

# **Habermas' Diskursethik**

**Christoph Lumer**

*Erschienen in:*

Zeitschrift für philosophische Forschung 51 (1997). S. 42-64.

## 1. Die Kernthesen der Diskursethik

Habermas' Diskursethik ist, metaethisch gesehen, der Versuch, eine spezielle, nämlich diskurstheoretische Form des *Kognitivismus* zu entwickeln. In seinem zentralen Aufsatz *Diskursethik - Notizen zu einem Begründungsprogramm* (=DE) <sup>1</sup> versucht Habermas - nach einer prinzipiellen Verteidigung des ethischen Kognitivismus (DE 58-64)<sup>2</sup> - zu zeigen, *wie* moralische Fragen kognitiv entschieden werden können und worin ihre Kognitivität genau besteht. Dieser Versuch wird im folgenden einer Kritik unterzogen.

Habermas spezifiziert den kognitiven Gehalt normativer Äußerungen wie folgt: Deren Kognitivität liege darin, daß man für sie argumentieren bzw. sie argumentativ kritisieren und auf diese Weise zu einem diskursiv erzielten Konsens über sie gelangen könne (vgl. DE 61). Dieser Kognitivität will Habermas sich dadurch vergewissern, daß er die "Logik" bzw. die "Regeln" solcher Diskurse und Argumentationen über moralische Normen aufklärt (DE 72 f.). Die zentrale Argumentationsregel solcher praktischen Diskurse sei das Universalisierungsprinzip U (DE 73):

U: "Jede gültige Norm [muß] der Bedingung genügen, daß die Folgen und Nebenwirkungen, die sich jeweils aus ihrer *allgemeinen* Befolgung für die Befriedigung der Interessen eines *jeden* Einzelnen (voraussichtlich) ergeben, von *allen* Betroffenen akzeptiert (und den Auswirkungen der bekannten alternativen Regelungsmöglichkeiten vorgezogen) werden können." (DE 75 f.)

U ist - entgegen Habermas' Behauptung (DE 73; 76) - *keine* Argumentationsregel. In ihm ist von Argumentationen oder Diskursen ja überhaupt nicht die Rede. Tatsächlich handelt es sich bei dem Universalisierungsprinzip um Habermas' (vorläufiges) Moralkriterium für Normen. Habermas behauptet zudem, U lasse sich prinzipiell nicht monologisch anwenden. (DE 77 f.) Auch diese Behauptung ist falsch: Sicherlich kann eine Vielzahl von Personen mehr Informationen (über Folgen, Nebenwirkungen, Präferenzen der Individuen etc.) zusammentragen als eine einzelne Person; wenn diese Informationen aber einmal zusammengetragen sind und der einsame Entscheider über hinreichend starke Nutzentheorien und große Rechenkapazitäten verfügt, dann kann er auch den je individuellen Nutzen der einzelnen Normen berechnen.

Daß Habermas den Status seines Universalisierungsprinzips U falsch einschätzt (Argumentationsregel, dialogisch), verurteilt seine grundsätzliche Position jedoch nicht zum Scheitern: U könnte immer noch in einem nichttrivialen Sinne wahr sein; und dies wäre ein

<sup>1</sup> Der Aufsatz *Diskursethik* (vgl. Literaturverzeichnis) wird im folgenden als "DE" zitiert; ein zusätzlicher Index (z. B. "DE<sup>1</sup>" oder "DE<sup>MS</sup>") gibt im Falle von Textabweichungen die Auflage an, bzw. ob es sich um das Manuskript handelt. - Dieser Aufsatz ist - trotz einiger Ergänzungen und Modifikationen (Habermas 1991b; Habermas 1992) - nach wie vor die zentrale Darstellung von Habermas' Diskursethik.

<sup>2</sup> Diese Verteidigung des Kognitivismus ist nicht besonders plausibel; aber darauf kann ich an dieser Stelle nicht eingehen.

ethisches Ergebnis oberster Wichtigkeit. Deshalb sind in den nächsten Abschnitten U und seine Begründung zu prüfen.<sup>3</sup>

Neben dem Universalisierungsgrundsatz U entwickelt Habermas noch ein zweites, endgültiges Moralprinzip, das *diskurs-ethische Prinzip D*:

D1 "Eine Norm [darf] nur dann Geltung beanspruchen, wenn alle von ihr möglicherweise Betroffenen als *Teilnehmer eines praktischen Diskurses* Einverständnis darüber erzielen (bzw. erzielen würden), daß diese Norm gilt." (DE 76.)

Das Verhältnis von U zu D bleibt in Habermas' Erläuterungen einigermaßen dunkel: D sei sparsamer als U (DE 103). In D spreche sich bereits die Grundvorstellung einer Diskursethik aus. D setze jedoch schon voraus, daß Normen begründet werden könnten; und diese Voraussetzung werde erst durch U erfüllt. (DE 76.) - Diese Erläuterungen lassen sich am besten so rekonstruieren: D ist

---

<sup>3</sup> In jüngster Zeit hat Habermas den Geltungsbereich von U mehrfach erheblich eingeschränkt:

1. Nur wenn eine Norm schon allgemein befolgt werde, sei sie gültig. Anderenfalls sei die Norm nicht zumutbar, sondern stelle schon einen supererogatorischen Anspruch dar. Erst die rechtliche Institutionalisierung könne die allgemeine Befolgung moralisch gültiger Normen sichern. (Habermas 1991b, 136; 199; Habermas 1992, 146; 148; 566 f.) - Das Problem der moralischen Überforderung hat Habermas richtig gesehen. Sein *spezieller* Lösungsvorschlag jedoch hat die inakzeptable Konsequenz, daß moralische Pflichten nicht über rechtliche hinausgehen und diesen sogar widersprechen können.

2. Habermas: Bei der Normenbegründung anhand von U könnten nur *typische*, nicht aber alle möglichen Fälle berücksichtigt werden. Deshalb gälten die nach U gültigen Normen nur *prima facie*. Ergäben sich bei der Anwendung im Einzelfall Normenkollisionen, so müsse nach einem "Prinzip der Angemessenheit" entschieden werden. (Habermas 1991b, 139 f.; Habermas, 1992, 200 f.; 266-268; 598.) Habermas führt dieses "Prinzip der Angemessenheit" nicht aus, sondern beruft sich auf Klaus Günther (vor allem: Günther 1988), der an der entscheidenden Stelle nur äußerst kurz und vage festsetzt: Im Fall von Normenkonflikten müsse betrachtet werden, "welche Norm sich [...] am besten rechtfertigen läßt" (ibid. 307). - Da die Normrechtfertigung nach Habermas' Konzeption aber anhand von U erfolgen sollte und U keine *Grade* der Rechtfertigung unterscheidet, ist der Vorschlag nicht mehr als eine Ausflucht. Dies bedeutet, daß bestimmte bekannte Probleme des Deontologismus (Normenkollisionen und kontraintuitive Prinzipienstrenge) auch in Habermas' Diskursethik nicht gelöst sind.

3. Moral, deren Prinzip ja in U ausgedrückt werde, stehe *neben* und nicht über dem Recht (Habermas 1992, 135; 137). Denn Rechtsnormen würden im Gegensatz zu moralischen Normen auch mit kontingenten, nichtmoralischen Gründen gerechtfertigt, beruhten auf Kompromissen und seien provinzieller als moralische Normen (ibid. 139; 189-194). "Je mehr sich die Moral verinnerlicht und autonom wird, desto mehr zieht sie sich in private Bereiche zurück." (Ibid. 566.) - Nach diesem Verständnis beschränkt sich die Moral auf den Mikrobereich unmittelbarer Interaktionen. Diese Konzeption widerspricht dem üblichen Moralverständnis, nach dem *alle* sozialen Verhältnisse, auch der Meso- und Makrobereich, moralisch beurteilt werden können, insbesondere auch Rechtsnormen oder ganze Rechtssysteme (z. B. wenn wir von einem "Unrechtsstaat" sprechen). Daß Rechtsnormen aufgrund von Kompromissen zwischen den politischen Parteien *entstehen*, bedeutet ja nicht, daß das Resultat nicht moralisch *beurteilt* werden könnte. Habermas verschenkt hier gewissermaßen moralisches Terrain und wird dadurch juristisch konservativ. Der entscheidende Grund für Habermas' Privatisierung der Moral ist, daß er - in Anlehnung an Kant - nur ganz bestimmte, nämlich *verallgemeinerbare* Interessen in die Moralbegründung einfließen lassen will; für die *individualistischeren* Interessen wären dann die rechtlichen Normen zuständig. Dieser Ansatz ist jedoch zu eng. In einer aufgeklärten Ethik müßten zunächst die individuellen Nutzen alternativer Regelungen unter Einbeziehung *aller* Interessen der Subjekte bestimmt werden; anschließend müßten diese Nutzenverteilungen *moralisch* bewertet werden.

deshalb sparsamer als U und spricht die Grundvorstellung der Diskursethik aus, weil D sozusagen nur einen *Spezialfall* einer allgemeinen diskursiven Konsens Theorie der Gültigkeit formuliert. Nach dieser Theorie ist genau das gültig, was in einem idealen Diskurs den Konsens aller Diskursteilnehmer finden würde (vgl. Habermas 1973a). D formuliert nur den Spezialfall der Gültigkeit, Richtigkeit von (moralischen) *Normen*; daneben gibt es noch die Gültigkeit, Wahrheit von Propositionen und die Gültigkeit, Wahrhaftigkeit von Äußerungen.<sup>4</sup> D kann man jedoch deshalb nicht unmittelbar aus der allgemeinen Diskurstheorie der Gültigkeit ableiten, weil dazu bereits vorausgesetzt werden muß, daß moralische Normen überhaupt diskursiv konsensfähig sind. Daß sie diskursiv konsensfähig sind, dies genau zeigt U. - Auf diese Rekonstruktion werde ich noch zurückkommen.

## 2. Habermas' Begründung der diskursethischen Prinzipien

Habermas entwickelt eine ausführliche Begründung des Universalisierungsprinzips U, an deren Ende U schließlich ziemlich schnell aus zwei Prämissen abgeleitet wird (DE 103). Der *größte* Teil der Argumentation ist hingegen der Begründung der ersten dieser beiden Prämissen gewidmet, nämlich der Begründung gewisser Regeln idealer Diskurse (DE 88-94; 96-102). Von diesen Regeln verwendet Habermas nachher für die Ableitung von U tatsächlich nur folgende drei - in verkürzter Version -:

3.1 Jeder darf an Diskursen teilnehmen; 3.2 jeder darf jede Behauptung aufstellen oder problematisieren oder seine Einstellungen und Wünsche äußern; 3.3 niemand darf durch Zwang an der Ausübung der in (3.1) und (3.2) festgelegten Rechte gehindert werden (DE 99).

Habermas' Begründung für diese Regeln ist zwar ausführlich, aber nicht triftig<sup>5</sup>; und die Regeln selbst halte ich für falsch.<sup>6</sup> Für die folgende Diskussion seien die Regeln 3.1 bis 3.3 jedoch Habermas als Prämissen geschenkt.

<sup>4</sup> Die Parallelisierung von Wahrheit, Richtigkeit und Wahrhaftigkeit sowie die allgemeine Diskurstheorie der Gültigkeit und die Diskurstheorie der Wahrheit sind problematisch und m. E. falsch. Hier ist jedoch nicht der Ort für eine Kritik. S. aber: Lumer 1990a, 149-158; 291-296; 313 f.

<sup>5</sup> Habermas verwendet für diese Begründung u. a. die von Apel und Kuhlmann erfundene Argumentationsfigur des "Nachweises eines performativen Widerspruchs" (DE 90-93; 105). Diese Argumentationsfigur ist aber nicht gültig; sie beweist nicht das, was sie beweisen soll. Ausführliche Kritik der Habermasschen Begründung der Diskursregeln: Lumer 1990a, 453 f.

<sup>6</sup> Die Anzahl der Diskursteilnehmer zu erhöhen (vgl. Regel 3.1) z. B. hat vor allem den Sinn, daß dadurch die Wahrscheinlichkeit, daß Begründungsfehler aufgedeckt werden, erhöht wird. Die Erhöhung der Teilnehmerzahl erschwert andererseits den Diskurs ziemlich schnell, erzeugt also hohe Kosten. Diese beiden Gesichtspunkte müßten in einer sinnvollen Diskursregel gegeneinander abgewogen werden. (Vgl. Lumer 1990a, 310-313.) Die Regel 3.1 jedoch, daß *jeder* am Diskurs teilnehmen darf, läßt den Kostengesichtspunkt völlig außer acht, ja läßt sogar Teilnehmer zu, die überhaupt nicht die Wahrscheinlichkeit der Fehleraufdeckung erhöhen.

Der Kern der Gesamtbegründung ist Habermas' Ableitung des Universalisierungsprinzips U; genauer muß man jedoch sagen: Habermas' Ableitungen des Universalisierungsprinzips. Denn mittlerweile habe ich sechs verschiedene Versionen dieser Ableitung gefunden, von denen allerdings zwei anscheinend inhaltlich identisch sein sollen, so daß es sich um *fünf* verschiedene Argumentationen für U handelt. Die nach den üblichen philologischen Kriterien zunächst einmal maßgebliche Version ist die Argumentation in der jüngsten, nämlich fünften Auflage des Aufsatzes *Diskursethik* (von 1991); diese ist identisch mit den Argumentationen in der zweiten bis vierten Auflage:

Mit den von Alexy vorläufig aufgestellten Regeln "verfügen wir, in Verbindung mit einem schwachen, d. h. nicht-präjudizierenden Begriff von Normenrechtfertigung, über hinreichend starke Prämissen für die Ableitung von 'U'." (DE<sup>2-5</sup> 102 f.) Es folgt die eigentliche Argumentation:

"[1] Wenn jeder, der in Argumentationen eintritt, u. a. Voraussetzungen machen muß, deren Gehalt sich in Form der Diskursregeln (3.1) bis (3.3) darstellen läßt; [2.2] und wenn wir ferner wissen, was es heißt, hypothetisch zu erörtern, ob Handlungsnormen in Kraft gesetzt werden sollen; [3] dann läßt sich jeder, der den ernsthaften Versuch unternimmt, normative Geltungsansprüche *diskursiv* einzulösen, intuitiv auf Verfahrensbedingungen ein, die einer impliziten Anerkennung von 'U' gleichkommen. [4] Aus den genannten Diskursregeln ergibt sich nämlich, [5] daß eine strittige Norm unter den Teilnehmern eines praktischen Diskurses Zustimmung nur finden kann, [6] wenn 'U' gilt, d. h. [7.2] wenn die Folgen und Nebenwirkungen, die sich aus einer *allgemeinen* Befolgung der strittigen Norm für die Befriedigung der Interessen eines *jeden Einzelnen* voraussichtlich ergeben, von allen *zwanglos* akzeptiert werden können." (DE<sup>2-5</sup> 103; Zusätze in eckigen Klammern hier und im folgenden immer von mir, C. L.)

Schon vor der Schlüssigkeitsprüfung kann festgestellt werden, daß diese Argumentation gar nicht auf die eigentliche These, U, sondern etwas anderes zielt; sie ist also eine Ignoratio elenchi. Denn das Ergebnis der "Ableitung" wäre immer nur die Verbindung von Satz 5 und 7.2, also: 'Eine strittige Norm kann unter den Teilnehmern eines praktischen Diskurses nur Zustimmung finden, wenn die Folgen und Nebenwirkungen ... von allen akzeptiert werden können.' (Vgl. Schema 1, die erste Konklusion: 5→7.2.) Dies ist aber nicht U, das Universalisierungsprinzip; dieses lautet ja, vereinfacht: U = '[10] Eine Norm ist nur dann *gültig*, [7.1] wenn die Folgen und Nebenwirkungen ... von allen akzeptiert und alternativen Regelungsmöglichkeiten vorgezogen werden können.' Habermas behauptet in Satz 6 zwar, Satz 7.2 sei identisch mit U; Satz 7.2 ist aber nur die zweite Hälfte von U - bzw. eine Vereinfachung dieser zweiten Hälfte. (Dies ist im Schema 1 gut zu sehen.) Um aus Habermas' Konklusion, also Satz 5→7.2 (bzw. 5→7.1), U selbst ableiten zu können, fehlt noch eine Prämisse, die aussagenlogisch sehr einfach zu konstruieren ist, nämlich: '[5] Eine strittige Norm kann unter den Teilnehmern eines praktischen Diskurses genau dann Zustimmung finden, wenn [10] sie gültig ist.' (Vgl. Schema 1, die von mir ergänzte Prämisse 5↔10.) Diese Behauptung ist schon eine Variante des Grundprinzips einer diskursiven Konsensstheorie der praktischen Gültigkeit; ich nenne sie deshalb kurz "*diskursives Konsensprinzip*". Dieses diskursive

Konsensprinzip hat zumindest Ähnlichkeiten mit dem diskursethischen Prinzip D. Darauf werde ich später noch zurückkommen; es wird sich herausstellen, daß die - für die Schlüssigkeit der Habermasschen Argumentation notwendige - Verwendung des diskursiven Konsensprinzips petitiös ist.

Ist die Ableitung der Implikation  $5 \rightarrow 7.2$  schlüssig? Aus dem Vorspann zu Habermas' Argumentation ergibt sich, daß die Diskursregeln (3.1) bis (3.3) und der Satz 2.2 die Prämissen der Argumentation sein sollen (vgl. Schema 1). Die Prämisse 2.2 ist eigenartig; Habermas geht an dieser Stelle davon aus, daß der Leser selbst schon die richtige Prämisse einsetzen wird, obwohl diese Prämisse nicht aus dem Zusammenhang rekonstruiert werden kann. Dies ist in einer gültigen Argumentation selbstverständlich nicht zulässig. Zum Glück hat Habermas diese Argumentation an anderer Stelle wiederholt (Habermas 1991b, 133 f.), und dort auch die Prämisse 2.2 erläutert: '[2.3] Der (schwache) Sinn [moralischer] Normenbegründungen sei, daß sie "einen Dissens über Rechte und Pflichten, d. h. über die Richtigkeit entsprechender normativer Aussagen auflösen"' (Habermas 1991b, 134; vgl. Schema 1, Prämisse 2.3).

Der Satz  $5 \rightarrow 7.2$  folgt offensichtlich nicht logisch aus den Sätzen 1 und 2.3. Folgt er bei Hinzufügung von geeigneten, *allgemein* (insbesondere dann auch für Habermas) akzeptablen Prämissen? Nach einer viel schwächeren (nicht schon wie bei Habermas als Wahrheits- und Gültigkeitstheorie fungierenden) Diskurstheorie ist der Sinn von erkenntnisorientierten Diskursen, die Wahrheit von Thesen argumentativ zu prüfen und dabei Begründungsfehler auszuschalten. Diskurse sind so geregelt, daß eine These unter den Diskursteilnehmern wahrscheinlich nur dann konsensuelle Zustimmung findet, wenn sie wahr ist. (Vgl.: Lumer 1988; Lumer 1991.) In ethischen Diskursen wird die These, daß eine strittige Norm  $n$  moralisch ist, d. h. dem Moralkriterium genügt, ebenfalls wahrscheinlich nur dann Zustimmung finden, wenn die Norm dem Moralkriterium genügt. Dies sei als Prämisse  $5^* \rightarrow 10^*$  konzediert (vgl. Schema 2, Satz  $5^* \rightarrow 10^*$ ). Um aus diesem Satz eine abgeschwächte Version von Habermas' Konklusion  $5 \rightarrow 7.2$  abzuleiten, benötigt man als Zusatzprämisse die Proposition:  $10^* \leftrightarrow 7.1$ : '[10\*] Eine Norm  $n$  ist moralisch / gültig genau dann, wenn [7.1] die Folgen [...], die sich aus der allgemeinen Befolgung von  $n$  für [...] jeden einzelnen [...] ergeben, wahrscheinlich von allen zwanglos akzeptiert (und den Auswirkungen der [...] alternativen Regelungsmöglichkeiten vorgezogen) werden können.' (Vgl. Schema 2, Satz  $10^* \leftrightarrow 7.1$ .) Diese Proposition ist aber nichts anderes als eine *Verstärkung von U*, statt der bloßen Subjunktion wird eine Bisubjunktion gefordert. (In Schema 2 wird diese Bisubjunktion als "U\*" bezeichnet.) Da Habermas aber U beweisen wollte, enthält diese rekonstruierte Argumentation also ebenfalls eine ganz klare *Petitio principii*.

Mit U begründet Habermas dann den diskursethischen Grundsatz D, das Kernstück seiner Diskursethik:

"[8] Ist nun aber gezeigt, wie der Universalisierungsgrundsatz auf dem Wege der transzendentalpragmatischen Ableitung aus Argumentationsvoraussetzungen begründet werden kann, [9] kann *die Diskursethik selbst* auf den sparsamen Grundsatz 'D' [D =  $10 \rightarrow 11$ ] gebracht werden, [10] daß nur die Normen Geltung beanspruchen dürfen, [11] die

die Zustimmung aller Betroffenen als Teilnehmer eines praktischen Diskurses finden (oder finden könnten)." (DE 103.)

Habermas versteht den Übergang zum Diskursprinzip anscheinend als Ableitung von D aus U oder aus dem Satz 5→7.2. Ich kann hier jedoch nicht den Schimmer einer Ableitung entdecken. Ich habe zwar eine Fülle von Spekulationen darüber angestellt, wie Habermas zu D gekommen sein mag; kein einziger der dabei vermuteten Wege ist jedoch korrekt. Deshalb ist es müßig, diese Spekulationen hier zu präsentieren.

Wie bereits dargelegt, habe ich außer der gerade analysierten bei Habermas noch vier andere Argumentationen für das Universalisierungsprinzip gefunden. 1. Der Aufsatz *Diskursethik* ist im Herbst 1983 in der ersten Auflage erschienen. Im März 1982 hatte ich von Habermas ein Manuskript dieses Aufsatzes erhalten (=DE<sup>MS</sup>), in dem die zentrale Begründung von U noch etwas anders konzipiert war. Ursprünglich, in dem Manuskript, hatte Habermas vorgesehen, U *alleine* aus den Diskursregeln (3.1) bis (3.3) (= Satz 1) abzuleiten, also ohne die Prämisse 2.2 bzw. 2.3. (Vgl. Schema 1.) Wenn die obige Analyse stimmt, daß U nicht aus den Diskursregeln und der Prämisse 2 zusammen folgt, dann kann U erst recht nicht aus den Diskursregeln alleine folgen. Diese ursprüngliche Argumentation war also nicht gültig. Interessant daran ist hier nur, daß Habermas' anfängliche Idee war, U aus den Diskursregeln *alleine* abzuleiten. Damit war diese Argumentation rein transzendentalpragmatisch im Sinne Apels: Die Quelle der Moral sollten alleine die - angeblich bei jeder Argumentation vorausgesetzten - Diskursregeln sein.

2. Habermas scheint dann die fehlende Schlüssigkeit seiner Argumentation entdeckt zu haben. In der ersten Veröffentlichung des Aufsatzes *Diskursethik* hat er in die fragliche Argumentation eine zusätzliche Prämisse eingebaut. Diese lautet:

"[2.1] Mit gerechtfertigten Normen [verbinden] wir den Sinn [...], daß diese gesellschaftliche Materien im gemeinsamen Interesse der möglicherweise Betroffenen regeln." (DE<sup>1</sup> 103)

(Vgl. Schema 2, Satz 2.1.) Diese Prämisse nennt er die "Idee der Rechtfertigung von Normen" (DE<sup>1</sup> 102). In der obigen Rekonstruktion der Habermasschen Argumentation aus den späteren Auflagen (DE<sup>2-5</sup>) hatte ich kritisiert, daß zur Ableitung des Satzes 5→7.2, bzw. einer abgeschwächten Version davon (5\*→7.1), die Prämisse 10\*↔7.1 erforderlich wäre: "[10\*] Eine Norm n ist gültig [etc.] genau dann, wenn [7.1] die Folgen [...], die sich aus der allgemeinen Befolgung von n für [...] jeden einzelnen [...] ergeben, wahrscheinlich von allen zwanglos akzeptiert (und den Auswirkungen der [...] alternativen Regelungsmöglichkeiten vorgezogen) werden können." (vgl. Schema 2: Satz 10\*↔7.1). Die Prämisse 2.1 scheint nun auch in etwa dies zu besagen (vgl. Schema 2: 2.1 , 10\*↔7.1). Nur führte die Verwendung dieser Prämisse zu einer *Petitio principii*, weil sie - wie bereits erläutert - schon eine Verstärkung des Universalisierungsprinzips darstellt (vgl. Schema 2: 10\*↔7.1 = U\*).

3. Diesen Fehler hat Habermas nachher bemerkt und die Prämisse 2.1 ab der zweiten Auflage (1985) des Aufsatzes *Diskursethik* ohne jeden Hinweis auf die Veränderung zu der oben zitierten Version 2.2 abgeschwächt (vgl. Schema 1, Satz 2.2), die er als "schwachen, d. h. nicht-

präjudizierenden Begriff von Normenrechtfertigung" bezeichnet (DE<sup>2-5</sup> 102 f.). Später (1986) jedoch kommentierte Habermas diese Korrektur so:

"Allerdings darf die Idee der Rechtfertigung von Normen nicht zu stark sein und nicht schon das in die Prämisse einführen, worauf doch erst geschlossen werden soll: daß gerechtfertigte Normen die Zustimmung aller Betroffenen müßten finden können. Dieser Fehler ist mir unterlaufen in J. Habermas (1983) [=DE<sup>1</sup>], 102 f.; er wurde in der zweiten Auflage (1985) [=DE<sup>2</sup>] korrigiert." (Habermas 1986, 13; ähnlich: Habermas 1991b, 134.)

Diese korrigierte Version ist im Laufe der jetzigen Darstellung ja als erstes präsentiert und kritisiert worden: Wegen der Abschwächung der Prämisse (von 2.1 zu 2.2 bzw. 2.3) ist die Argumentation zwar nicht mehr petitiös, nun aber wieder nicht mehr schlüssig (s. Schema 1).

4. Inzwischen scheinen Habermas aber wieder Zweifel gekommen zu sein, ob auch diese - bis zur fünften Auflage noch als "Ableitung" bezeichnete (DE<sup>1-5</sup> 103) - Argumentation so gültig ist. Jedenfalls titulierte er sie in den Erläuterungen zur Diskursethik nur noch als "Begründungsprogramm" (Habermas 1991b, 134); und er verweist auf einen Aufsatz von William Rehg Discourse and the Moral Point of View. Deriving a Dialogical Principle of Universalization (Rehg 1991), in dem ein "detaillierter Vorschlag zur Durchführung dieses Begründungsprogramms" gemacht werde (Habermas 1991b, 134). Nach den Ankündigungen konnte man erwarten, Rehg habe etwas mehr Kenntnis von formaler Logik und seine Argumentation sei entsprechend aufgebaut. Leider ist dies nicht so: Seine Argumentation ist zwar erheblich länger, aber ziemlich verworren, nicht schlüssig, petitiös und hat auch nicht U als Konklusion. Dies kann ich aus Platzgründen hier nicht im Detail darlegen.<sup>7</sup>

5. In seinem jüngsten großen Werk, Faktizität und Geltung (Habermas 1992), verweist Habermas zur Begründung von U an der zentralen Stelle nur noch auf den gerade diskutierten Aufsatz Rehgs und die zugehörige Ausarbeitung in Rehgs Dissertation (von 1991) (Habermas 1992, 140, Anm 32); seine eigene (oben zuerst diskutierte) Begründung aus dem Aufsatz Diskursethik erwähnt er dort mit keiner Silbe mehr.<sup>8</sup> An der gleichen Stelle führt Habermas zudem aus: "In moralischen Begründungsdiskursen nimmt dann das Diskursprinzip die Form eines Universalisierungsgrundsatzes an." (ibid. 140); und: "Das Moralprinzip [=U] ergibt sich nämlich erst aus einer Spezifizierung des allgemeinen Diskursprinzips für solche Handlungsnormen, die *allein* unter dem Gesichtspunkt gleichmäßiger Interessenberücksichtigung gerechtfertigt werden können" (ibid. 139). Soll dies nun heißen, U sei nur ein Spezialfall des Diskursprinzips D und müsse deshalb aus

<sup>7</sup> Ein Beispiel: Rehg führt u. a. eine wesentliche neue Prämisse ein: Man könne eine Norm genau dann mit guten Gründen akzeptieren, wenn man rationalerweise davon überzeugt sei, daß die generelle Befolgung der Norm "die Handlungen gemäß einem Wert koordiniert, der für alle Priorität hat" (Rehg 1991, 41). - Diese Prämisse ist *falsch*; sie setzt *jede* begründete Akzeptanz von Normen mit der *moralischen* Akzeptanz gleich. Zum anderen ist die Verwendung dieser Prämisse *petitiös*: Sie enthält schon ein bestimmtes Kriterium für moralische Akzeptabilität, nämlich daß moralische Normen in einer gewissen Weise *für alle optimal* sein müssen. Dies ist aber das Moralverständnis, das in U ausgedrückt wird.

<sup>8</sup> An weniger prominenter Stelle (Habermas 1992, 282) verweist er jedoch wieder auf seine eigene Begründung aus der ersten Auflage des Aufsatzes Diskursethik.

diesem abgeleitet werden, also genau in der umgekehrten Begründungsrichtung wie in *Diskursethik*? Dann müßte Habermas aber zum einen zuerst das Diskursprinzip D begründen und könnte es auf keinen Fall aus U ableiten; zum anderen sehe ich nicht, wie solch eine nichtzirkuläre Begründung aussehen könnte, vor allem nicht, wenn U bereits eine Geltungsvoraussetzung von D sein soll (DE 76; s. o., Abschn. 1).

Diese ganze Entwicklung der Habermasschen Argumentationen für U und D weist zwei durchgehende Tendenzen auf: Zum einen nimmt Habermas seinen Begründungsanspruch immer weiter zurück, von einer "Ableitung" von U (DE 103) bis hin zu einem bloßen "Begründungsprogramm" (Habermas 1991b, 134). Zum anderen fügt er in diese Argumentation immer mehr und immer stärkere Prämissen ein, die unsere moralischen *Intuitionen* beschreiben sollen. Im gleichen Maße verlieren die *Diskursregeln* als Prämissen an Gewicht. D. h., Habermas verlagert dabei auch seinen Begründungsansatz von der Transzendentalpragmatik zum Intuitionismus. In den *Erläuterungen zur Diskursethik* beruft er sich auch schon auf Rawls' Methode der Intuitionenrekonstruktion (Habermas 1991b, 125 f.) - allerdings weiterhin auch auf die Methode der Transzendentalpragmatik (ibid. 185 f.). Damit gibt er den starken, universellen Begründungsanspruch tendenziell auch methodisch auf.

Als Fazit der bisherigen Diskussion kann man auf jeden Fall festhalten: Habermas verfügt über keine tragfähige Begründung für U und D. Dabei sind dies *die* Kernthesen der Diskursethik. Und anscheinend zweifelt er selbst daran, daß er - oder einer seiner Schüler, wie etwa Rehg - diese Begründung schon geliefert hätte. Trotzdem hält Habermas an diesen Kernthesen fest, auch ohne Begründung. Und dies ist wohl eher aus seinem *politischen* Ideal zu verstehen als aus philosophischen Erwägungen.

### 3. Kritik des Universalisierungsprinzips U

Daraus, daß Habermas bislang noch keine tragfähige Begründung für U und D geliefert hat, folgt natürlich nicht, daß diese Prinzipien falsch sind. Ob sie es sind, soll nun untersucht werden.

Habermas' Universalisierungsprinzip lautet:

U: "Jede gültige Norm [muß] der Bedingung genügen, daß die Folgen und Nebenwirkungen, die sich jeweils aus ihrer *allgemeinen* Befolgung für die Befriedigung der Interessen eines *jeden* Einzelnen (voraussichtlich) ergeben, von *allen* Betroffenen akzeptiert (und den Auswirkungen der bekannten alternativen Regelungsmöglichkeiten vorgezogen) werden können." (DE<sup>1-5</sup> 75 f.; nahezu identische Formulierung: Habermas 1991b, 134.)<sup>9</sup>

<sup>9</sup> An weniger prominenten Stellen finden sich inhaltlich abweichende Formulierungen: "Jede gültige Norm muß der Bedingung genügen, daß die Folgen und Nebenwirkungen, die sich aus ihrer *allgemeinen* Befolgung für die Befriedigung der Interessen *jedes* Einzelnen voraussichtlich ergeben, von *allen* Betroffenen zwanglos akzeptiert werden können." (Habermas 1984, 219; ähnlich: Habermas 1980, 532.) Die inhaltliche Abweichung besteht darin, daß Habermas hier nur *Akzeptabilität* und nicht *Optimalität* fordert. - Formale Einwände gegen diese Abschwächung liegen sofort nahe: "Akzeptabilität" ist so *vage*, daß U dadurch leer wird; in einem

Dieses Prinzip ist ziemlich unklar und mehrdeutig. Habermas diskutiert auch kein einziges Beispiel, bei dem er eine Moralnorm mit U begründet; er lehnt dies sogar explizit ab: Dies sei Sache der Beteiligten (Habermas 1985, 225 f.). Es fragt sich allerdings, in welchem Forum die "Beteiligten" den Begründungsdiskurs führen sollen, wenn nicht vornehmlich auch in der Moralphilosophie?!

a) Die erste Deutung des Universalisierungsprinzips ist, daß U für Normen, die moralisch begründet oder gültig sein sollen, fordert, daß die soziale Geltung dieser Norm für jeden einzelnen Betroffenen optimal ist; diese allgemeine Optimalität nenne ich "*Paretoperfektheit*":

Ua Eine Norm ist genau dann moralisch gültig, wenn ihre allgemeine Befolgung für alle Betroffenen optimal ist, d. h. wenn die Folgen dieser Befolgung für alle Betroffenen besser sind als die Folgen der Befolgung alternativer Regelungen.

Wenn eine Norm paretoperfekt ist, wird ihr - innerhalb säkularer Ethiken - wohl niemand die Moralität absprechen. Denn paretoperfekte Normen sind z. B. nach so gegensätzlichen Moralkriterien wie dem utilitaristischen oder dem Rawlsschen Maximinkriterium oder dem Leximinkriterium moralisch. Dies liegt einfach daran, daß sich bei Paretoperfektheit für niemanden mehr eine Verbesserung erzielen läßt, selbst nicht auf Kosten anderer; oder umgekehrt: Jede Abweichung von einer paretoperfekten Norm wäre immer für irgendjemanden eine Verschlechterung und für niemanden eine Verbesserung. Nur gibt es wegen tatsächlicher Interessenkonflikte keine paretoperfekten Normen; für die verschiedenen Individuen ist jeweils die allgemeine Befolgung unterschiedlicher Normen optimal. Und deshalb ist ein Moralkriterium, das alle nicht paretoperfekten Normen für unmoralisch oder unbegründet erklärt, viel zu stark und unbrauchbar.

b) Die zweite Deutung von U ist, daß die anvisierten je individuellen Optimalitäten gar keine *prudentiellen*, sondern *moralische* Optimalitäten sein sollen:

Ub Eine Norm ist genau dann moralisch gültig, wenn ihre Befolgung  $\alpha$ ) nach den faktischen Maßstäben aller Betroffenen,  $\beta$ ) nach einem bestimmten vorausgesetzten moralischen Wünschbarkeitskriterium,  $\gamma$ ) nach den verallgemeinerungsfähigen Interessen der Individuen moralisch optimal ist.

Zu  $\alpha$ ): Wenn sich die moralischen Optimalitäten aus den *faktisch* je individuell akzeptierten moralischen Wünschbarkeitskriterien ergeben sollen, dann ist realistischerweise wieder keine Einigung zu erwarten, daß auch nur irgendeine Norm in allen ihren Details nach allen diesen moralischen Wünschbarkeitskriterien besser ist als die jeweiligen Alternativen: Der Streit zwischen verschiedenen Moralauffassungen um die richtige moralische Bewertung ist uns nur allzu geläufig. Dann gäbe es aber keine einzige moralisch gültige Norm. Zu  $\beta$ ): Alternativ könnte sich die moralische Optimalität der Norm nach einem *bestimmten*, von Habermas unterstellten moralischen Wünschbarkeitskriterium ergeben. Daß Habermas fordert, man müsse sich bei der Optimalitätsfeststellung in alle anderen hineinversetzen (DE 76 f.), zeigt ja, daß Habermas in dieser

---

gewissen Sinn von "akzeptabel" sind immer *viele* alternative Normen für alle Betroffenen akzeptabel, so daß U nicht genügend selektiv ist.

Beziehung gewisse Vorstellungen hat. U würde bei dieser Interpretation aber nicht mehr das volle Moralkriterium für Normen liefern; U würde nicht *mehr* besagen als: Normen sind moralisch gültig genau dann, wenn sie moralisch optimal sind (wogegen als solches nichts einzuwenden ist) - ohne aber die 'moralische Wünschbarkeit' zu definieren. Da Habermas bislang keinerlei Definitionsversuche in dieser Richtung unternommen hat, hätte er also die Hauptarbeit bei der Entwicklung eines Moralkriteriums noch vor sich. Zu  $\gamma$ ): Neuerdings schreibt Habermas, in moralischen Diskursen ginge es nur um "verallgemeinerungsfähige Interessen" (Habermas 1992, 190; 200). Welche Art von Interessen dies sein sollen, bleibt unklar. Sie sollen anscheinend verschieden sein von den zufälligen Interessen und unabhängig von besonderen kulturellen Prägungen (ibid. 202); eine nähere Spezifizierung gibt Habermas jedoch nicht, so daß der Gehalt der Deutung U $\gamma$  genauso offen bleibt wie der der Deutung U $\beta$ . Eine Spezifizierung der "verallgemeinerungsfähigen Interessen" führt zudem zu folgendem Dilemma: Entweder die "verallgemeinerungsfähigen Interessen" werden von einer *abstrakten* Person aus konstruiert und sind deshalb bei allen Individuen gleich; dann geht die von Habermas immer wieder - gegen Kant - reklamierte (z. B. DE 75 f.; 87) Intersubjektivität und Individualität der zu berücksichtigenden Interessen verloren, und diese Interessen hätten mit den realen Personen wenig bis gar nichts zu tun. Oder die "verallgemeinerungsfähigen Interessen" werden jeweils von den *realen* Personen aus konstruiert und sind deshalb intersubjektiv verschieden; dann gibt es wieder keine Norm, die für *alle* optimal, also paretopfekt ist. - Alle Deutungen von U, in denen nur eine *moralische* Optimalität und nicht eine prudentiell definierte, also alle Interessen der Person einschließende Optimalität gefordert wird, werfen zudem das Motivationsproblem auf: Wenn moralisch gültige, begründete Normen nur einem winzigen Bruchteil der Interessen der Subjekte gerecht werden, warum sollten die Subjekte dann diesen Normen zustimmen, ihre soziale Geltung fördern und sich ihnen unterwerfen?

c) Habermas spricht häufig davon, daß gültige moralische Normen "im gleichmäßigen Interesse" aller (möglicherweise) Betroffenen und "gleichermaßen gut" für alle Betroffenen seien (DE 76; Habermas 1991b, 130; 134; 164; 169 f.; 176; Habermas 1992, 133; 200). Wenn man diese Idee mit der Optimalitätsidee kombiniert, ergibt sich als weitere Deutung von U:

U $c$  Eine Norm  $n$  ist genau dann moralisch gültig, wenn gilt: Die Wünschbarkeit von  $n$  ist für alle von  $n$  Betroffenen gleich; und - falls es mehrere solcher Normen gibt - unter allen gleichermaßen wünschbaren Normen ist  $n$  die beste.

Angenommen, das Problem des intersubjektiven Nutzenvergleichs sei gelöst und die 'Wünschbarkeit von  $n$  für ein Subjekt  $s$ ' sei definiert als die Differenz der Erwartungswünschbarkeit für  $s$  der Welten, in denen  $n$  sozial gilt, und der 'natürlichen' Welten, in denen weder  $n$  noch eine Alternative zu  $n$  sozial gilt, dann wäre U $c$  moralisch kontraintuitiv: Die Wünschbarkeit eines gleichmäßigen Tötungsverbots beispielsweise ist intersubjektiv verschieden in Abhängigkeit davon, wie wahrscheinlich es ist, daß jemand in den 'natürlichen' Welten getötet werden würde. Diese Wahrscheinlichkeiten sind bei körperlich starken und gewandten, intelligenten Personen geringer als bei schwachen und wenig intelligenten Menschen. Die Wünschbarkeit des gleichmäßigen Tötungsverbots ist für jene also geringer als für diese. Um eine intersubjektiv gleiche

Wünschbarkeit zu erzielen, müßte die Norm so geändert werden, daß die Tötung von schwachen und weniger intelligenten Menschen gelegentlich erlaubt ist. Allgemeiner besteht das Problem der so interpretierten Egalitätsforderung darin, daß nach ihrer natürlichen Wünschbarkeitsunterschiede konserviert würden, während unsere intuitive Moral hier einen gewissen Ausgleich herzustellen versucht.

d) Es gibt jedoch noch eine weitere Deutungsmöglichkeit des Universalisierungsprinzips, nach der U und D quasi zusammenfallen: Die Formulierungen von U in *Diskursethik* enden nicht mit 'daß die Folgen der Normbefolgung für alle Betroffenen "akzeptabel sind"', sondern mit 'daß diese Folgen von allen Betroffenen "akzeptiert werden können"' (DE 75 f.; 103). Die bloße *Möglichkeit* der Akzeptanz zu fordern ist zunächst einmal eine völlig inakzeptable Abschwächung des Moralkriteriums: Was für mögliche Welten kann man sich nicht alles vorstellen, in denen die Betroffenen die grausamsten und absurdesten Normen einvernehmlich akzeptieren?! In Analogie zu Habermas' Ausführungen an anderen Stellen (z. B. Habermas 1973a, 219; 239 f.) mag die possibilistische Formulierung mit "können" auch als verunglückte Formulierung eines potentialen Konditionalsatzes verstanden werden:

Ud Eine Norm ist genau dann moralisch gültig, wenn gilt: Wenn die von der Normbefolgung Betroffenen über die Normgeltung einen praktischen Diskurs führen würden, dann würden sie nachher alle die allgemeine Befolgung dieser Norm akzeptieren und der Geltung alternativer Normen vorziehen.

Nach dieser Deutung liegt U schon sehr nahe bei D; und D liefert erst die korrekte Interpretation von U. Wenn U jedoch (quasi) identisch ist mit D, dann kann U nicht mehr die ihm zugedachte Funktion erfüllen: U sollte ja eine Voraussetzung von D sein, nämlich belegen, daß moralische Fragen überhaupt diskursiv konsensfähig sind (s. o., Abschn. 1). Ohne ein *zusätzliches* U steht also auch D auf tönernen Füßen.

#### 4. Kritik des Diskursprinzips D

Vor der substantiellen Kritik des Diskursprinzips ist zunächst leider wieder einige Interpretationsarbeit erforderlich. In den diskutierten Schriften finden sich (mindestens) fünf inhaltlich verschiedene Formulierungen "des" Diskursprinzips:

D1 "Eine Norm [darf] nur dann Geltung beanspruchen, wenn alle von ihr möglicherweise Betroffenen als *Teilnehmer eines praktischen Diskurses* Einverständnis darüber erzielen (bzw. erzielen würden), daß diese Norm gilt." (DE 76.)

D2 "Nur die Normen [dürfen] Geltung beanspruchen [...], die die Zustimmung aller Betroffenen als Teilnehmer eines praktischen Diskurses finden (oder finden könnten)." (DE 103; ähnlich: Habermas 1991b, 154 f.)

D3 "Gültig sind genau die Handlungsnormen, denen alle möglicherweise Betroffenen als Teilnehmer an rationalen Diskursen zustimmen könnten." (Habermas 1992, 138.)

- D4 "Eine Handlungsnorm hat Gültigkeit nur dann, wenn alle von ihr (und den Nebenwirkungen ihrer Anwendung) möglicherweise Betroffenen als Teilnehmer eines praktischen Diskurses ein (rational motiviertes) Einverständnis darüber erzielen würden, daß die Norm in Kraft treten (bleiben), also soziale Geltung erhalten (behalten) soll." (Habermas 1980, 531.)
- D5 "Jede gültige Norm müßte die Zustimmung aller Betroffenen, wenn diese nur an einem praktischen Diskurs teilnehmen würden, finden können." (Habermas 1984, 219.)
- D2 und D3 enthalten wieder das viel zu schwache und deshalb unbrauchbare "können" bzw. "könnten"; D1 und D4 hingegen enthalten u. a. das Konsequens eines potentialen Konditionalsatzes: die Betroffenen "*würden* ein Einverständnis erzielen". D5 schließlich ist ein vollständiger potentialer Konditionalsatz mit dem Antezedens: "wenn die Betroffenen an einem praktischen Diskurs [über die Norm] teilnehmen würden". Vermutlich ist immer dies gemeint. - D1, D2, D4 und D5 sind einfache Subjunktionen - "eine Norm darf *nur* Geltung beanspruchen, wenn" etc. -; D3 hingegen ist eine Bisubjunktion - "Normen sind *genau* dann gültig, wenn". Als Moralkriterium ist selbstverständlich nur eine Bisubjunktion zu gebrauchen. - In D1 und D2 wird der unklare Ausdruck "darf Geltung beanspruchen" verwendet; verständlicher sind die Formulierungen aus D3, D4 und D5 "ist gültig" bzw. "hat Gültigkeit".

Diese Mehrdeutigkeiten waren schnell zu entscheiden. Die letzte relevante Mehrdeutigkeit hingegen ist nicht aufzulösen, sondern ergibt sich aus einem fundamentalen Problem in Habermas' Ethik. Diese Mehrdeutigkeit betrifft die Natur des gemeinten Gesprächs: Als mögliche Deutungen bieten sich zwei Idealformen von Gesprächen an: a) Diskurse und b) (Vertrags-)Verhandlungen. Zu a): *Diskurse* sind argumentative Gespräche mit dem Ziel, durch Begründungen und Kritiken die Wahrheit einer These kooperativ zu prüfen und sich so über die Wahrheit zu vergewissern. Die Beteiligung anderer Personen hat dabei den Sinn, mögliche Begründungsfehler besser auszuschließen - nach dem Prinzip "vier Augen sehen mehr als zwei" - und damit die Fehlerwahrscheinlichkeit in den Begründungen und auch der These selbst zu verringern. Dazu müssen die hinzuzuziehenden Personen argumentativ qualifiziert sein und über die im fraglichen Wissensbereich nötigen Hintergrundinformationen verfügen. (Lumer 1990a, 304-314.) - Für diese Deutung spricht, daß Habermas selbst die gemeinten Gespräche durchgängig als "Diskurse" bezeichnet und daß er in D1 auch die *These* nennt, um deren Prüfung es dabei gehen könnte, nämlich: daß eine bestimmte Norm gilt, also moralisch gültig ist. Um differierende Diskursergebnisse auszuschließen, die sich daraus ergeben könnten, daß die Betroffenen unter verschiedenen Umständen und zu unterschiedlichen Zeiten diskutieren, müßte zusätzlich spezifiziert werden, daß die Betroffenen einen *idealen* Diskurs führen unter Bedingungen, wie Habermas sie in *Wahrheitstheorien* beschrieben hat (Habermas 1973a, 255 f.). Zu b): (Vertrags-)Verhandlungen, wie sie von der kooperativen Spieltheorie modelliert werden, sind Gespräche mit dem Ziel, zwischen den Gesprächspartnern Handlungen zu koordinieren; bei erfolgreichen Verhandlungen vereinbaren sie dann schließlich, wer von ihnen welche Handlung ausführen wird. Die Beteiligung anderer Personen ist deshalb erforderlich, weil die isolierte Ausführung der möglicherweise vereinbarten Handlungen für den Handelnden normalerweise suboptimal ist und deshalb nicht zustandekäme, während die allseitige Ausführung der vereinbarten Handlungen normalerweise

paretosuperior ist gegenüber den Alternativen, die die Beteiligten individuell realisieren könnten. - Für diese Deutung spricht, daß an den anvisierten Gesprächen genau die von der Norm *Betroffenen* teilnehmen sollen (und nicht beliebige kompetente Fachleute), daß Habermas in D2, D3 und D5 davon spricht, daß die Betroffenen der *Norm selbst* - und damit ist in diesem Fall wohl eine Handlungsweise gemeint - zustimmen sollen (und nicht einer *These* über sie) und vor allem daß nach D4 das Ergebnis des Gesprächs eine Art konsensueller legislativer Akt sein soll, mit dem die soziale Geltung der Norm gesichert würde.

Keine dieser beiden Deutungen führt zu einem sinnvollen Moralprinzip: Bei der Deutungsalternative a wird zunächst einmal eine These mit klaren Gültigkeitsbedingungen benötigt. Denn Diskurse dienen ja der kooperativen Prüfung der Wahrheit (bzw. in der Habermasschen Verallgemeinerung: der Gültigkeit) einer These. Allgemeiner: "Konsens" bedeutet bei einem Diskurs nicht die Abstimmung von Handlungsabsichten, sondern Einhelligkeit der Ansichten, Meinungen; und dies bedeutet, daß alle Teilnehmer von derselben These überzeugt sind. Ohne These mit klaren Gültigkeitsbedingungen, über deren Erfüllung im Diskurs gestritten werden kann, gäbe es also nichts, *worüber* überhaupt ein diskursiver Konsens zustandekommen könnte. Diese These ist in Anlehnung an D1 zunächst einmal: 'Die Norm ist moralisch gültig.' Die Alternative a läuft dann auf ein *diskursives Konsensprinzip* der Moral hinaus:

Da: *Diskursives Konsensprinzip der Moral*: Eine Norm ist moralisch gültig genau dann, wenn in einem (idealen) Diskurs unter den Diskursteilnehmern Konsens über die These erzielt werden würde, daß diese Norm moralisch gültig ist.

Dieses Prinzip wurde schon bei der Ableitung von U präsupponiert (s. Schema 1, Satz 5↔10). Da anschließend D aus U abgeleitet werden sollte, ist die Begründung von D auf jeden Fall zirkulär; und U selbst ruht auf einer unbegründeten Prämisse. Kritik an *Da*: 1. *Da* ist zunächst einmal offen zirkulär: Im Konsequens des Kriteriums wird der erst zu bestimmende Begriff 'moralische Gültigkeit' schon substantiell verwendet. 2. Diese Zirkularität läßt sich dadurch beseitigen, daß man in *Da* ein *materiales* Kriterium für die moralische Gültigkeit von Normen einsetzt; dieses materiale Kriterium wäre U. Der Schluß von *Da* müßte also lauten: "... Konsens über die These erzielt werden würde, (daß diese Norm moralisch gültig ist, d. h.) daß diese Norm die Bedingungen von U erfüllt." Dann wäre aber nicht mehr *Da* das Moralkriterium, sondern U; der hypothetische Konsens im idealen Diskurs wäre vielleicht (annähernd) material äquivalent mit der Moralität und somit vielleicht ein sekundäres Moralkriterium, ein (einigermaßen sicheres) Indiz für die Moralität, er wäre aber nicht *das* (primäre) Moralkriterium für Normen. 3. Ein allgemeineres Argument führt zu dem gleichen Ergebnis: Der diskursive Konsens kann nicht das Moralkriterium sein, sondern setzt ein solches bereits voraus. Denn in Diskursen wird mit Hilfe von Argumentationen die Wahrheit von Thesen gezeigt oder bestritten - hier also der These, daß die Norm moralisch gültig, begründet ist. Solche Argumentationen setzen *materiale Wahrheitskriterien* für die These voraus und legen letztlich nur dar, daß die Bedingungen dieser Kriterien erfüllt sind (Lumer 1990a, 43-49). Diese Kriterien müssen konsensunabhängig sein; denn sonst könnte man ja nicht *innerhalb* des Diskurses, also schon *vor* einem möglichen Konsens für die fragliche These argumentieren. Die Diskursteilnehmer gelangen je individuell zu einer (begründeten) Überzeugung von der These

selbst, weil sie glauben, daß die Wahrheitsbedingungen der These erfüllt sind; die Gemeinsamkeit dieser Überzeugungen macht dann den begründeten Konsens aus. Das konsensunabhängige Kriterium für die moralische Gültigkeit, Begründetheit einer Norm ist dann das eigentliche moralische Gültigkeitskriterium für Normen.<sup>10</sup> (Ein Stück weit gibt Habermas diesem Einwand sogar recht, wenn er sagt: Ohne Begründungsregel (U) bleibe der diskursethische Kognitivismus eine bloße Behauptung (DE 76; Habermas 1991b, 133.))<sup>11</sup> 4. Ein diskursives Konsensprinzip der Moral führt auch zu einem pragmatischen Dilemma: Die Bedingungen eines idealen Diskurses sind nach Habermas' eigenen Worten (Habermas 1973a, 257) prinzipiell nie erfüllt; man kann also nie die Probe auf's Exempel machen. Entweder wir verfügen nun über ein *sekundäres Kriterium*, welche Normen denn in einem idealen Diskurs für gültig befunden werden würden; dann ist das diskursive Konsensprinzip überflüssig, und statt seiner kann das sekundäre Kriterium verwendet werden. Oder wir verfügen über kein solches sekundäres Kriterium, dann wissen wir nie, welche Normen denn moralisch sind. 5. Konsens im Diskurs erhöht zwar die Wahrscheinlichkeit, daß die für gültig gehaltene Norm auch wirklich moralisch gültig ist; sie *garantiert* diese Gültigkeit aber nicht. Deshalb ist der Konsens nach einem idealen Diskurs nicht einmal sicher *material* äquivalent mit der moralischen Gültigkeit der Norm. 6. Ein Konsens im praktischen Diskurs impliziert nur, daß eine bestimmte Meinung über die Norm wahrscheinlich wahr ist. Diese Meinung und der Konsens sind jedoch zunächst einmal rein *theoretisch* im Sinne von "akademisch": Alle stimmen zu, die Norm n wäre moralisch. Völlig offen bleibt jedoch die *praktische Relevanz* dieser Meinung und Zustimmung: Warum sollte man sich nach dieser Norm richten oder sich für ihre Geltung einsetzen? In Habermas' Ethik fehlt an dieser Stelle eine Theorie, die praktische, motivierende Gründe dafür liefern würde, so zu handeln, wie die moralisch gültige Norm fordert.

<sup>10</sup> Ausführlichere Kritik der Habermasschen Konsentstheorie der Wahrheit: Lumer 1990a, 291-296.

<sup>11</sup> Habermas hat auf einen ähnlichen Einwand (von Peters, (1991, 257-261) und von Wellmer (1986, 70-73)) erwidert: i. Es gebe keine a) Bewertungskriterien und b) Evidenzen, die dem Diskurs schon vorauslägen und nicht selbst diskursiv begründet werden müßten. ii. Weil es in substantiellen Fragen keine "letzten" Evidenzen oder "schlagende" Argumente gebe, müßten wir auf Diskurse als Verfahren rekurrieren. (Habermas 1991b, 164 f.) - Beide Argumente versuchen nur zu zeigen, daß man Diskurse benötigt; auf den Kern des Einwandes, daß innerhalb von Diskursen wieder Kriterien, insbesondere also Moralkriterien benötigt würden, gehen sie überhaupt nicht ein. Zu i.a: Auch moralische Bewertungskriterien sind mehr oder weniger gut und begründungsbedürftig. Daraus folgt aber weder, daß man bei der moralischen Bewertung keine Bewertungskriterien benötigt, noch, daß die Begründung moralischer Bewertungskriterien nicht durch individuelle Einsicht, sondern nur im Diskurs erfolgen könnte. Zu i.b: Wenn etwas durch Erfahrung evident ist, kann man es in Argumentationen (also mit verbalen Mitteln) nur schwächer begründen als durch die Erfahrung, nämlich durch erkenntnisgenetische Argumentationen, die im Prinzip nur beschreiben, wie jemand die Erfahrungserkenntnis gemacht hat (Lumer 1990a, 246-260). Wenn eine bei der moralischen Bewertung verwendete Information hingegen *keine* Evidenz, sondern begründungsbedürftig ist, gilt das gleiche wie im Fall i.a. Zu ii.: U. a. wegen der Fülle zu berücksichtigender Informationen können viele moralische Urteile nicht sicher begründet werden. Durch die Beteiligung anderer kann die Informationsbasis verbreitert werden und können Begründungsfehler aufgedeckt werden. Ob die von den anderen gelieferten Informationen wahr und relevant sind und ob es sich tatsächlich um Begründungsfehler handelt, bemißt sich aber wieder an den Gültigkeitskriterien für die These bzw. für die Informationen. Die bloße Tatsache, daß sich noch jemand anderes zu der fraglichen These äußert, ändert ja nichts an deren kognitivem Status.

Die andere Deutung des Diskursprinzips läuft auf eine Verhandlungslösung der Moral hinaus:

Db *Verhandlungsprinzip der Moral*: Eine Norm ist moralisch gültig genau dann, wenn alle von der Normgeltung Betroffenen sich in einer rationalen Verhandlung darauf einigen würden, die Norm in Kraft zu setzen (oder zu belassen) und zu befolgen.

In der kooperativen Spieltheorie wird geklärt, welchen Prinzipien eine Einigung in rationalen Verhandlungen genügen muß. Insofern liefert die verhandlungstheoretische Deutung des Diskursprinzips - im Gegensatz zur diskursiven - ein *klares* Gültigkeitskriterium für Normen. 1. Dieses Gültigkeitskriterium ist jedoch moralisch erheblich kontraintuitiv: Eine Grundbedingung für rationale Verhandlungslösungen ist, daß sich keiner der Kooperationspartner durch die Kooperation gegenüber dem Nutzenniveau, das er auch ohne Kooperation erreichen könnte, verschlechtern darf. Damit konservieren Verhandlungslösungen die ungleichen Ausgangsbedingungen und lassen keine echten Umverteilungen zu, wie sie moralisch gefordert wären. 2. Die kooperative Spieltheorie setzt externe Instanzen voraus, die die Einhaltung der Vereinbarung erzwingen können. Moralische Normen sind aber zunächst einmal keine Rechtsnormen, die durch staatliche Erzwingungsinstanzen geschützt werden. Ein einfaches Verhandlungsmodell der moralischen Gültigkeit von Normen kann also nicht erklären und begründen, wieso man moralisch gültige Normen befolgen soll.

Diese Kritiken der beiden Deutungen des Diskursprinzips sind ziemlich vernichtend. Habermas mag nun behaupten, daß er keine dieser beiden Deutungen gemeint hat, sondern etwas anderes. Ich sehe bislang jedoch keine weiteren sinnvollen Gesprächsformen, die zwischen diesen beiden Typen lägen und eine Synthese aus ihnen darstellen würden. (Sicherlich kann in Vertragsverhandlungen auch diskutiert werden; dies ist im Prinzip aber nur eine Unterbrechung der eigentlichen Verhandlung und schafft keinen prinzipiell neuen Gesprächstyp.) Wenn Habermas einen anderen Gesprächstyp, eine Synthese aus Diskurs und Vertragsverhandlung gemeint haben sollte, dann müßte er diesen Gesprächstyp zunächst einmal hinreichend explizieren. Ich habe Zweifel, ob ein derartiger Zwitter eine sinnvolle Gesprächsform wäre. In der bisherigen Formulierung ist das Diskursprinzip jedenfalls nur ein *Formelkompromiß*, der all das verbal zusammenstellt, was Habermas am Herzen liegt: einerseits Diskurs, Kognitivismus und Wahrheit, andererseits Demokratie, Machtteilhabe und politische Autonomie. Diese Seiten zu vereinbaren ist auch ein politisch-institutionelles Problem. Zu dessen Lösung trägt ein bloßer Formelkompromiß aber nichts bei.

## Literatur

- Günther, K. (1988): *Der Sinn für Angemessenheit - Anwendungsdiskurse in Moral und Recht*, Frankfurt.
- Habermas, J. (DE): *Diskursethik - Notizen zu einem Begründungsprogramm*, MS 1982. Veröffentlichungen in: ders., *Moralbewußtsein und kommunikatives Handeln*, Frankfurt <sup>1</sup>1983; <sup>2</sup>1985; <sup>3</sup>1988; <sup>4</sup>1991; <sup>5</sup>1993, 53-125.
- (1973a): *Wahrheitstheorien*, in: H. Fahrenbach (Hg.), *Wirklichkeit und Reflexion*, Pfullingen, 211-264.
  - (1980): *Replik auf Einwände*, in: ders., *Vorstudien und Ergänzungen zur Theorie des kommunikativen Handelns*, Frankfurt 1984, 475-570.
  - (1984): *Über Moralität und Sittlichkeit - Was macht eine Lebensform 'rational'?*, in: H. Schnädelbach (Hg.), *Rationalität - Philosophische Beiträge*, Frankfurt, 218-235.

- (1985): Ein Interview mit der *New Left Review*, in: ders., *Die Neue Unübersichtlichkeit, Kleine Politische Schriften V*, Frankfurt, 213-257.
  - (1986): Treffen Hegels Einwände gegen Kant auch auf die Diskursethik zu?, in: Habermas 1991a, 9-30.
  - (1991a): Erläuterungen zur Diskursethik, Frankfurt.
  - (1991b): Erläuterungen zur Diskursethik, in: Habermas 1991a, 119-226.
  - (1992): Faktizität und Geltung, Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats, Frankfurt.
- Lumer, Ch. (1988): The Disputation - A Special Type of Cooperative Argumentative Dialogue, in: *Argumentation* 2, 441-464.
- (1990a): Praktische Argumentationstheorie - Theoretische Grundlagen, praktische Begründung und Regeln wichtiger Argumentationsarten, Braunschweig.
  - (1990b): Geltung, Gültigkeit, in: H. J. Sandkühler (Hg.), *Europäische Enzyklopädie zu Philosophie und Wissenschaften*, Hamburg, Bd. 2, 258-262.
  - (1991): Argumentative Dialoge - Eine philosophische Rekonstruktion von Disputationen, in: S. Stati, E. Weigand, F. Hundsnurscher (Hg.), *Dialoganalyse III, Teil 1*, Tübingen, 353-367.
- Peters, B. (1991): *Rationalität, Recht und Gesellschaft*, Frankfurt.
- Popitz, H. (1980): *Die normative Konstruktion von Gesellschaft*, Tübingen.
- Rehg, W. (1991): Discourse and the Moral Point of View - Deriving a Dialogical Principle of Universalization, in: *Inquiry* 34, 27-48.
- Wellmer, A. (1986): *Ethik und Dialog - Elemente des moralischen Urteils bei Kant und in der Diskursethik*, Frankfurt.

**Schema 1: Ableitung von U in: DE<sup>2-5</sup>, 103**

**1** Diskursregeln (3.1) bis (3.3): Jeder darf am Diskurs teilnehmen; jeder darf jede  
*geschenkt* Behauptung aufstellen oder problematisieren oder seine Einstellungen und Wünsche äußern; niemand darf an der Ausübung dieser Rechte gehindert werden. (DE 103.)

**2.2** Wir wissen, was es heißt, hypothetisch zu erörtern, ob Handlungsnormen in Kraft  
*?* gesetzt werden sollen. (DE<sup>2-5</sup> 103.)

**2.3** Der (schwache) Sinn [moralischer] Normenbegründungen ist, daß sie "einen Dissens  
*wahr* über Rechte und Pflichten, d. h. über die Richtigkeit entsprechender normativer Aussagen auflösen". (Habermas 1991b, 134.)

*nicht schlüssig*

---

**5→7.2[&7.1]** [5] Wenn eine [...] Norm unter den Teilnehmern eines praktischen Diskurses  
*(falsch)* Zustimmung finden kann, dann [7.2] können die Folgen [...], die sich aus der all gemeinen Befolgung der [...] Norm für [...] jeden einzelnen [...] ergeben, von allen zwanglos akzeptiert werden (DE 103) [[7.1] (und den Auswirkungen der [...] alternativen Regelungsmöglichkeiten vorgezogen) werden (DE 75 f.).

*Von mir, C. L., angehängte Ergänzung:*

**5↔10** (= Diskursives Konsensprinzip [=Da]): [5] Eine [...] Norm kann unter den Teil  
*petitiös* nehmern eines praktischen Diskurses genau dann Zustimmung finden, wenn [10] sie gültig ist.  
*(falsch)*

*schlüssig*

---

**10→7.1** (= U): [10] Wenn eine Norm gültig ist, dann [7.1] können die Folgen [...], die sich  
*(falsch)* [...] aus ihrer allgemeinen Befolgung für [...] jeden einzelnen [...] ergeben, von allen Betroffenen akzeptiert (und den Auswirkungen der [...] alternativen Regelungsmöglichkeiten vorgezogen) werden. (DE 75 f.)

**Schema 2: Verbesserte Rekonstruktion der Ableitung einer abgeschwächten Version von Satz 5→7.1 in: DE<sup>2-5</sup> 103 bzw. DE<sup>1</sup> 103**

**5\*→10\*** [5\*] Wenn die These, daß die Norm n moralisch ist / gültig ist, unter den  
*wahr* Teilnehmern eines ethischen Diskurses Zustimmung findet, dann [10\*] ist n wahr-  
scheinlich moralisch / gültig. (Von mir, C. L., eingefügt.)

**10\*↔7.1** (= U\*): [10\*] Eine Norm n ist moralisch / gültig genau dann, wenn [7.1] die Folgen  
*petitiös* [...], die sich aus der allgemeinen Befolgung von n für [...] jeden einzelnen [...]  
*falsch* ergeben, wahrscheinlich von allen zwanglos akzeptiert (und den Auswirkungen der  
[...] alternativen Regelungsmöglichkeiten vorgezogen) werden können. (Von mir, C.  
L., eingefügt.)

**2.1** "[2.1] Mit gerechtfertigten Normen [verbinden] wir den Sinn [...], daß diese  
*wahr?* gesellschaftliche Materien im gemeinsamen Interesse der möglicherweise Betroffe-  
nen regeln." (DE<sup>1</sup> 103.)

*schlüssig*

---

**5\*→7.1** [5\*] Wenn die These, daß die Norm n moralisch / gültig ist, unter den Teilnehmern  
*falsch* eines ethischen Diskurses Zustimmung findet, dann [7.1] können die Folgen [...], die  
sich aus der allgemeinen Befolgung der [...] Norm für [...] jeden einzelnen [...]  
ergeben, wahrscheinlich von allen zwanglos akzeptiert (und den Auswirkungen der  
[...] alternativen Regelungsmöglichkeiten vorgezogen) werden. (Formulierung von  
mir, C. L.)