Theorienreduktion in den Sozialwissenschaften: eine Fallstudie am Beispiel der Balancetheorien

Manhart, Klaus

THEORIENREDUKTION IN DEN SOZIALWISSENSCHAFTEN

Eine Fallstudie am Beispiel der Balancetheorien

KLAUS MANHART

SUMMARY. *Theory Reduction in the Social Sciences. The example of balance theories.* A central topic both in philosophy of science as well as in the empirical sciences is the reconstruction of the relations between theories. In the past comparisons of theories by means of traditional linguistic methods have proved to be extremely difficult and complicated. In this article the reconstruction of intertheoretical relations based on model-theoretical terms is propagated, as formulated within the structuralist view of theories. The efficiency of a model theoretical based comparison of theories is demonstrated by two theory elements from the social science research program of balance theories: The basic element by Heider and the transitivity theory by Holland and Leinhardt. First of all both theory elements are introduced informally and reconstructed in the structuralist format. On the basis of these reconstructions can be shown, that the Heider theory can be formally reduced to the Holland-Leinhardt theory and that the theory younger in history means an improvement. However, an integration of all balance theoretical elements into a theory net is not possible.

Key words: balance theories, intertheoretical relations, reduction, social sciences, structuralist reconstruction

1. EINLEITUNG

Wissenschaftliche Theorien treten selten isoliert auf, sondern sind meist eingebunden in ein komplexes Netz theoretischer Strukturen und unterschiedlicher Anwendungen. Einzelne Theorien stehen dabei – mehr oder weniger explizit – in bestimmten „intertheoretischen Relationen" zu anderen Theorie-Elementen. Viele theoretische Strömungen in den Natur- und Sozialwissenschaften sind beispielsweise dadurch gekennzeichnet, daß es ausgezeichnete Basiselemente gibt, die im historischen Ablauf in verschiedene Richtungen ausdifferenziert werden. Historisch jüngere Theorien können etwa Spezialfälle oder Verallgemeinerungen von älteren Theorien sein. Einen breiten Raum nimmt in der wissenschaftstheoretischen

Journal for General Philosophy of Science **29**: 301–326, 1998.

Literatur die Reduktionsproblematik ein und damit die Frage, ob und wie eine historische ältere und einfachere Theorie – wie die klassische Mechanik – auf eine jüngere, komplexere – wie die relativistische Mechanik – zurückgeführt werden kann.

Die exakte Rekonstruktion von Beziehungen zwischen einzelnen Theorien bildet einen wichtigen Baustein wissenschaftstheoretischer Untersuchungen. Nicht nur lassen sich damit intertheoretischen Relationen zwischen Theorie-Elementen präzise angeben, es kann auf der Basis solcher Vergleiche auch entschieden werden, ob etwa historisch jüngere Theorien allgemeiner sind als ältere und damit einen Fortschritt darstellen. Theorienvergleiche bilden auch die Grundlage der Rekonstruktion ganzer Theoriennetze und der Analyse von Forschungsprogrammen (Lakatos, 1982).

Stegmüller (1986, S. 288–289) betrachtet das Studium intertheoretischen Relationen als besonders zukunftsträchtiges und erfolgversprechendes Unternehmen. Zum einen besteht hier ein enormer wissenschaftstheoretischer Nachholbedarf, zum anderen stieß die bisherige Wissenschaftstheorie auf diesem Gebiet auch drastisch an die Grenzen ihrer Methoden. Der Grund ist, daß in der herkömmlichen Wissenschaftstheorie – etwa der Carnap-Schule – Theorienvergleich auf einer linguistischen, mikrologischen Ebene stattfand, bei dem Term-für-Term-Vergleiche angestellt wurden. Für viele intertheoretische Relationen ist dies aber äußerst kompliziert und mühsam.

Stegmüller propagiert einen Theorienvergleich auf modelltheoretischer Basis, wie er insbesondere in der strukturalistischen Theorienkonzeption ausgearbeitet wurde (Sneed, 1971; Balzer, 1982; Stegmüller, 1980, 1986; Balzer *et al.*, 1987). Während die modelltheoretische Methode zur Rekonstruktion intertheoretischer Beziehungen schon auf naturwissenschaftliche Theorien angewendet wurde, bildet ein solches Vorhaben in den Sozialwissenschaften ein Desiderat. Der Beitrag soll die Leistungsfähigkeit eines modelltheoretisch basierten Theorienvergleichs für den Vergleich sozialwissenschaftlicher Theorien aufzeigen. Dabei ist zu bedenken, daß Theorienvergleiche in den Sozialwissenschaften insofern problematisch sind, als sozialwissenschaftliche Theorien gewöhnlich sehr vage und kaum formalisiert sind. Eine Formalisierung ist aber eine notwendige Voraussetzung dafür, über intertheoretische Beziehungen überhaupt eine Aussage machen zu können.

Die strukturalistische Theorienauffassung und das damit verknüpfte modelltheoretische Vorgehen scheint uns aus mehreren Gründen für die Untersuchung der Beziehungen sozialwissenschaftlicher Theorien geeignet:

- Erstens ist es nicht notwendig, Theorien formalsprachlich in einen Kalkül zu übersetzen, sondern zur Präzisierung werden „informelle“ Mengenlehre und einfachste Mittel der Prädikatenlogik verwendet. Diese „sanfte“ Formalisierung erleichtert die Rekonstruktion sozialwissenschaftlicher Theorien insofern, als keine vollständige Formalisierung nötig ist.

- Zweitens stehen beim makrologischen, modelltheoretischen Vorgehen der Strukturalisten statt sprachlicher Konstrukte globale Strukturen im Vordergrund, bei denen von vornherein mit modelltheoretischen Entitäten operiert wird. Bei Vergleichen werden unmittelbar derartige globale Einheiten zueinander in Beziehung gesetzt. Die Erfahrung lehrt, daß ein solcher Umgang mit globalen Strukturen viel einfacher ist als es die entsprechenden mikrologischen Verfahrensweisen sind (Stegmüller, 1986, S. 73–74).

- Drittens verfügt die strukturalistische Theorienkonzeption über allgemeine und gut ausgearbeitete Konzepte der Entwicklung empirischer Theorien und stellt einen differenzierten und präzisen Begriffsapparat zur diachronen Theorienanalyse zur Verfügung. Die Ergebnisse von Theorienvergleichen können deshalb leicht in die diachrone Analyse und die Untersuchung ganzer Forschungsprogramme eingebaut werden (z.B. Balzer *et al.*, 1987).

Wir verwenden als Beispiel ein Theorienpaar aus der Sozialpsychologie bzw. der Soziologie: die Gleichgewichtstheorie von Fritz Heider (1946) und die Transitivitätstheorie von Holland und Leinhardt (1971). Beide Elemente gehören zum Forschungsprogramm der Balancetheorien, einer breiten Theorieströmung, welche die Sozialwissenschaften in den fünfziger bis siebziger Jahren stark dominiert hat. Die ursprüngliche Balancetheorie in der Fassung von Heider (1946) hat dabei eine entscheidende Initialisierungsrolle gespielt. Zum einen hat sie eine große Zahl von Experimenten angeregt und die Formulierung anderer – komplexerer – Gleichgewichtstheorien initiiert, zum andern hat sie die Anregung gegeben, genauer über Formalisierungsmöglichkeiten in den Sozialwissenschaften nachzudenken (Sukale, 1971, S. 48).

Wir gehen im folgenden so vor, daß wir zunächst beide Theorie-Elemente vorstellen und gleichzeitig im strukturalistischen Format rekonstruieren. Auf der Basis der beiden Rekonstruktionen soll dann das modelltheoretische Vorgehen bei der Ableitung der intertheoretischen Relationen demonstriert werden.

2. HEIDERS THEORIE KOGNITIVER BALANCE

Heiders Gleichgewichtstheorie beschäftigt sich mit der Widersprüchlichkeit in den Beziehungen, die eine Person zwischen sich und anderen Elementen ihrer Umwelt wahrnimmt. Ausgangspunkt ist Fritz Heiders Artikel „Attitudes and Cognitive Organization" aus dem Jahr 1946. In diesem Aufsatz macht Heider die grundlegende Annahme, daß soziale Wahrnehmung gestaltähnlichen Strukturprinzipien folgt und ausgeglichene oder balancierte Zustände gegenüber unausgeglichenen oder unbalancierten Zuständen präferiert werden. Gleichgewichtige Systeme werden als angenehm und kohärent im Sinne der Gestaltpsychologie empfunden: „Mit Gleichgewichtszustand ist eine Situation gemeint, in der die Relationen zwischen den Größen harmonisch zueinander passen; es gibt keinen Drang zu einer Veränderung" (Heider, 1977, S. 238).

Die Zustände „Ausgeglichenheit" und „Unausgeglichenheit" werden in der Heider-Theorie auf kognitive Triaden – eine Konfiguration bestehend aus drei Elementen – angewendet. Eine solche Triade wird gebildet aus einer wahrnehmenden Person p, einer anderen Person o und einem impersonalen Objekt x. x kann dabei „eine Situation, ein Ereignis, eine Idee oder irgendein Ding etc." sein (Heider, 1946, S. 107). Genau genommen unterscheidet Heider noch andere Systeme, nämlich dyadische Beziehungen mit zwei Elementen und triadische Konfigurationen mit drei Personen. In Anlehnung an Sukale (1971, S. 49) wollen wir diese einfacheren Systeme aber außer Betracht lassen.

Die in einer Triade bestehenden Relationen unterteilt Heider in positive und negative Einheits- oder Gefühlsbeziehungen („unit" oder „sentiment relations"). Einheitsrelationen U bestimmen, ob je zwei Einheiten zusammengehören oder nicht, wie etwa p besitzt x oder p und o sind sich ähnlich. Gefühlsrelationen L beziehen sich auf eine Bewertung von o oder x durch p und sind Sympathie- oder Antipathierelationen, z.B. p schätzt x oder o verabscheut x. Die Identifikation als Einheits- bzw. Gefühlsrelation ist für die Entscheidung, ob ein System als balanciert wahrgenommen wird oder nicht, aber nicht ausschlaggebend. Bestimmend ist nur die Anzahl der beteiligten positiven und negativen Beziehungen.

Auf der Basis der eben vorgestellten informellen Konzepte können bereits einige strukturalistische Prädikate angegeben werden. Ein charakteristischer Grundzug der strukturalistischen Theorienkonzeption ist, die „präsystematisch" vorgegebene Theorie durch „informelle Axiomatisierung" mengensprachlich zu rekonstruieren. Mit „informeller Axiomatisierung" ist gemeint, daß empirische Theorien mit Konzepten der Logik und Mengenlehre in Form von Axiomen (neu) beschrieben und präzisiert werden. Dabei wird keine formale Kunstsprache zugrundegelegt, sondern nur Logik und informelle (naive) Mengenlehre (Stegmüller, 1986, S. 21). Definition 1 legt die Grundbegriffe einer Heider-Triade durch Einführung eines Prädikats „...ist ein Heider-Graph" mengensprachlich fest.

Definition 1

x ist ein *Heider-Graph* ($x \in HG$) genau dann, wenn es O, P und N gibt, so daß gilt:

(1) $x = \langle O,P,N \rangle$

(2) O ist eine nicht-leere Menge der Kardinalität 3

(3) $P \subseteq O \times O$

(4) $N \subseteq O \times O$

(5) $P \cap N = \varnothing$

(6) Für alle $x \in O$: $\neg \langle x,x \rangle \in P \cup N$

(7) Wenn $x,y \in O$ dann $\langle x,y \rangle \in P \cup N$ oder $\langle y,x \rangle \in P \cup N$

Die Axiome besagen, daß O eine dreielementige Menge ist (2) und P und N auf O definierte Relationen (3,4) sind, welche disjunkt (5) und irreflexiv sind (6). In Axiom (7) wird die von Heider implizit getroffene Annahme der Vollständigkeit der Struktur formuliert, d.h. daß zwischen allen Paarelementen aus O entweder eine P- oder N-Relation existieren muß. Die uninterpretierten Grundbegriffe sollen folgende intendierte Interpretation haben: O soll eine Menge von Personen und Nicht-Personen sein und P bzw. N die Menge der positiv bzw. negativ verknüpften Objektpaare. $\langle x,y \rangle \in P$ ist zu lesen als: zwischen x und y besteht eine positive Relation und entsprechend $\langle x,y \rangle \in N$: zwischen x und y besteht eine negative Relation.

Auf der Basis dieser Grundbegriffe lassen sich verschiedene abgeleitete Beziehungen definieren, die zur weiteren Verwendung nützlich sind. Zunächst wird definiert, was allgemein eine Relation R ist und dann, was eine Nicht-Person (NP) und eine Person (PE) ist (vgl. hierzu auch Sukale 1971, S. 53).

Definition 2

Wenn $x = \langle O,P,N \rangle \in HG$ dann

(1) $R: = P \cup N$

(2) $x \in NP$ genau dann wenn (gdw) $x \in O$ und es existiert kein y, so daß $y \in O$ und xRy

(3) $x \in PE$ gdw $x \in O$ und $\neg\, x \in NP$

(4) $TR: = R \times R \times R$

Definition 2-1 legt R als Vereinigung von P und N fest. Die Definitionen 2-2 und 2-3 bestimmen Personen und Nicht-Personen. Personen werden von Nicht-Personen rein formal dadurch unterschieden, daß alles, was kein Anfangspunkt einer Relation ist, eine Nicht-Person ist und Personen diejenigen Elemente, die keine Nicht-Person sind. Definition 2-4 legt die Triadenmenge TR als Menge von Tripeln, also als dreifaches Cartesisches Produkt über R fest.

Die Definitionen charakterisieren mengensprachlich eine der Triaden bzw. Heider-Graphen, wie sie in Abbildung 1 dargestellt sind. In Abbildung 1 sind alle acht möglichen Permutationen von P- und N-Relationen grafisch aufgelistet. Es muß nun festgelegt werden, welche davon ausgeglichen und welche unausgeglichen sind. Hierzu liefert Heider eine einfache Regel, die auf dem eingangs erwähnten gestaltpsychologischen Prinzip basiert: „Eine Triade ist im Gleichgewicht, wenn alle drei Relationen positiv sind oder wenn zwei Relationen negativ sind und eine positiv ist. Ungleichgewicht tritt dann auf, wenn zwei Relationen positiv sind und eine negativ ist" (Heider, 1977, S. 240). Der Fall mit drei negativen Relationen wird von Heider nicht eindeutig entschieden und als mehrdeutig und unbestimmt bezeichnet (Heider 1946, S. 110; 1977, S. 240). Von den acht Fällen aus Abbildung 1 sind die Triaden (a–d) somit im Gleichgewicht, die Triaden (e–g) im Ungleichgewicht und (h) ist die indefinite Triade.

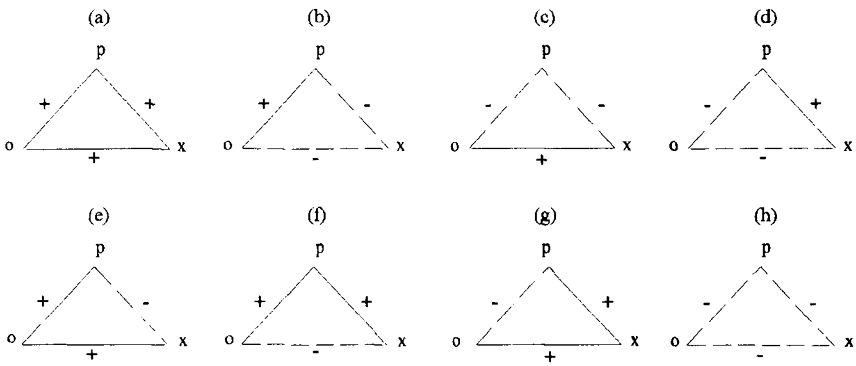

Abb. 1. Balancierte und unbalancierte Triaden nach Heider. Durchgezogene, mit „+" markierte Linien bedeuten positive, gestrichelte, mit „–" markierte Linien, negative Relationen (nach: Witte, 1989, S. 327). Die Triaden (a–d) sind balanciert, die Triaden (e–g) sind unbalanciert und Triade (h) ist „unbestimmt".

Die folgenden Situationen wären Beispiele für balancierte und unbalancierte Triaden:

(1) p ist mit o befreundet und p wählt eine Partei x, die o ablehnt.
(2) zwei Personen p und o sind befreundet und haben die gleiche Einstellung zur Kirche x.

Konfiguration (1) korrespondiert dabei der unbalancierten Triade (f), (2) der balancierten Triade (a) oder (b), je nachdem, ob p und o eine beidseitig positive oder beidseitig negative Einstellung zur Kirche haben.

Definition 3 legt balancierte, unbalancierte und indefinite Triaden als Teilmengen I (Indefinit), G (Gleichgewicht) und U (Ungleichgewicht) von TR fest.

Definition 3

Wenn $x = \langle O,P,N \rangle \in HG$ dann wird TR wie folgt in Teilmengen $G, U, I \subseteq TR$ zerlegt

(1) $I = \{\langle a,b,c \rangle \mid a,b,c \in N\}$

(2) $G = \{\langle a,b,c \rangle \mid a,b,c \in P \vee a \in P \wedge b,c \in N \vee b \in P \wedge a,c \in N \vee c \in P \wedge a,b \in N\}$

(3) $U = \{\langle a,b,c \rangle \mid a \in N \wedge b,c \in P \vee b \in N \wedge a,c \in P \vee c \in N \wedge a,b \in P\}$

Das grundlegende inhaltliche Axiom (das „Balanceprinzip") von Heider ist, daß Ungleichgewichtszustände dazu neigen, gleichgewichtig zu werden. Wenn wir „neigen" dahingehend interpretieren, daß unbalancierte Strukturen über die Zeit hinweg in balancierte überführt werden, dann ist implizit in dieser Behauptung ein Zeitindex enthalten. Damit muß ein Zeitindex eingeführt werden, so daß eine Menge von Zeitpunkten und eine entsprechende Ordnungsrelation auf ihr als weiterer Grundbegriff zu den schon festgelegten hinzukommt. Diese Terme bilden nun das vollständige Begriffsinventar, das zur mengensprachlichen Beschreibung der Heider-Theorie nötig ist. Die oben festgelegten Heider-Graphen können dann auf einzelne Zeitpunkte bezogen werden, so daß es möglich ist, diese Triaden zu unterschiedlichen Zeitpunkten zu betrachten.

Eine solche Folge von Heider-Graphen bildet in strukturalistischer Terminologie eine Menge von möglichen oder *potentiellen Modellen* M_p. Potentielle Modelle sind Strukturen, die in der Begrifflichkeit der Theorie beschreibbar sind, in denen die theoretischen Gesetze aber nicht notwendigerweise erfüllt sein müssen. Das Heidersche Balanceprinzip muß in den potentiellen Modellen also nicht unbedingt gelten. Definition 4 legt formal die möglichen Modelle der Heider-Theorie fest, indem der in Definition 1 determinierte Heider-Graph auf die Zeit bezogen wird.

Definition 4

x ist ein *potentielles Modell* der Heider-Theorie ($x \in M_p(HT)$) gdw es O, T, <, P, N gibt, so daß gilt:

(1) $x = \langle O, T, <, P, N \rangle$

(2) O ist eine dreielementige Menge

(3) $\langle T,<\rangle$ ist eine endliche, lineare Ordnung
(4) P: $T \rightarrow Pot(O \times O)$ und N: $T \rightarrow Pot(O \times O)$
(5) Für alle $t \in T$: $\langle O, P(t),N(t)\rangle \in HG$

Axiom (4) besagt, daß P und N Funktionen sind, die jedem Zeitpunkt $t \in T$ genau ein Element aus der Potenzmenge, also der Menge aller Teilmengen $O \times O$, zuordnen. Axiom (5) fordert, daß das Tripel $\langle O, P(t),N(t)\rangle$ für alle betrachteten Zeitpunkte t ein Heider-Graph ist. In diesem Axiom ist die Forderung enthalten, daß alle Objekte über die verstrichenen Zeiteinheiten die gleichen bleiben müssen. Ansonsten wäre es möglich, zwei ganz verschiedene Grundmengen zu einem Balancesystem zu verbinden, die gar nichts miteinander zu tun haben.

Mit Definition 5 ist die Festlegung der Grundbegriffe und abgeleiteten Begriffe abgeschlossen.

Definition 5
$x(t) := \langle O, P(t),N(t)\rangle$

Das Balanceprinzip von Heider läßt sich nun leicht rekonstruieren. Da gleichgewichtige Systeme als angenehm und harmonisch empfunden werden, sind diese über die Zeit stabil und werden nicht verändert. Unbalancierte Systeme werden hingegen als disharmonisch und spannungsvoll empfunden, sind instabil und tendieren zu einer Änderung in gleichgewichtige Systeme: „If no balanced state exists, then forces towards this state will arise ... If a change is not possible, the state of imbalance will produce tension“ (Heider, 1947, S. 107–108). Heiders fundamentale theoretische Aussage besteht also darin zu behaupten, daß ungleichgewichtige Zustände einen psychischen Druck erzeugen und dazu neigen, gleichgewichtig zu werden.

Die Herstellung von Gleichgewicht geschieht im einfachsten Fall durch Änderung der Relationen. In dem Beispiel (1) könnte ein Gleichgewichtszustand dadurch erzeugt werden, daß p versucht, o zur Wahl einer anderen Partei zu bewegen. Eine zweite Möglichkeit wäre, daß p seinerseits seine Parteienpräferenz überdenkt. Eine dritte Möglichkeit schließlich ist, daß p seine Relation zu o ändert. Konkrete Vorhersagen darüber, *welche* Veränderungen in einem unbalancierten System zu erwarten sind, macht die Balancetheorie allerdings nicht. Sie behauptet nur, *daß* unbalancierte Triaden dazu tendieren, gleichgewichtig zu werden.

Wir können nun das potentielle Modell durch Hinzufügen des „eigentlich inhaltlichen“ Axioms zu einem *Modell* ergänzen. Dieses inhaltliche Axiom drückt die theoretische Gesetzmäßigkeit, in diesem Fall das Balanceprinzip, aus. Die Menge der Modelle wird mit „M“ bezeichnet und bildet grundsätzlich eine (möglicherweise unechte) Teilmenge der potentiellen Modelle.

Definition 6

x ist ein *Modell* von HT (x $\in$ M(HT)) gdw es O, T, <, P und N gibt, so daß gilt:

(1) x = $\langle$O,T,<,P,N$\rangle$

(2) x $\in$ M_p(HT)

(3) Für alle t $\in$ T und für alle a: Wenn t < max(T) und a $\in$ $TR_{x(t)}$ und a $\in$ $U_{x(t)}$, dann gibt es ein t' $\in$ T so daß gilt: t < t' und a $\in$ $G_{x(t')}$ und für alle t'' > t': a $\in$ $G_{x(t'')}$

Definition 6 repräsentiert mit Axiom (3) Heiders Kernaussage, nach der unbalancierte Triaden über eine Zeitperiode hinweg in balancierte überführt werden. Definition 6–3 bildet das eigentlich inhaltliche Axiom und Fundamentalgesetz der Balancetheorie. U und G sind dabei die in Definition 3 eingeführten ungleichgewichtigen und gleichgewichtigen Heider-Graphen (Triaden). In unserer Rekonstruktion besagt das Gesetz informell, daß für alle Zeitpunkte t und alle Triaden a gilt: Wenn t kleiner ist als das Maximum von T und a zu t unbalanciert ist, dann gibt es einen Zeitpunkt t' > t, bei dem a balanciert ist und für alle Zeitpunkte t'', die größer sind als t', bleibt a balanciert. Die folgende Entwicklung der Heider-Triaden über sechs Zeitpunkte würde danach die Modelldefinition erfüllen:

t_1	t_2	t_3	t_4	t_5	t_6
U	U	U	G	G	G

Das zweite Konjunktionsglied nach dem Existenzquantor (für alle t'' > t': a $\in$ $G_{x(t'')}$) ist in unserer Interpretation von Heider nötig. Ansonsten könnte nämlich nach dem erstmaligen Wechsel einer unbalancierten Triade in eine balancierte, diese zu einem späteren Zeitpunkt wieder unbalanciert werden. Das folgende Beispiel zeigt eine solche Situation:

t_1	t_2	t_3	t_4	t_5	t_6
U	U	U	G	U	U

Wir meinen, Heider so zu interpretieren, daß dies ausgeschlossen ist; andernfalls könnte einfach das zweite Konjunktionsglied weggelassen werden.

Das Fundamentalgesetz (6–3) ist bewußt vage gehalten, dürfte aber die Intention Heiders genau ausdrücken. Zum einen wird nicht spezifiziert, welche Relationen geändert werden, zum andern wird die Zeitspanne, in der diese Änderung erfolgen soll, nicht festgelegt. An diesem Axiom kann die „Liberalität" des strukturalistischen Theorienkonzepts verdeutlicht werden. Obwohl das Axiom inhaltlich sehr vage ist, kann der Zusammenhang in strukturalistischer Deutung völlig exakt angegeben werden. Das strukturalistische Vorgehen unterscheidet sich damit wesentlich von Formalisierungsansätzen, in denen versucht wird, qualitative Begriffe zu quantifizieren oder verbale Aussagen über einen bestimmten Zusammenhang durch Angabe einer exakten Funktionsregel zu „präzisieren" (Westermann, 1987, S. 21).

Damit ist der mathematische Strukturkern der Theorie vollständig festgelegt (auf die bei strukuralistischen Rekonstruktionen übliche Diskussion der T-theoretischen Terme und die damit verbundene Definition von partiellen potentiellen Modellen wird hier nicht eingegangen, da sie in den vorliegenden Theorien keine Rolle spielen). Zu den bislang vorliegenden Strukturklassen M_p und M fehlt nun noch die Information, auf *was* diese Strukturen angewendet werden sollen. Da HT eine empirische Theorie sein soll, muß sie einen bestimmten Realitätsausschnitt erfassen. Zu den bereits eingeführten Strukturen kommt deshalb nun als dritte Strukturklasse die Information darüber hinzu, welche Weltausschnitte von der Heider-Theorie überhaupt behandelt werden. Diese für die Theorie vorgesehenen empirischen Systeme bezeichnet man als die *Menge der intendierten Anwendungen* und benutzt dafür das Symbol I. Die intendierten Anwendungen I lassen sich nicht formal definieren wie der Theoriekern, sondern werden rein pragmatisch mit Beispielen angegeben. In der Regel gibt der Urheber einer Theorie explizit Fälle in einer paradigmatischen Menge I_0 an, für die erfolgreiche Anwendungen gelungen sind. I setzt sich dann zusammen aus der Menge I_0 und einer sukzessive erweiterten Menge I^* von realen Systemen, die denen von I_0 „hinreichend ähnlich" sind: $I := I_0 \cup I^*$.

Zwei Eigenschaften von I sind besonders hervorzuheben:

1. Die intendierten Anwendungen einer Theorie sind unabhängig von der mathematischen Struktur gegeben und werden nicht mit dieser automatisch mitgeliefert.
2. Die Menge I ist eine offene Menge, die im historischen Verlauf in der Regel größer, aber auch – bei hartnäckigem Versagen einer Theorie – kleiner werden kann.

Intendierte Anwendungen der Balancetheorie I(HT) finden sich vor allem in Heider (1958/1977). In diesem Buch werden eine Vielzahl von Beispielen und Experimenten genannt, welche als erfolgreiche Anwendungen des

Kerns von HT interpretiert werden können:

I = {Partnerwahl, Xenophobie, Besitz, ...}.

Experimente von Jordan (1953), Landy und Aronson (1969) sowie Lerner und Simmons (1966) sind ebenfalls Beispiele für intendierte und erfolgreiche Anwendungen der Heider-Theorie. Allgemein lassen sich die von Heider betrachteten empirischen Strukturen als (triadische) Einstellungssysteme charakterisieren.

Resümierend und allgemein formuliert bestehen Theorien nach Auffassung der Strukturalisten also nicht aus Mengen von Aussagen, wie in der traditionellen Wissenschaftslehre, sondern sind adäquater als Gebilde erfaßbar, deren wichtigste Bestandteile mathematische Strukturen sind. Zwei dieser mathematischen Strukturklassen wurden vorgestellt. Die Modelldefinition drückt das Fundamentalgesetz einer Theorie aus und die Menge der Modelle M wird von allen Entitäten gebildet, die das Fundamentalgesetz oder inhaltliche Axiom erfüllen. Die Menge der potentiellen Modelle M_p repräsentiert das für M benötigte Begriffsgerüst, also das, was verbleibt, wenn man aus M das eigentliche Gesetz wegstreicht. Die beiden mathematischen Strukturklassen $\langle M, M_p \rangle$ werden als „Theoriekern" K einer Theorie T bezeichnet (tatsächlich gibt es mehr als zwei mathematische Strukturklassen, aber diese können für unsere Zwecke ausgeklammert werden). Der Kern, zusammen mit den intendierten Anwendungen, bildet dann die vollständige Beschreibung einer Theorie: $T = \langle K, I \rangle$. In neueren Darstellungen wird das Tupel $\langle K, I \rangle$ allerdings meist nicht als Theorie, sondern als Theorie-Element bezeichnet. Dadurch soll deutlicher zum Ausdruck gebracht werden, daß erst das Zusammenfügen derartiger elementarer Bausteine etwas ergibt, was gemeinhin als wissenschaftliche Theorie gilt. Das historisch erste Theorie-Element einer Folge von ähnlichen Elementen wird dabei auch als Basiselement bezeichnet. Die Heider-Theorie $T(HT) = \langle K(HT), I(HT) \rangle$ bildet also das erste balancetheoretische Theorie-Element.

3. DIE THEORIE TRANSITIVER STRUKTUREN

In den siebziger Jahren entwickelten Paul W. Holland und Samuel Leinhardt eine Theorie, von der sie behaupteten, sie enthielte die Heider-Theorie (sowie weitere) als Spezialfall (Holland und Leinhardt, 1971). Diese balancetheoretische Variante war zunächst als Modell interpersonaler Beziehungen formuliert und enthielt als zentrale Annahme, daß in sozialen Netzen eine Tendenz zu transitiven Beziehungen besteht. Ist a mit b befreundet und b mit c, dann – so behauptet die Transitivitätsregel – besteht die Tendenz, daß auch a mit c befreundet ist. Holland und Leinhardt verweisen darauf, daß Transitivität bereits bei Heider eine fundamentale Rolle spielt. In seinem ersten Artikel charakterisiert er L- und U-Relationen als psychologisch transitiv: „Logically, L is not transitive but there exists a psychological tendency to make it transitive when implications between U relations do not interfere with transitivity. The relation U, too, seems to be in this sense psychologically transitive" (Heider, 1946, S. 109–110).

Intendierte Anwendungen der Holland-Leinhardt-Theorie sind Freundschaftsnetze, also Strukturen, in denen gerichtete (Freundschafts-) Beziehungen als Grundrelation existieren und die z.B. in einem soziometrischen Test gemessen werden können. Formal betrachtet die Holland-Leinhardt-Theorie also Graphen mit gerichteten Relationen. Definition 1 legt solche Graphen mengentheoretisch fest.

Definition 1
x ist ein *Holland-Leinhardt-Graph* ($x \in HLG$) genau dann, wenn es X und R gibt, so daß gilt:
(1) $x = \langle X,R \rangle$
(2) X ist eine endliche, nicht-leere Menge
(3) $R \subseteq X \times X$

Auf der Basis der Relation R lassen sich drei neue Relationen definieren, indem für alle Knotenpaare festgestellt wird, ob die Wahl erwidert wird, einseitig ist oder überhaupt keine Wahl existiert.

Definition 2
Wenn $x = \langle X,R \rangle \in HLG$, dann gilt für alle $x, y \in X$:
(1) xMy gdw xRy und yRx
(2) xAy gdw xRy und $\neg$ yRx
(3) xNy gdw $\neg$ xRy und $\neg$ yRx

M bezeichnet also die gegenseitige („mutual") Wahl, A die einseitige („asymmetric") Wahl und N die beidseitige Nicht-Wahl.

Der Begriff der „Transitivität" löst nun den Begriff der „Balance" ab. Die Transitivitätsforderung wird in Definition 3 festgelegt.

Definition 3
Wenn $x = \langle X,R \rangle \in HLG$, dann gilt: x ist ein *transitiver Graph* ($x \in$ T-Graph) gdw für alle $x,y,z \in X$ gilt:
(1) xRx (Reflexivität)

(2) wenn xRy und yRz dann xRz (Transitivität)

Ein transitiver Graph ist also dadurch charakterisiert, daß die dem Graphen unterliegende Relation R reflexiv und transitiv ist. (2) ist die zentrale Forderung, welche – inhaltlich betrachtet – besagt, daß, wenn a b wählt und b c wählt, auch a c wählen muß. Ein vollständig transitives Netz hat einige interessante formale Eigenschaften, die in dem folgenden Theorem zusammengefaßt sind.

Theorem

Wenn $x = \langle X,R \rangle \in$ HLG, $x \in$ T-Graph, dann kann X in M-Cliquen partitioniert werden so daß gilt:

(1) innerhalb jeder M-Clique sind alle Paare von Individuen durch M-Kanten verbunden;

(2) zwischen zwei verschiedenen M-Cliquen sind alle Paare von Individuen entweder verbunden durch A-Kanten in der gleichen Richtung oder durch N-Kanten;

(3) die M-Cliquen bilden eine partielle Ordnung.

Jeder Teil dieses Theorems hat eine einfache soziometrische Interpretation. In (1) wird die interne Struktur jeder M-Clique durch gegenseitige positive Wahlen für jedes Mitgliedspaar gekennzeichnet. (2) charakterisiert die Relationen zwischen Paaren von M-Cliquen entweder durch eine Statusordnung von einer Clique über die andere (A-Kanten) oder durch keine Statusordnung. Teil (3) schließlich charakterisiert das ganze System von M-Cliquen als konsistente Struktur im Sinn einer partiellen Ordnung. Abbildung 2 zeigt die Konsequenzen transitiver Strukturen in einer grafischen Darstellung.

Das T-Graph-Modell zeigt, daß in einem Graphen ohne intransitive Wahlen *notwendigerweise* Muster von hierarchisch geordneten M-Cliquen entstehen. Zwischen psychologischen (Balance-)Zuständen von Individuen und dadurch bedingten Implikationen für die soziologische Gruppenstruktur besteht somit ein eindrucksvoller Mikro-Makro-Link (Hallinan und Felmlee 1975, S. 196).

Das Theorem gilt für vollständig transitive Netze. Empirisch sind komplett transitive Freundschaftsnetze die Ausnahme und bilden allenfalls den Grenzfall einer Entwicklung. Überträgt man das Balanceprinzip von Heider auf die allgemeineren HL-Graphen, so besagt das Gleichgewichtsprinzip in diesem Fall, daß in empirischen Systemen (wie Freundschaftsnetzen) eine *Tendenz* zu transitiven Relationen besteht. Die Modelle von HLT lassen sich also dadurch ausdrücken, daß für einen fixierten HL-Graphen –

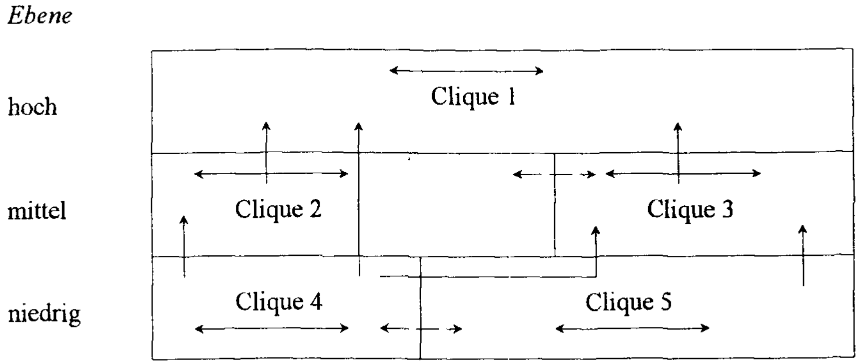

Abb. 2. Cliquen in hierarchischen Ebenen (nach Davis und Leinhardt, 1972). Durchgezogene Doppelpfeile bedeuten beidseitige Wahlen (M), gestrichelte Doppelpfeile beidseitige Ablehnung (N) und einfache Pfeile einseitige Wahlen (A).

also z.B. eine bestimmte Freundschaftsgruppe – ein Index der Transitivität definiert wird, der über ein hinreichend langes Zeitintervall zunehmen muß (oder allenfalls gleich bleiben darf). Das Vorgehen bei der Definition des Modells und potentiellen Modells ist also völlig analog zu Heider. Wir beziehen zunächst in der Definition des potentiellen Modells den HL-Graphen wieder auf die Zeit und fordern in der Modelldefinition, daß der Index der Transitivität über die Zeit zunehmen muß (oder gleich bleiben darf). Die Definitionen 4 und 5 entsprechen den Heider-Definitionen und brauchen nicht kommentiert zu werden.

Definition 4
x ist ein *potentielles Modell* der Holland-Leinhardt-Theorie ($x \in M_p(HLT)$) gdw es X, T, < ,R gibt, so daß gilt:
(1) $x = \langle X, T, <, R\rangle$
(2) X ist eine endliche, nicht-leere Menge
(3) $\langle T,<\rangle$ ist eine endliche, lineare Ordnung
(4) $R: T \rightarrow Pot(X \times X)$
(5) Für alle $t \in T$: $\langle X, R(t)\rangle \in HLG$

Definition 5
$x(t) := \langle X, R(t)\rangle$

Da wir transitivere und weniger transitive Graphen unterscheiden müssen, brauchen wir einen Index, der Informationen über das Ausmaß von Transitivität in einer gegebenen Struktur gibt. Für unsere Zwecke genügt ein einfacher Index, der die Abweichung einer empirischen Struktur von einem

deterministischen Modell kompletter Transitivität bestimmt, wie er etwa von Hallinan und Felmlee (1975) vorgeschlagen wird. Dieser Balance-Index läßt sich durch den Vergleich der Zahl intransitiver Tripel mit der Gesamtzahl aller möglichen Tripel konstruieren. Der Balance-Index soll maximal sein, wenn es keine intransitiven Tripel gibt, er ist minimal, wenn es ausschließlich intransitive Tripel gibt. Ist x der HL-Graph, dann kann der Transitivitätsindex TRX(x) damit bestimmt werden als:
1 – Anzahl aller intransitiven Triaden / Anzahl aller möglichen Triaden.

Die Anzahl aller möglichen Triaden in einem Graphen der Kardinalität n ist dabei gegeben durch
$n \cdot (n-1) \cdot (n-2) / 6$.

Definition 6 legt Kardinalität und Transitivitätsindex eines Graphen fest.

Definition 6
Wenn $x = \langle X,R \rangle \in M_p(HLT)$, dann werden n und TRX(x) wie folgt definiert:

(1) $n = \mathrm{card}(X)$

(2) $\mathrm{TRX}(x) := 1 -$

$$\frac{card\{\langle x,y,z\rangle \mid x,y,z \in X \wedge xRy \wedge yRz \wedge \neg xRz\} \cdot 6}{n \cdot (n-1) \cdot (n-2)}$$

Die Modelle der Holland-Leinhardt-Theorie lassen sich nun in einfacher Weise unter Bezugnahme auf den Transitivitätsindex definieren.

Definition 7
x ist ein *Modell* von HLT ($x \in M(HLT)$) gdw es X, T, <, R gibt, so daß gilt:
(1) $x = \langle X,T,<,R \rangle$
(2) $x \in M_p(HLT)$
(3) Für alle $t, t' \in T$: wenn $t < t'$ dann $\mathrm{TRX}(x_t) \leq \mathrm{TRX}(x_{t'})$

Das Fundamentalgesetz (3) drückt die „Tendenz zu Transitivität" aus. Es besagt, daß – in einem bestimmten HL-Graphen – für zwei beliebige Zeitpunkte t und t′ mit t′ größer t der Transitivitätsindex zu t′ größer ist als der Index für den früheren Zeitpunkt t oder gleich bleibt. Mit anderen Worten: Transitivität (und damit Balance) ändert sich über eine Zeitperiode hinweg entweder nicht oder nimmt zu. Tabelle 1 veranschaulicht die Modelldefinition mit drei fiktiven Beispielen. Die Zahlen in den Tabellenzeilen sollen die Entwicklung der Transitivitätsindizes über sechs Zeitpunkte repräsentieren. Nach der Modelldefinition sind die potentiellen Modelle S1 und S2 Modelle von HLT, S3 ist hingegen kein Modell von HLT.

TABELLE I

Drei Beispiele für mögliche zeitliche Entwicklungen der Transitivität. Die potentiellen Modelle S1 und S2 sind Modelle, S3 ist kein Modell von HLT

	Zeit					
M_p	t_1	t_2	t_3	t_4	t_5	t_6
S1	0.3	0.3	0.5	0.6	0.6	0.7
S2	0.6	0.6	0.6	0.6	0.6	0.6
S3	0.4	0.2	0.2	0.4	0.4	0.5

Holland und Leinhardt (1971, S. 107–109) wählen Freundschaftsnetze (kleine Gruppen mit „Sentiment" (Gefühls-) Relationen) als intendierte Anwendung I_0 aus. X wird als Menge von Individuen interpretiert und R als Gefühlsrelation, welche operational in einem soziometrischen Test gemessen werden kann. xRy würde dann bedeuten, daß Person x Person y in einem soziometrischen Test wählt. Diese Menge I_0 bildet die einzige, von den Urhebern ausgezeichnete Anwendung. Für dieses intendierte System sozialer Gruppen sind erfolgreiche Applikationen des Theoriekerns gelungen. Davis (1970) testete das T-Graph-Modell erfolgreich mit einer Analyse von 742 empirischen Soziomatrizen kleiner Freundschaftsgruppen. Ein ähnlicher Test mit 51 anderen Soziomatrizen wurde mit dem gleichen Ergebnis von Hallinan (1974) durchgeführt: „The null hypothesis that friendship choices are distributed randomly ist rejected for every group, and the hypothesis of transitivity is supported. It is apparent that the structure of sentiment relations in these groups approximates the structure of the transitivity model, namely, hierarchies of ranked clusters of cliques" (Hallinan, 1974, S. 366–368). Ein identisches Ergebnis zeigte sich bei Holland und Leinhardt (1975). Die HL-Theorie sagt strukturelle Trends für Freundschaftsnetze somit signifikant besser voraus als das Zufallsmodell. Die Anwendung der HL-Theorie auf andere empirische Systeme wird in Manhart (1994, 1995) diskutiert.

4. INTERTHEORETISCHE BEZIEHUNGEN

Die Transitivitätstheorie wird als eine „allgemeinere" Balancetheorie betrachtet (Holland und Leinhardt, 1971, S. 49). Dabei bleibt unklar, was dies genau bedeutet und wie sich der logische Zusammenhang zur Heider-Theorie exakt darstellt. Auf der Grundlage der erfolgten mengentheoretischen Axiomatisierung soll nun geklärt werden, welche intertheoretischen Beziehungen zwischen diesem Theorienpaar vorliegen. An diesem Beispiel wird der Vorteil der strukturalistischen Theorienkonzeption besonders klar, da sich die Relationen zwischen Theorien – im Gegensatz zu traditionellen linguistischen Vorgehensweisen – relativ einfach auf modelltheoretischer Basis untersuchen lassen.

Im strukturalistischen Standardwerk (Balzer *et al.* 1987) werden fünf grundlegende Arten intertheoretischer Relationen unterschieden: Spezialisierung, Theoretisierung, Reduktion, Äquivalenz und Approximation. Von diesen fünf intertheoretischen Relationen sind in unserem Zusammenhang nur die Spezialisierung und Reduktion relevant.

Spezialisierungs- und Reduktionsrelation können als strukturalistische Explikation und Differenzierung des substanzwissenschaftlichen Generalisierungsbegriffs verstanden werden. Die Spezialisierung ist dabei der einfachere Fall. Unter der *Spezialisierung* eines Theorie-Elementes versteht man das Hinzufügen neuer, spezieller Axiome, deren Gültigkeit nur für einen Teilbereich der intendierten Anwendungen behauptet wird. Spezialgesetze werden in der Regel als zusätzliche definitorische Bedingung zum vorgegebenen mengentheoretischen Prädikat hinzugefügt und verschärfen damit das Prädikat. Das neue Theorie-Element ist dann „spezialisierter" als das ursprüngliche Element.

Modelltheoretisch wird die Operation der Spezialisierung auf zwei einfache Bedingungen zurückgeführt: Identität zwischen der Klasse der potentiellen (und, falls T-theoretische Terme vorkommen, partiellen potentiellen) Modelle einerseits und Teilmengenrelation für Modelle und intendierte Anwendungen andererseits. Man schneidet bildlich gesprochen aus der Menge der Modelle jene Teilmenge aus, die durch ein mengentheoretisches Prädikat mit restriktiveren Bedingungen bestimmt wird und auf einen begrenzteren empirischen Bereich angewendet werden soll. Die Spezialisierung läßt sich damit wie folgt definieren (Balzer 1982, S. 294):

(D-1) Es seien $T = \langle M, M_p, I \rangle$ und $T^* = \langle M^*, M^*_p, I^* \rangle$ Theorie-Elemente.
T^* ist eine Spezialisierung von T (abgekürzt: $T^* \sigma T$) gdw
(1) $M^*_p = M_p$, (2) $M^* \subseteq M$ und (3) $I^* \subseteq I$

Bei der Spezialisierungsrelation σ bleibt das Begriffsinventar also gleich. Entscheidend ist, daß die speziellere Theorie T^* eventuell weniger Modelle hat als die allgemeinere Theorie T. Die Modelle von T^* sollen zudem nur für einen Teilbereich I^* der ursprünglichen Anwendungen I gelten.

Eine Spezialisierung der Heider-Theorie bestünde z.B. darin, zum inhaltlichen (Balance-)Axiom ein weiteres Axiom hinzuzufügen, das festlegt, in *welche* balancierte Triaden unbalancierte Triaden überführt werden. Die Modelle dieser spezialisierten Heider-Theorie wären dann eine Teilmenge der allgemeineren Theorie, in der diese Festlegungen nicht erfolgen.

Führt man ausgehend von einem Basiselement T Spezialisierungen T*, T**, T***... durch, die ebenfalls wieder spezialisiert werden können usf., so erhält man ein *Theoriennetz* von Theorie-Elementen in Form einer hierarchischen Ordnung. Ein Theoriennetz, das bestimmten zusätzlichen Bedingungen gehorcht, kann als Baum dargestellt werden (Balzer *et al.*, 1987, S. 172–204). In einem solchen Baum ist das oberste (Basis-)Element das allgemeinste Theorie-Element, von dem spezialisiertere Elemente abzweigen. Grundsätzlich müssen Theoriennetze nicht auf die intertheoretische Relation der Spezialisierung beschränkt sein, sondern zwischen den einzelnen Elementen eines solchen Netzes können auch andere Beziehungen bestehen.

Die zweite und bekannteste intertheoretische Relation ist die *Reduktion*. Grob gesprochen versteht man unter Reduktion das Zurückführen einer einfacheren Theorie T auf eine reichere, komplexere Theorie T*. T* enthält dabei in irgendeiner Form T, so daß T* erfolgreich alle jene Phänomene behandelt, die auch von T gemeistert wurden, sowie zusätzliche, von T nicht bearbeitete Fälle. Wenn wir sagen, daß T reduziert wird auf T*, so ist im folgenden immer angenommen, daß das zweite Relationsargument T* die komplexere oder „bessere" Theorie ist und das erste Argument T die einfachere. Dabei heißt T* die reduzierende, T die reduzierte Theorie.

Reduktionsbeziehungen zwischen Theorien wurden sowohl in der Wissenschaftstheorie als auch in den Substanzwissenschaften ausführlich und kontrovers behandelt. Von einigen Wissenschaftsphilosophen wie Kuhn und Feyerabend wurde die Auffassung vertreten, daß Theorien aufgrund intertheoretischer Bedeutungsvarianz einzelner Terme grundsätzlich „inkommensurabel" und damit nicht vergleichbar sind. Bekanntestes Beispiel ist die unterschiedliche Bedeutung des Massebegriffs in der klassischen und relativistischen Mechanik. Eine andere Thematik betraf die vom älteren und neueren Positivismus ausgelöste Debatte, inwieweit ganze Disziplinen auf andere reduzierbar seien, also z.B. Soziologie auf Psychologie und diese wiederum auf Physiologie. Diese Fragestellungen können hier nicht weiter verfolgt werden. Wir halten zumindest die zweite für spekulativ, die erste wird in Balzer *et al.* (1987, S. 313–320) behandelt.

Bei der Reduktionsrelation lassen sich verschiedene Unterfälle unterscheiden, die aber alle in dem gleichen formalen Rahmen behandelt werden können (Balzer *et al.*, 1987, S. 253–255). *Historische Reduktion* liegt vor, wenn eine ältere Theorie T auf eine neuere, erfolgreichere Theorie T* reduziert wird. Die wesentlichen Errungenschaften der Vorgängertheorie T müssen dabei nach T* überführt werden, so daß diese auch als Erfolge von T* betrachtet werden können. Bei der *praktischen Reduktion* benutzt man eine einfachere, reduzierte Form T einer komplexen Theorie T* aus dem praktischen Grund, Schwierigkeiten bei der Anwendung der komplexen Theorie T* auf eine gegebene Problemstellung zu meiden und diese nur grob mit T zu lösen. Ein weiterer Unterschied betrifft exakte und approximative Reduktion. Unter *exakter Reduktion* versteht man vereinfacht gesagt, daß nach entsprechender Übersetzung die Axiome der komplexeren Theorie T* jene der einfachen Theorie T implizieren. *Approximative Reduktion* liegt hingegen dann vor, wenn bei diesem Prozeß irgendwelche Näherungen durchgeführt werden müssen. Wir beschränken uns im folgenden auf exakte, historische Reduktion.

Die zentrale Forderung beim Reduktionskonzept ist, daß die Gesetze der einfacheren Theorie T aus den Gesetzen der komplexeren Theorie T* logisch ableitbar sein müssen. Fallstudien von naturwissenschaftlichen Theorien, die in der Literatur als reduzierbar angesehen werden, zeigen jedoch, daß eine unmittelbare Ableitung i.a. aus zwei Gründen nicht möglich ist. Erstens verwenden T und T* in der Regel – im Gegensatz zur Spezialisierung – einen *unterschiedlichen* Begriffsapparat. Dies bedeutet, daß eine Verbindung oder ein Brückenprinzip hergestellt werden muß zwischen jenen Termen, die in beiden Theorien nicht identisch vorkommen. Traditionell könnte man auch von der Übersetzung der beiden Theoriesprachen sprechen, aber im Strukturalismus vermeidet man – wie eingangs angedeutet – den Bezug auf Sprachen zugunsten einer modelltheoretischen Sicht. Ein zweiter Grund, warum eine unmittelbare Ableitung meist nicht möglich ist, liegt darin, daß die Basisgesetze von T* allein zu schwach sind für die Deduktion der Gesetze von T. In vielen Fällen muß daher T* noch spezialisiert werden. Erst wenn beide Operationen durchgeführt wurden – die Verbindung des Begriffsapparates und die Spezialisierung der komplexeren Theorie – ist in der Regel eine Herleitung möglich.

Die formale Behandlung des Reduktionskonzepts erfolgt dadurch, daß eine Übersetzungsrelation ρ definiert wird. In dem modelltheoretischen Rahmen bezieht diese Reduktionsrelation ρ die potentiellen Modelle beider Theorien aufeinander. ρ muß zwei Bedingungen genügen: einmal soll es zu jedem Modell der reduzierten Theorie eine Übersetzung in ein Modell der reduzierenden Theorie geben und zum anderen sollen sich die Modelleigenschaften der reduzierten Theorie bei Übersetzung ableiten lassen. Wir geben im folgenden eine vereinfachte Definition nach Balzer *et al.* (1987,

S. 277):

(D-2) Es seien T = $\langle M, M_p, I\rangle$ und T* = $\langle M^*, M^*_p, I^*\rangle$ Theorie-Elemente.
ρ reduziert T auf T* (abgekürzt: T ρ T*) gdw
(1) $\rho \subseteq M^*_p \times M_p$
(2) Für alle x, x*: wenn x ρ x* und x* $\in$ M*, dann x $\in$ M

Die zweite Bedingung der Definition drückt dabei modelltheoretisch die zentrale Forderung aus, daß bei Vorliegen der Reduktionsrelation die Gesetze von T aus den Gesetzen von T* ableitbar sind. Im Fall der historischen Reduktion ist eine zusätzliche pragmatische Forderung, daß die neue Theorie T* fähig sein sollte, erfolgreich mit all jenen Anwendungen umzugehen, welche die alte Theorie T schon meisterte. Vereinfacht gesagt wird verlangt, daß jede intendierte Anwendung der reduzierten Theorie T durch die Übersetzungsrelation ρ in eine intendierte Anwendung der reduzierenden Theorie T* übersetzt werden kann (Balzer, 1982, S. 300).

5. Zur Reduktion der Heider-Theorie auf die Holland-Leinhardt-Theorie

HLT ist im Vergleich zu HT die komplexere Theorie und behandelt mehr Anwendungen. Zunächst liegt es nahe, HT als Spezialisierung von HLT aufzufassen. Bei einer Spezialisierung muß der Begriffsapparat beider Theorien identisch bleiben. Ein kurzer Blick auf die potentiellen Modelle beider Theorien genügt, um diese Annahme zu verwerfen: beispielsweise gibt es in HT zwei Relationen, in HLT hingegen eine Grundrelation und drei definierte Relationen. HT kann also infolge des unterschiedlichen Begriffapparates kaum eine Spezialisierung von HLT sein. Es soll nun bewiesen werden, daß zwischen HT und HLT eine historische Reduktionsrelation vorliegt. Hierzu fassen wir die wichtigsten Definitionen noch einmal zusammen.

Potentielle Modelle x von HT bestehen aus einer (3-elementigen) Menge O von Objekten, einer Menge T von Zeitpunkten, der Kleiner-Relation $<$, sowie zweier Funktionen P: T $\rightarrow$ Pot(O $\times$ O) und N: T $\rightarrow$ Pot(O $\times$ O): x = $\langle O, T, <, P, N\rangle$.

Das Axiom, welches in Modellen von HT gelten muß, drückt das Balanceprinzip aus:

(A1) Für alle $t \in T$ und für alle a: Wenn $t < \max(T)$ und $a \in TR_{x(t)}$ und $a \in U_{x(t)}$, dann gibt es ein $t' \in T$, so daß gilt: $t < t'$ und $a \in G_{x(t')}$ und für alle $t'' > t'$: $a \in G_{x(t'')}$

Die potentiellen Modelle x* von HLT bestehen aus einer endlichen Menge X von Objekten, einer Menge T von Zeitpunkten, der Kleiner-Relation < und der Funktion R: $T \to Pot(X \times X)$. Potentielle Modelle von HLT haben also die Form:

$x^* = \langle X, T, <, R \rangle$.

Das Balancegesetz, welches in Modellen von HLT gelten muß, lautet:

(A2) Für alle $t, t' \in T$: wenn $t < t'$ dann $TRX(x^*_t) \leq TRX(x^*_{t'})$

Zur Definition der Reduktionsrelation ρ (D-2.1) sind die potentiellen Modelle und damit die Konzepte beider Theorien aufeinander zu beziehen. Die Begriffe müssen so ineinander überführt werden, daß intransitive Relationen aus HLT ungleichgewichtigen Heider-Triaden in HT entsprechen und transitive HLT-Relationen gleichgewichtigen Heider-Triaden. Es ist unmittelbar einsichtig, daß bei der Reduktion die jeweils ersten drei Tupelelemente von $x = \langle O, T, <, P, N \rangle$ und $x^* = \langle X, T, <, R \rangle$ einander korrespondieren müssen. Die Zahl der Objekte X wird damit eingeschränkt auf n = 3. Weniger offensichtlich ist die Zuordnung der Relationen. In HT gibt es *zwei* Grundrelationen P und N, in HLT hingegen nur *eine* Relation R. Wenn wir jedoch die P-Relation von HT mit der M-Relation von HLT identifizieren und die N-Relation von HT mit der von HLT, so stehen wir mit dieser Deutung in Einklang mit der Behandlung der historisch älteren Theorien bei Holland und Leinhardt (1971): Holland und Leinhardt setzen die positiven Relationen in den bewerteten Graphen der Vorgängertheorien mit M und die negativen Relationen mit N gleich.

Auf dem Hintergrund dieser Vorbemerkungen läßt sich die Reduktionsrelation $\rho \subseteq M_p(HLT) \times M_p(HT)$ wie folgt definieren.

(D-2.1′) $\langle x^*, x \rangle \in \rho$ gdw
(1) $x^* = \langle X,T,<,R \rangle \in M_p(HLT)$ und
$x = \langle O,T',<',P,N' \rangle \in M_p(HT)$
(2) $X = O$, $T = T'$, $< = <'$, $N = N'$, $M = P$

Wir müssen nun die zentrale Forderung (2) aus Definition D-2 nachweisen, nach der unter ρ das Balancegesetz von HT aus dem von HLT folgt, oder modelltheoretisch: für alle x, x* gilt: wenn $x \rho x^*$ und $x^* \in M^*$, dann $x \in M$.

Zum Beweis benötigen wir einen einfachen Hilfssatz, der eine Beziehung zwischen den Heider-Triaden und dem Transitivitätsindex herstellt:

(Th-1) Ist $\langle x^*, x\rangle \in \rho$ dann gilt für alle $a \in TR$:
(1) $a \in U$ gdw $TRX(x^*) = 0$
(2) $a \in G$ oder $a \in I$ gdw $TRX(x^*) = 1$

Beweis von Th-1.1:
Sei $\langle x^*, x\rangle \in \rho$.
Angenommen $\langle a,b,c\rangle \in U$. Dann ist nach Def. 3 von HT entweder $a \in N \wedge b,c \in M$ oder $b \in N \wedge a,c \in M$ oder $c \in N \wedge a,b \in M$. In der Definition des Transitivitätsindex (Def. 6 von HLT) wird der Ausdruck
$\mathrm{card}\{\langle x,y,z\rangle \mid x,y,z \in X \wedge xRy \wedge yRz \wedge \neg xRz\} = 1$
und wegen $n = 3$ $TRX(x^*) = 0$.
Angenommen umgekehrt $TRX(x^*) = 0$. Dann ist wegen $n = 3$
$\mathrm{card}\{\langle x,y,z\rangle \mid x,y,z \in X \wedge xRy \wedge yRz \wedge \neg xRz\} = 1$,
d.h. es gibt eine Triade x,y,z mit $x,y,z \in X \wedge xRy \wedge yRz \wedge \neg xRz$. Da es unter ρ nur M- und N-Relationen gibt, muß gelten: $xMy \wedge yMz \wedge xNz$. Nach Def. 2.1′ von ρ ist damit $\langle x,y\rangle, \langle y,z\rangle \in P$ und $\langle x,z\rangle \in N'$ und damit $\langle x,y,z\rangle \in U$.
Beweis von Th-1.2 analog □

Wir müssen nun unter Nutzung von (Th-1) zeigen, daß (A1) aus (A2) folgt. Eine einfache Überlegung zeigt, daß eine unmittelbare Deduktion nicht möglich ist. (A2) erlaubt nämlich, daß der Transitivitätsindex über die betrachtete Zeitperiode gleich bleibt, was für die Heider-Theorie die Konsequenz hätte, daß ein unbalancierter Zustand nicht verlassen werden muß. Das inhaltliche Axiom von HLT ist also zu schwach, um (A1) herleiten zu können und muß verschärft werden.

Eine naheliegende Spezialisierung von (A2) wäre die Verschärfung auf echte Ungleichheit der Transitivitätsindizes:

(A2′) Für alle $t, t' \in T$: wenn $t < t'$ dann $TRX(x^*_t) < TRX(x^*_{t'})$

Ist das ursprüngliche Axiom zu schwach, so ist diese Spezialisierung viel zu stark und zu unrealistisch. Der Grund ist, daß unter (A2′) für zwei beliebig dicht aufeinanderfolgende Zeitpunkte *immer* eine Transitivitätszunahme stattfinden müßte. Für die Herleitung von (A1) genügt die wesentlich schwächere Annahme (A2″), daß es zwei Zeitpunkte geben muß, bei denen der zweite Index echt größer als der erste ist.

Schließlich ist noch zu beachten, daß ein Transitivitätsindex 1 auch einer indefiniten Heider-Triade entspricht. Deren Herstellung muß jedoch ausgeschlossen werden, so daß sich als weitere Forderung ergibt, daß die Menge der M-Relationen nicht leer sein darf. Das ursprüngliche Axiom (A2) wird damit um die zwei Bedingungen (2) und (3) verschärft:

(A2$''$) (1) Für alle t, $t' \in T$: wenn $t < t'$ dann $TRX(x^*_t) \leq TRX(x^*_{t'})$ und
(2) Es gibt t, $t' \in T$: $t < t'$ und $TRX(x^*_t) < TRX(x^*_{t'})$ und
(3) Für alle t: $M_t \neq \varnothing$

Mit diesen Spezialisierungen können wir nun zeigen, daß (A1) aus (A2$''$) formal folgt.
Angenommen $x \rho x^*$ und $x^* \in M^*$.
Angenommen weiter, $t < \max(T)$ und $a \in TR_{x(t)}$ und $a \in U_{x(t)}$
dann ist wegen Th-1.1 $TRX(x_t^*) = 0$
Nach (A2$''$-2) gibt es $t' > t$, so daß $TRX(x^*_{t'}) > TRX(x^*_t)$ und damit $TRX(x^*_{t'}) = 1$. Also ist $a \in G_{x(t')}$ (da $M_t \neq \varnothing$ für alle t, ist $a \notin I_{x(t')}$).
Sei weiter $t'' > t'$, dann (nach A2$''$-1) $TRX(x^*_{t''}) \geq TRX(x^*_{t'})$. Da $TRX(x^*_{t'}) = 1$ ist die Ungleichung erfüllt gdw $TRX(x^*_{t''})$ ebenfalls 1 ist (dies gilt für alle $t'' > t'$). Damit ist $a \in G_{x(t'')}$ für alle $t'' > t'$.
Gilt also $x \rho x^*$ und $x^* \in M^*$ so folgt:
Wenn $t < \max(T)$ und $a \in TR_{x(t)}$ und $a \in U_{x(t)}$, dann gibt es ein $t' \in T$ so daß gilt:
$t < t'$ und $a \in G_{x(t')}$ und für alle $t'' > t'$: $a \in G_{x(t'')}$ □

Wir haben damit also gezeigt, daß die Heider-Theorie bei entsprechender Übersetzung und Spezialisierung auf die Holland-Leinhardt-Theorie formal reduzierbar ist. Transitivität ist damit tatsächlich das, was Heider Balance nennt. Jede Anwendung der Heider-Theorie kann durch die Reduktionsrelation in eine Anwendung der Holland-Leinhardt-Theorie übersetzt werden. HLT behandelt jene Fälle erfolgreich, die HT erfolgreich behandelt (nämlich triadische Einstellungssysteme), darüberhinaus meistert HLT zusätzliche Fälle, auf die HT nicht anwendbar ist (nämlich 3- und mehr-elementige Einstellungs- und Gruppenstrukturen).

6. AUSBLICK

Die Holland-Leinhardt-Theorie ist eine allgemeinere Variante der Heider-Theorie, in der sich zweifellos ein Fortschritt ausdrückt: sie behandelt alle Fälle der Heider-Theorie sowie zusätzliche, von Heider nicht intendierte Systeme. Es läßt sich nun weiter fragen, welche intertheoretischen Relationen zwischen der Vielzahl anderer balancetheoretischer Varianten bestehen (z.B. Abelson und Rosenberg, 1958; Cartwright und Harary, 1956; Davis 1963, 1967; Gollob, 1974; Morrissette, 1958; Osgood und Tannenbaum, 1955; Mohazab und Feger, 1985) und inwieweit diese verschiedenen Varianten ein Theoriennetz formen. Obwohl dies genauer untersucht werden müßte, läßt sich der informellen Literatur entnehmen, daß die Theoriekerne relativ unverbunden nebeneinander liegen (Frey, 1987, S. 57; Stahlberg und Frey, 1987, S. 219). Gestützt wird diese Vermutung, zumindest für einen zentralen Teilbereich dieser Theorien, durch eine Untersuchung von Koukkanen (1992). Koukkanen entwickelt ein formales Relationensystem zur strukturalistischen Rekonstruktion sozialpsychologischer Balancetheorien von Heider (1946), Cartwright und Harary (1956), Osgood und Tannenbaum (1955) und Morrissette (1958). Da die Entstehungsgeschichte und der formale Rahmen dieser Rekonstruktionen völlig unterschiedlich ist zu den hier vorgelegten, ist ein unmittelbarer Vergleich unmöglich. Nach Koukkanen sind jedoch die Heider- und die Osgood-Tannenbaum-Theorie inkompatibel, d.h. erstere ist weder eine Spezialisierung oder Reduktion letzterer noch umgekehrt. Die Cartwright-Harary-Theorie wiederum ist zwar eine strukturalistische Spezialisierung der Balancetheorie von Morrissette, zwischen dieser und der Balancetheorie von Osgood-Tannenbaum besteht aber wiederum eine Inkommensurabilität.

Obwohl also eine formale Untersuchung des ganzen balancetheoretischen Programms aussteht, lassen sich doch deutlich Hinweise erkennen, daß man von einem Theoriennetz im strengen Sinn nicht sprechen kann. Ein solches läßt sich allenfalls in einem Teilbereich von Balancetheorien rekonstruieren, die sich mit der Analyse soziometrischer Strukturen beschäftigen (Cartwright und Harary, 1956; Davis 1967; Holland und Leinhardt, 1971). Bei diesen Theorien läßt sich auch von einem echten Fortschritt sprechen (Manhart, 1995). Viele andere balancetheoretische Versuche haben aber den Charakter von „disorganisierter“ Forschung in unterschiedliche Richtungen, wie sie normalerweise bei Beginn eines Forschungsprogramms auftreten.

LITERATUR

Abelson, R.P. und Rosenberg, M.J.: 1958, Symbolic Psycho-Logic: A Model of Attitudinal Cognition, *Behavioral Science* **3**, 1–13.

Balzer, W.: 1982, *Empirische Theorien: Modelle – Strukturen – Beispiele. Die Grundzüge der modernen Wissenschaftstheorie*, Vieweg, Braunschweig.

Balzer, W., Moulines, C.U. und Sneed, J.D.: 1987, *An Architectonic for Science. The Structuralist Program*, Reidel, Dordrecht.

Cartwright, D. und Harary, F.: 1956, Structural Balance: A Generalisation of Heider's Theory, *Psychological Review* **63**, 277–293.

Davis, J.A.: 1963, Structural Balance, Mechanical Solidarity, and Interpersonal Relations, *American Journal of Sociology* **68**, 444–462.

Davis, J.A.: 1967, Clustering and Structural Balance in Graphs, *Human Relations* **20**, 181–187.

Davis, J.A.: 1970, Clustering and Hierarchy in Interpersonal Relations: Testing two Graph Theoretical Models on 742 Sociogramms, *American Sociological Review* **35**, 27–33.

Davis, J.A. und Leinhardt, S.: 1972, The Structure of Positive Interpersonal Relations in Small Groups, in: Berger, J. (Ed.), *Sociological Theories in Progress* (Vol. 2), Houghton-Mifflin, Boston, pp. 218–253.

Frey, D.: 1987, Kognitive Theorien, in: Frey, D. und Greif, S. (Hrsg.), *Sozialpsychologie* (2.Aufl.), Psychologie Verlags Union, München, S. 51–67.

Gollob, H.F.: 1974, The subject-verb-object approach to social cognition, *Psychological Review* **81**, 286–321.

Hallinan, M.T.: 1974, A Structural Model of Sentiment Relations, *American Journal of Sociology* **80**(2), 364–378.

Hallinan, M.T. und Felmlee, D.: 1975, *An Analysis of Intransitivity in Sociometric Data*, Sociometry **38**, 195–212.

Heider, F.: 1946, Attitudes and Cognitive Organization, *Journal of Psychology* **21**, 107–112.

Heider, F.: 1977, *Psychologie der interpersonalen Beziehung*, Klett, Stuttgart (Original erschienen 1958: The Psychology of Interpersonal Relations).

Holland, P.W. und Leinhardt, S.: 1970, A Method for Detecting Structure in Sociometric Data, *American Journal of Sociology* **70**, 492–513.

Holland, P.W. und Leinhardt, S.: 1971, Transitivity in Structural Models of Small Groups, *Comparative Group Studies* **2**, 107–124.

Holland, P.W. und Leinhardt, S.: 1975, Structural Sociometry. Papier präsentiert auf dem Advanced Research Symposium on Social Networks, Mathematical Social Science Board, Hanover, New Hampshire (September).

Irle, M.: 1975, *Lehrbuch der Sozialpsychologie*, Hogrefe, Göttingen.

Jordan, N.: 1953, Behavioral Forces that are a Function of Attitude and Cognitive Organization, *Human Relations* **6**, 273–287.

Koukkanen, M.: 1992, The Continuity Problem of Scientific Theories: An Example of Social-Psychological Balance Theorizing, in: Westmeyer, H. (Ed.), *The Structuralist Program in Psychology*, Huber, Bern.

Lakatos, I.: 1982, *Die Methodologie der wissenschaftlichen Forschungsprogramme*. Vieweg, Braunschweig.

Landy, D. und Aronson, E.: 1969, The Influence of the Character of the Criminal and his Victim on the Decisions of Simulated Jurors, *Journal of Experimental Social Psychology* **5**, 141–152.

Lerner, M.J. und Simmons, C.H.: 1966, Observer's Reaction to the „Innocent Victim“: Compasion of Rejection?, *Journal of Personality and Social Psychology* **4**, 203–210.

Manhart, K.: 1994, Strukturalistische Theorienkonzeption in den Sozialwissenschaften. Das Beispiel der Theorie vom transitiven Graphen, *Zeitschrift für Soziologie* **2**, 111–128.

Manhart, K.: 1995, *KI-Modelle in den Sozialwissenschaften. Logische Struktur und wissensbasierte Systeme von Balancetheorien*, Oldenbourg, München.

Mohazab, F. und Feger, H.: 1985, An Extension of Heiderian Balance Theory for Quantified Data, *European Journal of Social Psychology* **15**, 147–165.

Morrissette, J.: 1958, An Experimental Study of the Theory of Structural Balance, *Human Relations* **11**, 239–254.
Osgood, C.E. und Tannenbaum, P.H.: 1955, The Principle of Congruity in the Prediction of Attitude Change, *Psychological Review* **62**, 42–55.
Sneed, J.D.: 1971, *The Logical Structure of Mathematical Physics*, Reidel, Dordrecht.
Stahlberg, D. und Frey, D.: 1987, Konsistenztheorien, in: Frey, D. und Greif, S. (Hrsg.), *Sozialpsychologie* (2.Aufl), Psychologie Verlags Union, München, S. 214–221.
Stegmüller, W.: 1980, *Neue Wege der Wissenschaftsphilosophie*, Springer, Berlin.
Stegmüller, W.: 1986, *Theorie und Erfahrung. Die Entwicklung des neueren Strukturalismus seit 1973*, Springer, Berlin.
Sukale, M.: 1971, Zur Axiomatisierung der Balancetheorie. Eine wissenschaftstheoretische Fallstudie, *Zeitschrift für Sozialpsychologie* **2**, 40–57.
Westermann, R.: 1987, *Strukturalistische Theorienkonzeption und empirische Forschung in der Psychologie. Eine Fallstudie*, Springer, Berlin.
Witte, E.H.: 1989, *Sozialpsychologie. Ein Lehrbuch*, Psychologie Verlags Union, München.

Eisenacher Str. 10
80804 München