Shared Decision Making als Verhandlung von Ungewissheiten

Dissertation
zur Erlangung des Doktorgrades
am Fachbereich Humanwissenschaften
der Universität Osnabrück

Friedemann Geiger

Osnabrück, Juni 2007

„Die Selbstbeschreibung der modernen Gesellschaft als Wissensgesellschaft weist gerade darauf hin, dass Wissen zum Problem geworden ist und dass nicht das Wissen ein knappes Gut ist, sondern die Sicherheit, die wir dem Wissen einst entnommen haben.“

Armin Nassehi

Danksagung

Mein erster Dank geht an meinen Doktorvater Prof. Jürgen Kriz. Er hat mir bei dieser Arbeit auch bei richtungweisenden Entscheidungen viel Verantwortung übertragen und mich zugleich kritisch begleitet. Das mir entgegengebrachte Vertrauen hat mich von Beginn an ermutigt.

Bei Dr. Jürgen Kasper ist kaum zu entscheiden, ob ich mehr von seinen fachlichen Ratschlägen oder der freundschaftlichen Verbundenheit profitiert habe. Seine Ideen und Impulse sind an vielen Stellen dieser Dissertation zu erkennen. Zugleich wurde die Arbeit auch gerade von der persönlichen Motivation getragen, die ich aus unserer Zusammenarbeit schöpfen konnte.
In diesem Zusammenhang sei auch Prof. Ingrid Mühlhauser gedankt für die institutionelle Anbindung an die Fachrichtung Gesundheit des IGTW an der Universität Hamburg sowie die partielle Freistellung von Herrn Kasper.

Diese Arbeit wäre nicht möglich gewesen ohne die große Einsatzbereitschaft von Diplomandinnen und Doktorandinnen, namentlich Saskia Freiberger, Angelika Schmidt, Anna Südkamp, Sina Schneider und Janina Müller. Die Zusammenarbeit war in meinen Augen eine klassische „Win-Win-Situation“.

Wertvolle Kommentare zu inhaltlichen und formalen Fragen verdanke ich ferner Dr. Daniel Wollschläger und Jane Eggers.

Ein ganz besonderer Dank gilt meiner Frau Katja, die mir während der vergangenen Jahre den nötigen Freiraum gelassen hat, um dieses Projekt durchzuführen. Ihre Unterstützung in allem, was mir wirklich am Herzen liegt, trägt mich tagtäglich.

Schließlich bedanke ich mich beim Evangelischen Studienwerk Villigst e.V., das diese Arbeit durch ein Stipendium gefördert hat. Über die finanzielle Zuwendung hinaus

haben mich insbesondere die anregenden interdisziplinären Diskussionen mit den übrigen Villigstern auf Tagungen und Treffen bereichert.

Inhaltsverzeichnis

I Theoretischer Hintergrund **11**

1 Hintergrund **13**

1.1 Paternalistic Decision Making 14

1.2 Professional as Agent 15

1.3 Informed Decision Making 15

1.4 Shared Decision Making 16

1.4.1 Vorzüge von Shared Decision Making 17

1.4.2 Kontraindikationen von Shared Decision Making 21

2 Forschungsstand: Die Messung von Shared Decision Making **23**

2.1 Inhaltliche Beschreibungen des Shared Decision Making 23

2.2 Beobachtungsverfahren zur Beurteilung des Shared Decision Making 24

2.3 Unabhängige Variablen im Shared Decision Making 26

2.4 Abhängige Variablen im Shared Decision Making 28

2.5 Manipulation Check-Variablen im Shared Decision Making 30

2.6 Weitere Outcomevariablen im Zusammenhang mit Shared Decision Making 31

2.7 Unzulänglichkeiten bisheriger Messverfahren im Bereich Shared Decision Making 32

3 Theoretische Ansätze zur Entscheidungsfindung **39**

3.1 Normative Entscheidungstheorien 39

3.1.1 Expected Value Theory 40

3.1.2 Expected Utility Theory und Subjective Expected Utility Theory . . . 40
3.2 Deskriptive Entscheidungstheorien . . . 41
3.2.1 Rahmungseffekte . . . 42
3.2.2 Entscheidungsheuristiken . . . 44
3.2.3 Die Rationalität menschlicher Entscheidungen . . . 45

4 Die Rolle von Ungewissheit bei der medizinischen Entscheidungsfindung 47
4.1 Definition und Abgrenzung von Ungewissheit . . . 49
4.2 Taxonomien von Ungewissheit . . . 50
4.3 Shared Decision Making als Verhandlung von Ungewissheiten . . . 52

II Der QUICC – Ein Fragebogen zu Qualitäten von Ungewissheit in chronischen Erkrankungen 55

5 Die Entwicklung des QUICC 57
5.1 Kategorisierung des Datenmaterials . . . 58
5.1.1 Der Grounded Theory-Ansatz . . . 58
5.1.2 Die Bildung der Kategorien von Ungewissheit . . . 60
5.1.3 Die empirisch fundierten Kategorien von Ungewissheit . . . 62
5.1.4 Überprüfung der ‚Reliabilität' des Kategoriensystems . . . 68
5.2 Prinzipien der Itemgenerierung für den QUICC . . . 71
5.3 Vorstudie mit dem QUICC . . . 73
5.3.1 Stichprobe . . . 73
5.3.2 Ergebnisse und Diskussion . . . 74

6 Die Validierung des QUICC 77
6.1 Bestimmung der Reliabilität . . . 78
6.2 Itemanalyse und Itemselektion . . . 79
6.2.1 Itemtrennschärfe . . . 79
6.2.2 Itemschwierigkeit . . . 79
6.3 Die Bestimmung der Validität . . . 80
6.3.1 Interne Validität . . . 80
6.3.2 Externe Validität . . . 89

6.4 Weitere Überlegungen zur Bewertung des QUICC 90
6.5 Validierungsstudie 1 91
6.5.1 Stichprobe 92
6.5.2 Ergebnisse 93
6.5.3 Diskussion 97
6.5.4 Zusammenfassung 103
6.6 Validierungsstudie 2 103
6.6.1 Modifikation des QUICC 103
6.6.2 Annahmen über den QUICC 106
6.6.3 Stichprobe 106
6.6.4 Ergebnisse 107
6.6.5 Diskussion 110
6.7 Validierungsstudie 3 111
6.7.1 Stichprobe 112
6.7.2 Ergebnisse 113
6.7.3 Diskussion 117
6.8 Validierungsstudie 4 120
6.8.1 Herleitung der Methode 120
6.8.2 Stichprobe 121
6.8.3 Ergebnisse 121
6.8.4 Diskussion 121
6.9 Validierungsstudie 5 123
6.9.1 Stichprobe 125
6.9.2 Ergebnisse 125
6.9.3 Diskussion 127

III Gesamtdiskussion 131

7 Rekapitulation vorangegangener Schritte 133
7.1 Theoretische Präzisierung des Shared Decision Making 133
7.2 Die Qualitäten von Ungewissheit 134
7.3 Die Entwicklung und Validierung des QUICC 136

8 Mögliche Einschränkungen des Geltungsbereichs 139
8.1 Die Spezifität der Operationalisierung von Shared Decision Making 139

8.2 Die Allgemeingültigkeit der Ungewissheitskategorien 140

9 Integration weiterführender Gedanken in die vorgelegte Theorie 143

9.1 Die Verhandlung von Ungewissheiten als Ordnungsbildungsprozess . 143

9.2 Interaktionale Konzeption des Shared Decision Making 146

9.2.1 Konkordanz als interaktionales Effektmaß 147

9.2.2 Die Ungewissheiten der Ärzte 147

9.3 Shared Decision Making als interaktionaler Ordnungsbildungsprozess 150

Literatur 153

Appendices 169

A Formeln 169

A.1 Cohens κ 169

A.2 Fleiss' κ 170

A.3 Cronbachs α 171

A.4 AIC und BIC 172

B Leitfaden zur Kategorisierung des Interviewmaterials 173

C Die QUICC-Version aus der Vorstudie 177

D Die QUICC-Version aus Validierungsstudie 1 187

E Die QUICC-Version ab Validierungsstudie 2 195

F Die Ungewissheitstoleranzskala (UGTS) 201

G Zusammenfassung 203

H Lebenslauf 207

I Erklärung 209

Teil I

Theoretischer Hintergrund

Kapitel 1

Hintergrund

> „Den begonnenen Weg zu einer stärkeren Patientenpartizipation setzen wir mit dem Ziel fort, die Informations- und Beteiligungsrechte der Patientinnen und Patienten auszubauen und die Transparenz zu erhöhen. Die Rechtssicherheit von Patientenverfügungen wird gestärkt.“

So lautet der Abschnitt „Patientenrechte“ im jüngst besiegelten Koalitionsvertrag zwischen CDU/CSU und SPD (CDU/CSU/SPD, 2005, S. 86). Die Stärkung der Patientenpartizipation ist jedoch keine Initiative der neuen Bundesregierung, sondern führt eine seit mehreren Jahren erkennbare Linie der nationalen wie internationalen Gesundheitspolitik fort. Als nationale Wegmarke ist dabei etwa der Förderschwerpunkt des Bundesgesundheitsministeriums „Der Patient als Partner im medizinischen Entscheidungsprozess“ (BMGS, 2000) zu nennen. International hatte die Einbindung des Patienten[1] als Ziel bereits lange vorher Eingang in die Richtlinien der WHO gefunden (Waterworth und Luker, 1990). Um zu verstehen, was mit „stärkerer Patientenpartizipation“ gemeint ist, lohnt es, die Prototypen verschiedener Modelle der Arzt-Patient-Kommunikation näher zu beleuchten. Die Modelle unterscheiden sich vor allem darin, wie das medizinische Wissen des Arztes und der Wille des Patienten[2] kombiniert werden, um eine Behandlungsentscheidung zu treffen. An diese

[1] Wenn in dieser Arbeit von ‚Arzt‘ oder ‚Patient‘ die Rede ist, so bezieht sich das auf die Rolle dieser Person innerhalb der medizinischen Zusammenarbeit. ‚Arzt‘ steht somit gleichermaßen für ‚Ärztin‘ wie für ‚Arzt‘. Im Dienste der LeserInnenfreundlichkeit wird auf die Verwendung beider Formen verzichtet. Stattdessen werden beide Formen im Wechsel verwendet. Trotz der Gleichbehandlung von Ärzten und Ärztinnen sowie Patienten und Patientinnen in diesem Text kann es allerdings für die Interaktion der Beteiligten in einem SDM-Gespräch gleichwohl bedeutsam sein, ob beispielsweise eine Brustkrebspatientin einer Ärztin oder einem Arzt gegenübersitzt.

[2] Selten wird bei diesen Modellen zusätzlich auch *das medizinische Wissen des Patienten* und *der Wille des Arztes* berücksichtigt.

Betrachtung kann sich eine Bewertung ihrer jeweiligen konzeptionellen Vor- und Nachteile anschließen.

1.1 Paternalistic Decision Making

Das ‚Paternalistic Decision Making' (PDM) stellt die älteste Form der Arzt-Patent-Interaktion dar. Wie die Bezeichnung bereits vermuten lässt, übernimmt der Arzt dabei eine ‚väterlich-fürsorgende' Rolle. Er allein entscheidet in diesem Modell uneigennützig nach bestem Wissen und Gewissen, wie detailliert der Patient wann aufgeklärt wird. Darüber hinaus entscheidet der Arzt auch über die Art der Behandlung(en) allein. Alternativlos ist paternalistisches Handeln des Arztes bei der Akutversorgung von Unfallopfern. Die dazu komplementäre Rolle ist die des Patienten. Das Wort wurde im 16. Jahrhundert vom lateinischen ‚patiens' substantiviert, was soviel wie ‚(er)duldend, leidend' bedeutet (Duden, 2001). Diese Krankenrolle entbindet einerseits von gewissen gesellschaftlichen Pflichten, etwa am Arbeitsplatz oder in der Familie. Im Gegenzug hat der Patient den Anweisungen und Empfehlungen des Arztes Folge zu leisten, um die schnellstmögliche Genesung zu gewährleisten (Parsons, 1951). In dieser ursprünglichen Form des PDM wird also der passive, abhängige Patient dem Arzt als bestimmendem Experten gegenübergestellt. Historisch wurde eine solche Schieflage häufig begünstigt durch Statusunterschiede zwischen Arzt und Patient in Bezug auf Bildung, Einkommen oder Geschlecht (Charles et al., 1999a). Weniger extrem ist eine Auslegung des Modells, bei der der Arzt den Patienten mit jenen ausgewählten Informationen versorgt, die dieser braucht, um die Entscheidung des Arztes nachvollziehen, annehmen und umsetzen zu können. Dafür hat sich auch der Begriff ‚Informed Consent' eingebürgert. Zwar hat in solchen Fällen der Patient formal das letzte Wort, allerdings verfügt er über keinerlei Möglichkeiten, eine tatsächliche Alternative zur Schlussfolgerung des Arztes abzuwägen, was ihn weiterhin in Abhängigkeit hält. Durch die kennzeichnende Asymmetrie im Verhältnis der beiden Akteure ist auch diese Interaktion als paternalistisch zu bezeichnen (Emanuel und Emanuel, 1992). Das folgende Zitat von Weston (2001, S. 438) illustriert selbstkritisch das Bild des objektiv urteilenden, wohlmeinenden Arztes paternalistischen Zuschnitts:

> „Thirty-five years ago, when I was a medical student, we were taught to be paternalistic. We were trained to withhold information from patients,

> especially painful information about a potentially terminal illness; we were also instructed to take charge of interviews and to avoid getting ‚sidetracked' by patients' ‚irrelevant' concerns. The handbook that my university developed to teach the clinical method referred to the interview as ‚the interrogation'. Patients who did not comply with ‚doctor's orders' were called defaulters, untrustworthy, unreliable or faithless".

1.2 Professional as Agent

Charakteristisch für die Entscheidungsfindung zwischen Arzt und Patient ist die Ungleichverteilung der verschiedenen entscheidungsrelevanten Informationen. Der Arzt verfügt über (mehr) Informationen zu Krankheiten und Therapieoptionen, wohingegen der Patient (genauer) weiß, wie wichtig ihm Heilung, Vermeidung von Nebenwirkungen etc. sind (Hurley et al., 1992). Nur durch die Kombination beider Komponenten kann jedoch eine ausgewogene Entscheidung gefällt werden. Diesem Ungleichgewicht begegnen die hier aufgeführten Modelle auf unterschiedliche Weise. Während im PDM der Arzt sein Informationsdefizit bezüglich der Patientenpräferenzen durch im Einzelfall meist unvalidierte Annahmen kompensiert, verfolgt das ‚Professional as Agent-Modell' (PAM) (Charles et al., 1997) einen konsequenteren Ansatz. Hier obliegt es dem Arzt, sich im Gespräch ein umfassendes Bild der Patientenpräferenzen zu erarbeiten. Damit soll er in der Lage sein, eine Entscheidung zu treffen, als wäre er der Patient (Scheibler et al., 2003). Statt eines a priori ‚allwissenden Vaters' übernimmt der Arzt also die Rolle eines Agenten oder Vertreters in Sachen Patientenwille.

1.3 Informed Decision Making

Die zum Professional as Agent-Modell komplementäre Lösung zur Überbrückung der Informationsdefizite beschreibt das ‚Informed Decision Making' (IDM). Statt den Arzt zum Experten für die Patientenpräferenzen zu machen, soll hier der Patient zum Fachmann in den entscheidungsrelevanten medizinischen Aspekten werden, um anschließend allein entscheiden zu können. Dadurch wird der Arzt zum Lieferanten

von validen Informationen degradiert, es entsteht eine Beziehung wie zwischen Kellner und Gast. Diese Konzeption kann man als Gegenbewegung zum PDM ansehen (Deber, 1994; Levine et al., 1992), die im Einklang mit der erstarkenden Bürgerrechtsbewegung der sechziger Jahre sowie der Konsumentenbewegung der siebziger Jahre steht (Klemperer, 2003).

1.4 Shared Decision Making

Beim Professional as Agent-Modell kennt im Extremfall nur der Arzt vor der Auswahl einer Behandlung sowohl die medizinischen Möglichkeiten als auch die Patientenpräferenzen, beim PDM meint er sie zu kennen. Beim IDM dagegen weiß unter Umständen nur der Patient über beide Informationskomponenten bescheid. Das ‚Shared Decision Making' (SDM) stellt eine Synthese zwischen diesen Extremen dar. Erklärtes Ziel ist die Herstellung einer Situation, in der beide Interaktionspartner in der Lage sind, gemeinsam im Konsens eine Behandlung auszuwählen. Das setzt voraus, dass sich der Arzt über die Patientenpräferenzen informiert und der Patient erfährt, was Vor- und Nachteile der verfügbaren Therapieoptionen sind. Dies haben Charles et al. (1997) als „Two way exchange of information" bezeichnet. Das Wort „shared" bezieht sich also gleichermaßen auf die Entscheidung selbst wie auf die dazu nötigen Informationen.[3]

Die hier präsentierten Modelle sind alle als Prototypen anzusehen. Mischformen sind eher die Regel als die Ausnahme, zumal auch ein und derselbe Arzt von Patient zu Patient oder gar innerhalb einer Konsultation den Kommunikationsstil wechseln kann. Etwa könnte eine Patientin umso weniger mitentscheiden wollen, je klarer ihr die Lebensbedrohlichkeit ihrer Erkrankung wird. Inwieweit und unter welchen Umständen man als Arzt diesem Bedürfnis nachkommt, muss im Einzelfall abgewogen werden (vgl. auch 1.4.2).

[3] In diesem Punkt gibt es allerdings auch andere Auffassungen. So werden nach McNutt (2004) beim SDM nur die Informationen geteilt, während die Entscheidung allein der Patient trifft. Nach der hier vorgestellten Taxonomie wäre dies eher als IDM einzustufen.

1.4.1 Vorzüge von Shared Decision Making

SDM erscheint insgesamt als das am besten begründbare Kommunikationsmodell. Dieses Gesamturteil fußt auf zumeist theoretischen Überlegungen, die im Folgenden ausgeführt werden; mit den empirischen Nachweisversuchen befasst sich Kapitel 2 ausführlicher. Gleichwohl gibt es Konstellationen, die SDM weder erfordern, noch ermöglichen. Diese Einschränkungen werden in Abschnitt 1.4.2 aufgegriffen.

Der Patient als Experte. Während einer Operation oder nach Unfällen kann es durch hohen Blutverlust erforderlich werden, eine Bluttransfusion zu verabreichen. Da die Präparate routinemäßig auf verschiedene Krankheiten wie HIV oder Hepatitis getestet werden, ist das gesundheitliche Risiko äußerst gering, der medizinische Nutzen dagegen immens und unter Umständen lebensrettend. Der paternalistische Arzt würde also nach bestem Wissen und Gewissen dafür votieren. Nur der Patient kann jedoch wissen, ob durch seine persönlichen Einstellungen selbst in solch einem Fall eine Transfusion nicht in Frage kommt. Ein Zeuge Jehovas etwa müsste sich aus religiösen Überzeugungen dagegen entscheiden. Dieses Beispiel überzeugt weniger durch seine praktische Relevanz als vielmehr seine didaktische: Auch in scheinbar eindeutigen Situationen muss die Sichtweise des Patienten einbezogen werden, denn er allein ist der Experte für die eigenen Präferenzen. Die gleiche Schlussfolgerung erlaubt auch ein weniger exotisches Beispiel aus der hausärztlichen Praxis. Für die ‚Volkskrankheit' Bluthochdruck gibt es verschiedene, ähnlich wirksame Therapiemöglichkeiten, die von Tabletten über Ernährungsumstellung zu kontinuierlicher sportlicher Aktivität reichen (Donner-Banzhoff, 2001). Auch hier ist die Experteneinschätzung des Patienten unentbehrlich, ob er zu regelmäßigem Sport bereit ist, oder ob er die bequemere, aber gleichzeitig nebenwirkungsbehaftete Tabletteneinnahme vorzieht. McNutt (2004, S. 2517) bringt es auf den Punkt:

> „[Decision-making] is so difficult, that only a patient can do it".

Bei diesem Aspekt scheidet das PDM grundsätzlich aus, da es die Sicht des individuellen Patienten zugunsten von Annahmen über den hypothetischen ‚Durchschnittspatienten' weitgehend ausklammert.

Wichtigkeit einer vertrauensvollen Beziehung. Die demonstrative Offenlegung der verfügbaren medizinischen Informationen durch den Arzt und die explizite Einladung zur Mitbestimmung schaffen Vertrauen auf Seiten des Patienten (Charles et al., 1999a). Ferner betonten die Patientinnen in einer qualitativen Studie (Charles et al., 1999b) die Wichtigkeit, als Individuen gesehen zu werden. Dies bringen die Autoren in Zusammenhang mit der Berücksichtigung des persönlichen Entscheidungshintergrundes der Patientinnen (s.o.). Dieser Aspekt wird von vielen Praktikern demnach auch als „geheime Botschaft des SDM“ bezeichnet (Kasper et al., 2003).

Identifikation mit der Behandlungsentscheidung. Haben Patienten am Entscheidungsprozess aktiv teilgenommen, steigt dadurch die Identifikation mit der Entscheidung, was wiederum die Therapietreue (‚Adherence‘[4]) erhöht (Young und Klingle, 1996; Haynes et al., 1996). Kombiniert mit dem durch SDM vermittelten Vertrauen zum Arzt führt dies zu weniger Überweisungen und Laboruntersuchungen (Stewart et al., 1995). Gleichzeitig mag sich allerdings die durch SDM unterstützte Mündigkeit von Patienten auch gelegentlich im Einholen einer Zweitmeinung, einem wohlüberlegten Abbruch oder dem Wechsel einer Therapie ausdrücken.

Erhöhung der Handlungskompetenz. Patienten erleben häufig einen starken Kontrollverlust, wenn sie die Diagnose einer lebensbedrohlichen Krankheit wie Krebs erhalten (Breitbart, 1995; Cole und Pargament, 1999). Dieses Gefühl wird durch hochtechnisierte, stationäre Behandlungen wie etwa eine Strahlentherapie weiter genährt. Die Mitwirkung am Entscheidungsprozess kann dabei helfen, den Patienten wenigstens teilweise wieder vom behandelten Objekt zum handelnden Subjekt zu machen. Die positiven Effekte dessen sind aus der Grundlagenforschung der Psychologie lange bekannt (vgl. ‚Self-efficacy‘ (Bandura, 1977), ‚Internal locus of control‘ (Weiner et al., 1978)). Diese wiederum wirken sich günstig auf den Krankheitsverlauf aus (Lin, 1998; Buckelew et al., 1996).

Geteilte Verantwortung. Solcherart ‚Patientenempowerment‘ beim SDM birgt dabei auch für die behandelnden Ärzte Vorteile. Im paternalistischen Modell sind

[4] Das englische Pendant zu Therapietreue im PDM ist ‚Compliance‘. Sie bezeichnet die Folgsamkeit des Patienten gegenüber ärztlichen Verordnungen, etwa regelmäßige Tabletteneinnahme. Dagegen spricht man im SDM eher von Adherence, um zu verdeutlichen, dass es um ein Festhalten an einer selbst- oder zumindest mitbestimmten Maßnahme geht.

sie allein für die Behandlungswahl und deren Folgen verantwortlich. Darüber hinaus haben sie für die Orientierung in der Komplexität fachlicher Information, für die Aufklärung des Patienten sowie die Verarbeitung der Krankheit und deren Folgen Sorge zu tragen. Eine Interaktion nach dem Informed Decision Making-Modell bürdet hingegen vieles davon allein dem Patienten auf. Beim SDM lassen sich diese Zuständigkeiten gleichmäßig auf beide Parteien verteilen, so dass eine Verantwortungsschieflage vermieden wird (Kasper et al., 2003; Elwyn et al., 1999). Wie schon oben erwähnt, ist aber nicht jeder Patient zu jeder Zeit bereit, einen Teil der Verantwortung mitzutragen, sondern schwankt oft zwischen aktiv-kritischem Verbraucher und passiv-indolentem Konsumenten (Lupton, 1997) (vgl. wiederum Abschnitt 1.4.2).

Rechtsethische Aspekte. Das grundgesetzlich verbriefte Selbstbestimmungsrecht des Patienten (Art. 1 Abs. 1; Art. 2 Abs. 1 GG) geht dem Gesundheits- und Lebensschutz vor. Damit kann ein Patient jede Behandlung ablehnen. Diese Entscheidungsmöglichkeit muss ihm also in jedem Fall eingeräumt werden. Nur im Eilfall darf stattdessen – wie im PDM - nach dem *mutmaßlichen Willen* gehandelt werden (Lübbe, 2003; Laufs und Uhlenbruck, 2002). Umgekehrt hat auch der Arzt außerhalb von Notfällen das Recht, eine Behandlung aus fachlichen, ethischen oder ökonomischen Gründen abzulehnen.

Als Basis für eine solcherart konsensuelle Entscheidung fungiert eine vollumfängliche Aufklärung des Patienten, wie auch der Bundesgerichtshof jüngst erneut klargestellt hat:

> „Auch die Aufklärung über bestehende unterschiedliche Behandlungsmöglichkeiten dient dem Selbstbestimmungsrecht des Patienten und ist daher Voraussetzung einer rechtmäßigen Behandlung“ (Bundesgerichtshof, 15. März 2005, S. 1).

Und weiter:

> „Erst eine nach vollständiger und gewissenhafter Aufklärung des Patienten wirksame Einwilligung (‚informed consent‘) macht den Eingriff in seine körperliche Integrität rechtmäßig“ (ebd., S. 9).

Ein rechtliches Relikt bildet indes die heute noch begrenzt gültige Rechtsfigur des sog. ‚Therapeutischen Privilegs‘, wonach der Arzt nach pflichtgemäßem Ermessen

auf eine vollständige Aufklärung verzichten können soll, wenn diese dem Patienten schaden könnte (Lübbe, 2003). Der Anwendung dieses Privilegs sind dabei allerdings durch die Rechtssprechung des Bundesgerichtshofs enge Grenzen gesetzt (Giesen, 1987).

Während in Deutschland Gerichte und Gesetze eine Arzt-Patient-Interaktion fordern, die zumindest große Übereinstimmungen mit dem SDM haben, greift die WHO SDM in ihren Richtlinien sogar explizit auf (Waterworth und Luker, 1990). Auch die Europäische Patientenrechtscharta (European Charter of Patients' Rights, 2002) formuliert eindeutig. So heißt es zum Teilen von Informationen:

> „Every individual has the right of access to all information that might enable him or her to actively participate in the decisions regarding his or her health; this information is a prerequisite for any procedure and treatment [...] (ebd, S. 5)"

und zum Teilen der Entscheidung:

> „Each individual has the right to freely choose from among different treatment procedures and providers on the basis of adequate information. [...] Each individual has the right to diagnostic or therapeutic programmes tailored as much as possible to his or her personal needs" (ebd, S. 5-7).

Kosten. Unter anderen hegen die Krankenkassen die Hoffnung, dass seltener unnötige Behandlungen durchgeführt werden, wenn der Patient als mündige Kontrollinstanz fungiert. Geringeres ‚Doctor shopping' von eingebundenen Patienten (Charles et al., 1999a; Stewart et al., 1995) verspricht ebenfalls Kostensenkungen. Jedoch wäre es auch vorstellbar, dass gut informierte Patienten eine Behandlung einfordern, die vom Arzt wegen geringer Erfolgschance nicht durchgeführt worden wäre. Zudem ist SDM wenigstens bei der ersten Konsultation meist zeitaufwändiger und damit teurer. Dass die Unterschiede im zeitlichen Aufwand bei versierten Ärztinnen und Ärzten im zu vernachlässigenden Rahmen bleiben, zeigen indes die Studien von Roter et al. (1995), Clark et al. (1998) und Loh et al. (2005).

Bisher ist noch nicht geklärt, ob SDM auf diese Weise direkt zu einer Kostensenkung im Gesundheitswesen beitragen kann (Büchi et al., 2000; Charles et al., 1999a).

Eine indirekte Kostensenkung durch eine qualitative Verbesserung der Behandlung gemäß den eben aufgezählten Punkten erscheint jedoch plausibel. Insbesondere sei die Adherence hier erneut unter ökonomischen Gesichtspunkten erwähnt: Nach Volmer und Kielhorn (1998) errechnet sich der idealisierte Nutzwert eines Medikaments (Efficiency) aus einer Verknüpfung von dessen Wirksamkeit (Efficacy) und der Therapietreue – wenn es nicht oder zu selten eingenommen wird, kann das potenteste Mittel nicht helfen:

$$Nutzwert(Efficiency) = Wirksamkeit(Efficacy) \times Therapietreue\ (Adherence)$$

Entscheidet sich also ein Patient aus persönlichen Gründen für ein weniger effektives Medikament, kann dies bei entsprechender Adherence letztlich wirksamer und mithin kostengünstiger sein als der jeweilige Goldstandard.

1.4.2 Kontraindikationen von Shared Decision Making

Angesichts dieser Vielzahl von Vorzügen einer am SDM orientierten Arzt-Patient-Interaktion soll im Folgenden erörtert werden, unter welchen Umständen SDM *nicht* indiziert ist. Dabei lässt sich unterscheiden nach situativen Einflüssen einerseits und Patientenpräferenzen andererseits.

Situative Variablen. Es gibt Situationen, die kein SDM zulassen. Dazu zählt ärztliches Handeln in akuten Notfällen, in denen der Zeitdruck keinen gemeinsamen Entscheidungsprozess zulässt. Darüber hinaus ist SDM mit bewusstlosen, geistig oder psychisch stark eingeschränkten Patienten nicht möglich. Während eine Ärztin im ersten Fall notgedrungen nach dem mutmaßlichen Willen entscheiden muss (und darf), kann (und muss) in den letztgenannten Beispielen ein stellvertretendes SDM mit Angehörigen erfolgen, um dem tatsächlichen Willen des Patienten möglichst nahe zu kommen (vgl. die Bemerkungen zu den rechtlichen Aspekten in Abschnitt 1.4.1).

Vielfach wird Zeitmangel auch außerhalb akuter Notfälle als Argument gegen SDM angeführt: Eine Einbeziehung des Patienten könne man sich im Alltag einer Arztpraxis schlicht nicht erlauben (Isfort et al., 2002; Elwyn et al., 1999). Einerseits wurden

oben diesbezüglich bereits anders lautende Befunde angeführt. Andererseits argumentieren Howie et al. (1997), dass sich der Einsatz von SDM aus den genannten Gründen der Effektivitätssteigerung dennoch lohne.

Patientenpräferenzen. Wie schon erwähnt, wollen nicht alle Patienten in vollem Maße am Entscheidungsprozess partizipieren, sondern je nach Studie etwa jeder zweite (Kjellgren et al., 2000; Fallowfield, 2001). Während zwar ein großer Teil der Patienten in vollem Umfang über die Erkrankung und ihre Therapiemöglichkeiten *informiert* werden will, möchte nur ein deutlich geringerer Prozentsatz auch über die Therapie *mitentscheiden* (Beisecker und Beisecker, 1990; Fallowfield et al., 1994; Strull et al., 1984; Elwyn et al., 1999). Man muss also zwischen Informations- und Partizipationsbedürfnis differenzieren. Patienten gegen ihren erklärten Willen einbinden zu wollen, ist generell kontraproduktiv und belastet die Beziehung.

Angesichts dieser Befunde haben verschiedene Studien untersucht, von welchen Faktoren es abhängt, ob ein Patient sich am Entscheidungsprozess beteiligen möchte. So gibt es mehrere Hinweise, dass ein höherer Schulabschluss (Beaver et al., 1996; Rothenbacher et al., 1997; Veronesi et al., 1999; Wallberg et al., 2000), ein geringeres Lebensalter (ebd.), ein leichter Krankheitsverlauf (Elwyn et al., 1999) sowie ein höheres Aktivitätsniveau (Rothenbacher et al., 1997; Blanchard et al., 1988) mit einem größeren Mitentscheindungsbedürfnis einhergehen. Außerdem korrelieren meist trotz oben genannter Differenzen Informations- und Partizipationsbedürfnis (Davidson et al., 1999; Stewart et al., 2000). Dies ist insofern nicht verwunderlich, als die Partizipationswilligen zumeist eine Untermenge der Informationswilligen darstellen. Allerdings konnten Ende et al. (1989) diesen Zusammenhang nicht bestätigen.

Trotz dieser größtenteils übereinstimmenden Befunde können Ärzte daraus keine Heuristik ableiten, ob eine vor ihnen sitzende 53-jährige Hypertoniepatientin mit mittlerer Reife mitentscheiden möchte, welche Therapie sie bekommt. Dazu ist der Zusammenhang zwischen den diskutierten Parametern und den Patientenpräferenzen zu schwach, sprich die Varianzaufklärung der Prädiktoren zu gering. Stattdessen bedarf es für jede Konsultation einer Validierung des Eindrucks von den Beteiligungswünschen der Patienten.

Kapitel 2

Forschungsstand: Die Messung von Shared Decision Making

In den vorangegangenen Ausführungen wurden Begriffe wie „SDM“ oder „Patientenpartizipation“ mit scheinbarer Selbstverständlichkeit verwendet. Auch wurden mögliche Vorteile von SDM diskutiert. Dies ist insofern rückwirkend kritisch zu beleuchten, als bis heute kein Konsens erzielt wurde, wie SDM zu definieren sei. Noch weiter entfernt ist man von einer Einigung, wie SDM *gemessen* werden könnte. Wenn aber SDM noch nicht befriedigend operationalisiert werden kann, ist fraglich, wie man Effekte von SDM überhaupt valide bestimmen können sollte.

Dieses Kapitel stellt verschiedene Operationalisierungsversuche aus der aktuellen SDM-Forschung vor. Sie können im Wesentlichen gruppiert werden in inhaltliche Beschreibungen, Beobachtungsverfahren sowie in Betrachtungen im Sinne einer unabhängigen oder einer abhängigen Variable. Zusätzlich fokussieren einige Verfahren auf den Prozess der Entscheidungsfindung. Eine ausführliche Auseinandersetzung mit den verschiedenen Ansätzen erfolgt im Anschluss in Abschnitt 2.7.

2.1 Inhaltliche Beschreibungen des Shared Decision Making

Die grundlegende Beschreibung des SDM stammt von Charles et al. (1997). Demnach ist SDM folgendermaßen charakterisiert:

- SDM umfasst mindestens zwei teilnehmende Parteien, nämlich Arzt und Patient.
- Beide Parteien beteiligen sich am Entscheidungsprozess.
- Der Austausch von Informationen ist eine notwendige Bedingung für SDM.
- Es wird eine Entscheidung getroffen, indem beide Parteien sich einigen.

Diese Mindestanforderungen an ein SDM-Gespräch liegen auch den Ausführungen in Abschnitt 1.4 zugrunde.

Towle und Godolphin (1999) liefern darüber hinaus eine Schritt für Schritt-Anleitung für Ärzte zur Gestaltung eines SDM-Gesprächs:

1. Entwickle eine Partnerschaft mit dem Patienten
2. Erarbeite oder erfrage die Informationswünsche des Patienten (Umfang/Art)
3. Erörtere, welche Rolle der Patient im Entscheidungsprozess einnehmen möchte
4. Ermittle und berücksichtige die Patientenpräferenzen und -prioritäten in Bezug auf die Therapieoptionen
5. Stelle daraus eine Vorauswahl zusammen, sammele dazu die klinische Evidenz und passe sie auf den individuellen Patienten an
6. Sprich diese Informationen mit dem Patienten durch und hilf ihm, die Vor- und Nachteile der verschiedenen Optionen in Bezug auf seine eigenen Präferenzen zu erkennen
7. Handelt gemeinsam eine Entscheidung aus, die diesen Konflikt löst
8. Einigt Euch auf einen Behandlungsplan inkl. Folgeuntersuchungen

2.2 Beobachtungsverfahren zur Beurteilung des Shared Decision Making

Elwyn et al. (2003) stellten eine ähnliche Anleitung für Ärzte zusammen. Sie formulierten die Abfolge der Schritte dabei so, dass diese als Items für ein videobasiertes

Beobachtungsverfahren (OPTION Scale, Elwyn et al., 2005) benutzt werden können. Die Items der überarbeiteten Version lauten:

1. Der Arzt lenkt die Aufmerksamkeit auf ein bestimmtes Problem, das einer Entscheidung bedarf.
2. Der Arzt stellt klar, dass es mehr als einen Weg gibt, um mit dem identifizierten Problem umzugehen. (‚Equipoise').
3. Der Arzt klärt, in welcher Art und Weise der Patient Informationen erhalten möchte, um eine Entscheidung treffen zu können (z. B. im Gespräch, durch das Lesen von Informationsmaterial, durch die Präsentation graphisch aufbereiteter Daten, durch Videos oder andere Medien).
4. Der Arzt zählt Optionen auf (einschließlich derjenigen, nichts zu tun).
5. Der Arzt erläutert dem Patienten die Vor- und Nachteile der verschiedenen Optionen (nichts zu tun ist ebenfalls eine Option).
6. Der Arzt exploriert die Erwartungen (oder Ideen) des Patienten, wie mit dem Problem (den Problemen) umgegangen werden soll.
7. Der Arzt exploriert die Sorgen (Befürchtungen) des Patienten im Hinblick auf den Umgang mit dem Problem.
8. Der Arzt vergewissert sich, dass der Patient die Informationen verstanden hat.
9. Der Arzt bietet dem Patienten explizit Möglichkeiten an, während des Entscheidungsprozesses Fragen zu stellen.
10. Der Arzt findet heraus, in welchem Ausmaß der Patient bei der Entscheidungsfindung beteiligt werden möchte.
11. Der Arzt weist darauf hin, dass es notwendig ist, jetzt eine Entscheidung zu treffen (oder aufzuschieben).
12. Der Arzt weist darauf hin, dass es notwendig ist, noch einmal auf die Entscheidung zurückzukommen (oder auf das Aufschieben der Entscheidung).

Seit der Revision von OPTION wird die Ausprägung des mit dem Item beschriebenen Verhaltenselements auf einer fünfstufigen Likert-Skala bewertet.

Während das Verfahren augenscheinvalide ist, zeigt auch die überarbeitete Version Schwächen bei der Konstruktvalidität. Wenn die Items gemäß der Theorie auf einen

Faktor[1] gruppiert werden, so klärt dieser gerade 28% der Varianz in den Daten auf (Elwyn et al., 2005). Dies mag zum Teil auf die eher durchschnittliche Reliabilität des Verfahrens zurückzuführen sein. Die Retestreliabilitäten von den einzigen beiden untersuchten Ratern liegen immerhin bei 0,68 bzw. 0,82, wohingegen die Interraterreliabilität zwar bei manchen Items Cohens κ-Werte von 0,98 erreicht, bei anderen Items dagegen 0,08 beträgt. Solch große Diskrepanzen sind vermutlich neben der Qualität der Items dem extremen Bodeneffekt in der angeführten Stichprobe geschuldet (auf die oberen drei der fünf Antwortstufen entfallen *kumuliert* 0,6% der Urteile). Insgesamt überzeugt die von den Urhebern selbst publizierte Studie nur bedingt. Dennoch wird OPTION als das einzige Messinstrument seiner Art derzeit in verschiedenen laufenden Studien eingesetzt. Es existiert auch eine von den Verfassern autorisierte, aber nicht veröffentlichte deutsche Übersetzung, aus der die obigen Itemformulierungen stammen.

Ein Beobachtungsverfahren, mit dem das Verhalten des Patienten beurteilt werden könnte, ist noch nicht publiziert worden. Ebenso fehlen bislang Studien zur Beobachtung der Beziehung zwischen Arzt und Patient, wie es beispielsweise mit der Structural Analysis of Social Behaviour (SASB) von Benjamin (1993) möglich wäre.

2.3 Unabhängige Variablen im Shared Decision Making

Unter der Überschrift der SDM-Messinstrumente werden häufig Verfahren angeführt, die unter Verwendung von Termini der Versuchsplanung zuweilen als ‚unabhängige Variablen' bezeichnet werden (Scheibler et al., 2003). Denn sie fokussieren auf Vorbedingungen auf Patientenseite für eine sich anschließende SDM-Kommunikation.[2] Darunter fallen Vorstellungen eines Patienten zu Art und Umfang der

[1] In der Bewertung der Messverfahren zum SDM werden grundlegende Vokabeln aus der Fragebogenkonstruktion als bekannt vorausgesetzt. Ausführlicher befasst sich Abschnitt 6.3 mit Begrifflichkeiten und Logik dieser Methoden.

[2] Da diese Vorbedingungen nicht vom Versuchleiter erzeugt werden, müssten sie allerdings konsequenterweise als *quasiexperimentelle* Variablen bezeichnet werden. Und selbst das könnte man insofern für unpräzise halten, als diese Variablen vermutlich nicht nur das SDM-Gespräch beeinflussen, sondern zirkulär auch davon wiederum beeinflusst werden. Es wäre somit zu überlegen, inwiefern die in diesem Abschnitt diskutierten Variablen eher als ‚Traits' oder ‚States' aufzufassen sind.

von der Ärztin zu präsentierenden Information sowie zu Form und Ausmaß der Mitbestimmung an der Entscheidung. Durch Kenntnis dieser Patientenwünsche soll die Arzt-Patient-Kommunikation möglichst gut auf den individuellen Patienten zugeschnitten werden können. Nachfolgend werden zwei der international am weitesten verbreiteten Messverfahren zu Patientenpräferenzen vorgestellt.

Autonomy Preference Index – API (Ende et al., 1989). Der API umfasst mit seinen zwei Subskalen ‚Preference for Information Seeking' und ‚Preference for Decision Making' die beiden eben skizzierten Arten von Patientenwünschen. Die acht bzw. fünfzehn Items der Subskalen werden durch die Patienten selbst beantwortet, größtenteils in fünfstufigem Likert-Format.

Sowohl die Reliabilität als auch die Validität des API hat sich in bisherigen Untersuchungen als zufriedenstellend erwiesen. Gleichwohl sind weitere Belege dazu wünschenswert (Giersdorf et al., 2003). Insbesondere bleibt abzuwarten, ob die aktuell in verschiedenen Studien eingesetzte deutsche Version ebenfalls überzeugen wird.

Control Preferences Scale (Degner und Sloan, 1992). Bei der kompakten Control Preferences Scale bringen Patienten fünf Aussagen in eine Rangfolge, die das Ausmaß der Übereinstimmung mit der eigenen Auffassung abbildet. Dies kann zum Beispiel durch eine Kartensortieraufgabe geschehen. Die Aussagen beschreiben eine graduell unterschiedlich ausgeprägte Beteiligung des Patienten an der Entscheidungsfindung.

Während in der Originalarbeit das zur Prüfung der Reliabilität solcher Instrumente gebräuchliche Coombs-Kriterium[3] in zwei großen Stichproben erfüllt war, konnten Beaver et al. (1996) dies nicht belegen. Zudem wiesen Entwistle et al. (2001) Validitätsprobleme der Skala nach.

[3]Dabei handelt es sich um eine Prüfgröße aus der Unfolding-Analyse, die die Eindimensionalität einer Skala kontrolliert.

2.4 Abhängige Variablen im Shared Decision Making

Während im vorangegangenen Abschnitt Verfahren vorgestellt wurden, die wichtige Informationen im Vorfeld einer am SDM orientierten Arzt-Patient-Kommunikation liefern sollen, wird nun der Fokus auf die gängigsten Messinstrumente gelenkt, die sich mit den *Auswirkungen* von SDM beschäftigen. Versuchsplanerisch kann man analog von ‚abhängigen Variablen' sprechen.

Decisional Conflict Scale – DCS (O'Connor, 1995). Wie in Abschnitt 1.4 und 1.4.1 ausgeführt wurde, sollten Patienten mit einer Behandlungsentscheidung eher einverstanden sein, wenn sie selbst dazu beigetragen haben. Entsprechend untersucht die DCS, ob nach einer Entscheidung ein ‚Entscheidungskonflikt' verbleibt. Darunter versteht die Autorin die Unsicherheit, welcher Weg einzuschlagen ist. Entscheidungskonflikte sollen wahrscheinlicher in Situationen hohen Risikos und großer Ungewissheit sein. Daher werden neben der vier Items umfassenden Skala zur Beurteilung der getroffenen Entscheidung (Effective Decision) zusätzlich die wahrgenommene Unsicherheit in Bezug auf den Entscheidungsgegenstand (Decision Uncertainty, drei Items) und die dazu beitragenden Faktoren (Factors Contributing to Uncertainty, neun Items) mit eigenen Skalen erfasst.

In der 909 Personen umfassenden Validierungsstichprobe zeigte die DCS in den Subskalen recht gute interne Konsistenzen (Cronbachs α zwischen 0,78 und 0,92) sowie eine erfreuliche Retestreliabilität von .81. Auch einige unterstützende Befunde zur Validität werden berichtet. Allerdings räumt die Autorin selbst ein, dass zusätzliche Validitätsuntersuchungen bei anderen Patientengruppen in schwierigeren Entscheidungssituationen erforderlich sind. In der Tat kamen bei zwei Studien mit Krebspatienten, die sich vermutlich Entscheidungen höherer Komplexität gegenüber sahen, Zweifel an der faktoriellen und der Konstruktvalidität auf (Koedoot et al., 2001).

Satisfaction with Decision Scale – SwDS (Holmes-Rovner et al., 1996). Die SwDS misst anhand von nur sechs Items die Zufriedenheit von Patienten mit ihrer Entscheidung insgesamt, der vorangegangenen Informationsvermittlung sowie der Beteiligung am Entscheidungsprozess.

Zwar konnten die Autoren sowohl die interne Konsistenz als auch die konkurrente Validität der Skala belegen. Allerdings fehlen bislang überzeugende Validierungsstudien jenseits des ursprünglichen Entwicklungskontextes postmenopausaler Patientinnen.

Post-Decision Satisfaction Scale – PDSS (Sainfort und Booske, 2000). Die Zufriedenheit mit einer Entscheidung und den dafür verfügbaren Informationen bei Menschen allgemein erfasst die PDSS anhand von gerade acht Items, die zusätzlich in drei Subskalen unterteilt sind.

Die Gütekriterien stellen trotz des eigenwilligen Verhältnisses zwischen Item- und Skalenanzahl zufrieden. Allerdings hinterfragen die Autoren selbst die Übertragbarkeit ihrer Ergebnisse auf andere Kontexte. Schließlich wurde die einzige Untersuchung mit der PDSS bei 197 Angestellten durchgeführt, die sich zwischen unterschiedlichen Krankenkassenleistungen entscheiden mussten.

Combined Outcome Measure for Risk Communication and Treatment Decision Making Effektiveness – COMRADE (Edwards et al., 2003). Eine Mischform aus Outcome- und Prozessmessung (vgl. 2.5) stellt COMRADE dar. Aus einem explorativ-faktorenanalytischen Vorgehen entstanden die eher prozessorientierte Skala ‚Risk communication' sowie die ergebnisbezogene Skala ‚Confidence in decision' mit je zehn Items. Ein Großteil davon ist anderen Fragebögen entnommen (z.B. SwDS, DCS). Während Umfang und Zusammenstellung der Validierungsstichprobe überzeugen, erscheint die stark durch die Faktorladungen bestimmte Itemauswahl und -gruppierung problematisch. Entsprechend kritisieren die Autoren selbst die inhaltliche Heterogenität insbesondere der Skala ‚Risk communication' trotz statistischer Homogenität. Dieser Aspekt ist anhand weiterer Stichproben genauer zu untersuchen. Auch blieben in der Validierungsstudie einige postulierte Korrelationen mit anderen Maßen aus. Dennoch stellt COMRADE aufgrund der kombinierten Erfassung von Prozess- und Outcomemerkmalen ein viel versprechendes Verfahren dar.

2.5 Manipulation Check-Variablen im Shared Decision Making

Neben Vorbedingungen und Konsequenzen werden auch die Merkmale von SDM *während* des Arzt-Patient-Gesprächs betrachtet. Um den Vergleich zur Versuchsplanung erneut zu bemühen, könnte man diese als ‚Manipulation Check-Variablen' bezeichnen, da sie überprüfen sollen, ob SDM überhaupt stattgefunden hat.

Patients' Perceived Involvement in Care Scale – PICS (Lerman et al., 1990). Da die Beteiligung der Patienten von vielen als zentrales Element des SDM angesehen wird, erscheint es konsequent von Lerman et al. (1990), die Patienten nach einem Entscheidungsprozess einfach zu fragen, ob sie sich beteiligen konnten. Dazu wurde die PICS entwickelt, die auf drei Subskalen die Patientenaktivierung durch die Ärzte (Doctor Facilitation Scale, fünf Items), das Informationsverhalten des Patienten (Patient Information Scale, vier Items) sowie dessen Entscheidungsteilnahme (Patient Decision Making Scale, vier Items) erfasst. Die Gesamtskala wies im Original zufrieden stellende Reliabilität auf (Cronbachs $\alpha = 0,73$). Eine von Pfaff et al. (2001) untersuchte deutsche Übersetzung übertraf diese Werte sogar leicht, was allerdings auf die erheblich größere Stichprobe von 6614 Nierenkranken zurückgehen dürfte. Gleichzeitig enthüllte diese jedoch Unstimmigkeiten bei der ‚Patient Decision Making Scale', die die Autoren auf Unterschiede zwischen den Gesundheitssystemen in den USA und Deutschland zurückführen. Insofern ist eine Überarbeitung für deutsche Verhältnisse erforderlich.

Man-Son-Hing Scale (Man-Son-Hing et al., 1999). Die Man-Son-Hing Scale wurde nicht in erster Linie zur Messung von SDM selbst entwickelt, sondern um den Einfluss einer Entscheidungshilfe bei Fibromyalgie zu erfassen. Daher standen bei der Veröffentlichung der Skala deren Testgütekriterien nicht im Vordergrund. Bis heute gibt es keine Angaben zu Reliabilität und Validität. Zudem enthält die Skala einige Items mit offenem Antwortformat, was zusätzlich die Auswertungsobjektivität unterminiert. Eine Stärke des Verfahrens liegt allerdings darin, dass auch eine Version für die Einschätzung des Arztes vorliegt.

Fragebogen zur Partizipativen Entscheidungsfindung – PEF-FB (Giersdorf et al., 2004). Der bislang einzige originär in Deutschland entwickelte Fragebogen zum SDM ist aus dem Methodenprojekt des BMGS-Schwerpunktes ‚Der Patient als Partner im medizinischen Entscheidungsprozess' hervorgegangen. Der PEF-FB besteht aus neun dichotomen und 14 vierstufigen Items. Diese erfassen Prozessmerkmale, die den Schritten der oben aufgeführten OPTION-Items ähneln. Die interne Konsistenz des eindimensionalen Bogens beträgt 0,85 (Simon et al., 2005a). In einem Kontrollgruppendesign gaben jene Patienten durchschnittlich höhere Werte für SDM an, deren Ärzte zuvor eine entsprechende Schulung durchlaufen hatten, worin die Autoren die Konstruktvalidität und Veränderungssensibilität belegt sehen (Simon et al., 2005b). Einschränkend ist zu sagen, dass der PEF-FB bislang nur bei Depressiven eingesetzt wurde, und die Items sich im Sinne der Testtheorie als zu leicht erwiesen haben. Derzeit werden sowohl eine gekürzte und in zwei Subskalen aufgeteilte Version wie auch eine modifizierte eindimensionale Variante bei verschiedenen Indikationen im Rahmen des Förderschwerpunktes geprüft, so dass noch keine abschließende Bewertung erfolgen kann.

2.6 Weitere Outcomevariablen im Zusammenhang mit Shared Decision Making

Die in den vorangegangenen Abschnitten aufgeführten Verfahren beanspruchen, direkt im Zusammenhang mit SDM stehende Variablen zu erfassen wie etwa ‚Patientenbeteiligung' oder ‚Partizipationspräferenzen'. Darüber hinaus verspricht man sich von SDM einige Effekte, die teilweise bereits in Abschnitt 1.4.1 erwähnt wurden. So wird z.B. in Studien untersucht, wie sich SDM auswirkt auf die Lebensqualität, den allgemeinen oder krankheitsspezifischen Gesundheitszustand, die Adherence, den Behandlungserfolg, die Behandlungszufriedenheit etc. Eine Übersicht ist dem Review von Kennedy zu entnehmen (Kennedy, 2003).

Ferner wird häufig der Einsatz von Entscheidungshilfen (Decision Aids) in Verbindung mit SDM gebracht. Dabei handelt es sich um Schautafeln, Videos, interaktive Computerprogramme o.ä., die die Vermittlung von relevanten Informationen erleichtern, den Entscheidungsprozess unterstützen und nicht zuletzt dem Patienten zur aktiv-kritischen Teilnahme am Gespräch verhelfen sollen. Sie werden womöglich

häufiger in partizipativen Arzt-Patient-Gesprächen eingesetzt, stellen aber kein eigenständiges Messinstrument dar (Kennedy, 2003).

2.7 Unzulänglichkeiten bisheriger Messverfahren im Bereich Shared Decision Making

Wie im vorangegangenen Abschnitt deutlich wurde, existieren durchaus einige Messinstrumente für verschiedene Aspekte des SDM. „Eine methodisch zufrieden stellende Validitätsstudie liegt allerdings für keines der bisher entwickelten Messinstrumente vor“, wie Giersdorf et al. (2003, S. 81) in ihrer systematischen Übersichtarbeit urteilen. Sie verweisen zudem auf die überwiegend nordamerikanische Provenienz der Verfahren. Nicht zuletzt bedingt durch Unterschiede in den Gesundheitssystemen lasse sich jedoch kein Messinstrument durch bloße Übersetzung für den Einsatz im deutschen Sprachraum anpassen. Neben dem Mangel an Validierungsstudien in verschiedenen Systemen und Kulturen liegen zudem für die wenigsten Instrumente Erfahrungen in verschiedenen Krankheitsbildern vor. Den Mangel an befriedigenden Messverfahren im Bereich SDM beklagen analog auch schon Sainfort und Booske (2000), Elwyn et al. (2001), Simon et al. (2005a) sowie Kasper et al. (2003).

Nun könnte man einwenden, diese Probleme seien leicht lösbar, indem die vorhandenen Verfahren übersetzt, validiert und gegebenenfalls den deutschen Verhältnissen angepasst würden. Auch die Durchführung indikationsübergreifender Validierungsstudien stellt schließlich kein grundsätzliches Problem dar. Allerdings treten zu diesen eher empirisch-statistisch basierten Unzulänglichkeiten der Verfahren einige *theoretisch-inhaltliche* Zweifel hinzu, die die Messbarkeit von SDM anhand der vorgestellten Instrumente in Frage stellen.

Inhaltliche Beschreibungen. Wirft man zunächst einen Blick auf die obigen inhaltlichen SDM-Beschreibungen von Charles et al. (1997) und Towle und Godolphin (1999), so wird schnell klar, dass sie zu grob sind, um Ärztinnen als Anleitung zu dienen. Genauso wenig taugen sie für Gesprächsbeobachter als Beurteilungskriterien, inwieweit SDM umgesetzt wurde. Außerdem sind die Schritte von Towle und

Godolphin nicht theoriebasiert abgeleitet worden, z.B. aus kommunkationstheoretischen Überlegungen. Nicht wesentlich mehr Fundierung besitzen die Charakteristika von Charles et al.

Beobachtungsverfahren. Obgleich ähnlich wenig nachvollziehbar theoriebasiert entwickelt, zeichnet sich die OPTION Scale (Elwyn et al., 2003) durch eine deutlich präzisere Ausdifferenzierung der Anforderungen an einen SDM praktizierenden Arzt aus. Wie bereits erwähnt, krankt das Verfahren allerdings auch in seiner überarbeiteten Version an ungenügender Testgüte. Als mögliche Ursachen sehen Giersdorf und Mitarbeiter (2003) die bei Beobachtungsverfahren unumgängliche Beschränkung auf beobachtbare Verhaltensweisen der Beteiligten. So werden z.B. Motivationen oder verschwiegene Unstimmigkeiten bei Arzt und Patient übersehen. Es bleibt bei reinen Beobachtungsverfahren offen, ob sich ein Patient von seiner Ärztin respektiert und eingebunden *gefühlt* hat – unabhängig davon, wie eng sie sich an Gesprächsleitfäden gehalten haben mag. Auch ‚*spürt*' man als Beobachter häufig wichtige Aspekte, die aber nicht *explizit* gesagt oder getan wurden. Unausweichlich entsteht das Dilemma, wie viel Interpretation bei der Auswertung solcher Gesprächsvideos zulässig ist. Nicht offen ausgesprochene und gleichwohl hoch bedeutsame Interaktionen dürften insbesondere bei länger bestehenden Arzt-Patient-Beziehungen anzutreffen sein (Kasper et al., 2003).

Unabhängige Variablen. Die als unabhängige Variablen bezeichneten Maße für die Patientenpräferenzen fokussieren dagegen auf solcherart Motivationen und Wünsche von Patienten. Jedoch liefern sie für sich genommen keinen Anhaltspunkt über die Qualität des sich anschließenden Gesprächs. Ein paternalistischer Entscheidungsprozess wird nicht dadurch weniger direktiv, dass im Vorhinein die Beteiligungswünsche eines Patienten erhoben werden, solange diese nicht berücksichtigt werden. So mag die Kenntnis der Patientenpräferenzen etwa durch den Autonomy Preference Index (API) (Ende et al., 1989) die Umsetzung von SDM zwar erleichtern, ist damit aber nicht gleichzusetzen. Zudem mangelt es an theoretischer Präzisierung dahingehend, wie stabil diese Präferenzen sein können. So ist in der Regel eher davon auszugehen, dass die vermeintlich ‚unabhängigen' Variablen unter dem Einfluss von Krankheitsverlauf, Person des Arztes, Art der Entscheidung etc. variieren, so dass mit einer einmaligen Messung im Vorwege keine valide Einschätzung für die gesamte Arzt-Patient-Kooperation gewonnen werden kann.

Abhängige Variablen. Werden bestimmte Variablen dagegen in Abhängigkeit von einem zuvor abgeschlossenen Entscheidungsprozess erfasst, so ist prinzipiell ein Rückschluss auf dessen Art möglich. Dabei ist jedoch zu hinterfragen, wie eindeutig dieser Rückschluss im Einzelfall ist. Wird etwa nach der ‚Zufriedenheit' mit einer Entscheidung gefragt (z.B. Post Decision Satisfaction Scale (PDSS) (Sainfort und Booske, 2000)), so ist diese sicherlich abhängig von den Informationen, die als Grundlage der Entscheidung gedient haben. Blendet eine Ärztin widersprüchliche oder zu komplex erscheinende Informationen für den Patienten aus, so dürfte diesem seine daraufhin getroffene Entscheidung ‚befriedigender' und ‚schlüssiger' vorkommen, als hätte er zusätzlich erfahren, wie fraglich die Stichhaltigkeit der verfügbaren Informationen sein mag. Informationen müssen schließlich nicht korrekt sein, um Zweifel auszuräumen (Brashers, 2001). Durch den kaum einzuholenden Wissensvorsprung könnte eine Ärztin vermutlich beinahe jeden Entscheidungsprozess so steuern, dass der Patient eine hohe Zufriedenheit mit seiner Entscheidung erlebt (Cleary, 1998). Eine gewisse Vorsortierung der Informationen darf dabei nicht ausschließlich kritisch gesehen werden. Denn Selektion ist angesichts des Missverhältnisses von verfügbaren Informationen zur Verarbeitungskapazität von Patienten (und Ärzten) unumgänglich. Den starken Einfluss selbst von vermeintlich harmloser Vorsortierung und Aufbereitung von Informationen beleuchtet später Abschnitt 3.2.1 etwas anschaulicher.

Das Problem ändert sich nicht wesentlich, wenn statt der ‚Zufriedenheit' nach dem verbliebenen ‚Entscheidungskonflikt' (DCS) gefragt wird. Beide Ansätze sind zusätzlich insofern kritisch zu betrachten, als SDM insbesondere in Situationen indiziert ist, die sich durch fehlende Evidenz oder zumindest vordergründige Gleichwertigkeit verschiedener Therapieoptionen auszeichnen. Daher konstatiert McNutt (2004, S. 2518):

> „However, no one should suggest that the goal of medical decision making is to make decision making easy. On the contrary, the goal should be to make it difficult; to make sure that patients understand that every decision is influenced by uncertainty and risk. The ideal is not to reduce decisional conflict, but to maximize it".

Unter diesem Aspekt erscheinen ‚Entscheidungszufriedenheit' und geringer ‚Entscheidungskonflikt' geradezu verdächtig.

Manipulation Check-Variablen. Es verbleiben die als Manipulation Check-Variablen zusammengefassten Ansätze. Entsprechend der bei den Beobachtungsverfahren angemeldeten Kritik konzentrieren diese sich auf die Wahrnehmung der Patienten, was ihre Stärke ausmacht. Ferner beziehen sie sich auf den Entscheidungsprozess selbst, nicht auf dessen Vorläufer oder Ergebnisse. Spätestens an dieser Stelle kommt man indes nicht um eine entscheidende Frage herum: Was ist der Kern des Shared Decision Making? Spontan möchte man antworten, der Kern geteilter Entscheidungsfindung sei das Teilen der Entscheidung. Mithin wäre es folgerichtig, SDM zu messen, indem das Ausmaß des Teilens der Entscheidung bestimmt würde, wie es etwa die Patients' Perceived Involvement in Care Scale (PICS) von Lerman et al. (1990) tut. Dieser Ansatz hat durchaus seine Berechtigung. Allerdings hat er auch Grenzen. Nicht selten kommt es z.B. vor, dass ein Patient *bewusst* nicht mitentscheiden *möchte*, sei es aus Überforderung mit der Krankheitssituation, aus mangelnder ‚Einsatzbereitschaft' oder auch aus dem gewachsenen Vertrauen heraus, dass sein Arzt die für ihn passende Entscheidung treffen werde. Die PICS würde in diesem Fall die für SDM (vermeintlich) zu geringe Patientenbeteiligung anzeigen. Und der behandelnde Arzt könnte nicht beanspruchen, die SDM-Leitlinien befolgt zu haben, obwohl er den Wünschen und Vorstellungen des Patienten genau entsprochen hat. Dieser Widerspruch führt zu der zunächst paradox erscheinenden Konsequenz, dass eine im Sinne des SDM gemeinsame Entscheidung nicht (immer) gemeinsam getroffen werden muss. Es geht also weniger darum, wer entscheidet, sondern darum, *wer entscheidet, wer entscheidet.*

Die Priorisierung der Patientenpartizipation gegenüber anderen Aspekten des SDM ist noch aus einem weiteren Grund unangebracht. Denn um dem Patienten die Entscheidung für die seinen persönlichen Präferenzen entsprechende Therapieoption zu ermöglichen, muss in erster Linie sichergestellt werden, dass er die dazu nötigen Informationen bekommt, versteht und zu einer Entscheidung verarbeiten kann. Mag Patientenpartizipation aus ethischen oder anderen Erwägungen heraus für sich genommen bereits einen Wert haben, so ist sie in Bezug auf die Güte des Entscheidungsprozesses kein Selbstzweck. Sie ist im Arzt-Patient-Gespräch vielmehr entweder *Ausdruck* einer vertieften Informationsverarbeitung oder ein *Vehikel*, um diese zu erzeugen (Kasper, 2005).

Weitere Outcomevariablen. Ähnlich unspezifisch in der Messung von SDM-Charakteristika in Abgrenzung zu anderen Kommunikationsstilen sind die oben erwähnten Variablen Behandlungszufriedenheit oder die davon insbesondere bei ernsthaften Erkrankungen eng zusammenhängende Lebensqualität. Untersuchungen zur Lebensqualität als Outcomevariable haben zusätzlich den Nachteil, dass diese Größe während einer Erkrankung ihre Bedeutung verändern kann. Mag nämlich eine Patientin mit Multipler Sklerose zu Beginn der Erkrankung ihre Lebensqualität weitgehend von körperlicher Funktionsfähigkeit bestimmt sehen, gewinnen häufig im Laufe einer solchen chronischen und mithin nicht heilbaren Erkrankung andere Aspekte wie etwa die Verlässlichkeit von sozialen Beziehungen an Einfluss für die Beurteilung der Lebensqualität, während die eingeschränkte körperliche Funktionsfähigkeit als kaum veränderlich akzeptiert wird. Gleichzeitig wäre jedoch auch eine umgekehrte Prioritätenverschiebung denkbar, indem der Wert der Funktionsfähigkeit steigt, je stärker er durch die fortschreitende Erkrankung bedroht ist. Ein Ergebnisparameter, dessen Definition sich zwischen verschiedenen Messpunkten verändern kann, ist dabei nicht nur in Bezug auf SDM von fraglicher Validität.

Nicht nur unspezifisch wie die Bestimmung der Lebensqualität, sondern geradezu konträr zu den Grundsätzen des SDM ist dagegen die Erhebung der gewählten Behandlung selbst als Outcomevariable einer SDM-Intervention, wie sie in manchen Studien erhoben wird. Denn SDM proklamiert ja gerade, dass es keine universell *richtigen* Entscheidungen gibt. Vielmehr wird die Mitteilung der *prinzipiellen Gleichwertigkeit* aller verfügbaren Therapieoptionen – die so genannte ‚Equipoise-Botschaft' – als Auftakt und notwendige Voraussetzung von SDM gesehen (Elwyn et al., 2005). Neben prinzipiellen Überlegungen werden auch deshalb die Präferenzen der Patienten im SDM so stark berücksichtigt (vgl. 1.4.1). Damit kann aber eine Bevorzugung einer Therapieoption gegenüber den übrigen kein Indiz für SDM und den dabei gebotenen Respekt vor der Entscheidungskompetenz der Patienten sein. Im Gegenteil: Steht bereits vor Anhörung der Patientensicht eine Behandlungsmethode als angestrebtes, vermeintlich gemeinsam ausgehandeltes Gesprächsziel fest, so zeugt dies eher von paternalistisch verhafteter Denkweise.

Aus demselben Grund verbietet es sich, den Behandlungserfolg als abhängige Variable einer SDM-Intervention zu erfassen. Zwar erscheint es zunächst plausibel, dass eine in Behandlungsentscheidungen einbezogene Patientin über die in 1.4.1 diskutierten Wirkmechanismen eher von einer Therapie profitiert, als wenn man

ihr eine Behandlung gegen ihre Überzeugung aufgedrängt hätte. Es können allerdings auch andere Effekte wirksam werden. Man stelle sich vor, jene Patientin litte an einem Karzinom im präfinalen Stadium. Unter Umständen ließen sich ihre verbleibende Lebenszeit von sechs auf neun Monate verlängern und ihre Schmerzen lindern, wenn sie sich einer stationären Chemotherapie unterzöge. Im Rahmen eines SDM-Gesprächs mit ihrer Ärztin könnte sie sich allerdings gegen eine Behandlung im Krankenhaus entscheiden, um stattdessen die ihr verbleibende Zeit zu Hause mit ihrer Familie zu verbringen. In diesem Beispiel würden die dem SDM inhärente vollumfängliche Aufklärung über Nutzen *und* Schaden einer Behandlung und die Umsetzung der Patientenentscheidung also sogar zu einer verkürzten Lebenszeit und vermutlich stärkeren Schmerzen führen, als wenn man die Patientin paternalistisch zu der rein medizinisch gerechtfertigten Chemotherapie gedrängt hätte. So gesehen scheidet eine Messung von Behandlungserfolg als Erfolgsmaß für SDM in der Regel aus (Kennedy, 2003).

Fazit. Zusammenfassend lässt sich festhalten, dass keiner der im Rahmen des SDM entwickelten Messansätze sowohl in methodischer als auch in theoretischer Hinsicht überzeugt (Elwyn et al., 2001; Giersdorf et al., 2003; Kasper et al., 2003; Simon et al., 2005a). Daher soll in den folgenden Kapiteln zunächst ein theoretisch fundiertes Konzept von SDM vorgestellt werden (Kapitel 4), um davon in Teil II dieser Arbeit ein Messinstrument abzuleiten und zu validieren.

Kapitel 3

Theoretische Ansätze zur Entscheidungsfindung

Mehrfach wurde im Verlauf der bisherigen Erörterungen die Vermutung geäußert, dass sich die unbefriedigende Validität der Messmethoden im Bereich SDM zu einem Gutteil aus dem Mangel an Präzision im Konzept des SDM speist. Mit anderen Worten: Man kann kaum erwarten, dass Messverfahren Charakteristika reliabel und eindeutig erfassen, wenn keine Klarheit darüber besteht, was diese Charakteristika überhaupt sind. Um diesbezüglich Abhilfe zu schaffen, gibt dieses Kapitel einen Überblick über Theorien der Entscheidungsfindung im Allgemeinen. Das folgende Kapitel 4 knüpft daran an, um daraus eine theoretische Basis für SDM abzuleiten.

3.1 Normative Entscheidungstheorien

Wie treffen Menschen Entscheidungen? Wonach entscheiden sie, welche Hose sie anziehen, ob sie ihren Arbeitsvertrag kündigen oder ob sie sich den Blinddarm entfernen lassen?

Normative Entscheidungstheorien bieten eine theoretische Leitlinie, wie sich Menschen entscheiden *sollten*. Sie entwerfen dabei den ultimativ rational urteilenden Menschen. Allen normativen Entscheidungstheorien ist gemein, dass sie jene Entscheidungsoption herausfinden helfen sollen, die die günstigste Kombination aus ihrem Nutzen und dessen Auftretenswahrscheinlichkeit bieten.

3.1.1 Expected Value Theory

Wenn der Wert einer Entscheidung respektive deren Resultat bekannt ist, und dieses Resultat mit hundertprozentiger Sicherheit eintritt, sollte man sich im Dienste der Wertmaximierung für die Option mit dem höchsten Wert entscheiden. Damit ist ein Spezialfall der ‚Expected Value Theory' (EVT) beschrieben, bei dem die Auftretenswahrscheinlichkeit der Ereignisse 1 beträgt. Die allgemeine Variante beinhaltet die Möglichkeit, dass der erhoffte Wert einer jeden Entscheidungsoption nur mit einer gewissen, bekannten Wahrscheinlichkeit auftritt. Beim Roulette etwa beträgt der Gewinn (Wert) beim Setzen auf Rot das Doppelte des Einsatzes. Die Eintrittswahrscheinlichkeit des Gewinns liegt bei $\frac{18}{37}$ (18 rote Felder stehen 18 schwarzen und einem grünen gegenüber). Anhand der allgemeinen Formel der EVT lässt sich ableiten, ob eine Wette auf Rot sinnvoll ist. Sie lautet:

$$EV = \sum_i p_i v_i$$

wobei EV den erwarteten Gewinn, p_i die Wahrscheinlichkeit eines jeden Resultats i und v_i deren jeweiligen Wert bezeichnet. Da bei Eintreten des Ereignisses Rot der Gewinn so hoch ist wie der Einsatz s und der Verlust im anderen Fall ebenso groß, ergibt sich:

$$EV = \sum_{i=1}^{2} \frac{18}{37}s + \frac{19}{37}(-s) = -\frac{1}{37}s$$

Unabhängig von der Höhe des Einsatzes ist der erwartete Gewinn also negativ, weshalb man Roulettetische nach der EVT meiden sollte. Dies gilt im Übrigen analog für alle anderen Wetten des Roulettes.

3.1.2 Expected Utility Theory und Subjective Expected Utility Theory

Eine starke Idealisierung der EVT liegt in der Annahme, dass der *Wert* der möglichen Resultate objektiv feststellbar sei. Diese Restriktion lässt sich überwinden,

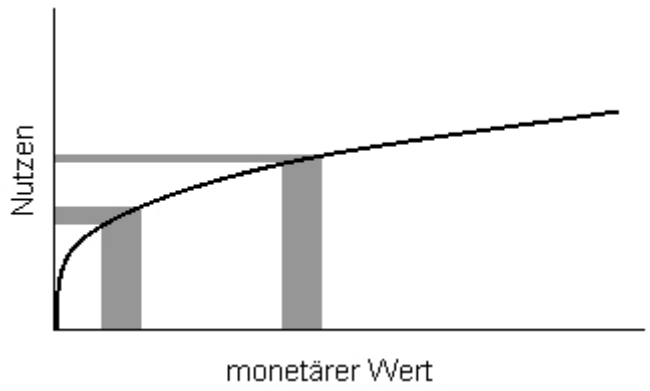

Abbildung 3.1: Die Beziehung zwischen monetärem Wert und dessen subjektiv erlebtem Nutzen. Dessen Zuwachs ist bei höherem monetären Wert geringer.

indem der Wert v_i (Value) einer jeden Option durch einen individuell zu bestimmenden Nutzen (Utility) ersetzt wird, wie es die ‚Expected Utility Theory' (EUT) umsetzt (Neumann und Morgenstern, 1947). Die ‚Subjective Expected Utility Theory' (SEUT) (Savage, 1954) erlaubt zusätzlich, subjektive Einschätzungen der Eintrittswahrscheinlichkeiten zu berücksichtigen.

Die Unterscheidung zwischen Value und Utility ermöglicht es dabei, durch Aussagen über die Beziehung zwischen diesen beiden Größen den Anforderungen alltäglicher Entscheidungsaufgaben näher zu kommen. Nähme man eine lineare Beziehung zwischen dem (monetären) Wert eines Gewinns und dessen (subjektivem oder objektivem) Nutzen an, so ergäbe sich die Situation, dass sich der Teilnehmer an einem Preisausschreiben über den Gewinn von sieben Kaffeemaschinen siebenmal so intensiv freuen müsste wie über nur eine Maschine. Stattdessen kann für viele Situationen eine annähernd logarithmische Beziehung zwischen Wert und Nutzen angenommen werden, die der nachlassenden Attraktivität jeder weiteren Kaffeemaschine Rechnung trägt (3.1).[1]

3.2 Deskriptive Entscheidungstheorien

Während die normativen Entscheidungstheorien Leitlinien bieten, wie man sich aus rationaler Sicht entscheiden *sollte*, wurden in den vergangenen Jahrzehnten unzäh-

[1] Dieser Gedanke geht bereits auf Bernoulli (1738) zurück, der ihn aus dem sog. ‚Sankt Petersburg-Paradox' entwickelt hat.

lige Situationen untersucht, in denen Menschen von diesen idealen Entscheidungsregeln abweichen (z.B. Kriz, 1982). Die Abweichungen sind dabei keine Folge von Rechenungenauigkeiten bei den Experimentsteilnehmern, sondern verweisen einerseits auf die Empfindlichkeit von Menschen gegenüber den die eigentliche Entscheidungssituation begleitenden Faktoren. Andererseits enthüllen sie die menschlichen Strategien, die bei vielen unterschiedlichen Entscheidungstypen angewandt werden.

3.2.1 Rahmungseffekte

Der sachliche Gehalt einer Information zu Vor- und Nachteilen von Handlungsoptionen ist leitend für eine Entscheidung zwischen diesen Optionen. Dass es daneben auch wesentlich auf die kognitive und emotionale Einordnung jener Sachinformation ankommt, ist eindrucksvoll von Tversky und Kahneman (1981) demonstriert worden. Sie definieren die ‚Rahmung' (Framing) einer Entscheidung als die „Konzeption [eines Entscheidenden] über die Handlungen, Handlungsergebnisse und Kontingenzen, die mit einer bestimmten Wahlmöglichkeit verbunden sind".[2] Dabei können bereits kleine Unterschiede in der Präsentation von Informationen zu stark unterschiedlichen Rahmungen und mithin Entscheidungen führen, wie das folgende Beispiel zeigt:

> **Darbietungsform 1**
> Die Vereinigten Staaten bereiten sich auf den Ausbruch einer ungewöhnlichen asiatischen Krankheit vor, die vermutlich 600 Menschen töten wird. Zwei Aktionsprogramme wurden zur Bekämpfung der Krankheit vorgeschlagen. Die exakten wissenschaftlichen Schätzungen der Konsequenzen beider Programme lauten:
>
> Wenn Programm A durchgeführt wird, werden 200 Menschen gerettet.
> Wenn Programm B durchgeführt wird, gibt es eine $\frac{1}{3}$-Wahrscheinlichkeit, dass 600 Menschen gerettet werden, und eine $\frac{2}{3}$-Wahrscheinlichkeit, dass niemand gerettet wird.
>
> Welches Programm würden Sie bevorzugen?

[2] Übersetzung aus Solso (2005, S. 400)

Darbietungsform 2
Die Vereinigten Staaten bereiten sich auf den Ausbruch einer ungewöhnlichen asiatischen Krankheit vor, die vermutlich 600 Menschen töten wird. Zwei Aktionsprogramme wurden zur Bekämpfung der Krankheit vorgeschlagen. Die exakten wissenschaftlichen Schätzungen der Konsequenzen beider Programme lauten:

Wenn Programm A durchgeführt wird, werden 400 Menschen sterben.
Wenn Programm B durchgeführt wird, gibt es eine $\frac{1}{3}$-Wahrscheinlichkeit, dass niemand stirbt, und eine $\frac{2}{3}$-Wahrscheinlichkeit, dass 600 Menschen sterben werden.

Welches Programm würden Sie bevorzugen?

Dem aufmerksamen Leser wird aufgefallen sein, dass einerseits der sachliche Gehalt der Informationen in den beiden Darbietungsformen identisch ist: Bei der ersten Version wird die Information als ‚Anzahl Geretteter' angegeben, in der zweiten Version wird der Informationsgehalt über das Gegenereignis, nämlich die ‚Anzahl nicht Geretteter' transportiert. Andererseits ist obendrein der Erwartungswert dieser beiden Anzahlen für Programm A und B jeweils gleich, wie man leicht anhand der obigen Formel der EVT berechnen kann.
Obschon also die durchschnittlich zu erwartenden Konsequenzen für die Betroffenen in allen vier Varianten identisch sind, entschieden sich in der ersten Darbietungsform etwa drei Viertel der Versuchspersonen für Programm A und ein Viertel für Programm B, während die Verhältnisse bei Darbietungsform 2 genau umgekehrt waren. Offensichtlich macht es einen großen Unterschied, ob sich Menschen die Konsequenzen einer Handlung anhand eines Ereignisses oder anhand dessen Komplements vergegenwärtigen.

Auch in einem anderen Beispiel bewirkt die unterschiedliche Formulierung desselben mathematischen Sachverhalts stark divergierende Vorstellungen, die handlungsleitend sein können: Verspricht ein Pharmahersteller, dass sein Medikament das Risiko eines Herzinfarkts halbieren kann, so verkauft er vermutlich mehr, als wenn er angäbe, dass das Risiko von 1:10000 auf 0,5:10000 sinkt. Die Formulierung in Form relativer Risikoreduktion führt dabei zu einer anderen Rahmung als bei der Beschreibung als absolute Risikoreduktion.

3.2.2 Entscheidungsheuristiken

Zudem scheinen Entscheidende häufig auf verschiedene Daumenregeln oder Heuristiken zurückzugreifen, anstatt die verfügbaren Informationen sachlich zu bewerten. Wieder hilft ein Experiment von Tversky und Kahneman (1973) bei der Veranschaulichung:

> **Häufigkeit von Buchstaben**
> Die Auftretenshäufigkeit von Buchstaben in der englischen Sprache ist untersucht worden. Ein typischer Text wurde ausgewählt, um die Häufigkeiten festzustellen, mit denen verschiedene Buchstaben des Alphabets an erster und an dritter Stelle eines Wortes vorkommen. Wörter mit weniger als drei Buchstaben wurden ausgeschlossen.
>
> Wir bitten Sie, anhand der nachfolgend dargebotenen Buchstaben jeweils zu entscheiden, ob diese häufiger an erster oder an dritter Stelle auftreten. Außerdem schätzen Sie bitte das Verhältnis zwischen diesen Häufigkeiten.

In diesem Experiment überschätzen Teilnehmer systematisch – also unabhängig vom Buchstaben – die Häufigkeit des Auftretens am Wortanfang. Tversky und Kahneman sehen darin den Hinweis auf eine Heuristik, wie Menschen Schätzungen über Wahrscheinlichkeiten machen. Nach der ‚Verfügbarkeitsheuristik' (Availability Heuristic) überschätzen Entscheidende die Wahrscheinlichkeiten von Dingen, die sie sich leichter vorstellen können, die also kognitiv leichter verfügbar sind. Und da es leichter fällt, Wörter z.B. mit „f" am Anfang aufzuzählen als solche mit „f" an dritter Stelle – wie man leicht ausprobieren kann – wirkt sich diese Entscheidungsheuristik in der gestellten Aufgabe aus.

Eine andere Heuristik ist die ‚Repräsentativitätsheuristik' (Representativeness Heuristic, Tversky und Kahneman, 1974). Fragt man Versuchsteilnehmer, ob bei einem zufälligen Würfelexperiment die Abfolge 1-2-3-4-5-6 oder die Abfolge 3-2-5-5-1-4 wahrscheinlicher ist, so entscheiden sie sich mehrheitlich für die zweite Variante. Tatsächlich sind beide Varianten gleich wahrscheinlich, nämlich $(\frac{1}{6})^6$. Die Autoren erklären solche Effekte damit, dass eine ungeordnete Zahlenfolge eher dem entspricht, was man sich unter einem Zufallsexperiment vorstellt. Sie ist ‚repräsentativ' und wird wegen ihrer Plausibilität als häufiger angenommen.

3.2.3 Die Rationalität menschlicher Entscheidungen

Die in diesem Kapitel vorgenommenen Betrachtungen über Beeinflussungen von Entscheidungen durch Rahmungen und Heuristiken beanspruchen keine Vollständigkeit. Vielmehr sollen sie andeuten, wie man sich die Brücke zwischen den idealen, normativen Entscheidungstheorien und der alltäglichen Entscheidungstätigkeit von Menschen und ihren Möglichkeiten im Umgang mit Werten und Wahrscheinlichkeiten vorstellen kann. Weitere Rahmungseffekte, Beeinflussungen und Heuristiken sind dem Sammelband von Kahneman, Slovic und Tversky (1982) zu entnehmen. Den Versuch einer Integration vieler Befunde unternehmen Kahneman und Tversky mit ihrer ‚Prospect Theory' (1979).

Doch auch dieser kurze Einblick in die Informationsverarbeitungsprozesse bei Entscheidenden regt zu der Frage an, wie rational oder gar sinnvoll diese agieren. Angesichts offensichtlicher Fehleinschätzungen wie im obigen Beispiel zur Repräsentativitätsheuristik (vgl. 3.2.2) könnte man hier zu einem negativen Urteil kommen. Andererseits hat diese Diskussion Parallelen: In der Wahrnehmungspsychologie wurde zuweilen an der Zuverlässigkeit der Sinnesorgane gezweifelt, da sich doch etwa das visuelle System durch teilweise erstaunlich simpel konstruierte ‚Optische Täuschungen' wie die Müller-Lyer-Figur oder das Kanizsa-Dreieck irreführen lässt. Doch die Bewertung seiner Leistungsfähigkeit hängt eng mit der Definition seiner Aufgaben und Funktionen zusammen. Wer es für die Aufgabe des visuellen Systems hält, ein veridikales Abbild der Außenwelt zu liefern, der wird dessen Leistung für ungenügend halten. Definiert man dagegen als oberste Priorität eine möglichst flexible, reaktionsschnelle und Ressourcen sparende Anpassung an die gängigsten Lebensanforderungen, dann nimmt man damit einhergehende Fehleinschätzungen in ungewöhnlichen, artifiziellen Situationen billigend in Kauf. Dabei wird auch die Leistung des Wahrnehmungssystems erst ermöglicht durch den Einsatz von internalisierten Heuristiken – etwa die sog. ‚Grey World Assumption' (Buchsbaum, 1980; Helmholtz, 1896) oder die Gestaltgesetze.

Ganz ähnlich argumentiert Gigerenzer, die Effekte von Tversky und Kahneman beruhten zum Teil auf der mangelnden Gegenständlichkeit des experimentellen Materials und verringerten sich, wenn etwa „16 von 100 Personen" statt „Wahrscheinlichkeit von $\frac{4}{25}$" als Umschreibung verwendet würde (Gigerenzer, 1996). Wiederum analog zur Diskussion um die visuelle Wahrnehmung führe die Untersuchung von Urteilen und Entscheiden unter Laborbedingungen zu „cognitive illusions" (Cohen, 1981).

Als Gegenentwurf präsentierten Gigerenzer und Goldstein (1996) eine Gruppe von Heuristiken, die weniger präzise als normative Modelle, aber dennoch (oder gerade deswegen) im Wesentlichen hilfreich in alltagsnahen Entscheidungssituationen waren – die sog. ‚Fast and Frugal'-Heuristiken.

Als Fazit ist festzuhalten, dass die Befunde von Tversky, Kahneman und anderen für sich genommen inhaltsvalide sind. Sie enthüllen, dass Entscheidende unter bestimmten Bedingungen nicht nur graduell, sondern systematisch von einer unter bestimmten Axiomen idealen rationalen Entscheidungsstrategie abweichen, und leisten damit einen unverzichtbaren Beitrag zum Verständnis menschlicher Entscheidungsfindung. Dazu sind *gerade* artifizielle Situationen hilfreich. Daraus sollte jedoch nicht der Schluss gezogen werden, die menschliche Entscheidungsfähigkeit sei grundsätzlich suboptimal. Sie ist vielmehr auf andere Zielparameter optimiert als die normativen Entscheidungstheorien. Evans (1989) formuliert es so:

> „The view that I wish to argue here is that errors of thinking occur because of, rather than in spite of, the nature of our intelligence. In other words, they are an inevitable consequence of the way in which we think and a price to be paid for the extraordinary effectiveness with which we routinely deal with the massive information-processing requirements of everyday life."[3]

Wie immer man das menschliche Entscheidungsvermögen bewerten und ob man es rational oder irrational nennen mag, kann doch als Konsequenz für die Einbindung von Patienten festgehalten werden, dass die Einschätzung von Heilungswahrscheinlichkeiten und Rückfallrisiken keinesfalls so selbstverständlich ist, wie es eine Prozentzahl suggeriert.[4] Dies gilt umso mehr, als Framing-Effekte unter dem Einfluss chronischer oder gar lebensbedrohlicher Krankheiten eher häufiger zu erwarten sind als bei der Schätzung der Häufigkeit von Buchstaben.

[3] Analog äußerte sich vor 150 Jahren Hermann von Helmholtz in Bezug auf die visuelle Wahrnehmung: „Das Sinnesorgan täuscht uns dabei nicht, es wirkt in keiner Weise regelwidrig, im Gegenteil, es wirkt nach seinen festen, unabänderlichen Gesetzen und es kann gar nicht anders wirken. Aber wir täuschen uns im Verständnis der Sinnesempfindung" (Helmholtz, 1855, S. 100).

[4] Zumal nicht einmal die Annahme selbstverständlich ist, dass ein direkter Zusammenhang zwischen wissenschaftlich bestimmter, ‚objektiver Wahrscheinlichkeit' und deren subjektiver Repräsentation besteht. Kriz (1982) vermutet vielmehr, dass die etwa von der Prospect Theory postulierten Beziehungen zwischen objektiven und subjektiven Wahrscheinlichkeiten ein Artefakt aus der Untersuchung zu simpler Entscheidungssituationen sind.

Kapitel 4

Die Rolle von Ungewissheit bei der medizinischen Entscheidungsfindung

Letztlich enthält fast jeder Ansatz im Bereich SDM Elemente der normativen Entscheidungstheorien. So wird betont, dass einerseits die Präferenzen des Patienten eruiert werden sollen; damit wird implizit die Variable Utility angesprochen. Aufgabe des Arztes ist es andererseits, die medizinisch fundierten, für den ‚Durchschnittspatienten' idealisierten Wahrscheinlichkeiten zu den einzelnen Handlungsergebnissen bereitzustellen und verständlich zu erklären, die durch den Parameter p_i in den Gleichungen von EVT und EUT ausgedrückt werden (vgl. 3.1.1 bzw. 3.1.2). Auch wenn auf der konkreten Handlungsebene zwischen Arzt und Patient selten explizit über „Nutzen" oder „Erwartungswert" gesprochen wird, so ist doch die Aufgabe, die die beiden Beteiligten zu bewältigen haben, abstrakt analog zur Bestimmung der Parameter v_i und p_i.

Die Situation medizinischer Entscheidungsfindung ist jedoch ungleich komplexer als beim Roulette: So ist einerseits davon auszugehen, dass Patienten ihre eigenen Präferenzen und Prioritäten in der Regel nicht kennen, da sie meist noch nie darüber haben nachdenken müssen. Und selbst wenn es sich um eine wiederholt auftretende Krankheitssituation handelt – etwa die Frage einer möglichen Operation zur Behandlung eines akuten Schubes einer chronisch entzündlichen Darmkrankheit –, werden sich die Präferenzen wahrscheinlich verschoben haben. Demnach müssen Patienten im gemeinsamen Gespräch mit dem Arzt, den Angehörigen oder für sich selbst erst

zu klären versuchen, worin ihre Prioritäten bestehen. Die Bestimmung der Parameter v_i ist also gekennzeichnet von der Auseinandersetzung mit *Ungewissheit*.

Substantielle Ungewissheit ist darüber hinaus auch bei der Festlegung der Wahrscheinlichkeitsparameter p_i unumgänglich. Denn nach Charles et al. (1997) ist SDM ja gerade dann besonders angebracht, wenn keine eindeutige Überlegenheit einer bestimmten Behandlungsoption feststellbar ist, da zu den jeweiligen Optionen die medizinische Evidenz uneindeutig oder schlicht nicht vorhanden ist. Damit ist ungewiss, inwieweit ein Medikament wirkt. Doch selbst bei gesicherter Evidenz bleibt es für einen Patienten ungewiss, ob ein Medikament *bei ihm persönlich* wirken wird, wenn es erfahrungsgemäß bei 70 von 100 Patienten wirkt. Dieser vermeintliche Deduktionsschluss von einem idealisierten ‚Durchschnittspatienten' auf einen Einzelnen ist nicht zwingend. Dasselbe gilt für die Nebenwirkungen.

Ferner ist bereits in der historischen Entwicklung der normativen Entscheidungstheorien ein Problem ins Blickfeld geraten, das an einem Beispiel erörtert werden soll. Gegeben sind die folgenden drei Situationen:

1. In einer Urne befinden sich schwarze und weiße Kugeln, deren Zahlenverhältnis unbekannt ist.

2. In einer zweiten Urne befinden sich schwarze und weiße Kugeln in unbekanntem Zahlenverhältnis. Bei einer Stichprobe aus dieser Urne werden beträgt das Verhältnis zwischen den Farben 1:1.

3. In einer dritten Urne befinden sich gleich viele schwarze und weiße Kugeln.

Für jede der Urnen müsste man eine Chance von $\frac{1}{2}$ annehmen, bei einmaligem Ziehen eine weiße Kugel zu entnehmen.[1] Die ‚Verlässlichkeit' oder ‚Gewissheit' dieser Erwartung ist jedoch von Beispiel zu Beispiel unterschiedlich.

Übertragen auf medizinische Entscheidungen bedeutet dies, dass nicht nur die Heilungswahrscheinlichkeiten für verschiedene Therapien bestimmt werden müssen. Wichtig ist darüber hinaus auch die Einschätzung der Verlässlichkeit der Quellen dieser

[1] Hier wird bewusst nicht von „Wahrscheinlichkeit" gesprochen, da es sich nicht in allen Fällen um dieselbe Art von Wahrscheinlichkeit im mathematischen Sinne handelt. Vielmehr ist die Wahrscheinlichkeit in den ersten beiden Urnen womöglich nicht $\frac{1}{2}$; da dies aber unbekannt ist, ist es sinnvoll, eben dies *anzunehmen*.

Information sowie der Information selbst. Dabei ist einerseits die Ungewissheit bezüglich der medizinischen Evidenz zu beurteilen, wobei randomisiert-kontrollierte Studien mehr Anhalt geben könnten als mündliche Erfahrungsberichte. Anderseits ist auch die Ungewissheit bezüglich des Arztes selbst relevant, der schließlich eine wesentliche Vorauswahl, Gewichtung und Interpretation von den insgesamt verfügbaren Daten vornimmt.

Schließlich ist nachvollziehbar, dass gerade bei schweren Krankheiten die Unterstützung von Angehörigen wichtig ist. Insofern stellt die Ungewissheit bezüglich der Verlässlichkeit sozialer Bindungen eine wichtige Einflussgröße dar. So könnte beispielsweise eine Partnerschaft von fraglicher Stabilität zu einer anderen Prioritätensetzung bei der Beurteilung der Behandlungsoptionen führen als eine belastbare Beziehung.

Obschon hier teilweise nur hypothetische Beispiele diskutiert wurden, lässt sich festhalten, dass Ungewissheit bei medizinischen Entscheidungen eine entscheidende Rolle spielt (Babrow et al., 1998; Brashers, 2001; Brashers et al., 2003; Bottorff et al., 1998; Kasper et al., 2003). Daher widmet sich der folgende Abschnitt dem Thema systematischer.

4.1 Definition und Abgrenzung von Ungewissheit

„Uncertainty exists when the probability of an event occurring is not 0 or 1“ (Good, zit. nach Kozine, 2004, S. 3). Ungewissheit beschreibt also einen Zustand des ‚Nicht-Wissens‘ oder ‚Nicht-Sicher-Wissens‘ in Bezug auf einen Sachverhalt. Dabei ist wenigstens prinzipiell eine Information denkbar, die zur Auflösung dieses Zustands führt. Das Erleben von Ungewissheit ist rein (intra-) subjektiv zu sehen: Selbst wenn viele Informationen und ausreichend Rezeptionszeit zur Verfügung stehen, mag eine Person Ungewissheit erleben, eine andere in derselben Situation dagegen nicht (Brashers, 2001).

Da sich das Konstrukt Ungewissheit primär auf die kognitive Einschätzung bezieht, ist es klar abzugrenzen von emotionalen *Bewertungen* derselben Situation, die sich in ‚Unsicherheit‘, ‚Stress‘ oder ‚Angst‘ äußern können. Gleichzeitig kann jedoch das Erleben von Ungewissheit zu Unsicherheit führen – oder auch zu Hoffnung, Neugier oder Nervenkitzel. Damit wird deutlich, dass Ungewissheit an sich weder positive

noch negative Valenz besitzt, sondern erst durch zusätzliche Attributionen als angenehm oder unangenehm wahrgenommen wird. Diese Attributionen variieren dabei abhängig von Situation und Disposition (Brashers et al., 2000; Sorrentino und Roney, 2000).

4.2 Taxonomien von Ungewissheit

Die Vielschichtigkeit von Ungewissheit hat zu verschiedenen Taxonomien geführt. In einer eher theoretisch und technisch inspirierten Herangehensweise unterscheidet Rowe (1994) ‚metrische', ‚strukturelle', ‚temporale' und ‚Übertragungsungewissheit'.[2] Metrische Ungewissheit bezieht sich auf Ungenauigkeiten in der Messung oder Bestimmung relevanter Einflussgrößen. Strukturelle Ungewissheit entsteht durch eine nicht (in Gänze) überschaubare Komplexität eines Sachverhalts. Temporäre Ungewissheit geht auf nicht rekonstruierbare Zustände in der Vergangenheit oder unvorhersehbare Entwicklungen in der Zukunft dieses Sachverhaltes zurück. Übertragungsungewissheit schließlich entsteht, wenn Analysen und Einschätzungen von verschiedenen Personen mit womöglich unterschiedlichen Perspektiven oder divergierendem Sprachgebrauch zu einer Gesamtbewertung integriert werden sollen. Obgleich ursprünglich nicht in erster Linie für den medizinischen Kontext entwickelt, lassen sich alle vier Typen von Ungewissheit leicht auf Therapieentscheidungen zwischen Arzt und Patient übertragen. Auch dort trifft die Charakterisierung von Rowe zu, dass in den meisten risikobehafteten, kritischen Entscheidungssituationen der Anteil der verfügbaren Information kleiner ist als der der fehlenden.

Aus einer qualitativen Untersuchung des Ungewissheitserlebens von HIV-Patienten entwickelten Brashers et al. (2003) die folgende Aufteilung:

- ‚Medizinische Ungewissheit' speist sich aus unzureichenden Informationen zu Diagnose, Prognose, der Komplexität des Krankheitsgeschehens und dessen Behandlung.
- ‚Persönliche Ungewissheit' entsteht in der Auseinandersetzung mit der Selbstkonzeption als Kranker bzw. als ‚normaler Mensch' sowie durch unklare finanzielle Konsequenzen der Erkrankung.

[2] Im englischen Original lauten die Bezeichnungen „metrical", „structural", „temporal" und „translational uncertainty".

- ‚Soziale Ungewissheit' resultiert bei HIV-Patienten aus der Unvorhersagbarkeit interpersonaler Reaktionen auf die Krankheit, der Spanne zwischen Akzeptanz und Abweisung oder Stigmatisierung sowie der latenten Gefahr, andere zu infizieren.

Manche der hier aufgeführten Ungewissheitsquellen sind in dieser Form nur für HIV-Patienten relevant. Dennoch zeigt die Einteilung gewisse Parallelen zur nachfolgenden.

Aus einer qualitativen Untersuchung mit Eltern chronisch kranker Kinder extrahierte Cohen (1993) die Kategorisierung von Ungewissheit in ‚ereignisbezogen', ‚zeitlich', ‚ätiologisch', ‚behandlungsbezogen' und ‚prognostisch'.

Mishel geht noch einen entscheidenden Schritt weiter: Mit ihrer ‚Mishel Uncertainty in Illness Scale' (MUIS) hat sie erstmals ein Messinstrument vorgelegt, das Ungewissheit quantitativ erfasst. Nach einer ersten Version mit zwei mäßig überzeugend entwickelten Faktoren (Mishel, 1981) sind die 34 Items aktuell zu vier Faktoren gruppiert (Mishel et al., 1984). Diese sind mit ‚Ambiguity', ‚Complexity', ‚Lack of information' und ‚Unpredictability' überschrieben. Die MUIS ist ein weit verbreitetes, im Wesentlichen bewährtes Instrument, das in verschiedene Sprachen übersetzt wurde. Eine deutsche Version existiert jedoch nicht.

Bei der Betrachtung der hier vorgestellten Taxonomien fällt auf, dass zwei unterschiedliche Ebenen von Ungewissheit miteinander vermischt werden. Teilweise werden *Themen* von Ungewissheit gruppiert, etwa ‚medizinisch', ‚persönlich' oder ‚behandlungsbezogen'. An anderer Stelle wird unterschieden nach *Auslösern* von Ungewissheit, z.B. ‚Ambiguity' oder ‚Unpredictability'. Die Beziehung zwischen diesen beiden Ebenen ist klar: Jedes Ungewissheitsthema kann durch unterschiedliche Auslöser gespeist werden. So mag beispielsweise die Ungewissheit bezüglich des derzeitigen Krankheitszustands u.a. aus der Ambiguität der Symptome sowie dem Mangel an aktuellen diagnostischen Informationen hervorgehen.

4.3 Shared Decision Making als Verhandlung von Ungewissheiten

Wie aber lässt sich aus der Betrachtung von Ungewissheit die in Kapitel 2 in Aussicht gestellte theoretische Basis für SDM ableiten? In den vorangegangenen Abschnitten wurde sowohl theoretisch als auch praktisch hergeleitet, dass der Umgang mit Ungewissheit ein zentrales Element bei Entscheidungsprozessen ist. Arzt und Patient müssen sich in irgendeiner Weise gegenüber Ungewissheiten verhalten, sei es durch aktive Informationssuche, Akzeptanz oder Ignoranz (Mishel, 1999). Aufgrund der Komplexität des Entscheidungsgegenstands sowie der begrenzt verfügbaren Informationen kann es nicht Ziel, Gegenstand oder Ergebnis von SDM sein, Ungewissheiten um jeden Preis zu reduzieren oder gar auszuräumen[3] (Brashers, 2001). Stattdessen wird hier postuliert, dass SDM in der *konsensuellen Validierung* der jeweils von Arzt und Patient erlebten Ungewissheiten besteht. Charles' ‚Two way exchange' (1997) wäre also zu präzisieren als ein wechselseitiger Austausch über die dem Entscheidungsgegenstand inhärenten Aspekte von Ungewissheit.

Wenn SDM als eine Verhandlung von Ungewissheiten beschrieben werden kann (Kasper et al., 2003), stellt sich aus messmethodischer Perspektive die Frage, worin das Ergebnis eines erfolgreichen SDM-Prozesses besteht. Da, wie schon erwähnt, die Ungewissheit nicht als Ganzes reduzierbar ist, kann mit einer eindimensionalen Konzeption von Ungewissheit kein SDM-Effekt abgebildet werden. Denn diese könnte nur das Ausmaß von Ungewissheit insgesamt widerspiegeln. Auch eine Differenzierung von Ungewissheit nach verschiedenen ‚Auslösern' wie Ambiguität, Unvorhersagbarkeit etc. (vgl. 4.2) hilft hier nicht wesentlich weiter. Denn auch für diese Aspekte gilt das Axiom der globalen Nichtreduzierbarkeit. Stattdessen bietet eine Kategorisierung unterschiedlicher ‚Themen' oder ‚Felder' von Ungewissheit die Chance, einer Reduktion von Ungewissheit in Teilbereichen Rechnung zu tragen. So könnte es beispielsweise durch den intensiven Austausch zwischen Arzt und Patient sehr wohl gelingen, dass die Ungewissheit bezüglich der Vertrauenswürdigkeit des Arztes reduziert wird, während gerade durch dessen offene Aufklärung auch über die Zweifel an den verfügbaren Therapieoptionen die Prognoseungewissheit sogar bis auf einen der medizinischen Realität angemessenen Wert *anwachsen* könn-

[3] Wenngleich die Reduktion des Entscheidungskonflikts, wie ihn die Decisional Conflict Scale erfasst, genau darauf abzielt (vgl. 2.7).

te (Grant et al., 2001). Dieses Beispiel macht gleich mehreres deutlich: Neben der partiellen Reduktion von Ungewissheit erwächst aus einem erfolgreichen SDM-Gespräch bei den Beteiligten auch eine Differenzierung zwischen zu akzeptierenden, unveränderlichen Ungewissheiten und solchen, die durch Informations- oder Handlungsprozesse günstig beeinflusst werden können. Damit ist in der Regel nach einer SDM-Intervention eine Veränderung im *Ungewissheitsprofil* zu erwarten, d.h. das Ausmaß erlebter Ungewissheit variiert zwischen den einzelnen Ungewissheitsqualitäten stärker. Eine indifferente, global erlebte und damit lähmende Ungewissheit wird in handhabbare, differenzierte Ungewissheitsqualitäten transformiert, für die unter Umständen auch jeweils unterschiedliche Umgangsweisen (‚Coping') denkbar sind. Darüber hinaus mag bereits der Konsens zwischen Arzt und Patient über die vorhandenen Ungewissheiten an sich eine entlastende Wirkung auf beide Beteiligten haben im Vergleich zum vorher meist diffus gebliebenen Ungewissheitserleben.

Zusammenfassend lässt sich also sagen, dass der Effekt von SDM als eine differenziertere Repräsentation der entscheidungsrelevanten Ungewissheitsqualitäten beschrieben werden kann. Um diese Repräsentation empirisch abzubilden, bedarf es eines multidimensionalen, veränderungssensiblen Messinstruments zur Erfassung von Ungewissheiten. Teil II schildert die Entwicklung und Validierung eines solchen Instruments für die von Patienten erlebten Ungewissheiten.

Teil II

Der QUICC – Ein Fragebogen zu Qualitäten von Ungewissheit in chronischen Erkrankungen

Kapitel 5

Die Entwicklung des QUICC

Im vorangegangenen Kapitel wurde gezeigt, inwiefern Ungewissheit ein Charakteristikum von medizinischen Entscheidungen darstellt. Ferner wurde klargestellt, warum eine ‚eindimensionale' Betrachtung von Ungewissheit unzureichend ist. Und schließlich wurden die wenigen bisher vorgestellten Ansätze zur mehrdimensionalen Konzeption geprüft. Diese zeichnen sich dadurch aus, dass sie rein theoretisch motiviert waren. Die Kategorisierung von unterschiedlichen Ungewissheitsqualitäten erfolgte in einem Top down-Prozess. Während solch ein Vorgehen zwar die Erzeugung einer kohärenten Struktur erleichtert, läuft man dabei dennoch Gefahr, dass das Ergebnis der Erlebniswelt der Beforschten, also der Patienten, zu wenig entspricht. Das allein ist für eine Theorie kein prinzipielles Problem, soll sie doch in erster Linie explanatorisch vereinheitlichen und empirisch prüfbare Vorhersagen ermöglichen. Ob eine Theorie für die Betrachteten augenscheinvalide ist, muss dabei kein Kriterium sein. Geht es uns hier aber gleichzeitig um die Entwicklung eines Messinstruments, das Patienten zur Selbstbeurteilung vorgelegt werden können soll, so ist es sehr wohl erforderlich, dass das Abgefragte möglichst eng an die Erfahrungen und Sprechweisen der Ausfüllenden anknüpft, um über das richtige Verständnis der Items valide Messungen zu gewährleisten.[1]

Die Nähe zur Erlebniswelt der Patienten wäre demnach schon für sich genommen Grund genug, zur Entwicklung des Messinstruments ein Bottom up-Vorgehen zu wählen. Zusätzlich erscheint uns das Wissen um das Erleben von Ungewissheiten,

[1] Bei Fragebogenverfahren lässt sich häufig auch umgekehrt argumentieren: Je weniger augenscheinvalide Items konstruiert sind, desto schwieriger ist eine zielgerichtete Verfälschung der Ergebnisse durch die Ausfüllenden, die in einigen diagnostischen Bereichen ebenso die Validität bedroht. Da die Verfälschungsmotivation im vorliegenden Fall als recht gering eingestuft wird, wird der vorgenannten Argumentationslinie gefolgt.

wie es in Kapitel 4 referiert wurde, zu rudimentär, um davon direkt ein Instrument abzuleiten. Vielmehr ist der Fragebogenentwicklungsprozess auch als theorie*bildend* anzusehen: Welche Qualitäten von Ungewissheit Patienten selbst – und nicht theoretisierende Wissenschaftler – erleben und inwieweit sie diese differenzieren, muss erst noch gezeigt werden. Daher ist es wichtig, das Datenmaterial für die Itemgenerierung mit Verfahren zu gewinnen und zu ordnen, die den Forscher zu weitgehender Ergebnisoffenheit disziplinieren.

Nachfolgend werden das Verfahren zur Datenkategorisierung (5.1.1) und das daraus folgende Vorgehen bei der Datengewinnung (5.1.2) dargestellt. Beide Arbeitsschritte stehen im Dienste der Theoriebildung. Deren Produkt – eine Konzeption über die von Patienten erlebten Ungewissheiten – wird in verschiedenen Schritten wiederum empirisch validiert und weiter präzisiert (Kapitel 6.5ff.).

5.1 Kategorisierung des Datenmaterials

5.1.1 Der Grounded Theory-Ansatz

Das Verfahren. Die Grounded Theory (GT) geht auf Glaser und Strauss (1967) zurück. Der Name beschreibt dabei das *Produkt* des Forschungsprozesses, nämlich eine Theorie, die ‚grounded', d.h. in den Daten verankert ist. Dagegen bildet die Grounded Theory als *Verfahren* vielmehr einen Leitfaden für einen Forschungsprozess von der Planung über die Datenerhebung bis zur Auswertung. Sie versteht sich somit als „empirisch fundierte Theoriebildung" (Kelle, 1994). In einem GT-Prozess wird zunächst das Material, etwa das Transkript eines Interviews zu Ursachen der Fremdenfeindlichkeit, in inhaltlich geschlossene Abschnitte (Properties) unterteilt. Anschließend wird willkürlich einer dieser Abschnitte aus dem Gesamtbestand ausgewählt und einer ihm übergeordneten Kategorie (Category) zugeordnet. Diese Kategorien werden nicht a priori sondern während der Auseinandersetzung mit dem Datenmaterial gebildet.[2] Der jeweils nächste Interviewabschnitt kann entweder in eine schon bestehende Kategorie integriert werden oder begründet eine neue. Die

[2] Dabei sind die Kategorien hinreichend spezifisch zu definieren. Anderenfalls wäre den Anforderungen der Grounded Theory *formal* genüge getan, wenn eine einzige Kategorie etwa mit der Bezeichnung ‚Aspekte zu Ursachen der Fremdenfeindlichkeit' gebildet würde. *Inhaltlich* dagegen wären solche allumfassenden Kategorien unbrauchbar.

Schlüssigkeit des Kategoriensystems wird dabei fortwährend geprüft (Constant Comparative Process). Sobald es in diesem Prozess etwa durch neues Datenmaterial zu inhaltlichen Inkonsistenzen kommt, werden alle Zuordnungen einer Revision unterzogen, indem Kategorien geteilt oder zusammengelegt und neu definiert werden. Die Kategorien sind ausreichend abgesichert (saturated), wenn sich eine größere Zahl von Daten widerspruchsfrei in die bisherigen Kategorien integrieren lässt. In einem letzten Schritt werden die Beziehungen (Relations) zwischen den Kategorien beschrieben.

Kritische Einwände. Aus diesem Vorgehen wird der Anspruch abgeleitet, das resultierende Kategoriensystem sei kein Produkt der Vorannahmen des Forschers, sondern habe objektive Gültigkeit. Denn schließlich seien die Kategorien nicht einfach ausgedacht, sondern aus den Daten ‚emergiert' (Glaser und Strauss, 1967). Um dies zu gewährleisten, ist der Forscher gehalten, „literally to ignore the literature of theory and fact on the area under study, in order to assure that the emergence of categories will not be contaminated" (ebd, S. 37). Diese Vorstellung einer von subjektiven Einflüssen unbefleckten Theorie hat in den vergangenen Jahren mancherlei Kritik hervorgerufen. Während Anselm Strauss zuletzt von diesem Anspruch abrückte und die anschließende empirische Überprüfung von Grounded Theories befürwortete (Strauss und Corbin, 1996), vertritt Barney Glaser noch immer die Auffassung, dass „the abstractions that emerge become independent of the researcher bias" (Glaser, 2002, S. 3). Er liefert sich damit einen polemisch ausufernden Streit insbesondere mit Kathy Charmaz, die für eine konstruktivistisch fundierte Neufassung der Grounded Theory eintritt (Charmaz, 2000). Zur Illustration des Disputs ein kurzer Ausschnitt:

> „The researcher composes the story; it does not simply unfold before the eyes of an objective viewer. The story reflects the viewer as well as the viewed." (Charmaz, 2000, S. 522)

> „Again, absolutely NO, the GT researcher does not ‚compose' the ‚story'. GT is not description, and the unfolding is emergent from the careful tedium of the constant comparative method and theoretical sampling – fundamental GT procedures. [...] only people who can conceptualize should do GT." (Glaser, 2002, S. 3, Hervorhebung im Original)

Fazit. Das geschilderte Vorgehen der GT eignet sich gut, um aus unstrukturierten Interviews zum Ungewissheitserleben Kategorien von Ungewissheit zu extrahieren. Gleichwohl leuchten die Einwände gegen die von Glaser postulierte Objektivität von Grounded Theories ein. Um einen möglichen Bias bei der Kategorienbildung weitgehend zu vermeiden oder zumindest aufzudecken, werden einerseits die Empfehlungen von Glaser und Strauss (Glaser und Strauss, 1967; Strauss und Corbin, 1996) zu ‚Gegenmaßnahmen' berücksichtigt (zum Theoretischen Sampling und zum Permanenten Vergleich s. 5.1.2). Andererseits lässt sich außerhalb des GT-Prozesses in Übereinstimmung mit Seale (1999) die intersubjektive Nachvollziehbarkeit der Kategorisierung belegen (5.1.4). Zusätzliche Schritte werden im weiteren Prozess der Fragebogenentwicklung und -validierung in den folgenden Kapiteln dargestellt. Dieses Vorgehen verspricht die Kombination der Vorteile einer ergebnisoffenen qualitativen Perspektive mit denen des statistisch fundierten quantitativen Ansatzes.

5.1.2 Die Bildung der Kategorien von Ungewissheit

Stichprobe. Als Ausgangsdatenmaterial wurden unstrukturierte Tiefeninterviews mit Krebspatienten geführt, aufgezeichnet und transkribiert. Die Patienten wurden über das Tumorzentrum der Universitätsklinik Schleswig-Holstein, Campus Kiel, rekrutiert. Sie wurden gefragt, ob sie zu einem Interview über ihre Krebserkrankung bereit seien, das der Entwicklung eines Fragebogens zu chronischen Erkrankungen diene. Im Falle des Einverständnisses wurde ein gesonderter Termin vereinbart. Nach dem GT-Prinzip des ‚maximalen Kontrasts' (Strauss und Corbin, 1996) wurden fünf Interviewpartner ausgewählt, die sich in Alter, Geschlecht, Tumorart und Behandlungsphase möglichst stark unterschieden. Damit sollte gewährleistet werden, dass das Datenmaterial die Extraktion *exhaustiver* Kategorien von Ungewissheit gestattet.[3]

Ablauf der Interviews. Den Teilnehmenden wurde erklärt, dass es um das Erleben von Ungewissheiten im Rahmen ihrer Krebserkrankung gehe. Daraufhin wurden sie gebeten, Situationen ihrer Erkrankung zu beschreiben, in denen sie Ungewissheiten

[3] Die Probanden gezielt auszuwählen, statt eine Zufallsstichprobe zu ziehen, wird in der GT als ‚Theoretical Sampling' bezeichnet (Strauss und Corbin, 1996).

erlebten oder erlebt hatten. Explizit wurden dabei erlebte Ungewissheiten in verschiedenen Stationen des Krankheitsverlaufs herausgearbeitet, etwa erste Symptome, Diagnosestellung, Behandlungsplanung, Behandlung, Nachsorge. Um Missverständnisse zu minimieren, wurden wiederholt Aussagen der Patienten paraphrasiert und um evtl. nötige Korrektur gebeten. Die Interviews dauerten ca. 90 bis 150 Minuten. Durch einen permanenten Supervisionsprozess konnte die Interviewtechnik fortlaufend optimiert werden. Es handelte sich also um ein ‚lernendes' Interviewsystem, bei dem die Erfahrungen früherer Interviews bezüglich der Strukturierung und der Rahmenbedingungen in die Gestaltung der weiteren bewusst eingebunden wurden.

Ablauf des Kategorisierungsprozesses. In einem ersten Arbeitsschritt wurden die Transkripte der Interviews durch vier unabhängig voneinander arbeitende Mitglieder der Forschungsgruppe in ungewissheitsbezogene Sinnabschnitte unterteilt, nachfolgend ‚Elemente' genannt. Im Falle divergierender Elementarisierung wurde ein Konsens gesucht. Uneindeutige Elemente wurden ausgeschlossen. Ein Beispiel für ein Element lautet:

> „Ja, am 20. Januar bin ich aufgewacht, und dann hatte ich hier an der Brust so einen Schmerz und dann war so eine rote Stelle. Das tat unheimlich weh, und tja...ich dachte, das wäre irgendwie Allergie, oder was weiß ich..."

Die insgesamt 400 Elemente aus den fünf Interviews wurden anschließend sukzessive in zufälliger Reihenfolge gezogen, um dafür jeweils passende (vorläufige) Kategorien von Ungewissheit zu bilden. Sobald ein Element nicht zweifelsfrei einer bestehenden oder neu zu schaffenden Kategorie zugeteilt werden konnte, wurde die gesamte bis dato entstandene Kategorienstruktur mitsamt ihren Elementen einer Revision unterzogen (*Constant Comparative Process*). Entsprechend wurde auch verfahren, wenn zwei Elemente zwar jeweils gut zu einer Kategorienüberschrift passten, untereinander aber zu unterschiedlich waren. Da das Kategoriensystem als Basis für die Entwicklung von Fragebogendimensionen dienen sollte, wurde im Dienste der Praktikabilität eine Obergrenze von acht Kategorien festgelegt. Zwischenzeitlich tolerierte feinere Unterteilungen wurden durch Fusion von Kategorien aufgelöst.

An dieser diskursiv geprägten Kategorisierungsprozedur waren vier Mitglieder der Forschungsgruppe beteiligt.

Zum Ende der Kategorisierung der 400 Elemente kam es zu einer Sättigung (Saturation), die vorher als „mindestens 50 aufeinander folgende, widerspruchsfrei integrierte Elemente“ definiert worden war.[4]

An dieser Stelle sei angemerkt, dass die inhaltliche Verknüpfung der Kategorien in Form von ‚Relations‘ – ein wesentlicher Schritt im GT-Prozess – im Rahmen dieser Arbeit aus ökonomischen Gründen zunächst noch nicht erfolgt. Für die Entwicklung eines Messinstruments sind in erster Linie die Kategorien selbst erforderlich. Gleichwohl sind die Beziehungen zwischen den einzelnen Ungewissheitsqualitäten von hohem Interesse, insbesondere dann, wenn sich die Kategorien als empirisch nützlich erweisen.

5.1.3 Die empirisch fundierten Kategorien von Ungewissheit

Aus dem geschilderten Prozedere gingen acht Kategorien oder Qualitäten von Ungewissheit hervor. Sie werden nachfolgend beschrieben und durch prototypische Ungewissheitsformulierungen illustriert. Weitere Beispiele können Appendix B entnommen werden. Die Reihenfolge ist inhaltlich nicht bedeutsam.

1. **Soziale Einbindung**
   Diese Kategorie versammelt Ungewissheiten, die die Verlässlichkeit der Beziehungen zum sozialen Umfeld angesichts der Erkrankung betreffen.

   Aus verschiedenen Gründen werden durch eine chronische Erkrankung Ungewissheiten bezüglich der sozialen Einbindung aktualisiert. Chronisch Kranke brauchen meist mehr Unterstützung als Gesunde. Das umfasst sowohl emotionale als auch ganz praktische Aspekte wie etwa die Erledigung belastender Haushaltstätigkeiten. Dies erfordert vom sozialen Umfeld des Patienten ein erhöhtes Engagement. Das wiederum kann Schuldgefühle bei den Patienten und Überforderung bei den Bezugspersonen hervorrufen. Die Frage, ob man trotz der Krankheit noch im wahrsten Wortsinne ‚tragbar‘ ist, steht über dieser Kategorie. Neben den besonders nahe stehenden Personen sind darüber hinaus

[4] Wäre innerhalb dieser 400 Elemente keine Sättigung eingetreten, hätte dies die Unzulänglichkeit des Kategoriensystems verdeutlicht und die Erhebung weiteren Datenmaterials nach sich gezogen.

auch Freunde und Bekannte oder Kollegen von chronischen Erkrankungen berührt. Eventuell muss ein früher als unternehmungslustig bekannter Patient durch seine Erkrankung häufiger Veranstaltungen meiden oder Einladungen ausschlagen. Dadurch mögen sich soziale Beziehungen verringern. Gleichwohl mag auch eine Krebspatientin, die dieser Tendenz des sozialen Rückzugs widersteht und trotz ihrer durch die Chemotherapie bedingten Kahlköpfigkeit zu einer Feier geht, über die Belastbarkeit der sozialen Einbindung nachdenken.

> „Ob meine Frau wohl auch dann noch zu mir hält, wenn ich weitere körperliche Einschränkungen erleide?“

2. **Krankheitszustand**
   Die beiden folgenden Kategorien bezeichnen Ungewissheiten, die die Einschätzung des aktuellen Krankheitszustandes betreffen.

   a) **Diagnose – Prognose**
   Chronische Erkrankungen begleiten Patienten definitionsgemäß über viele Jahre und zeichnen sich dabei durch eine häufig veränderliche Intensität aus. Das bedeutet, dass sich letztlich in beinahe jedem Lebensmoment eine Verschlechterung des Krankheitszustandes ankündigen könnte, auf die womöglich schnell reagiert werden muss. So steigen die Heilungschancen von Krebspatienten deutlich, je früher ihr Rezidiv erkannt und behandelt wird. Doch auch wenn bei anderen Krankheiten keine sichere Behandlungsoption in einer akuten Phase besteht – etwa bei Multipler Sklerose (Heesen et al., 2004) – müssen womöglich trotzdem Maßnahmen zum angepassten Umgang ergriffen werden. Chronisch Kranke befinden sich demnach in einem permanenten Selbstdiagnoseprozess, in dem sie auf kleinste Anzeichen ihrer Erkrankung achten, diese von ‚normalen‘ Empfindungen abgrenzen, interpretieren und daraus Konsequenzen ableiten müssen. Dieser Prozess setzt sich auch während einer Behandlung fort, da präzise Rückmeldungen von Patienten meist unentbehrlich für eine optimale Anpassung der Behandlungsstrategie sind. In dieser Kategorie zusammengefasste Ungewissheiten betreffen demnach den Versuch von Patienten, sich ein Bild vom aktuellen Krankheitszustand zu machen. Dies schließt auch die Schwierigkeit ein, unter Einbezug medizinischer Informationen die daraus erwachsende weitere Prognose einzuschätzen.

> „Ob meine Müdigkeit noch normal ist, oder hat das was zu bedeuten?“

b) **Informationsentschlüsselung**
Auch bei den in dieser Kategorie versammelten Ungewissheiten geht es darum, sich einen validen Eindruck vom aktuellen und zukünftigen Krankheitszustand zu machen. Allerdings liegt der Fokus hier ausschließlich auf solchen Ungewissheiten, die bei der Interpretation von ärztlichem Handeln und Sprechen bewusst werden. Die Interviewpartner berichteten von diversen Situationen, wo scheinbar belanglose Ereignisse wie ein Telefonanruf während der Visite oder tuschelnde Krankenschwestern als Beleg für die schlechte Prognose interpretiert wurden. Zuweilen deuteten Patienten es als schlechtes Zeichen, wenn ihre Ärztin sehr kurz angebunden war – vermeintlich wegen ungünstiger Entwicklungen. Oder aber es schien ihnen verdächtig, dass sich ihre Ärztin – vermeintlich aus Sorge – besonders viel Zeit für ein Gespräch nahm. Weiterhin werden in diese Kategorie Ungewissheiten einsortiert, die das Verständnis medizinischer Aussagen betreffen. Fragen wie „Was bedeutet ‚gute Heilungschancen‘?“ gehören ebenso dazu wie solche nach der Bedeutung allgemeiner statistischer Aussagen für die persönliche Situation. Diese Kategorie von Ungewissheit scheint auch deshalb bedeutsam, weil das medizinische Personal neben den eigenen Körperempfindungen des Patienten (Kategorie 2a) durch Rückgriff auf Blutwerte, Röntgenbilder etc. einen weiteren Zugang zu dessen Zustand hat, der diesem in der Regel nur durch das Nadelöhr der Kommunikation mit den Ärzten offen steht.

> „Ist es wirklich Routine, dass ich noch zu einem Facharzt überwiesen werde?“

3. **Bewältigung von Anforderungen**
Hier werden Ungewissheiten subsumiert, die die eigene Anpassungsfähigkeit (Adaptation) betreffen.

Wie unter Kategorie 1 schon angeklungen, haben chronisch Kranke häufig auch in relativ ruhigen Krankheitsphasen nicht die volle Leistungskraft. Dennoch müssen und wollen sie weiterhin am sozialen und beruflichen Leben teilnehmen. Manche Interviewpartner wollten sich durch einen vollen Wiedereinstieg

in den bisherigen Beruf auch selbst beweisen, dass sie nicht von der Krankheit dominiert sind. Insofern stehen die Patienten vor einem fortwährenden Abwägungsprozess zwischen Fordern und Überfordern. Teilweise tragen auch ökonomische Zwänge dazu bei, dass die Anforderungen nicht im gewünschten Ausmaß reduziert werden können. Und nicht zuletzt kann auch die Behandlung selbst eine zu bewältigende Aufgabe sein. Beschrieben wurden etwa die regelmäßigen Fahrten zur Chemotherapie, die Belastungen durch deren Nebenwirkungen oder die Trennung von der Familie bei stationärer Therapie, womöglich gar in infektionsverhindernder Isolation.

> „Werde ich wohl nach meiner Behandlung wieder in meinen Beruf zurückkehren können?“

4. **Kausalattribution**
   Hierunter werden Ungewissheiten verstanden, die sich bei dem Versuch auftun, das eigene Krankheitsgeschehen in der Vergangenheit, Gegenwart oder Zukunft sinnstiftend kognitiv zu integrieren.

   Durch eine ernsthafte Erkrankung kommen häufig Gedanken auf, was diese verursacht haben könnte. Ein Interviewteilnehmer mit Bronchialkarzinom dachte über seinen früheren Nikotinkonsum nach, ein anderer brachte beruflichen Stress als Auslöser ins Spiel. Neben diesen physischen oder auch psychischen Faktoren berichteten Patienten auch davon, dass sie die Erkrankung als einen Wink des Schicksals ansehen. Daraus erwuchs bei ihnen die Ungewissheit, worin denn die ‚Lehre‘ aus der Erkrankung bestehen könne. Die hier subsumierten Ungewissheiten bekommen damit auch eine zukunftsbezogene Komponente, wenn sie zu der Frage führen, wie nach der persönlichen Kausalattribution die Zukunft zu gestalten ist. Weiterhin gehören auch Versuche religiöser Erklärungen in diese Kategorie, die die Krankheit als eine Strafe, Prüfung oder zumindest Fügung Gottes deuten.

   > „Warum bin wohl gerade ich krank geworden?“

5. **Umgang mit dem Arzt**
   Die beiden folgenden Kategorien beinhalten Ungewissheiten im Umgang mit dem Behandlungspersonal.

   a) **Rollenpräferenz**
      Einer der Kernpunkte medizinischer Entscheidungsfindung ist die Patien-

tenbeteiligung. So verwundert es nicht, dass vielfach von Ungewissheiten berichtet wurde, die bei der Bestimmung der eigenen Rolle in der Behandlung aktualisiert werden. Darf oder muss man als Patientin eine passive Krankenrolle einnehmen und der Ärztin die Entscheidungen überlassen? Könnte oder müsste man sich bei der Behandlungsauswahl beteiligen – sei es, um die eigene Meinung berücksichtigt zu wissen oder um die Qualität der Behandlung zu überwachen? Letzterer Aspekt scheint insbesondere bei komplexeren Krankheiten in stationärer Behandlung ein wichtiger Beweggrund für aufmerksame Beteiligung zu sein, da Patienten manchmal Befürchtungen hatten, durch häufigen Wechsel des medizinischen Personals könnten wichtige Aspekte übersehen werden.

„Sollte ich mich stärker in die Behandlungsplanung einmischen?“

b) **Qualität des Arztes**
Damit eng verknüpft ist die Frage, ob man dem Arzt insgesamt vertrauen kann. Eine verlässliche Beziehung zum Arzt wurde von allen Interviewten als sehr wichtig angesehen, so dass die Einschätzung des Arztes die Patienten häufig beschäftigte. Das umfasst einerseits dessen fachliche Kompetenz, von der die Güte der Behandlung und damit auch die Heilungschancen abhängen. Andererseits erlebten die befragten Patienten Ungewissheiten, ob sie auch menschlich bei ihrem Arzt gut aufgehoben sind. In diesem Zusammenhang wurden auch Ungewissheiten berichtet, ob eine Ärztin schlechte Neuigkeiten womöglich verheimlichen würde. Dieser Aspekt grenzt an die unter Kategorie 2b versammelten Ungewissheiten an.

„Kann ich meinem Arzt vertrauen?“

6. **Behandlung**
Diese Kategorie umfasst Ungewissheiten darüber, welche medizinischen und ‚privaten‘ Handlungen die Krankheit in gewünschter Weise beeinflussen.

Ein bestimmendes Thema der Interviews war die Ungewissheit, welche Behandlung zu wählen ist. Womöglich stehen unterschiedliche Optionen mit unterschiedlichen Vor- und Nachteilen zur Wahl, zu denen unterschiedliche Ärzte unter Umständen auch noch unterschiedliche Meinungen haben. Damit wird nachvollziehbar, dass Patienten Ungewissheit darüber erleben, wie sie sich

entscheiden sollen. Mit der Wahl einer bestimmten Option hört diese Ungewissheit jedoch nicht immer auf. Auch während der Behandlung fragen sich Patienten häufig, ob diese bei ihnen wirken wird und ob es lohnt, etwaige Nebenwirkungen weiterhin in Kauf zu nehmen. Und selbst nach einer erfolgreich erscheinenden Behandlung mag eine Ungewissheit verbleiben, ob mit einer anderen Therapie nicht das Rückfallrisiko geringer gewesen wäre. Neben den ‚offiziellen' Behandlungen sind aber auch die ‚privaten' Maßnahmen Gegenstand von Überlegungen. So fragten sich Patientinnen, ob ihnen die regelmäßige selbstständige Kontrolle auf Knoten in der Brust insgesamt gut tut. Oder es ging um die Frage, wieviel private Aktivität hilfreich für die Heilung sein mag.

„Was wohl in Fällen wie meinem die optimale Behandlung wäre?"

Als Ergebnis des GT-Prozesses ist festzuhalten, dass man beim Ungewissheitserleben Krebskranker sinnvoll bis zu acht verschiedene Qualitäten unterscheiden kann. Die Kategorienstruktur erscheint als eine exhaustive Beschreibung des Ungewissheitserlebens. Dieses wird als inter- und intraindividuell variabel angesehen, d.h. Patienten desselben Alters mit derselben Diagnose erleben unter Umständen unterschiedliche (Ausprägungen von) Ungewissheitsqualitäten, und das Ungewissheitserleben kann sich im Laufe der Erkrankung verändern. Darüber hinaus wird angenommen, dass die Kategorien ausreichend disjunkt sind, so dass Patienten zwischen den einzelnen Kategorien differenzieren können. Sie sind dabei aber nicht voneinander unabhängig. So mag etwa eine Ungewissheit in der Beziehung zur Ärztin (Kategorie 5b) einer Ungewissheit über eine mit dieser Ärztin ausgewählte Therapie (Kategorie 6) Vorschub leisten etc.

Die Bezeichnungen mit a und b bei vier Kategorien weisen darauf hin, dass es sich zwar um eigenständige Kategorien handelt, dass diese aber einen gemeinsamen Hintergrund haben, nämlich Ungewissheiten bezüglich des Krankheitszustandes bzw. des Umgangs mit dem Arzt. Da sich auch diese verwandten Kategorien theoretisch gut abgrenzen lassen sowie eindeutige Kategorisierungen der Elemente zuließen, bleibt bis auf weiteres die Unterscheidung von acht Kategorien bestehen. Diese Frage wird im folgenden sowie im Abschnitt 6.8 erneut aufgegriffen.

5.1.4 Überprüfung der ‚Reliabilität' des Kategoriensystems

Wie in Abschnitt 5.1.1 begründet, erscheint es erforderlich, die ‚Reliabilität' der Kategorisierungsprozedur zu belegen. Während der Begriff Reliabilität im Sinne von ‚Reproduzierbarkeit' in der quantitativen Forschung verschiedene Operationalisierungen aufweist – etwa Retest- oder Paralleltestreliabilität – fehlt bei qualitativem Vorgehen eine direkte Entsprechung (Steinke, 1999). Seale (1999) schlägt daher vor, das Konzept der *Interraterreliabilität* auf die Kategorisierung qualitativen Datenmaterials anzuwenden, um die intersubjektive Nachvollziehbarkeit zu belegen. Kriz (in: Lisch und Kriz, 1978, S. 100f.) weist allerdings darauf hin, dass damit lediglich die interindividuelle Reproduzierbarkeit des Kodierverhaltens in einer kleinen Stichprobe von Beurteilern belegt werden kann. Dies wäre zumindest eine notwendige Bedingung für die Überzeugungskraft des Kategoriensystems. Hinreichend wäre es dagegen in diesem Fall, wenn neues Interviewmaterial in einer unabhängigen Auswertung – idealerweise von unabhängigen Forschern – zur Bildung desselben Kategoriensystems führte. Wegen des großen zeitlichen Aufwandes des GT-Prozesses muss dieser Nachweis hier aus ökonomischen Gründen ausbleiben. In Anbetracht der unter 5.1.1 begründeten Zweifel an der ‚Objektivität' von Grounded Theories erscheint diese Forderung zudem kaum erfüllbar. Daher wird hier nur die notwendige Reproduzierbarkeit der Kodierungen überprüft, während sich die inhaltliche Stimmigkeit der Kategorien im späteren Einsatz des daraus abgeleiteten Fragebogens zu erweisen hat.

Erhebung des Datenmaterials. Zu diesem Zweck wurde gemäß dem unter 5.1.2 beschriebenen Verfahren ein sechstes Interview geführt und elementarisiert. Dabei wurden bewusst Äußerungen vermieden, die einen Hinweis auf eine mögliche Kategorienzuordnung erlaubt hätten.

Ablauf der Kategorisierung für die Interraterreliabilitätsprüfung. Die so entstandenen 145 Elemente wurden wiederum von denselben vier Forschern zuzüglich eines neuen Mitglieds der Forschungsgruppe unabhängig voneinander den acht Ungewissheitsqualitäten zugeordnet. Auslassungen von Elementen oder die Zuordnung zu mehr als einer Kategorie waren unzulässig. Als Hilfsmittel wurde ein Leitfaden mit Beispielen zu den einzelnen Kategorien bereitgestellt (s. Appendix B).

Bestimmung der Interraterreliabilität. Bei dem hier vorliegenden kategorialen Datenmaterial scheiden alle korrelativen Übereinstimmungsmaße aus, da sie Intervallskalenniveau voraussetzen. Stattdessen eignen sich nominal basierte Kennwerte wie Cohens κ (Cohen, 1960) sowie dessen Erweiterung für mehr als zwei Rater von Fleiss (1971), nachfolgend Fleiss' κ genannt. Beide basieren auf der prozentualen Übereinstimmung zwischen den Ratern und berücksichtigen dabei den Anteil zufällig zu erwartender Urteilsgleichheit. Damit sind sie den Koeffizienten z.B. von Weymann oder Ritsert überlegen. Zudem sind sie weniger anfällig gegenüber ungleichen Randsummenverteilungen als Scotts π (Wirtz und Caspar, 2002). Daher werden in dieser Studie die Koeffizienten von Cohen und Fleiss verwendet. Gleichwohl verbleiben auch bei allen erwähnten Kennwerten einige logische Unzulänglichkeiten wie z.B. die Nichtberücksichtigung von Übereinstimmung im Ausschluss von Kategorien (vgl. ausführlicher Kriz in: Lisch und Kriz, 1978, S. 98). Die Berechnungsformeln sind in Appendix A.1 bzw. A.2 nachzulesen.

Ergebnisse. Werden acht Kategorien von Ungewissheit unterschieden, so liegt Fleiss' κ bei 0,54. Cohens κ rangiert für die $\binom{5}{2} = 10$ verschiedenen Raterpaarungen zwischen 0,44 und 0,69 bei einem Median von 0,55.

Fasst man jeweils die Subkategorien a und b von Kategorie 2 und Kategorie 5 zusammen, so dass 6 Kategorien entstehen, ergibt sich ein Fleiss' κ von 0,61. Die Werte für Cohens κ liegen zwischen 0,56 und 0,66 bei einem Median von 0,61.

Diskussion. Mit Wirtz und Caspar (2002) sind die berichteten Reliabilitätskennwerte als „akzeptabel" bis „gut" zu bezeichnen. Zur Orientierung sei angemerkt, dass die Maße ähnlich wie ein Korrelationskoeffizient theoretisch Werte zwischen -1 und 1 annehmen könnten.

Die Unterschiede zwischen den Kennwerten der sechs- und der achtkategoriellen Zuordnung scheinen statistisch eher im Zufallsbereich zu liegen und nicht für eine Ununterscheidbarkeit der jeweiligen Subkategorien zu sprechen. Insbesondere sind sie kein Artefakt der Kategorienanzahl an sich, da es sich bei beiden Koeffizienten um basisratenkorrigierte Maße handelt. Inhaltlich spricht dagegen einiges dafür, die Kategorien 2a und 2b resp. 5a und 5b nicht zusammenzulegen. Trotz einiger Entsprechungen auf abstrakter Ebene sollte aus den Kategorienbeschreibungen in Abschnitt 5.1.3 deutlich geworden sein, dass es für Patienten z.B. einen Unterschied

macht, ob eine Ungewissheit aus der Unvorhersagbarkeit des individuellen Erkrankungsverlaufs (Kategorie 2a) oder aus den uneindeutigen Worten des Behandlers (Kategorie 2b) resultiert. So zieht letzteres womöglich eine erhöhte Ungewissheit in Bezug auf die Person und Kompetenz des Arztes (Kategorie 5b) nach sich, ersteres dagegen mag die Beziehung zum Arzt sogar eher festigen, wenn sich Arzt und Patient durch die gemeinsame Auseinandersetzung mit der unberechenbaren Erkrankung solidarisieren. Aus diesen Gründen wird die Variante mit acht Kategorien präferiert. Weitere empirische Argumente für diese Entscheidung liefert zudem Validierungsstudie 4 in Abschnitt 6.8.

Für die intersubjektive Nachvollziehbarkeit der Kategorienzuordnungen spricht auch die Tatsache, dass die paarweisen Cohens κ-Werte der neu in die Forschungsgruppe gekommenen Raterin sogar höher ausgefallen sind als die zwischen denjenigen vier Forschern, die über lange Zeit gemeinsam das Kategoriensystem entwickelt hatten (Median 0,58 vs. 0,55 bei sechs und 0,63 vs. 0,60 bei acht Kategorien). Offensichtlich basiert die Kategorienzuordnung auf vermittelbaren, inhaltlichen Gesichtspunkten und nicht auf eingeübten, undurchsichtigen Heuristiken.

Schließlich ist festzuhalten, dass das Datenmaterial zuweilen mehrere plausible Inhaltsinterpretationen und mithin Kategorienzuschreibungen zuließ. Die in Abschnitt 5.1.2 erwähnte Vorauswahl von Elementen bezog sich schließlich nur auf die Frage, ob ein Element *überhaupt* Ungewissheit thematisiert, nicht jedoch, ob eventuell mehrere verschiedene Ungewissheitsqualitäten anklingen. Zur Illustration seien zwei der kategorisierten Elemente aufgeführt:

> „Ich bin gleich zum Arzt gegangen. Ich hab gleich gesagt: ‚Jetzt macht weiter! Jetzt will ich das wissen, was los ist‘“.

Dass der Patient hier über Ungewissheit spricht, ist leicht nachvollziehbar. Versteht man die Aussage als Ausdruck einer Ungewissheit über den aktuellen Krankheitszustand, so wäre sie in Kategorie 2a einzuordnen. Hört man dagegen eher die Entschlossenheit des Patienten heraus, die in seinen Augen nötigen diagnostischen Maßnahmen einzufordern, so ist dies eine Aussage, die die Frage der eigenen Rolle in der Arzt-Patient-Dyade betrifft, wenngleich hier als Gewissheit formuliert. Demnach könnte sie folgerichtig auch Kategorie 5a zugewiesen werden.

> „Das war eigentlich – und das ist das, was ich vorher sagte – als wenn irgendeine Macht mir gesagt hätte: ‚Du musst jetzt hin und Schluss. Das

muss raus‘.“

Auch dieses Element ist mehrdeutig: Bei Betonung auf „irgendeine Macht“ könnte man es als Antwort auf die Frage nach der Bedeutung dieses Schicksalswinks verstehen (Kategorie 4). Fokussiert man hingegen eher auf den Entschluss „Das muss raus“, so ist dies Ausdruck einer in jenem Moment geklärten Ungewissheit in Bezug auf die Behandlungswahl (Kategorie 6) – in diesem Fall die riskante operative Entfernung des Tumors statt einer Fortführung der Chemotherapie.

Die anhand dieser Beispiele erkennbare Mehrdeutigkeit, die sich beim Rating dieser beiden Beispiele entsprechend in divergierenden Urteilen zeigte, liegt dabei nicht in der Konzeption der Kategorien, denn die zu den Beispielen vorgeschlagenen Interpretationen sind eindeutig einer Kategorie zuzuordnen. Vielmehr liegen die Ambiguitäten im Datenmaterial selbst. Durch Nachfragen während des Interviews hätte dahingehend Eindeutigkeit geschaffen werden können. Solch häufig unzureichender Kontextinformation eingedenk scheinen die berichteten Reliabilitätskennwerte die prinzipielle intersubjektive Nachvollziehbarkeit eher zu unterschätzen.

Zur Unterschätzung der tatsächlichen Übereinstimmung trägt auch die auf Seite 69 erwähnte Blindheit der Koeffizienten gegenüber übereinstimmend ausgeschlossenen Kodierungen bei: In beiden Beispielen waren sich alle Rater einig, welche sechs Kategorien *nicht* in Frage kommen. Diese Einigkeit wird jedoch durch die Koeffizienten nur unzureichend berücksichtigt.

5.2 Prinzipien der Itemgenerierung für den QUICC

Die Erfindung eines geeigneten Itemformats basierte gemäß den Ausführungen in Kapitel 4 auf folgenden Anforderungen:

Ein Item sollte

- ...eine bestimmte Situation oder Situationsklasse seiner Kategorie ansprechen, etwa die Ungewissheit der Vertrauenswürdigkeit des Arztes, Kategorie 5b zugehörig.

- ...es den Patienten ermöglichen, pro Item zu äußern, in welchem *Ausmaß* sie Ungewissheit in diesem Punkt erleben. Als Antwortformat wurden daher fünfstufige, unipolare Likert-Skalen gewählt. Deren Abstufungen sind verbal gekennzeichnet, um für alle Teilnehmer eine möglichst ähnliche Repräsentation der jeweiligen Antwortkategorie zu erzeugen (Rost, 2004).

- ...deutlich machen, dass es nicht um eine Wahrscheinlichkeitsaussage geht – etwa wie sicher man sich sein könnte, dem Arzt zu vertrauen. Vielmehr interessiert die *Beschäftigung* der Patienten mit der Frage der Vertrauenswürdigkeit. Um diesen Charakter einer gedanklichen Auseinandersetzung mit der jeweiligen Ungewissheit zu betonen, wurden bewusst grammatikalisch unvollständige Fragen gebildet und etwas vage anmutende Itemformulierungen gewählt. Außerdem wurde diese Unterscheidung zwischen Wahrscheinlichkeitseinschätzung und Ungewissheitserlebnis explizit in der Einleitung aufgegriffen und anhand eines Beispiels erläutert.

- ...prinzipiell universell einsetzbar sein für möglichst alle chronischen Erkrankungen. Daher wurden trotz der onkologischen Provenienz des Kategoriensystems krebsassoziierte Formulierungen vermieden. Dennoch wird in der späteren Anwendung zu prüfen sein, ob ein Item bei Betroffenen verschiedener Erkrankungen vergleichbare Antwortreaktionen erzeugt.

- ...prinzipiell universell einsetzbar sein in allen Phasen einer Erkrankung, also in der frühen Diagnostik, während der Behandlung sowie auch lange nach Abschluss der Behandlung. Gerade diese Anforderung an die Items machte zuweilen umständlichere Formulierungen nötig, wie das untere Beispiel in Abbildung 5.1 zeigt: Das Item thematisiert die nach Auskunft vieler Patienten ungewissheitsbefrachtete Situation der Behandlungsauswahl. Für Patienten, die sich in eben jener Situation befinden, wäre eine Formulierung wie „Welche Behandlung ich wohl wählen soll?“ treffend und verständlich. Bei Patienten in der Tumornachsorge hingegen riefe diese Formulierung Unverständnis und vielleicht sogar Verunsicherung hervor, ob womöglich eine weitere Behandlung nötig ist. Stattdessen wurde den Patienten im Fragebogen zugemutet, die Entscheidungssituation(en) während ihrer aktuellen oder zurückliegenden Behandlung quasi ‚aus der Distanz‘ zu beurteilen. Gewiss besteht darin ein Unterschied zu obiger Variante, jedoch wird in beiden Fällen die Beschäftigung mit Ungewissheit bezüglich der Behandlungswahl adressiert.

Nach diesen Kriterien entstand ein Fragebogen, dessen Qualität in verschiedenen Phasen geprüft und optimiert werden soll. So folgen zunächst eine Prüfung der Verständlichkeit (Readability) sowie der Vollständigkeit des Itempools. Daran schließen sich umfangreiche empirische Validierungen an unterschiedlichen Stichproben an.

Ob ich dem Arzt vertrauen kann?

	gar nicht	etwas	mittel	ziemlich	sehr
Diese Ungewissheit beschäftigt mich	☐	☐	☐	☐	☐

Wie sicher man sich bei der Entscheidung über eine medizinische Maßnahme sein sollte?

	gar nicht	etwas	mittel	ziemlich	sehr
Diese Ungewissheit beschäftigt mich	☐	☐	☐	☐	☐

Abbildung 5.1: Zwei Beispiele aus der ersten Version des Itemformats im Entwicklungsprozess des QUICC.

5.3 Vorstudie mit dem QUICC

Vor einem ersten Einsatz des QUICC in einer Stichprobe sollten dessen Lesbarkeit, Verständlichkeit und Ökonomie sichergestellt sein. Dazu bietet sich nach Prüfer und Rexroth (1996) insbesondere die Think Aloud-Methode (Van Someren et al., 1994) an. Aus den laut geäußerten Gedanken von Probanden beim Bearbeiten des Fragebogens kann geschlossen werden, ob die Instruktionen und Items im Sinne der Autoren verstanden werden. Durch die Form des semistrukturierten Interviews kann dabei in Zweifelsfällen direkt nachgefragt werden.

Den Probanden wird der QUICC mit 76 Items vorgelegt. Die hohe Zahl ermöglicht eine spätere großzügige Eliminierung von Items aufgrund der Ergebnisse dieses Validierungsschritts. Der Fragebogen ist in Anhang C zu finden.

5.3.1 Stichprobe

Zwei der für die Kategorienbildung interviewten Krebspatienten wurden für die Think Aloud-Prozedur gewonnen.

5.3.2 Ergebnisse und Diskussion

Der QUICC wurde insgesamt als vollständig (exhaustiv) und verständlich bewertet. Gleichzeitig wurde er auch als „bedeutsam“ eingeschätzt, d.h. die Patienten fühlten sich verstanden und zum Nachdenken angeregt. Dennoch gab es mancherlei kritische Anmerkungen, die zu folgenden Konsequenzen im Fragebogendesign geführt haben:

- Die im Antwortformat enthaltene Formulierung „Diese Ungewissheit beschäftigt mich“ (s. Abbildung 5.1) führte zu Irritationen, da das Wort Ungewissheit einerseits als zu elaboriert beurteilt wurde, und andererseits manche Themen als völlig gewiss eingeschätzt würden. Daher sei es zuweilen unpassend. „Ungewissheit“ wurde deshalb durch den neutraleren und gebräuchlicheren Ausdruck „Frage“ ersetzt.

- Durch die geäußerten Gedanken wurde ersichtlich, dass die Patienten bei manchen Items das aktuelle Ungewissheitserleben einschätzten, bei anderen wiederum an eine frühere Situation im Krankheitsverlauf zurückdachten. Um Klarheit zu schaffen, wurde in jede Antwortphrase der Ausdruck „zur Zeit“ eingefügt.

- Manchmal schienen die Patienten vom intendierten Verständnis der Items abzuweichen. Da die Operationalisierung von erlebter Ungewissheit als ‚Beschäftigung mit einer Frage‘ eine recht abstrakte Itemkonstruktion nach sich zieht, wurde eine ausführlichere Einleitung für den QUICC verfasst, die das gewünschte Verständnis der Items sicherstellen soll. So führt die Einleitung zunächst kurz in das Thema ein (siehe Abbildung 5.2), um anhand eines Beispiels mit dem Itemformat vertraut zu machen. Um sicherzustellen, dass die Ausfüllenden nicht die empfundene Wichtigkeit des jeweiligen Itemstamms einschätzen, sondern das Ausmaß der gedanklichen Beschäftigung mit dieser Frage, wurde eben dieser Unterschied explizit formuliert. Außerdem wurden Veränderungen am Drucksatz vorgenommen, um die Phrase „Diese Frage beschäftigt mich“ stärker zu betonen.

- Ferner trugen die Anmerkungen der Interviewteilnehmer und einiger zusätzlich informell befragter Fachkollegen dazu bei, die avisierte Reduktion des Itempools vorzunehmen. Als für die erste quantitative Validierungsphase angemessener Kompromiss zwischen Ökonomie, Messgenauigkeit und Testung

Liebe Patientin, lieber Patient,

Betroffene einer Krebserkrankung haben mit unterschiedlichen Ungewissheiten zu tun. Fragen im Zusammenhang mit der Erkrankung lassen sich manchmal nicht beantworten. So wissen Sie womöglich nicht genau, wie Ihre Erkrankung weiter verlaufen wird. Oder Sie fragen sich, warum gerade Sie erkrankt sind.
Von solchen Ungewissheiten handeln die folgenden Fragen.

Bevor Sie beginnen, machen Sie sich bitte anhand dieses **Beispiels** mit der Frageform vertraut:

Ob ich meinem Arzt wohl vertrauen kann?

	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐

Hier geht es **nicht** darum, wie sehr Sie Ihrem Arzt vertrauen, oder wie wichtig Ihnen dieses Vertrauen ist.

Uns interessiert, **ob Sie mit der Suche nach der Antwort auf diese Frage beschäftigt sind.**

Vielen Dank für Ihre Mühe!

Abbildung 5.2: Die Instruktion zur Unterstützung des intendierten Verständnisses der Items.

möglichst vieler Items erschien eine Kürzung auf sechs bis sieben Items für jede der acht Kategorien.

Die Anordnung der Items erfolgte nicht zufällig, sondern in systematischem

Wechsel der Kategorien. Dabei wurde deren Reihenfolge in der Mitte des Bogens invertiert, um etwaige Positionseffekte gleichmäßig auf die Kategorien zu verteilen.

Die aus der Think Aloud-Prozedur hervorgegangene QUICC-Version ist in Anhang D zu sehen.

Kapitel 6

Die Validierung des QUICC

In Teil I wurde herausgearbeitet, warum es sinnvoll sein könnte, verschiedene Qualitäten entscheidungsbezogener Ungewissheit zu differenzieren. Auf diesem Ansatz fußend war die Entwicklung und Überprüfung eines Systems aus acht Kategorien von Ungewissheit (Kapitel 5) erfolgt, auf Basis dessen unter Rückgriff auf weitere theoretische Überlegungen (vgl. 5.2 bzw. 4) ein erster Entwurf des QUICC zusammengestellt und nach einer Vorstudie modifiziert wurde.

Die folgenden Abschnitte 6.5 bis 6.8 befassen sich mit unterschiedlichen Aspekten der Fragebogenvalidierung. Um die Motivation hinter diesen einzelnen Schritten zu erläutern, erfolgt an dieser Stelle zunächst ein Überblick allgemeiner Prinzipien. Da sich manche Gütekriterien von Fragebögen teilweise nur auf Kosten der übrigen optimieren lassen, schließt sich jeweils die Argumentation für die in dieser Arbeit gewählten Kompromisse und Validierungsstrategien an.

Abbildung 6.1 zeigt schematisch einen typischen Fragebogenkonstruktionsprozess nach Meyer (2004). Die drei Felder der ersten Reihe – Konstruktdefinition, Hypothesengenerierung und Testzusammenstellung – wurden in den vorangegangenen Kapiteln bereits weitgehend behandelt. Die anstehende quantitative Validierung des QUICC muss demnach neben der Datenerhebung die Schritte Reliabilitätsbestimmung, Itemanalyse und Validitätsprüfung beinhalten. In der Regel ist diese Abfolge mehrfach zu absolvieren, und zwar nicht nur bei suboptimaler Reliabilität, wie es die Abbildung suggeriert. Auch mangelnde Validität kann ein Grund für Modifikationen und nachfolgend erneute Datenerhebung sein. Die Testeichung ist für den QUICC in dieser Arbeit nicht vorgesehen.[1]

[1] Dies hat einerseits ökonomische Gründe. Andererseits ist es fraglich, ob für eine zeitlich meist

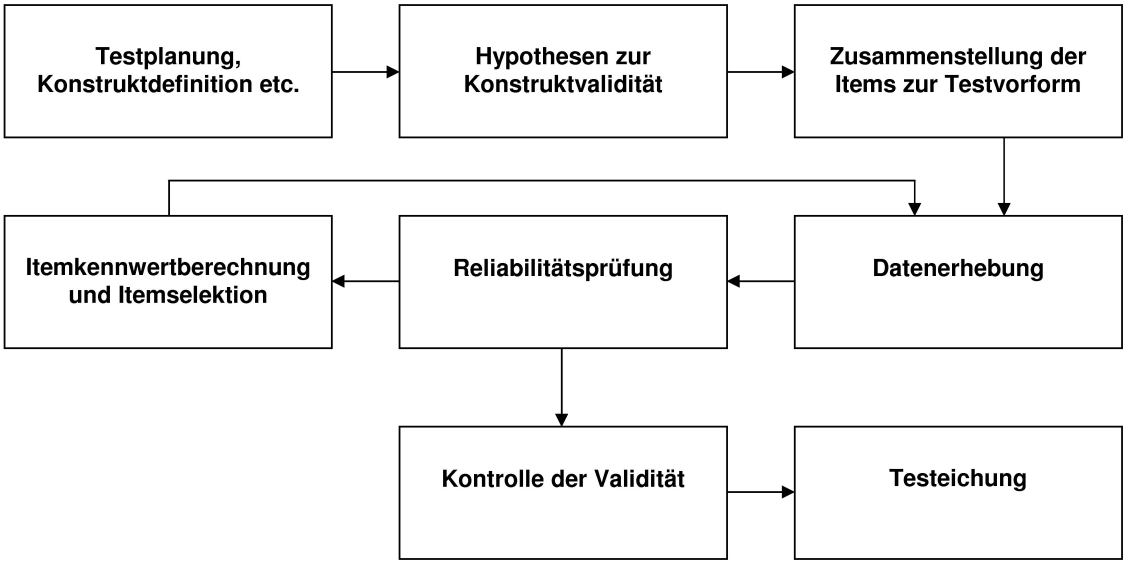

Abbildung 6.1: Der übliche Ablauf der Testkonstruktion nach Meyer (2004). Die Schritte in der ersten Reihe wurden beim QUICC bereits weitestgehend absolviert, die der zweiten und dritten Reihe erfolgen größtenteils in diesem Kapitel.

6.1 Bestimmung der Reliabilität

Reliabilität oder Messgenauigkeit eines Tests beinhaltet in der klassischen Testtheorie die Reproduzierbarkeit der Messwerte unter vergleichbaren Bedingungen. Wenn trotz gleicher Kontextbedingungen – was immer man darunter im Einzelfall verstehen mag – Unterschiede in den Messungen auftreten, deutet dies auf eine Ungenauigkeit des Messinstruments hin. Das berührt indes noch nicht die Frage, *was* gemessen wird, sondern bezieht sich nur darauf, *wie präzise* es gemessen wird. Berechnet wird die Reliabilität allgemein über eine Korrelation von Testwerten mit einem geeigneten parallelen Maß. Gebräuchlich sind dabei Retests, Paralleltests, die Halbtest-Methode sowie die Methode der internen Konsistenz. Alle vier Methoden haben ihre Vor- und Nachteile (vgl. ausführlicher z.B. Rost, 2004; Meyer, 2004).

Als brauchbarer Kompromiss aus ökonomischen und theoretischen Überlegungen hat sich die Methode der internen Konsistenz anhand von Cronbachs α etabliert, die auch in dieser Arbeit verwendet werden soll. In diesem Kennwert wird ein ge-

recht instabile, d.h. veränderliche und von vielen Kontextfaktoren abhängige Variable wie das Ungewissheitserleben die Bildung relevanter Vergleichsgruppen überhaupt möglich ist.

wichtetes Mittel der Korrelationen zwischen den Items einer Skala gebildet (vgl. A.3), um einen Hinweis zu bekommen, inwieweit die Messungen dieser Items übereinstimmen.

6.2 Itemanalyse und Itemselektion

Während sich der Begriff der Reliabilität im obigen Sinne auf die Messgenauigkeit und damit die Brauchbarkeit einer Skala als Ganzes bezieht, ist es ferner sinnvoll, eine Betrachtung der Qualität einzelner Items vorzunehmen. Daraus ergibt sich die Möglichkeit, ungeeignete Items zu eliminieren, um den Test im Hinblick auf ein bestimmtes Kriterium – etwa seine Reliabilität – zu optimieren.

6.2.1 Itemtrennschärfe

Die Trennschärfe eines Items gibt an, wie es mit dem Summenscore der zugehörigen Skala zusammenhängt. Ein Item mit hoher Trennschärfe ist damit ein guter Repräsentant seiner Skala. Zusätzlich differenzieren Items mit hoher Trennschärfe gut zwischen Personen mit hoher und jenen mit geringer Merkmalsausprägung. Daher werden allgemein hohe Trennschärfen angestrebt.

Die Trennschärfe wird in dieser Arbeit wie üblich operationalisiert als die Korrelation eines Items mit dem Summenscore der zugehörigen Skala, der ohne den Einfluss des betreffenden Items berechnet wird (sog. ‚Part-Whole-Korrektur', Bühner, 2005).

6.2.2 Itemschwierigkeit

Die Itemschwierigkeit gibt Aufschluss, wie viele Personen einer Population ein Item lösen bzw. bejahen. Schwierige Items differenzieren gut zwischen Personen mit hoher Merkmalsausprägung und schlecht zwischen jenen mit geringer Ausprägung. Bei leichten Items ist es entsprechend umgekehrt. Damit entstehen unerwünschte Decken- oder Bodeneffekte. Den größten Informationsgehalt haben Items mittlerer Schwierigkeit. Außerdem begünstigen diese hohe Trennschärfen, da zwischen Schwierigkeit und Trennschärfe eine umgekehrt U-förmige Beziehung besteht. Gleichzeitig

darf die Streuung der Items nicht zu gering sein, da sich dies wiederum negativ auf die Trennschärfe auswirkt (Bühner, 2005).

Während die Itemschwierigkeit bei dichotomen Items einheitlich als Prozentsatz derjenigen operationalisiert wird, die ein Item lösen bzw. bejahen, sind bei mehrstufigen Ratingskalen unterschiedliche Indices gebräuchlich. In dieser Arbeit wird wegen des darin erhalten bleibenden Bezugs zur Antwortskala das arithmetische Mittel eines Items in der Stichprobe verwendet. Bei der im QUICC verwendeten Antwortskala von 0 bis 4 bedeutet ein Mittelwert von 2 demnach eine mittlere Itemschwierigkeit.

Um einen möglichst hohen Informationsgehalt der Items einerseits und die Basis für hohe Trennschärfen andererseits zu gewährleisten, werden beim QUICC Schwierigkeitsindices zwischen 1 und 3 angestrebt.

Die beiden Kennwerte Schwierigkeit und Trennschärfe spielen neben anderen Erwägungen eine wichtige Rolle bei der Selektion von Items zur Testoptimierung.

6.3 Die Bestimmung der Validität

6.3.1 Interne Validität

Ein Test gilt als intern valide, wenn ein theoretisch begründetes Testmodell auf die empirisch erhobenen Daten passt (Rost, 2004). Während mit „erhobenen Daten“ schlicht die Antworten von Patienten auf die Items des QUICC gemeint sind, wäre zu präzisieren, wie ein „theoretisch begründetes Testmodell“ definiert werden kann. Dazu wird in diesem Abschnitt etwas ausführlicher zunächst der weit verbreitete faktorenanalytische Ansatz diskutiert und anschließend einem probabilistischen Modell gegenübergestellt.

Der faktorenanalytische Ansatz

Direkt zu beobachten sind beim Einsatz eines Tests lediglich die Itemantworten. Möchte man ein Testmodell spezifizieren, muss postuliert werden, wie diese Itemantworten zustande gekommen sind. Im faktorenanalytischen Modell nähert man sich

diesem Problem, indem man annimmt, dass die Itemantworten vom Einfluss theoretischer oder ‚latenter Variablen', abhängen, den sog. ‚Faktoren'. Formal wird diese Abhängigkeit folgendermaßen präzisiert: Der Messwert X_{vi} einer Person v bei Item i setzt sich additiv zusammen aus einer gewichteten Summe von Faktorwerten F_{vi}, die diese Person auf den zu postulierenden h Faktoren hat. Die Gewichte oder ‚Ladungen' werden gemeinhin durch die Koeffizienten a_{ij} ausgedrückt, die angeben, in welchem Ausmaß ein Faktor j die Ausprägung des Items i bestimmt. Zusätzlich wird ein personen- und itemspezifischer Fehler E_{vi} in das Modell integriert; das Modell ist also nicht deterministisch. Daraus ergibt sich die Gleichung

$$X_{vi} = \sum_{j=1}^{h} a_{ij} F_{vj} + E_{vi}$$

In der Praxis liegt zur Bestimmung dieser Parameter als Basis nur die Korrelationsmatrix der Items vor, aus der man die Ladungsmatrix Λ bestimmen muss, die den Einfluss der h Faktoren beschreibt. Für die Bestimmung werden einige einschränkende Annahmen über die Eigenschaften dieser Matrix gemacht (vgl. ausführlich z.B. Schönemann und Borg, 1996). Zusätzlich wird angenommen, dass die Daten intervallskaliert und Faktoren und Messfehler sowie Messfehler untereinander unkorreliert sind. Damit lässt sich eine Ladungsmatrix konstruieren, mithilfe derer sich eine theoretische Korrelationsmatrix R berechnen lässt, die eine möglichst gute Übereinstimmung mit der ursprünglichen empirischen Korrelationsmatrix haben sollte. In diesem Fall könnte man sagen, dass das faktorenanalytische Testmodell im obigen Sinne gut auf die Daten passt.

Es gibt zumindest zwei grundsätzliche Einwände gegen diesen Ansatz. Zum einen steigt die Übereinstimmung von theoretischer und empirischer Korrelationsmatrix mit der Anzahl der verwendeten Faktoren. Im Extremfall sind Faktoren- und Itemanzahl identisch, was zu einer perfekten Reproduktion des Ausgangsdatensatzes führt. Anpassungsgüte und Sparsamkeit des Modells müssen also gegeneinander abgewogen werden. Wenn sich aus der Definition des zu messenden Konstrukts keine anzunehmende Faktorenanzahl ableiten lässt, richtet man sich meist nach dem Kaiser-Kriterium, das nur Faktoren mit Eigenwerten größer 1 einbezieht. Das ist allerdings reine Konvention. Auch die graphische Interpretation des Eigenwerteverlaufs (‚Scree-Plot') führt weder zu eindeutigen Entscheidungen, noch hat sie eine theoretisch

nachvollziehbare Begründung (Bühner, 2005).

Zum anderen ist auch bei festgelegter Faktorenanzahl trotz der getroffenen restriktiven – und keineswegs trivialen – Annahmen (siehe dazu z.B. Harman, 1976) die konstruierte Ladungsmatrix als Lösung nicht eindeutig. Vielmehr lassen sich unendlich viele weitere Ladungsmatrizen konstruieren, die identische Anpassungsgüte (Varianzaufklärung) erreichen (Schönemann und Borg, 1996). Diese können durch sog. ‚Rotationen' ineinander überführt werden. Faktoren und deren Einflüsse werden also nicht „gefunden", wie häufig formuliert wird, sondern eher „ausgewählt".

Die Unterbestimmtheit in der Festlegung von der Faktorenanzahl und der Zuordnung der Items zu Faktoren könnte man im vorliegenden Fall zwar überwinden, da aus dem Grounded Theory-Prozess acht Faktoren mit a priori zugeordneten Items begründet werden können. Dazu wäre eine sog. ‚konfirmatorische Faktorenanalyse' anzuwenden. Die dafür verwendeten Strukturgleichungsmodelle aber „tragen die Probleme der Faktorenanalyse in verschärfter Form fort" (Schönemann und Borg, 1996, S. 250), d.h. die dafür zusätzlich nötigen Annahmen machen das Modell noch fragwürdiger.

Neben diesen Einschränkungen des *Verfahrens* der Faktorenanalyse ist zusätzlich vor jeder Anwendung zu hinterfragen, ob das *Modell* der Faktorenanalyse grundsätzlich zum Untersuchungsgegenstand passt. Um dies zu beurteilen, müsste präzisiert werden, wie man sich das ‚Zusammenspiel' von Faktoren und Items inhaltlich vorstellt.

Folgt man einer ‚substantiellen Interpretation' dieses Zusammenspiels, so müsste man annehmen, dass die in obiger Gleichung definierten Linearkombinationen direkte Entsprechungen in physiologischen oder psychologischen Prozessen haben. Identifizierte man gemäß der gängigen Praxis der Faktorenanalyse im vorliegenden Fall die Kategorien von Ungewissheit als Faktoren, müsste man ihnen eine wie auch immer geartete ontologische Qualität zusprechen, damit sie die Itemantworten von Personen auf die im Modell beschriebene Weise *verursachen* können. Behandlungsungewissheit (Kategorie 6) existiert ‚real' jedoch genauso wenig wie etwa ‚Extraversion' oder ‚Zeitgeist'. Sie stellen Konzepte dar, die sich bestenfalls zur Beschreibung menschlichen Verhaltens als nützlich erweisen. Diese Beschreibung jedoch als direkten kausalen Wirkungszusammenhang aufzufassen, wäre schlicht naiv. Zumal diese Einwirkung auch noch auf genau jene additiv-multiplikative Weise aus der obigen Formel erfolgen müsste.

Alternativ könnte man die Faktorenanalyse auch als ein Verfahren zur Identifikation eines Vorhersagemodells im Sinne der multiplen Regression verwenden, bei dem es nur auf Sparsamkeit und Präzision der Vorhersagegleichung ankommt. Damit wäre es unwichtig, ob es reale Entsprechungen zu den Faktoren gibt, solange das Modell nur hinreichend verlässlich arbeitet. Für die Entwicklung von Einstellungstests etwa ist in erster Linie deren Vorhersagekraft entscheidend, während die vermittelnden Wirkprozesse zunächst irrelevant sind. Damit wäre auch die Festlegung auf die Art der Verknüpfung der Modellvariablen unkritisch, da sie nicht inhaltlich begründet werden müsste, sondern eine frei zu bestimmende Eigenschaft des Modells wäre. Es bliebe lediglich zu überlegen, ob eine andere Verknüpfung eine noch bessere Varianzaufklärung lieferte. Während diese ‚Regressionsinterpretation' des faktorenanalytischen Modells einige der oben genannten Unstimmigkeiten der substantiellen Interpretation vermeidet, büßt sie dagegen ihren ‚explanatorischen Charme' ein. Außerdem entstehen durch die Regressionsinterpretation neue Schwachpunkte. Wenn es nur um eine optimale Vorhersage geht, ist unverständlich, warum diese ausschließlich linear modelliert werden sollte. Ferner dienen die gebräuchlichen Rotationsverfahren dazu, leichter zu interpretierende Ladungsmatrizen zu erzeugen. Bei der Regressionsinterpretation sollte stattdessen die Korrelationsmatrix im Mittelpunkt des Interesses stehen, was aber die gebräuchlichen Ansätze nicht vorsehen.

Was bedeuten diese Ausführungen für die Beurteilung der internen Validität mittels Faktorenanalyse? Oben war interne Validität definiert worden als das Ausmaß, in dem ein theoretisch begründetes Testmodell auf einen Datensatz passt. Noch bevor man sich dieser Frage empirisch nähert, ist also zu klären, ob sich das faktorenanalytische Modell im vorliegenden Fall *grundsätzlich* theoretisch begründen lässt. Die Annahmen des Testmodells in der geschilderten substantiellen Auslegung sind für die Kategorien von Ungewissheit nicht haltbar, da eine materielle Entsprechung zu den Faktoren und deren Wirkungen fehlt. Die Bewertung der internen Validität ist auf diese Weise also nicht möglich.

Nicht wesentlich anders fällt das Urteil bei einer Regressionsinterpretation aus. Selbst wenn man bereit wäre, die verschiedenen Unstimmigkeiten bei der Regressionsinterpretation sowie der Faktorenanalyse im Allgemeinen hinzunehmen, hätte man nicht viel gewonnen: Es könnte bestenfalls gezeigt werden, dass ein hinreichend sparsames und präzises Vorhersagemodell aus den Daten entwickelt werden kann. Sofern man allerdings wie im vorliegenden Fall nicht an reinen Vorhersagen inter-

essiert ist, liefert dies keinen Hinweis auf die Validität des Fragebogens, geschweige denn der Theorie. Schließlich würden Faktoren und Ladungen zur Rechtfertigung des Testmodells gerade um jegliche inhaltliche Bedeutung bereinigt.

Aus diesen Gründen wird auf die Anwendung des faktorenanalytischen Modells zur Bestimmung der internen Validität in dieser Arbeit verzichtet.

Der probabilistische Ansatz

Die probabilistische Testtheorie (PTT) macht Annahmen darüber, wie die Antwort einer Versuchsperson auf ein Item mit bestimmten Eigenschaften der Person (Personenparameter) und solchen des Items (Itemparameter) zusammenhängt. Dieser Zusammenhang wird nicht deterministisch, sondern probabilistisch konzipiert und kann bei dichotomen Itemantworten sehr übersichtlich mit einer sog. ‚Itemfunktion' oder ‚Item Characteristic Curve' (ICC) dargestellt werden (Abbildung 6.2). Darin ist abzulesen, dass mit steigender Ausprägung des Personenparameters θ – etwa Intelligenz oder Behandlungsungewissheit – die Wahrscheinlichkeit p wächst, das Item zu lösen bzw. zu bejahen.

Wie aber kommt man von dieser plausiblen allgemeinen Beziehung zur Festlegung der Parameter, d.h. wie wird aus der fiktiven Abbildung 6.2 links eine ‚reale'? Dazu werden zunächst theoretische Annahmen über die Itemfunktionen eingeführt, um nachfolgend die freien Parameter dieser Funktionen anhand der Daten zu schätzen.[2]

Zunächst zu den Annahmen: Aus der gut nachvollziehbaren Formulierung „je höher die Fähigkeit desto höher die Lösungswahrscheinlichkeit" lässt sich ableiten, dass der Funktionsverlauf *monoton steigend* sein muss.[3] In der klassischen Testtheorie (KTT) nimmt man einen linearen Zusammenhang zwischen Personenparametern und Lösungswahrscheinlichkeiten an. Eine lineare Formalisierung bringt jedoch das Problem mit sich, dass ab einer bestimmten Fähigkeitsausprägung die Lösungswahrscheinlichkeit außerhalb des definierten 0-1-Intervalls liegt, z.B. links vom Schnitt-

[2] Die Modellparameter sind grundsätzlich bezogen auf eine Population. Daher können sie nicht ‚bestimmt' oder ‚berechnet' werden, sondern nur anhand einer Stichprobe geschätzt werden (vgl. Mittelwert vs. Erwartungswert in der klassischen Testtheorie).

[3] In manchen Anwendungsbereichen lassen sich auch nicht-monotone Funktionsverläufe begründen, etwa U-förmige bei Präferenzmessungen oder stufenförmige in vielen Bereichen der Wahrnehmung. Hier wie auch bei den meisten Leistungstests spielen sie dagegen keine Rolle.

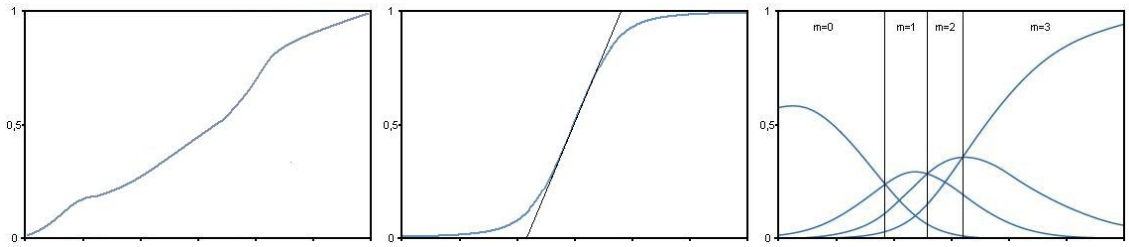

Abbildung 6.2: Der Zusammenhang zwischen der Ausprägung des Personenparameters θ und der Wahrscheinlichkeit, ein Item zu lösen (links). Solch ein Zusammenhang könnte z.B. linear oder wie beim Raschmodell logistisch angenähert werden (Mitte). Bei mehreren Antwortmöglichkeiten m_i wird für jede Kategorie eine eigene Funktion eingeführt, die jeweils innerhalb eines bestimmten Bereichs die höchste Wahrscheinlichkeit aufweist (rechts).

punkt der Geraden mit der x-Achse. Sinnvoll erscheint daher die Annahme eines Graphen, der in den Extrembereichen abgeflacht ist, während er im Mittelbereich annähernd linear ansteigt. Damit könnte die psychologisch einleuchtende Überlegung operationalisiert werden, dass die Lösungswahrscheinlichkeit im steilen Mittelbereich am stärksten mit zunehmender Fähigkeit wächst. Ist ein Item dagegen deutlich zu leicht oder zu schwer, dann verändert sich die Lösungswahrscheinlichkeit auch nur unwesentlich mit der Fähigkeit. Entsprechend hat Rasch (1960) die ‚logistische Itemfunktion' vorgeschlagen, die auf einer logarithmischen Transformation der Lösungswahrscheinlichkeit beruht. Die Lage der Kurve auf der x-Achse gibt dabei die Itemschwierigkeit an, während man die Steigung im Wendepunkt als Trennschärfe interpretieren kann (vgl. ausführlich Rost, 2004). Das dichotome Raschmodell lässt sich auch auf den Fall von mehr als zwei Antwortkategorien verallgemeinern, wie Masters (1982) gezeigt hat (vgl. Abbildung 6.2 rechts).

Im vorliegenden Fall lässt sich noch eine weitere Annahme begründen. Setzte man wie bei der KTT Intervallskalenniveau voraus, hieße das u.a., dass der Abstand zwischen den Antwortkategorien 1 und 2 genauso groß sein müsste wie der zwischen 2 zu 3. Diese strikte Anforderung ist kaum zu halten. Da allerdings beim QUICC alle Items anhand desselben Antwortformats bearbeitet werden, ist es plausibel, dass die Abstände der Antwortkategorien *zwischen* den Items identisch sind – aber eben nicht *innerhalb* der Items. Eine solche Festlegung führt zum sog. ‚Ratingskalenmodell' (Andrych, 1978).

Mit diesen Annahmen lassen sich nun die Modellparameter aus einer Datenmatrix

heraus mithilfe der ‚Maximum-Likelihood-Methode' iterativ schätzen. Vereinfacht gesprochen handelt es sich dabei um ein ‚fortwährend verfeinertes Probieren', um zu beurteilen, wie wahrscheinlich es ist, dass eine Datenmatrix aufgrund einer jeweiligen Parameterkombination zustande gekommen ist.

Auf diese Weise kann man einerseits entscheiden, ob die im Ratingskalenmodell theoretisch begründete Restriktion der Antwortschwellenparameter empirisch angemessen ist. Für den aktuellen Kontext noch entscheidender ist indes die Möglichkeit, die empirische Passung des postulierten achtdimensionalen Modells des Ungewissheitserlebens mit der einer eindimensionalen Konzeption zu vergleichen, wie sie in bisherigen Ansätzen zum Ungewissheitserleben verbreitet ist.

Dazu sind einerseits inferenzstatistische Modellgeltungstests gebräuchlich, etwa anhand des χ^2-Tests. Damit die Prüfstatistik wenigstens annähernd einer χ^2-Verteilung folgt, muss die Stichprobe allerdings mindestens so viele Personen enthalten, wie es verschiedene potentielle Antwortmuster gibt (Rost, 2004). Selbst wenn man die 52 QUICC-Items dichotomisierte, wären also immer noch 2^{52} d.h. etwa 4,5 Billiarden Teilnehmer nötig. Daher wird auf einen inferenzstatistischen Modelltest in dieser Arbeit verzichtet.

Andererseits kann man ein anderes Gütekriterium zur Beurteilung von Theorien heranziehen: deren Sparsamkeit. Analog zur Erhöhung der Faktorenanzahl bei der Faktorenanalyse führt auch im probabilistischen Ansatz die Verwendung zusätzlicher Dimensionen, verbunden mit einer Erhöhung der Freiheitsgrade, grundsätzlich zu einem besseren Modellfit. Daher liegt es nahe, letzteren in Relation zur Anzahl investierter Modellparameter zu setzen. Zwei verschiedene Arten dieser Verhältnisbildung sind die sog. ‚Informationskriterien' AIC (Akaike's Information Criterion) und BIC (Bayes' Information Criterion). Die beiden Indices unterscheiden sich nur darin, wie stark die Parameteranzahl gewichtet wird, wie man den Berechnungsformeln im Anhang A.4 leicht entnehmen kann. Beide Gewichtungen sind jedoch eine Konvention und nicht inhaltlich begründet. Damit hat man einen gewissen Ermessensspielraum, um wie viel etwa ein AIC-Wert besser sein müsste, um ein Modell einem anderen vorzuziehen (Rost, 2004). Spätestens an dieser Stelle kommen wieder theoretische Argumente zum Tragen. Außerdem kann man mit den Informationskriterien nur zu entscheiden versuchen, welches Modell aus einer gegebenen Menge am besten auf die Daten passt. Dadurch ist allerdings nicht ausgeschlossen, dass es andere, noch besser passende Modelle geben könnte.

Insgesamt scheint nach dem hier Ausgeführten ein probabilistisch fundierter Vergleich des vom Grounded Theory-Prozess abgeleiteten achtdimensionalen Modells mit einem konventionellen eindimensionalen Modell angemessen, *eine* Antwort auf die Frage der internen Validität zu geben. Der Vorzug gegenüber dem faktorenanalytischen Modell liegt einerseits im möglichen Verzicht auf die Annahme von Intervallskalenniveau und Normalverteiltheit der untersuchten Variablen. Andererseits lassen sich die Annahmen über die Personen- und Itemparameter besser begründen und flexibler auswählen. Der für probabilistische Modelle meist höhere Rechenaufwand hat seit der Verfügbarkeit von entsprechender Software an Relevanz verloren. Gleichwohl ist für diese Art der Modelltestung eine recht große Stichprobe nötig, so dass sie erst in Validierungsstudie 4 (Abschnitt 6.8) erfolgen kann.

Weitere Überlegungen zur internen Validität

Bereits für die erste Erhebung lässt sich aber zusätzlich noch anders argumentieren. Der QUICC beansprucht, acht inhaltlich differenzierbare Qualitäten von Ungewissheit zu erfassen. Für jeden dieser Bereiche sind mehrere Items formuliert worden. Vom theoretischen Standpunkt her wäre ein Item pro Kategorie ausreichend, wenn es denn den Inhalt der Kategorie hinreichend treffend erfasste. Mehrere Items werden einerseits verwendet, um Fehlerschwankungen durch parallele Messungen auszugleichen.[4] Dafür werden verschiedene Formulierungen desselben Aspekts dargeboten. Andererseits kann sich die prototypische ‚Idee' einer Kategorie in verschiedenen charakteristischen Situationen verwirklichen, von denen nicht jede auf eine Person zutreffen muss. Um daher die Ausprägung der Ungewissheit bezüglich der Kategorie nicht zu unterschätzen, erfragen die Items auch unterschiedliche Aspekte derselben Ungewissheitskategorie.

Die intendierte Ähnlichkeit der Itemformulierung im ersten Fall sollte zu ähnlichen Antworten bei jedem Teilnehmer führen. Diese wiederum spiegelten sich in positiven Iteminterkorrelationen innerhalb jeder Skala wider. Wenn stattdessen im zweiten Fall Patienten nur manchen Items einer Kategorie zustimmen, anderen dagegen nicht, mindert dies die Korrelationen zwischen eben diesen Items. Betrachtet man

[4] Als Fehlerschwankungen kann man auch Antwortdifferenzen aufgrund interindividuell unterschiedlichen Verständnisses von Formulierungen bezeichnen. Diese sind keine Zufallsschwankungen und insofern vermutlich stabil. Allerdings zählen sie zu den nicht intendierten Unterschieden, da sie nicht auf die Ausprägung der latenten Variable zurückgehen.

dagegen alle Iteminterkorrelationen einer Kategorie zusammen, wäre zu erwarten, dass sich solche lokalen Gegebenheiten ausgleichen.

Zusammenfassend kann man also sagen, dass hohe Korrelationen innerhalb einer Kategorie für die Homogenität der Skala sprechen. Cronbachs α stellt einen Quotienten dar, der diese handlich auf einen einzigen Wert reduziert. Zwar lässt sich Cronbachs α leicht durch die Verwendung vieler Items künstlich erhöhen, ohne dass dafür hohe Interkorrelationen nötig wären.[5] Allerdings dürfte dieser Gesichtspunkt bei den gegebenen Itemanzahlen nicht relevant werden. Cronbachs α wird hier also sowohl als Hinweis auf die Reliabilität als auch die interne Validität verwendet.

Neben der Homogenität der Skalen wäre es zudem Ausdruck der internen Validität des QUICC, wenn sich die im Grounded Theory-Prozess (vgl. Abschnitt 5.1) gebildete Differenzierung des Ungewissheitserlebens in acht Kategorien[6] empirisch-statistisch untermauern ließe. Aus der inhaltlichen Differenzierbarkeit unkorrelierte Skalen des QUICC zu folgern, wäre aber dennoch unangemessen. Denn nach den in den Kapiteln 4 und 5.1.3 ausgebreiteten Gedanken treten verschiedene Qualitäten von Ungewissheit meist gleichzeitig und einander wechselseitig beeinflussend auf. Hohe Skaleninterkorrelationen als Anzeichen von gleichzeitiger Präsenz verschiedener Ungewissheitsqualitäten sollten demnach eher die Regel als die Ausnahme sein. Dies ist nebenbei ein weiterer Grund, keine Faktorenanalyse zur Bestätigung der Kategorienstruktur zu verwenden, da diese doch gerade auf den Korrelationen der Items basiert. Auch ein obliques Modell käme hier kaum zu eindeutigen Ergebnissen.

Ein starker Beleg für die postulierte Theorie sowie die Qualität des Messinstruments wäre es dagegen, wenn sich zeigen ließe, dass durch eine SDM-Intervention die Variabilität in der Ausprägung der Ungewissheitsqualitäten bei Patienten zunimmt (vgl. Abschnitt 4.3). Solch eine Variabilität stellt jedoch kein absolutes Maß dar, denn die Ungewissheitsrepräsentation hängt sowohl von individuellen wie krankheitsspezifischen Aspekten ab. Insofern kann lediglich die Veränderung gegenüber

[5] Diesen Aspekt kann man sich leicht anhand der in Anhang A.3 aufgeführten Formel klarmachen.

[6] In dieser Arbeit werden die Begriffe ‚Ungewissheitsqualität', ‚-kategorie' und ‚-skala' verwendet. Alle drei verweisen auf verschiedene Ebenen. Während ‚Qualität' sich am ehesten auf die Erlebnisebene von Patienten anwenden lässt, bezieht sich ‚Kategorie' mehr auf die in dieser Arbeit vorgeschlagene thematische Trennung, wie sie dem Grounded Theory-Prozess entsprungen ist. ‚Skala' wird dagegen im Zusammenhang mit der quantitativen Umsetzung im QUICC verwendet. Zuweilen gelten Aussagen jedoch auf mehreren Ebenen gleichzeitig. In diesen Fällen wird meist dennoch nur die im jeweiligen Kontext relevanteste genannt.

dem Ausgangszustand bei jedem Patienten beurteilt werden. Dazu ist ein Design mit mindestens zwei Messzeitpunkten und einer SDM-Intervention erforderlich, wie es in Abschnitt 6.9 vorgestellt wird. Die Variabilität wird dabei operationalisiert als Varianz zwischen den acht QUICC-Skalen pro Patient und Messzeitpunkt. Damit kann man die Hypothese zunehmender Variabilität prüfen, indem die durchschnittliche Varianz vor und nach der SDM-Intervention verglichen wird. Um bei dieser Prüfung keine Verteilungsannahmen hinsichtlich der Variable Varianz machen zu müssen, erfolgt die Prüfung anhand des parameterfreien Wilcoxon-Tests für abhängige Stichproben.

6.3.2 Externe Validität

Zusätzlich ist die externe Validität des QUICC zu prüfen. Nach Rost (2004) gilt ein Test als extern valide, wenn sich theoretisch plausible Zusammenhänge mit Variablen nachweisen lassen, die nicht Teil des Tests selbst sind. Gäbe es z.B. bereits ein etabliertes Pendant zum QUICC, ließen sich die Resultate beider Tests vergleichen (konkurrente Validität).

Ohne größeren Aufwand kann das Behandlungsstadium der Patienten erfasst werden, d.h. ob sie ihre Behandlung bereits abgeschlossen haben und in der Nachsorge waren oder ob die Behandlung noch läuft. Es ist anzunehmen, dass die Ungewissheitsqualitäten insgesamt in der Nachsorge weniger salient sind als während der Behandlung. Dies lässt sich dadurch begründen, dass die Chance mit zunehmender Erkrankungsdauer steigt, Antworten auf die Ungewissheiten zu finden. So könnte sich etwa die Partnerschaft während der Erkrankung als tragfähig erwiesen haben (Kategorie 1). Oder die Rückkehr ins Berufsleben ist geglückt (Kategorie 3). Die größte Reduktion ist indes für die Behandlungsungewissheit (Kategorie 6) zu erwarten, da sie am spezifischsten von der Gruppierungsvariable betroffen ist und entsprechende Fragen nach Abschluss der Behandlung meist nicht mehr relevant sind. Der geringste Unterschied sollte bei der Kausalattribution (Kategorie 4) auftreten. Denn Fragen nach dem Sinn oder den lebensplanerischen Konsequenzen einer Krebserkrankung sind einerseits während der Behandlung vermutlich noch weniger präsent als solche nach der passenden Behandlungsstrategie, da zunächst der Kampf gegen den Krebs im Vordergrund steht. Daraus ergibt sich ein vergleichsweise niedri-

ger Ausgangswert.[7] Andererseits erledigt sich die Frage der Kausalattribution selten durch externe Entwicklungen oder Erfahrungen, so dass diese Ungewissheit länger persistieren dürfte. Insgesamt sollte dies in einer geringeren Differenz resultieren.

Diese Effekte können mit dem t-Test für unabhängige Stichproben auf Signifikanz geprüft werden. Etwaige Verletzungen der Verteilungsvoraussetzungen können in Anbetracht der ausreichend großen Stichprobe vernachlässigt werden.[8]

Außerdem wird in Validierungsstudie 1 der Zusammenhang zwischen Ungewissheitserleben und der sog. ‚Ungewissheitstoleranz' untersucht. Diese lässt sich beschreiben als die Fähigkeit von Menschen, ungewisse, ambigue, komplexe oder unvorhersehbare Situationen hinzunehmen (ausführlich z.B. in Dalbert, 1996; 1999). Dabei wäre zu erwarten, dass denjenigen Patienten Ungewissheiten mehr durch den Kopf gehen, die weniger ungewissheitstolerant sind und die diese Ungewissheiten folglich als belastender empfinden. Daraus ergäbe sich eine negative Korrelation zwischen Ungewissheitserleben und Ungewissheitstoleranz. Letztere wird in dieser Studie anhand der acht Items umfassenden ‚Ungewissheitstoleranzskala' erfasst (UGTS, Dalbert, 1999; 2002). Die UGTS ist in Anhang F zu sehen.

6.4 Weitere Überlegungen zur Bewertung des QUICC

Es ist wichtig festzuhalten, dass die Beurteilung von Reliabilität, Validität etc. mit den geschilderten Verfahren stichprobenabhängig ist. Das bedeutet, dass sich z.B. ein Fragebogen trotz hoher Reliabilitätskennwerte in der einen Stichprobe in einer anderen durchaus als unreliabel erweisen könnte. Nun kann eine Fragebogenvalidierung kaum an jeder denkbaren Population erfolgen. Um aber wenigstens etwas sicherer sein zu können, dass die Eigenschaften des QUICC nicht auf eine Stichprobe

[7] Direkte Vergleiche zwischen den Kategorien sind dabei nicht zulässig, da die gemessene Ausprägung von der Itemschwierigkeit abhängig ist.

[8] Die Prüfung der Voraussetzungen Normalverteiltheit und Varianzenhomogenität etwa mit Hilfe des Kolmogorov-Smirnov-Tests und des Levene-Tests wird hier nicht vorgenommen, da die Hypothesenstruktur bei diesen Tests der inferenzstatistischen Logik widerspricht; dadurch wird der β-Fehler entscheidend, ohne dass man ihn kontrollieren oder abschätzen kann. Stattdessen sollte man daher auf die Robustheit des t-Tests im Falle der Verletzung dieser Annahmen vertrauen, wenn die Daten als zufällig und unabhängig erhoben angesehen werden können.

beschränkt sind, werden im Laufe dieser Validierung Stichproben aus verschiedenen chronischen Krankheiten betrachtet.

Neben den Hauptgütekriterien Reliabilität und Validität werden indes in allen Studien die Objektivität und Praktikabilität des QUICC betrachtet.

6.5 Validierungsstudie 1

Die folgende Studie bei Krebspatienten dient der quantitativen Validierung des QUICC in seinem aktuellen Entwicklungsstand. Nach den Ausführungen in Abschnitt 6.3 werden dazu folgende Annahmen geprüft:

- Reliabilität
  - Die acht Skalen des QUICC sind reliabel, d.h. Cronbachs $\alpha > 0,8$.
- Validität
  - Die Skalen des QUICC sind homogen, d.h. Cronbachs $\alpha > 0,8$.
  - Patienten in Behandlung erleben mehr Ungewissheit als solche in der Nachsorge, d.h. sie haben höhere Skalenmittelwerte. Der größte Unterschied besteht bei der Behandlungsungewissheit (Skala 6), der geringste bei der Kausalattribution (Skala 4).
  - Je geringer die Ungewissheitstoleranz laut Ungewissheitstoleranzskala (UGTS), desto intensiver ist das Ungewissheitserleben, d.h. der Mittelwert der UGTS korreliert negativ mit den Skalenmittelwerten im QUICC.
- Itemschwierigkeit
  - Die Items des QUICC sind von mittlerer Schwierigkeit, d.h. die Schwierigkeitsindices liegen zwischen 1 und 3.
- Trennschärfe
  - Die Items des QUICC sind trennscharf, d.h. die korrigierten Trennschärfeindices sind größer als 0,5.
- Objektivität

  - Der QUICC ist durchführungs- und auswertungsobjektiv, d.h. die Testergebnisse hängen nicht vom Versuchsleiter oder anderen äußeren Gegebenheiten während der Erhebung oder der Auswertung ab.

- Praktikabilität
  - Der QUICC ist praktikabel, d.h. die Bearbeitungsdauer liegt unterhalb von 30 Minuten.

6.5.1 Stichprobe

Als Validierungsstichprobe wurden Krebspatienten ausgewählt, die sich zwischen Dezember 2004 und Januar 2005 zur Behandlung im Wartebereich der Strahlentherapeutischen Ambulanz des Universitätsklinikums Schleswig-Holstein, Campus Kiel, einfanden. Diese Wahl resultierte aus der zu erwartenden Mischung von Patienten in aktueller Behandlung und solchen in anschließender Nachsorge sowie aus dem hohen Patientendurchlauf, der eine ökonomische Datenerhebung versprach.

Von den 341 erschienenen Patienten wurden 319 um eine Teilnahme an der Validierungsuntersuchung gebeten, die übrigen 22 mussten aus organisatorischen Gründen ausgelassen werden. 50 der angesprochenen 319 Patienten (15,7%) verweigerten direkt die Teilnahme aus persönlichen, terminlichen oder gesundheitlichen Gründen. Weitere 28 (8,8%) kamen ihrer Ankündigung nicht nach, den Fragebogen in einem bereits frankierten Umschlag zurückzusenden. Somit gingen zunächst 241 Patienten (75,5%) in die Untersuchung ein.

Aus dieser Gruppe wurden einerseits nachträglich 7 Teilnehmer ausgeschlossen, die etwa durch Überblättern einer Seite zu viele fehlende Werte produzierten. Andererseits ließ es sich wegen der Priorität des straffen Behandlungsablaufs nicht verhindern, dass 107 Patienten während des Ausfüllens zum Behandlungsgespräch gerufen wurden. Obwohl alle diese Patienten den Fragebogen anschließend zu Ende ausfüllten, wurden sie aus der Auswertung ausgeschlossen, da sich ihre Antworten nach dem Gespräch auf ein vermutlich durch diese Intervention verändertes Ungewissheitserleben bezogen. In diesem organisatorischen Problem der Datenerhebung liegt zwar auch eine Chance: So könnte man anhand der beiden Testhälften zu überprüfen versuchen, ob sich durch das Gespräch das Ungewissheitserleben tatsächlich verändert hat. Allerdings ließ sich in der Regel nicht festhalten, an welcher Stelle

Bösartige Neubildungen betreffend ...	Anteil in Prozent
Brustdrüse	38,6 (62,8)
Männliche Geschlechtsorgane	13,4 (38,6)
Lippe, Mundhöhle, Pharynx	8,7
Lymphatisches, blutbildendes und verwandtes Gewebe	7,9
Verdauungsorgane	7,1
Atmungs- und sonstige intrathorakale Organe	5,5
Weibliche Geschlechtsorgane	5,5 (8,9)

Tabelle 6.1: Die häufigsten Krebsarten unter den eingeschlossenen Teilnehmern. Gegebenenfalls sind Anteile im jeweils allein betroffenen Geschlecht in Klammern angegeben.

des QUICC die Unterbrechung stattfand, d.h. ob z.B. nach der dritten oder der dreißigsten Frage. Daher schied diese Möglichkeit aus. Somit verblieben 127 verwertbare Fragebögen.

Von diesen 127 Patienten waren 34,6% männlich und 61,4% weiblich.[9] Das Alter lag durchschnittlich bei 59,6 Jahren ($SD = 12,8$). Diese Kennwerte differieren kaum von denen der Gesamtstichprobe (Altersdurchschnitt 61,5 ($SD = 12,7$); 36,1% männlich und 62,5% weiblich).

Die häufigsten Krebserkrankungen unter den letztendlich eingeschlossenen 127 Patienten waren bösartige Neubildungen der Brustdrüse (ICD-10 C50) mit einem Anteil von 38,6% (62,8% der eingeschlossenen Frauen) sowie bösartige Neubildungen der männlichen Geschlechtsorgane (C60-C63.9: 13,4% bzw. 38,6% der Männer). Die Häufigkeiten weiterer Diagnosen sind Tabelle 6.1 zu entnehmen.

6.5.2 Ergebnisse

Praktikabilität. Die Bearbeitungszeit für jeden Patienten lag bei 20-30 Minuten. Die Patienten schienen kaum Bedarf zu haben, zwischendurch Rückfragen zu klären, so dass die Bearbeitung weitestgehend eigenständig erfolgen konnte. Insbesondere Patienten mit einem schlechteren Gesundheitszustand und daraus resultierendem geminderten Konzentrationsvermögen schienen jedoch teilweise etwas überfordert zu sein, was sich in längerer Bearbeitungszeit äußerte.

[9] An 100% fehlende Anteile gehen bei dieser Studie auf lückenhafte Erfassung der betreffenden Daten zurück.

Objektivität. Im praktischen Einsatz erschien der QUICC durch die mit einem Beispielitem versehene Einleitung und die selbständige Beantwortung durchführungsobjektiv. Das gebundene Antwortformat gewährleistete die Auswertungsobjektivität. Speziell bei dieser Erhebung wurden allerdings zusätzlich zur gedruckten Instruktion bereits bei der Bitte um Teilnahme an der Untersuchung einige erläuternde Worte zu deren Zweck gesprochen, die nicht als standardisiert gelten können.

Itemschwierigkeit. Nach einer Transformation der verbalen Antwortstufen „gar nicht“, „etwas“, „mittel“, „ziemlich“ und „sehr“ in numerische Kodierungen von 0 bis 4 ergaben sich Itemschwierigkeiten zwischen 1,02 und 2,64. Der Median[10] betrug 1,88. Eine ausführliche Darstellung dieser und der nachfolgenden Parameter findet sich in Tabelle 6.2.

Skala	Item	Schwierigkeit	Trennschärfe	Streuung	N	Reliabilität
1	1	1,45	0,49	1,57	127	0,85
	7	1,43	0,61	1,43	127	
	13	1,02	0,64	1,28	125	
	20	1,78	0,62	1,45	126	
	27	1,73	0,64	1,69	124	
	34	1,36	0,58	1,39	125	
	43	1,13	0,67	1,24	127	
2a	2	2,37	0,62	1,30	127	0,90
	14	2,64	0,81	1,26	126	
	21	1,96	0,71	1,46	126	
	31	2,31	0,75	1,45	127	
	41	2,08	0,79	1,37	127	
	47	2,46	0,76	1,42	127	
2b	8	1,87	0,73	1,50	126	0,90
	17	1,55	0,77	1,49	126	
	28	2,11	0,67	1,62	126	
	36	1,41	0,57	1,35	124	
	44	1,61	0,79	1,51	127	
	51	1,93	0,83	1,57	126	
3	3	2,22	0,74	1,30	125	0,91
	9	1,39	0,64	1,51	126	

[10] Wegen des ungewissen Skalenniveaus wird als Maß der zentralen Tendenz hier und im Folgenden stets der Median statt des Mittelwertes angegeben.

Skala	Item	Schwierigkeit	Trennschärfe	Streuung	N	Reliabilität
	15	2,03	0,72	1,46	125	
	23	1,58	0,68	1,41	127	
	29	1,87	0,72	1,44	126	
	37	1,93	0,75	1,40	127	
	45	1,83	0,72	1,51	127	
4	4	1,62	0,53	1,47	124	0,89
	10	1,88	0,74	1,48	126	
	16	1,52	0,79	1,40	124	
	24	1,64	0,70	1,42	125	
	30	1,69	0,77	1,61	126	
	38	1,41	0,72	1,58	126	
	48	1,46	0,68	1,40	121	
5a	5	1,33	0,63	1,38	126	0,90
	18	1,75	0,75	1,516	126	
	25	2,07	0,74	1,59	127	
	35	1,80	0,82	1,49	125	
	42	1,71	0,81	1,50	126	
	49	2,30	0,70	1,60	127	
5b	11	1,98	0,79	1,64	127	0,95
	22	2,04	0,86	1,57	127	
	32	2,04	0,86	1,68	125	
	39	2,16	0,79	1,58	127	
	46	1,94	0,83	1,74	126	
	52	2,05	0,88	1,69	127	
6	6	1,93	0,79	1,46	127	0,92
	12	1,93	0,81	1,48	123	
	19	2,03	0,66	1,34	126	
	26	1,94	0,76	1,58	124	
	33	2,10	0,67	1,46	126	
	40	1,86	0,78	1,42	125	
	50	2,23	0,73	1,56	127	

Tabelle 6.2: Die Itemparameter Schwierigkeit, Trennschärfe und Streuung der 52 QUICC-Items sowie die Skalenreliabilitäten. Variierende Stichprobengrößen gehen auf einzelne fehlende Werte zurück.

Skala	$M_{Behandlung}$	$M_{Nachsorge}$	d	t	df	p
1	1,76	1,14	0,54	3,24	115	0,01
2a	2,71	1,96	0,72	3,69	115	0,01
2b	2,21	1,39	0,73	3,70	115	0,01
3	2,28	1,47	0,66	3,96	115	0,01
4	1,81	1,37	0,39	2,09	115	0,02
5a	2,35	1,46	0,72	4,01	115	0,01
5b	2,59	1,62	0,66	3,65	115	0,01
6	2,52	1,60	0,77	4,32	115	0,01

Tabelle 6.3: Unterschiede in den Skalenmittelwerten zwischen Patienten während ($N = 44$) und nach einer Behandlung ($N = 77$). Die Irrtumswahrscheinlichkeiten sind einseitig. Bei sechs Patienten lag keine Angabe zur Krankheitsphase vor.

Trennschärfe. Die Part-Whole-korrigierten Trennschärfen der 52 Items rangierten zwischen 0,49 und 0,88. Der Median betrug 0,74.

Reliabilität. Die mediane Reliabilität der acht Skalen lag bei 0,90 (Cronbachs α). Den niedrigsten Wert lieferte die Skala Soziale Einbindung (Skala 1) mit 0,85, der höchste Wert stammte mit 0,95 von der Skala 5b, die Ungewissheiten in Bezug auf den Arzt erfragt.

Validität. Die im Rahmen der Bewertung der internen Validität als Zeichen der Homogenität der Skalen zu betrachtenden Cronbachs α-Werte wurden bereits erwähnt.

Zur externen Validität fand sich das postulierte Muster der Mittelwertsunterschiede zwischen den Patienten in Behandlung und denen mit abgeschlossener Behandlung. Letztere wiesen auf allen acht Skalen des QUICC niedrigere Mittelwerte auf ($p < 0,02$). Erwartungskonform war dabei die größte Effektstärke[11] bei der Behandlungsungewissheit (Skala 6) zu verzeichnen, die kleinste bei der Kausalattribution (Skala 4). Die Behandlung lag in der Nachsorgegruppe durchschnittlich 27,31 Monate zurück ($SD = 25,06$). Die Werte sind im Detail Tabelle 6.3 zu entnehmen.

Der erwartete Zusammenhang zwischen Ungewissheitserleben und Ungewissheitstoleranz ließ sich ebenfalls nachweisen. Der Gesamtscore der UGTS korrelierte durch-

[11] Die Effektstärke d errechnet sich aus der Differenz der Mittelwerte relativiert an der Standardabweichung.

weg negativ mit denen der Ungewissheitsskalen, d.h. ungewissheitstolerante Patienten beschäftigten sich tendenziell weniger mit Ungewissheit. Die Korrelationen lagen zwischen -0,08 und -0,25 ($p < 0,05$ für die Skalen 3 und 6; $p < 0,01$ für die Skalen 2b, 5a und 5b).

6.5.3 Diskussion

Die hier vorgestellten Ergebnisse belegen, dass der QUICC in dieser Stichprobe ein praktikables und sehr reliables Messinstrument ist. Alle Reliabilitätsmaße übertrafen deutlich die vorgegebene Untergrenze von 0,8. Bei der Itemanalyse verfehlte nur das erste Item die erhoffte minimale Trennschärfe von 0,5. Hier wäre zu prüfen, ob es sich um einen Positionseffekt handelte. Die Itemschwierigkeitsparameter gruppierten sich recht eng um den Mittelpunkt der Skala, so dass die Items gut im mittleren Ausprägungsbereich differenzierten, wobei alle Items im angepeilten Bereich zwischen 1 und 3 lagen. Decken- oder Bodeneffekte sind somit zumindest in diesem Patientenkollektiv unwahrscheinlich.

Weiterhin gingen aus der qualitativen Fragebogenentwicklung offensichtlich sehr homogene Skalen hervor, wie die außerordentlich hohen Cronbachs α-Werte belegen. Das gilt umso mehr, als die Itemanzahlen pro Skala nicht übermäßig groß sind. Dies unterstreicht die interne Validität des QUICC. Dabei lassen die Kennwerte eine weitere Reduktion des Itempools zu. Dass bei dieser Studie eine große Zahl von Patienten ausgeschlossen werden musste, weil diese während des Ausfüllens ihr Arztgespräch hatten, erscheint nicht als Bedrohung dieser und der folgenden Schlussfolgerungen. Zwar wurde dadurch die Stichprobengröße reduziert, jedoch können daraus keine *systematischen* Einflüsse auf die Ergebnisse abgeleitet werden.

Die zur Prüfung der externen Validität postulierten Unterschiede zwischen aktuell behandelten Patienten und jenen in der Nachsorge konnten durchweg belegt werden. Nachsorgepatienten beschäftigen sich insgesamt weniger mit krankheitsrelevanten Ungewissheiten als Akutpatienten. Auch die erwartete Rangordnung der Unterschiede war nachweisbar. Die Robustheit dieser Befunde ist dabei insofern überzeugend, als durch die gewählte Gruppierungsvariable ‚Behandlungsstadium' teilweise grundverschiedene Krebsformen, Schweregrade, Behandlungsformen etc. aggregiert wurden, die zu erhöhten Intragruppenvarianzen geführt haben dürften. So schwankte

etwa die Zeit seit Abschluss der Behandlung in der Nachsorgegruppe zwischen 1 und 110 Monaten.

An dieser Stelle mag eingewendet werden, dass in früheren Kapiteln die globale Nichtreduzierbarkeit von krankheitsrelevanten Ungewissheiten begründet wurde, während sich hier eindeutige Reduktionen ergeben haben. Dazu sei auf die unterschiedlichen Zeiträume hingewiesen. Die in Abschnitt 4.3 beschriebene Nichtreduzierbarkeit steht im Kontext der Entscheidungsfindung zwischen verschiedenen Therapievarianten, also *vor* der betreffenden Therapie oder dem Therapieabschnitt. Zu diesem Zeitpunkt lassen sich viele Ungewissheiten in der Tat nicht oder nicht wesentlich reduzieren. Einige Monate oder gar Jahre nach Abschluss der Therapie stehen allerdings neue Informationen zur Verfügung, die die Ungewissheiten reduzieren können. Wenn beispielsweise zwei Jahre nach Abschluss der Therapie kein Rezidiv aufgetreten ist und die Nebenwirkungen der Behandlung erträglich gewesen sind, kann dies Patienten in der Meinung bestärken, dass sie eine gute Behandlungsentscheidung getroffen haben. Auch wenn sie nicht wissen können, ob in Zukunft Rezidive auftreten und wie andere Behandlungen verlaufen wären, mag diese Erfahrung eine Reduktion der Behandlungsungewissheit (Kategorie 6) bewirken. Nichtreduzierbarkeit von Ungewissheiten während der Entscheidungsphase steht also nicht im Widerspruch zu den Ungewissheitsreduktionen nach Abschluss der Behandlung.[12]

Obschon nicht Teil der Validierung des QUICC ist es in diesem Zusammenhang interessant, einen Blick auf den Verlauf des Ungewissheitserlebens zu werfen. Abbildung 6.3 fasst Nachsorgepatienten bis zu drei Jahre nach ihrer Behandlung gruppiert nach Halbjahresabschnitten zusammen. Zwar liegen nur Querschnittsdaten vor, so dass es sich hierbei auch um Kohorteneffekte handeln könnte. Dennoch lässt sich mit aller Vorsicht etwas über die Entwicklung des Ungewissheitserlebens mutmaßen: Während sich im ersten Jahr nach Abschluss der Behandlung tatsächlich die Ungewissheiten reduzieren, erleben Patienten zwischen 12 und 18 Monaten nach ihrer Behandlung sogar mehr Ungewissheit als währenddessen. Dies lässt sich womöglich so interpretieren, dass Patienten manche Ungewissheit erst mit einigem Abstand zur akuten Behandlung bewusst wird. Nachdem währenddessen vielleicht eher das ‚Funktionieren' im medizinischen System im Vordergrund stand, schließt sich eine Phase stärkerer Reflexion des Geschehenen an. Diese Deutung ist auch verträglich

[12] Dabei sei noch einmal betont, dass eine Behandlung nicht unbedingt wie im obigen Beispiel erfolgreich verlaufen sein muss, um die Behandlungsungewissheit zu verringern. Auch ein offensichtlich gescheiterter Heilungsversuch bringt diesbezüglich Gewissheit.

mit dem unter 4.3 Ausgeführten über die Spezifität der Auseinandersetzung mit Ungewissheiten im SDM. Die in der vorliegenden Studie erfassten Patienten haben keine kontrollierte SDM-Intervention erhalten, sondern wurden nach den gängigen Standards behandelt, die vermutlich nur selten dem unter 1.4 skizzierten SDM-Ideal entsprachen. Die – so könnte man spekulieren – während der Entscheidungsphase suboptimale Auseinandersetzung mit Ungewissheiten wird daraufhin später nachgeholt. Diese Hypothese wäre allerdings in einer prospektiven Studie mit einer kontrollierten SDM-Intervention zu überprüfen.

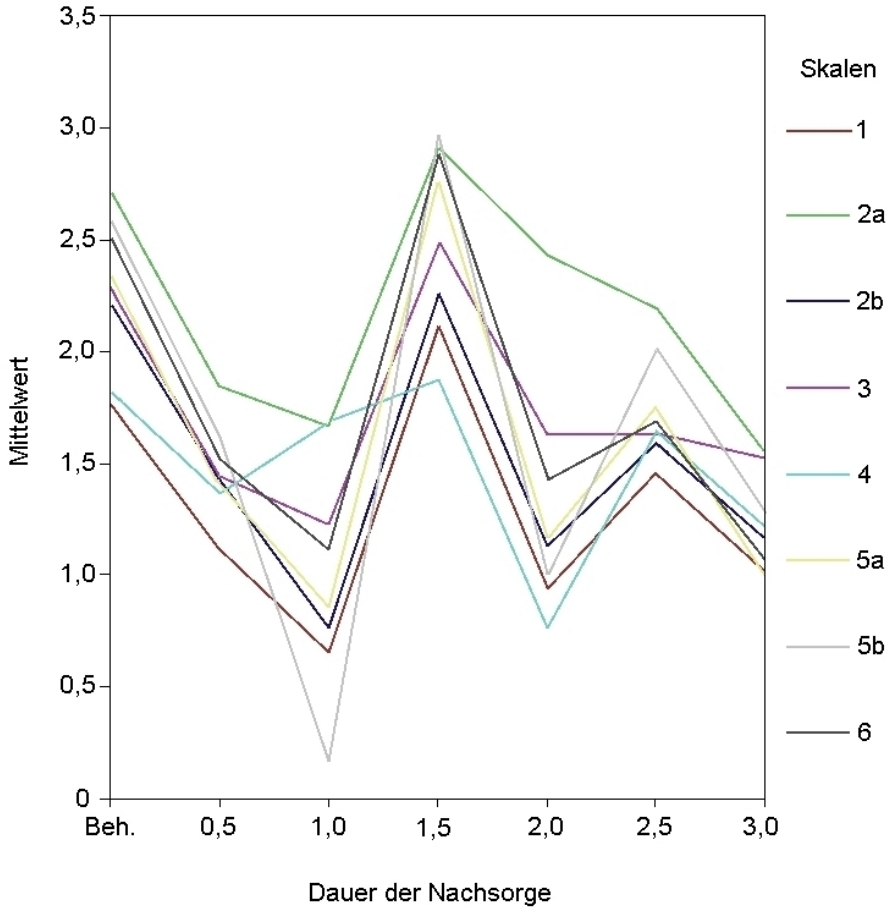

Abbildung 6.3: Querschnittsbetrachtung des Ungewissheitserlebens während der Behandlung und drei Jahre danach ($N = 97$).

Über den Einfluss der Behandlungsphase hinaus konnte bei der Frage der externen Validität des QUICC gezeigt werden, dass Ungewissheiten umso stärker erlebt werden, je geringer die Ungewissheitstoleranz ist. Die berichteten Korrelationen zu die-

ser Annahme sind zwar größtenteils signifikant, aber dennoch teilweise sehr gering, sodass von einem *inhaltlich* bedeutsamen Zusammenhang nicht unbedingt gesprochen werden kann. Dazu lassen sich zwei mögliche Erklärungen anführen. Zum einen wies die Ungewissheitstoleranzskala in dieser Untersuchung anders als im bisherigen Einsatz (Dalbert, 1999; 2002) nur eine ungenügende Reliabilität auf (Cronbachs $\alpha = 0,50$). Ein Grund dafür mag die alltagsnahe und allgemeine Formulierung der UGTS-Items sein, die im Kontext einer Krebsbehandlung insbesondere auch direkt nach Konfrontation mit den Items des QUICC zu verzerrtem Verständnis geführt haben könnten. So hat das Item 3 „Ich mag es, wenn unverhofft Überraschungen auftreten“ für Krebspatienten im Krankenhaus vermutlich eine andere Bedeutung als für gesunde Menschen im Alltag. Dabei dürfte klar sein, dass die Reliabilität der UGTS ein limitierender Faktor für deren Korrelation mit anderen Variablen wie etwa den Ungewissheitsskalen ist.[13] Neben diesem statistischen Aspekt könnte auch noch ein inhaltlicher die Korrelationen verringert haben: Zwar ist es einerseits möglich, dass Ungewissheitsintoleranten erwartungskonform die vorhandenen Ungewissheiten mehr durch den Kopf gehen. Andererseits gäbe es für diese Patienten auch den Ausweg, die Ungewissheiten zu verdrängen oder herunterzuspielen. Nach dieser Überlegung müssten Ungewissheitsintolerante im Fragebogen konsequenterweise *geringeres* Ungewissheitserleben angeben, um sich die als bedrohlich erlebten Ungewissheiten nicht eingestehen zu müssen. Unterschiedliche Umgangsweisen mit der individuellen Ungewissheitstoleranz mit jeweils konkurrierenden Einflüssen auf das Ungewissheitserleben mögen demnach bei der hier erfolgten Gruppenbetrachtung zu den uneindeutigen Korrelationen geführt haben.

Bei den Überlegungen über das zu postulierende Testmodell wurde auf Seite 88 die Frage der zu erwartenden Korrelationen zwischen den Skalen des QUICC erörtert. Wegen der wechselseitigen inhaltlichen Beeinflussung der Ungewissheitsqualitäten sowie deren gemeinsamer Abhängigkeit von Drittvariablen wie Krankheitsverlauf, Familienverhältnissen etc. war statistische Unabhängigkeit der Skalen nicht als Kriterium der Konstruktvalidität im Sinne differenzierbarer Ungewissheitsqualitäten definiert worden. Wirft man dennoch einen Blick auf die $\binom{8}{2} = 28$ Skaleninterkorrelationen, so stellt man außerordentlich hohe Werte fest, die zwischen 0,64 und 0,91 liegen (alle $p < 0,01$). Dies ist per se aus den genannten Gründen kein Manko des Fragebogens. Allerdings erscheinen dadurch die hohen Werte für Cronbachs α in etwas anderem Licht, da bei solch insgesamt hohen Korrelationen zwischen

[13] In der Tat gilt $Val_{max} = \sqrt{Rel}$.

den Items die Intraskalenkorrelationen weniger beeindruckend sind: Wenn bereits die Items unterschiedlicher Skalen hoch miteinander korrelieren – oder zumindest die aus ihnen gebildeten Mittelwerte – dann überraschen auch hohe Korrelationen innerhalb einer Skala nicht.

Zur Bewertung dieses Faktums stark korrelierender Ungewissheitsskalen seien erneut denkbare Ursachen betrachtet:

1. **Unzulänglichkeiten des Messinstruments.** Trotz sorgfältiger Entwicklung und umfangreicher Vortests wäre es möglich, dass der QUICC in seiner vorliegenden Form nicht in der Lage ist, Ungewissheitsqualitäten differenziert zu erfassen. So könnte etwa die Formulierung der Items uneindeutig sein. Oder die Bearbeitung des Bogens erfordert so viel Zeit und Konzentration, dass beim Ankreuzen nicht mehr zwischen den Items differenziert wird. Auch ein mangelndes oder beim Ausfüllen nachlassendes Verständnis der Fragebogeninstruktion könnte zu höherer Fehlerstreuung und damit zu weniger klar abgrenzbaren Skalen geführt haben.

2. **Besonderheiten der Stichprobe.** Die Wahl der Stichprobe beruhte auf der Vielfältigkeit von Krebsdiagnosen und Behandlungsphasen und -formen am Untersuchungsort sowie dem hohen Patientenaufkommen. Neben diesen Vorteilen könnten mit dem Setting der Strahlentherapeutischen Ambulanz aber auch gravierende Nachteile verbunden sein. So waren womöglich Patienten nicht willens oder fähig, die Items hinreichend spezifisch zu beantworten, wenn sie gerade auf die Ergebnisse ihrer jährlichen Rezidivkontrolle oder ähnlicher entscheidender Untersuchungen warteten. Zudem wäre es denkbar, dass die akute oder durch die Umgebung kognitiv und emotional aktualisierte Krebserkrankung es Patienten unmöglich machte, die QUICC-Items differenziert zu betrachten. Dann könnten die hohen Skaleninterkorrelationen als Besonderheit der Situation Krebskranker interpretiert werden, die womöglich auf den potentiell lebensbedrohlichen Charakter der Krankheit zurückzuführen ist.

3. **Ungewissheit ist unidimensional.** Andererseits könnte es auch sein, dass nicht nur Krebspatienten, sondern Patienten insgesamt nicht so deutlich zwischen Qualitäten von Ungewissheit unterscheiden, wie es in den vorangegangenen Ausführungen postuliert wurde. Zwar könnte es sein, dass ein Kategoriensystem wie das hier vorgestellte auch für Patienten theoretisch-inhaltlich

nachvollziehbar ist, dass aber dennoch im *Erleben* der Patienten nicht hinreichend trennbar ist, welche Ungewissheitsqualität mehr oder weniger stark ausgeprägt ist. In diesem Fall müsste man entgegen der bisherigen Annahmen von einem eher einheitlichen, ‚unidimensionalen' Ungewissheitserleben sprechen. Entsprechend wäre es kaum möglich, an einer veränderten Repräsentation von Ungewissheit den Effekt einer SDM-Intervention abzuleiten.

4. **Interkorrelationen markieren gleichzeitige Präsenz von Ungewissheiten.** Schließlich wäre es auch möglich, dass die vorliegenden hohen Skaleninterkorrelationen weder aus Unzulänglichkeiten auf empirischer noch auf theoretischer Ebene hervorgehen, sondern dass sie die realen Verhältnisse des Ungewissheitserlebens widerspiegeln. Wiederholt wurde in dieser Arbeit darauf hingewiesen, dass sich die erlebten Ungewissheitsqualitäten gegenseitig verstärken können und dass mehrere Qualitäten in ähnlicher Weise von externen Variablen modifiziert werden können. Resultat davon wäre, dass die verschiedenen Ungewissheitsqualitäten gleichzeitig und in ähnlicher Intensität – ob hoch oder niedrig – erlebt werden. Ein valides Messinstrument sollte dann bei jedem Patienten ähnlich hohe oder niedrige Werte auf den Ungewissheitsskalen messen. Daraus resultierten dann bei entsprechender Messgenauigkeit und ausreichender Varianz in den Daten hohe Skaleninterkorrelationen, ohne dass dies auf Undifferenzierbarkeit der erlebten Qualitäten hindeutete. Die Korrelationen wären schlicht Ausdruck gleichzeitiger Präsenz mehrerer Ungewissheitsqualitäten.[14]

   Diese Interpretation stünde darüber hinaus im Einklang mit dem postulierten Effekt von SDM auf das Ungewissheitserleben. SDM soll nach den Ausführungen in Abschnitt 4.3 zu einer erhöhten Varianz zwischen den erlebten Ungewissheitsqualitäten führen. Bei Patienten, die keine SDM-Intervention erfahren haben, sollte die Varianz zwischen den Qualitäten entsprechend gering sein. Demnach wären bei den hier untersuchten, nicht nach SDM-Prinzipien oder auch noch gar nicht behandelten Patienten hohe Skaleninterkorrelatio-

[14] Kriz (05.04.2006, pers. Mitteilung) hält für diese Erklärung eine einleuchtende Analogie bereit: Würde man die drei Variablen Länge, Breite und Höhe anhand einer Population von Postpaketen erheben, so ergäben sich vermutlich sehr hohe Korrelationen zwischen diesen Variablen. Eine exploratorische Faktorenanalyse würde sicherlich einen einzigen Generalfaktor zur Beschreibung der Daten nahe legen. Dennoch kann man ohne weiteres zwischen diesen Dimensionen differenzieren und sie *prinzipiell* weitgehend unabhängig voneinander manipulieren. Ihre Korrelationen sind einzig den typischen Proportionen von Postpaketen geschuldet.

nen keine Infragestellung, sondern vielmehr eine Bestätigung der inhaltlichen Validität des QUICC.

6.5.4 Zusammenfassung

Insgesamt kann man die Ergebnisse der ersten quantitativen Validierung des QUICC als vielversprechende Basis für weitere Entwicklungsschritte ansehen. Neben der weitgehenden Bestätigung der Erwartungen in Bezug auf Objektivität, Praktikabilität, Reliabilität und Validität bleiben einige Fragen offen. Dazu zählen insbesondere die diskutierten Hypothesen zur Entstehung der extrem hohen Skaleninterkorrelationen. Diese werden in weiteren Studien systematisch überprüft.

6.6 Validierungsstudie 2

Eine zweite Validierungsstudie dient der Klärung von zwei Fragen: 1. Lassen sich die erwartungskonformen Ergebnisse der ersten Studie replizieren? 2. Lassen sich jene unerwarteten Resultate durch Modifikationen am Fragebogen beseitigen? Zum zweiten Punkt wurden unter 6.5.3 verschiedene Hypothesen zur Interpretation der hohen Skaleninterkorrelationen des QUICC diskutiert. In der folgenden Studie wird als erster Schritt genauer geprüft, inwieweit mögliche Unzulänglichkeiten des QUICC selbst (vgl. Punkt 1 auf Seite 101) dazu beigetragen haben könnten.

6.6.1 Modifikation des QUICC

Als ein möglicher Schwachpunkt des QUICC war erwogen worden, dass etwa die Instruktion des QUICC nicht ausreichend verständlich ist bzw. im Laufe der Bearbeitung in Vergessenheit geraten könnte. Unter anderem nach Beratung mit dem Zentrum für Umfragen, Methoden und Analysen in Mannheim (ZUMA) wird zum einen das Beispielitem in der Instruktion ausgetauscht. Statt der womöglich zum Einstieg etwas verunsichernden Frage nach der Vertrauenswürdigkeit des Arztes steht stattdessen ein etwas ‚harmloser' anmutendes Item zum Umgang mit krankheitsbedingten Einschränkungen an erster Stelle (vgl. für diese und die folgenden Modifikationen die überarbeitete QUICC-Version in Anhang E).

Zudem wird das Itemformat leicht variiert, um die ‚Beschäftigung' auch optisch stärker zu betonen. Dagegen wird davon ausgegangen, dass der variable inhaltliche Teil der Items auch in kleinerem Druck von den Patienten beachtet wird, da ohne diesen keine Beantwortung möglich ist. Das bisherige und das überarbeitete Itemformat sind in Abbildung 6.4 gegenübergestellt. Dabei sei darauf hingewiesen, dass bei der aktuellen Version die Phrase „Mich beschäftigt zur Zeit die Frage ..." nur noch zwei mal pro Seite auftaucht, um sie durch diese Variation im Verlauf der Bearbeitung im Bewusstsein zu halten.

Zusätzlich werden die Teilnehmer am Fuß jeder Seite gebeten, jenes Item anzukreuzen, das sie am stärksten beschäftigt. Die Antwort auf diese Aufforderung wird nicht ausgewertet. Vielmehr stellt sie eine Möglichkeit dar, die Patienten kurz aus der etwas monotonen Beantwortung herauszuführen, um die Konzentration zu erhalten, sowie den Zielparameter des QUICC, nämlich die Beschäftigung, auf leicht veränderte Weise zu aktualisieren.

Um die abgestufte Einschätzung der Beschäftigung zu erleichtern, sind die anzukreuzenden Kästchen in der aktuellen Version durch nummerierte Kreise ersetzt. Dagegen sind nur noch die Extrempole jedes Items verbal beschriftet, um eine mögliche Verlockung zur Undifferenziertheit durch die ehemals mit „mittel" beschriftete Antwortkategorie zu vermeiden.

Schließlich wird der QUICC auch aufgrund der Patientenrückmeldungen im Umfang gekürzt, um auch am Ende des Bogens bei den Patienten ausreichende Konzentration gewährleisten zu können. Eine Itemselektion ist prinzipiell einerseits anhand statistischer Kennwerte wie Trennschärfe oder Itemschwierigkeit möglich, andererseits nach inhaltlichen Gesichtspunkten. Da keines der Items in der ersten Erhebung bezüglich der statistischen Parameter eindeutig inakzeptable Werte aufwies, wurden Items nach inhaltlichen Aspekten eliminiert. Auswahlkriterien waren Verständlichkeit, Eindeutigkeit, Prototypizität in Bezug auf die jeweilige Skala, Überlappungen mit anderen Items derselben Skala sowie Reaktionen der Patienten bei der vorherigen Erhebung. Als Ziel waren fünf Items für jede der acht Skalen angepeilt, wobei Skala 4 zunächst sechs Items behielt, da die Entscheidung zur Eliminierung eines weiteren Items auf Basis der bisherigen Informationen nicht getroffen werden konnte. Der QUICC wurde somit von ursprünglich 52 auf nunmehr 41 Items gekürzt. Damit geht gleichzeitig eine Erhöhung der Praktikabilität einher.

Was wohl in Fällen wie meinem die optimale Behandlung wäre?

Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr
	☐	☐	☐	☐	☐

Mich beschäftigt zur Zeit die Frage...					
	gar nicht				**sehr**
... , was wohl in Fällen wie meinem die optimale Behandlung wäre.	⓪	①	②	③	④

Abbildung 6.4: Das Antwortformat aus Validierungsstudie 1 (oben) und das modifizierte Format (unten).

Die Auswirkungen der Modifikationen am QUICC sollen einigermaßen isoliert betrachtet werden. Daher wird darüber hinaus nichts Wesentliches am Vorgehen gegenüber der Validierungsstudie 1 geändert. Das heißt insbesondere, dass die Stichprobe erneut aus derselben Population gezogen wird. Zudem wird eine vergleichbare Stichprobengröße angestrebt, da die zu untersuchenden Werte eng damit zusammenhängen.

Da die Modifikationen am QUICC nicht zuletzt dazu beitragen sollen, den Teilnehmern das intendierte Verständnis der Instruktion zu ermöglichen, wird für diese Untersuchung ein Kurzinterview eingeführt: Das Beispielitem aus der Einleitung beantworten die Patienten probehalber im Beisein des Versuchsleiters und begründen kurz ihre Einschätzung durch Verweis auf ihre Situation. Um sicherzustellen, dass die Instruktion komplett verstanden wird, werden die Teilnehmenden gebeten, die Lage eines fiktiven anderen Patienten zu skizzieren, der den entgegengesetzten Itempol ankreuzen würde. Eine solch gründliche Einführung ist sicherlich nur in Ausnahmefällen zu leisten. Dennoch wird dieses Vorgehen gewählt, um das größtmögliche Instruktionsverständnis zu erreichen und so diesen Aspekt als Fehlerquelle auszuschließen.

6.6.2 Annahmen über den QUICC

In dieser Studie stehen jene Parameter im Mittelpunkt der Betrachtungen, die durch die obigen Modifikationen betroffen sind. Wenn die Skaleninterkorrelationen wesentlich von den ab Seite 101 diskutierten möglichen Unzulänglichkeiten des QUICC abhingen, sollten sich nun Reduktionen der Korrelationen ergeben.

Außerdem gilt es zu beobachten, ob sich die bisher hohen Reliabilitätskennwerte durch die Reduktion des Itempools wesentlich verringern. Zwar kann Cronbachs α grundsätzlich auch anwachsen, wenn die Items auf diesen Aspekt hin eliminiert werden. Bei der hier erfolgten inhaltsbasierten Itemreduktion ist aber eher von leichten Verringerungen auszugehen, da sich Messfehler bei wenigen Items stärker auswirken. Für diese Studie werden wegen der sehr hohen Werte in der vorherigen Untersuchung dennoch erneut Cronbachs α-Werte von über 0,80 angestrebt.

Diese Argumentation gilt analog für die Trennschärfen, für die trotz der zu erwartenden Reduktion der Skalenreliabilitäten Indices oberhalb von 0,5 erwartet werden.

Da die Modifikationen keinen Effekt auf die Schwierigkeit der verbliebenen Items haben dürften, werden erneut Werte zwischen 1 und 3 angestrebt.

6.6.3 Stichprobe

In den Monaten Mai und Juni des Jahres 2005 wurden Patienten im Wartebereich der Strahlentherapeutischen Ambulanz des Universitätsklinikums Schleswig-Holstein, Campus Kiel, um die Teilnahme an der Untersuchung gebeten.

Von den 225 im Untersuchungszeitraum erschienenen Patienten wurden 190 um eine Teilnahme gebeten, die übrigen 35 mussten aus organisatorischen Gründen ausgelassen werden. 57 der angesprochenen 190 Patienten (30,0%) verweigerten direkt die Teilnahme aus persönlichen, terminlichen oder gesundheitlichen Gründen. Weitere 4 (2,1%) kamen ihrer Ankündigung nicht nach, den Fragebogen in einem bereits frankierten Umschlag zurückzusenden. Somit gingen zunächst 129 Patienten (66,3%) in die Untersuchung ein.

Aus dieser Gruppe wurden einerseits nachträglich 7 Teilnehmer ausgeschlossen, die etwa durch versehentliches Überblättern einer Seite zu viele fehlende Werte produziert hatten. Andererseits ließ es sich auch bei dieser Untersuchung nicht immer ver-

Bösartige Neubildungen betreffend ...	Anteil in Prozent
Brustdrüse	24,5 (43,3)
Männliche Geschlechtsorgane	12,3 (28,3)
Verdauungsorgane	15,1
Lymphatisches, blutbildendes und verwandtes Gewebe	12,3
Lippe, Mundhöhle, Pharynx	9,4
Weibliche Geschlechtsorgane	5,7 (10,0)

Tabelle 6.4: Die häufigsten Krebsarten unter den eingeschlossenen Teilnehmern. Gegebenenfalls sind Anteile im jeweils allein betroffenen Geschlecht in Klammern angegeben.

hindern, dass Patienten während des Ausfüllens zum Behandlungsgespräch gerufen wurden. Diese 16 Patienten wurden auch dann aus der Auswertung ausgeschlossen, wenn sie den Fragebogen anschließend zu Ende ausfüllten (vgl. Seite 92). Somit verblieben 106 verwertbare Fragebögen.

Von diesen 106 Patienten waren 43,4% männlich und 56,6% weiblich. Das Alter lag durchschnittlich bei 58,7 Jahren ($SD = 13,8$). Diese Kennwerte differieren kaum von denen der Gesamtstichprobe (Altersdurchschnitt 59,4 ($SD = 13,5$); 43,8% männlich und 56,2% weiblich).

Die häufigsten Krebserkrankungen waren bösartige Neubildungen der Brustdrüse (ICD-10 C50) mit einem Anteil von 24,5% (das entspricht 43,3% der eingeschlossenen Frauen) sowie bösartige Neubildungen der Verdauungsorgane (C15-C26). Die Häufigkeiten weiterer Diagnosen sind Tabelle 6.4 zu entnehmen.

6.6.4 Ergebnisse

Itemschwierigkeit. Die Schwierigkeitsindices hatten einen Median von 2,02. Am leichtesten war das Item 10 mit einem Itemmittelwert von 2,85, während Item 1 mit 1,02 am schwierigsten war. Eine Gesamtübersicht dieser und der nachfolgend berichteten Parameter ist Tabelle 6.5 zu entnehmen.

Trennschärfe. Die Trennschärfe variierte in dieser Stichprobe zwischen 0,32 bei Item 1 und 0,85 bei Item 34 um einen Median von 0,68.

Reliabilität. Der Median der Reliabilitäten der acht Skalen lag bei 0,85 (Cronbachs α). Den niedrigsten Wert lieferte die Skala Soziale Einbindung (Skala 1) mit 0,76, der höchste Wert stammte mit 0,91 von der Skala 5b, die Ungewissheiten in Bezug auf den Arzt erfragt.

Skala	Item	Schwierigkeit	Trennschärfe	Streuung	N	Reliabilität
1	1	1,05	0,32	1,22	103	0,76
	9	1,72	0,44	1,44	103	
	17	1,73	0,64	1,67	103	
	32	1,96	0,57	1,41	103	
	40	1,89	0,69	1,73	103	
2a	2	2,51	0,69	1,32	105	0,84
	10	2,85	0,67	1,34	105	
	18	2,31	0,62	1,46	105	
	31	2,18	0,63	1,41	105	
	39	2,57	0,61	1,43	105	
2b	6	1,71	0,72	1,46	104	0,87
	14	1,41	0,71	1,43	104	
	23	2,15	0,70	1,59	104	
	27	1,45	0,62	1,38	104	
	35	1,97	0,74	1,42	104	
3	4	2,18	0,63	1,43	104	0,83
	12	1,65	0,46	1,56	104	
	20	2,19	0,76	1,48	104	
	29	1,95	0,58	1,46	104	
	37	2,16	0,72	1,36	104	
4	5	1,87	0,56	1,59	103	0,83
	13	1,81	0,58	1,42	103	
	24	1,90	0,52	1,34	103	
	28	2,14	0,71	1,62	103	
	36	1,91	0,68	1,61	103	
	41	1,51	0,59	1,47	103	
5a	3	1,55	0,54	1,32	105	0,87
	11	1,68	0,67	1,46	105	
	19	2,19	0,75	1,48	105	
	30	2,02	0,78	1,49	105	
	38	2,43	0,77	1,59	105	
5b	7	2,12	0,80	1,60	104	0,91

Skala	Item	Schwierigkeit	Trennschärfe	Streuung	N	Reliabilität
	15	2,15	0,80	1,65	104	
	22	2,27	0,71	1,57	104	
	26	1,74	0,74	1,66	104	
	34	2,10	0,85	1,63	104	
6	8	2,37	0,69	1,57	105	0,85
	16	2,47	0,62	1,35	105	
	21	1,94	0,63	1,39	105	
	25	2,35	0,68	1,34	105	
	33	2,25	0,73	1,39	105	

Tabelle 6.5: Die Itemparameter Schwierigkeit, Trennschärfe und Streuung der 41 QUICC-Items sowie die Skalenreliabilitäten. Variierende Stichprobengrößen gehen auf einzelne fehlende Werte zurück.

Skaleninterkorrelation. Der Median der Skaleninterkorrelationen betrug 0,72 mit einer Spannweite zwischen 0,56 und 0,87. Damit lag der Median knapp unterhalb dessen aus der vorherigen Erhebung. Von den 28 Korrelationen waren 16 im Vergleich in dieser Studie niedriger, 12 höher. Laut gerichtetem Vorzeichentest ist diese Veränderung nicht signifikant ($p < 0,29$). Eine komplette Übersicht über die Interkorrelationen liefert Tabelle 6.6.

Skala	1	2a	2b	3	4	5a	5b	6
1		0,63	0,74	0,72	0,72	0,70	0,71	0,67
2a	0,68		0,69	0,81	0,69	0,67	0,66	0,76
2b	0,77	0,72		0,75	0,72	0,90	0,86	0,84
3	0,75	0,70	0,65		0,71	0,74	0,74	0,77
4	0,78	0,64	0,75	0,63		0,69	0,63	0,74
5a	0,76	0,72	0,87	0,60	0,66		0,88	0,86
5b	0,73	0,63	0,82	0,56	0,68	0,82		0,84
6	0,72	0,75	0,84	0,66	0,69	0,81	0,80	

Tabelle 6.6: Die Skaleninterkorrelationen. Oberhalb der Diagonalen sind die Werte aus Validierungsstudie 1 eingetragen, unterhalb jene der aktuellen Untersuchung.

6.6.5 Diskussion

Anhand dieser Untersuchung konnten die Ergebnisse der vorherigen, nach den zusätzlich erhobenen Daten als recht ähnlich anzusehenden Stichprobe repliziert werden, die den QUICC als ein hoch reliables und praktikables Messinstrument erscheinen lassen. Trotz der Reduktion des Itempools um über 20% wurden erneut sehr hohe Reliabilitäten erreicht, auch wenn jeder einzelne Wert merklich zurückgegangen ist und Skala 1 mit 0,79 knapp unter dem avisierten Minimum von 0,80 lag. Gleichzeitig konnte der QUICC dadurch erwartungsgemäß auch deutlich schneller von den Patienten ausgefüllt werden. Die durchschnittliche Bearbeitungszeit lag trotz des einführenden Kurzinterviews bei gut 20 Minuten. Der Austausch mit den Teilnehmenden während dieses Interviews lässt dabei den Schluss zu, dass die Instruktion von den meisten recht gut verstanden wurde. Tendenziell schienen eher ältere Menschen Schwierigkeiten beim Verständnis zu haben.

Wie erwartet reduzierten sich mit der geringeren Itemanzahl auch die Trennschärfekoeffizienten. Lag deren Median in der ersten Validierungsstudie noch bei 0,74, so betrug er in dieser nur noch 0,68. Allerdings ist festzuhalten, dass auch die aktuellen Koeffizienten noch vollends zufrieden stellend sind. Im Allgemeinen werden Trennschärfen oberhalb von 0,5 als hoch angesehen. Hier wird dieser Grenze nur von drei Items leicht unterschritten. Wie in der ersten Untersuchung fällt dabei das an erster Stelle stehende Item durch den geringsten Koeffizienten auf.[15] Offensichtlich müssen sich die Patienten trotz des einleitenden Beispiels erst noch mit dem QUICC vertraut machen. Allerdings ist selbst dieser Wert mit 0,32 als akzeptabel zu bezeichnen.

Ähnlich positiv fällt das Fazit zu den Itemschwierigkeiten aus. Die Itemmittelwerte lagen sämtlich im angestrebten Intervall von 1 bis 3, womit erneut die Ergebnisse der ersten Untersuchung bestätigt werden konnten. Das trifft auch auf die Itemstreuungen zu. Die Items des QUICC lassen sich demnach als angemessen schwierig und somit gut geeignet für diese Population bezeichnen.

Keinen nennenswerten Einfluss hatten die vielfältigen Modifikationen am QUICC offensichtlich auf die Skaleninterkorrelationen. Deren minimale durchschnittliche Verringerungen sind vermutlich eher zufälliger Natur und könnten darüber hinaus von

[15] Dabei handelt es sich jedoch nicht um dasselbe Item wie in Validierungsstudie 1.

den gesunkenen Reliabilitäten begünstigt worden sein. Damit lässt sich schlussfolgern, dass die Interkorrelationen in dieser und der vorangegangenen Studie nicht auf mangelndes Verständnis der Instruktion oder eine Überforderung durch den Umfang des Fragebogens zurückzuführen sind. Die ab Seite 101 diskutierten Zweifel an der Güte des Messinstruments erscheinen damit als weitgehend ausgeräumt. Es verbleiben demnach die ebenda erwogenen übrigen Interpretationen, die es in weiteren Untersuchungen zu prüfen gilt.

6.7 Validierungsstudie 3

Nach den Ergebnissen der Validierungsstudie 2 haben die Modifikationen am QUICC dessen Praktikabilität gesteigert, ohne gleichzeitig die Reliabilität bedenklich einzuschränken. Dennoch bleibt die Frage des Ursprungs der hohen Skaleninterkorrelationen durch die bisher getesteten Vermutungen aus Abschnitt 6.5.3 ungeklärt. Unter Punkt 2 wurde auf Seite 101 erwogen, ob die Interkorrelationen ein valider Ausdruck des Ungewissheitserlebens speziell von Krebspatienten sind, das z.B. auf den potentiell lebensbedrohlichen Charakter oder die für viele Menschen besondere Bedeutung der Diagnose Krebs zurückgehen könnte. Um diese Vermutung wiederum empirisch zu prüfen, wird der QUICC in einer multizentrischen Studie an verschiedenen nichtonkologischen Stichproben chronisch Kranker eingesetzt, um das Ungewissheitserleben Krebskranker mit dem von anderen chronisch Erkrankten zu vergleichen.

Dies ermöglicht über die Betrachtung der Skaleninterkorrelationen hinaus auch eine Bewertung der Einsetzbarkeit des QUICC außerhalb seines ursprünglichen Entwicklungsfeldes, was einen wesentlichen weiteren Schritt der Validierung darstellt (vgl. Abschnitt 6.3.1). Deshalb ist es in dieser Studie von besonderem Interesse, wie die in Abschnitt 6.3 begründeten und operationalisierten Testgütekriterien sowie die Schwierigkeits- und Trennschärfeparameter in den jeweiligen Teilstichproben ausfallen.

6.7.1 Stichprobe

Für eine indikationsübergreifende Betrachtung wurden mehrere Teilstichproben chronischer Erkrankungen untersucht.

Chronisch-entzündliche Darmkrankheiten. Die Subgruppe der Patienten mit einer chronisch-entzündlichen Darmkrankheit (CED) stammte zum einen Teil aus der CED-Ambulanz der Universitätsklinik Schleswig-Holstein, Campus Kiel. Der andere Teil wurde bei Selbsthilfegruppen in Hamburg, Ulm, München, Münster, Schwerin, Rostock, Wismar, Düsseldorf und Berlin gewonnen. Insgesamt wurden 187 Patienten in die Untersuchung eingeschlossen. Davon litten 86,1% an Morbus Crohn (ICD-10 K50) und 12,8% an Colitis Ulcerosa (K51).[16] Das Alter der Patienten variierte zwischen 18 und 72 Jahren bei einem Mittelwert von 38,7 Jahren ($SD = 11,5$). Das Geschlechterverhältnis betrug 40,1% Männer zu 58,3% Frauen.

Chronische Schmerzpatienten. An der Klinik für Neurologie des Universitätsklinikums Schleswig-Holstein, Campus Kiel, wurden 48 Patienten zur Teilnahme an der Untersuchung gewonnen, die an chronischen, neuropathischen Schmerzen litten. Dazu zählten in dieser Subgruppe Zentral bedingter Schmerz, Trigeminus-Neuralgie, Phantomschmerz, CRPS (Complex Regional Pain Syndrome), Posttraumatische Neuralgie, Postzosterische Neuralgie und Schmerzhafte Polyneuropathie. Das Alter der Patienten lag im Mittel bei 59,8 Jahren ($SD = 12,6$). Der jüngste Patient war 24, der älteste 85 Jahre alt. Von den Patienten waren 31,3% männlich und 62,5% weiblich.

Depression. Als dritte Subgruppe wurden am Zentrum für Integrative Psychiatrie in Kiel 45 stationär behandelte Patienten in die Studie aufgenommen, die an einer unipolaren depressiven Störung (ICD-10 F32-F34) litten. Deren Alter lag zwischen 19 und 80 Jahren bei einem Mittelwert von 46,9 Jahren ($SD = 15,5$). Der Anteil männlicher Teilnehmer betrug 37,8%, der der weiblichen belief sich auf 62,2%.

[16] An 100% fehlende Anteile gehen bei dieser und den folgenden Angaben auf fehlende Werte zurück.

Für keine dieser Subgruppen liegen Angaben darüber vor, wie viele der angesprochenen Patienten die Teilnahme verweigerten, da die Dokumentation dessen das mit der Rekrutierung befasste Personal zeitlich überfordert hätte.

6.7.2 Ergebnisse

Itemschwierigkeit. Tabelle 6.7 zeigt die Schwierigkeiten und Itemstreuungen der 41 QUICC-Items in den drei Teilstichproben sowie in der Gesamtstichprobe. Der Großteil der Schwierigkeitsindices bewegte sich in einem Korridor zwischen 1 und 3 um den Skalenmittelwert. In der Gesamtstichprobe lag der niedrigste Wert bei 1,24 (Item 26), der höchste bei 2,85 (Item 2).

In den Teilstichproben verließen einige Indices den genannten Mittelbereich. So wiesen bei den Depressionspatienten die Items 2, 4 und 25 Werte knapp über 3 auf, während Item 27 einen Mittelwert unter 1 lieferte.

Die Schmerzpatienten hatten vergleichsweise hohe Werte in der Behandlungsungewissheit (Items 6, 16, 25) sowie bei der Frage der weiteren Entwicklung ihrer Erkrankung (Item 10).

Dagegen zeichneten sich die CED-Patienten durch insgesamt recht moderate Schwierigkeitsindices aus, die zwischen 1,17 und 2,77 lagen.

Skala	Item	CED		Depression		Schmerz		Gesamt	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
1	1	1,68	1,35	1,95	1,34	1,06	1,24	1,64	1,35
	9	2,46	1,31	2,77	1,31	1,86	1,40	2,43	1,34
	17	1,98	1,49	1,88	1,71	1,43	1,48	1,88	1,53
	32	1,69	1,34	2,63	1,45	2,00	1,46	1,90	1,42
	40	1,77	1,54	2,30	1,52	1,97	1,42	1,90	1,53
2a	2	2,77	1,26	3,05	1,09	2,97	0,86	2,85	1,18
	10	2,43	1,38	2,84	1,19	3,31	0,90	2,63	1,33
	18	2,27	1,25	1,42	1,42	1,94	1,45	2,07	1,34
	31	2,70	1,21	2,19	1,28	2,60	1,31	2,59	1,24
	39	2,26	1,35	1,95	1,54	2,43	1,38	2,23	1,39
2b	6	1,36	1,35	1,72	1,37	1,91	1,60	1,51	1,40
	14	1,64	1,40	1,60	1,26	1,63	1,65	1,63	1,41
	23	2,10	1,23	1,72	1,22	2,29	1,49	2,06	1,27

Skala	Item	CED		Depression		Schmerz		Gesamt	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
	27	1,60	1,33	0,95	1,36	1,49	1,36	1,47	1,36
	35	1,41	1,28	1,86	1,37	2,03	1,40	1,58	1,33
3	4	2,42	1,40	3,07	1,16	2,69	0,99	2,58	1,33
	12	1,86	1,39	1,86	1,64	1,26	1,48	1,77	1,46
	20	2,12	1,44	2,88	1,18	2,60	1,24	2,33	1,40
	29	2,57	1,28	2,33	1,43	2,26	1,48	2,48	1,33
	37	2,40	1,31	2,37	1,25	2,54	1,20	2,41	1,28
4	5	1,88	1,47	1,49	1,52	1,74	1,60	1,79	1,50
	13	2,31	1,23	2,67	1,29	2,06	1,30	2,34	1,26
	24	1,91	1,44	2,88	1,28	2,60	1,33	2,18	1,45
	28	1,84	1,52	2,23	1,54	2,00	1,51	1,93	1,52
	36	1,51	1,38	2,02	1,65	1,37	1,46	1,58	1,46
	41	1,83	1,46	1,95	1,53	1,91	1,46	1,86	1,47
5a	3	2,02	1,31	1,93	1,40	2,11	1,32	2,02	1,32
	11	1,94	1,22	1,56	1,42	1,40	1,48	1,79	1,31
	19	2,40	1,26	2,12	1,45	2,14	1,29	2,31	1,30
	30	1,85	1,29	1,88	1,33	1,97	1,44	1,87	1,31
	38	1,62	1,35	2,19	1,26	2,69	1,41	1,88	1,40
5b	7	1,79	1,30	2,02	1,41	2,17	1,52	1,89	1,36
	15	2,01	1,43	1,72	1,33	1,89	1,60	1,94	1,44
	22	1,84	1,29	1,19	1,37	2,51	1,50	1,82	1,38
	26	1,17	1,33	1,33	1,41	1,46	1,34	1,24	1,34
	34	1,49	1,40	1,42	1,45	2,00	1,41	1,55	1,42
6	8	2,20	1,57	2,53	1,40	3,11	1,23	2,39	1,53
	16	1,91	1,52	2,56	1,26	3,03	0,98	2,19	1,46
	21	1,56	1,43	1,72	1,47	2,03	1,44	1,65	1,44
	25	2,60	1,27	3,16	1,02	3,03	1,04	2,76	1,22
	33	2,07	1,38	1,74	1,29	2,46	1,40	2,07	1,38

Tabelle 6.7: Die Itemschwierigkeiten und Streuungen der 41 QUICC-Items in den verschiedenen Teilstichproben sowie in der Geamtpopulation.

Trennschärfe und Reliabilität. Die Trennschärfen der Items sowie die Reliabilitätskennwerte der Skalen in den einzelnen Teilstichproben können Tabelle 6.8 ent-

nommen werden.

In der CED-Gruppe lagen die korrigierten Trennschärfeindices zwischen 0,32 und 0,80 bei einem Median von 0,57. Die Reliabilitäten, operationalisiert als Cronbachs α, erreichten einen Median von 0,79 bei Schwankungen zwischen 0,74 (Skala 2a) und 0,86 (Skala 6).

Die Teilstichprobe der Depressiven Patienten wies Trennschärfen zwischen 0,28 und 0,79 auf, wobei der Median 0,60 betrug. Die Reliabilitätskennwerte reichten von 0,67 (Skala 2a) bis 0,86 (Skala 6) bei einem Median von 0,80.

Bei den Schmerzpatienten variierten die Trennschärfen zwischen 0,27 und 0,76. Der Median lag bei 0,56. Die mediane Reliabilität betrug 0,78, wobei Skala 2a mit 0,68 den niedrigsten Wert lieferte. Das Maximum bildete Skala 5b mit 0,84.

Betrachtet man alle Teilstichproben zusammen, so zeigten sich Trennschärfen zwischen 0,33 und 0,78 mit einem Median von 0,57. Cronbachs α erreichte den höchsten Wert bei Skala 6 mit 0,86, den niedrigsten bei Skala 2a mit 0,71. Der Median betrug 0,80.

Skala	Item	CED	Depression	Schmerz	Gesamt
1	1	0,38	0,39	0,27	0,37
	9	0,56	0,69	0,45	0,57
	17	0,63	0,65	0,46	0,59
	32	0,54	0,53	0,56	0,53
	40	0,68	0,65	0,64	0,66
	α	0,78	0,80	0,72	0,77
2a	2	0,54	0,31	0,55	0,48
	10	0,44	0,64	0,32	0,42
	18	0,66	0,41	0,55	0,57
	31	0,54	0,42	0,56	0,52
	39	0,37	0,40	0,33	0,37
	α	0,74	0,67	0,68	0,71
2b	6	0,50	0,50	0,61	0,51
	14	0,72	0,79	0,51	0,68
	23	0,51	0,60	0,52	0,52
	27	0,69	0,51	0,45	0,60
	35	0,68	0,65	0,61	0,65
	α	0,82	0,82	0,77	0,81

Skala	Item	CED	Depression	Schmerz	Gesamt
3	4	0,55	0,38	0,73	0,54
	12	0,54	0,34	0,48	0,48
	20	0,69	0,60	0,62	0,65
	29	0,65	0,55	0,64	0,61
	37	0,71	0,58	0,56	0,66
	α	0,83	0,72	0,81	0,80
4	5	0,32	0,46	0,40	0,33
	13	0,65	0,54	0,69	0,63
	24	0,41	0,68	0,42	0,45
	28	0,69	0,60	0,73	0,68
	36	0,72	0,64	0,76	0,71
	41	0,44	0,54	0,39	0,45
	α	0,78	0,81	0,79	0,79
5a	3	0,57	0,41	0,42	0,51
	11	0,55	0,44	0,51	0,50
	19	0,53	0,51	0,75	0,55
	30	0,63	0,28	0,75	0,59
	38	0,46	0,66	0,66	0,49
	α	0,77	0,70	0,82	0,76
5b	7	0,50	0,49	0,67	0,53
	15	0,43	0,71	0,75	0,53
	22	0,52	0,67	0,58	0,56
	26	0,65	0,65	0,57	0,63
	34	0,72	0,67	0,62	0,70
	α	0,79	0,83	0,84	0,80
6	8	0,73	0,77	0,52	0,73
	16	0,80	0,77	0,50	0,78
	21	0,64	0,60	0,56	0,63
	25	0,65	0,68	0,32	0,62
	33	0,59	0,64	0,53	0,58
	α	0,86	0,86	0,72	0,85

Tabelle 6.8: Die korrigierten Trennschärfen der 41 QUICC-Items sowie die Reliabilitätskoeffizienten der Skalen in den verschiedenen Teilstichproben.

Skala	1	2a	2b	3	4	5a	5b	6
1		0,64	0,68	0,65	0,77	0,74	0,62	0,49
2a	0,76		0,67	0,70	0,69	0,64	0,57	0,66
2b	0,53	0,69		0,59	0,76	0,80	0,65	0,46
3	0,73	0,72	0,60		0,70	0,68	0,66	0,68
4	0,64	0,69	0,65	0,79		0,79	0,63	0,47
5a	0,57	0,65	0,78	0,57	0,68		0,71	0,54
5b	0,66	0,71	0,75	0,47	0,44	0,73		0,70
6	0,52	0,73	0,76	0,59	0,61	0,70	0,69	
1		0,61	0,59	0,59	0,57	0,62	0,54	0,48
2a	0,62		0,67	0,42	0,38	0,74	0,59	0,66
2b	0,62	0,68		0,36	0,53	0,86	0,82	0,72
3	0,66	0,65	0,55		0,21	0,36	0,34	0,31
4	0,73	0,62	0,69	0,64		0,48	0,45	0,39
5a	0,67	0,67	0,80	0,60	0,70		0,84	0,84
5b	0,58	0,62	0,71	0,56	0,57	0,75		0,74
6	0,45	0,67	0,54	0,62	0,49	0,60	0,70	

Tabelle 6.9: Die Interkorrelationen der Ungewissheitsskalen. Im oberen Teil der Tabelle sind oberhalb der Diagonalen die Werte aus der CED-Teilstichprobe eingetragen, unterhalb diejenigen der Depressionsgruppe. Im unteren Teil der Tabelle stehen oberhalb der Diagonalen die Koeffizienten der Schmerzpatienten, während unterhalb jene der Gesamtstichprobe zu sehen sind.

Skaleninterkorrelation. Die Korrelationen der QUICC-Skalen untereinander sind Tabelle 6.9 zu entnehmen.

In der CED-Gruppe lagen die Koeffizienten zwischen 0,46 und 0,80 mit einem Median von 0,66. Die Subgruppe der Depressiven wies Werte zwischen 0,44 und 0,79 auf. Dort betrug der Median 0,68. Bei den Schmerzpatienten streuten die Korrelationen von 0,21 bis 0,86 um den Median von 0,58. Wurden alle Subgruppen aggregiert, so ergaben sich Koeffizienten zwischen 0,45 und 0,80 mit einem Median von 0,63.

6.7.3 Diskussion

In der vorliegenden Studie wurde der QUICC erstmals außerhalb des ursprünglichen, onkologischen Entwicklungskontextes eingesetzt. Auch in diesen teilweise grundlegend anderen Patientenkollektiven hat er sich als reliables Messinstrument erwiesen. Das belegen zum einen die akzeptablen bis hohen Werte für Cronbachs α,

zum anderen auch die berichteten Trennschärfekoeffizienten, die fast ausschließlich oberhalb von 0,3 liegen. Dabei ist weiterhin festzustellen, dass nur bei Item 1 die Trennschärfen in allen drei Stichproben kleiner als 0,4 sind, was womöglich auch ein Positionseffekt sein könnte. So war bereits in der ersten Validierungsstudie das in der damaligen Version erste Item durch die geringste Trennschärfe ausgefallen. Dies wiederum lässt vermuten, dass die geringeren Werte bei manchen Items entweder zufällige Schwankungen sind oder aber auf Besonderheiten eben jener Stichprobe zurückgehen und somit keine grundsätzliche Schwäche einzelner Items markieren.

Die berichteten Werte sind umso überzeugender, wenn man bedenkt, dass sie teilweise an Subgruppen von unter 50 Personen gewonnen wurden. Bei der den Stichprobenumfängen aus den ersten beiden Studien vergleichbaren CED-Gruppe etwa liegen die Reliabilitäten und Trennschärfen dennoch unterhalb der damaligen Werte (vgl. Abschnitte 6.5.2 und 6.6.4). Insgesamt sind solcherart Vergleiche jedoch mit großer Vorsicht anzustellen. Denn die Homogenität der Skalen hängt nicht nur von der Indikation ab, sondern auch in hohem Maße von anderen Faktoren, z.B. der jeweiligen Krankheitsphase. So mögen etwa die Items 2 („..., wie mein aktueller Krankheitszustand ist“) und 10 („..., wie sich meine Krankheit weiter entwickeln wird“) für eine Krebspatientin mit ganz frischer Diagnose inhaltlich sehr eng miteinander verwoben sein. Für jemanden mit gerade erfolgter Nachsorgeuntersuchung ohne jeglichen Befund ist dagegen die erste Frage nicht relevant, die zweite sehr wohl. Da in dieser multizentrischen Studie kaum zusätzliche Variablen erhoben wurden, lässt sich dieser Einfluss demnach nicht kontrollieren.

Zusätzlich kann allerdings auch in einer ähnlichen Krankheitsphase bei einer anderen Indikation die Homogenität unterschiedlich ausfallen. Im Vergleich mit besagter Krebspatientin könnte auch ein jüngst erkrankter Depressiver die Frage nach dem aktuellen Krankheitszustand (Item 2) weniger beschäftigen, weil er bei seiner Krankheit einen etwas verlässlicheren subjektiven Eindruck seines Befindens hat; während er recht genau weiß, wie schlecht es ihm geht, fehlt der Krebskranken ein ‚Rezeptor‘ für das Wachstum entarteter Zellen. Gleichwohl ist auch für den Depressiven die Frage nach der zukünftigen Krankheitsentwicklung (Item 10) höchst relevant. Während also für Krebspatienten beide Items meist zwei Seiten derselben Medaille darstellen, muss das bei Depressiven nicht unbedingt der Fall sein. Diese Überlegungen enthüllen die grundsätzlichen Schwierigkeiten bei der universellen Verwendung eines einzigen Fragebogens in verschiedensten Situationen.

Gleiches gilt für die Bewertung der Itemschwierigkeiten. Auch wenn mittlere Itemschwierigkeiten als Ziel angepeilt werden, so ist zu akzeptieren, dass diese nicht in jeder Stichprobe übereinstimmen. Vielmehr *sollen* Items ja gerade von Menschen in unterschiedlichen Situationen auch als unterschiedlich schwer wahrgenommen und entsprechend unterschiedlich beantwortet werden. Da die Itemmittelwerte aber dennoch in allen Teilstichproben weitgehend innerhalb des angepeilten Korridors zwischen 1 und 3 lagen, spricht auch von dieser Seite her nichts gegen den Einsatz des QUICC in diesen Populationen.

Insgesamt betrachtet lässt sich der QUICC also grundsätzlich gut in den untersuchten Indikationen anwenden. Allerdings sind die erwähnten Unterschiede in weiteren Studien auch im Hinblick auf weitere Variablen genauer zu prüfen. Insbesondere wäre dabei auch der Mangel an Informationen über Anzahl und Beweggründe nicht teilnehmender Patienten zu beheben, um die Repräsentativität der Daten abschätzen zu können.

Neben der Frage der Einsetzbarkeit des QUICC außerhalb der Onkologie diente diese Untersuchung insbesondere der Klärung der Frage, ob die bisherigen hohen Skaleninterkorrelationen eine Besonderheit der Indikation Krebs mit ihrem potentiell lebensbedrohlichen Charakter sind. Diese Frage ist klar zu verneinen. Zwar lagen die Interkorrelationen aus Tabelle 6.9 größtenteils unterhalb der in den Abschnitten 6.5.2 und 6.6.4 berichteten Werte, allerdings gehen diese graduellen Unterschiede wohl eher auf die ebenfalls niedrigeren Reliabilitäten der Skalen sowie die geringeren Umfänge der Teilstichproben zurück. Ein wesentlicher *qualitativer* Unterschied zwischen den Indikationen ist dagegen nicht auszumachen.

Diese Erkenntnis führt zurück auf die in Abschnitt 6.5.3 diskutierten möglichen Ursachen der hohen Skaleninterkorrelationen. Nachdem die erwogenen Unzulänglichkeiten des Messinstruments sowie die Indikation Krebs in dieser und der vorangegangenen Studie als Ursachen ausgeschlossen werden konnten, ist noch die mögliche Erklärung auszuräumen, dass Ungewissheit entgegen der eingangs entwickelten Theorie nicht sinnvoll als ein mehrdimensional erlebtes Konstrukt aufgefasst werden kann. Dazu dient die folgende Studie.

6.8 Validierungsstudie 4

Im Rahmen dieser Arbeit wurde zunächst theoretisch begründet, warum es sinnvoll sein kann, verschiedene Qualitäten von Ungewissheit zu unterscheiden. Auf dieser Annahme aufbauend wurde ein Kategoriensystem entwickelt, das acht einander beeinflussende, aber dennoch inhaltlich trennbare Ungewissheitskategorien präzisierte. Die daraus abgeleiteten Skalen korrelierten in verschiedenen Stichproben sehr hoch miteinander. Dies wird konventionell meist als Beleg dafür gewertet, dass das untersuchte Merkmal als eindimensional aufzufassen ist (vgl. etwa Bühner, 2005). Auch wenn auf Seite 102 unter Punkt 4 bereits inhaltlich begründet wurde, warum dieser Schluss nicht alternativlos ist, soll dem Einwand in dieser Studie empirisch nachgegangen werden.

Die Datenbasis aus den vorangegangenen Untersuchungen ermöglicht ein eindimensionales Ungewissheitsmodell, der Grounded Theory-Prozess legt ein achtdimensionales Modell nahe. Es bedarf also eines geeigneten statistischen Verfahrens, um diesen Widerspruch entscheiden zu helfen. Dabei handelt es sich um eine Frage der Konstruktvalidität. Auf Seite 84ff wurde ausgeführt, aus welchen Gründen sich hier ein Modelltest nach der probabilistischen Testtheorie anbietet. Die beiden zu vergleichenden Modelle werden als Ratingskalenmodelle konzipiert, d.h. es werden für alle Items einer Skala identische Antwortschwellenabstände festgelegt (vgl. ebd.).

6.8.1 Herleitung der Methode

In dieser Studie wird zur Berechnung das Programm ConQuest (Wu et al., 2003) eingesetzt. Es eignet sich nach Rost (2004) gerade bei höherdimensionalen Modellen besser als etwa MULTIRA (Carstensen und Rost, 2003), da es für die postulierten Dimensionen eine gemeinsame Normalverteilung annimmt, was die Anzahl der zu schätzenden Parameter drastisch reduziert. Zwar gibt man damit die Möglichkeit der Berechnung verteilungsfreier Datensätze preis. Allerdings konnte Walter (2005) in Simulationen belegen, dass die Ergebnisse von ConQuest recht robust gegenüber Verteilungsverletzungen sind. Daher erfolgt hier keine weitere Prüfung dieser Annahme, zumal kein geeigneter Test zur Verfügung steht (vgl. die analoge Argumentation auf Seite 90). Der Vergleich des Modellfits erfolgt anhand der Informationskriterien AIC und BIC (vgl. Seite 86).

6.8.2 Stichprobe

Als Analysestichprobe dienen die aggregierten Daten der Validierungsstudien 1 bis 3 mit den Indikationen Krebs, chronisch-entzündliche Darmkrankheiten, chronischer Schmerz und Depression. Details zu diesen Teilstichproben sind in den Abschnitten 6.6.3 und 6.7.1 nachzulesen. Daraus ergibt sich ein Stichprobenumfang von $N = 489$.[17]

6.8.3 Ergebnisse

Die Anpassung des eindimensionalen Modells konvergierte nach 84 Iterationen bei einem Konvergenzkriterium von 0,001. Mit den in diesem Modell geschätzten 45 Parametern resultierten ein AIC von 54247,42 und ein BIC von 54436,08.

Das achtdimensionale Modell konvergierte nach 22 Iterationen. Bei einer Parameteranzahl von 80 ergaben sich Werte von 53259,28 für den AIC und 53594,67 für den BIC.

6.8.4 Diskussion

Die Informationskriterien AIC und BIC unterscheiden sich jeweils deutlich zwischen dem eindimensionalen und dem achtdimensionalen Modell. Da bei diesen Indices kleinere Werte eine bessere Modellanpassung signalisieren, kann man als Schlussfolgerung festhalten, dass die postulierten acht Dimensionen von Ungewissheit die Daten nicht nur besser erklären, sondern dass es sich auch lohnt, die höhere Komplexität dieser Konzeption in Kauf zu nehmen. Wie auf Seite 86 erläutert wurde, steht zwar kein Mittel zur Verfügung, um die AIC- und BIC-Werte inferenzstatistisch zu vergleichen. Auch aus anderen Untersuchungen kann man keinen Hinweis ableiten, ob Unterschiede von jeweils grob 1000 Punkten als bedeutsam einzustufen sind. Diese Differenzen besagen allerdings, dass das achtdimensionale Modell auch dann noch besser abgeschnitten hätte, wenn statt der 80 Parameter bis zu 504 (BIC) bzw. 574 (AIC) hätten investiert werden müssen, wie man sich durch

[17] Bei den Fragebögen aus Validierungsstudie 1 wurden nur die in den späteren Versionen enthaltenen 41 Items verwendet. Die zwischenzeitlich erfolgten übrigen Modifikationen werden nach den berichteten Ergebnissen nicht als so folgenreich angesehen, dass sie einer Aggregation im Wege stünden.

Umformung der Gleichungen leicht ausrechnen kann. In Anbetracht dessen erscheint das klare Votum für das theoriekonforme Modell gerechtfertigt.

Wie ist diese scheinbare Diskrepanz zwischen den Unidimensionalität suggerierenden Skaleninterkorrelationen einerseits und dem deutlichen Beleg für Multidimensionalität im ConQuest-Modell andererseits zu erklären? Zum einen ist anzumerken, dass bei der Betrachtung von Korrelationen der Skalenmittelwerte die Verhältnisse *innerhalb* der Skalen unberücksichtigt bleiben. So müsste man die Korrelationen für eine ausgewogene Bewertung etwa mit den internen Konsistenzen der Skalen in Beziehung setzen; dafür steht jedoch keine allgemeingültige Vergleichsformel oder ähnliches zur Verfügung, zumal es sich im einen Fall um Korrelationen zwischen Skalen, im anderen um solche zwischen Items handelt. Die Skaleninterkorrelationen dienen auch aus diesem Grund also nur als Hinweis auf *mögliche* Unidimensionalität.

Zum anderen wird im ConQuest-Modell die Passung eines Modells nicht auf Basis von Korrelationen beurteilt, wie es etwa bei der Faktorenanalyse erfolgt. Vielmehr werden hier die Schwellen- und Trennschärfeparameter geschätzt. Nach den Informationskriterien lässt sich der vorliegende Datensatz also dadurch deutlich besser reproduzieren, dass für die Items verschiedener Skalen auch unterschiedliche Itemparameter geschätzt werden, während sie innerhalb einer Skala übereinstimmen. Diese Ähnlichkeit innerhalb und Verschiedenheit außerhalb der Skalen bildet dabei die Basis der Entscheidung für das postulierte multidimensionale Modell.

Die Überlegenheit des theoriekonformen Modells geht indes nicht nur auf die erwartete Unzulänglichkeit der eindimensionalen Variante zurück. In den Abschnitten 5.1.3 und 5.1.4 wurde ein sechsdimensionales Modell diskutiert, das aus der Zusammenlegung der Kategorien 2a und 2b sowie 5a und 5b hervorgeht. Dieses inhaltlich nicht gänzlich unplausible Modell ist zwar nach den Informationskriterien ebenfalls dem eindimensionalen Pendant überlegen, erreicht aber auch nicht die Anpassungsgüte des favorisierten QUICC-Modells ($AIC = 53652,51$; $BIC = 53925,01$; 65 Parameter). Damit wird die ebd. inhaltlich begründete Bevorzugung der Lösung mit acht Kategorien auch empirisch untermauert.

Dass die eingangs begründete Annahme eines Ratingskalenmodells nicht nur theoretisch, sondern auch statistisch gerechtfertigt ist, zeigt ein Vergleich mit einem achtdimensionalen ordinalen Raschmodell, das für jedes Item individuelle Schwellenabstände zulässt. Mit einem AIC von 54257,38 bzw. einem BIC von 54580,19

ist letzteres nicht nur dem achtdimensionalen, sondern auch dem eindimensionalen Ratingskalenmodell unterlegen.

Man könnte in der hier erfolgten Zusammenlegung von Stichproben unterschiedlicher Provenienz eine Einschränkung der Aussagekraft vermuten. Dem ist einerseits entgegenzuhalten, dass die Definition des Anwendungskontextes des QUICC als „chronische Krankheiten“ per definitionem eine große Heterogenität einkalkuliert, der sich der QUICC zu stellen hat. Andererseits kann man gerade die damit einhergehende Erhöhung der Varianz in den Daten bzw. der Vielzahl unterschiedlicher Antwortpattern als Stärkung der Aussagekraft begreifen. Daneben ist nicht einsichtig, inwiefern die Datenaggregation gerade das hier favorisierte achtdimensionale Modell begünstigt haben sollte.

6.9 Validierungsstudie 5

In den vorangegangenen Validierungsstudien wurde anhand gängiger Testgütekriterien nachgewiesen, dass der QUICC ‚in sich‘ ein präzises Messinstrument ist. Diese Beurteilung ist aber noch in einem wesentlichen Punkt unvollständig. Der QUICC ist dazu entwickelt worden, die postulierten Effekte von SDM nachzuweisen. Gemäß den Ausführungen in Abschnitt 4.3 führt SDM dazu, dass Patienten aus einem Zustand indifferent erlebter Ungewissheiten herausgeführt werden in einen Zustand, in dem manche Ungewissheiten reduziert, andere hingegen bestehen geblieben oder gar angewachsen sind – etwa dadurch, dass Patienten das Fehlen von verlässlicher Evidenz zu ihrer Behandlung realisieren. Diese Differenzierung zwischen verschiedenen Ungewissheitsqualitäten war auf Seite 88 operationalisiert worden als ein Zuwachs an Varianz zwischen den Mittelwerten der Ungewissheitsskalen.

Um diesen Effekt nachzuweisen ist ein Studiendesign mit mindestens zwei Messzeitpunkten erforderlich, denn die abhängige Variable ist weniger absolut zu beurteilen, sondern eher relativ zum Ausgangszustand. Schließlich ist die Varianz wesentlich von der untersuchten Stichprobe und der zugrunde liegenden Indikation, dem Schweregrad etc. abhängig. Zusätzlich bedarf es zum Nachweis einer SDM-Intervention, die den unter 1.4 aufgeführten Merkmalen entspricht. In Validierungsstudie 3 war ein Kollektiv von depressiven Patienten vorgestellt worden, die zu Beginn ihrer stationären Behandlung im Zentrum für Integrative Psychiatrie in Kiel einen QUICC

ausfüllten. Zum Behandlungsstandard in dieser Klinik gehört der Besuch einer Psychoedukationsgruppe, deren erklärtes Ziel darin besteht, den Patienten zum Experten seiner eigenen Erkrankung zu machen, um ihn so als Partner in den Behandlungsprozess einbinden zu können. Man bezeichnet dies gemeinhin als ‚Patientenempowerment'. Zwar handelt es sich hierbei nicht um eine klassische SDM-Intervention, die *vor* der initialen Entscheidung für oder gegen eine bestimmte Therapie erfolgt und in eine solche von Arzt und Patient geteilte Entscheidung mündet (vgl. die Definition von Charles, 1997). Dennoch ist die Zielsetzung dieser Gruppe erklärtermaßen identisch mit der des SDM. Entsprechend soll die Edukationsgruppe zu einer fortwährenden Annäherung an das SDM-Ideal führen. Diese Operationalisierung von SDM erscheint auch insofern gerechtfertigt, als in dieser Arbeit wiederholt darauf verwiesen wurde, dass SDM nicht in phänomenologischen Aspekten wie etwa der entscheidungsbezogenen Patientenbeteiligung an sich besteht (vgl. Seite 35), sondern im Kern als eine Kommunikation über Ungewissheiten aufzufassen ist. Und eben dies geschieht in besagter Psychoedukationsgruppe, in der von fachlicher und von Patientenseite Fragen zur Erkrankung und deren Behandlungsoptionen erörtert werden. Man kann daher durchaus von einer gruppenbasierten Risikokommunikation zum Zwecke geteilter Entscheidungsfindung sprechen.

Neben der Hypothese über die Erhöhung der Varianz zwischen den Skalen ist in diesem Setting noch ein weiterer Effekt zu erwarten. In Validierungsstudie 1 war gezeigt worden, wie sich durch zusätzliche Informationen im Laufe einer Erkrankung viele Ungewissheiten reduzieren lassen (vgl. Seite 98). Auch vor und während der Laufzeit der Depressionsgruppe machen die Patienten Erfahrungen mit den Wirkungen und Nebenwirkungen der Behandlung (Kategorie 6), mit der Verlässlichkeit ihrer Partner und Familien (Kategorie 1), der Verlässlichkeit des Bahandlungspersonals (Kategorie 5b) etc. Daher ist davon auszugehen, dass viele Fragen, mit denen sich Patienten zu Beginn der Gruppenphase beschäftigten, nach deren Abschluss geklärt sind. Dies geht teilweise kausal auf die Inhalte des Gruppenprogramms selbst zurück, teilweise hat es externe Ursachen wie die eben beschriebenen. Zusätzlich ist gerade beim Krankheitsbild Depression bei erfolgreich einsetzender Therapie mit einer Abnahme der Rumination zu rechnen. Fasst man diese unterschiedlichen Einflüsse zusammen, so ist anzunehmen, dass sich beim zweiten Messzeitpunkt die mit dem QUICC erhobenen Ungewissheiten insgesamt reduziert haben werden.

6.9.1 Stichprobe

Von den 45 Patienten mit einer unipolaren depressiven Störung (ICD-10 F32-F34) aus Validierungsstudie 3 füllten 29 nach Abschluss der Psychoedukationsgruppe ein zweites Mal den QUICC aus. Davon waren 7 männlich (24,1%) und 22 weiblich (75,9%). Das Alter der Patienten lag zwischen 22 und 80 Jahren mit einem Mittelwert von 46,2 ($SD = 17,2$). Die übrigen 16 Patienten konnten nicht eingeschlossen werden, da sie die Edukationsgruppe zwecks Entlassung vorzeitig abbrachen oder gar nicht erst begannen.

6.9.2 Ergebnisse

Reduktion der Ungewissheiten. Zum ersten Messzeitpunkt betrug der Median der Skalenmittelwerte 2,16. Den höchsten Wert lieferte die Skala 6 mit einem Mittelwert von 2,43. Am niedrigsten war der Wert bei Skala 2b mit 1,56. Zum zweiten Messzeitpunkt lag der Median der Skalenmittelwerte bei 1,75. Hier wies die Skala 4 mit 2,14 den höchsten Wert auf, während mit 1,27 erneut bei Skala 2b der niedrigste Mittelwert zu verzeichnen war. Damit reduzierten sich alle Ungewissheitsqualitäten signifikant während der Gruppenphase ($p < 0,05$). Einzig bei Skala 2b war lediglich ein Trend zu verzeichnen ($p = 0,07$). Die Ergebnisse zeigt Tabelle 6.10 im Detail.

Skala	$M_{prä}$	M_{post}	d	t	df	p
1	2,14	1,71	0,60	3,21	28	0,01
2a	1,83	1,34	0,61	3,26	28	0,01
2b	1,56	1,27	0,28	1,50	28	0,07
3	2,24	1,82	0,45	2,44	28	0,01
4	2,39	2,14	0,38	2,05	28	0,02
5a	2,10	1,78	0,33	1,76	28	0,04
5b	2,19	1,63	0,56	3,02	28	0,01
6	2,43	1,97	0,61	3,29	28	0,01

Tabelle 6.10: Mittlere Ausprägung der Ungewissheitsqualitäten vor und nach der Psychoedukationsgruppe ($N = 29$).

Differenzierung der Ungewissheiten. In Anbetracht der eben berichteten Mittelwertsunterschiede zwischen den Messzeitpunkten erfordert die Untersuchung der

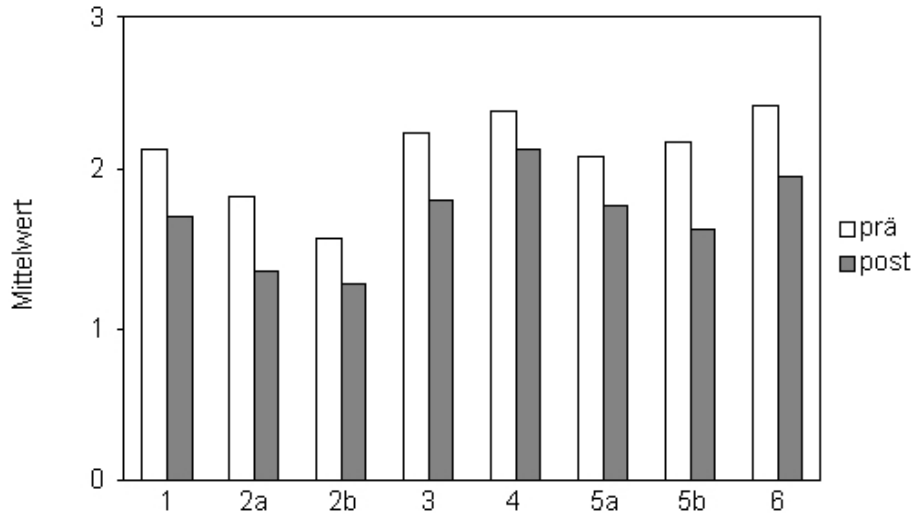

Abbildung 6.5: Durchschnittliche Ausprägung der Ungewissheitsqualitäten vor und nach der Gruppenintervention ($N = 29$).

Varianz zwischen den Skalen zunächst eine geeignete Transformation, die diese Unterschiede nivelliert. Schließlich ist eine Varianz von 0,8 bedeutsamer, wenn sie bei einem Mittelwert von 1 auftritt statt bei einem Mittelwert von 3. Um nicht gleichzeitig auch die Varianz der Daten zu verändern, wurde statt üblicher Methoden wie etwa der z-Transformation folgendes Verfahren verwendet: Separat für jeden Messzeitpunkt wurden die acht Skalenmittelwerte jedes Patienten durch dessen Mittelwert aus allen Skalen (‚individual grand mean') geteilt. Gegenüber einer gruppenbasierten z-Transformation wurden die Daten also stattdessen spezifisch für jeden Patienten reskaliert. Nach dieser Stauchung wurden alle Werte durch Subtraktion von 1 auf einen Mittelwert von 0 verschoben. Positive Werte zeigen somit Abweichungen vom individuellen Ungewissheitsdurchschnitt nach oben an, während negative Werte auf unterdurchschnittlich ausgeprägte Ungewissheitsqualitäten hinweisen.

Neben der Vergleichbarkeit der Varianzen vor und nach der Intervention ermöglicht es die beschriebene Reskalierung zudem, die quadrierten Abweichungen vom individuellen Gesamtmittelwert zu betrachten. Diese sind graphisch anschaulicher als Varianzen, da man sie separat für jede Skala berechnen kann. Die Quadrierung erlaubt dabei die gleichzeitige Betrachtung von positiven wie negativen Abweichungen. Abbildung 6.6 zeigt die durchschnittlichen Abweichungsquadrate zu den zwei Messzeitpunkten.

Der Median der Varianz zwischen den transformierten Skalenwerten beträgt vor der Gruppenintervention 0,06 bei einem Range von 0 bis 0,36.[18] Nach der Gruppenphase ist der Median auf 0,09 angewachsen, wobei die Werte einzelner Patienten zwischen 0,003 und 0,53 schwanken. Dieser hypothesenkonforme Zuwachs ist im exakten gerichteten Wilcoxon-Test signifikant ($N = 29; p = 0,02$).

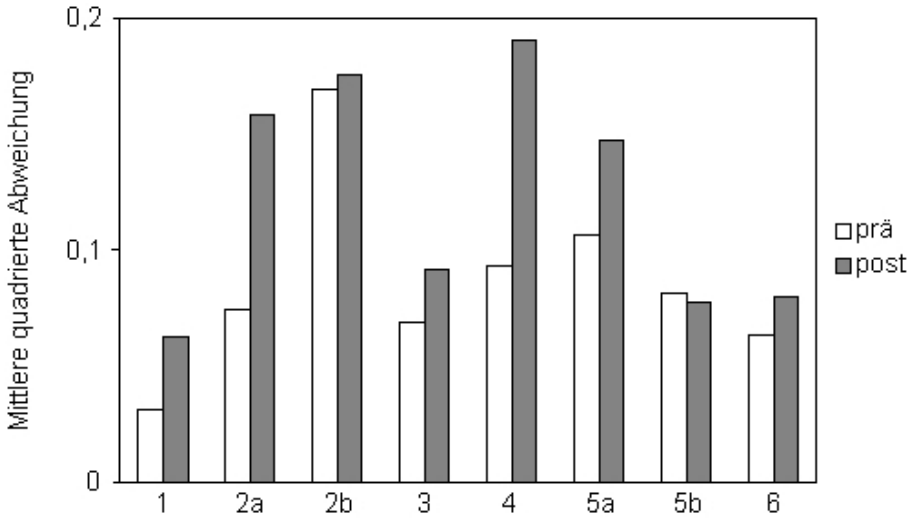

Abbildung 6.6: Durchschnittliche quadrierte Abweichung vom reskalierten individuellen Ungewissheitsniveau vor und nach der Gruppenintervention ($N = 29$).

6.9.3 Diskussion

In dieser Studie wurde der QUICC erstmals zur Evaluation einer SDM-Maßnahme eingesetzt, also in seinem originär intendierten Sinn. Dabei haben sich die vielversprechenden Ergebnisse der bisherigen Studien bestätigt. So erwies sich der QUICC einerseits als sensibel für behandlungsbedingte Veränderungen im Ausmaß erlebter Ungewissheiten. Andererseits ist nach allen indirekten Validierungsbefunden aus den vorangegangenen Untersuchungen insbesondere festzuhalten, dass sich die Effekte von SDM-Interventionen tatsächlich in einer signifikanten Erhöhung der Varianz zwischen den Ungewissheitsskalen zeigen. Dies unterstreicht nicht nur erneut die Sensibilität des Bogens, sondern spricht zusätzlich auch für die Plausibilität der in

[18] Es sei darauf hingewiesen, dass die hier berichteten Werte durch die Reskalierung eine andere Größenordnung als die Originaldaten haben und somit nicht inhaltlich auf die von 0 bis 4 reichende Antwortskala bezogen werden können.

dieser Arbeit entwickelten theoretischen Konzeption von SDM als Verhandlung von Ungewissheiten und dem daraus abgeleiteten Ergebnisparameter. Nach den bislang vorliegenden Untersuchungen besteht Anlass zur Annahme, dass mit dem QUICC endlich ein Messinstrument vorliegt, welches sowohl in psychometrischer als auch in theoretischer Hinsicht überzeugt.

Gleichwohl seien an dieser Stelle mögliche Einschränkungen bezüglich der Belastbarkeit der in dieser Validierungsstudie vorgestellten Daten diskutiert. So ist zwar unstrittig, dass der QUICC die Differenzierung im Ungewissheitserleben der Patienten belegt. Allerdings mag man einerseits hinterfragen, inwieweit die untersuchte Edukationsgruppe dem unter 1.4 beschriebenen SDM-Prototyp entspricht. Anderseits ermangelt dieses Studiendesign einer Kontrollgruppe, bestehend etwa aus einer paternalistischen Arzt-Patient-Kommunikation. Erst das Ausbleiben der Varianzzunahme in der Kontrollgruppe wäre neben der nachgewiesenen Sensitivität als notwendige Voraussetzung auch ein hinreichender Beleg für die Spezifität der vorgestellten Messmethodik. Weiterhin wäre mit einem Kontrollgruppendesign auch auszuschließen, dass die beschriebenen Effekte auf konfundierte Einlüsse zurückgehen wie etwa die während dieser Zeit geführten klärenden Gespräche mit Ärzten und Angehörigen etc.

Desweiteren fehlen zur erschöpfenden Beurteilung der Ergebnisse Daten über die Veränderung der Krankheitsintensität. Wie schon erwähnt, stehen gerade die im Rahmen einer erfolgreichen Depressionstherapie zunehmende Entscheidungsfähigkeit der Patienten sowie die Abnahme der Rumination in engem Zusammenhang mit dem Ungewissheitserleben, wie es der QUICC erfragt. Dabei sind Kausalitäten in beiden Richtungen denkbar. Man könnte auch sagen, das eine sei nicht Konsequenz sondern eher *Ausdruck* des anderen. Beide Lesarten würden die interne Validität der vorliegenden Untersuchung insofern mindern, als dann die Abnahme und Differenzierung des Ungewissheitserlebens nicht mehr nur und direkt auf die SDM-Intervention an sich zurückzuführen wären. Das Zusammenspiel von Depressivität und Ungewissheitserleben könnte man durch korrelative Betrachtungen näher beleuchten, indem der Anteil geteilter Varianz von Ungewissheitsdifferenzierung bzw. –reduktion einerseits und Rückgang der depressiven Symptomatik andererseits ermittelt würde. Leider gehört die kontinuierliche Diagnostik mit ensprechenden Messverfahren wie ‚Beck-Depressions-Inventar' (BDI) oder ‚Hamilton Rating Scale for Depression' (HAMD) nicht zum Alltagsrepertoire in der untersuchten Stichprobe, so

dass diese Frage hier unbeantwortet bleiben muss. Doch selbst wenn die Behandlungfortschritte Einfluss auf das Ungewissheitserleben nehmen, was ja plausibel ist und entsprechend bereits in der Hypothese zur Ungewissheitsreduktion sowie der Reskalierung der Daten Berücksichtigung fand, ist damit der postulierte und belegte Zusammenhang zwischen SDM und dem Differenzierungsgrad des Ungewissheitserlebens keineswegs in Frage gestellt. Vielmehr wurde mit den hier vorgestellten Überlegungen sensu Rog (1994) eine Möglichkeit der moderierenden Prozesse im Gesamtkomplex Behandlung skizziert.

Teil III

Gesamtdiskussion

Kapitel 7

Rekapitulation vorangegangener Schritte

7.1 Theoretische Präzisierung des Shared Decision Making

SDM ist derzeit vielfach Gegenstand der Diskussion im Gesundheitswesen – einerseits im Hinblick auf Patientenrechte, andererseits im Rahmen der Debatte um die Finanzierbarkeit des Gesundheitssystems. Dieser großen praktischen Relevanz steht der erstaunliche Mangel an einer hinreichenden theoretischen Fundierung gegenüber, der eingangs aufgezeigt wurde. So sind bislang weder überzeugende Konzeptionen der zu erwartenden Effekte von SDM vorgelegt worden, noch der vermittelnden Prozesse.

In dieser Arbeit wurde SDM als Verhandlung von Ungewissheiten präzisiert. Da SDM insbesondere in Situationen mit unklarer oder mangelnder Evidenz zu den verfügbaren Therapieoptionen indiziert ist, muss eine Auseinandersetzung mit den Ungewissheiten erfolgen, die aus diesem Mangel erwachsen. Doch selbst bei gesicherter Evidenzbasis können immer nur Aussagen über den idealisierten ‚Durchschnittspatienten' getroffen werden – wie die Behandlung beim individuell Betroffenen verläuft, kann man nie mit Sicherheit sagen. Außerdem lässt sich aus keiner noch so gut kontrollierten Wirksamkeitsstudie ableiten, wie sich etwa eine Partnerschaft im Zuge der Erkrankung und deren Behandlung entwickelt. Da Ungewissheiten nach diesen Überlegungen unausweichliche Charakteristika von Behandlungsentscheidungen

sind, erfolgte die Abkehr vom Anspruch einer globalen Reduktion von Ungewissheiten durch eine SDM-Intervention, wie sie z.B. in den Ansätzen von O'Connor (1995) und Mishel (1999) angestrebt wird. Trotz dieser Nichtreduzierbarkeit oder gerade deswegen sind Ungewissheiten Grund und Gegenstand von Risikokommunikation.

Aus dieser Konzeption konnte somit ein Zielparameter von SDM abgeleitet werden. Statt globaler Reduktion besteht dieser in einer Differenzierung zwischen den verschiedenen Ungewissheitsqualitäten. Damit können einzelne Ungewissheiten reduziert, andere allerdings auch vergrößert werden. Während eine diffus erlebte Ungewissheit bedrohlich wirken kann (Kasper et al., 2003), erlaubt eine differenzierte Sicht, zwischen veränderlichen und unveränderlichen Ungewissheiten zu unterscheiden, um gegebenenfalls aktive Schritte zur Reduktion einzelner Ungewissheiten zu unternehmen. Ferner eignen sich für unterschiedliche Ungewissheiten häufig auch unterschiedliche Copingstrategien.[1] Eine effektive SDM-Kommunikation führt demnach zu einer erhöhten Varianz in der Ausprägung der einzelnen Ungewissheitsqualitäten.

Diese Operationalisierung fällt in die Kategorie Ergebnisparameter für SDM, die in Abschnitt 2.4 diskutiert wurde. Gegenüber den anderen dort aufgeführten Variablen bietet sie jedoch den Vorteil einer direkten inhaltlichen Verknüpfung mit dem *Prozess* des SDM, indem dieser als Verhandlung oder konsensuelle Validierung der erlebten Ungewissheiten charakterisiert wird.

7.2 Die Qualitäten von Ungewissheit

Dass sich plausible Ungewissheitsqualitäten nicht nur theoretisch ableiten, sondern auch empirisch belegen lassen, wurde durch die ausführlichen Tiefeninterviews in Kombination mit dem Grounded Theory-Prozess gezeigt. Mittels dieses theoriebildenden Bottom up-Vorgehens unter ausschließlicher Einbeziehung der eigentlichen

[1] Bezüglich des ‚Copings von Ungewissheiten' stellt sich allerdings die Frage, was man als Auslöser und Reaktion bezeichnen möchte. Ist z.B. Ablenkung eine Copingstrategie gegen die Beschäftigung mit Ungewissheit? Oder ist die Beschäftigung mit Ungewissheit selbst als Coping der ungewissen Situation anzusehen? Kriz (1997) verweist bei einer analogen Thematik auf das Beispiel Fieber. Dieses kann man einerseits als zu behandelnde Krankheit betrachten, andererseits stellt Fieber selbst einen Bewältigungsversuch des Körpers etwa auf eine Entzündung dar.

Experten – der Patienten – erhalten die Kategorien oder Qualitäten von Ungewissheit eine deutlich größere Überzeugungskraft gegenüber rein theoretisch begründeten Ungewissheitstaxonomien wie sie in Abschnitt 4.2 referiert wurden. Die intersubjektive Nachvollziehbarkeit der aus dem Grounded Theory-Prozess resultierenden Kategorienlösung (vgl. Abschnitt 5.1.3) wird durch zwei Merkmale gestützt: Zum einen erreichte der Prozess innerhalb der ersten fünf von sechs Interviews eine informative Sättigung, was die Vernachlässigung wesentlicher Aspekte von Ungewissheit in der untersuchten Population unwahrscheinlich macht. Zum anderen war die Interraterübereinstimmung bei der Zuordnung von Äußerungen aus den Interviews zu den Kategorien befriedigend, was auf eine hinreichende inhaltliche Disjunktivität der Kategorien schließen lässt.

Über die *empirische* Eignung der Ungewissheitskategorien als Basis für ein Messinstrument hinaus wird damit ein weiterer, für das *theoretische* Verständnis von SDM wichtiger Aspekt deutlich: Zwar scheint es selbstverständlich, dass für eine Entscheidung Ungewissheiten über Diagnose, Prognose und Behandlungsmöglichkeiten (Kategorien 2a, 6 und teilweise 2b) erörtert werden müssen. Zusätzlich ging allerdings aus den Patienteninterviews hervor, dass auch Fragen z.B. zur sozialen Unterstützung oder der Bewältigung alltäglicher Anforderungen (Kategorien 1, 3 und 4) die Entscheidungsfindung massiv beeinflussen oder gar behindern können und folglich berücksichtigt werden müssen. Und schließlich wird erkennbar, dass auch jene Ungewissheiten entscheidungsrelevant sind, die die Kommunikation und den Behandler selbst betreffen, etwa Fragen zur gewünschten Rolle im Entscheidungsprozess (Kategorie 5b) sowie solche, die die Interpretation von Aussagen modulieren (Kategorien 2b, 5a). Diese Gruppe von Ungewissheiten hat damit eine Filterfunktion für alle weiteren Erörterungen zwischen Arzt und Patient und erfordert zusätzliche, metakommunikative Fertigkeiten vom Arzt, um die Kommunikation selbst zum Gegenstand der Kommunikation zu machen. Diese Erkenntnisse enthüllen die Komplexität effektiver SDM-Kommunikation, geben aber gleichzeitig erste Hinweise, worauf bei der Umsetzung von SDM in der Praxis zu achten ist.

Obwohl Ungewissheit wiederholt Gegenstand von Publikationen im Bereich des SDM sowie der Bewältigung chronischer Krankheiten war (z.B. Mishel, 1999; Babrow et al., 1998; Brashers et al., 2000), geht diese Arbeit zumindest in zwei wesentlichen Punkten über den bisherigen Stand der Forschung hinaus. So wurde zum einen erstmals im deutschen Sprachraum systematisch untersucht, welche Ungewissheiten

Patienten überhaupt erleben. Die in Abschnitt 4.2 berichteten Studien sind alle angloamerikanischen Ursprungs. Zum anderen fußt diese Ungewissheitstaxonomie als eine der wenigen auf einem qualitativen, empirischen Ansatz, der zudem anschließend quantitativ abgesichert wurde.

7.3 Die Entwicklung und Validierung des QUICC

Durch die vom Grounded Theory-Prozess erzwungene inhaltliche Präzision der Kategorien fiel es leicht, so viele potentiell geeignete Items zu generieren, dass mithilfe der in Kapitel 5.3 geschilderten Think Aloud-Prozedur nach Maßgabe der Patientensicht großzügig aussortiert werden konnte. Dabei wurde bereits der ursprüngliche Itempool als gut verständlich und exhaustiv bewertet.

In den anschließenden quantitativen Validierungsstudien 1 bis 5 erwies sich der QUICC mit seinen acht Skalen als ein hoch reliables Instrument zur Messung der Ausprägung der Ungewissheitsqualitäten. Bereits in der ersten Fragebogenversion wurden durchgängig hohe Werte für Cronbachs α erreicht, was erneut auch die Stimmigkeit des theoretischen Unterbaus unterstreicht. Selbst nach der Reduktion des Itempools blieben die Reliabilitäten der Skalen auf erfreulich hohem Niveau. Dieser Gesamteindruck spiegelt sich auch auf Itemebene anhand hoher Trennschärfekoeffizienten wider. Die Itemschwierigkeiten variieren naturgemäß recht stark zwischen den einzelnen Stichproben. Gleichwohl kann festgehalten werden, dass sich fast alle Itemmittelwerte in dem angestrebten Korridor der mittleren drei Antwortkategorien bewegen. Damit sind Boden- oder Deckeneffekte für die untersuchten Populationen recht unwahrscheinlich.

Die Validität des QUICC wurde auf verschiedene Weise untersucht. Für die Beurteilung der externen Validität wurde der Zusammenhang zwischen Untersuchungszeitpunkt und Ungewissheitsniveau betrachtet. Wie erwartet ergaben sich bei der Skala Behandlungsungewissheit die größten Unterschiede zwischen Patienten in Behandlung gegenüber jenen nach deren Abschluss. Dagegen war die Differenz bei der Skala zur Kausalattribution erwartungsgemäß am geringsten; diese Effektstärke war nur etwa halb so groß. Dabei waren alle acht Differenzen signifikant ($p<0{,}02$), was die Sensibilität des QUICC für verschiedene Behandlungsphasen unterstreicht.

Zusätzlich konnte der erwartete Zusammenhang zwischen Ungewissheitserleben und Ungewissheitstoleranz nachgewiesen werden. Dieser fiel jedoch recht schwach aus, was auf konkurrierende Einflüsse zwischen den beteiligten psychologischen Konstrukten sowie die geringe Reliabilität der UGTS zurückzuführen sein dürfte (vgl. Kapitel 6.5).

Die Beurteilung der inhaltlichen Validität des QUICC erwies sich als etwas aufwändiger. Zunächst wurde die häufig zur Beantwortung dieser Frage verwendete Faktorenanalyse aus methodischen und inhaltlichen Gründen ausgeschlossen. Zu diesen Gründen zählen z.B. die hohen Voraussetzungen an das Datenmaterial, die Uneindeutigkeit möglicher Lösungen sowie die unpassende Art der Verknüpfung der Modellvariablen. Stattdessen wurde zum einen die anhand der Cronbachs α-Werte ablesbare hohe Homogenität der Skalen als Beleg für die Übereinstimmung mit dem theoretisch postulierten Testmodell interpretiert. Zum anderen konnte in Validierungsstudie 4 gezeigt werden, dass eben jenes Testmodell, das im Grounded Theory-Prozess wurzelt, die Daten deutlich besser erklären kann als ein alternatives, eindimensionales Modell, das man auf Basis bestehender Literatur hätte erwarten können. Dabei konnte auch belegt werden, dass es berechtigt ist, für den QUICC ein Ratingskalenmodell anzunehmen, d.h. für die Items einer Skala untereinander identische Antwortschwellenabstände vorzugeben. Dies stimmt mit der intuitiv plausiblen Annahme überein, dass gleiche verbale oder numerische Antwortkategorienbezeichnungen innerhalb einer Skala auch zu gleichen wahrgenommenen Abstufungen in der Merkmalsausprägung führen sollten.

Die Skalen des QUICC korrelieren untereinander durchweg sehr hoch. Darin ist nach mehreren Studien jedoch keine Bedrohung der inhaltlichen Validität zu sehen. Denn einerseits sind hohe Interkorrelationen theoretisch einleuchtend. So kann sich etwa eine Ungewissheit über die Vertrauenswürdigkeit der Ärztin auf die Ungewissheit bezüglich der von ihr vorgeschlagenen Therapie auswirken und umgekehrt. Auch die gemeinsame Abhängigkeit des Ungewissheitserlebens von Drittvariablen wie Krankheitsverlauf oder Ungewissheitstoleranz trägt vermutlich zu hohen Interkorrelationen bei. Andererseits konnten mögliche Erklärungen der Interkorrelationen wie Schwächen am Layout, Stichproben- oder Ermüdungseffekte sowie die Unidimensionalität des Ungewissheitskonstkrukts systematisch ausgeschlossen werden, die die Inhaltsvalidität tatsächlich infrage gestellt hätten.

Validierungsstudie 5 integriert die bis hierhin rekapitulierten Aspekte der Eignung

des QUICC. So ließ sich im konkreten Anwendungsfall zeigen, dass der Bogen in der Tat die postulierte Differenzierung des Ungewissheitserlebens durch eine SDM-Intervention abzubilden vermag.

Die Nebengütekriterien Praktikabilität und Objektivität sind angesichts kurzer Bearbeitungsdauer und offenkundig verständlicher Instruktion erfüllt. Das schließt allerdings die Interpretationsobjektivität nicht mit ein, für die allgemein das Vorliegen von Normwerten gefordert wird (z.B. Rost, 2004). Auch wenn im Rahmen dieser Arbeit annähernd 500 Patienten mit verschiedenen Indikationen den QUICC bearbeitet haben, sind deren Ergebnisse nicht als Normstichprobe zu verwenden. Es stellt sich ohnehin die Frage, inwieweit man beim QUICC sinnvoll mit Normen arbeiten kann. Denn das Ungewissheitserleben ist der Definition nach nicht als Trait-, sondern eher als Statevariable anzusehen und damit in der Ausprägung sehr instabil. Daher fiele es grundsätzlich sehr schwer, geeignete Vergleichsgruppen bezüglich aller relevanten Aspekte zu finden, die sehr individuell (z.B. Ungewissheitstoleranz) und situativ sind (z.B. Einfluss der Wortwahl des Arztes). Ferner wäre für jede Untersuchungsfrage erneut zu klären, was z.B. eine ‚überdurchschnittliche Soziale Ungewissheit' inhaltlich bedeuten oder welche Konsequenz daraus abgeleitet werden könnte. Sinnvoller als solche Querschnittsvergleiche scheinen in der Regel Verlaufsbetrachtungen, wie sie weiter unten skizziert werden (vgl. Seite 146).

Alles in allem können die Erwartungen an den QUICC in den verschiedenen Stichproben als erfüllt angesehen werden. Er eignet sich sowohl vom testtheoretischen als auch vom inhaltlichen sowie vom anwendungspraktischen Standpunkt her nach den bisherigen Erfahrungen sehr gut für einen Einsatz zur Messung der Effekte von SDM. Dieses Fazit wurzelt nicht zuletzt in der insgesamt großen und heterogenen Stichprobe, die eine solide Datenbasis darstellt. Gleichwohl muss sich diese Erwartung erst noch in zukünftigen Untersuchungen bestätigen.

Kapitel 8

Mögliche Einschränkungen des Geltungsbereichs

8.1 Die Spezifität der Operationalisierung von Shared Decision Making

In den Eingangskapiteln war die Unzulänglichkeit bisheriger Operationalisierungsversuche von SDM als eine Motivation für die Entwicklung der vorliegenden Konzeption beschrieben worden. So kranken aktuelle Messverfahren unter anderem an mangelnder Spezifität, d.h. man kann anhand der Messergebnisse nicht verlässlich zwischen SDM und anderen Formen der Arzt-Patient-Kommunikation unterscheiden. Zwar konnte anhand der Ergebnisse aus Validierungsstudie 5 belegt werden, dass die vorgeschlagene Messmethodik die Effekte einer SDM-Intervention hypothesenkonform abbildet. Beim aktuellen Stand der Entwicklung kann jedoch noch nicht entschieden werden, ob die Operationalisierung mithilfe des Ungewissheitsprofils tatsächlich eindeutig ist, oder ob beispielsweise auch eine paternalistische Kommunikation in einem ähnlich differenzierten Ungewissheitserleben resultiert. Nach den Ausführungen in Abschnitt 4.3 spricht einiges für die Spezifität dieses Ansatzes. Letztlich wird sich aber in einem Kontrollgruppendesign empirisch erweisen müssen, ob der QUICC uneingeschrängt als Ergebnisparameter von SDM geeignet ist.

Dem Einwand noch nicht erwiesener Spezifität der Messmethodik mag man indes auch noch anders begegnen. So könnte man die Verhandlung und Differenzierung entscheidungsrelevanter Ungewissheiten unabhängig von SDM-Idealen auch *per se* als

erstrebenswert ansehen, etwa weil davon die Wiederherstellung der Handlungsfähigkeit von Patienten zu erwarten ist. In diesem Fall wäre in weiteren Untersuchungen primär weniger die Frage interessant, ob der Ergebnisparameter SDM-spezifisch ist, sondern vielmehr, welcher der – ethisch vertretbaren – Kommunikationsstile dem Ziel der Ungewissheitsdifferenzierung am ehesten dienlich ist. Zwar spricht nach den berichteten theoretischen und empirischen Aspekten einiges dafür, dass SDM diesbezüglich am effektivsten ist. Gleichwohl würde dieser Ansatz dazu ermutigen, den Ungewissheiten von Patienten je nach Typ und Situation (vgl. Seite 21f) unter Umständen auch einmal z.B. im Sinne eines ‚Professional as agent´ entgegenzutreten.

8.2 Die Allgemeingültigkeit der Ungewissheitskategorien

In Abschnitt 5.1.1 war bereits ausführlich diskutiert worden, wie eine Grounded Theory einzuordnen ist: Im Einklang mit Charmaz (2000) ist sie nicht als objektive Wahrheit anzusehen, sondern als *eine mögliche*, plausibel entwickelte Konzeption zu verstehen. Es steht außer Frage, dass durch Einbeziehung weiterer Interviewpartner mit anderen chronischen Krankheiten wie etwa Morbus Crohn oder Diabetes[1] möglicherweise eine andere Kategorienstruktur entstanden wäre. Nicht das Kriterium der Richtigkeit, sondern das der Nützlichkeit ist hier also anzulegen. Eben diese Nützlichkeit wird sich über die bisherigen Untersuchungen hinaus erweisen müssen.

Zusätzlich könnte das Ungewissheitserleben bei verschiedenen Indikationen auch mit demselben Ansatz systematisch beleuchtet werden. Dazu könnte man entweder Elemente aus entsprechenden Tiefeninterviews in die bestehende Kategorienstruktur einzuordnen versuchen oder gänzlich ohne Restriktionen ein neues Kategoriensystem auf demselben Weg erzeugen und mit dem bisherigen vergleichen.

Die Divergenz zwischen Richtigkeit und Nützlichkeit bei der Grounded Theory findet eine Entsprechung in der quantitativen Validierung: Dass die gebildeten acht Skalen reliabel messen, heißt nicht, dass eine andere Skalenstruktur nicht ebenso

[1] und eventuell sogar anderen Patienten mit derselben Diagnose, aber einer anderen Sichtweise ihrer Situation

gute oder sogar bessere Eigenschaften aufweisen könnte. Auch bei dem in Validierungsstudie 4 vorgestellten Modelltest wurden nur zwei plausible von unzähligen möglichen Konzeptionen verglichen. Abgesehen von der vermutlich nicht zu entscheidenden Frage, was ein ‚richtiges Kategoriensystem' sein sollte oder wonach man es als ‚richtig' klassifizieren könnte, handelt es sich hierbei jedoch nicht um eine spezifische Eigenschaft der hier gewählten Vorgehensweise, sondern um ein allgemeines Problem. Es wird jedoch durch zuweilen unreflektierte Mischung von ‚der Realität' und den Sichtweisen auf diese Realität übersehen.

Kapitel 9

Integration weiterführender Gedanken in die vorgelegte Theorie

In dieser Arbeit ließen sich alle wesentlichen theoretisch begründeten Erwartungen über das Ungewissheitserleben chronisch Kranker überzeugend empirisch belegen. Daher erscheint es angemessen, zum Ende noch einige neue, weiterführende Gedanken in die vorgelegte Theorie zu integrieren, aus denen sich verschiedene anzuschließende Untersuchungen ableiten lassen.

9.1 Die Verhandlung von Ungewissheiten als Ordnungsbildungsprozess

Die Nützlichkeit der vorgestellten SDM-Konzeption erweist sich auch in der Anschlussfähigkeit an bestehende Ansätze in anderen Disziplinen, z.B der Selbstorganisationstheorie. Wären medizinische oder selbst alltägliche Entscheidungen so überschaubar wie die aus der Spieltheorie entlehnten Situationen, die im Rahmen der Deskriptiven Entscheidungstheorien diskutiert werden (vgl. Abschnitt 3.2), dann könnten Menschen tatsächlich ihre Entscheidungen durch eine Verrechnung von Eintrittswahrscheinlichkeiten und Ergebnisfolgeerwartungen treffen.[1] Es wurde aber gezeigt, dass solch ein Versuch in realen Situationen scheitern muss (s. Abschnitt 4),

[1] Dass Versuchspersonen das aber selbst in relativ einfachen Situationen nicht zwangsläufig auch tun, zeigte Kriz bereits 1982.

denn es handelt sich im mathematischen Sinne um ein ‚ill posed problem' in zweifacher Hinsicht: Zum einen ist die Entscheidung *unterdeterminiert*, d.h. wesentliche Informationen über den Entscheidungsgegenstand und die Handlungsoptionen sind unbekannt oder *ungewiss*. Gleichzeitig ist die Entscheidung *überdeterminiert*, d.h. es stehen trotz partiellen Informationsmangels in anderen Bereichen so viele komplexe und teilweise widersprüchliche Informationen zur Verfügung, dass selektiv nur eine Untermenge daraus verarbeitet werden kann. Zeitdruck kann die Lage noch zusätzlich verschärfen. Aus beiden Charakteristika komplexer Entscheidungssituationen folgt, dass sie nicht ohne weiteres ‚rechnerisch' gelöst werden können, wie es beim Roulette möglich wäre. Wie also kann man sich dann einen Entscheidungsprozess vorstellen?

Wieder hilft eine Analogie zur visuellen Wahrnehmung. Haken (1996) hat in diesem Bereich mit dem Konzept des ‚Synergetischen Computers' ein formales Modell entwickelt, welches zeigt, wie auf Basis eines visuellen Inputs Muster erkannt und unterschieden werden können, etwa Gesichter oder Gegenstände. Zwischen synergetisch modellierter Mustererkennung und Entscheidungsfindung lassen sich nach Haken und Schiepek (2006) nun vielfältige, charakteristische Parallelen aufzeigen:

- Mustererkennung und Entscheidungsfindung können (prinzipiell) in eine eindeutige Identifizierung bzw. Entscheidung münden.

- Durch Fokussierung auf unterschiedliche Details kommen jeweils unterschiedliche Aspekte und Detailinformationen in den Blick. Die wegen der meist unvollständigen Datenlage nötige Komplettierung des Datensatzes kann als ‚Muster*bildungs*prozess' beschrieben werden, bei dem Interessen, Vorlieben, Erwartungen und Emotionen Einfluss nehmen, was in selektiver Informationssuche und -verarbeitung resultiert (Kriz, 2001).

- Die Resultate – Muster und Entscheidungen – zeichnen sich häufig durch eine gewisse Instabilität aus, wie man sie von Kippfiguren wie dem Necker-Würfel kennt. Im Sinne einer ‚Symmetriebrechung' geben oft Kleinigkeiten oder zufällige Schwankungen den Ausschlag für die eine und gegen die andere Lösung. In Verbindung mit den bereits beschriebenen Einflüssen von Voreingenommenheiten führen diese häufig zu dem Eindruck, die Entscheidung oder Wahrnehmung sei ‚irrational'.

- Auch nach der Festlegung auf eine Lösung muss diese Instabilität nicht überwunden sein. Vielmehr gewinnt die verworfene Interpretation häufig gerade nach der Festlegung an Plausibilität oder Attraktivität, wie es auch in der Theorie der kognitiven Dissonanz (Festinger, 1978) beschrieben wird.
- In beiden Fällen lassen sich ‚Hystereseeffekte' festmachen: Wer sich einmal mit Erfolg in bestimmter Weise entschieden hat, neigt dazu, sich zukünftig genauso zu entscheiden, selbst wenn sich die Umstände geändert haben. Die visuelle Variante ist in Abbildung 9.1 zu sehen.
- Für beide Anwendungsbereiche lassen sich im synergetischen Modell Aufmerksamkeitsparameter operationalisieren, die die Fokussierung auf einzelne Ausschnitte der Datenbasis abbilden. Erweist sich eine visuelle Interpretation oder eine Entscheidung als falsch, fällt der zugehörige Parameter auf null, so dass andere, vorher weniger plausible Lösungen in Betracht gezogen werden.

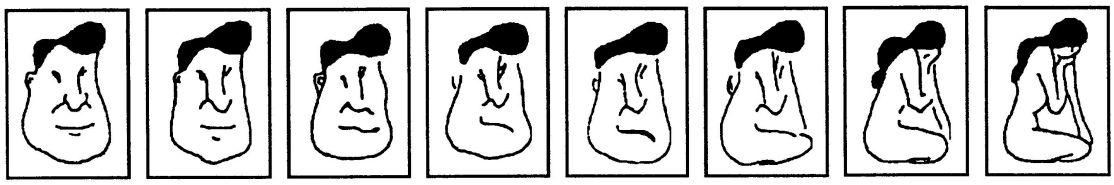

Abbildung 9.1: Hysterese. Innerhalb der Bilderfolge transformiert sich die Abbildung von einem Gesicht zu einer sitzenden Frau. Je nach ‚Systemgeschichte', d.h. bei Beginn von rechts oder links schlägt die dominante Interpretation jeweils bei einem anderen Bild um. Die einmal gewählte Variante wird möglichst lange beibehalten (Abbildung aus Schiepek, 1999).

Diese Analogiebildung verweist auf die zugrunde liegende Theorie der Synergetik, nach der man die gesamte Entscheidungsfindung von der Informationssuche bis hin zur endgültigen Entscheidung (und darüber hinaus) als ‚selbstorganisierenden Ordnungsbildungsprozess' ansehen kann:

> „Die durch den Einfluss bestimmender Kontrollparameter (emotional relevante Motive und Ziele) getriggerte Vorbereitungsphase ist durch kritische Instabilität (geringer Wirkungsgrad des Systems) und Informationssuche (kognitive Aktivität bei der Musterergänzung) gekennzeichnet, bis sich ein Ordner (die gewählte Option) durchsetzt. Dieser prägt dann das Handeln und weist eine gewisse Stabilität auf, d.h. ablenkende Einflüsse und immer wieder mögliche Alternativen haben geringe Chancen, die

> Umsetzung der Entscheidung (Handlungsrealisierung) zu stören“ (Haken und Schiepek, 2006, S. 215).

Die Konzeption von Entscheidungsfindung als Ordnungsbildungsprozess hat dabei neben der interdisziplinären Anschlussfähigkeit den Vorteil, dass in der Synergetik empirische Methoden zum Nachweis von Ordnungsbildung entwickelt worden sind (Schiepek und Strunk, 1994), die damit *prinzipiell* auch bei Entscheidungsprozessen eingesetzt werden könnten.

An dieser Stelle schließt sich der Kreis zur vorliegenden Arbeit: Wenn man Shared Decision Making als Verhandlung von Ungewissheiten beschreiben und das Ungewissheitserleben mit dem QUICC erfassen kann, so wäre zu überlegen, ob sich anhand einer Messreihe mit dem QUICC während eines SDM-Prozesses eine Ordnungsbildung im Sinne der Synergetik nachweisen ließe. Gewiss sind dabei bereits jetzt Schwierigkeiten auszumachen. So verwendet man etwa in der Physik Zeitreihendaten mit mehreren tausend Messzeitpunkten, um Ordnungsbildungsphänomenen auf die Spur zu kommen. Aufgrund der Zumutbarkeit und des Einflusses von Erinnerungseffekten ist mit dem QUICC vermutlich höchstens eine Messung pro Tag möglich. Somit kommen nur Anwendungskontexte in Frage, die Entscheidungsprozesse über längere Zeiträume ermöglichen. Dass trotz einiger versuchsplanerischer Hindernisse grundsätzlich der Nachweis von Ordnungsbildungsphänomenen in psychosozialen Kontexten anhand von Fragebogendaten möglich ist, zeigen indes Anwendungen im Rahmen der Psychotherapieprozessforschung (Schiepek, 1999; Haken und Schiepek, 2006).

9.2 Interaktionale Konzeption des Shared Decision Making

Die bisher vorgestellte Theorie ist gewissermaßen nur die halbe Wahrheit. In einem wichtigen Punkt ist die vorgestellte Theorie des SDM noch nicht über die Unzulänglichkeit anderer Ansätze hinausgekommen: Es wurde lediglich über das Ungewissheitserleben *des Patienten* und die entsprechenden Effekte von SDM diskutiert. Wenn aber SDM, wie eingangs berichtet, als *geteilte* Entscheidungsfindung definiert wird, dann sollte eine Theorie über SDM sicherlich auch *beide* teilenden Protagonisten umfassen, also auch den Arzt. Dazu bieten sich zwei Wege an:

9.2.1 Konkordanz als interaktionales Effektmaß

Die erste Möglichkeit behält die Fokussierung auf die patientenseitigen Ungewissheiten bei. Dies lässt sich insofern rechtfertigen, als der Patient auch im SDM die Therapieentscheidung selbst letztlich alleine trifft, auch wenn der Entscheidungsfindungsprozess bis dahin gemeinsam mit dem Arzt erfolgte (vgl. Seite 16f bzw. 19f). Da ein Hauptmerkmal des SDM der Konsens zwischen Arzt und Patient ist und das Ergebnis von SDM in einer veränderten Repräsentation der entscheidungsrelevanten Ungewissheiten besteht, müsste man SDM interaktional definieren als die *Konkordanz* von Arzt und Patient bei der Ungewissheitswahrnehmung. Der ‚Two way exchange of information' nach Charles et al. (1997) wäre dann als gelungen zu bezeichnen, wenn Arzt und Patient in der Einschätzung der Ungewissheiten übereinstimmen. Diese Übereinstimmung ließe sich leicht operationalisieren, indem zusätzlich zum Patienten auch der Arzt einen QUICC ausfüllt, mit dem er das Ungewissheitserleben des Patienten einschätzt. An der Korrelation beider Fragebögen wäre demnach der Effekt von SDM abzulesen (Kasper et al., 2006).

Diese Operationalisierung geht dabei über die Übereinstimmung bzgl. der reinen Sachinformationen hinaus. Denn eine valide Einschätzung der patientenseitigen Ungewissheiten durch den Arzt ist nur möglich, wenn dieser auch eine gewisse Nähe zum Patienten hat und sich entsprechend gut einfühlen kann. Damit kommt man auch dem wichtigen Beziehungsaspekt des SDM näher (Makoul und Clayman, 2005; Siminoff und Step, 2005).

9.2.2 Die Ungewissheiten der Ärzte

Die zweite Möglichkeit, dem interaktionalen Charakter von SDM gerecht zu werden, besteht darin, die Ungewissheiten der Ärzte einzubeziehen. Auch dies lässt sich aus der Definition des SDM heraus begründen: Denn wenn SDM als eine gemeinsame Verhandlung von Ungewissheiten beschrieben wird, ist es plausibel, das Ungewissheitserleben auch bei beiden Beteiligten zu untersuchen. A priori kann man erwarten, dass es teilweise Überlappungen mit den Ungewissheiten der Patienten gibt, etwa im Hinblick auf die Diagnose oder die Behandlung. An anderer Stelle kann man auch Unterschiede vermuten. So dürfte z.B. die Frage der Kausalattribution (Kategorie 4) für Ärzte eher nebensächlich sein.

In einer Pilotstudie wurde dieser Frage bereits nachgegangen. Hausärzte wurden eingeladen, ihren Umgang mit Ungewissheiten näher zu betrachten. Im Zuge dessen wurden Videoaufnahmen von sieben Gesprächen analysiert, bei denen eine Therapieentscheidung getroffen oder vorbereitet wurde. In einer ersten Videoanalysephase stellte der Autor gemeinsam mit einem weiteren Experten für Risikokommunikation Hypothesen auf, welche Ungewissheitsqualitäten der Arzt während des Gesprächs erlebt haben könnte. Im zweiten Schritt wurde der jeweilige Arzt nach einer ausführlichen thematischen Einführung gebeten, sich sein Video im Beisein der Experten anzusehen und ebenfalls die während des Gesprächs erlebten Ungewissheiten herauszuarbeiten bzw. sich derer zu erinnern. Sofern damit nicht alle der zuvor aufgestellten Ungewissheitshypothesen abgedeckt waren, wurden dem Arzt die verbliebenen anhand der entsprechenden Videosequenz vorgestellt, um sie so an seiner Innensicht zu validieren. Das sequentielle Vorgehen erlaubte zusätzlich, den Ärzten in einer letzten Analysephase die berichteten Ungewissheitsqualitäten ihrer Kollegen vorzustellen, um unabhängig vom jeweils analysierten Video zu klären, ob es sich dabei um typische bzw. verbreitete Ungewissheiten handelt.

Die bislang extrahierten Ungewissheitsqualitäten lauten wie folgt:
Anliegen und Auftrag. Welche Zielvorstellung hat der Patient? Kann ich dem Auftrag des Patienten (noch) gerecht werden?
Vertrauen des Patienten. Hat der Patient Vertrauen in mich und meine Kompetenz?
Verständnis des Patienten. Hat der Patient alle relevanten Informationen verstanden?
Informationsvermittlung. Wann möchte der Patient Informationen in welchem Ausmaß?
Diagnose. Was ist der aktuelle Krankheitszustand und wie wird der weitere Verlauf sein?
Diagnostik. Welche (weiteren) diagnostischen Maßnahmen sind berechtigt bzw. notwendig?
Behandlung. Was ist in diesem Fall die optimale Behandlung?

Anhand dieser Kurzübersicht wird deutlich, dass sich die erlebten Ungewissheiten bei Ärzten und Patienten teilweise entsprechen. In Tabelle 9.1 werden die Ungewissheitstaxonomien einander gegenübergestellt. Dabei ist zu beachten, dass die

korrespondierenden Ungewissheiten aus den beiden Perspektiven nicht vollständig deckungsgleich sind, selbst wenn die Bezeichnungen zum Teil übereinstimmen.

Arzt	Patient
	Soziale Einbindung
Diagnose	Diagnose – Prognose
Verständnis des Patienten	Informationsentschlüsselung
Diagnostik	
	Bewältigung von Anforderungen
	Kausalattribution
Vertrauen des Patienten	Qualität des Arztes
Informationsvermittlung	Rollenpräferenz
Anliegen und Auftrag	
Behandlung	Behandlung

Tabelle 9.1: Gegenüberstellung der entscheidungsbezogenen Ungewissheiten bei Ärzten und Patienten.

Dem Arzt obliegt im SDM nicht die Therapieentscheidung selbst, sondern die Gestaltung des Entscheidungsprozesses. Insofern ist es seine Aufgabe, zur Klärung *aller* entscheidungsrelevanten Ungewissheiten beizutragen, d.h. denen des Patienten und seiner eigenen. Daher wurde in besagter Pilotstudie auch festgehalten, wie die Ärzte mit den identifizierten Ungewissheiten jeweils umgehen. Dabei fiel auf, dass jeder Arzt meist nur einige wenige Umgangsweisen für verschiedene Ungewissheiten anwendete, während das Spektrum der Copingstrategien in der Gesamtgruppe der Studienteilnehmer recht breit war. Nun ist davon auszugehen, dass die Qualität ärztlichen Ungewissheitsmanagements mit der Anzahl der potentiell verfügbaren Umgangsstrategien wächst, weil damit die Chance steigt, eine für die jeweilige Situation passende auswählen zu können. Es erhöht sich somit der ärztliche Handlungsspielraum. Daher sollen in Zukunft aus dem bisherigen und weiter zu erhebenden Datenmaterial alle bei den Ungewissheitsqualitäten offensichtlich anwendbaren Copingstrategien herausgearbeitet werden, um damit ein Trainingscurriculum zu entwickeln. Leitend ist dafür die Überzeugung, dass Ungewissheit A nicht schematisch am besten mit Strategie B begegnet werden sollte, sondern dass die Anzahl der Wahlmöglichkeiten auf ärztlicher Seite zu erhöhen ist. Die Auswahl der am ehesten

passenden Strategie trifft weiterhin der Arzt selbst.

9.3 Shared Decision Making als interaktionaler Ordnungsbildungsprozess

Nach den beiden vorangegangenen Abschnitten liegt es nahe, die dort vorgestellten Ansätze zu kombinieren. So wurde unter 9.1 zunächst erläutert, inwiefern man die Entscheidungsfindung als Ordnungsbildungsprozess in den patientenseitigen Ungewissheiten auffassen kann. In Abschnitt 9.2 wurde dagegen klargestellt, wie wichtig die interaktionale Perspektive beim SDM ist. Beides lässt sich integrieren. So wäre einerseits denkbar, nicht nur beim Patienten, sondern auch beim Arzt geeignete Maße zu entwickeln, anhand derer man dessen Entscheidungsfindung als Ordnungsbildungsprozess beschreiben könnte. Da SDM aber interaktional definiert wird, sind nicht nur die Ordnungsbildungsprozesse *innerhalb* der beiden Beteiligten interessant, sondern insbesondere auch derjenige *zwischen* Arzt und Patient. Damit wäre eine sehr treffende Präzisierung des für SDM charakteristischen Konsenses skizziert, die die Schwierigkeit der empirischen Umsetzung nach sich zieht. In einem Folgeprojekt soll diesbezüglich überprüft werden, ob sich dieser Konsens zwischen Arzt und Patient anhand der ‚Structural Analysis of Social Behavior' (SASB, Benjamin, 1993) abbilden lässt.

Schlussbemerkung

Der vorliegende Ansatz erscheint im Wesentlichen aus drei Gründen weiterführend: Erstens wurde ein tragfähiges theoretisches Fundament für die Präzisierung von SDM entwickelt. Zweitens konnte daraus ein anhand vielfältiger Methoden empirisch fundiertes Messinstrument entwickelt werden. Und drittens lässt sich die Konzeption von SDM als Verhandlung von Ungewissheiten gut mit anderen aktuellen Forschungsansätzen verbinden, woraus sich weitere Untersuchungsperspektiven ergeben.

Dass diese Konzeption trotz der enormen wissenschaftlichen Fortschritte gerade auf *Un*gewissheiten aufbaut, ist dabei vielleicht ein Charakteristikum dieser Zeit. An die Stelle unkritischen Fortschrittsglaubens ist (glücklicherweise) eine Offenheit für die blinden Flecken der Wissenschaft getreten. Der Soziologe Armin Nassehi, von dem bereits das Eingangszitat dieser Arbeit stammt, bringt es folgendermaßen auf den Punkt:

> „Das Neue ist doch, dass wir heute einsehen können, dass Nichtwissen etwas ist, was man sich zumuten muss, was nicht vermieden werden kann, sondern produktiv eingesetzt werden muss“ (Nassehi, 2005).

Literatur

Andrych, D. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. Applied Psychological Measurement, 2, 581-594.

Babrow, A. S., Kasch, C. R. und Ford, L. A. (1998). The many meanings of uncertainty in illness: toward a systematic accounting. Health Communication, 10(1), 1-23.

Bandura, A. (1977). Self-efficacy. Psychological Review, 84, 191-215.

Beaver, K., Luker, K., Owens, R., Leinster, S., Degner, L. F. und Sloan, J. (1996). Treatment decision making in women newly diagnosed with breast cancer. Cancer Nursing, 19, 8-19.

Beisecker, A. und Beisecker, T. (1990). Patient information-seeking behaviours when communicating with doctors. Medical Care, 28(1), 19-28.

Benjamin, L. A. (1993). SASB - Die Strukturale Analyse Sozialen Verhaltens. Heidelberg: Asanger.

Bernoulli, D. (1738). Exposition of a New Theory on the Measurement of Risk, 1954 (Originaltitel: Comentarii Academiae Scientiarum Imperialis Petropolitanae). Econometrica, 22, 23-36.

Blanchard, C., Labrecque, M., Ruckdeschel, J. und Blanchard, E. (1988). Information and decision-making preferences of hospitalized adult cancer patients. Social Science and Medicine, 27(11), 1139-1145.

BMGS. (2000). Bekanntmachung des Bundesministeriums für Gesundheit im Rahmen des Modellprogramms zur Förderung der medizinischen Qualitätssicherung "Der Patient als Partner im medizinischen Entscheidungsprozeß".

Bottorff, J. L., Ratner, P. A., Johnson, J. L., Lovato, C. Y. und Joab, S. A. (1998). Communicating cancer risk information: the challenges of uncertainty. Patient Education and Counseling, 33, 67-81.

Brashers, D. E. (2001). Communication and uncertainty management. Journal of Communication, 51(3), 477-497.

Brashers, D. E., Neidig, J. L., Haas, S. M., Dobbs, L. K., Cardillo, L. W. und Russell, J. A. (2000). Communication in the management of uncertainty: The case of persons living with HIV or AIDS. Communication Monographs, 67, 63-84.

Brashers, D. E., Neidig, J. L., Russell, J. A., Cardillo, L. W., Haas, S. M., Dobbs, L. K., Garland, M., McCartney, B. und Nemeth, S. (2003). The medical, personal, and social causes of uncertainty in HIV illness. Issues in Mental Health Nursing, 24, 497-522.

Breitbart, W. (1995). Identifying patients at risk for, and treatment of major psychiatric complications of cancer. Supportive Care in Cancer, 3(1), 45-60.

Büchi, M., Bachmann, L. M., Fischer, J. E., Peltenburg, M. und Steurer, J. (2000). Alle Macht den Patienten? Vom ärztlichen Paternalismus zum Shared Decision Making. Schweizerische Ärztezeitung, 81(49), 2776-2780.

Buchsbaum, G. (1980). A spatial processor model for object colour perception. Journal of The Franklin Institute, 310, 1-26.

Buckelew, S. P., Huyser, B., Hewett, J. E., Parker, J. C., Johnson, J. C., Conway, R. und Kay, D. R. (1996). Self-efficacy predicting outcome among fibromyalgia subjects. Arthritis Care Research, 9(2), 97-104.

Bühner, M. (2005). Einführung in die Test- und Fragebogenkonstruktion. München: Pearson Studium.

Bundesgerichtshof. (15. März 2005). VI ZR 313/03. www.bundesgerichtshof.de.

Carstensen, C. H. und Rost, J. (2003). MULTIRA: Ein Programmsystem für mehrdimensionale Rasch-Modelle. (Version v1.65). Kiel: IPN.

CDU/CSU/SPD. (2005). Gemeinsam für Deutschland - mit Mut und Menschlichkeit. Koalitionsvertrag zwischen CDU, CSU und SPD. www.spd.de.

Charles, C., Gafni, A. und Whelan, T. (1997). Shared decision-making in the medical encounter: what does it mean? (Or it takes at least two to tango). Social Science and Medicine, 44(5), 681-692.

Charles, C., Gafni, A. und Whelan, T. (1999a). Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. Social Science and Medicine, 49, 651-661.

Charles, C., Whelan, T., Gafni, A. und Lindsay, E. (1999b). Shared treatment decision-making among physicians and women with early stage breast cancer. Unveröffentlichte Studie, National Cancer Institute of Canada, Toronto.

Charmaz, K. (2000). Grounded Theory: Objectivist and Constructivist Methods. In N. K. Denzin und Y. S. Lincoln (Hrsg.), Handbook of Qualitative Research (2. ed., S. 509-535). Thousand Oaks: Sage.

Clark, N. M., Gong, M., Schork, A., Evans, D., Roloff, D. und Hurwitz, M. (1998). Impact of education for physicians on patient outcomes. Pediatrics, 101, 831-836.

Cleary, P. D. (1998). Satisfaction may not suffice! A commentary on 'A patient's perspective'. International Journal of Technology Assessment in Health Care, 14(1), 24-34.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 10(1), 37-46.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? Behavioral and Brain Sciences, 4, 317-370.

Cohen, M. H. (1993). The unknown and the unknowable - managing sustained uncertainty. Western Journal of Nursing Research, 15(1), 77-96.

Cole, B. und Pargament, K. (1999). Re-creating your life: a spiritual/psychotherapeutic intervention for people diagnosed with cancer. Psycho-Oncology, 8(5), 395-407.

Dalbert, C. (1996). Ungewißheitstoleranz und der Umgang mit Ungerechtigkeit. In dieselbe (Hrsg.), Über den Umgang mit Ungerechtigkeit. Eine psychologische Analyse (S. 189-230). Bern: Verlag Hans Huber.

Dalbert, C. (1999). Die Ungewissheitstoleranzskala: Skaleneigenschaften und Validierungsbefunde. Halle: Martin-Luther-Universität Halle-Wittenberg, FB Erziehungswissenschaften - Pädagogik.

Dalbert, C. (2002). UGTS. Ungewissheitstoleranzskala. In E. Brähler, J. Schumacher und B. Strauß (Hrsg.), Diagnostische Verfahren in der Psychotherapie (S. 374-377). Göttingen: Hogrefe.

Davidson, J., D, B. M. und Feldman-Stewart, D. (1999). Lung cancer treatment decisions: patients' desires for participation and information. Psycho-Oncology, 8, 511-520.

Deber, R. (1994). Physicians in health care management. The patient-physician partnership: changing roles and the desire for information. Canadian Medical Association Journal, 151(2), 171-176.

Degner, L. F. und Sloan, J. (1992). Decision making during serious illnesses: what role do patients really want to play? Journal of Clinical Epidemiology, 45(9), 941-950.

Donner-Banzhoff, N. (2001). Perspektiven der Hypertoniebehandlung in der hausärztlichen Praxis. Zeitschrift für ärztliche Fortbildung und Qualitätssicherung, 95(5), 333-338.

Duden. (2001). Das Herkunftswörterbuch. Etymologie der deutschen Sprache. (Vol. 7). Mannheim: Dudenverlag.

Edwards, A., Elwyn, G., Hood, K., Robling, M., Atwell, C., Holmes-Rovner, M., Kinnersley, P., Houston, H. und Russell, I. (2003). The development of COMRADE - a patient-based outcome measure to evaluate the effectiveness of risk communication and treatment decision making in consultations. Patient Education and Counseling, 50, 311-322.

Elwyn, G., Edwards, A. und Kinnersley, P. (1999). Shared decision-making in primary care: the neglected second half of the consultation. British Journal of General Practice, 49(443), 477-482.

Elwyn, G., Edwards, A., Mowle, S., Wensing, M., Wilkinson, C., Kinnersley, P. und Grol, R. (2001). Measuring the involvement of patients in shared decision-making: a systematic review of instruments. Patient Education and Counseling, 43(1), 5-22.

Elwyn, G., Edwards, A., Wensing, M., Hood, K., Atwell, C. und Grol, R. (2003). Shared decision making: developing the OPTION scale for measuring patient involvement. Quality and Safety in Health Care, 12, 93-99.

Elwyn, G., Hutchings, H., Edwards, A., Rapport, F., Wensing, M., Cheung, W.-Y. und Grol, R. (2005). The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks. Health Expectations, 8, 34-42.

Emanuel, E. J. und Emanuel, L. L. (1992). Four models of the physician-patient relationship. Journal of the American Medical Association, 267, 2221-2226.

Ende, J., Kazis, L., Ash, A. und Moskowitz, M. A. (1989). Measuring Patients' Desire for Autonomy: Decision Making and Information-seeking Preferences among Medical Patients. Journal of General Internal Medicine, 4(1), 23-30.

Entwistle, V., Skea, Z. und O'Donnell, M. (2001). Decisions about treatment: interpretations of two measures of control by women having a hysterectomy. Social Science and Medicine, 53(6), 721-732.

European Charter of Patients' Rights. (2002). www.activecitizenship.net/health/ec.htm.

Evans, J. S. (1989). Bias in human reasoning: Causes and consequences. Brighton: Erlbaum.

Fallowfield, L. (2001). Participation of patients in decions about treatment for cancer. British Medical Journal, 323(7322), 1144.

Fallowfield, L., Hall, A., Maguire, P., Baum, M. und A'Hern, R. P. (1994). Psychological effects of being offered choice of surgery for breast cancer. British Medical Journal, 309, 448.

Festinger, L. (1978). Theorie der kognitiven Dissonanz. Bern: Huber.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. Psychological Bulletin, 76(5), 378-382.

Giersdorf, N., Loh, A., Bieber, C., Caspari, C., Deinzer, A., Doering, T., Eich, W., Hamann, J., Heesen, C., Kasper, J., Leppert, K., Müller, K., Neumann, T., Neuner, B., Rohlfing, H., Scheibler, F., Oorschot, B. v., Spies, C., Vodermaier, A., Weiss-Gerlach, E., Zysno, P. und Härter, M. (2004). Entwicklung eines Fragebogens zur Partizipativen Entscheidungsfindung. Das Gesundheitswesen, 47, 969-976.

Giersdorf, N., Loh, A. und Härter, M. (2003). Quantitative Messverfahren des shared decision-making. In F. Scheibler und H. Pfaff (Hrsg.), Shared decision-making. Der Patient als Partner im medizinischen Entscheidungsprozess. Weinheim: Juventa.

Giesen, D. (1987). Zwischen Patientenwohl und Patientenwille. Juristenzeitung, 42(6), 282-290.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. Psychological Review, 103, 592-596.

Gigerenzer, G. und Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. Psychological Review, 103, 650-669.

Glaser, B. G. (2002). Constructivist Grounded Theory? Forum: Qualitative Social Research, 3(3).

Glaser, B. G. und Strauss, A. L. (1967). The discovery of grounded theory: strategies of qualitative research. New York: Aldine Publishing Company.

Grant, F. C., Laupacis, A., O'Connor, A. M., Rubens, F. und Robblee, J. (2001). Evaluation of a decision aid for patients considering autologous blood donation before open-heart surgery. Canadian Medical Association Journal, Apr 17; 164(8), 1139-1144.

Haken, H. (1996). An Application of Synergetics. Decision Making as Pattern Recognition. Zeitschrift für Wissenschaftsforschung, 9/10, 45-72.

Haken, H. und Schiepek, G. (2006). Synergetik in der Psychologie. Selbstorganisation verstehen und gestalten. Göttingen: Hogrefe.

Harman, H. H. (1976). Modern factor analysis. Chicago: University of Chicago Press.

Haynes, R., McKibbon, K. und Kanani, R. (1996). Systematic review of randomised trials of interventions to assist patients to follow prescriptions for medications. Lancet, 348(9024), 383-386.

Heesen, C., Schwickert, Y., Kasper, J., Köpke, S., Scheunemann, D., Meyer, G. und Mühlhauser, I. (2004). Immuntherapien der Multiplen Sklerose. Hamburg.

Helmholtz, H. v. (1855). Über das Sehen des Menschen, Vorträge und Reden (4. ed., Vol. 1). Braunschweig: Vieweg.

Helmholtz, H. v. (1896). Handbuch der Physiologischen Optik (2. Aufl.). Hamburg: Voss.

Holmes-Rovner, M., Kroll, J., Rovner, D., Rother, M. L., Padonu, G. und Talarczyk, G. (1996). Patient satisfaction with health care decisions: the satisfaction with decision scale. Medical Decision Making, 16, 58-64.

Howie, J., Heaney, D. und Maxwell, M. (1997). Measuring quality in general practice. Pilot study of a needs, process and outcome measure, Royal College of General Practicioners. London.

Hurley, J., Birch, S. und Eyles, J. (1992). Information efficiency and decentralization within health care systems, CHEPA Working Paper. Hamilton: McMaster University.

Isfort, J., Floer, B., Koneczny, N., Vollmar, H. C. und Butzlaff, M. (2002). Wer entscheidet – Arzt oder Patient? Deutsche Medizinische Wochenschrift, 127, 2021-2024.

Kahneman, D., Slovic, P. und Tversky, A. (Hrsg.). (1982). Judgement under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press.

Kahneman, D. und Tversky, A. (1979). Prospect Theory: An analysis of decisions under risk. Econometrica, 47, 263-291.

Kasper, J. (2005, 23.-25.05.2005). Welchen Beitrag leistet der BMGS Förderschwerpunkt 'Patient als Partner' zum Bereich der Entscheidungshilfen? Gemeinsam entscheiden - neue Anstöße zur Förderung der Patientenbeteiligung, Berlin.

Kasper, J., Geiger, F., Südkamp, A. und Schneider, S. (2006). Entwicklung eines Fragebogens zur Messung des Erlebens krankheitsbezogener Ungewissheit. Patient und Arzt als Team. Von der wissenschaftlichen Erkenntnis zur praktischen Umsetzung. 4. deutscher Kongress zum Thema Partizipative Entscheidungsfindung in der Medizin, Heidelberg.

Kasper, J., Kuch, C. und Heesen, C. (2003). Shared decision-making als Interaktionsstil. In F. Scheibler und H. Pfaff (Hrsg.), Shared decision-making. Der Patient als Partner im medizinischen Entscheidungsprozess (S. 34-45). Weinheim: Juventa.

Kelle, U. (1994). Die Methode der empirisch fundierten Theoriebildung von Glaser und Strauss. In derselbe (Hrsg.), Empirisch begründete Theoriebildung (S. 283-349). Weinheim: Deutscher Studien Verlag.

Kennedy, A. D. (2003). On what basis should the effectiveness of decision aids be judged? Health Expectations, 6(3), 255-268.

Kjellgren, K. I., Svensson, S., Ahlner, J. und Saljo, R. (2000). Antihypertensive treatment and patient autonomy – the follow-up appointment as a resource for care. Patient Education and Counseling, 40(1), 39-49.

Klemperer, D. (2003). Wie Ärzte und Patienten Entscheidungen treffen. Konzepte der Arzt-Patient-Kommunikation. Berlin: Wissenschaftszentrum Berlin für Sozialforschung.

Koedoot, N., Molenaar, S., Oosterveld, P., Bakker, P., de Graeff, A., Nooy, M., Varekamp, I. und de Haes, H. (2001). The decisional conflict scale: further validation in two samples of Dutch oncology patients. Patient Education and Counseling, 45(3), 187-193.

Kozine, I. (2004). Survey of decision-making theories. Riso: Riso National Laboratory.

Kriz, J. (1982). Subjektive Wahrscheinlichkeiten und Entscheidungen: zur Problematik von Methodenartefakten. Frankfurt/Main: R. G. Fischer.

Kriz, J. (1997). Systemtheorie. Eine Einführung für Psychotherapeuten, Psychologen und Mediziner. Wien: Facultas-Universitäts-Verlag.

Kriz, J. (2001). Self-Organization of Cognitive and Interactional Processes. In M. Matthies, H. Malchow und J. Kriz (Hrsg.), Integrative Systems Approaches to Natural and Social Dynamics (S. 517-537). Berlin: Springer.

Laufs, A. und Uhlenbruck, W. (2002). Handbuch des Arztrechts. München: Beck.

Lerman, C. E., Brody, D. S., Caputo, G. C., Smith, D. G., Lazaro, C. G. und Wolfson, H. G. (1990). Patients' Perceived Involvement in Care Scale: Relationship to Attitudes about Illness and Medical Care. Journal of General Internal Medicine, 5, 29-33.

Levine, M. N., Gafni, A., Markham, B. und MacFarlane, D. (1992). A bedside decision instrument to elicit a patient's preference concerning adjuvant chemotherapy for breast cancer. Annals of Internal Medicine, 117, 53-58.

Lin, C.-C. (1998). Comparison of the Effects of Perceived Self-Efficacy on Coping with Chronic Cancer Pain and Coping with Chronic Low Back Pain. Clinical Journal of Pain, 14(4), 303-310.

Lisch, R. und Kriz, J. (1978). Grundlagen und Modelle der Inhaltsanalyse. Bestandsaufnahme und Kritik. Reinbek: Rowohlt.

Loh, A., Simon, D., Bermejo, I. und Härter, M. (2005). Miteinander statt Nebeneinander – Der Patient als Partner in der Depressionsbehandlung. In M. Härter, A. Loh und C. Spies (Hrsg.), Gemeinsam entscheiden – erfolgreich behandeln. Neue Wege für Ärzte und Patienten im Gesundheitswesen (S. 165-174). Köln: Deutscher Ärzte-Verlag.

Lübbe, A. (2003, 10.-12.10.). Ist der Patientenwille alles? Patientinnen und Patienten, Ärztinnen und Ärzte im Recht, Bad Boll.

Lupton, D. (1997). Consumerism, reflexivity and the medical encounter. Social Science and Medicine, 45(3), 373-381.

Makoul, G. und Clayman, M. L. (2005). An integrative model of shared decision making in medical encounters. Patient Education and Counseling, 60, 301-312.

Man-Son-Hing, M., Laupacis, A., O'Connor, A. M., Biggs, J., Drake, E., Yetisir, E. und Hart, R. (1999). A patient decision aid regarding antithrombotic therapy for stroke prevention in atrial fibrillation: a randomized controlled trial. Journal of the American Medical Association, 282(8), 737-743.

Masters, G. M. (1982). A Rasch model for partial credit scoring. Psychometrika, 47, 149-174.

McNutt, R. A. (2004). Shared Medical Decision Making. Problems, Process, Progress. Journal of the American Medical Association, 292(20), 2516-2518.

Meyer, H. (2004). Theorie und Qualitätsbeurteilung psychometrischer Tests. Stuttgart: Kohlhammer.

Mishel, M. H. (1981). The Measurement of Uncertainty in Illness. Nursing Research, 30(5), 258-263.

Mishel, M. H. (1999). Uncertainty in chronic illness. Annual Review of Nursing Research, 17, 269-294.

Mishel, M. H., Hostetter, T., King, B. und Graham, V. (1984). Predictors of psychosocial adjustment in patients newly diagnosed with gynecological cancer. Cancer Nursing, 7(4), 291-299.

Nassehi, A. (26.06.2005). Zu viel Wissen schadet nur (Interview). Frankfurter Allgemeine Sonntagszeitung.

Neumann, J. v. und Morgenstern, O. (1947). Theory of games and economic behavior (2. Aufl.). Princeton: Princeton University Press.

O'Connor, A. M. (1995). Validation of a decisional conflict scale. Medical Decision Making, 15(1), 25-30.

Parsons, T. (1951). The Social Systems. Glencoe, Illinois: Free Press.

Pfaff, H., Freise, D. C., Mager, G. und Schrappe, M. (2001). Der Kölner Patientenfragebogen (KPF): Entwicklung und Validierung eines Fragebogens zur Erfassung der Einbindung des Patienten als Kotherapeuten (Forschungsbericht 1). Köln: Abteilung der Medizinischen Soziologie der Universität zu Köln.

Prüfer, P. und Rexroth, M. (1996). Verfahren zur Evaluation von Survey-Fragen: Ein Überblick. ZUMA-Nachrichten, 39, 95-116.

Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests (2. Aufl.). Chicago: University of Chicago Press, 1980.

Rog, D. J. (1994). Constructing Natural Experiments. In J. S. Wholey, H. P. Hatry und K. E. Newcomer (Hrsg.), Handbook of practical program evaluation (S. 119-132). New York: Wiley.

Rost, J. (2004). Lehrbuch Testtheorie - Testkonstruktion. Bern: Verlag Hans Huber.

Roter, D. L., Hall, J. A., Kern, D. E., Barker, L. R., Cole, K. A. und Roca, R. P. (1995). Improving physicians' interviewing skills and reducing patients' emotional distress. A randomized clinical trial. Archives of Internal Medicine, 155, 1877-1884.

Rothenbacher, D., Lutz, M. P. und Porzsolt, F. (1997). Treatment decisions in palliative cancer care: Patients' preferences for involvement and doctors' knowledge about it. European Journal of Cancer, 33, 1184-1189.

Rowe, W. D. (1994). Understanding uncertainty. Risk Analysis, 14(5), 743-750.

Sainfort, F. und Booske, B. (2000). Measuring post-decision satisfaction. Medical Decision Making, 20(1), 51-61.

Savage, L. J. (1954). The foundations of satistics. New York: Wiley.

Scheibler, F., Janssen, C. und Pfaff, H. (2003). Shared decision making: Ein Überblick über die internationale Forschungsliteratur. Sozial- und Präventivmedizin, 48(1), 11-23.

Schiepek, G. (1999). Selbstorganisation in psychischen und sozialen Prozessen: Neue Perspektiven der Psychotherapie. In K. Mainzer (Hrsg.), Komplexe Systeme in Natur und Gesellschaft. Komplexitätsforschung in Deutschland auf dem Weg ins nächste Jahrhundert (S. 280-317). Berlin: Springer.

Schiepek, G. und Strunk, G. (1994). Dynamische Systeme. Grundlagen und Analysemethoden für Psychologen und Psychiater. Heidelberg: Asanger.

Schönemann, P. H. und Borg, I. (1996). Von der Faktorenanalyse zu den Strukturgleichungsmodellen. In E. Erdfelder, R. Mausfeld, T. Meiser und G. Rudinger (Hrsg.), Handbuch Quantitative Methoden (S. 241-268). Weinheim: Psychologie Verlags Union.

Seale, C. (1999). The quality of qualitative research. London: Sage.

Siminoff, L. A. und Step, M. M. (2005). A Communication Model of Shared Decision Making: Accounting for Cancer Treatment Decisions. Health Psychology, 24(4(Suppl.)), 99-105.

Simon, D., Loh, A. und Härter, M. (2005a). Messung der partizipativen Entscheidungsfindung. In M. Härter, A. Loh und C. Spies (Hrsg.), Gemeinsam entscheiden - erfolgreich behandeln. Neue Wege für Ärzte und Patienten im Gesundheitswesen (S. 239-247). Köln: Deutscher Ärzte-Verlag.

Simon, D., Loh, A. und Härter, M. (2005b, 21.-24.09.2005). Validierung eines Fragebogens zur Partizipativen Entscheidungsfindung (PEF-FB). 4. Deutscher Kongress für Versorgungsforschung. Die Zukunft der Gesundheitsversorgung in Deutschland: Bedarfsgerechtigkeit, Innovation, Systemgestaltung, Berlin.

Sorrentino, R. M. und Roney, C. J. R. (2000). The uncertain mind: Individual differences in facing the unknown. Philadelphia: Taylor and Francis.

Steinke, I. (1999). Kriterien qualitativer Forschung. Ansätze zur Bewertung qualitativ-empirischer Sozialforschung. Weinheim: Juventa.

Stewart, D. E., Wong, F., Cheung, A. M., Dancey, J., Meanar, M., Cameron, J. I., McAndrews, M. P., Bunston, T., Murphy, J. und Rosen, B. (2000). Information needs and decisional preferences among women with ovarian cancer. Gynecological Oncology, 77(3), 357-361.

Stewart, M., Brown, J., Donner, A., McWhinney, I., Oates, J. und Weston, W. W. (1995). The impact of patient-centered care on patient outcomes in family practice. London, Ontario: Centre for Studies in Family Medicine.

Strauss, A. L. und Corbin, J. (1996). Grounded theory: Grundlagen qualitativer Sozialforschung. Weinheim: Beltz, PVU.

Strull, W. M., Lo, B. und Charles, G. (1984). Do Patients Want to Participate in Medical Decision Making? Journal of the American Medical Association, 252(21), 2990-2994.

Towle, A. und Godolphin, W. (1999). Framework for teaching and learning informed shared decision making. British Medical Journal, 319(7212), 766-771.

Tversky, A. und Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-232.

Tversky, A. und Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Tversky, A. und Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211, 453-458.

Van Someren, M. W., Barnard, Y. F. und Sandberg, J. A. C. (1994). The think aloud method. A practical guide to modelling cognitive processes. London: Academic Press.

Veronesi, U., von Kleist, S., Redmond, K., Costa, A., Delvaux, N., Freilich, G., Glaus, A., Hudson, R., McVie, J. G., Macnamara, C., Meunier, F., Pecorelli, S. und Serin, D. (1999). Caring about women and cancer (CAWAG): A European survey of the perspectives and experiences of women with female cancers. European Journal of Cancer, 35(12), 1667-1675.

Volmer, T. und Kielhorn, A. (1998). Compliance und Gesundheitsökonomie. In F. Petermann (Hrsg.), Compliance und Selbstmanagement (S. 103-138). Göttingen: Hogrefe.

Wallberg, B., Michelson, H., Nystedt, M., Bolund, M., Degner, L. und Wilking, N. (2000). Information needs and preferences for patient participation in treatment decisions among Swedish breast cancer patients. Acta Oncologica, 39(4), 467-476.

Walter, O. (2005). Kompetenzmessung in den PISA-Studien: Simulationen zur Schätzung von Verteilungsparametern und Reliabilitäten. Lengerich: Pabst Science Publishers.

Waterworth, S. und Luker, K. (1990). Reluctant collaborators: do patients want to be involved in decisions concerning care? Journal of Advances in Nursing, 15, 971-976.

Weiner, B., Russell, D. und Lerman, D. (1978). Affective consequences of causal ascriptions. In J. H. Harvey, W. J. Ickes und R. F. Kidd (Hrsg.), New directions in attribution research (Vol. 2). Hillsdale: Erlbaum.

Weston, W. W. (2001). Informed and shared decision-making: the crux of patient-centred care. Canadian Medical Association Journal, 165(4), 438-439.

Wirtz, M. und Caspar, F. (2002). Beurteilerübereinstimmung und Beurteilerreliabilität - Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen: Hogrefe.

Wu, M., Adams, R. J. und Haldane, S. (2003). ConQuest - Generalised Item Response Modelling Software (Version 28 August 2003). Camberwell: Australian Council for Educational Research.

Young, M. und Klingle, R. (1996). Silent partners in medical care: a cross-cultural study of patient participation. Health Communication, 8(1), 29-53.

Appendices

Anhang A

Formeln

A.1 Cohens κ

$$\kappa = \frac{P_{obs} - P_{exp}}{1 - P_{exp}}$$

$$P_{exp} = \frac{1}{N^2} \sum_{j=1}^{s} n_{.j} n_{j.} = \sum_{i=1}^{s} h_{.j} h_{j.}$$

P_{obs} :	Prozentualer Anteil beobachteter Übereinstimmungen
P_{exp} :	Prozentualer Anteil zufällig zu erwartender Übereinstimmungen
N :	Anzahl der beurteilten Objekte
s :	Anzahl der Kategorien
$n_{.j}, n_{j.}$:	Anzahl der zu Kategorie j zugewiesenen Objekte (Randsummen)

A.2 Fleiss' κ

$$\kappa = \frac{\overline{P} - \overline{P}_{exp}}{1 - \overline{P}_{exp}}$$

$$\overline{P} = \frac{1}{N} \sum_{i=1}^{N} P_i$$

$$P_i = \frac{1}{n(n-1)} \sum_{j=1}^{k} n_{ij}(n_{ij} - 1)$$

$$\overline{P}_{exp} = \sum_{j=1}^{k} p_j^2$$

$$p_j = \frac{1}{Nn} \sum_{i=1}^{N} n_{ij}$$

N :	Anzahl der Objekte
n :	Anzahl der Ratings pro Objekt
k :	Anzahl der Kategorien
i :	Laufindex Objekte
j :	Laufindex Kategorien
n_{ij} :	Anzahl der Rater bei Objekt i und Kategorie j
P_j :	Prozentuale Übereinstimmung der Rater bei Objekt j
P_{exp} :	Prozentualer Anteil zufällig zu erwartender Übereinstimmungen
p_j :	Wahrscheinlichkeit der Kategorien bei allen Zuordnungen

A.3 Cronbachs α

$$\alpha = \frac{N\overline{r}}{1 + (N-1) - \overline{r}}$$

N :	Anzahl der Items
$\overline{r}$:	Mittlere Korrelation zwischen allen Items

A.4 AIC und BIC

$$AIC = -2logL + 2n_p$$

$$BIC = -2logL + (logN)n_p$$

L :	Likelihood
n_p :	Anzahl der geschätzten Parameter
N :	Anzahl der Versuchspersonen

Anhang B

Leitfaden zur Kategorisierung des Interviewmaterials

Der untenstehende Leitfaden wurde im Rahmen der Kategorisierung des Interviewmaterials entwickelt. Die hochgestellten Buchstaben a bis g verweisen auf nahe verwandte und daher besonders aufmerksam abzugrenzende Leitfragen mit jeweils identischer Kennzeichnung.

1. **Soziale Einbindung**

   Diese Kategorie versammelt Ungewissheiten, die die Verlässlichkeit der Beziehungen zum sozialen Umfeld angesichts der Erkrankung betreffen.

   > Wie stehen meine *Mitmenschen* zu meiner Krankheit und ihren Folgen? (Mitmensch≠Arzt)
   > Wird mich die Gesellschaft mit meiner Erkrankung unterstützen/benachteiligen?[a]
   > Was erwarten sie von mir?
   > Wie reagieren sie?
   > Wie ist ihr Verhalten zu interpretieren?

2. **Krankheitszustand**

   Die beiden folgenden Kategorien bezeichnen Ungewissheiten, die die Einschätzung des aktuellen Krankheitszustandes betreffen.

   a) **Diagnose – Prognose**

Was ist der aktuelle *Istzustand* bezüglich meiner Erkrankung?
Was bedeuten die Symptome? Oder bilde ich sie mir nur ein?
Wie wird der *weitere Verlauf* meiner Erkrankung prognostiziert?
Werde ich sterben? Werde ich den Krebs besiegen?[b]
Was kommt da auf mich zu? (im Sinne von 'Wie wird es mir wohl bei der Behandlung ergehen?' – ohne dass die Behandlung in Frage gestellt wird)[c]
Wird meine Erkrankung so schlimm werden, dass es dazu kommen könnte, dass andere mich benachteiligen?[a]

b) **Informationsentschlüsselung**

Was bedeutet es, wenn die Ärztin etwas andeutet/nichts andeutet?
Was bedeutet: 'Es sieht ganz gut aus'/'Es besteht ein Verdacht'/'Ich habe für Sie lieber mal einen Termin in der Uniklinik gemacht'?
Was bedeutet es, wenn der Befund auf sich warten lässt?
Was bedeuten irgendwelche Allgemeinstatistiken für mich persönlich?
Was bedeutet es, wenn Befunde oder Beurteilungen einander widersprechen? (im Sinne von 'Was ist da in meinem Körper los?'/'Warum können die Ärzte da nichts Eindeutiges feststellen?' – ohne dass die Kompetenz der Ärzte in Frage gestellt wird)[d]
Was passiert hier mit mir? (wenn sich der Patient die Frage stellt, was die Ärzte in Zukunft mit ihm vorhaben und ob ihr Verhalten bedeuteten könnte, dass es doch schlimmer um die eigene Erkrankung steht)[e]

3. **Bewältigung von Anforderungen**

Hier werden Ungewissheiten subsumiert, die die eigene Anpassungsfähigkeit (Adaptation) betreffen.

Wie werde ich zukünftige Herausforderungen in Anbetracht meiner (überstandenen) Erkrankung *bewältigen*?
Kann ich wieder arbeiten?
Was wird sich sonst noch in meinem Leben ändern (müssen)?

Werde ich mit dem Krebs umgehen können? Kann ich damit leben?[b]
Werde ich es bewältigen, wenn andere mich aufgrund meiner Erkrankung benachteiligen?[a]
Werde ich die Belastungen der Behandlung bewältigen? (ohne die Behandlungswahl zu hinterfragen)

4. **Kausalattribution**

Hierunter werden Ungewissheiten verstanden, die sich bei dem Versuch auftun, das eigene Krankheitsgeschehen in der Vergangenheit, Gegenwart oder Zukunft sinnstiftend kognitiv zu integrieren.

Womit hängt meine Krankheit/Gesundheit zusammen? z.B.: Wie hängt es mit meiner Krankheit zusammen, dass ich früher viel geraucht habe?[f]
Wodurch wird/wurde sie gefördert? Wodurch gebremst?
Ist es eine Fügung, die mir etwas sagen will?
Soll ich wieder arbeiten?

5. **Umgang mit dem Arzt**

Die beiden folgenden Kategorien beinhalten Ungewissheiten im Umgang mit dem Behandlungspersonal.

a) **Rollenpräferenz**

Wie viel Verantwortung/Kontrolle soll ich in der Behandlung übernehmen/übergeben? (Partizipationsgrad)[g]
Wie viel Kontrolle möchte ich über die Entscheidungsfindung haben?
Wie genau möchte ich über Details informiert sein?
Möchte ich schlechte Nachrichten erfahren?
Soll ich mich gegen die Meinung der Ärzte durchsetzen?

b) **Qualität des Arztes**

Ist der Arzt ausreichend kompetent?[d]
Gibt meine Ärztin sich genug Mühe?
Kennt sie auch ihre Grenzen?

Bin ich auch menschlich bei meinem Arzt gut aufgehoben?
Ist er aufrichtig? Erfahre ich auch schlechte Nachrichten?

6. **Behandlung**

Diese Kategorie umfasst Ungewissheiten darüber, welche medizinischen und ‚privaten' Handlungen die Krankheit in gewünschter Weise beeinflussen.

In welchem Verhältnis stehen Nutzen und Nebenwirkungen einer (jeden) Behandlungsoption?[c]
Hätte ich mich doch anders entscheiden sollen? Sollte ich die Behandlung abbrechen?[e]
Mit welcher Haltung kann ich mir helfen? Welche Einstellung wirkt sich gut, welche hinderlich auf den Verlauf aus?
Was kann ich selbst tun, damit es mir besser geht? z.B. mit dem Rauchen aufhören[f]
Wie viele Kontrolluntersuchungen sind gut?[g]

Anhang C

Die QUICC-Version aus der Vorstudie

Qualitäten von Ungewissheiten in chronischen Krankheiten

Bogen für Betroffene von Krebserkrankungen (QUICC-CA)

Datum:	Code:

Liebe Patientin, lieber Patient,

mit folgenden Fragen möchten wir erfahren, wie Sie sich und Ihre Situation bezüglich Ihrer Krebserkrankung wahrnehmen. Manche der Fragen werden Ihnen vielleicht ähnlich vorkommen, aber sie beleuchten unterschiedliche Aspekte. Bitte lassen Sie keine Frage aus, und setzen Sie jeweils nur ein Kreuz. Einige der angesprochenen Aspekte empfinden Sie vielleicht je nach Tagesform und Situation als schwankend.

Antworten Sie deshalb bitte ganz auf den jetzigen Moment bezogen.

Zur Illustration ein einfaches Beispiel:

Ob das Wochenende wohl verregnet sein wird?

Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr
	☐	☐	☐	☐	☐

Natürlich kann man das Wochenendwetter nicht mit völliger Sicherheit vorhersagen, es besteht also eine Ungewissheit.
Uns interessiert dabei nicht die Wahrscheinlichkeit schlechten Wetters, sondern **inwieweit Sie sich über diese Ungewissheit Gedanken machen**: Wenn Sie z.B. zu einer Gartenparty geladen haben, dann würden Sie wahrscheinlich „sehr“ ankreuzen, weil Sie auf gutes Wetter angewiesen sind.
Wenn Sie sich stattdessen für jenes Wochenende die Neuordnung Ihrer Briefmarkensammlung vorgenommen haben, würden Sie womöglich „gar nicht“ ankreuzen, weil dabei die Frage des Wetters für Sie nicht entscheidend ist.

Vielen Dank für Ihre Mühe!

Inwieweit erleben Sie in den folgenden Bereichen Ungewissheit?

Ob mich meine Angehörigen wohl auch unterstützen, wenn es mir mal schlechter geht?

Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	
	☐	☐	☐	☐	☐	01

Wie wohl mein aktueller Krankheitszustand ist?

Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	
	☐	☐	☐	☐	☐	02

Ob ich wohl zukünftig die Auswirkungen meiner Erkrankung bewältigen kann?

Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	
	☐	☐	☐	☐	☐	03

1

Ob ich verstehe, was mir der Arzt über meine Situation mitteilen will?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	04
	☐	☐	☐	☐	☐	
Ob ich meinen finanziellen Verpflichtungen trotz der Erkrankung nachkommen kann?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	05
	☐	☐	☐	☐	☐	
Ob sich der Arzt ausreichend mit meiner Erkrankung auskennt?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	06
	☐	☐	☐	☐	☐	
Wie sich wohl meine Erkrankung weiterentwickeln wird?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	07
	☐	☐	☐	☐	☐	
Ob ich mich intensiver mit Informationen über meine Erkrankung auseinandersetzen sollte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	08
	☐	☐	☐	☐	☐	
Ob mir der Arzt wichtige Informationen verschweigen würde?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	09
	☐	☐	☐	☐	☐	
Ob mich krankheitsbezogene Gedanken und Ängste wohl überfordern können?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	10
	☐	☐	☐	☐	☐	
Welchen Unterschied es gemacht hätte, wenn die Erkrankung zu einem anderen Zeitpunkt diagnostiziert worden wäre?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	11
	☐	☐	☐	☐	☐	
Woran ich merken kann, dass die Behandlung, für die ich mich entschieden habe, gut ist?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	12
	☐	☐	☐	☐	☐	
Was wohl in Fällen wie meinem die optimale Behandlung wäre?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	13
	☐	☐	☐	☐	☐	
Wie viel Unterstützung von Seiten meiner Mitmenschen ich wohl in Anspruch nehmen kann?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	14
	☐	☐	☐	☐	☐	

Ob es ein schlechtes Zeichen für meinen Zustand ist, wenn sich Fachleute in meinem Fall uneinig sind?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	15
	☐	☐	☐	☐	☐	
Inwieweit ich meinem Arzt Bedenken mitteilen sollte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	16
	☐	☐	☐	☐	☐	
Wie sich mein Körper im Laufe der Erkrankung wohl noch verändern wird?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	17
	☐	☐	☐	☐	☐	
Ob ich zukünftig mit den Einschränkungen umgehen kann, die ich aufgrund der Krankheit hinnehmen muss?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	18
	☐	☐	☐	☐	☐	
Was ich wohl zusätzlich zu der medizinischen Behandlung noch tun kann?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	19
	☐	☐	☐	☐	☐	
Welchen Einfluss meine Erkrankung auf meine sozialen Beziehungen hat oder noch haben wird?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	20
	☐	☐	☐	☐	☐	
Ob mir wohl etwas entgangen ist, das der Arzt mir mitteilen wollte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	21
	☐	☐	☐	☐	☐	
Wie ich für eine medizinische Maßnahme das Verhältnis von Nutzen und Risiken bewerten soll?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	22
	☐	☐	☐	☐	☐	
Ob sich wohl seit der letzten Untersuchung etwas in meinem Körper verändert hat?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	23
	☐	☐	☐	☐	☐	
Ob mich der Arzt mit meinen Sorgen ernst nimmt?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	24
	☐	☐	☐	☐	☐	
Ob ich wohl zukünftig trotz Erkrankung meinem Alltag gewachsen bin?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	25
	☐	☐	☐	☐	☐	

3

Was andere denken, wenn sie Zeichen meiner Erkrankung bemerken?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	26
	☐	☐	☐	☐	☐	
Welche Rückschlüsse ich auf meinen Zustand ziehen kann aus der Art und Weise, wie der Arzt mit mir umgeht?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	27
	☐	☐	☐	☐	☐	
Ob ich auch menschlich bei meinem Arzt gut aufgehoben bin?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	28
	☐	☐	☐	☐	☐	
Wie die Erkrankung wohl mit meinem bisherigen Leben zusammenhängt?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	29
	☐	☐	☐	☐	☐	
Ob ich den Erwartungen anderer trotz meiner Erkrankung gerecht werde?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	30
	☐	☐	☐	☐	☐	
Woran ich eine gute Entscheidung bezüglich medizinischer Maßnahmen erkennen kann?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	31
	☐	☐	☐	☐	☐	
Wie aktiv ich mich in die Gespräche mit meinem Arzt einbringen sollte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	32
	☐	☐	☐	☐	☐	
Was es wohl bedeutet, wenn sich Aussagen oder Befunde zu meinem Zustand widersprechen?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	33
	☐	☐	☐	☐	☐	
Wie viel Belastung durch meine Erkrankung meine Mitmenschen ertragen können?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	34
	☐	☐	☐	☐	☐	
Was wohl Ursachen für meine Erkrankung sind?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	35
	☐	☐	☐	☐	☐	
Inwieweit mir die Beziehungen zu nahestehenden Personen gut tun?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	36
	☐	☐	☐	☐	☐	

Ob ich vom Arzt auch über schlechte Nachrichten informiert werden möchte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	37
	☐	☐	☐	☐	☐	
Ob bestimmte Belastungen in meinem Leben die Krankheit mit bedingt haben?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	38
	☐	☐	☐	☐	☐	
Ob der Arzt den neuesten Forschungsstand zu meiner Erkrankung kennt?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	39
	☐	☐	☐	☐	☐	
Was allgemeine Statistiken für meinen persönlichen Krankheitsverlauf bedeuten?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	40
	☐	☐	☐	☐	☐	
Was bestimmte Beschwerden wohl über meinen Krankheitszustand aussagen?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	41
	☐	☐	☐	☐	☐	
Bis zu welchem Punkt meine mir liebsten Menschen hinter mir stehen?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	42
	☐	☐	☐	☐	☐	
Ob ich wirklich alles richtig verstanden habe bezüglich meiner Erkrankung und den Folgen?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	43
	☐	☐	☐	☐	☐	
Wie ich zukünftig trotz der Erkrankung mit dem Leben zurechtkommen werde?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	44
	☐	☐	☐	☐	☐	
Inwieweit bei meiner Erkrankung wohl die Psyche eine Rolle spielt?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	45
	☐	☐	☐	☐	☐	
Inwieweit ich mich wohl an Behandlungsentscheidungen beteiligen sollte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	46
	☐	☐	☐	☐	☐	
Was wohl bezüglich der Erkrankung gerade in meinem Körper vorgeht?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	47
	☐	☐	☐	☐	☐	

5

Inwieweit mich meine Erkrankung sozial isoliert?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	48
	☐	☐	☐	☐	☐	
Ob mir der Arzt auch schlechte Nachrichten mitteilen würde?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	49
	☐	☐	☐	☐	☐	
Warum ich wohl gerade diese Krankheit bekommen habe?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	50
	☐	☐	☐	☐	☐	
Ob ich trotz meiner Erkrankung das erreichen kann, was ich mir vornehme?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	51
	☐	☐	☐	☐	☐	
Inwieweit wohl allgemeine Statistiken bezüglich meiner Krankheit auf mich persönlich zutreffen?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	52
	☐	☐	☐	☐	☐	
Ob meine wichtigen Beziehungen der Erkrankung standhalten werden?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	53
	☐	☐	☐	☐	☐	
Ob mir der Arzt die ganze Wahrheit sagt?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	54
	☐	☐	☐	☐	☐	
Welchen körperlichen Anforderungen ich mit meiner Erkrankung gewachsen bin?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	55
	☐	☐	☐	☐	☐	
Ob meine Erkrankung wohl eine Fügung ist?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	56
	☐	☐	☐	☐	☐	
Wie ich mein Leben verändern kann, um auf meine Erkrankung günstig einzuwirken?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	57
	☐	☐	☐	☐	☐	
Was es wohl bedeutet, wenn ich an bestimmten Stellen im Körper Beschwerden habe?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	58
	☐	☐	☐	☐	☐	

Ob sich der Arzt wohl genug Mühe mit mir gibt?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	59
	☐	☐	☐	☐	☐	
Wie andere mit meiner Diagnose zurechtkommen?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	60
	☐	☐	☐	☐	☐	
Warum gerade bei mir die Krankheit entstanden ist?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	61
	☐	☐	☐	☐	☐	
Welches Maß an Aufmerksamkeit für meinen eigenen Körper mir wohl gut tut?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	62
	☐	☐	☐	☐	☐	
Inwieweit ich in Gesprächen mit dem Arzt meinen eigenen Standpunkt vertreten soll?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	63
	☐	☐	☐	☐	☐	
Inwieweit ungewohnte Körperempfindungen der Beachtung wert sind?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	64
	☐	☐	☐	☐	☐	
Ob ich dem Arzt vertrauen kann?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	65
	☐	☐	☐	☐	☐	
Was wohl andere von mir erwarten?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	66
	☐	☐	☐	☐	☐	
Ob es wichtig ist, immer alles ganz genau zu verstehen?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	67
	☐	☐	☐	☐	☐	
Welche Lebensaufgabe ich wohl in dieser Erkrankung sehen könnte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	68
	☐	☐	☐	☐	☐	
Wie sicher man sich bei der Entscheidung über eine medizinische Maßnahme sein sollte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	69
	☐	☐	☐	☐	☐	

7

Ob ich kommende medizinische Maßnahmen (Behandlung/Nachsorge) wohl bewältigen werde?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	70
	☐	☐	☐	☐	☐	
Ob ich die Bedeutung von medizinischen Aussagen und Befunden richtig verstehe?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	71
	☐	☐	☐	☐	☐	
Inwieweit ich meine Wünsche gegenüber dem Behandlungsteam äußern sollte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	72
	☐	☐	☐	☐	☐	
Mit welcher Einstellung ich meine Erkrankung am ehesten günstig beeinflussen könnte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	73
	☐	☐	☐	☐	☐	
Ob der Arzt wohl ausreichend kompetent ist?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	74
	☐	☐	☐	☐	☐	
Ob ich es wohl rechtzeitig bemerke, wenn meine Erkrankung weiter fortschreitet?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	75
	☐	☐	☐	☐	☐	
Wie detailliert ich über meine Erkrankung informiert sein sollte?						
Diese Ungewissheit beschäftigt mich	gar nicht	etwas	mittel	ziemlich	sehr	76
	☐	☐	☐	☐	☐	

Anhang D

Die QUICC-Version aus Validierungsstudie 1

Qualitäten von Ungewissheiten in chronischen Krankheiten

Bogen für Betroffene von Krebserkrankungen (QUICC-CA)

Datum:	Code:

Liebe Patientin, lieber Patient,

Betroffene einer Krebserkrankung haben mit unterschiedlichen Ungewissheiten zu tun. Fragen im Zusammenhang mit der Erkrankung lassen sich manchmal nicht beantworten. So wissen Sie womöglich nicht genau, wie Ihre Erkrankung weiter verlaufen wird. Oder Sie fragen sich, warum gerade Sie erkrankt sind.
Von solchen Ungewissheiten handeln die folgenden Fragen.

Bevor Sie beginnen, machen Sie sich bitte anhand dieses **Beispiels** mit der Frageform vertraut:

Ob ich meinem Arzt wohl vertrauen kann?

	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐

Hier geht es **nicht** darum, wie sehr Sie Ihrem Arzt vertrauen, oder wie wichtig Ihnen dieses Vertrauen ist.

Uns interessiert, **ob Sie mit der Suche nach der Antwort auf diese Frage beschäftigt sind.**

Vielen Dank für Ihre Mühe!

2

Inwieweit beschäftigen Sie zur Zeit die folgenden Fragen?

	gar nicht	etwas	mittel	ziemlich	sehr	
Ob mich meine Angehörigen wohl auch unterstützen, wenn es mir mal schlechter geht?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	01
Wie wohl mein augenblicklicher Krankheitszustand ist?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	02
Ob ich wohl künftig die Auswirkungen meiner Erkrankung bewältigen kann?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	03
Welchen Unterschied es gemacht hätte, wenn die Erkrankung früher oder später festgestellt worden wäre?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	04
Ob ich mich intensiver mit Informationen über meine Erkrankung auseinandersetzen sollte?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	05
Woran ich merken kann, dass die Behandlungsmethode gut ist?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	06
Wie die Erkrankung meine sozialen Beziehungen beeinflusst?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	07
Ob ich verstehe, was mir der Arzt über meine Situation mitteilen will?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	08
Ob ich meinen finanziellen Verpflichtungen trotz der Erkrankung nachkommen kann?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	09
Was wohl Ursachen für meine Erkrankung sind?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	10

Ob der Arzt ausreichend über meine Erkrankung bescheid weiß?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	11
Was wohl in Fällen wie meinem die optimale Behandlung wäre?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	12
Was andere denken mögen, wenn sie Anzeichen meiner Erkrankung bemerken?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	13
Wie sich wohl meine Erkrankung weiter entwickeln wird?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	14
Ob ich wohl meinem Alltag künftig trotz Erkrankung gewachsen bin?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	15
Wie die Erkrankung wohl mit meinem bisherigen Leben zusammenhängt?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	16
Was für Rückschlüsse über meinen Zustand ich aus dem Verhalten des Arztes ziehen kann?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	17
Inwieweit ich meinem Arzt Bedenken mitteilen sollte?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	18
Was ich wohl zusätzlich zu meiner medizinischen Behandlung noch tun kann?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	19
Wie viel Belastung durch die Erkrankung meine Mitmenschen ertragen können?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	20
Wie sich mein Körper im Laufe der Erkrankung wohl noch verändern wird?						
	gar nicht	**etwas**	**mittel**	**ziemlich**	**sehr**	
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	21

4

	gar nicht	etwas	mittel	ziemlich	sehr	
Ob mich der Arzt mit meinen Sorgen ernst nimmt?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	22
Ob ich den Erwartungen anderer trotz meiner Erkrankung gerecht werde?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	23
Ob bestimmte Belastungen in meinem Leben die Krankheit mit bedingt haben?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	24
Inwieweit ich über meine Behandlung mitentscheiden sollte?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	25
Wie ich das Verhältnis von Nutzen und Risiken einer medizinischen Maßnahme bewerten soll?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	26
Bis zu welchem Punkt meine mir liebsten Menschen hinter mir stehen?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	27
Was es wohl bedeutet, wenn Aussagen oder Befunde über meinen Zustand einander widersprechen?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	28
Ob ich auch mit meiner Erkrankung das erreichen kann, was ich mir vornehme?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	29
Warum ich wohl gerade diese Krankheit bekommen habe?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	30
Ob sich wohl seit der letzten Untersuchung etwas in meinem Körper verändert hat?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	31
Ob ich auch menschlich bei meinem Arzt gut aufgehoben bin?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	32

	gar nicht	etwas	mittel	ziemlich	sehr	
Wie ich mein Leben verändern kann, um meine Erkrankung günstig zu beeinflussen?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	33
Wie andere mit meiner Diagnose zurechtkommen?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	34
Wie nachdrücklich ich in Gesprächen mit dem Arzt meinen eigenen Standpunkt vertreten soll?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	35
Was allgemeine Statistiken für meinen persönlichen Krankheitsverlauf bedeuten?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	36
Welchen körperlichen Anforderungen ich trotz meiner Erkrankung gewachsen bin?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	37
Warum gerade ich krank geworden bin?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	38
Ob der Arzt den neuesten Forschungsstand zu meiner Erkrankung kennt?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	39
Wie viel Aufmerksamkeit für meinen Körper mir wohl gut tut?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	40
Was wohl bestimmte Beschwerden über meinen Krankheitszustand aussagen?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	41
Inwieweit ich meine Wünsche gegenüber dem Behandlungsteam äußern sollte?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	42
Was wohl meine Mitmenschen von mir erwarten?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	43

6

	gar nicht	etwas	mittel	ziemlich	sehr	
Ob ich wirklich alles richtig verstanden habe hinsichtlich meiner Erkrankung und der Folgen?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	44
Ob ich künftige medizinische Maßnahmen (Behandlung/Nachsorge) bewältigen werde?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	45
Ob mir der Arzt die ganze Wahrheit sagt?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	46
Ob ich es wohl rechtzeitig bemerke, wenn meine Erkrankung fortschreitet?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	47
Welche Lebensaufgabe ich wohl in dieser Erkrankung sehen könnte?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	48
Wie genau ich über meine Erkrankung informiert sein sollte?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	49
Wie sicher man sich bei der Entscheidung über eine medizinische Maßnahme sein sollte?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	50
Ob ich die Bedeutung von medizinischen Aussagen und Befunden richtig verstehe?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	51
Ob sich der Arzt wohl genug Mühe mit mir gibt?						
Diese Frage beschäftigt mich zur Zeit	☐	☐	☐	☐	☐	52

Anhang E

Die QUICC-Version ab Validierungsstudie 2

Liebe Patientin, lieber Patient,

Betroffene einer chronischen Erkrankung haben mit unterschiedlichen Ungewissheiten zu tun. Fragen im Zusammenhang mit der Erkrankung lassen sich manchmal nicht beantworten. So wissen Sie womöglich nicht genau, wie Ihre Erkrankung weiter verlaufen wird. Oder Sie fragen sich, warum gerade Sie erkrankt sind.

Von solchen „Ungewissheiten“ handeln die folgenden Fragen.

Bevor Sie beginnen, machen Sie sich bitte anhand dieses **Beispiels** mit der Frageform vertraut:

Mich beschäftigt zur Zeit die Frage...					
	gar nicht				**sehr**
... , wie ich künftig mit Einschränkungen durch meine Erkrankung zurecht komme.	⓪	①	②	③	④

Hier geht es **nicht** darum, wie wichtig es Ihnen ist, zurecht zu kommen. Uns interessiert, **ob Sie mit der Suche nach der Antwort auf diese Frage beschäftigt sind.**

Am Ende jeder Seite werden Sie zusätzlich aufgefordert, die Frage zu markieren, über die Sie am meisten nachgrübeln.

Vielen Dank für Ihre Mühe!

3

Mich beschäftigt zur Zeit die Frage...

	gar nicht				sehr	
... , was andere denken mögen, wenn sie Anzeichen meiner Erkrankung bemerken.	⓪	①	②	③	④	01
... , wie mein augenblicklicher Krankheitszustand ist.	⓪	①	②	③	④	02
... , ob ich mich intensiver mit Informationen über meine Erkrankung auseinandersetzen sollte.	⓪	①	②	③	④	03
... , ob ich wohl künftig die Auswirkungen meiner Erkrankung bewältigen kann.	⓪	①	②	③	④	04
... , welchen Unterschied es gemacht hätte, wenn die Erkrankung früher oder später festgestellt worden wäre.	⓪	①	②	③	④	05

Mich beschäftigt zur Zeit die Frage...

	gar nicht				sehr	
... , ob ich verstehe, was mir der Arzt über meine Situation mitteilen will.	⓪	①	②	③	④	06
... , ob mich der Arzt mit meinen Sorgen ernst nimmt.	⓪	①	②	③	④	07
... , was in Fällen wie meinem die optimale Behandlung wäre.	⓪	①	②	③	④	08
... , wie die Erkrankung meine sozialen Beziehungen beeinflusst.	⓪	①	②	③	④	09
... , wie sich meine Erkrankung weiter entwickeln wird.	⓪	①	②	③	④	10

Welche Frage auf dieser Seite beschäftigt Sie am stärksten?
Bitte am Rand ankreuzen.

Mich beschäftigt zur Zeit die Frage...						
	gar nicht				**sehr**	
... , inwieweit ich meinem Arzt Bedenken mitteilen sollte.	⓪	①	②	③	④	11
... , ob ich meinen finanziellen Verpflichtungen trotz der Erkrankung nachkommen kann.	⓪	①	②	③	④	12
... , wie die Erkrankung mit meinem bisherigen Leben zusammenhängt.	⓪	①	②	③	④	13
... , welche Rückschlüsse über meinen Zustand ich aus dem Verhalten des Arztes ziehen kann.	⓪	①	②	③	④	14
... , ob ich auch menschlich bei meinem Arzt gut aufgehoben bin.	⓪	①	②	③	④	15

Mich beschäftigt zur Zeit die Frage...						
	gar nicht				**sehr**	
... , was ich wohl zusätzlich zu meiner medizinischen Behandlung noch tun kann.	⓪	①	②	③	④	16
... , ob mich meine Angehörigen wohl auch unterstützen, wenn es mir mal schlechter geht.	⓪	①	②	③	④	17
... , ob sich seit der letzten Untersuchung etwas in meinem Körper verändert hat.	⓪	①	②	③	④	18
... , inwieweit ich über meine Behandlung mitentscheiden sollte.	⓪	①	②	③	④	19
... , ob ich meinem Alltag künftig trotz Erkrankung gewachsen bin.	⓪	①	②	③	④	20

Welche Frage auf dieser Seite beschäftigt Sie am stärksten?
Bitte am Rand ankreuzen.

6

Mich beschäftigt zur Zeit die Frage...

	gar nicht				sehr	
... , wie ich das Verhältnis von Nutzen und Risiken einer medizinischen Maßnahme bewerten soll.	⓪	①	②	③	④	21
... , ob der Arzt den neuesten Forschungsstand zu meiner Erkrankung kennt.	⓪	①	②	③	④	22
... , was es wohl bedeutet, wenn Aussagen oder Befunde über meinen Zustand einander widersprechen.	⓪	①	②	③	④	23
... , ob bestimmte Belastungen in meinem Leben die Krankheit mit bedingt haben.	⓪	①	②	③	④	24
... , wie ich mein Leben verändern kann, um meine Erkrankung günstig zu beeinflussen.	⓪	①	②	③	④	25

Mich beschäftigt zur Zeit die Frage...

	gar nicht				sehr	
... , ob mir der Arzt die ganze Wahrheit sagt.	⓪	①	②	③	④	26
... , was allgemeine Statistiken für meinen persönlichen Krankheitsverlauf bedeuten.	⓪	①	②	③	④	27
... , warum ich wohl gerade diese Krankheit bekommen habe.	⓪	①	②	③	④	28
... , ob ich auch mit meiner Erkrankung das erreichen kann, was ich mir vorgenommen habe.	⓪	①	②	③	④	29
... , wie nachdrücklich ich in Gesprächen mit dem Arzt meinen eigenen Standpunkt vertreten soll.	⓪	①	②	③	④	30

Welche Frage auf dieser Seite beschäftigt Sie am stärksten?
Bitte am Rand ankreuzen.

Mich beschäftigt zur Zeit die Frage...						
	gar nicht				**sehr**	
... , was wohl bestimmte Beschwerden über meinen Krankheitszustand aussagen.	⓪	①	②	③	④	31
... , wie viel Belastung durch die Erkrankung meine Mitmenschen ertragen können.	⓪	①	②	③	④	32
... , wie sicher man sich bei der Entscheidung über eine medizinische Maßnahme sein sollte.	⓪	①	②	③	④	33
... , ob sich der Arzt genug Mühe mit mir gibt.	⓪	①	②	③	④	34
... , ob ich wirklich alles richtig verstanden habe hinsichtlich meiner Erkrankung und deren Folgen.	⓪	①	②	③	④	35

Mich beschäftigt zur Zeit die Frage...						
	gar nicht				**sehr**	
... , warum gerade ich krank geworden bin.	⓪	①	②	③	④	36
... , welchen körperlichen Anforderungen ich trotz meiner Erkrankung gewachsen bin.	⓪	①	②	③	④	37
... , wie genau ich über meine Erkrankung informiert sein sollte.	⓪	①	②	③	④	38
... , ob ich es rechtzeitig bemerke, wenn meine Erkrankung fortschreitet.	⓪	①	②	③	④	39
... , bis zu welchem Punkt meine mir liebsten Menschen hinter mir stehen.	⓪	①	②	③	④	40
... , welche Lebensaufgabe ich wohl in dieser Erkrankung sehen könnte.	⓪	①	②	③	④	41

Welche Frage auf dieser Seite beschäftigt Sie am stärksten?
Bitte am Rand ankreuzen.

8

Anhang F

Die Ungewissheitstoleranzskala (UGTS)

Abschließend bitten wir Sie, zu verschiedenen Themen Stellung zu nehmen. Entscheiden Sie, inwieweit jede Aussage auf Sie persönlich zutrifft.

Bitte lassen Sie keine Frage aus, und setzen Sie jeweils nur ein Kreuz.

Ich probiere gerne Dinge aus, auch wenn nicht immer etwas dabei herauskommt.	stimmt überhaupt nicht	stimmt weitgehend nicht	stimmt eher nicht	stimmt ein wenig	stimmt weitgehend	stimmt genau
01	☐	☐	☐	☐	☐	☐
Ich beschäftige mich nur mit Aufgaben, die lösbar sind.	stimmt überhaupt nicht	stimmt weitgehend nicht	stimmt eher nicht	stimmt ein wenig	stimmt weitgehend	stimmt genau
02	☐	☐	☐	☐	☐	☐
Ich mag es, wenn unverhofft Überraschungen auftreten.	stimmt überhaupt nicht	stimmt weitgehend nicht	stimmt eher nicht	stimmt ein wenig	stimmt weitgehend	stimmt genau
03	☐	☐	☐	☐	☐	☐
Ich lasse die Dinge gerne auf mich zukommen.	stimmt überhaupt nicht	stimmt weitgehend nicht	stimmt eher nicht	stimmt ein wenig	stimmt weitgehend	stimmt genau
04	☐	☐	☐	☐	☐	☐
Ich habe es gerne, wenn die Arbeit gleichmäßig verläuft.	stimmt überhaupt nicht	stimmt weitgehend nicht	stimmt eher nicht	stimmt ein wenig	stimmt weitgehend	stimmt genau
05	☐	☐	☐	☐	☐	☐
Ich warte geradezu darauf, dass etwas Aufregendes passiert.	stimmt überhaupt nicht	stimmt weitgehend nicht	stimmt eher nicht	stimmt ein wenig	stimmt weitgehend	stimmt genau
06	☐	☐	☐	☐	☐	☐
Wenn um mich herum alles drunter und drüber geht, fühle ich mich so richtig wohl.	stimmt überhaupt nicht	stimmt weitgehend nicht	stimmt eher nicht	stimmt ein wenig	stimmt weitgehend	stimmt genau
07	☐	☐	☐	☐	☐	☐
Ich weiß gerne, was auf mich zukommt.	stimmt überhaupt nicht	stimmt weitgehend nicht	stimmt eher nicht	stimmt ein wenig	stimmt weitgehend	stimmt genau
08	☐	☐	☐	☐	☐	☐

Anhang G

Zusammenfassung

Einführung. Das Paradigma des ‚Shared Decision Making' (SDM) hat sich aus rechtsethischen wie theoretischen Erwägungen heraus zum Ideal der Entscheidungsfindung zwischen Arzt und Patient bei der Auswahl einer Therapie entwickelt. Beim SDM erfolgt ein wechselseitiger Informationsaustausch über die medizinischen Sachverhalte sowie die persönlichen Präferenzen der Beteiligten. Auf dieser Basis geteilten Wissens wird im Konsens eine Therapieentscheidung getroffen und umgesetzt.

Bislang fehlen überzeugende Theorien über SDM und die dabei beteiligten psychologischen Prozesse. Infolgedessen ermangelt es ferner geeigneter Messverfahren zu den postulierten Effekten von SDM sowie der vermittelnden Variablen, so dass dessen klinischer Nutzen nicht zweifelsfrei belegt werden kann.

Theoretischer Hintergrund. SDM gilt insbesondere dann als indiziert, wenn mehrere Entscheidungsoptionen plausibel begründet werden können. Diese Situation resultiert häufig aus dem Mangel an tragfähiger Evidenz zu den Handlungsalternativen. Doch selbst bei fundierten durchschnittlichen Heilungswahrscheinlichkeiten kann nicht mit Sicherheit vorhergesagt werden, welchen Nutzen und welchen Schaden der individuelle Patient von der Behandlung haben wird. Daher erscheint die Vorstellung illusorisch, eine Entscheidung auf dem Wege der Herstellung von Gewissheit herbeiführen zu wollen. Vielmehr besteht die Herausforderung beim SDM darin, sich mit verschiedenen Arten von *Ungewissheit* auseinander zu setzen mit dem Ziel, diese zu differenzieren und zu ordnen. Diese Ordnung ermöglicht einerseits einen *jeweils* angemessen Umgang und andererseits eine elaborierte Entscheidung unter Berücksichtigung der verschiedenen Ungewissheiten. Aus dieser Überlegung wurde

als Outcomeparameter für SDM die Differenzierung von entscheidungsrelevanten Ungewissheiten abgeleitet.

Methodik. Zur Präzisierung und Überprüfung dieses Konzepts von SDM wurde zunächst anhand von Tiefeninterviews mit Krebspatienten ergebnisoffen ermittelt, welche verschiedenen Arten von Ungewissheit erlebt werden. Aus den Transkripten wurden durch einen Grounded Theory-Prozess acht Ungewissheitskategorien abstrahiert. Anschließend erfolgte zur Prüfung der Nachvollziehbarkeit dieses Kategoriensystems eine Bestimmung der Interraterreliabilität bei der Zuordnung eines weiteren Tiefeninterviews. Zu den acht Kategorien wurde ein Itempool für den Fragebogen QUICC (Qualities of Uncertainty in Chronic Conditions) gebildet, der durch eine Think aloud-Prozedur auf zunächst 52 Items reduziert wurde. Zur Bestimmung der Testgüte wurde dieser Fragebogen nachfolgend – begleitet von kleineren Modifikationen und Kürzungen – verschiedenen Populationen vorgegeben. Abschließend erfolgte eine SDM-Interventionsstudie, bei der die postulierte Differenzierung im Ungewissheitserleben als Zunahme von Interskalenvarianz operationalisiert wurde.

Ergebnisse. Die Interraterreliabilität lag bei 0,54 (Fleiss-Koeffizient) bzw. 0,55 (medianes Cohens κ). Die Think aloud-Interviews belegten, dass der QUICC die entscheidungsrelevanten Ungewissheiten disjunktiv und exhaustiv auf verständliche Weise abfragt, wobei noch einige Anregungen zu Modifikationen gewonnen werden konnten.

In den nachfolgenden multizentrischen Validierungsstudien ($N = 616$) bei den Indikationen Krebs, Chronisch-entzündliche Darmkrankheiten, Chronischer Schmerz sowie Depression erwies sich der QUICC als reliabel ($0,76 <$ Cronbachs $\alpha < 0,95$). Die Items wiesen durchweg mittlere Schwierigkeiten und hohe Trennschärfen auf.

Im Rahmen der Betrachtung der externen Validität zeigte sich erwartungskonform ein intensiveres Ungewissheitserleben bei Akut- gegenüber Nachsorgepatienten ($p < 0,02$) sowie eine negative Korrelation mit der individuellen Ungewissheitstoleranz ($p < 0,05$).

In Bezug auf die interne Validität des QUICC war das postulierte achtdimensionale Testmodell trotz hoher Skaleninterkorrelationen ($0,56 < r < 0,87$) in einem pro-

babilistischen Modelltest einer eindimensionalen Alternative gemäß AIC und BIC überlegen.

In der abschließenden SDM-Interventionsstudie ($N = 29$) konnte mit dem QUICC neben einer generellen Reduktion im Ungewissheitserleben ($0,01 < p < 0,07$) die erwartete Zunahme an Interskalenvarianz belegt werden ($p = 0,02$).

In allen Untersuchungen erwies sich der QUICC zudem als recht praktikabel und ökonomisch.

Diskussion. Die vorliegenden Ergebnisse unterstützen sowohl die theoretische Konzeption von SDM als Verhandlung von Ungewissheiten als auch die messtheoretische Qualität des QUICC. Damit wurden wesentliche Voraussetzungen für die überzeugende Messung der Effekte von SDM geschaffen. Gleichwohl gilt es in zukünftigen Untersuchungen zu klären, inwieweit neben SDM auch andere Kommunikationsstile zur Differenzierung des Ungewissheitserlebens beitragen können. In diesem Fall eignete sich der QUICC nicht als spezifischer Outcomeparameter für SDM, sondern könnte ganz allgemein Aufschluss darüber geben, ob sich durch eine Risikokommunikation die Repräsentation der entscheidungsrelevanten Ungewissheiten beim Patienten so verändern ließ, dass ihm eine elaborierte Entscheidung möglich ist.

Anhang H

Lebenslauf

Persönliche Daten

- Name: Friedemann Geiger
- Geburtsdatum: 04. März 1977
- Geburtsort: Buchholz in der Nordheide
- Familienstand: verheiratet
- Nationalität: deutsch
- Adresse: Weißenburgstraße 11, 24116 Kiel

Schulausbildung

- 1983–1987: Grundschule Klein Hehlen, Celle
- 1987–1989: Orientierungsstufe Klein Hehlen, Celle
- 1989–1996: Hermann-Billung-Gymnasium, Celle
- Juni 1996: Erlangung der allgemeinen Hochschulreife (Abitur)

Hochschulstudium

- 1997–1998: Christian-Albrechts-Universität zu Kiel, Magisterstudiengang Philosophie
- 1998–2002: Christian-Albrechts-Universität zu Kiel, Diplomstudiengang Psychologie
- September 2002: Diplom

Promotion

- 2004–2006: Promotionsstipendium des Evangelischen Studienwerks Villigst e.V.
  Betreuung durch Prof. Dr. Jürgen Kriz, Universität Osnabrück

Anhang I

Erklärung

Hiermit erkläre ich, Friedemann Geiger, dass ich die Dissertation „Shared Decision Making als Verhandlung von Ungewissheiten“ selbstständig verfasst und keine Hilfsmittel außer den angegebenen verwendet habe. Diese Arbeit war bisher weder teilweise noch ganz Gegenstand eines Prüfungsverfahrens an einer wissenschaftlichen Hochschule.

(Friedemann Geiger)