein.

### Literatur

GÖDEL 31: K. Gödel, Über formal unentscheidbare Sätze der Principia Mathematicae und verwandter Systeme I, Monatshefte f. Mathem. 38 (1931), S. 173-198

SMORENSKY: C. Smorensky, The incompleteness theorems, in: J. Barwise (ed.), Handbook of mathem. logic, North Holland Amsterdam 1977, S. 821 ff.

(Weitere Literatur bei Wa, War, und den einzelnen Kritiken.)

### Adresse

Prof. Dr. Manfred Wolff, Eberhard-Karls-Universität Tübingen, Mathematisches Institut, Auf der Morgenstelle 10, D-7400 Tübingen 1