

Institut für Community Medicine, Ernst-Moritz-Arndt-Universität, Greifswald

## Psychometrische Probleme in der Gesundheitsökonomie

### Psychometric problems in health economics

Uwe Konerding

#### Zusammenfassung

*Bei dem Versuch, die für das Gesundheitswesen zur Verfügung stehenden begrenzten Mittel möglichst sinnvoll einzusetzen, stellen sich in verschiedener Hinsicht psychometrische Probleme. Eines dieser Probleme ist die Erfassung gesundheitsbezogener Lebensqualität, ein anderes die Erfassung der Ursachen gesundheitsbezogenen Verhaltens. Wie gesundheitsbezogene Lebensqualität erfasst werden könnte, ist innerhalb der Gesundheitsökonomie bereits recht umfassend diskutiert worden. Im vorliegenden Beitrag werden deshalb die wichtigsten Methoden vorgestellt, die dort zu diesem Zweck verwendet werden. Es wird untersucht, inwieweit Bedarf an weiterer Forschung zu diesen Methoden besteht. Wie die Ursachen gesundheitsbezogenen Verhaltens erfasst werden könnten, wird innerhalb der Gesundheitsökonomie bisher kaum diskutiert. Im vorliegenden Beitrag werden deshalb zwei Ansätze vorgestellt, mit denen in der Psychologie Verhaltensursachen erfasst werden. Es wird überlegt, inwieweit diese Ansätze in die Gesundheitsökonomie integriert werden könnten.*

#### Schlüsselwörter

Psychometrie; Gesundheitsökonomie; gesundheitsbezogene Lebensqualität; gesundheitsbezogenes Verhalten

#### Abstract

*Trying to attain maximal effect with the limited resources which are available for the health care system produces several psychometric problems. One of these problems consists of assessing health-related quality of life. A further problem consists of assessing the causes of health-related behaviour. The assessment of health-related quality of life*

Die Arbeiten zu diesem Beitrag wurden vom Ministerium für Schule und Weiterbildung, Wissenschaft und Forschung des Landes Nordrhein-Westfalen gefördert (Aktenzeichen des Bewilligungsbescheids: 516-40001399). Außerdem möchte ich mich bei Markus Längen, Ute Schmid, sowie zwei Gutachtern für kritische Durchsichten früherer Versionen dieses Aufsatzes bedanken.

*has already extensively discussed within health-economics. Therefore, within this contribution, the most important methods applied for this purpose are presented. The need of further research concerning these methods is investigated. The assessment of the causes of health-related behaviour has, hitherto, hardly ever been discussed within health-economics. Therefore, within this contribution two approaches which are applied in psychology for assessing the causes of behaviour are presented. The possibility of integrating these approaches into health-economics is discussed.*

#### **Key words**

Psychometrics; health economics; health-related quality of life; health-related behaviour

#### **Psychometrische Probleme in der Gesundheitsökonomie**

Die Mittel, die für das Gesundheitswesen zur Verfügung stehen, sind begrenzt. Damit stellt sich die Frage, wie mit der Begrenztheit dieser Mittel umzugehen ist. Lösungen für diese Problematik zu finden, ist das grundlegende Anliegen der Gesundheitsökonomie. Hierzu sind verschiedene Lösungsansätze denkbar. Ein Ansatz besteht darin, die Verteilung der Mittel zu optimieren. Ein anderer darin, das Gesundheitssystem und die damit zusammenhängenden Bereiche so zu gestalten, dass mit den vorhandenen Mitteln ein möglichst großes Ergebnis erzielt wird. Beide Lösungsansätze können sich sinnvoll ergänzen. Insofern empfiehlt es sich, innerhalb der Gesundheitsökonomie beide Lösungsansätze zu verfolgen und die Ergebnisse dieser Bemühungen miteinander in Beziehung zu setzen. Dies bietet sich insbesondere dann an, wenn sich unter beiden Lösungsansätzen ähnliche Probleme stellen. In diesem Sinne werden im Folgenden zwei Probleme erörtert, die zwar in unterschiedlicher Weise mit diesen beiden Lösungsansätzen verbunden sind, die aber eines gemeinsam haben: In beiden Fällen geht es darum, menschliches Denken und Fühlen in systematischer Weise zu erfassen. Mit anderen Worten, in beiden Fällen handelt es sich um psychometrische Probleme.

Das erste dieser beiden psychometrischen Probleme ist die Erfassung gesundheitsbezogener Lebensqualität. Dieses Problem stellt sich unter Voraussetzung beider Lösungsansätze. Sowohl die Verteilung von Mitteln, als auch die Gestaltung des Gesundheitssystems setzen nämlich ein Optimierungskriterium voraus. In beiden Fällen muss es eine Zielgröße geben, die maximiert werden soll. In der Gesundheitsökonomie werden zu diesem Zweck vor allem zwei Größen verwendet: 1) die Lebensdauer und 2) die gesundheitsbezogene Lebensqualität. Um Mittel so verteilen zu können, dass die gesundheitsbezogene Lebensqualität maximiert wird, und um das Gesundheitssystem und die damit zusammenhängenden Teile der Gesellschaft so gestalten zu können, dass mit den vorhandenen Mittel möglichst viel gesundheitsbezogene Lebensqualität erreicht wird, muss diese Größe möglichst sinnvoll erfasst werden.

Das zweite psychometrische Problem, das hier betrachtet wird, ist die Erfassung der Ursachen gesundheitsbezogenen Verhaltens. Innerhalb der Gesundheitsökonomie wird dieses Problem bisher kaum diskutiert. Es stellt sich aber unter Voraussetzung des zweiten Lösungsansatzes. Um nämlich das Gesundheitssystem und die damit zusammenhängenden Teile der Gesellschaft bestmöglich in Hinblick auf eine effiziente Nutzung der verfügbaren Mittel zu gestalten, müssen die Parameter gezielt beeinflusst werden, von denen das Verhältnis zwischen Aufwand und Ergebnis abhängt. Einer dieser Parameter ist das Verhalten der Nutzer des Gesundheitswesens. Je mehr diese Personen aktiv an den medizinischen Therapien mitarbeiten und je mehr sie sich darum bemühen, Krankheit zu vermeiden, desto mehr Gesundheit wird das Gesundheitssystem bei gegebenem Budget

produzieren. Um dieses gesundheitsbezogene Verhalten möglichst gezielt beeinflussen zu können, müssen die Ursachen dieses Verhaltens möglichst wirklichkeitstreu erfasst werden.

Inwieweit es gelingt, zu einem sinnvollen Umgang mit der Begrenztheit der für das Gesundheitswesen verfügbaren Mittel zu kommen, hängt damit wesentlich von den Methoden ab, die zur Bearbeitung der beiden eben genannten psychometrischen Probleme zur Verfügung stehen. Das Problem, wie gesundheitsbezogene Lebensqualität erfasst werden könnte, ist innerhalb der Gesundheitsökonomie bereits sehr ausführlich diskutiert worden. Das Problem, wie die Ursachen gesundheitsbezogenen Verhaltens erfasst könnten, dagegen kaum. Überlegungen zur Weiterentwicklung der Methodologie müssen deshalb in beiden Fällen an anderer Stelle ansetzen. Bei der Erfassung der gesundheitsbezogenen Lebensqualität geht es darum, die bestehenden Methoden kritisch zu hinterfragen und gegebenenfalls weiteren Forschungsbedarf aufzuzeigen. Bei der Erfassung der Ursachen gesundheitsbezogenen Verhaltens geht es dagegen zunächst einmal darum, geeignete Methoden in die Gesundheitsökonomie einzuführen.

Im folgenden Beitrag werden deshalb als Erstes die wichtigsten der Verfahren, die zurzeit in der Gesundheitsökonomie zur Erfassung von Lebensqualität verwendet werden, kritisch diskutiert. Als Zweites werden zwei Ansätze aus der Psychologie vorgestellt, die in der Gesundheitsökonomie bisher noch weitgehend unbekannt sind, die aber dazu verwendet werden könnten, die Ursachen gesundheitsbezogenen Verhaltens zu erfassen. Als Drittes wird überlegt, in welcher Richtung und in welcher Weise die auf psychometrische Probleme bezogenen Methoden der Gesundheitsökonomie weiterentwickelt werden könnten.

## **1 Die Erfassung gesundheitsbezogener Lebensqualität**

Was unter gesundheitsbezogener Lebensqualität zu verstehen ist und wie diese Größe dementsprechend erfasst werden könnte, ist in der Gesundheitsökonomie bereits recht umfassend erörtert worden. Gesundheitsbezogene Lebensqualität wird dort meist mit der Bewertung von Krankheitszuständen, vornehmlich durch den durchschnittlichen Nutzer des Gesundheitswesens, gleichgesetzt. Vor allem zwei Arten von Verfahren werden verwendet, um Lebensqualität in diesem Sinne zu erfassen: 1) Verfahren zur direkten Bewertung von Krankheitszuständen und 2) Indexverfahren. Beide Arten von Verfahren werden jetzt nacheinander betrachtet.

### **1.1 Verfahren zur direkten Bewertung von Gesundheitszuständen**

Bei den Verfahren zur direkten Bewertung von Krankheitszuständen werden Menschen mehr oder minder direkt nach Bewertungen über den zu beurteilenden Gesundheitszustand gefragt. Da es sehr verschiedene Möglichkeiten gibt, Bewertungen auszudrücken, sind auch sehr verschiedene Verfahren zur Erfassung von Bewertungen denkbar. In der Gesundheitsökonomie werden vor allem drei Verfahren verwendet: 1) das Urteilsskalenverfahren (rating scale procedure), 2) das Standardspiel (standard gamble) und 3) das Zeitausgleichsverfahren (time trade-off) (vgl. Drummond et al., 1997, Kap. 6.3). Diese drei Verfahren werden im Folgenden zunächst dargestellt und dann kritisch gewürdigt.

#### *1.1.1 Darstellung*

Beim Urteilsskalenverfahren wird der befragten Person eine visuelle oder kategorielle Urteilskala vorgegeben. Die beiden Begrenzungen der Skala werden dabei meistens mit dem best und dem schlechtest denkbaren Gesundheitszustand identifiziert. Die befragte Person soll den zu beurteilenden Gesundheitszustand auf dieser Skala so einordnen, dass

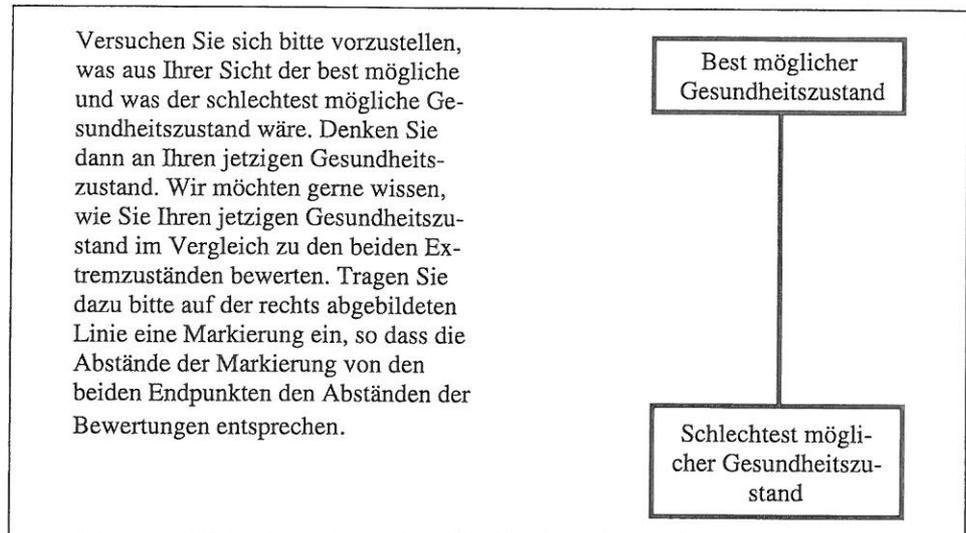


Abbildung 1: Bewertung des eigenen Gesundheitszustandes mit dem Urteilsskalenverfahren

dies den Abständen der Bewertungen der dazugehörigen Gesundheitszustände entspricht (siehe Abbildung 1). Die Bewertung des zu beurteilenden Gesundheitszustandes wird dann über die dazugehörige Position auf der Urteilsskala bestimmt. Dabei wird – meist implizit – die Annahme vorausgesetzt, dass die befragten Personen so etwas wie Abstände zwischen Bewertungen intern repräsentiert haben und auf der Urteilsskala unverzerrt zum Ausdruck bringen können. Die Urteile auf der Urteilsskala werden dann entsprechend als eine lineare Abbildung der tatsächlichen Bewertungen interpretiert. Beim Standardspiel wird die befragte Person mit zwei fiktiven Alternativen konfrontiert. Im einfachsten Fall bestehen diese Alternativen darin, entweder 1) in dem zu beurteilenden Gesundheitszustand zu leben oder 2) sich einem völlig schmerz- und kostenfreien Eingriff zu unterziehen, der mit einer bestimmten Wahrscheinlichkeit zu völliger Gesundheit und mit der entsprechenden Gegenwahrscheinlichkeit zum sofortigen Tod führt. Die befragte Person soll zwischen diesen beiden Alternativen wählen (siehe Abbildung 2).

Stellen Sie sich bitte vor, Sie würden vor die Wahl zwischen zwei Möglichkeiten gestellt.

- 1) Sie leben weiterhin in Ihrem jetzigen Gesundheitszustand.
- 2) Sie unterziehen sich einer völlig schmerz- und kostenfreien Operation, die zwei mögliche Ausgänge hat.  
Ausgang 1: völlige Gesundheit.  
Ausgang 2: Tod.  
Völlige Gesundheit tritt dabei in 95 % aller Fälle ein,  
der Tod in 5% aller Fälle.

Welche von beiden Möglichkeiten würden Sie wählen?

Abbildung 2: Bewertung des eigenen Gesundheitszustandes mit dem Standardspiel

Die Wahrscheinlichkeiten bei der zweiten Alternative werden so lange verändert, bis die befragte Person beide Alternativen für gleichwertig hält. Die Bewertung des zu beurteilenden Gesundheitszustandes wird dann auf der Grundlage der Wahrscheinlichkeiten bestimmt, bei denen Gleichwertigkeit besteht.

Bei der Auswertung des Standardspiels wird die Annahme vorausgesetzt, dass sich die befragten Personen gemäß dem Modell des subjektiv erwarteten Nutzens verhalten. Gemäß diesem Modell ergibt sich die Gesamtbewertung  $u$  einer Wahlmöglichkeit  $W$  mit mehreren unsicheren Ausgängen  $x_i$  als Summe der Produkte der Wahrscheinlichkeiten dieser Ausgänge  $p_i$  mit den dazugehörigen Bewertungen  $u(x_i)$ , d. h.

$$u(W) = \sum_{i=1}^n p_i * u(x_i). \quad (1)$$

Dabei ist  $n$  die Anzahl der unsicheren Ereignisse. Setzt man den Wert für den Tod per definitionem gleich null und den für die Gesundheit gleich eins, folgt bei Gleichwertigkeit beider Alternativen aus dem Modell des subjektiv erwarteten Nutzens, dass die Bewertung des zu beurteilenden Gesundheitszustand  $u(GZ)$  gleich der Wahrscheinlichkeit für die vollständige Gesundung  $p_{GES}$  ist, d. h.

$$u(GZ) = p_{GES}. \quad (2)$$

Beim Zeitausgleichsverfahren wird die befragte Person ebenfalls mit zwei Alternativen konfrontiert. Im einfachsten Fall bestehen diese beiden Alternativen darin, entweder 1) eine bestimmte Zeit in dem zu beurteilenden Gesundheitszustand zu leben und dann sofort zu sterben oder 2) eine etwas kürzere Zeit in völliger Gesundheit zu leben und dann sofort zu sterben. Die befragte Person soll zwischen diesen beiden Alternativen wählen (siehe Abbildung 3). Die Zeitdauer in der zweiten Alternative wird so lange verändert, bis die befragte Person beide Alternativen für gleichwertig hält. Die Bewertung des zu beurteilenden Gesundheitszustandes wird dann auf der Grundlage der beiden Zeitdauern bestimmt, bei denen Gleichwertigkeit besteht.

Bei der Auswertung des Zeitausgleichsverfahrens wird – eher implizit – die Annahme vorausgesetzt, dass sich die Bewertung  $u$  einer Folge  $F$  von Gesundheitszuständen  $x_i$  als Summe der Produkte aus den Zeitdauern der Zustände  $t_i$  mit deren Bewertungen  $u(x_i)$  ergibt, d. h.

$$u(F) = \sum_{i=1}^n t_i * u(x_i). \quad (3)$$

Dabei ist  $n$  die Anzahl der unterschiedlichen Zustände, die hintereinander folgen. Bei der einfachsten Version des Zeitausgleichsverfahrens ist  $n$  immer gleich eins. Setzt man wieder die Bewertung des Todes gleich null und die völliger Gesundheit gleich eins,

Stellen Sie sich bitte vor, Sie würden vor die Wahl zwischen zwei Möglichkeiten gestellt.

- 1) Sie leben von jetzt an 20 Jahre in Ihrem jetzigen Gesundheitszustand und sterben dann unmittelbar.
- 2) Sie leben von jetzt an 18 Jahre in völliger Gesundheit und sterben dann unmittelbar.

Welche von beiden Möglichkeiten würden Sie wählen?

Abbildung 3: Bewertung des eigenen Gesundheitszustandes mit dem Zeitausgleichsverfahren

folgt bei Gleichwertigkeit beider Alternativen aus dem eben beschriebenen Modell, dass die Bewertung des zu beurteilenden Gesundheitszustandes  $u(GZ)$  gleich der Zeitdauer für das Leben in völliger Gesundheit  $t_{GES}$  geteilt durch die Zeitdauer für den zu beurteilenden Gesundheitszustand  $t_{GZ}$  ist, d. h.

$$u(GZ) = t_{GES}/t_{GZ}. \quad (4)$$

Das in Gleichung 3 beschriebene theoretische Modell wird in den üblichen Darstellungen des Zeitausgleichsverfahrens nicht explizit genannt. Damit stellt sich die Frage, in welcher Weise es implizit die Voraussetzung für das Zeitausgleichsverfahren ist. Dies ergibt sich aus der explizit vorausgesetzten Annahme, dass die Bewertung eines Gesundheitszustandes, der eine bestimmte Zeit lang andauert,  $u(GZ, t)$ , gleich dem Produkt aus der Zeitdauer  $t$  und der zeitlich invarianten Bewertung dieses Zustandes,  $u(GZ)$ , ist. Damit gilt

$$u(GZ, t) = t * u(GZ). \quad (5)$$

Nun ist ein Zustand, der für die Zeit  $t$  dauert, nichts anderes als eine Folge  $F$  von  $n$  gleichartigen Zuständen, die jeweils die Zeit  $t/n$  dauern. Es sei  $x_i = GZ$  der  $i$ -te gleichartige Zustand in dieser Folge und  $t_i = t/n$  die Zeit, die dieser Zustand dauert. Dann gilt für die Folge  $F$  wegen Gleichung 5

$$u(F) = t * u(GZ) = n * (t/n) * u(GZ) = \sum_{i=1}^n t_i * u(x_i). \quad (6)$$

Letzteres ist die Annahme, von der hier behauptet wurde, dass sie beim Zeitausgleichsverfahren implizit vorausgesetzt wird.

### 1.1.2 Kritische Würdigung

Hinter jedem dieser Verfahren steht eine plausible Idee darüber, worin sich Bewertungen von Krankheitszuständen ausdrücken könnten. Sie sind damit alle als mögliche Ansätze zur direkten Bewertung von Krankheitszuständen ernst zu nehmen. Es gibt aber wenigstens zwei Probleme, die sich bei der Anwendung dieser Verfahren stellen. Das erste Problem ergibt sich dadurch, dass jedes der Verfahren auf bestimmten Annahmen über das menschliche Urteilsverhalten beruht. In dem Maße, wie diese Annahmen verletzt sind, ist fraglich, was die Ergebnisse bedeuten, die mit diesem Verfahren gewonnen werden. Das zweite Problem ergibt sich dadurch, dass die Verfahren sich auf unterschiedliche Formen menschlichen Urteilsverhaltens beziehen und damit auch auf inhaltlich unterschiedlichen theoretischen Annahmen beruhen. Damit stellt sich die Frage, inwieweit die drei Verfahren tatsächlich dasselbe erfassen. Im Folgenden werden beide Probleme etwas tiefer gehend erörtert.

Das Urteilsskalenverfahren beruht auf der Annahme, dass Menschen bezüglich der Größe, nach der gefragt wird, intern konsistent Abstände wahrnehmen und die Größe entsprechend diesen internen Abständen unverzerrt auf der Urteilsskala zum Ausdruck bringen können. Im Zusammenhang mit der Bewertung von Krankheitszuständen ist diese Annahme bisher kaum kritisch geprüft worden. Es gibt aber einige wenige Untersuchungen, in denen diese Annahmen bezüglich anderer Bewertungen geprüft worden sind. Diese Untersuchungen stammen alle aus den achtziger Jahren (Orth, 1982; Orth & Wegener, 1983; Westermann, 1984, 1985; Westermann & Hager, 1983, 1985). Als Ergebnis zeigt sich vor allem, dass große individuelle Unterschiede bestehen. Manche Personen scheinen konsistent Abstände zwischen Bewertungen beurteilen und auf der Urteilsskala zum Ausdruck bringen zu können, andere eher nicht.

Das Standardspiel beruht auf der Annahme, dass sich Menschen bei Entscheidungen gemäß dem Modell des subjektiv erwarteten Nutzens verhalten. Im Zusammenhang mit

Entscheidungen zwischen Krankheitszuständen ist diese Annahme bisher kaum kritisch geprüft worden. Wohl aber sehr ausführlich im Zusammenhang mit Wahlen zwischen monetären Lotterien. Hier zeigt sich unter anderem, dass Entscheidungen beim Standardspiel wesentlich davon abhängen, wie die Frage formuliert ist. Bei inhaltlich gleichen, aber unterschiedlich formulierten Alternativenpaaren entscheiden sich Menschen völlig unterschiedlich (Hershey, Kunreuther & Schoemaker, 1988). Des Weiteren verhalten sich Menschen bei Entscheidungen zwischen Lotterien, bei denen die möglichen Ausgänge mit sehr kleinen oder sehr großen Wahrscheinlichkeiten verbunden sind, kaum gemäß dem Modell des subjektiv erwarteten Nutzens. Wenn die möglichen Ausgänge mit mittleren Wahrscheinlichkeiten verbunden sind, ist dies dagegen eher der Fall. Menschen scheinen also große Schwierigkeiten beim rationalen Umgang mit extrem kleinen und extrem großen Wahrscheinlichkeiten zu haben (Kahneman & Tversky, 1979).

Das Zeitausgleichsverfahren beruht auf der Annahme, dass sich die Gesamtbewertung einer Folge von Gesundheitszuständen als Summe der Produkte aus den Bewertungen der einzelnen Gesundheitszustände mit den dazugehörigen Zeitdauern ergibt. Auch diese Annahme ist bisher nur wenig empirisch geprüft worden. Es gibt hier allerdings eine wichtige Untersuchung von Dolan (1996). In dieser Untersuchung wurden dieselben Gesundheitszustände einmal mit der Dauer von einem Monat, das zweite Mal mit der Dauer von einem Jahr und ein drittes Mal mit der Dauer von zehn Jahren mit Hilfe des Urteilsskalenverfahrens bewertet. Unter Voraussetzung der hier vorgestellten Annahmen für das Urteilsskalen- und das Zeitausgleichsverfahren müssten die Verhältnisse zwischen den Abständen der Bewertungen der verschiedenen Zeitdauern innerhalb der drei Zeitdauergruppen gleich sein. Dolan fand aber, dass bei längerer Dauer die schwereren Krankheitszustände im Vergleich zu den leichteren Krankheitszuständen schlechter bewertet werden als bei kürzerer Dauer. Darüber hinaus zeigt sich in anderen empirischen Untersuchungen, dass die Bevorzugungen zwischen Folgen von Zuständen von der Reihenfolge und der zeitlichen Entferntheit der Zustände abzuhängen (vgl. Jungermann, Pfister & Fischer, 1998, Kap. 3.4). Auch dies steht im Widerspruch zu der Annahme, auf der das Zeitausgleichsverfahren beruht.

Die Verfahren zur direkten Bewertung von Gesundheitszuständen beruhen, wie bereits gesagt, auf recht unterschiedlichen theoretischen Konzeptionen. Der wichtigste Unterschied besteht in dem erkenntnistheoretischen Prinzip, nach dem der Bewertungsbegriff festgelegt wird. Beim Urteilsskalenverfahren geschieht dies durch Rückgriff auf das Sprachverständnis der Befragten. Sie werden direkt nach Bewertungen gefragt. Was sie dabei aber tatsächlich unter der Bezeichnung „Bewertung“ verstehen, bleibt bis zu einem gewissen Grade offen. Infolgedessen ist auch die Bedeutung der mit diesem Verfahren gewonnenen Ergebnisse unklar. Beim Standardspiel und beim Zeitausgleichsverfahren wird dagegen überhaupt nicht nach Bewertungen gefragt, sondern lediglich nach Bevorzugungen zwischen zwei Alternativen. Der Begriff der Bewertung wird dann implizit durch die Modellannahmen festgelegt, nach denen diese Bevorzugungen interpretiert werden.

Der Unterschied zwischen dem Urteilsskalenverfahren auf der einen Seite und dem Standardspiel und dem Zeitausgleichsverfahren auf der anderen ist auch aus ökonomischer Sicht mehrfach problematisiert worden (siehe u. a. Green, Brazier & Deverill, 2000). Aus dieser Sicht äußert sich der Wert eines Gutes wesentlich darin, inwieweit man bereit ist, für dieses Gut etwas einzutauschen. Das Standardspiel und das Zeitausgleichsverfahren beruhen direkt auf dem Gedanken des Eintauschens, das Urteilsverfahren nicht. Damit wäre eher vom Standardspiel und vom Zeitausgleichsverfahren zu erwarten, dass sie Bewertungen in dem Sinne erfassen, wie sie in der Ökonomie verstanden werden.

Im Vergleich zum Urteilsskalenverfahren haben das Standardspiel und das Zeitausgleichsverfahren zwar einige wichtige Gemeinsamkeiten. Es bestehen aber auch Unterschiede. Beim Standardspiel geht es um Entscheidungen zwischen Wahlmöglichkeiten mit

unsicheren Ausgängen, beim Zeitausgleichsverfahren um Wahlmöglichkeiten zwischen zeitlichen Folgen sicherer Ereignisse. Das erstgenannte Verfahren beruht also auf Entscheidungen unter Unsicherheit, das zweite auf Entscheidungen unter Sicherheit. Manche Ökonomen (u. a. Drummond et al., 1997; Gold et al. 1996; Mehrez & Gafni, 1991) weisen diesem Unterschied eine große Bedeutung zu. Aus ihrer Sicht bilden Entscheidungen unter Unsicherheit deshalb eine validere Grundlage für die Messung von Bewertungen, weil unsere ganze Welt voller Unsicherheit sei. Bewertungen, die über Entscheidungen unter Unsicherheit bestimmt worden sind, werden dort oft auch als „utilities“ bezeichnet, Bewertungen, die auf Urteilen unter Sicherheit beruhen, dagegen als „values“.

Es gibt mittlerweile auch eine Vielzahl empirischer Untersuchungen zur Beziehung zwischen den drei hier diskutierten Verfahren. Auch die Ergebnisse dieser Untersuchungen deuten darauf hin, dass zwischen dem Urteilsskalenverfahren auf der einen Seite und den anderen beiden Verfahren auf der anderen der größte Unterschied besteht. So korrelieren Standardspiel- und Zeitausgleichswerte meistens hoch miteinander (vgl. Green, Brazier & Deverill, 2000). Urteilsskalenwerte scheinen dagegen nicht in einem linearen Zusammenhang mit Werten des Standardspiels und des Zeitausgleichsverfahrens zu stehen (Krabbe, Essink-Bot & Bonsel, 1999). Dolan und Sutton (1997) fanden allerdings, dass die Art des Zusammenhangs zwischen Urteilsskalenwerten und den Werten der anderen beiden Verfahren von der Schwere der zu beurteilenden Krankheitszustände abhängt. Green et al. (2000) schließen aufgrund einer sehr umfassenden Durchsicht der empirischen Befunde zu diesen drei Verfahren, dass Urteilsskalenwerte einen anderen Aspekt gesundheitsbezogener Lebensqualität widerspiegeln, als Werte, die mit dem Standardspiel oder dem Zeitausgleichsverfahren bestimmt worden sind.

## 1.2 Indexmessung

### 1.2.1 Darstellung

Wenn man mit Hilfe der direkten Bewertungsverfahren festzustellen versucht, wie Menschen den Gesundheitszustand bewerten, in dem sie sich gerade befinden, stellt sich abgesehen von den eben diskutierten Problem noch ein weiteres: der Gegenstand der Bewertung ist wegen der Art, in der die Bewertungen erhoben werden, äußerst schlecht definiert. Jeder Gesundheitszustand ist nämlich ein äußerst komplexes Gebilde, das durch sehr viele verschiedene Aspekte bestimmt ist. Es ist aber nicht zu erwarten, dass jemand zugleich an alle diese Aspekte denkt, wenn er zu seinem Gesundheitszustand befragt wird. Laut Gedächtnisforschung (Miller, 1956) können Menschen nur etwa 5 bis 9 Einheiten zur selben Zeit im Kurzzeitgedächtnis repräsentiert haben. Des Weiteren ist damit zu rechnen, dass unterschiedliche Menschen, die sich in demselben Gesundheitszustand befinden, bei der Bewertungsmessung an unterschiedliche Aspekte ihres Gesundheitszustandes denken. Außerdem ist zu erwarten, dass noch völlig andere Faktoren als der eigentliche Gesundheitszustand auf etwaige Bewertungsurteile wirken. Der wichtigste Faktor dürfte dabei das Ausmaß sein, in dem die befragten Menschen gelernt haben, mit ihrem Gesundheitszustand umzugehen.

Bewertungen, die mit direkten Bewertungsverfahren erhoben werden, spiegeln damit also vieles wider, was im Sinne des gesundheitsökonomischen Verständnisses gesundheitsbezogener Lebensqualität irrelevant ist. Als Reaktion auf diese Problematik ist ein völlig anderer Ansatz zur Bewertung von Gesundheitszuständen entwickelt worden: der Ansatz der Indexmessung. Dieser Ansatz beruht darauf, dass der zu bewertende Gesundheitszustand erst systematisch klassifiziert wird und dass dann der dabei resultierenden Klassifizierung nach einer fest vorgegebenen Zuordnungsregel ein Wert zugeordnet wird. Dieser Wert wird als Index bezeichnet. Er wird als Messwert für die Lebensqualität betrachtet,

die mit dem zu bewertenden Gesundheitszustand verbunden ist. Ein Indexmessverfahren besteht dementsprechend aus zwei Komponenten: 1) aus einer Sammlung von gesundheitsrelevanten Attributen zur Klassifikation von Krankheitszuständen und 2) einer Zuordnungsfunktion, mit der jeder möglichen Klassifizierung ein Indexwert zugeordnet wird.

Das am weitesten verbreitete und am umfassendsten beforschte Indexverfahren zur Messung gesundheitsbezogener Lebensqualität ist der EuroQol (vgl. Brooks et al., 1996). Der EuroQol wird üblicherweise als Fragebogen vorgegeben, auf dem die Betroffenen ihren gegenwärtigen Gesundheitszustand selbst einstufen sollen. Der EuroQol enthält dabei jeweils eine Frage zu fünf lebensqualitätsrelevanten Attributen. Diese Attribute sind „Mobilität“, „Fähigkeit zur Körperpflege“, „Fähigkeit zu allgemeinen Tätigkeiten“, „Schmerzen/körperliche Beschwerden“ und „Ängstlichkeit/Niedergeschlagenheit“. Bei jedem Attribut werden drei Antwortmöglichkeiten vorgegeben: „keine Probleme“, „einige Probleme“ und „extreme Probleme“. Jedes Antwortmuster eines vollständig ausgefüllten Fragebogens bildet eine Klassifizierung eines Gesundheitszustandes. Den EuroQol gibt es mittlerweile in zahlreichen Übersetzungen. Für einen großen Teil dieser Übersetzungen werden Zuordnungsfunktionen angeboten, mit denen jeder Klassifizierung ein Index zugewiesen werden kann.

Die fünf Attribute des EuroQol sind von einem Wissenschaftlergremium aus mehreren europäischen Ländern über Konsensfindung ausgewählt worden. Die Zuordnungsfunktionen für die verschiedenen sprachspezifischen Versionen sind dagegen meistens empirisch bestimmt worden. In den meisten Fällen wurde dabei dieselbe Strategie angewendet. Diese Strategie ist dadurch motiviert, dass es zu aufwändig wäre, alle 243 beim EuroQol möglichen Klassifizierungen empirisch bewerten zu lassen. Aus diesem Grunde werden lediglich einige ausgewählte Klassifizierungen mit Hilfe eines oder mehrerer der oben beschriebenen direkten Bewertungsverfahren bewertet. Die Indizes für alle möglichen Klassifizierungen werden dann aufgrund der vorliegenden Daten mit Hilfe mathematischer Modellannahmen bestimmt. Die umfassendste Studie dieser Art ist die Normierungsstudie zur englischen Version des EuroQol (Dolan, 1997). Eine Studie zur deutschen Version stammt von Schulenburg et al. (1998).

### 1.2.2 Kritische Würdigung

Genau wie bei den direkten Bewertungsverfahren gibt es auch bei den Indexverfahren Probleme, die die Aussagekraft dieser Verfahren gefährden. Eines dieser Probleme ergibt sich im Zusammenhang mit der Auswahl der Attribute, nach denen die Krankheitszustände klassifiziert werden. Durch die Auswahl der Attribute wird eine ganz bestimmte Sicht auf Gesundheitszustände definiert. Je nach Sicht können sich aber ganz verschiedene Rangreihen bei der Bewertung von Krankheitszuständen ergeben. Damit stellt sich die Frage nach der Auswahl der Attribute und nach den Kriterien der Auswahl. In der gegenwärtigen gesundheitsökonomischen Literatur wird dieses Problem bisher kaum diskutiert.

Weitere Probleme ergeben sich im Zusammenhang mit der Bestimmung der Funktion, nach der den Klassifizierungen der Krankheitszustände Bewertungen zugeordnet werden. Hier stellt sich zunächst einmal die Frage, mit welchen direkten Bewertungsverfahren sich diese Klassifizierungen in welcher Weise am sinnvollsten bewerten lassen. Mit anderen Worten, hier stellen sich all die Probleme, die bereits im Zusammenhang mit den direkten Bewertungsverfahren erörtert worden sind. Des Weiteren stellt sich aber auch die Frage, in welcher Weise man aufgrund der Bewertungen einer kleinen Auswahl der Klassifizierungen, die zu einem Indexverfahren gehören, empirisch valide Bewertungen für alle Klassifizierungen bestimmen kann, die zu diesem Verfahren gehören. In jedem Fall ist es dazu notwendig, die fehlende empirische Information durch theoretische An-

nahmen über das menschliche Urteilsverhalten zu ersetzen. Welche Annahmen dabei getroffen werden, kann große Auswirkungen auf die Gestalt der Zuordnungsfunktion haben. Einige Aspekte dieser Problematik sind bereits in der gesundheitsökonomischen Literatur diskutiert worden (Brooks et al. 1996; Busschbach, McDonnel, Essink-Bot & van Hout, 1999; Froberg & Kane, 1989 a,b; Kerim-Sade & Wasem, 1998). In vieler Hinsicht dürften hier aber noch tiefergehende Betrachtungen notwendig sein.

## 2 Die Erfassung der Ursachen gesundheitsbezogenen Verhaltens

Wenn es nicht nur darum gehen soll, die für das Gesundheitswesen zu Verfügung stehenden Mittel sinnvoll zu verteilen, sondern auch darum, das Gesundheitswesen und die damit zusammenhängenden Bereiche der Gesellschaft gezielt zu gestalten, stellt sich das Problem, gezielt auf das gesundheitsbezogene Verhalten der Nutzer des Gesundheitswesens einzuwirken. Daraus wiederum ergibt sich das Problem, die Ursachen dieses Verhaltens systematisch zu erfassen. In der Gesundheitsökonomie ist dieses Problem bisher kaum diskutiert worden. Damit stellt sich die Frage, ob nicht bereits an anderer Stelle Ansätze existieren, die zur Bearbeitung dieses Problems genutzt werden könnten. Der Bereich außerhalb der Gesundheitsökonomie, der dabei als nächstes in Betracht kommt, ist die Psychologie. In dieser Disziplin gibt es eine Vielzahl theoretischer Ansätze, die die grundlegenden Prinzipien der Steuerung menschlichen Verhaltens zum Gegenstand haben. Des Weiteren gibt es auch einige Ansätze, bei denen es speziell um menschliches Gesundheitsverhalten geht (vgl. Armitage & Conner, 2000). Empirische Gültigkeit vorausgesetzt, liefern diese Ansätze theoretische Grundlagen für psychometrische Verfahren, mit denen sich die Ursachen für ein bestimmtes Verhalten erfassen lassen.

Eine vollständige Diskussion all dieser Ansätze würde hier zu weit führen. Deshalb werden nur die beiden Ansätze eingehender erörtert, die in diesem Zusammenhang in der Psychologie die meiste Akzeptanz genießen. Es handelt sich dabei um 1) die Theorie des überlegten Handelns (theory of reasoned action) von Fishbein und Ajzen (1975; Ajzen & Fishbein, 1980) und 2) die Theorie des geplanten Verhaltens (theory of planned behavior) von Ajzen (1985, 1991). Beide Ansätze beruhen auf der Vorstellung, dass sich Menschen mehr oder minder bewusst mit den Vor- und Nachteilen möglicher Verhaltensweisen auseinandersetzen und sich dann gemäß diesen Überlegungen verhalten. In den beiden Ansätzen wird in verschiedener Weise konkretisiert, welche Überlegungen in welcher Weise das Verhalten bestimmen. Dabei bezieht sich die Theorie des überlegten Handelns vor allem auf solche Verhaltensweisen, bei denen es allein von der handelnden Person abhängt, ob sie diese Verhaltensweise ausführt oder nicht. Die Theorie des geplanten Verhaltens bezieht sich dagegen auch auf solche Verhaltensweisen, die nicht vollständig unter willentlicher Kontrolle stehen. Die Theorie des geplanten Verhaltens ist damit eine Verallgemeinerung der Theorie des überlegten Handelns.

In der Theorie des überlegten Handelns wird das Verhalten in drei Schritten auf kausal immer weiter vorgelagerte subjektive Urteile zurückgeführt (siehe Abbildung 4). Die direkte Ursache für das Verhalten ist dabei, die Absicht das Verhalten auszuführen. Diese Absicht wiederum wird gemäß der Theorie durch zwei Größen bestimmt: 1) die Einstellung zum Verhalten und 2) die subjektive Norm. Die Einstellung zum Verhalten ist dabei eine Bewertung des Verhaltens unter dem Aspekt des persönlichen Nutzens. Die subjektive Norm dagegen ist eine Einschätzung des gesammelten sozialen Druckes, das Verhalten auszuführen bzw. nicht auszuführen. Sowohl die Einstellung als auch die subjektive Norm werden weiter auf jeweils zwei unterschiedliche Arten subjektiver Urteile zurückgeführt. Bei der Einstellung sind das 1) Erwartungen über mögliche Konsequenzen des Verhaltens (Konsequenzerwartungen) und 2) Bewertungen dieser Konsequenzen (Konsequenzbewertungen). Bei der subjektiven Norm sind das 1) Erwartungen über Normen

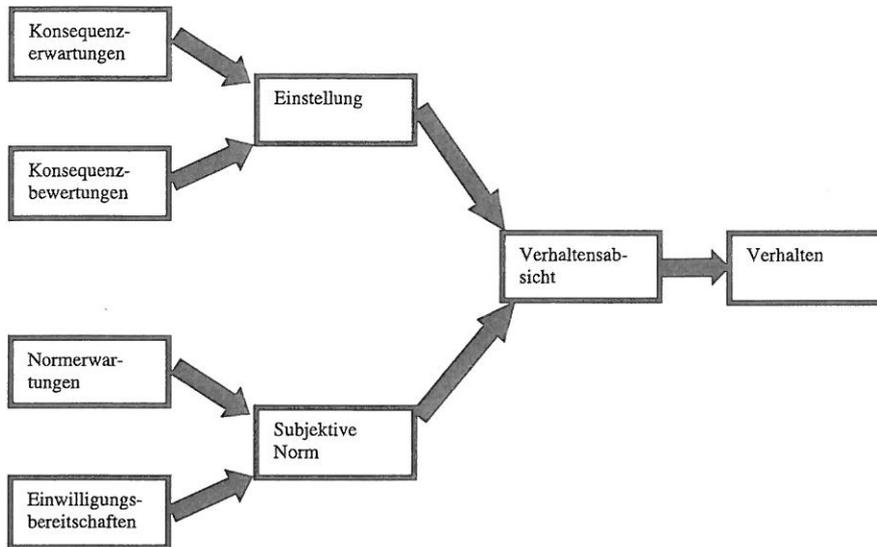


Abbildung 4: Die Theorie des überlegten Handelns von Fishbein und Ajzen

anderer Personen im näheren sozialen Umfeld (Normerwartungen) und 2) die Bereitschaften, den Normen dieser Personen nachzukommen (Einwilligungsbereitschaften).

Unter Verwendung dieser Theorie lässt sich zum Beispiel eine Studie zu der Frage planen und auswerten, warum manche Menschen Sport treiben und andere nicht. Man würde dazu zunächst einmal in Vorstudien bestimmen, an welche Konsequenzen die Befragten im Zusammenhang mit Sporttreiben denken und welche Personen zu ihrem näheren sozialen Umfeld gehören. Auf der Grundlage der Ergebnisse der Vorstudien würde man einen Fragebogen konstruieren, in dem systematisch die Erwartungen und Bewertungen zu den Konsequenzen und die Normerwartungen und Einwilligungsbereitschaften zu den wichtigen Bezugspersonen erfragt werden. Eine Befragung mit diesem Fragenbogen würde ein verhältnismäßig umfassendes Bild der internen subjektiven Urteile liefern, von denen es abhängt, ob Menschen Sport treiben. Mit Hilfe regressionsanalytischer Methoden ließe sich weiter eingrenzen, welche Erwartungen, Bewertungen, Normerwartungen und Einwilligungsbereitschaften in welcher Weise und in welchem Ausmaß über die Einstellung und die subjektive Norm auf die Verhaltensabsicht und damit auf das Verhalten einwirken. Würde man dabei feststellen, dass die Personen, die nicht Sport treiben, falsche Erwartungen über die Auswirkungen von Sport auf ihre Gesundheit haben, wäre hier der Ansatz für geeignete Aufklärungsmaßnahmen gegeben.

Die Theorie des geplanten Verhaltens ist als eine Erweiterung der Theorie des vernünftigen Handelns entstanden. Im Unterschied zu ihrer Vorgängerin zielt sie ausdrücklich darauf ab, auch solche Verhaltensweisen vorherzusagen und zu erklären, die zumindest zum Teil durch äußere Faktoren bestimmt sind. Zu diesem Zweck wird zusätzlich zu den Variablen der Theorie des überlegten Handelns eine weitere Variable hinzugefügt: die subjektive Verhaltenskontrolle (siehe Abbildung 5). Diese Variable spiegelt das Ausmaß wider, in dem die handelnde Person glaubt, das Verhalten willentlich steuern zu können. Gemäß der Theorie des geplanten Verhaltens wirkt die subjektive Verhaltenskontrolle zusammen mit der Einstellung und der subjektiven Norm auf die Verhaltensabsicht. Des Weiteren trägt sie unter bestimmten Voraussetzungen zusätzlich zur Verhaltensabsicht zur Vorhersage des Verhaltens bei. Genau wie die Einstellung und die

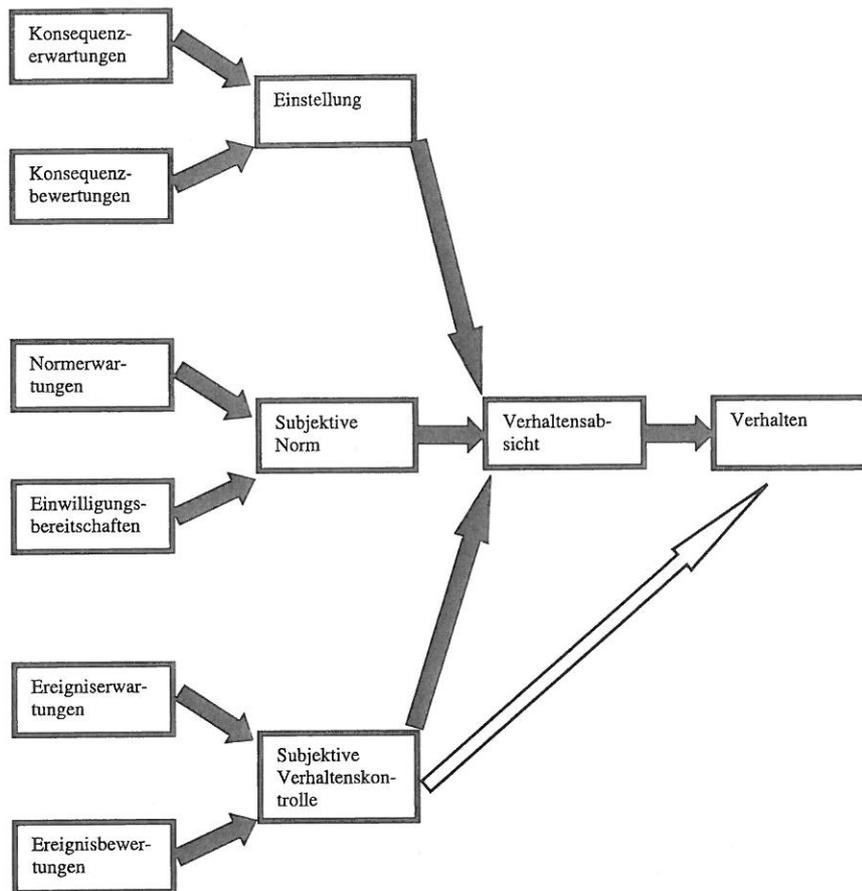


Abbildung 5: Die Theorie des geplanten Verhaltens von Ajzen

subjektive Norm wird auch die subjektive Verhaltenskontrolle wiederum auf zwei Arten kausal weiter vorgelagerter subjektiver Urteile zurückgeführt. Bei der subjektiven Verhaltenskontrolle sind dies 1) Einschätzungen über die Eintretenswahrscheinlichkeiten von Ereignissen, die das Verhalten behindern oder fördern könnten, (Ereigniserwartungen) und 2) Beurteilungen der Hinderlichkeit bzw. Förderlichkeit dieser Ereignisse (Ereignisbewertungen).

Auch diese Theorie kann als Vorlage zu einer Fragebogenstudie zu der Frage genutzt werden, warum manche Menschen Sport treiben und andere nicht. Im Wesentlichen entspricht das Vorgehen dabei dem Vorgehen bei der Theorie des vernünftigen Verhaltens. Es werden lediglich zusätzlich noch Fragen zur subjektiven Verhaltenskontrolle gestellt und – außerdem – Fragen zu den Ereignissen, die gemäß der Theorie die Verhaltenskontrolle bestimmen. Diese Fragen erweitern die Ansatzmöglichkeiten für Maßnahmen zur Beeinflussung des Verhaltens. So könnte sich ergeben, dass viele Menschen deshalb keinen Sport treiben, weil sie dort, wo sie wohnen, nur unter Schwierigkeiten Zugang zu den für sie interessanten Sportmöglichkeiten haben. In diesem Fall wäre über geeignete Veränderungen der Infrastruktur nachzudenken.

Tatsächlich sind sowohl die Theorie des vernünftigen Handelns, als auch die Theorie des geplanten Verhaltens in vielfältiger Form dazu genutzt worden, um Fragebogenstudien in

der eben beschriebenen Form zu planen und auszuwerten. Dabei gibt es auch viele Anwendungen im Bereich des gesundheitsrelevanten Verhaltens. So liegen mittlerweile mindestens 96 Datensätze zur Kondomnutzung (vgl. Albarracin et al., 2001) und mindestens 31 Studien zum Sporttreiben (vgl. Hausenblas, Carron & Mack, 1997) vor.

### 3 Weiterführende Überlegungen

In diesem Beitrag sind zwei verschiedene psychometrische Probleme betrachtet worden, die sich in unterschiedlicher Form innerhalb der Gesundheitsökonomie stellen: die Erfassung gesundheitsbezogener Lebensqualität und die Erfassung der Ursachen gesundheitsbezogenen Verhaltens. Gesundheitsbezogene Lebensqualität zu erfassen, ist sowohl in Hinblick auf eine möglichst sinnvolle Verteilung der für das Gesundheitswesen zur Verfügung stehenden Mittel notwendig, als auch in Hinblick auf eine gezielte Gestaltung des Gesundheitswesens. Die Ursachen gesundheitsbezogenen Verhaltens zu erfassen, ist vor allem in Hinblick auf eine gezielte Gestaltung des Gesundheitswesens von Interesse.

Die wichtigsten Methoden, die zurzeit in der Gesundheitsökonomie zur Erfassung gesundheitsbezogener Lebensqualität verwendet werden, sind kritisch diskutiert worden. Dabei hat sich gezeigt, dass in verschiedener Hinsicht noch Bedarf an weiterer Forschung besteht. Bei dieser weiteren Forschung wird ein ständiger Austausch zwischen normativer Setzung und erfahrungswissenschaftlicher Forschung stattfinden müssen. So wäre über normative Setzungen immer weiter zu spezifizieren, welche Form von Bewertung als Kriterium für die Verteilung der Mittel und die Gestaltung des Gesundheitssystems dienen soll. Die erfahrungswissenschaftliche Forschung hätte dann herauszufinden, wie die verschiedenen Verfahren zu gestalten, anzuwenden und auszuwerten sind, so dass die resultierenden Werte tatsächlich Bewertungen in dem über die normative Setzung festgelegten Sinn widerspiegeln.

Bei der Forschung zu den direkten Bewertungsverfahren könnte beispielsweise durch normative Setzung der Bewertungsbegriff zum Kriterium erklärt werden, der dem Modell des subjektiv erwarteten Nutzens und damit dem Standardspiel zugrunde liegt. Die erfahrungswissenschaftliche Forschung hätte dann herauszufinden, unter welchen Umständen welche Personen sinnvoll mit dem Standardspiel umgehen können. Dies ließe sich durch Rückgriff auf die Axiome entscheiden, die hinreichend und notwendig für die Gültigkeit des Modells des subjektiv erwarteten Nutzens sind (vgl. Jungermann, Pfister & Fischer, 1998; Kap. 6.1.2). Wenn die mit dem Standardspiel getroffenen Entscheidungen in Einklang mit den zugrundeliegenden Axiomen stehen, spricht das dafür, dass die Person das Standardspiel verstanden hat. Abweichungen von den Axiomen sprechen eher dagegen. Anhand dieses Kriteriums ließe sich eingrenzen, welche speziellen Variationen bei der Anwendung des Standardspiels sinnvoller sind und welche weniger. Die Standardspielwerte, die unter empirischer Gültigkeit der Axiome bestimmt worden sind, könnten dann als Referenzwerte für die Forschung zu den anderen direkten Bewertungsverfahren verwendet werden.

Bei der Forschung zu den Indexverfahren besteht das grundlegendste Problem in der Auswahl der Attribute, mit denen die Krankheitszustände klassifiziert werden. Auch hier sind zunächst normative Setzungen unumgänglich. Eine mögliche normative Setzung wäre, dass die Attribute möglichst die gesundheitsrelevanten Dimensionen erfassen sollten, die sich am meisten auf die subjektiv empfundene Lebensqualität auswirken. Welche Auswahl von Dimensionen diesem Kriterium entspricht, ließe sich dann weitgehend empirisch bestimmen. Weitere Forschungsfragestellungen ergeben sich aus dem Verfahren, mit dem üblicherweise die Zuordnungsfunktionen für Indexmessverfahren bestimmt werden. Gemäß diesem Verfahren werden zunächst einige der Klassifikationen, die mit

den vorliegenden Dimensionen gebildet werden können, empirisch bewertet und im Anschluss daran wird aus den empirisch gewonnenen Daten eine Zuordnungsfunktion für alle Klassifikationen geschätzt. Hier stellt sich die Frage, welche Klassifikationen ausgewählt werden sollten und nach welchen Prinzipien aufgrund der empirisch vorliegenden Daten eine Zuordnungsfunktion für alle Klassifikationen bestimmt werden kann. Um diese Fragen zu beantworten, ist noch Grundlagenforschung in verschiedener Hinsicht notwendig.

Wie die Ursachen gesundheitsbezogenen Verhaltens erfasst werden könnten, wird innerhalb der Gesundheitsökonomie bisher kaum diskutiert. Aus diesem Grund sind hier zwei Ansätze vorgestellt worden, die in der Psychologie dazu verwendet werden, Verhalten vorherzusagen und zu erklären, und die auch in der Gesundheitsökonomie zu diesem Zweck verwendet werden könnten. Als Nächstes wäre zu klären, in welcher Weise diese oder ähnliche Ansätze in das Methodeninventar der Gesundheitsökonomie integriert werden könnten. Auf der einen Seite werden Strategien entwickelt werden müssen, nach denen aus Erkenntnissen über die Ursachen gesundheitsbezogenen Verhaltens Empfehlungen für die Gestaltung des Gesundheitswesens abgeleitet werden können. Auf der anderen Seite werden die Methoden zur Erfassung der Ursachen gesundheitsbezogenen Verhaltens dahingehend verfeinert werden müssen, dass sie die für die Gestaltung des Gesundheitswesens benötigten Erkenntnisse liefern.

Die Theorien des vernünftigen Handelns und des geplanten Verhaltens bilden hier sicherlich einen guten Ansatzpunkt. Im Zusammenhang mit den speziellen Anforderungen der Gesundheitsökonomie könnten hier aber durchaus noch spezielle Weiterentwicklungen notwendig sein. Dies könnte sowohl die Auswahl der Variablen betreffen, die als mögliche Ursachen des Verhaltens in Betracht gezogen werden (vgl. Conner & Armitage, 1998), als auch die mathematische Modellierung der Zusammenhänge zwischen diesen Variablen (vgl. Konerding, 1999, 2001).

Sowohl Kenntnisse über die gesundheitsbezogene Lebensqualität, als auch Kenntnisse über die Ursachen gesundheitsbezogenen Verhaltens sind notwendig, um mit den vorhandenen finanziellen Mitteln ein möglichst bedarfsgerechtes Gesundheitswesen zu verwirklichen. Je besser die Kenntnisse über die Determinanten der gesundheitsbezogenen Lebensqualität sind, desto besser lassen sich sinnvolle konkrete Zielvorgaben für das Gesundheitssystem bestimmen. Je besser die Kenntnisse über die Ursachen gesundheitsbezogenen Verhaltens sind, desto besser lassen sich Wege finden, auf denen diese Zielvorgaben auf möglichst kostengünstige und angenehme Weise erreicht werden. Inwiefern aber derartige Kenntnisse gewonnen werden können, hängt wesentlich von den dazu verwendeten Methoden ab. Aus diesem Grunde sollten diese Methoden innerhalb der Gesundheitsökonomie immer wieder erörtert und weiterentwickelt werden.

### Literatur

- AJZEN, I. (1985): From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Hrsg.), *Action-control: From cognition to behavior* (S. 11–39). Heidelberg: Springer.
- AJZEN, I. (1991): The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, **50**, 179–211.
- AJZEN, I. & FISHBEIN, M. (1980): *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- ALBARRACIN, D., JOHNSON, B. T., FISHBEIN, M. & MUELLERLEILE, P. M. (2001): Theories of reasoned action and planned behavior as models of condom use: A Meta-Analysis. *Psychological Bulletin*, **127**, 142–161.
- ARMITAGE, C. J. & CONNER, M. (2000): Social cognition models and health behaviour: A structured review. *Psychology and Health*, **15**, 173–189.
- BROOKS, R. & the EuroQol Group (1996): EuroQol: the current state of play. *Health Policy*, **37**, 53–72.

- BUSSCHBACH, J. J. V., McDONNELL, J., & ESSINK-BOT, M. L., & HOUT, B. A. (1999): Estimating parametric relationships between health description and health valuation with an application to the EuroQol EQ-5D. *Journal of Health Economics*, **18**, 551–571.
- CONNER, M. & ARMITAGE, C. J. (1998): Extending the theory of planned behavior: A review and avenues for future research. *Journal of Applied Social Psychology*, **28**, 1429–1464.
- DOLAN, P. (1996): Modeling valuations for health states: the effect of duration. *Health Policy*, **38**(3), 189–203.
- DOLAN, P. (1997): Modeling valuations for EuroQol health states. *Medical Care*, **35**, 1095–1108.
- DOLAN, P. & SUTTON, M. (1997): Mapping visual analogue scale health state valuations onto standard gamble and time trade-off values. *Social Science in Medicine*, **44**, 1519–1530.
- DRUMMOND, M. F., O'BRIEN, B., STODDART, G. L. & TORRANCE, G. W. (1997): *Methods for the Economic Evaluation of Health Care Programmes*. Oxford, UK: Oxford University Press.
- FISHBEIN, M. & AJZEN, I. (1975): *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, Mass.: Addison – Wesley.
- FROBERG, D. G. & KANE, R. L. (1989, a): Methodology for measuring health-state preferences-I: Measurement strategies. *Journal of Clinical Epidemiology*, **42**, 345–354.
- FROBERG, D. G. & KANE, R. L. (1989, b): Methodology for measuring health-state preferences-III: Population and context effects. *Journal of Clinical Epidemiology*, **42**, 585–592.
- GOLD, M. R., SIEGEL, J. E., RUSSELL, L. B. & WEINSTEIN, M. C. (1996): *Cost-effectiveness in health and medicine*. New York: Oxford University Press.
- GREEN, C., BRAZIER, J., & DEVERILL, M. (2000): Valuing health-related quality of life: A review of health state valuation techniques. *Pharmacoeconomics*, **17**, 151–165.
- HAUSENBLAS, H. A., CARRON, A. V. & MACK, D. E. (1997): Application of the theories of reasoned action and planned behavior: A meta-analysis. *Journal of Sport and Exercise Psychology*, **19**, 36–51.
- HERSHEY, J. C., KUNREUTHER, H. C. & SCHOEMAKER, P. J. H. (1988): Sources of bias in assessment procedures for utility functions. In D. E. Bell, H. Raiffa & A. Tversky (Hrsg.), *Decision making: Descriptive, normative, and prescriptive interactions* (S. 422–442): Cambridge: Cambridge University Press.
- JUNGERMANN, H., PFISTER, H. R. & FISCHER, K. (1998): *Die Psychologie der Entscheidung*. Heidelberg, Berlin: Spektrum Akademischer Verlag.
- KAHNEMAN, D. & TVERSKY, A. (1979): A prospect theory: An analysis of decision under risk. *Econometrica*, **47**, 263–291.
- KERIM-SADE, C. & WASEM, J. (2000): An external control of validity of the German EuroQol-5D questionnaire. In X. Badia, M. Herdman & M. Roset (Hrsg.), *EuroQol: An instrument to value health: 16th plenary meeting of the EuroQol Group: Discussion papers* (S. 189–203): Sitges, Spanien: Institut de Salut Publica de Catalunya, Universitat de Barcelona.
- KONERDING, U. (1999): Formal models for predicting behavioral intentions in dichotomous choice situations. *Methods of Psychological Research*, **4**(2), 1–32. <http://www.mpr-online.de/>.
- KONERDING, U. (2001): Theory and methods for analyzing relations between behavioral intentions, behavioral expectations and behavioral probabilities. *Methods of Psychological Research*, **6**(1), 21–66. <http://www.mpr-online.de/>.
- KRABBE, P. F., ESSINK-BOT, M. L. & BONSEL, G. J. (1997): The comparability and reliability of five health-state valuation methods. *Social Science in Medicine*, **45**, 1641–1652.
- MEHREZ, A. & GAFNI, A. (1991): The healthy-years equivalents: How to measure them using the standard gamble approach. *Medical Decision-Making*, **11**, 140–146.
- MILLER, G. A. (1956): The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, **63**, 81–97.
- ORTH, B. (1982): Zur Bestimmung der Skalenqualität bei ‚direkten‘ Skalierungsverfahren. *Zeitschrift für experimentelle und angewandte Psychologie*, **29**, 160–178.
- ORTH, B. & WEGENER, B. (1983): Scaling occupational prestige by magnitude estimation and category rating methods: A comparison with the sensory domain. *European Journal of Social Psychology*, **13**, 417–431.
- SCHULENBURG, J. M. G., CLAES, C., GREINER, W., & UBER, A. (1998): Die deutsche Version des EuroQol-Fragebogens (The German Version of the EuroQol Questionnaire): *Zeitschrift für Gesundheitswissenschaften*, **6**, 3–20.
- WESTERMANN, R. (1984): Zur empirischen Überprüfung des Skalenniveaus von individuellen Einschätzungen und Ratings. *Zeitschrift für Psychologie*, **192**, 122–134.

- WESTERMANN, R. (1985): Empirical Tests of Scale Types for Individual Ratings. *Applied Psychological Measurement*, **9**, 265–274.
- WESTERMANN, R. & HAGER, W. (1983): Eine empirische Untersuchung zum Skalenniveau von Normwerten für die Bildhaftigkeit von Substantiven. *Psychologische Beiträge*, **25**, 112–125.
- WESTERMANN, R. & HAGER, W. (1985): Zur Konstruktion metrischer Skalen für die Schwereinschätzung von Delikten. *Diagnostica*, **31**, 153–163.

Uwe Konerding, Institut für Community Medicine, Ernst-Moritz-Arndt-Universität, Greifswald, Walther-Rathenau-Straße 48, 17487 Greifswald, Uwe.Konerding@Uni-Greifswald.de  
Tel.: 03834/86-7764