

KANTISCHER EXTERNALISMUS UND MOTIVE ZU MORALISCHEM HANDELN

von
Christoph Lumer

Zusammenfassung: (1) Kantianer möchten Moral begründen ohne Rekurs auf Motive, die diese Moral inhaltlich bestimmen (= fundativer Externalismus); gleichwohl soll diese Begründung Menschen zur Moralbefolgung bewegen (= praktische Forderung). Dies setzt voraus, 1. daß es motivationsverleihende Mechanismen gibt, die selbst inhaltlich einigermaßen neutral sind, aber moralischen Urteilen folgen, und 2. daß diese motivationsverleihenden Mechanismen auch einer externalistisch begründeten Moral motivationale Kraft verleihen. Thema des Artikels ist, ob diese beiden Bedingungen erfüllt sind. (2–7) Diverse logisch mögliche motivationsverleihende Mechanismen werden moralpsychologisch untersucht, u. a. ein Handeln aus reiner Vernunft und ein apriorisches Motiv der Achtung vor dem Sittengesetz, und bis auf einen, den Wunsch nach Selbstwertsteigerung, als empirisch inexistent oder zu eng verworfen. (8) Eine Skizze einer Psychologie des moralischen Urteils zeigt aber, daß dieser Mechanismus die zweite Bedingung nicht erfüllt. Ein praktischer fundativer Externalismus ist also aus empirischen Gründen nicht möglich.

Summary: (1) Kantians want to justify morals without relying on motives which would determine the morals' content (= foundational externalism); nonetheless this justification shall motivate subjects to comply with the morals (= practical demand). This presupposes 1. that there are mechanisms providing motivation and which are themselves fairly neutral but follow moral judgements and 2. that these mechanisms can provide motivational force even for externally justified morals. The topic of this paper is to see if these two conditions are fulfilled. (2–7) Various logically possible mechanisms of providing motivation are scrutinized psychologically, among others acting from pure reason and an *a priori* cognizable motive of respecting the laws of morals, and all but one, namely the desire to increase one's self-esteem, are rejected as empirically inexistent or being too narrow. (8) A sketch of a psychology of moral judgements however shows that this remaining mechanism does not fulfill the second condition. Therefore, a foundational externalism fulfilling the practical demand is empirically impossible.

* * *

1. Handlungstheoretische Voraussetzungen des fundativen Externalismus

Viele ethische Klassiker, z. B. Aristoteles, Hume oder (in sehr viel geringerem Maße) auch Kant, stützen ihre Ethik auf eine explizite empirische Handlungstheorie [Aristoteles, *NE* III 1–7; VI 1–11; Hume 1739/40, Buch 3; insbes. 3,

III, 3; Kant *GMS* insbes. BA 36f.; BA 63f.; BA 97f.; *KpV* insbes. A 128 bis 132].¹ Heutige Ethiker hingegen stützen ihre Ethik üblicherweise nicht durch eine explizite und elaborierte Handlungstheorie ab – obwohl sie sich oft auf wenig explizierte oder auch gänzlich implizite handlungstheoretische Annahmen verlassen, die dann häufig in keiner Weise wissenschaftlich abgesichert sind. Die Vernachlässigung der empirischen Handlungstheorie innerhalb der Ethik ist insofern problematisch, als sich jede brauchbare Ethik in wenigstens zwei Hinsichten auf starke handlungstheoretische Annahmen verlassen muß. Zum einen müssen Ethiken darauf Rücksicht nehmen, ob das von ihnen geforderte Handeln überhaupt möglich ist, insbesondere motivational möglich ist; Ethiken dürfen Menschen nicht motivational überfordern – wie dies z. B. ein radikaler Utilitarismus aber tut, der gebietet, immer das moralisch Beste zu tun. Zum anderen sollten triftige Moralbegründungen (kluge) Subjekte wenigstens ein Stück weit zur Befolgung der Moral motivieren (= *praktische Forderung für die Moralbegründung*). Deswegen setzen brauchbare Ethiken starke handlungstheoretische Annahmen über die Inhalte unserer Handlungsgründe und über unsere Entscheidungsmechanismen voraus, so daß folgende *handlungstheoretische Dependenzthese* gilt: Eine Rationalitätstheorie oder Ethik ist immer höchstens so gut wie die ihr zugrundeliegende (empirische) Handlungstheorie.

Im folgenden will ich nicht diese handlungstheoretische Dependenzthese verteidigen, sondern am Beispiel eines kantischen Externalismus zeigen, daß auch eine Ethik, die gerade keine empirischen Annahmen machen will, letztlich doch auf impliziten empirisch handlungstheoretischen, insbesondere moralpsychologischen Annahmen basiert. Und ich will zeigen, daß diese Ethik auch daran scheitert, daß diese Annahmen falsch sind. Die Stoßrichtung der folgenden Diskussion ist also vorwiegend kritisch. Die dabei verwendete *positive* Moralpsychologie ist ausführlicher dargestellt in meinem Aufsatz "Motive zu moralischem Handeln" [Lumer 2002] – dessen Kenntnis hier aber *nicht* vorausgesetzt wird. Neben der konkreten Kritik am kantischen Externalismus möchte ich aber auch indirekt ganz allgemein und im Sinne der handlungstheoretischen Dependenzthese eine Lanze für die empirische Handlungstheorie brechen, deren Bedeutung für die Ethik heutzutage viel zu wenig ernstgenommen wird.

Die Ethik, die hier kritisch diskutiert werden soll, ist näherhin der *praktische fundative Externalismus*, der durch zwei Eigenschaften charakterisiert

1. "Empirische Handlungstheorie" soll hier nur bedeuten, daß diese Handlungstheorie empirische Aussagen macht, nicht, daß diese Aussagen mit empirischen Methoden abgesichert sind.

ist. Zum einen ist er ein *fundativer Externalismus*, d. h. eine Form der Moralbegründung, die Moral nicht aus irgendwelchen Motiven zu moralischem Handeln heraus begründet, sondern extern begründet, z. B. aus reiner Vernunft.² Zum anderen ist er aber insofern *praktisch*, als er die oben bereits erwähnte *praktische* oder *Motivationsforderung* anerkennt: Triftige Moralbegründungen müssen kluge Subjekte wenigstens ein Stück weit zur Befolgung der Moral motivieren. Diese praktische Forderung besagt, daß Menschen nicht nur irgendwie zur Befolgung der externalistisch begründeten Moral motiviert werden können müssen, sondern daß die externalistische Moralbegründung als solche ein gewisses Maß an Motivation erzeugen muß. Die Erfüllung der praktischen Forderung ist unverzichtbar für die Moralbegründung, da sie das eigentlich praktische Element dieser Begründung charakterisiert, daß es nämlich bei diesen Begründungen – anders als bei theoretischen Begründungen – um mehr als einen bloßen Glauben an eine These geht [vgl. Lumer 2000, 34–44, insbes. 36–38; Lumer 1999; Lumer 1994].

Die prominentesten praktisch externalistischen Konzeptionen sind natürlich das Kantische Handeln aus reiner Vernunft, aus Achtung vor dem (moralischen) Gesetz oder aus (moralischer) Pflicht [z. B. Kant, *GMS* BA 28; 33; 36f.; 63f.; *KpV* A 56–58; 126–135; 139–146; 272; 279f.]; aktuelle Vertreter dieser Idee sind z. B. Habermas oder Korsgaard [Habermas 1991, 184; 144; 135; Habermas 1992, 202 (Kritik: Lumer 1997a); Korsgaard 1986, 143f.]. Insbesondere auch Kant erkennt die praktische Forderung insofern an, als er nicht nur möchte, daß Menschen die von ihm begründete Moral überhaupt befolgen, sondern daß sie sie aus den richtigen Gründen befolgen [z. B. Kant, *GMS* BA X; 26].

Die Anerkennung der praktischen Forderung macht den praktischen fundativen Externalismus in jedem Fall abhängig von empirischen Annahmen. Denn daß eine Moralbegründung (in der gewünschten Weise) "motiviert", heißt, daß sie eine entsprechende Motivation *verursacht*; und Verursachungs-

2. Man kann Ethiken nach verschiedenen Gesichtspunkten klassifizieren: 1. nach *fundativen* Gesichtspunkten, d. h. nach der Art der Begründung, insbesondere der Quelle, aus der die Moral geschöpft wird; 2. nach *metaethischen* Gesichtspunkten, d. h. nach der Art der Aussagen über die Bedeutung moralischer Ausdrücke; 3. nach *materialethischen* oder *kriteriologischen* Gesichtspunkten, d. h. nach dem resultierenden Moralkriterium, usw. Außer einem fundativen Externalismus gibt es noch einen metaethischen Externalismus, demzufolge in der Bedeutung moralischer Urteile nicht auf Motive zu moralischem Handeln Bezug genommen wird. Beide Theorien sind logisch voneinander unabhängig. Der fundative Externalismus ist m. E. falsch, der metaethische hingegen richtig.

beziehungen sind empirisch und (pace Kant) synthetisch aposteriori – für diese Behauptung werde ich hier aber nicht weiter argumentieren. Die konsistente Verbindung der beiden Komponenten des praktischen fundativen Externalismus setzt deshalb voraus, daß es motivationale Mechanismen gibt, die dem externalistisch begründeten moralischen Urteil folgen und ihm nachträglich motivierende Wirkung verleihen. Fundative *Internalisten*, die also Moral aus Motiven zu moralischem Handeln heraus begründen wollen, erkennen in aller Regel ebenfalls die praktische Forderung an, vertreten den fundativen Internalismus aber u. a. deshalb, weil sie diese Voraussetzung des praktischen fundativen Externalismus für nicht erfüllt halten.

Genauer geht der praktische fundative Externalismus von zwei starken empirischen *externalistischen Ausgangshypothesen* aus:

1. *Existenz von motivationsverleihenden Mechanismen*: Es gibt bestimmte motivationale Mechanismen, insbesondere moralische Motive i. e. S., die einer einmal akzeptierten Moral als solcher folgen, d. h. ihr motivationale Kraft verleihen, ohne selbst den Inhalt der Moral vorzugeben.

2. *Motivationale Zulieferfunktion externalistischer Begründungen*: Die jeweilige externalistische Moralbegründung kann solch eine Akzeptanz dieser Moral erzeugen, der bestimmte unter 1 genannte motivationale Mechanismen motivationale Kraft verleihen.

Thema dieses Artikels ist, ob diese beiden Ausgangshypothesen erfüllt sind. (Nach Korsgaards Terminologie [Korsgaard 1986, 121 f.] beschäftigt er sich also mit der *motivationalen Skepsis bezüglich praktischer Vernunft*, d. h. dem Zweifel, ob (reine) Vernunft überhaupt ein Handlungsmotiv sein kann oder allgemeiner, ob externalistische Moralbegründungen zur Befolgung der Moral motivieren können. Die *inhaltliche Skepsis bezüglich praktischer Vernunft*, ob sich ohne Rekurs auf Motive zu moralischem Handeln überhaupt bestimmte Moralprinzipien begründet auszeichnen lassen, teile ich zwar auch, wird hier aber nicht diskutiert [s. jedoch: Lumer 2000, 75 f.])

Der Rest des Artikels ist wie folgt aufgebaut: Es werden fünf Arten von prinzipiell möglichen oder von Kant angedachten motivationsverleihenden Mechanismen skizziert, also fünf Arten, wie die Forderung nach der Existenz von motivationsverleihenden Mechanismen erfüllt sein könnte. Diese fünf möglichen motivationsverleihenden Mechanismen werden jeweils einzeln kritisch betrachtet, ob sie empirisch existieren und hinreichend allgemein sind [Abschn. 2–3; 5–7]. Zwischengeschoben ist ein Abschnitt [Abschn. 4], der der Suche nach weiteren motivationsverleihenden Mechanismen gewidmet ist. Der fünfte Mechanismus existiert tatsächlich und ist hinreichend allgemein, wenn er auch nicht den Vorstellungen der Kantianer entspricht. Abschließend wird deshalb untersucht, ob die zweite Ausgangshypothese des praktischen

fundativen Externalismus erfüllt ist, ob also die externalistische Begründung solch eine Akzeptanz der Moral erzeugt, der der fünfte motivationsverleihende Mechanismus motivationale Kraft verleiht [Abschn. 8].

2. Erster motivationsverleihender Mechanismus: Handeln aus reiner praktischer Vernunft – analytisch unmöglich

2.1. Kants Handeln aus reiner praktischer Vernunft

Kants ursprüngliche Idee zur Erfüllung der praktischen Forderung ist, daß es ein Handeln aus Pflicht oder reiner praktischer Vernunft gebe, bei dem die von der reinen Vernunft ausgezeichneten Handlungen unmittelbar, ohne Vermittlung von Motiven, in die Tat umgesetzt würden: Die reine Vernunft erkenne bestimmte (moralische) Prinzipien oder Gesetze unabhängig von empirischen Neigungen und für alle Vernunftwesen gleichermaßen gültig als "objektiv notwendig" oder "als praktisch notwendig, d. i. als gut". Und die menschliche Autonomie bestehe darin, daß diese Erkenntnisse den Willen bestimmten – auch gegen empirische Neigungen. Dabei nötige die reine Vernunft den Willen. Da Menschen keinen durchweg guten Willen hätten und eben auch empirischen Neigungen unterlägen, sei diese Nötigung aber nicht immer zwingend, der Wille folge nicht immer der reinen Vernunft. [Kant, *GMS* BA 33; BA 36f.; BA 63f.; *KpV* A 56–58; A 126–128.]

2.2. Ein allgemeines, entscheidungstheoretisches Handlungsmodell

Was sagt die heutige empirische Handlungstheorie und Moralpsychologie zu dieser Idee eines Handelns aus reiner praktischer Vernunft? In der sozialpsychologischen Literatur gibt es eine Fülle von Untersuchungen zu moral-konformem oder prosozialem Verhalten. Umfassende Modelle solchen Verhaltens versuchen, die Theorie moral-konformen Handelns in allgemeine Handlungsmodelle zu integrieren, und kommen dabei zu dem naheliegenden Ergebnis: Moral-konforme Handlungen sind den gleichen Gesetzmäßigkeiten wie andere Handlungen unterworfen. Das Besondere an ihnen sind nur die ausgewählten Handlungen und vielleicht einzelne Motive. Das allgemeine Handlungsmodell, das bei solchen Erklärungen moral-konformer Handlungen verwendet wird, ist das am weitesten verbreitete handlungspsychologische Modell, nämlich ein i. w. S. entscheidungstheoretisches Handlungsmodell [Lynch/Cohen 1978; Heckhausen 1989, 301 f.]. Dessen m. E. plausibelste Interpretation ist die *Optimalitätstheorie*, der zufolge Entscheidungen in

Optimalitätsurteilen bestehen [Lumer 2000, 155–169, insbes. 155; Lumer i. V.a, Kap. 6; Lumer i. V.b]. Danach wählt der Handelnde eine Handlung, insbesondere eine moralische Handlung, weil er sie für die beste unter den in Betracht gezogenen Alternativen hält. Zur Bewertung der einzelnen Alternativen werden wiederum Bewertungen ihrer Folgen und eventuell ihrer folgenunabhängigen Eigenschaften ungefähr additiv zusammengefaßt. Dabei werden diverse Kosten und Nutzen gegeneinander abgewogen.³ Die verschiedenen Folgenbewertungen und Bewertungen der folgenunabhängigen Eigenschaften (handlungsphilosophisch auch als “Wünsche” bezeichnet) sind dann die *Motive* für oder gegen die Handlung.

Die Optimalitätsurteile basieren auf intrinsischen Bewertungen entweder von Folgen der Handlungen oder der Handlung selbst. Unter einer “*intrinsischen Bewertung*” eines Ereignisses wird dabei eine Bewertung verstanden, die in dem Sinne unbegründet ist, daß das Subjekt keinen weiteren Grund für diese Bewertung hat – außer vielleicht dem, daß dieses Ereignis ein bestimmtes Kriterium erfüllt, wobei das Subjekt dann aber über keine Gründe für dieses Kriterium verfügt. Intrinsische Bewertungen rekurren also insbesondere nicht auf Informationen über Folgen des Wertgegenstandes. Statt auf solchen kognitiven Gründen wie Informationen über die Folgen beruhen intrinsische Bewertungen vielmehr auf angeborenen oder erworbenen konativen Dispositionen, Gegenstände einer bestimmten Art gut oder schlecht zu finden. Diese intrinsischen Bewertungen entsprechen also Kants “Triebfedern”. Intrinsische Wünschbarkeiten sind die Basis für alle weiteren Bewertungen. Diverse *extrinsische Wünschbarkeiten* haben Gegenstände, insbesondere Handlungen, dadurch, daß sie intrinsisch wünschbare Folgen hervorrufen. Eigene Schmerzen werden beispielsweise als intrinsisch schlecht bewertet; eine schmerzverursachende Handlung ist dann in der Hinsicht extrinsisch schlecht, daß sie diesen Effekt hat; und sie ist in dieser Hinsicht in dem Maße extrinsisch schlecht, in dem der resultierende Schmerz intrinsisch schlecht ist. Die *Gesamtwünschbarkeit* eines Gegenstandes wird dann mehr oder weniger additiv aus seiner eventuellen intrinsischen und seinen diversen extrinsischen Wünschbarkeiten gebildet. In dem die Entscheidung ausmachenden *Optimalitätsurteil* schließlich werden die Gesamtwünschbarkeiten der aktuellen Handlungsalternativen verglichen.

3. Darüber, wie nicht sichere Handlungsfolgen in tatsächlichen Entscheidungen berücksichtigt werden, besteht heute keine Einigkeit unter den einschlägigen Forschern [Überblick: Camerer 1995]. Im vorliegenden Zusammenhang ist dieser Streit aber irrelevant. Entscheidend ist die Einigkeit über die Grundidee der Integration verschiedener Vor- und Nachteile von Alternativen.

Es gibt *ursprüngliche* intrinsische Bewertungen, d. h. Gegenstände dieses Typs werden ohne weitere Begründung, insbesondere ohne Berücksichtigung ihrer Folgen, aufgrund biologischer Prädispositionen (positiv oder negativ) bewertet. Daneben gibt es aber auch *verselbständigte* intrinsische Bewertungen, die ursprünglich einmal Gesamtbewertungen des Gegenstandes waren, woran sich das Subjekt aber nicht mehr erinnert, so daß es keine Begründung für diese Bewertung angeben kann, weshalb sie nach der Definition als intrinsisch gelten. Ursprünglich nicht intrinsische Bewertungen verselbständigen sich z. B. aufgrund von Verdrängungen oder von Autoritätsbegründungen: Insbesondere Kinder glauben Erwachsenen oft ohne weiteres, daß irgend etwas einen bestimmten Wert hat, ohne zu wissen warum; anschließend vergessen sie diese Autoritätsbegründung. [Lumer i. V.a, Kap. 9; Lumer 2000, Abschn. 3.6.] Verselbständigte intrinsische Bewertungen sind als solche nicht stabil, wenn das Subjekt über ihre Genese aufgeklärt wird. Sie sind deshalb auch nicht als Basis einer rationalen Ethik geeignet. Im folgenden interessieren aus diesem Grund unter den intrinsischen Bewertungen immer nur die ursprünglich intrinsischen.

Es gibt zwei grundlegend verschiedene Arten von ursprünglich intrinsischen Bewertungen. 1. *Zeitlich stabile intrinsische Bewertungen* sind situationsunabhängig; derselbe Gegenstand wird (bei gleicher Information über den Gegenstand) zu unterschiedlichen Zeiten immer gleich bewertet. Hedonistische Bewertungen der eigenen Gefühle z. B. sind zeitlich stabil: Daß ich morgen Hunger haben werde, bewerte ich heute schon negativ und gehe deshalb heute einkaufen. Wie an diesem Beispiel schon deutlich wird, sind solche zeitlich stabilen intrinsischen Bewertungen die Grundlage längerfristiger rationaler Planung. 2. *Affektinduzierte intrinsische motivationale Bewertungen* werden durch Affekte verursacht und treten nur während des Affektes auf; nach dem Abklingen des Affekts wird der Gegenstand wieder intrinsisch neutral bewertet [Lumer 1997b; Lumer 2000, 477–484]. So bewerten Menschen z. B. in Wut die Zerstörung oder Schädigung des Wutobjekts motivational intrinsisch positiv; wenn die Wut verraucht ist, ist auch dieser heiße Bestrafungswunsch – bis auf Nachklangeffekte – verschwunden. Affektinduzierte intrinsische Bewertungen sind die Grundlage für Affekthandlungen. Wegen ihrer zeitlichen Instabilität werden sie oft als irrational angesehen.

2.3. Die empirische und analytische Unmöglichkeit eines Handelns aus reiner praktischer Vernunft

Schon Kant fand die Idee eines Handelns aus reiner praktischer Vernunft problematisch: Wie ein Gesetz unmittelbarer Bestimmungsgrund des Willens

sein könne, sei ein für die menschliche Vernunft unauflösliches Problem [Kant, *KpV* A 128]. Nach dem gerade skizzierten allgemeinen Handlungsmodell jedenfalls ist ein Handeln aus reiner praktischer Vernunft, d. h. eine Bestimmung unseres Willens durch reine Vernunft unmöglich.⁴ Denn die die Willensbildung darstellenden Optimalitätsurteile basieren nach diesem Modell letztlich immer auf intrinsischen Bewertungen, die nicht rein kognitiv gefällt werden, sondern auf angeborenen – und nicht von der reinen Vernunft geschaffenen – konativen Dispositionen beruhen. Handlungen, die nicht auf intrinsischen Motiven, d. h. in Kants Terminologie: auf Triebfedern, beruhen, gibt es nicht.⁵

Diese Argumentation gegen ein Handeln aus reiner Vernunft ist empirisch, stützt sich auf die Optimalitätsurteiltstheorie der Entscheidung. Als Erwiderung auf diese Argumentation könnte ein Kantianer deshalb auch diese Theorie ablehnen und seine Hoffnungen auf eine "bessere" Handlungstheorie setzen. 1. Bislang ist aber noch keine halbwegs präzise (mit empirischen Handlungsgesetzen) und erklärungskräftige allgemeine empirische Handlungstheorie entwickelt worden, der zufolge ein Handeln aus reiner Vernunft möglich ist. Kant und seine Nachfolger behaupten solch eine Möglichkeit ja nur, entwickeln aber keine entsprechende Theorie. Anders gesagt: Wer be-

4. Die im folgenden skizzierte Kritik am Handeln aus reiner Vernunft kann selbstverständlich eine gründliche Diskussion der Kantischen Gedanken nicht ersetzen. Aber für letztere ist hier nicht der Ort und Platz.
5. Eine ähnliche empirische Argumentation, allerdings auf der Basis eines simpleren Handlungsmodells findet sich bereits bei Hume [1739/40, II, 3, 3]. – Korsgaard [1986, 124; 142] hat diese Argumentation Humes so uminterpretiert, daß Hume den motivationalen Skeptizismus bezüglich praktischer Vernunft (daß reine Vernunft alleine nicht zu Handlungen motivieren kann) über den inhaltlichen Skeptizismus (daß reine Vernunft überhaupt keine Handlungsalternativen auszeichnen kann) begründe. Auch wenn der Anfang von Humes Argumentation Korsgaards Interpretation Recht zu geben scheint, so sind doch die Hauptprämissen der Argumentation ein empirisches Handlungsmodell, in dem ein Handeln aus reiner Vernunft nicht vorgesehen ist: Handlungen erwachsen immer aus einer Kombination 1. von Wünschen, Passionen o. ä. und 2. dazu passenden Annahmen über Mittel; also könnten Erkenntnisse, welcher Art auch immer, alleine keine Handlungen verursachen. Flankierend zeigt Hume noch, wie spezielle Erkenntnisse eine Rolle bei den Handlungsentscheidungen spielen können. Humes Argumentation für den motivationalen Skeptizismus kommt also gänzlich ohne eine inhaltlich skeptizistische Prämisse aus. Falsch ist an dieser strukturell völlig korrekten Argumentation nur die zu simple Handlungstheorie, die aber – wie gerade geschehen [s. o., Abschn. 2.2.] – verbessert werden kann. Korsgaard übersieht also völlig die eigentliche Brisanz von Humes Argumentation.

hauptet, Urteile, insbesondere solche der reinen Vernunft, könnten ohne Zutun von konativen Faktoren, allein aus der Entscheidungskraft des (analytischen) Erkenntnisvermögens heraus Handlungen verursachen, der muß zum Beweis erst einmal eine Handlungstheorie mit entsprechenden Handlungsgesetzen entwickeln, nach der dies möglich ist. Nach allen gängigen Handlungspsychologien ist es unmöglich. 2. Daß eine solche Theorie je entwickelt werden wird, ist zudem aus analytischen Gründen auszuschließen. Schärfer gesagt: Ein Handeln aus reiner Vernunft ist schon analytisch (nicht erst empirisch) unmöglich. Denn nach der Idee der reinen praktischen Vernunft müßten ja ganz bestimmte Erkenntnisse über eine Handlung, etwa daß sie *F* ist, die Ausführung dieser Handlung *verursachen*, während andere Erkenntnisse, daß sie nicht *F* ist, daß sie *G*, *H* etc. ist, nicht diese Wirkung haben. Wenn dies aber so sein soll, dann muß es (in psychologischen Gesetzen beschreibbare) psychische Mechanismen geben, die bewirken, daß genau Erkenntnisse der Art, daß eine Handlung *F* ist, zur Ausführung dieser Handlung führen. Welche der vielen von der (reinen) Vernunft erkannten Eigenschaften einer Handlung für die Ausführung dieser Handlung relevant ist, ist dann durch diesen *Mechanismus* festgelegt und nicht durch die reine Vernunft. Die reine Vernunft kann nicht die praktische Relevanz ihrer Erkenntnisse verfügen. [Ausführlichere Kritik: Lumer 2000, 70–74.]

3. Zweiter motivationsverleihender Mechanismus: apriorische Achtung vor dem moralischen Gesetz – und ihre analytische Unmöglichkeit

Die Idee eines Handelns aus reiner Vernunft scheidet also ganz unabhängig von dem hier zugrunde gelegten Handlungsmodell. Handeln aus reiner Vernunft ist aber nur *eine* Form des praktischen fundativen Externalismus; andere Versionen sind durchaus mit dem entscheidungstheoretischen Handlungsmodell vereinbar. Kant war mit der ursprünglichen Konzeption des Handelns aus reiner Vernunft anscheinend selbst nicht zufrieden. (Er selbst schreibt, es könne kein apriorisches Vernunftinteresse geben [Kant, *GMS* BA 122 f.] und ein Handeln aus reiner Vernunft sei – für den menschlichen Verstand – unerklärlich [*ibid.* BA 123; 125; 128].) Jedenfalls wendet er sich in der *Kritik der praktischen Vernunft* erneut dem Problem zu, wie die reine Vernunft, insbesondere die Einsicht in das moralische Gesetz, praktisch werden, also zu entsprechendem Handeln motivieren kann, und er schlägt eine neue Lösung vor: Das moralische Gesetz (gemeint ist wohl die *Einsicht* in das moralische Gesetz) erzeuge einen Schmerz über die Niederschlagung unseres Eigendünkels, und es erzeuge ein Gefühl der Achtung vor diesem Gesetz, das seinerseits

wieder Motiv zu dessen Befolgung sei [Kant, *KpV* A 129f.; A 133; A 139 bis 147]. Sehr eigenwillig an Kants Erklärung ist dann, daß er – auf die Linie eines Handelns aus reiner Vernunft zurückfallend – annimmt, bei dieser Verursachung von Affekten handele es sich um eine "intellektuelle Kausalität", welche man a priori einsehen könne [*ibid.* A 129f.; A 132; A 141f.].

A priori könnte die Erkenntnis des Sittengesetzes selbstverständlich ganz andere Wirkungen zeitigen, z. B. Ärger über solch ein einschränkendes Gesetz und damit die entgegengesetzte Motivation, es gerade nicht zu befolgen. Der Zusammenhang zwischen der Einsicht in das moralische Gesetz und entsprechender Motivation kann also nicht apriorisch sein. (Eine a priori einsehbare Kausalität ist nach einem Humeschen Kausalitätsverständnis ohnehin analytisch unmöglich.)

4. Moralische Motive i.e.S. als weitere motivationsverleihende Mechanismen – ein Überblick

Kants Idee eines Handelns aus Achtung vor dem Gesetz kann jedoch durch eine kleine Uminterpretation deutlich verbessert werden: Der Zusammenhang zwischen der Einsicht in das moralische Gesetz und der entsprechenden Motivation könnte empirisch sein, nämlich dergestalt, daß bestimmte Arten von moralischen Einsichten aufgrund gewisser Motive (intrinsischer oder extrinsischer Wünsche) regelmäßig ein gewisses Maß an Motivation verursachen (im üblichen Sinne), die moralisch ausgezeichneten Handlungen auszuführen. Die Auszeichnung von Handlungen oder Maximen als 'moralisch' würde externalistisch, also ohne Rekurs auf vorhandene Wünsche vorgenommen, und anschließend würde sich wegen der motivationsverleihenden Motive die gewünschte Motivation zum moralischen Handeln einstellen. Allerdings wäre nur eine sehr spezifische Art von Moralitätsauszeichnung praktisch wirksam; und daß sie wirksam wäre, läge an dem – moralisch gesehen – kontingenten empirischen Faktum, daß wir über Motive (intrinsische oder extrinsische Wünsche) verfügten, genau diese Art von Moralität umzusetzen. Die Bedingungen der externalistischen Ausgangshypothese wären damit gleichwohl erfüllt, jedoch nicht auf eine apriorische Weise – was man im Bereich der Motivation, also der Kausalität, ohnehin nicht erwarten darf. Dem externalistischen Programm dürfte damit also kein Abbruch getan sein. Diese psychologisch realistischere Version des praktischen fundativen Externalismus soll nun genauer untersucht werden.

Die gesuchten Motive, durch die die erste externalistische Ausgangshypothese erfüllt werden würde, nenne ich "moralische Motive (i. e. S.)" und de-

finiere diesen Ausdruck so: Es sind Motive, bei denen ein moralisches Urteil, daß ein bestimmtes Handeln moralisch geboten (bzw. verboten) oder moralisch gut (bzw. schlecht) ist, den Ausgangspunkt einer moralkonformen Motivation bildet (also einer (nicht unbedingt starken) Motivation, das moralisch Gebotene bzw. Gute zu tun und das moralisch Verbotene bzw. Schlechte zu unterlassen).

Das oben skizzierte allgemeine, entscheidungstheoretische Handlungsmodell ist moralphilosophisch noch kaum einengend, weil es ziemlich formal ist, nämlich alle möglichen Arten von Einzelmotiven oder Einzelbewertungen zuläßt und nur angibt, wie aus diesen Einzelbewertungen Entscheidungen synthetisiert werden. Moralphilosophisch interessanter und für die Theoriebildung viel einschneidender ist die empirische Untersuchung der Arten von einzelnen Motiven oder Bewertungen, ob sie moralische Motive i. e. S. sind. Wenn man die sonstigen Wege menschlicher Motivation berücksichtigt, gäbe es im Prinzip drei grundsätzliche Typen solcher moralischen Motive:

1. *Ursprünglich intrinsische moralische Motive:* Die Befolgung von moralischen Geboten oder die Realisierung von moralisch Gutem wird motivational ursprünglich intrinsisch und situationsunabhängig positiv bewertet; Moralbefolgung ist also ein ursprünglich intrinsisches Motiv. Dies ist die psychologisch realistischere und Kant am nächsten liegende Interpretation der Kantischen Idee eines Handelns aus Achtung vor dem Sittengesetz [s. o., Abschn. 3]. Auch Nida-Rümelins "normativische" und "imperativische" Handlungsgründe, die sich nicht konsequentialistisch rekonstruieren ließen [Nida-Rümelin 1993, 173–178], könnten psychologisch realistisch als solche intrinsischen moralischen Motive interpretiert werden.

2. *Affektinduzierte moralische Motive:* Auch unter den oben [Abschn. 2.2] erläuterten affektinduzierten intrinsischen Motiven könnten sich moralische Motive finden. Diese würden so funktionieren: Ein Werturteil, daß es gut bzw. schlecht ist, daß bestimmte – externalistisch begründete – moralische Standards erfüllt bzw. verletzt worden sind, führt zu moralischen Affekten, die wiederum (in Abhängigkeit von der Intensität dieser Affekte) intrinsische Motive zu bestimmten moralischen Handlungen induzieren. Die wichtigsten derartigen Affekte sind moralische Schuldgefühle, Empörung, Entrüstung und moralische Wut. Moralische Schuldgefühle beispielsweise werden durch Urteile verursacht, gegen die eigenen moralischen Standards verstoßen zu haben; die Schuldgefühle induzieren dann das intrinsische moralische Motiv, sich selbst zu bestrafen oder den moralischen Schaden wiedergutzumachen.

3. *Extrinsische moralische Motive:* Die dritte Möglichkeit für moralische Motive i. e. S. ist: Die Befolgung der von einem selbst akzeptierten Moral hat

nach Ansicht des Handelnden irgendwelche anderen für ihn intrinsisch positiven Folgen. Empirisch sind die zentralen derartigen Folgen die Steigerung des Selbstwertgefühls und die Vermeidung des sonst drohenden Sinkens des Selbstwertgefühls. Moralbefolgung ist in diesem Fall nur ein extrinsisches Motiv.

5. Dritter motivationsverleihender Mechanismus: moralischer Nativismus (empirische Uminterpretation der Achtung vor dem Gesetz) – nicht erfüllt

Die gerade vorgestellten kombinatorischen Möglichkeiten moralischer Motive waren nur rein hypothetisch entwickelt in Anlehnung an sonstige motivationale Mechanismen der Menschen. Daß diese kombinatorischen Möglichkeiten auch empirisch vorkommen, ist bei den letzten beiden Möglichkeiten (affektinduzierte und extrinsische moralische Motive) ziemlich offensichtlich, nicht aber bei der ersten. Die These, daß die erste Möglichkeit empirisch realisiert ist, daß es also ursprünglich intrinsische moralische Motive gibt, nenne ich "*moralischen Nativismus*".⁶ Die Grundidee des moralischen Nativismus ist, daß schon das Erkennen, daß bestimmte bzw. die eigenen moralischen Standards bei einer bestimmten Handlung erfüllt sind, unmittelbar, nämlich intrinsisch, ein motivierender Grund ist, diese Handlung auszuführen.

Der moralische Nativismus ist, wie ich nun zeigen möchte, falsch; d. h. es gibt keine ursprünglich intrinsischen Motive zu moralischem Handeln. Zum einen gibt es keine psychologischen Evidenzen für den moralischen Nativismus. Der moralische Nativismus (oder eine nicht ausformulierte Vorform davon) wird meist von Philosophen vertreten; denn er ist vielleicht die für Ethiker schönste Moralpsychologie – leider aber keine wahre. Nur wenige (vor allem kognitionspsychologisch, nicht aber motivationspsychologisch orientierte) Psychologen in der Nachfolge Kohlbergs haben diese These überhaupt

6. "*Moralischer Nativismus*" bedeutet nicht, daß es ein angeborenes moralisches Empfinden gibt; dies ist lediglich *eine* Möglichkeit des moralischen Nativismus. Es bedeutet vielmehr, daß jemand ursprünglich intrinsische Motive hat, moralische Gebote situationsunabhängig umzusetzen oder moralisch Gutes situationsunabhängig zu realisieren. Der Ausdruck "Nativismus" leitet sich davon ab, daß nach dieser Theorie ursprünglich intrinsische Motive zu moralischem Handeln zu unserer biologischen Ausstattung gehören, auch wenn sie erst nach der Geburt auftreten oder sich sogar nur nach biologisch einigermaßen vorgegebenen Mustern entwickeln. (Im letzteren Sinne sagt man z. B. auch von Chomskys Theorie des Spracherwerbs, sie sei "nativistisch".)

vertreten, zudem ohne weitere empirische Belege und z. T. auch nur implizit [vgl. z. B. Kohlberg 1976, 125f.]. Selbst in der Kohlberg-Schule wird diese These nun massiv kritisiert; der Weg von moralischen Urteilen zu moralischem Handeln sei vielmehr über andere Motive vermittelt [z. B. Nunner-Winkler 1993, insbes. 336]. Das Gros der Moralpsychologen hingegen erklärt moralisches Handeln über andere Motive.

Allerdings scheint die Selbsterfahrung mancher Menschen den moralischen Nativismus zu bestätigen: Diese Menschen glauben, sie hätten eine bestimmte moralische Pflicht, und schon sind sie motiviert, entsprechend zu handeln. Dieser tatsächlich oft vorhandene Ablauf alleine bestätigt aber noch nicht den moralischen Nativismus; denn der Wunsch, die moralische Pflicht zu erfüllen, könnte ohne weiteres ein extrinsisches Motiv sein. Für die Wirksamkeit extrinsischer Motive ist es keineswegs erforderlich, daß wir uns die Verbindung vom extrinsisch Gewünschten (Pflichterfüllung) zum intrinsisch Gewünschten (z. B. hohes Selbstwertgefühl) vor Augen halten. Solche Verbindungen werden, nach einiger Übung, vielmehr unbewußt und automatisch hergestellt. Daß wir sie bewußt herstellen, uns also intrinsische Ziele vor Augen halten, ist sogar die Ausnahme. Damit das Pflichtgefühl ein intrinsisches Motiv ist, müßte vielmehr gelten, daß das Subjekt selbst nach längerer Überlegung und gezielter Nachfrage und bei Offenheit sich selbst gegenüber und entsprechender Reflexionsfähigkeit keine dahinterliegenden Ziele angeben kann, warum es die Pflicht zu erfüllen wünscht. An dieser Stelle werden aber schon die meisten Menschen, die die genannten Bedingungen erfüllen, fündig. Aber selbst wenn sie nicht fündig werden, wäre damit nur gezeigt, daß es sich um einen *intrinsischen* Wunsch handelt. Um zu zeigen, daß dieser Wunsch auch *ursprünglich* intrinsisch ist, muß man in einem tiefenpsychologischen Prozeß der Genese des Wunsches nach Pflichterfüllung nachgehen, und dabei dürften keine dahinter liegenden Wünsche zum Vorschein kommen (und – gegen einen einfachen Falsifikationismus – für den Mißerfolg müßte es noch eine theoretische Erklärung geben). Bislang haben Psychologen von keinem erfolgreichen Beleg dieser Art für ein ursprünglich intrinsisches Pflichtmotiv berichtet. Kurz, der von der alltäglichen Erfahrung ausgehende Versuch, den moralischen Nativismus zu beweisen, ist psychologisch zu naiv, bedenkt nicht die psychologisch äußerst gehaltvollen Bedingungen für ursprünglich intrinsische Wünsche.

Zum zweiten ist der moralische Nativismus nur eine isolierte These; bislang hat sich keiner seiner Vertreter (oder Vertreter von Protoformen des moralischen Nativismus) der Mühe unterzogen, ihn in eine ausgearbeitete Handlungspsychologie zu integrieren. Daß er in eine solche Theorie paßt, wäre also erst noch zu zeigen. Bis dahin ist die These ad hoc.

Zum dritten und vor allem widersprechen die Konsequenzen des moralischen Nativismus der Erfahrung: Auch die These des moralischen Nativismus selbst ist nie so weit präzisiert worden, daß angegeben wurde, welches Prädikat 'M' ein Handelnder genau einer Handlung zusprechen muß, damit er sie in motivierender Weise intrinsisch positiv bewertet, was also genau das subjektive Kriterium für eine solche positive Bewertung ist.

Die eine Möglichkeit ist, daß dieses Prädikat 'M' schon mit einem bestimmten Moralkriterium identisch ist, z. B. einem utilitaristischen (so daß gälte: 'M(x) = 'x hat unter den bekannten Alternativen die höchste utilitaristische Wünschbarkeit') oder einem Kantischen (so daß z. B. gälte: 'M(x) = 'ich kann nicht wollen, daß alle nach einer Regel handeln, nach der ein Unterlassen von x erlaubt wäre'); diese Möglichkeit kann man als "inhaltlichen moralischen Nativismus" bezeichnen. 1. Gegen den inhaltlichen Nativismus spricht die Vielfalt der faktisch akzeptierten Moralen und die völlige Diversität der individuellen Moralentwicklungen; denn zumindest die anderen ursprünglich intrinsischen Motive sind anthropologisch einheitlich. 2. Zudem spricht die Ontogenese unserer Moralvorstellungen dagegen: In der Regel bleiben Menschen nicht beim ersten Moralkriterium stehen und entwickeln die Nachfolger auch nicht einfach aufgrund von ihnen intuitiv neu zufallenden Kriterien weiter, sondern sie arbeiten vorhandene Kriterien weiter aus, modifizieren sie aufgrund von Überlegungen, machen sie z. B. kohärent, usw. Diese Entwicklung entspricht nicht dem Bild von ursprünglich intrinsischen Motiven. 3. Die meisten von Ethikern diskutierten Moralkriterien sind auch viel zu abstrakt, um von vielen Menschen (insbesondere auch Kindern), denen wir durchaus Moral zusprechen, überhaupt verstanden zu werden.

Die andere, interessantere Möglichkeit ist, daß das gesuchte Prädikat 'M' nur formal ist und noch inhaltlich offen, nämlich daß es das Prädikat ist 'ist moralisch geboten'; dies könnte man als "formalen moralischen Nativismus" bezeichnen. 1. Gegen den formalen moralischen Nativismus spricht zum einen, daß Kinder nach der kognitiven Vermittlung moralischer Normen zusätzlich zu deren Befolgung angehalten werden müssen, und zwar mit Hilfe anderer Motive wie etwa dem Wunsch nach Anerkennung oder mit Hilfe eines allgemeinen Autoritätsglaubens, daß die Befolgung auch irgendwie gut für das Kind selbst ist. Mit dem in Frage stehenden intrinsischen Motiv wäre dies nicht oder nicht in dem faktischen Umfang erforderlich, denn es würde ja mindestens einen Teil der erforderlichen Motivation liefern. 2. Zum anderen sehe ich keine Möglichkeit, wie man dieses Prädikat 'M' – in formaler Weise – anders explizieren kann als unter Rekurs auf andere Motive, etwa so: 'M(x) = 'x folgt aus gegebenen Situationsinformationen und einer hinsichtlich Zeit und Subjekten allgemeinen Regel, hinter der eine normative Erzwin-

gungsinstanz *e* steht, die die Befolgung dieser Regel verbindlich macht, d. h. von der eine motivierende Wirkung derart ausgeht, daß Subjekte, die erkennen, daß aus Situationsinformationen und der von der Erzwingungsinstanz *e* ausgezeichneten Regel folgt, daß sie *y* tun sollen, (ein Stück weit) motiviert sind, *y* zu tun.' Der Rekurs auf die anderen Motive ist erforderlich, weil moralische Gebote irgendeine Form von Verbindlichkeit implizieren. Angenommen, die moralische Erzwingungsinstanz sei die Empathie; dann müßte der Handelnde bei dieser Version des moralischen Nativismus z. B. in motivierender Weise werten: 'Meiner empathischen Motivation, *h* zu helfen, zu folgen ist intrinsisch gut.' Solch eine Art von intrinsischer Motivation ist logisch möglich, aber empirisch reichlich implausibel: 2.1. Zum einen wäre ein solches Motiv kaum von dem grundlegenden Motiv – in dem Beispiel also dem empathischen Motiv – zu unterscheiden. 2.2. Zum anderen hätte das moralische Motiv nur die Funktion, ein ohnehin schon vorhandenes Motiv zu verstärken; evolutionstheoretisch ist es ziemlich unwahrscheinlich, daß sich ein solches Motiv herausbildet, statt daß sich das grundlegende Motiv zu größerer Stärke entwickelt. (Dieses Problem besteht nicht bei den durch moralische Affekte – wie Empörung, Schuldgefühle etc. [s. u.] – induzierten moralischen Motiven: Diese Motive halten nicht zur Pflichtbefolgung an, sondern zu sanktionierenden Reaktionen auf Pflichtverletzungen; dadurch haben sie eine eigene Funktion.)

6. Vierter motivationsverleihender Mechanismus: affektinduzierte moralische Motive – nicht hinreichend allgemein

Wie affektinduzierte moralische Motive funktionieren, wird am besten deutlich durch eine Darstellung der einzelnen Motive. Bei moralischen Schuldgefühlen ist ein Urteil, gegen die eigenen moralischen Standards verstoßen zu haben, der Ausgangspunkt; die daraus erwachsenden Schuldgefühle (Affekte) induzieren dann das intrinsische moralische Motiv, sich selbst zu bestrafen oder den moralischen Schaden wiedergutzumachen. Empörung und moralische Wut werden durch Urteile verursacht, daß jemand anderes schuldhaft die moralischen Standards des Handelnden verletzt hat; diese Gefühle induzieren das intrinsische moralische Motiv, den anderen zu bestrafen oder – seltener – den moralischen Schaden wiedergutzumachen.

Die affektinduzierten moralischen Motive sind praktisch sehr wichtig bei der Moraldurchsetzung, sowohl beim Subjekt selbst als auch sozial. Aber sie sind zum einen affektabhängig, entstehen und vergehen mit dem Affekt; sie sind also zeitlich instabil und deshalb nicht als Basis einer langfristigen ratio-

nalen Entscheidung geeignet. Schon aus diesem Grund sind sie kaum als *der* zentrale motivationsverleihende Mechanismus der externalistischen Moralbegründung zu gebrauchen. Zum anderen motivieren die affektinduzierten moralischen Motive nur zu reaktiven, negativen moralischen Handlungen: Bestrafungen für Moralverletzungen oder Wiedergutmachung, nicht aber zu positiven Handlungen. Sie sind deshalb auch inhaltlich zu wenig allgemein, um als der zentrale motivationsverleihende Mechanismus des fundativen Externalismus zu fungieren.

Um einen Effekt für die *positive* Moralbefolgung zu haben, müssen affektinduzierte moralische Motive von anderen Motiven aufgegriffen werden: Erst der Wunsch, eigene Schuldgefühle mit anschließender Selbstbestrafung oder aus anderer Menschen Empörung etc. resultierende Bestrafung durch andere zu vermeiden, ist das zugehörige positive Motiv zur Moralbefolgung. Diese Motive gehören schon zu den nun zu untersuchenden extrinsischen moralischen Motiven.

7. Fünfter motivationsverleihender Mechanismus: extrinsische moralische Motive – Steigerung des Selbstwertgefühls als zentrales moralisches Motiv

Das möglicherweise einzige allgemeine, nach Ansicht vieler Sozialpsychologen aber auf jeden Fall zentrale extrinsische moralische Motiv und *das* i. e. S. moralische Motiv ist die gezielte Steigerung bzw. die gezielte Vermeidung des Absinkens der moralischen Komponente des Selbstwertgefühls. Dieses Motiv funktioniert wie folgt: Menschen entwickeln ideale Selbstbilder von sich und Standards, wie sie sein möchten, und bewerten sich selbst häufig anhand dieser Selbstbilder. *Eine* Komponente dieser idealen Selbstbilder und Standards ist moralischer Natur. Eine positive Selbstbewertung wegen guter oder Übererfüllung der selbst gesetzten Standards führt zu angenehmen Affekten der Selbstwertsteigerung: Stolz auf, Zufriedenheit mit sich selbst u. ä. Eine negative Selbstbewertung wegen Untererfüllung oder Verletzung der selbst gesetzten (moralischen) Standards führt zu unangenehmen Gefühlen der Selbstwertminderung: Ärger über, Unzufriedenheit mit sich selbst, Selbstverachtung etc. [Keshen 1996, 3–6; Schwartz 1977, 231]. Menschen wissen – schon früh (etwa 40% der 6–7jährigen) [Nunner-Winkler 1993, 322; 324] – um diese emotionalen Effekte ihres Handelns. Das extrinsische moralische Motiv besteht dann darin, sich aus hedonistischen Gründen möglichst angenehme Selbstwertgefühle zu verschaffen, indem man seinen moralischen Standards gerecht wird [*ibid.* 320f.; 326f.; Schwartz 1977, 226; Heckhausen 1989, 287f.].

Das Motiv der Optimierung der Selbstwertgefühle macht zunächst vielleicht keinen besonders respektablen Eindruck, weil es in gewisser Weise egoistisch, hedonistisch ist. Aber zum einen ist es zentral bei (zumindest einer Reihe von) Entscheidungen, die als moralisch besonders herausragend und lobenswert angesehen werden. Z. B. gaben von Yad Vashem anerkannte Judenretter es als Motiv für ihre Entscheidung an, etwa: "Mein Mann sagte mir, daß sie, wenn wir ihnen nicht hülften, alle umgebracht werden würden. Ich konnte diesen Gedanken nicht ertragen. Ich hätte es mir nie vergeben." [Oliner/Oliner 1988, 168.] Zum anderen ist es anscheinend der einzige motivationsverleihende Mechanismus, der die Bedingungen der ersten externalistischen Ausgangshypothese in zeitlich stabiler und inhaltlich uneingeschränkter Weise erfüllt: Einem breiten Spektrum von unabhängig von diesem Motiv begründeten Erkenntnissen über moralisch Gebotenes oder Gutes wird motivationale Wirkung verliehen.

8. Das Problem der motivationalen Zulieferfunktion externalistischer Moralbegründungen

Fazit der bisherigen Untersuchung ist, daß es zwar nicht die von Kant angedachten motivationsverleihenden Mechanismen gibt, daß es aber sehr wohl moralische Motive i. e. S. gibt, nämlich vor allem das extrinsische moralische Motiv der Optimierung der moralischen Komponente des Selbstwertgefühls. Dieses Motiv ist zwar nicht so hehr, wie man und insbesondere Kant sich das wünschen würde. Die erste externalistische Ausgangshypothese ist damit aber trotzdem erfüllt: Es gibt ein inhaltlich (einigermaßen) neutrales, dem moralischen Urteil folgendes Motiv.

Zu untersuchen bleibt noch, ob dieser motivationsverleihende Mechanismus die zweite externalistische Ausgangshypothese erfüllt: Können externalistische Moralbegründungen eine Akzeptanz der begründeten Moral erzeugen, der die Optimierung der moralischen Komponente des Selbstwertgefühls motivationale Kraft verleiht? Oder konkreter: Können externalistische Moralbegründungen dazu führen, daß wir unser ideales Selbst über die so begründete Moral definieren und die Befolgung dieser Moral positive Selbstwertgefühle erzeugt (bzw. die Nichtbefolgung negative Selbstwertgefühle)?

8.1. Zur Psychologie moralischer Urteile: psychologischer Internalismus

Die Beantwortung der Frage nach der motivationalen Zulieferfunktion externalistischer Moralbegründungen setzt eine ausgearbeitete Psychologie mora-

lischer Urteile voraus. In *Motive zu moralischem Handeln* [Lumer 2002, Abschn. 7] habe ich zentrale Hypothesen solch einer Theorie skizziert, die sich vor allem auf Material aus der entwicklungspsychologischen Schule in der Tradition Piagets stützt. Diese Theorie hat drei Teile:

1. *Heteronomer Beginn der Moral*: Ursprünglich erwerben Kinder moralische Regeln und Kriterien über zwei Mechanismen: Gratifikationsorientierung und Autoritätsglauben. *Gratifikationsorientierung* bedeutet, Strafen wegen Nichtbefolgung moralischer Normen zu vermeiden und Anerkennung und Belohnung für die Befolgung zu erhalten zu suchen. Die Akzeptanz moralischer Standards über den *Autoritätsglauben* funktioniert so, daß das Kind glaubt, diese von den Erwachsenen eingeführten Standards seien irgendwie gut und wichtig, auch gut und wichtig für es selbst, selbst wenn es sie nicht versteht. Dieser Beginn der moralischen Orientierung ist also völlig heteronom; die Akzeptanz der Moral hat nichts mit Einsichten in den wahren Wert der Moral zu tun.

2. *Kognitive Weiterentwicklung der Moral*: Der Bestand akzeptierter moralischer Standards wird dann dynamisiert durch drei Arten von kognitiven Fortschritten: a) *Wissensfortschritte* ermöglichen ein Verständnis des Sinns komplexer Normen, von sozialen Beziehungen, systemischen Wirkungen von Handlungen, von Regeln usw. b) Die akzeptierte Moral wird zunehmend *kohärenter*. Ältere Kinder und Jugendliche sind einerseits zu höheren Abstraktionsleistungen in der Lage, die es ihnen erlauben, immer abstraktere und allgemeinere Standards und Regeln zu akzeptieren und bisherige Standards als Spezialfälle dieser abstrakteren Standards zu verstehen. Andererseits entstehen dadurch mehr Möglichkeiten für Widersprüche zwischen diversen Standards, die nun getilgt werden müssen. c) Aufgrund des – vor allem ab der Pubertät einsetzenden – *Autoritätsverlustes* der bisherigen Leitfiguren werden auch die von ihnen propagierten moralischen Standards in Frage gestellt; und es wird nach primären Begründungen für sie gesucht. Ohne diese Begründung wird die Akzeptanz des fraglichen Standards ein Stück weit suspendiert.

3. *Psychologischer Internalismus*: Die kognitive Weiterentwicklung der Moral ist nur formal, also inhaltlich neutral. Es gibt aber auch autonome inhaltliche Quellen der moralischen Höherentwicklung, aus denen ganz neue moralische Prinzipien begründet werden oder die im Konfliktfall zwischen bisher akzeptierten moralischen Standards einen dieser Standards unterstützen. Und diese autonomen Quellen der Moral sind genau diejenigen Motive, auf die sich die diversen internalistischen Moralbegründungen im Laufe der Ethikgeschichte gestützt haben: zum einen moralnahe Motive, vor allem die Empathie, aber u. a. auch das aus Achtungsaffekten erwachsende Achtungsmotiv (das vor allem den Schutz des Achtungsobjekts zum Ziel hat),

zum anderen das – technisch ausgedrückt – spieltheoretisch informierte Eigeninteresse, mit dem allseitig vorteilhafte Kooperationen begründet werden. – Wohlgermerkt, auch dieser dritte Teil der Theorie besteht nicht aus ethischen Aussagen, sondern ist empirischer, moralpsychologischer Natur; und für die hier aufgelisteten autonomen Quellen der Moral sind in dem oben erwähnten Aufsatz [Lumer 2002, Abschn. 7] jeweils empirische biographische Beispiele angeführt, in denen das fragliche Motiv der Grund für eine Änderung der vom Subjekt akzeptierten moralischen Standards war. Andere autonome Quellen für moralische Standards als die genannten waren in dem von mir durchgesehenen Material nicht auszumachen. – Der psychologische Internalismus ist auf den ersten Blick erstaunlich, bedeutet er doch, daß autonome Moralkriterien und damit dann auch die nach ihnen gefällten Moralurteile letztlich doch wieder gewissen Motiven zu moralischem Handeln folgen und nicht umgekehrt. Auf den zweiten Blick ist dies schon weniger erstaunlich, weil eben reine Kognitionen keine Relevanzen festlegen können.

Die ersten beiden Teile dieser Moralpsychologie des moralischen Urteils sind vermutlich für die meisten Theoretiker akzeptabel und im vorliegenden Zusammenhang einer Kritik des praktischen fundativen Externalismus ohne Brisanz. In zweifacher Hinsicht problematisch (umstrittener und brisanter) ist hingegen der psychologische Internalismus. Außer den eben erwähnten, zugegebenermaßen nicht sehr starken empirischen Belegen lassen sich noch einige theoretische Gründe für ihn anführen: 1. Die im vorliegenden Zusammenhang relevante subjektive Akzeptanz eines Moralkriteriums muß mindestens in der Weise affektiv wirksam sein, daß unser Selbstwertgefühl auf die Einhaltung oder Verletzung dieses Kriteriums durch uns selbst reagiert. Da die einzelnen Typen von Emotionen jeweils durch sehr spezifische, ihnen von Natur aus zugehörige Arten von Urteilen verursacht werden [Solomon 1993, 228–307; Lumer 2000, 458–463; 465–471], muß also in irgendeiner Form *von seiten der Emotionen* vorgegeben sein, was praktisch relevant ist – und nicht z. B. von seiten der reinen Vernunft. 2. Gegen jede vorgeschlagene Moral, für die man keine motivierenden, internalistischen Gründe sieht, kann man einen internalistischen Einwand der offenen Frage vorbringen: „Moralität’ wird hier zwar so und so definiert; aber warum sollte das für mich relevant sein, warum sollte ich diese Moral als die meinige (affektiv und motivational) akzeptieren, mich nach ihr richten?“ 3. Die ersten beiden Argumente zeigen zugleich Schwierigkeiten für die Hypothese auf, externalistische Gründe könnten zu der erforderlichen affektiven Akzeptanz einer Moral führen. Diese Schwierigkeiten ließen sich durch einen affektiven moralischen Nativismus lösen, der besagt, daß wir eben natürlich disponiert seien, bestimmte Moralkriterien in affektiv wirksamer Weise zu akzeptieren, einfach

wenn sie uns vorgeschlagen oder in bestimmter – externalistischer – Weise begründet werden. Gegen diesen affektiven moralischen Nativismus sprechen aber die gleichen oder analoge Gründe wie schon oben gegen den (motivationalen) moralischen Nativismus: Der inhaltlichen Version dieses Nativismus widerspricht die Vielgestaltigkeit und Entwicklungsfähigkeit der faktisch akzeptierten Moralen. Die formale Version wäre wegen der Verbindlichkeitskomponente des Prädikats 'moralisch geboten' wieder auf unabhängige Motive zur Befolgung der Moral angewiesen, die die nativistische Moral evolutionär überflüssig machen würde.

8.2. Widerlegung eines Einwandes gegen den psychologischen Internalismus

Trotz der genannten theoretischen Gründe spricht *prima facie* gegen den psychologischen Internalismus, daß es in der philosophischen Ethik ja eine Vielzahl von externalistischen Moralbegründungen gibt, deren Autoren die so begründete Moral auch ernsthaft vertreten; ein Beispiel dafür ist Kant. Zum einen muß aber unterschieden werden, ob der jeweilige Autor diese Moral nur akademisch verteidigt oder ob er sie auch so akzeptiert, daß seine affektiv wirksamen Selbstbewertungen sich danach richten; im vorliegenden Zusammenhang ist selbstverständlich nur letzteres relevant. Allerdings wird man einer genügend großen Zahl externalistischer Ethiker, insbesondere Kant, zugutehalten, daß sie auch diese stärkere Bedingung erfüllen. Zum anderen werden wegen des heteronomen Beginns der Moral und der kognitiven Bearbeitung dieser heteronomen Quellen nicht nur ungeheuer viele, sondern z. T. auch äußerst abstruse Moralen akzeptiert. Dies steht überhaupt nicht in Widerspruch zu der skizzierten Psychologie moralischer Urteile. Diese Theorie behauptet nur, daß in solchen Fällen die jeweils akzeptierte Moral nicht stringent aus irgendwelchen autonomen Quellen der Moral entwickelt wurde. Drittens und vor allem ist zu unterscheiden, ob ein Autor die von ihm in affektiv wirksamer Weise akzeptierte Moral nur in seinen Publikationen externalistisch begründet oder ob er sie auch wegen dieser Begründung akzeptiert. Dem psychologischen Internalismus widerspricht nur die letztere Möglichkeit, für die ich aber keine empirischen Belege sehe. Gerade Kant selbst ist m. E. ein gutes Beispiel für diesen Unterschied: Erste Indizien dafür, daß Kant die von ihm philosophisch begründete Moral nicht aufgrund dieser externalistischen Begründung akzeptiert – und ich unterstelle wieder: in affektiv wirksamer Weise akzeptiert –, sind, daß er 1. mehrere Begründungen für 2. verschiedene Moralprinzipien vorlegt, die allerdings extensional gleich sein sollen. Die weiteren Argumente für meine Behauptung sind strittiger: 3.

Kants Begründungen der verschiedenen Versionen des Kategorischen Imperativs sind ziemlich grundlegend falsch; 4. die Formeln des Kategorischen Imperativs sind selbst unklar; und 5. aus ihnen folgen nicht die einzelnen Gebote (Pflicht, Versprechen einzuhalten, etc.), die Kant aus ihnen ableiten will.⁷ Dies legt folgende Interpretation nahe: Kant hat eine bestimmte Moral aus ganz anderen Gründen akzeptiert (zunächst heteronom angenommen, dann kognitiv weiterentwickelt etc.). Zu dieser Moral hat er in reifen Jahren eine Systematisierung gesucht, die zu seinem apriorischen Philosophiebegründungsprogramm paßt. Diese Systematisierung ist ihm nur recht schlecht gelungen; und die vielen Fehler zeigen, daß nicht die externalistische Begründung, sondern der vorher akzeptierte Inhalt den Primat hat. Damit ist Kant also kein Gegenbeispiel gegen den psychologischen Internalismus.⁸

8.3. Konsequenzen aus dem psychologischen Internalismus: Scheitern des fundativen Externalismus an der praktischen Forderung

Fazit aus dem psychologischen Internalismus für die Diskussion des praktischen fundativen Externalismus ist: Aufgrund des heteronomen Beginns der

7. Nichtkantianer werden diese Behauptungen vielleicht so akzeptieren, Kantianer hingegen selbstverständlich nicht. Hier ist aber nicht der Ort, eine Jahrhunderte alte und Bibliotheken füllende Kontroverse über Kants Moralbegründung zu entscheiden. Deshalb verzichte ich hier auf eine Begründung; vgl. aber z. B. Lumer (2000, 75f.).
8. Ein weiterer Einwand gegen den psychologischen Internalismus ist: Der psychologische Internalismus stütze sich ja nur auf tatsächlich vorgefundene Moralentwicklungen. Solche Beobachtungen sagten aber direkt nichts über weitere mögliche Entwicklungen. Ethische Begründungen könnten abstrakter und radikaler sein als die je im Alltag vorgefundenen und dann auch möglicherweise die praktische Forderung erfüllen. [Anton Leist, briefliche Mitteilung vom 20.7.2000.] – Der hier vorgestellte psychologische Internalismus besteht ja nicht einfach nur aus Beispielen, sondern ist eine Theorieskizze, die sich auf Beispiele stützt, die eben alle einem bestimmten Muster entsprechen. Aus einer Theorie kann aber selbstverständlich folgen, daß es bestimmte Phänomene nicht geben kann. Aus dem Energieerhaltungssatz folgt z. B., daß es kein Perpetuum mobile geben kann. Gewisse Anfangsbestätigungen für eine Theorie schließen selbstverständlich nicht aus, daß diese Theorie irgendwann falsifiziert wird. Aber dazu müssen konkrete Gegenbeispiele vorgebracht werden, die bislang fehlen; die bloße Möglichkeit der Falsifikation ist selbstverständlich keine Falsifikation. Und es muß eine bessere Theorie entwickelt werden, die nicht nur das Ausgangsmaterial, sondern auch die Gegenbeispiele erklärt. Eine solche Theorie fehlt bislang erst recht – obwohl fundative Externalisten sie dringend zur Stützung ihres Anspruchs, die praktische Forderung zu erfüllen, benötigen.

Moral und ihrer kognitiven Bearbeitung werden zwar auch externalistisch begründete Moralen in affektiv wirksamer Weise akzeptiert. Aber die externalistischen Moralbegründungen als solche können keine affektive Akzeptanz einer Moral mit den entsprechenden Folgen für das Selbstwertgefühl erzeugen. Externalistische Moralbegründungen erfüllen also nicht die motivationale Zulieferfunktion für das moralische Motiv i. e. S., die Optimierung des Selbstwertgefühls. Die zweite Ausgangshypothese des praktischen fundativen Externalismus ist also verletzt. Und dies bedeutet: Der fundative Externalismus (unabhängig von seiner je konkreten Ausformung!) ist nicht in der Lage, die praktische Forderung zu erfüllen (daß seine Moralbegründung kluge Subjekte ein Stück weit zur Befolgung der Moral motivieren muß).

Eine mögliche Reaktion auf dieses Ergebnis wäre, die praktische Forderung preiszugeben und nur darauf zu setzen, daß es möglich ist, Menschen *irgendwie* zur Einhaltung der externalistisch begründeten Moral zu motivieren. Diese Möglichkeit ist sicherlich gegeben, impliziert aber, 1. daß die gewünschte Motivation letztlich manipulativ erzeugt werden muß, weil sie sich nicht auf die eigentlichen Gründe für die externalistisch begründete Moral stützen kann, 2. was selbst *prima facie* nicht besonders moralisch ist und 3. mindestens langfristig vermutlich nicht besonders erfolgreich ist, und es impliziert 4. daß es auch keine motivierenden moralischen Gründe gibt, sich selbst oder andere zu der Moralbefolgung zu manipulieren. Eine zweite mögliche Reaktion ist, auf jegliche Form der Motivation zur Befolgung der externalistisch begründeten Moral zu verzichten. Der fundative Externalismus würde dadurch aber praktisch irrelevant. Eine dritte und die naheliegendste mögliche Reaktion ist: den fundativen Externalismus aufzugeben.

Prof. Dr. Christoph Lumer
 Università degli Studi di Siena
 Dipartimento di Filosofia
 Via Roma 47
 I-53100 Siena

LITERATUR

- Camerer, Colin [F.] (1995): Individual Decision Making, in: John H. Kagel/Alvin E. Roth (Hg.): *The Handbook of Experimental Economics*, Princeton/NJ: Princeton University Press, S. 587–703.
- Habermas, Jürgen (1991): Erläuterungen zur Diskursethik, in: ders.: *Erläuterungen zur Diskursethik*, Frankfurt/M.: Suhrkamp, S. 119–226.
- Habermas, Jürgen (1992): *Faktizität und Geltung. Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*, Frankfurt/M.: Suhrkamp.
- Heckhausen, Heinz (1989): *Motivation und Handeln*. 2., völlig überarb. u. erg. Aufl., Berlin [etc.]: Springer.
- Hume, David (1739/40): *Ein Traktat über die menschliche Natur*, übers. v. Theodor Lipps, hg. v. Reinhard Brändt, Hamburg: Meiner, 1978.
- Keshen, Richard (1996): *Reasonable Self-Esteem*, Montreal [etc.]: McGill-Queen's Univ. Press.
- Kohlberg, Lawrence (1976): Moralstufen und Moralerwerb. Der kognitiventwicklungstheoretische Ansatz, in: ders.: *Die Psychologie der Moralentwicklung*, hg. v. Wolfgang Althof unter Mitarbeit von Gil Noam und Fritz Oser, Frankfurt/M.: Suhrkamp, 1996.
- Korsgaard, Christine M. (1986): Skeptizismus bezüglich praktischer Vernunft, in: Stefan Gösepath (Hg.): *Motive, Gründe, Zwecke. Theorien praktischer Rationalität*, Frankfurt/M.: Fischer Taschenbuch Verlag, 1999, S. 121–145.
- Lumer, Christoph (1994): Was ist eine triftige Moralbegründung?, in: Georg Meggle/Ulla Wessels (Hg.): *Analyomen I. Proceedings of the 1st Conference "Perspectives in Analytical Philosophy"*, Berlin–New York: de Gruyter, S. 785–796.
- Lumer, Christoph (1997a): Habermas' Diskursethik, in: *Zeitschrift für philosophische Forschung* 51, S. 42–64.
- Lumer, Christoph (1997b): The Content of Originally Intrinsic Desires and of Intrinsic Motivation, in: *Acta Analytica – Philosophy and Psychology* 18, S. 107–121.
- Lumer, Christoph (1999): Begründung, in: Hans Jörg Sandkühler (Hg.): *Enzyklopädie Philosophie*, Bd. 1, Hamburg: Meiner, S. 149–156.
- Lumer, Christoph (2000): *Rationaler Altruismus. Eine prudentielle Theorie der Rationalität und des Altruismus*, Osnabrück: Universitätsverlag Rasch.
- Lumer, Christoph (2002): Motive zu moralischem Handeln, in: *Analyse & Kritik* 24, S. 163–188.

- Lumer, Christoph (i. V. a): *Kognitive Handlungstheorie. Empirische Handlungsgesetze, Freiheit und die Grundlagen praktischer Rationalität*, in Vorb.
- Lumer, Christoph (i. V. b): Intentions Are Optimality Beliefs – but Optimizing What? Erscheint in: *Erkenntnis*
- Lynch, John G. jr./Cohen, Jerry L. (1978): The Use of Subjective Expected Utility Theory as an Aid to Understanding Variables That Influence Helping Behavior, in: *Journal of Personality and Social Psychology* 36, S. 1138 bis 1151.
- Nida-Rümelin, Julian (1993): *Kritik des Konsequentialismus*, München: Oldenbourg.
- Nunner-Winkler, Gertrud (1993): Moralische Motivation und moralische Identität. Zur Kluft zwischen Urteil und Handeln, in: Detlef Garz/Fritz Oser/Wolfgang Althof (Hg.): *Moralisches Urteil und Handeln*, Frankfurt/M.: Suhrkamp, 1999, S. 314–339.
- Oliner, Samuel P./Oliner, Pearl M. (1988): *The Altruistic Personality. Rescuers of Jews in Nazi Europe*, New York–London: Free Press–Collier Macmillan.
- Schwartz, Shalom H. (1977): Normative Influences on Altruism, in: Leonhard Berkowitz (Hg.): *Advances in experimental social psychology*, Bd. 10, New York–San Francisco–London: Academic Press, S. 221–279.
- Solomon, Robert C. (1993): *The Passions. Emotions and the Meaning of Life* [1976]. *Revised edition*, Indianapolis: Hackett.