

## *Abriß empirischer Vorarbeiten*

Unternimmt man eine Literaturrecherche beispielsweise in den Datenbanken von PsychInfo, so stellt man fest, daß das Suchwort Wunsch (resp. wish) in der Regel zu psychoanalytischen Arbeiten oder zu Veröffentlichungen zu einzelnen Wünschen führt. Bei letzteren Studien befinden sich die Wünsche selbst eigentlich nicht im Mittelpunkt der Bemühungen. Wunschdaten werden in der Regel zur Indizierung anderer psychologischer Konstrukte benutzt, wie beispielsweise motivationale Entwicklung oder delinquentes Verhalten. Oder es sollen mit Hilfe von Wünschen klinische Phänomene diagnostiziert werden.

Die Pluralform erweist sich als Stichwort ergiebiger für das Auffinden von Literatur zu Wünschen an sich. Wünsche sind im letzten Dezennium immer wieder Gegenstand empirischer psychologischer Untersuchungen gewesen (Goodenough, 1925; Coghill, 1937; Jacobson, 1937; Gray, 1944; Cobb, 1954; Wheeler, 1963; Ables, 1972; Vandewiele, 1980, 1981; King, 1995). Diese Ergebnisse dieser Forschungsbemühungen sollen hier kurz dargestellt werden. Zuerst wird auf diejenigen Veröffentlichungen eingegangen, bei denen am stärksten zu vermuten ist, daß sie eine Begriffsexplikation des Wunsches beinhalten: Fragebogeninstrumente.

### Fragebogen zu Wünschen

Es lassen sich zwei Fragebogen zum Thema Wunsch nachweisen, Perrys (1997) „Standardized Wishes and Fear List“ und die „Reiss Profile of Fundamental Goals and Motivational Sensitivities (Reiss Profile)“ von Reiss und Havercamp (1998). Beide wurden für diagnostische Verwendungszwecke in der Psychotherapie entwickelt. Es geht dabei um das zentrale Problem, für die Einleitung einer effizienten therapeutischen Intervention das Anliegen des Psychotherapieklienten genau zu erfassen. Grosse Holtforth (1999) gibt einen Überblick über die bisherigen in diesem Bereich unternommenen Versuche einer standardisierten Erfassung von Motiven innerhalb des Prozesse der Psychotherapie.

Perry (1997) versucht mit seiner Liste von Wünschen und Befürchtungen eine Operationalisierung des Acht-Stufen-Modells lebenslanger psychischer Entwicklung nach Erikson (1950). Die Items der Liste stellen ein Kondensat umfangreicher klinischer Interviewtätigkeit, von Selbstberichten und projektiven Tests dar. Die abschließend gefundenen Formulierungen der 40 Items der Liste sind jeweils genau einer der spezifischen Eriksonschen Lebenskrisen zugeordnet. Bekanntlich ist im Sinne Eriksons die Erklímmung jeder von acht Entwicklungsstufen an die Lösung einer spezifischen Lebenskrise gekoppelt. Prinzipiell kann die Lebenskrise einen

günstigen oder aber einen ungünstigen Verlauf nehmen. Perry benutzt nun Wünsche als Maß der Selbsteinschätzung für die Antizipation eines günstigen Ausgangs durch den Klienten und, vice versa, Befürchtungen als Indikator für die Erwartung eines ungünstigen Ausgangs der jeweiligen Lebenskrise. Finden sich Entwicklungsstufen, die sich durch die Wünsche und Befürchtungen des Klienten als besonders betont werden, so hat der anwendende Psychotherapeut einen Ansatzpunkt für eine therapeutische Maßnahme identifiziert. Für die Wünsche selbst muß jedoch festgestellt werden, daß diese bei Perry ähnlich wie in den oben angesprochenen Studien zu Einzelwünschen eine bloße Indikatorfunktion für die eigentlich zu messenden theoretischen, hier Eriksonschen, Konstrukte innehaben. Eine eigentliche Klärung des Wunschbegriffs scheint mit dieser zugewiesenen Funktion offensichtlich als gar nicht notwendig erachtet worden zu sein.

Im Gegensatz zur Perrys Methode, die Items seiner Liste auf der Basis einer Inhaltsanalyse zu generieren, wählen Reiss und Havercamp (1998) für die Konstruktion der Reiss-Profile eine quantitative Methodik. Den Ausgangspunkt für die Entwicklung des Reiss-Profils bildet eine Verallgemeinerung eines Angstepfindlichkeitskonzeptes („anxiety sensitivity“, Reiss, 1991) zu einer Empfindlichkeitstheorie der Motivation durch Reiss und Havercamp (1996, 1997).

Der Angstepfindlichkeitsansatz beruht auf dem psychopathologisch relevanten Befund, daß bei Personen, die von Angsterlebnissen eine körperliche oder psychische Schädigung erwarten, eine stärkere Neigung zur Ausbildung von Panikstörungen zu beobachten ist, als bei Personen, die Angst zwar für unangenehm aber nicht für schädigend halten. Vor diesem Hintergrund muß die Aussage, daß alle Menschen Angstsituationen im selben Maß zu vermeiden motiviert sind, dahingehend spezifiziert werden, daß diese Motivation interindividuell unterschiedlich stark anzutreffen ist. Das Prinzip, daß Empfindlichkeit oder Empfänglichkeit für Außenreize weitreichende Konsequenzen für die Person selbst zur Folge hat, wird in den Bereich der Motivation übertragen. Den Autoren geht es insbesondere darum, ein Werkzeug zur Aufklärung der Motivation im psychotherapeutischen Prozeß zu schaffen. Die dazu aufgestellte Fragestellung ist aber viel weiter gefaßt: „The sensitivity theory of motivation may be viewed as a call for research on individual differences in what people want from there lives“ (Reiss & Havercamp, 1996). Diese Aussage und die oben angegebenen Alltagsbedeutung von Wünschen legen nahe, daß das ausgesprochene Forschungsinteresse des Sensitivity-Theory-Ansatzes sich auf Wünsche bezieht, diese zumindest aber mit einschließt. Dabei impliziert der Gedanke, daß zwischen Personen mit Hilfe der von ihnen präferierten Motive unterschieden werden kann, auch eine Differenzierung unter den Motiven. Unter Umständen läßt sich diese Differenzierung ebenfalls auf Wünsche mit dem Ziel einer Formulierung von Wunschklassen anwenden.

In einem lerntheoretischen Kontext werden interindividuelle Unterschiede für Verstärkerwirksamkeit oder motivationale Stärke eines entzogenen Verstärker angenommen. Die individuelle motivationale Stärke läßt sich demnach an vier Aspekten erkennen: 1) am Umfang der Verstärkung, die für die Befriedigung notwendig ist, 2) an der Intensität und Persistenz, mit der Personen den entsprechenden Verstärker nachsuchen, 3) am Ausmaß der Ungeduld, mit der Personen einen Verstärker erwarten, und 4) an der Höhe der Schwelle, ab der instrumentelles Verhalten zur Verstärkererlangung ausgelöst wird. Im weiteren beziehen sich Reiss und Havercamp auf eine bestimmte Art von Verstärkern, die sie als fundamentale Motivatoren (fundamental motivators) bezeichnen. Diese werden als „conceptually distinct reinforcers relevant to understand a significant amount of behavior displayed by a large percentage of all people“ definiert (Reiss & Havercamp, 1996, S. 624). Im Vordergrund steht also Verständnisgewinn relevanten Verhaltens möglichst vieler Personen.

Der fundamentale Charakter dieser Motive soll durch die Unabhängigkeit der gewählten Begriffe zu ihrer Bezeichnung gewährt werden. Insgesamt werden 25 fundamentale Motivatoren deklariert, die zum Teil auf dem Ansatz von Motiven nach Murray aber auch auf klinischer Erfahrung beruhen. Die Liste reicht von sozialer Anerkennung über Hilfsbedürfnis bis zu Rache (die vollständige Liste findet sich in Anhang 1).

Entsprechend der vier genannten Indikationsmöglichkeiten wurden zur Messung der motivationalen Stärke jeweils 8-18 Items für jedes einzelne der fundamentalen Motive formuliert. Die Items beginnen mit Anfangsphrasen wie beispielsweise „I like ...“ oder „I must have...“ um den subjektiven Bezug zu sichern und zu verhindern, daß allgemeine Einstellungen oder Werthaltungen gemessen werden. Insgesamt 328 Fragen bilden den Ausgangspunkt zur Entwicklung des Reiss-Profiles. Die Messung erfolgte vermutlich mit Hilfe von Ratingskalen. Leider finden sich dazu keine Angaben und demzufolge auch nicht zur Spanne der Skalen.

Argumente für die Konstruktion des Instrumentes werden aus der internen Konsistenz der Itemgruppen und der Faktorstruktur über alle Items geschöpft. In sechs empirischen Studien werden diese Argumente gesammelt und bei der Endversion des Reiss-Profiles verarbeitet.

Während die Reliabilitätsmaße, hier Maße der internen Konsistenz, für die resultierenden Itemgruppen durchweg günstig ausfallen, erweist sich jedoch die Faktorstruktur als problematisch. Gewählt wurden auf Maximum-Likelihood-Schätzung basierende exploratorische Faktormodelle mit schiefwinkliger Rotation. Die damit zugelassenen Korrelationen der Faktoren untereinander verbessert zwar in der Regel die Modellanpassung. Hier führen sie aber dazu, daß die zuvor postulierte Unabhängigkeit der fundamentalen Motive durch diese Faktormodelle implizit aufgegeben wird. Leider wird keine Maßzahl berichtet, an der zu ersehen

wäre, wie hoch die interfaktoriellen Beziehungen maximal ausfallen. Deshalb ist es leider auch nicht nachvollziehbar, inwieweit die zunächst konzeptionell geforderte Distinktheit der einzelnen Motive durch die Wahl des Faktorenmodells aufgeweicht worden ist.

Ein noch größeres konzeptionelles Problem ergibt sich durch den Unterschied zwischen der Anzahl angenommener Motive und der Anzahl extrahierter Faktoren. Erwartet wurden nämlich 25 Faktoren entsprechend der Anzahl von fundamentalen Motiven. Die Bestimmung der Anzahl zu extrahierender Faktoren ist oftmals eine schwierige Forschungsentscheidung, bei der formale Kriterien nicht immer zu verwertbaren Ergebnissen führen. Es ist festzustellen, daß eine Struktur mit 25 Faktoren nicht aufgefunden werden kann. Vielmehr sind zehn Modelle mit einer schrittweise von 10 auf 20 erhöhten Faktorenzahl erprobt worden. Mit der Begründung, daß dies am besten interpretierbar sei, fiel die Wahl schließlich auf ein 15-Faktormodell.

Auf der Grundlage dieses Modells wurden 110 Items mit der Eigenschaft bezüglich des Ladungsmusters ausgewählt, auf genau einen Faktor mit einem Betrag  $\lambda \geq \lambda_{krit} = .30$  und auf allen anderen Faktoren mit einem Betrag von  $\lambda < \lambda_{krit} = .30$  zu laden. In sechs folgenden Revisionsstudien halten die Autoren an der Struktur mit 15 Faktoren fest, sie reformulierten jedoch von den erhaltenen Items bis zu 42 Stück und fügten darüber hinaus dem Fragebogen bis zu 78 neue Items pro Studie hinzu. Bei der Itemselektion gilt das genannte Grenzkriterium der Ladungsmuster. Schließlich umfaßt die Endversion 100 Items. Dieses Vorgehen wirft eine ganze Reihe von Fragen auf. Zunächst zur Anzahl der Faktoren. Es wurde dabei ausdrücklich auf eine formal begründete Entscheidung über die Faktorzahl, wie beispielsweise anhand des Kaiser-Kriteriums (Auswahl aller Faktoren mit einem Eigenwert größer 1), für die 15 Faktoren verzichtet.

Könnte man nicht aber auch zu einer Entscheidung für zwei Faktoren kommen, wenn man der heuristischen Entscheidungshilfe eines Screeplots folgt? (Siehe dazu den Screeplot der Eigenwertverteilung der 15-Faktorenlösung in Studie 3 von Reiss und Haverkamp, 1998, siehe Abbildung 5). Allerdings wäre damit eine Reduktion der aufgeklärten Varianz von 52,9% bei 15 Faktoren auf 19,9% bei zwei Faktoren verbunden.

Vermutlich soll das genannte formale Kriterium  $\lambda_{krit} = .30$  helfen, Items im Sinne einer Einfachstruktur zu identifizieren. Die hinreichende Eignung dieses Kriteriums für die Itemselektion bleibt jedoch zu bezweifeln, denn bei der Ladung eines Items auf einen Faktor im näheren Bereich des Kriteriums liegt die gemeinsame Varianz unter 10%. Interpretiert man die Ladung als Cosinus des Winkels  $\alpha$  zwischen Faktor und Itemvektor, so läßt sich diese ungünstige Konstellation geometrisch durch  $\alpha \approx 80^\circ$  bei  $\lambda_{krit} = .30 = \cos \alpha$  veranschaulichen.

Faktor und Itemvektor stehen fast orthogonal aufeinander. Dies spricht mehr für Unabhängigkeit als für eine hinreichend gute Erklärung des Items durch den Faktor.

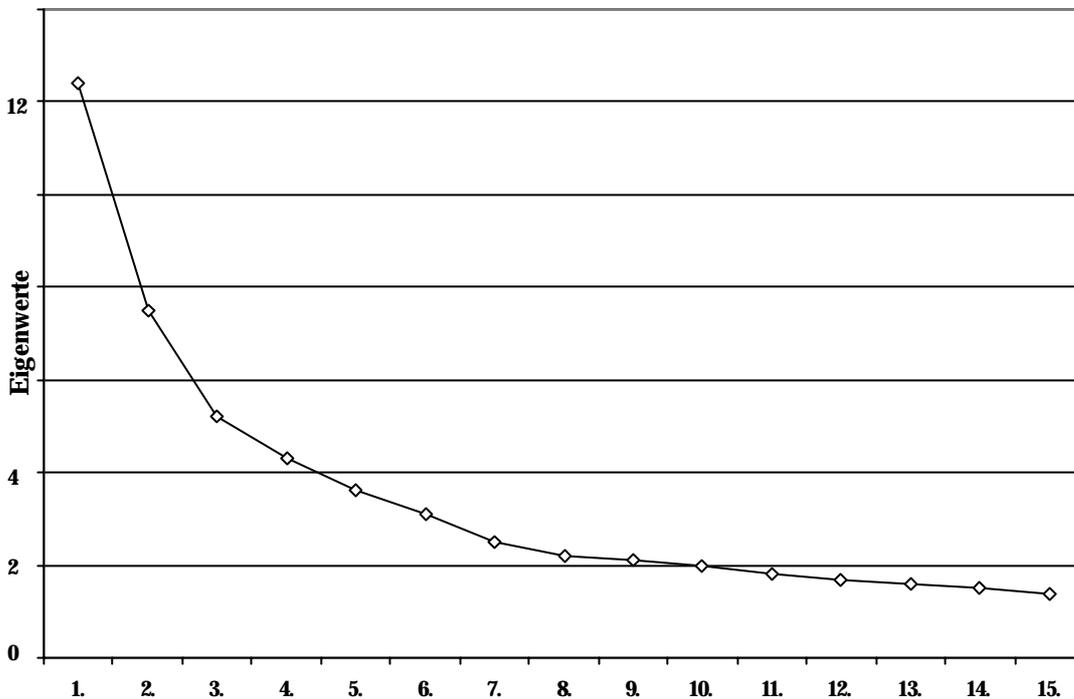


Abbildung 5: Eigenwertverteilung eines 15-Faktormodells. Liegen vielleicht doch nur zwei Faktoren vor? (nach Studie 3 bei Reiss & Havercamp 1998)

Leider gibt es kein formal abgeleitetes Kriterium für die Itemselektion im Sinne einer Einfachstruktur anhand von Itemladungen. Die Faustregel, daß die Ladung auf dem relevanten Faktor mindestens doppelt so groß sein sollte, wie auf allen anderen irrelevanten Faktoren (Muthén, 2000), leitet aber unter Umständen die Itemauswahl zu einer deutlicheren Einfachstruktur als beim von Reis und Havercamp (1998) gewählten Vorgehen. So übersteigen 34% der berichteten Itemladungen auf den relevanten Faktoren nicht das Doppelte der kritischen Grenze von  $\lambda_{krit} = .30$ . Dies ist eine Datenlage, die die vorgeschlagene Interpretation der Faktorstruktur für das Reiss-Profil beeinträchtigt.

Des Weiteren ist nicht sichtbar, ob Items aus dem ursprünglichen Pool, dessen Zusammenhänge die 15-Faktorlösung begründen, über die Revisionsschritte hinweg bis zur Endfassung erhalten geblieben sind. Es ist nicht auszuschließen, daß mit dem Konstruktionsprozeß nicht nur eine offensichtliche konzeptionelle Reduktion und Aufweichung einherging, sondern auch eine inhaltliche Verschiebung. Dieser Verdacht erhält durch den Wandel der Bedeutung des Motivs „Rache“ bei Reiss und Havercamp Nahrung. In den theoretischen Arbeiten (Reiss & Havercamp, 1996, 1997) wird das Motiv „Rache“ quasi beiläufig aufgelistet. In der empiri-

schen Arbeit (Reiss & Havercamp, 1998) fungiert „Rache“ jedoch als Bezeichner für den stärksten Faktor mit 12,4% Varianzaufklärung (Faktor 1 in Abbildung 5). Ist das Reiss-Profil somit vielleicht mehr ein spezielles Meßinstrument für Rachsucht, dagegen aber weniger eines für fundamentale Motive oder Wünsche? Obwohl hier nicht bestritten wird, daß das Reiss-Profil klinisch verwertbare Informationen liefert, so müssen doch Zweifel an dessen Vermögen angemeldet werden, fundamentale Motivatoren oder Wünsche erschöpfend erheben zu können.

Neben den eher technischen Aspekten bei der Beurteilung des Reiss-Profiles sollte nicht aus den Augen verloren werden, daß strukturerkennende Verfahren wie eine explorative Faktoranalyse die Güte der Eingangsdaten nicht verbessern können. Es ist zu hinterfragen, ob Ratingskalen die geeignete Form der Messung der Stärke von Motiven oder Wünschen darstellen. Die Vorteile bei diesem Vorgehen liegen in der hohen Standardisierung und in der gewissermaßen natürlichen Verbindung zu den auch tatsächlich bei den Reiss-Profilen angewendeten Auswertungsmethoden des ALM. Beide Umstände kommen dabei dem Gütekriterium Objektivität zugute. Ernstzunehmen ist jedoch das Bedeutungsproblem der resultierenden Scores. Beinhalten diese wirklich valide Information oder sind sie Ausdruck bloßen Ankreuzverhaltens beim Ausfüllen eines Fragebogens?

Vorausgesetzt, man hat es mit auskunftswilligen Personen zu tun, so fällt dem Verständnis des Antwortenden für den Iteminhalt eine moderierende Rolle für die schließlich gegebene Antwort zu. In diesem Sinn läßt der Kriminalautor G.K. Chesterton seinen Protagonisten, den Detektiven Pater Brown, feststellen: „Ist Ihnen schon einmal aufgefallen, daß die Menschen nie wirklich auf das antworten, was man sie fragt? Sie beantworten das, was man meint – oder vielmehr das, was sie glauben, das man meint.“ (1987, S.90 f.).

Eine Alternative zu Ratingskalen bietet das offene Frageformat. Bevor die Ergebnisse von Studien zu Wünschen mit dieser alternativen Datengewinnung berichtet werden, folgt zunächst eine allgemeine Betrachtung zu den in der Literatur bekannten Vor- und Nachteilen bei der Entscheidung für ein geschlossenes oder offenes Frageformat.