

## **IV DISKUSSION DER ERGEBNISSE**

In dieser Arbeit wurden zunächst theoretische Perspektiven und empirische Befunde zur Lebensqualität im Alter dargestellt. Vor diesem Hintergrund wurde eine qualitative, empirische Vorstudie (Studie 1) zum Entwurf eines Fragebogens durchgeführt. Darauf aufbauend wurde anschließend ein Fragebogen zur Lebenszufriedenheit multimorbider älterer Menschen entwickelt und in einer Pilotuntersuchung (Studie 2) hinsichtlich verschiedener Validitätsaspekte getestet. Die erhobenen Daten wurden sowohl unter quantitativen als auch qualitativen Gesichtspunkten ausgewertet: Es fand eine explorative Abschätzung der Konstruktvalidität und Inhaltsvalidität des neu entwickelten Fragebogens zur Lebensqualität multimorbider älterer Menschen (FLQM) statt; aus Daten von Studie 1 und den qualitativen Angaben im Rahmen des FLQM in Studie 2 wurden mit Hilfe von qualitativen Inhaltsanalysen Kategoriensysteme von subjektiv bedeutsamen Bereichen der Lebensqualität abgeleitet. Neben einer Gesamtauswertung wurde möglichen inhaltlichen Unterschieden zwischen Altersgruppen, Geschlechtern und Bildungsniveaus nachgegangen.

Zunächst werden die Ergebnisse mit Bezug auf die Konstrukt- und Inhaltsvalidität des FLQM diskutiert, anschließend Verständlichkeit und Durchführbarkeit. Das Kategoriensystem der subjektiv für Lebensqualität bedeutsamen Bereiche wird besprochen und mögliche Schlussfolgerungen über die Entwicklung von Komplexität und Heterogenität der subjektiven Definitionen von Lebensqualität im Alter wird diskutiert. Schließlich werden Einschränkungen und Grenzen einerseits der Durchführung der Studien 1 und 2 insgesamt als auch des FLQM im Besonderen diskutiert.

### **4.1 Konstruktvalidität des FLQM**

Die empirischen Befunde zur Validität des FLQM sind ermutigend für eine Beibehaltung des Ansatzes. Dies gilt sowohl für die Verteilungsmerkmale des FLQM als auch für seine Zusammenhänge mit anderen Skalen. Die empirischen Werte sind normalverteilt, gleichzeitig ist die Verteilung leicht nach rechts, also zum positiven Pol der Skala, verschoben. Das bedeutet, dass die Teilnehmer trotz ihres eingeschränkten Gesundheitszustandes auf dem FLQM eine relativ hohe Zufriedenheit mit den ihnen im Leben wichtigen Bereichen berichten. Dieser Befund steht im Einklang mit den Kenntnissen über

das „Paradox des subjektiven Wohlbefindens“ (vgl. Staudinger, 2000). Erklärlich ist diese tendenzielle „Überschätzung“ der eigenen Zufriedenheit beispielsweise über vornehmlich akkomodative Anpassungsmechanismen, welche eine Abwertung oder Umdeutung selbstwert- oder zufriedenheitsabträglicher Informationen begünstigen (Brandtstädter & Rothermund, 1994). Im Falle der hier betrachteten Population gesundheitlich relativ stark beeinträchtigter Personen bedeutet dies, dass Einschränkungen in bestimmten Lebensbereichen entweder dazu führen können, dass diese Bereiche in ihrer Bedeutsamkeit abgewertet werden oder, dass ein geringerer Grad von „Zielerreichung“ als ausreichend für eine positive Bewertung angesehen wird. Dabei spielen soziale und temporale Vergleichsprozesse eine große Rolle. Soziale Vergleiche beziehen sich beispielsweise auf andere Menschen gleichen Alters mit geringeren Fähigkeiten, temporale Vergleiche beispielsweise auf das Selbst in jüngeren Jahren, also mit altersbedingt höheren Fähigkeiten. Durch diese Kontrastierungen, welche die eigenen Fähigkeiten und Möglichkeiten an stärker benachteiligten oder begünstigten Vergleichspersonen messen, wird vor allen Dingen die Veränderungen interner Standards gefördert (z. B. Beaumont & Kenealy, 2004; Filipp & Ferring, 1998).

Brickman und Campbell (1971) postulieren im Kontext der *adaptation level theory* eine grundlegende Veranlagung des Selbst, eine moderat positive Affektlage beizubehalten und prägen für die damit verbundenen Anpassungsprozesse den Begriff „*hedonic treadmill*“: Innerhalb eines gewissen Rahmens passt sich der Mensch an alle Veränderungen, Erfahrungen und Ereignisse, welche die affektive Balance in positiver oder negativer Richtung aus dem Gleichgewicht bringen, an und bewegt sich auf ein affektives Equilibrium zu (vgl. Brickman, Coates & Janoff-Bulman, 1978). Im Rahmen der Interpretation von Studie 2 bedeutet diese Annahme folgendes: Da viele der typischen Alterskrankheiten chronisch sind (vgl. Böhmer, 2003), waren die Teilnehmer zum Befragungszeitpunkt vermutlich keinen starken *neuen* Beeinträchtigungen und Einschränkungen unterworfen. Vielmehr hatten sie vermutlich bereits über einen längeren Zeitraum die Möglichkeit, sich an ihre Situation anzupassen und im Laufe der Krankheitsentwicklung selbstwert- und zufriedenheitsschützende Strategien zu entwickeln. Dafür spricht auch, dass es keine signifikanten Zusammenhänge der FLQM-Werte mit der Anzahl der selbstberichteten Erkrankungen oder dem Alter gibt. Die empirische Werteverteilung des FLQM zeigt im Gegensatz zu vielen Maßen der gesundheitsbezogenen Lebensqualität, wie sie in der Medizin gebräuchlich sind, keinen Bodeneffekt,

sondern ist um den über dem theoretischen Skalenmittel liegenden Mittelwert von  $M = 4,04$  annähernd normalverteilt. So wie die schiefe Verteilung für die inhaltliche Gültigkeit des FLQM spricht, weist die Normalverteilung auf die Diskriminierungsfähigkeit des FLQM innerhalb der Studienpopulation hin: Lebensqualität, gemessen als subjektives Konstrukt, verhält sich anders als Lebensqualität, gemessen als normativer Grad der Beeinträchtigung innerhalb festgelegter Fähigkeits- und Erlebensbereiche. Die Inhaltsvalidität des FLQM wird also vor dem Hintergrund theoretischer und empirischer Literatur dadurch gestützt, dass die vergleichsweise ungünstigen objektiven Bedingungen in der untersuchten Population multimorbider Älterer keinen starken direkten Einfluss auf die Bewertung der Lebensqualität zu haben scheinen.

Allgemeine Lebenszufriedenheit stellt vor dem Hintergrund des zu Grunde gelegten Modells von Campbell et al. (1976) eines der wichtigsten Konstrukte zur Validierung des FLQM dar. Die hohe Korrelation zwischen dem FLQM-Gesamtscore und dem Item zur allgemeinen Lebenszufriedenheit weist deutlich auf die Gültigkeit des FLQM hin. Gleichzeitig wirft die Tatsache, dass eine einfache Frage offenkundig einen großen Anteil der Information des viel komplexeren Fragebogens aufklären kann die Frage auf: Wie ist ein im Vergleich zu einer Einzelfrage nicht unbeträchtlicher Mehraufwand an Zeit und Ressourcen im Falle des FLQM zu rechtfertigen? Bowling (2005) vertritt beispielsweise die Auffassung, dass Einzelitems zur Erfassung von Konstrukten längeren Fragebögen grundsätzlich vorzuziehen seien, sofern sie psychometrischen Anforderungen genügen, sie argumentiert dabei in erster Linie ökonomisch. Bei der Abwägung spielen jedoch verschiedenste Faktoren eine Rolle: Einerseits natürlich – vordringlich in klinischen Kontexten – der tatsächliche erforderliche Zeit- und Personalaufwand sowie die mögliche Belastung von Patienten durch die Länge der Befragung, andererseits jedoch auch inhaltliche Merkmale des Fragebogens. Zunächst sollte konstatiert werden, dass der Anteil der gemeinsamen Varianz von FLQM und der Frage zur globalen Lebenszufriedenheit nicht überschätzt werden darf: Auch die hochsignifikante Korrelation von  $r = 0,63$  verweist auf eine Varianzaufklärung von lediglich knapp 40 % durch die Frage zur globalen Lebenszufriedenheit alleine. Auf welche anderen Variablen möglicherweise ebenfalls größere Varianzanteile entfallen, ließe sich nur im Rahmen einer multiplen Regressionsanalyse beantworten – angesichts der gegebenen Stichprobengröße von  $N = 44$  ist die Durchführung jedoch nicht sinnvoll.

Das bedeutendste inhaltliche Argument zu Gunsten des FLQM ist, dass mittels

dieses Fragebogens weit mehr Informationen erfasst werden, als ein reines *outcome*. Vielmehr werden vielschichtige Daten zur individuellen, subjektiven Konzeptualisierung von Lebensqualität generiert. Diese können auf der einen Seite quantitativ genutzt werden, beispielsweise um längsschnittliche Veränderungen der Lebensqualität differenzierter aufschlüsseln und bewerten zu können: Durch die Explikation der zur Bildung des Gesamturteils beitragenden Komponenten – den bereichsspezifischen Zufriedenheiten und den relativen Bedeutsamkeiten/Gewichten der Einzelbereiche – lassen sich Annahmen zu Ursachen von Veränderungen über die Zeit differenzierter überprüfen als bei Indices, welche lediglich aus *einem* Zahlenwert bestehen<sup>1</sup>. Auf der anderen Seite werden über den FLQM auch qualitativ-inhaltliche Aspekte der subjektiven Konzepte von Lebensqualität zugänglich gemacht. Diese können sowohl Ansatzpunkte für eine stärker patientenorientierte Behandlung und psychosoziale Versorgung liefern als auch wiederum Rückschlüsse auf diejenigen Prozesse ermöglichen, welche längsschnittliche Veränderungen in der berichteten Lebensqualität bedingen. Verschiedene Untersuchungen zur Akzeptanz individualisierter Maße lieferten Hinweise darauf, dass gerade die über eine Quantifizierung der Lebensqualität hinausgehenden, qualitativen Informationen von professionellen Anwendern (z. B. Ärzte oder Pflegefachkräfte) als erweiterte Grundlage zur Pflege- und Behandlungsplanung sehr geschätzt werden (Macduff, 2000; Mountain et al., 2004; Patel et al., 2003).

In der Population der alten und mehrfach erkrankten Menschen, so wurde argumentiert, besteht vermutlich auf Grund der heterogenen Profile von Beschwerden, Fähigkeiten und Möglichkeiten eine große interindividuelle Variabilität der Bereiche, die zur Lebensqualität beitragen. Eine einzelne Frage scheint gerade vor dem Hintergrund eines praxis- und interventionsorientierten Anspruchs an ein Messinstrument zu wenig, um dieser Heterogenität gerecht zu werden. Der Mehraufwand an Zeit und Ressourcen gegenüber einer einzelnen Frage zur globalen Lebenszufriedenheit sollte zu Gunsten der Fülle und Vielschichtigkeit der mit dem FLQM gewonnenen Informationen in Kauf genommen werden.

Die mittleren Zusammenhänge mit den beiden Subskalen des PGCMS „Lebenszufriedenheit“ und „Zufriedenheit mit dem eigenen Alter“ sowie mit dem

---

<sup>1</sup> Während für andere patientengenerierte, individualisierte Maße der Lebensqualität bereits Studien vorliegen, welche deren Bedeutung für die Evaluation eines *response shift* verdeutlichen (Schwartz & Sprangers, 1999; Sprangers & Schwartz, 1999; vgl. Abschnitt 4.1), steht dieser Nachweis für den FLQM selbstverständlich noch aus.

PGCMS-Gesamtscore sprechen erwartungskonform ebenfalls für die Konstruktvalidität des FLQM. Auch der nicht signifikante und ohnehin sehr geringe Zusammenhang mit negativem Affekt (gemessen mit dem PANAS) entspricht den Erwartungen, jedoch ist der hoch signifikante Zusammenhang zwischen FLQM-Score und positivem Affekt (gemessen mit dem PANAS) in Hinblick auf die Validität des FLQM ambivalent zu interpretieren. Das Modell der Urteilsbildung von Schwarz und Strack (1991) legt nahe, dass bereichsspezifische Urteile relativ unabhängig von der momentanen Affektlage getroffen werden, da differenzierte kognitive Prozesse in Bezug auf spezifische Lebensbereiche leicht zugänglich sind. Im Gegensatz dazu stehen globale Urteile, die deutlich stärker auf Grundlage affektiver Informationen getroffen werden. Zahlreiche Untersuchungen belegen diese Unterscheidung zwischen kognitiver („Zufriedenheit“) und affektiver Komponente („Emotion“) des Wohlbefindens (Diener & Emmons, 1984; Lawton et al., 1992b; Lucas, Diener & Suh, 1996; Watson et al., 1988).

Der enge Zusammenhang zwischen positivem Affekt und FLQM-Gesamtscore wirft die Frage auf, ob für den FLQM die Abkopplung der bereichsspezifischen Urteile von der momentanen Affektlage tatsächlich im gewünschten Umfang gelungen ist. Zwei Interpretationsmöglichkeiten bieten sich an: Die erste Interpretation besteht darin, dass der FLQM, entgegen der Intention, relativ sensitiv für Stimmungseinflüsse ist und somit sowohl kognitive *als auch* affektive Komponenten der Lebensqualität erfasst. Der nicht signifikante Zusammenhang mit negativem Affekt stünde dieser Interpretation nicht unbedingt entgegen, da erstens positiver und negativer Affekt als unabhängige Konstrukte angesehen werden und zweitens die empirische Verteilung von negativem Affekt in dieser Studie ausgesprochen homogen war und sehr wenig Variationsbreite aufwies – das dritte Quartil der Verteilung (Mittelwert  $M = 1,46$ ) lag bei nur 1,88 auf der Skala von 1 bis 5, die Standardabweichung bei  $SD = 0,52$ . Der geringe Zusammenhang zwischen FLQM und negativem Affekt kann also unter Umständen als Artefakt der Verteilungsform interpretiert werden. Die zweite Interpretation steht in Zusammenhang mit dem Verständnis des PANAS seitens der Studienteilnehmer: Obwohl die momentane Affektlage als Referenzgröße in der Instruktion genannt wird („*Wie fühlen Sie sich im Moment?*“), ist denkbar, dass zumindest einige der Teilnehmer die allgemeine Affektlage als Grundlage der Beantwortung verwendet hat. Aus mündlichen Äußerungen während der Erhebung und Nachfragen seitens der Teilnehmer kann geschlossen werden, dass nicht alle Teilnehmer jedes Item des PANAS in Bezug auf die *momentane* Stimmung beantworteten. Sollte diese

Annahme zutreffen, wären die Ausprägungen von positivem und negativem Affekt nicht als Zustands-Variablen, sondern im Sinne von Eigenschaften zu interpretieren. Persönlichkeitseigenschaften wiederum werden in den meisten integrativen Modellen zum Entstehen von Zufriedenheit und Wohlbefinden als maßgebliche Einflussgrößen angesehen: Subjektive Bewertungsprozesse (*appraisal*), wie sie auch der Bildung von Zufriedenheitsurteilen zu Grunde liegen, sind nicht zuletzt von Dispositionen wie beispielsweise Neurotizismus und Extraversion oder dispositionellem positivem und negativem Affekt abhängig (Brief et al., 1993; Clark & Watson, 1991; Costa & McCrae, 1980; Mroczek & Spiro, 2005). So wird auch im dem FLQM zu Grunde liegenden Modell der Lebenszufriedenheit nach Campbell et al. (1976) ein maßgeblicher Einfluss von Persönlichkeitsfaktoren auf alle Stufen des Prozesses der Urteilsbildung postuliert (vgl. Abbildung 2). Der hoch signifikante Zusammenhang zwischen FLQM-Ausprägung und positivem Affekt wäre somit theoriekonform, falls die überwiegende Zahl der Teilnehmer den PANAS im Sinne einer Eigenschaftsmessung beantwortet hätte. Zur Überprüfung dieser Annahme wären Daten zu weiteren Persönlichkeitsvariablen hilfreich, beispielsweise Neurotizismus und Extraversion (Borkenau & Ostendorf, 1993) oder Depressivität (z. B. Hautzinger, Bailer, Keller & Worall, 1995; Gauggel & Birkner, 1999). Eine zukünftige Studie zur Konstruktvalidierung des FLQM sollte diese Variablen mit einbeziehen. Das Zusammenhangsmuster zwischen FLQM und den PANAS Subskalen kann darüber hinaus als Hinweis interpretiert werden, dass der FLQM eher positive als negative Aspekte der Lebensqualität erfasst, also ressourcenorientiert angelegt ist, und nicht negative im Sinne einer Defizitorientierung des Fragebogeninhalts.

Der Fragebogen zur subjektiven Gesundheitseinschätzung SF-36 wurde in erster Linie hinsichtlich der körperlichen und psychischen Summenskalen ausgewertet. Die empirischen Werte der insgesamt acht Einzelskalen, welche teilweise lediglich aus zwei Items bestehen, wurden zwar der Vollständigkeit halber im Ergebnisteil dargestellt, jedoch nicht in die eigentliche Validierung mit einbezogen. Es ist diskussionswürdig, ob die Subskalen tatsächlich inhaltstvalid die in ihren Bezeichnungen implizierten Konstrukte abbilden.

Die körperliche Summenskala zeigt einen sehr viel engeren und auf einem höheren Niveau signifikanten Zusammenhang mit den Gesamtscores des FLQM als angenommen wurde. Diese Beobachtung steht – trotz der unerwartet hohen Korrelation – in Einklang mit anderen Forschungsarbeiten zum Zusammenhang von Lebenszufriedenheit und subjektiver

Gesundheit, denn subjektive Gesundheitseinschätzungen zeigen einen engeren Zusammenhang mit subjektivem Wohlbefinden als viele objektive Faktoren (Schneider et al., 2004; Smith et al., 2002; vgl. Abschnitt 1.2.3). Auch im Lichte des Urteilsmodells der Zufriedenheit von Schwarz & Strack (1991) ist dieser Befund nachvollziehbar: Eine frühe Modellstufe beinhaltet das „Abrufen der relevanten Informationen [zur Beurteilung der Zufriedenheit]“ (vgl. Abbildung 3); Bewertungen der eigenen Gesundheit fließen also direkt mit in Zufriedenheitsurteile ein – vorausgesetzt, Gesundheit stellt eine subjektiv für die Lebenszufriedenheit relevante Variable dar. Da 30 der 44 Teilnehmer direkt ihre eigene Gesundheit im Rahmen des FLQM nannten (und 20 die benachbarte Kategorie „Mobilität und körperliche Fähigkeiten“), ist ein enger Zusammenhang zwischen der körperlichen Summenskala und dem FLQM post hoc gut nachvollziehbar. In diesem Sinne dürfen die Ausprägungen der körperlichen Summenskala des SF-36 als Untermauerung der Validität des FLQM interpretiert werden.

Der nur mittlere Zusammenhang zwischen FLQM und psychischer Summenskala entspricht der Vorannahme innerhalb der Konstruktvalidierung. Diese mag zunächst kontraintuitiv erscheinen, wird aber unter *inhaltlichen* Gesichtspunkten verständlich: Die Items, welche am stärksten gewichtet in die psychische Summenskala eingehen, beziehen sich zum größten Teil auf Symptome einer psychiatrischen *Morbidität*. Beispielsweise verwenden sechs dieser Items das Fragegerüst „[...] konnten sie [... Tätigkeit oder Handlungsablauf nicht wie gewohnt durchführen], z. B. weil sie sich niedergeschlagen oder ängstlich fühlten“. Eine derartig gerahmte Frage kann schwerlich als Indikator psychischen *Wohlbefindens* angesehen werden, sondern eher als Indikator psychischer *Krankheit*. Die psychische Summenskala hat, wie der gesamte SF-36, eher den Charakter einer Bestandsaufnahme von Defiziten als einer Auslotung von Potenzialen und Grenzen, wodurch die Variationsbreite der möglichen Antworten von vorne herein einschränkt ist. Der SF-36 wird zwar mit dem Anspruch präsentiert, (gesundheitsbezogene) Lebensqualität zu erfassen. Da die Operationalisierung des Konstruktes jedoch fast ausschließlich defizitorientiert erfolgt, impliziert die bloße Abwesenheit der im Fragebogen erfassten Defizite oder Symptome hohe Lebensqualität – eine fast absurde Annahme (vgl. z. B. die WHO-Definition von Gesundheit als Wohlbefinden: (WHO, 1946). Eine weitgehende Unabhängigkeit der FLQM-Gesamtwerte von einem Maß psychischer Morbidität wäre dennoch erstaunlich, daher stützt der gefundene mittlere Zusammenhang weiter die Konstruktvalidität des FLQM. Es wäre wünschenswert das Verhältnis von psychischer

Summenskala und FLQM über die Einbeziehung weiterer Variablen psychischer Gesundheit, z. B. Depressivität, näher zu beleuchten.

Die explorative Konstruktvalidierung des FLQM anhand einer Pilotstichprobe unterstützt die Validität des Fragebogens. Hohe und mittlere Zusammenhänge finden sich zwischen dem FLQM den inhaltlich am stärksten verwandten Validierungsskalen: der Einzelfrage zur allgemeinen Lebenszufriedenheit sowie dem PGCMS und zwei seiner Teilskalen. Während der FLQM-Score im prognostizierten Rahmen mit der psychische Summenskala des SF-36 zusammenhing, war die Korrelation mit der physischen Summenskala höher als erwartet. Die wahrgenommene eigene Gesundheit und Funktionsfähigkeit im Alltag scheint auch bei Personen, die sich vermutlich schon über einen längeren Zeitraum an ihre Einschränkungen anpassen konnten, eine bedeutende Rolle zu spielen. Negativer Affekt hängt, wie erwartet, nicht signifikant mit dem Gesamtscore auf dem FLQM zusammen, was als Hinweis auf dessen Ressourcenorientierung gedeutet wird. Unerwartet hoch fiel dagegen die Korrelation mit positivem Affekt aus. Den möglichen Ursachen sollte nachgegangen werden – vermutet werden eine unintendierte Konfundierung des FLQM mit dem affektiven Zustand oder eine *trait*-Interpretation des PANAS seitens der Studienteilnehmer.

#### **4.2 Methodologische Gesichtspunkte**

Abgesehen von allgemeinen Validitätsgesichtspunkten, muss die spezifische Operationalisierung von Lebensqualitätswerte im FLQM genau reflektiert werden. Einige Aspekte des Vorgehens könnten als problematisch angesehen werden: Die Bildung einer Produktsumme als Index, die Art der verwendeten Antwortskalen sowie selbstverständlich die große Heterogenität der individuell generierten Items werden Hauptdiskussionspunkte sein.

Ganz allgemein ist eine Besonderheit von Variablen, die aus Produktsummen gebildet werden, dass sehr unterschiedliche, sogar gegensätzliche Ausprägungen der einzelnen Produktterme zum gleichen Ergebnis führen können. Eine geringe Zufriedenheit, gekoppelt mit einer hohen Priorität des Bereichs, resultiert im selben Gesamtwert wie eine hohe Zufriedenheit in einem relativ unwichtigen Bereich. Gleiche *Produktsummen* können folglich ebenfalls auf sehr unterschiedliche Weise zustande kommen. Impliziert wird damit, dass sowohl Zufriedenheit als auch Gewicht in gleichem Ausmaß zum Gesamtwert



beitragen (Lauver & Knapp, 1993). Da diese Annahme in in dieser Arbeit verwendeten Modell getroffen wird, ist die Bildung einer Produktsomme grundsätzlich zulässig. Viel entscheidender ist, dass durch die simple Summierung der Produkte eine extrem große Spannweite von Werten erzeugt wird, die lediglich ein Artefakt der Berechnungsvorschrift darstellt. Dadurch, dass die Gewichtungen für jeden Bereich neu völlig frei gewählt werden können, werden für alle Personen, welche viele Bereiche als eher bedeutsam ansehen (und damit hohe Gewichte in die Gleichung einbringen) hohe Gesamtwerte erzeugt. Ein hoher Indexwert auf dem FLQM steht für eine hohe Lebensqualität – die bloße Tatsache, dass den genannten Lebensbereichen im Schnitt eine hohe Bedeutsamkeit beigemessen wird, bedeutet aber keinesfalls, dass damit eine hohe Lebensqualität verbunden ist. Auch für Personen, die eine große Anzahl an Bereichen nennen (und damit auch mehr Terme in die Produktsomme einbringen) werden bei diesem Vorgehen künstlich höhere Werte erzeugt (vgl. Streiner & Norman, 2003). Im FLQM wird diesem Problem begegnet, indem die Produktsomme um die Summe der Gesamtgewichte korrigiert, d. h. durch diese Summe dividiert wird. So wird ein *individuell gewichteter* Mittelwert der Lebensqualität gebildet, der nur die *relative* Gewichtung der Bereiche zueinander reflektiert und weder durch die Anzahl der genannten Bereiche, noch durch die Höhe der vergebenen Gewichte verzerrt wird. Die Grundannahme, dass Gewichte und Zufriedenheiten in ähnlichem Maße zum Gesamtwert beitragen, bleibt bestehen.

Ein weiterer methodologischer Aspekt ist die Verzerrung durch systematische Messfehler auf Grund des Formats oder durch Antworttendenzen. In Bezug auf die erste mögliche Fehlerquelle – Verzerrungen, die durch das Antwortformat begünstigt werden – besteht bei den an die Schulnoten angelehnten Skalen, wie sie im FLQM verwendet werden, die Gefahr einer künstlich schiefen Verteilung: Der Mittelpunkt der Schulnotenskala wird subjektiv erfahrungsbasiert eher zwischen 4 („ausreichend“) und 5 („mangelhaft“) angesiedelt als beim tatsächlichen Skalenmittelpunkt zwischen 3 und 4 (vgl. Streiner & Norman, 2003). Zum einen deuten jedoch die Ergebnisse aus einer begleitenden Befragung der Hälfte der Studienteilnehmer mit einem Meta-Fragebogen darauf hin, dass die Teilnehmer sich stark an den konkreten Bezeichnungen der Skalenpunkte orientierten und weniger an den Zahlenwerten. Zum anderen bestehen aus verschiedenen Studien gute Erfahrungen mit dem Verständnis des Bewertungsprinzips seitens älterer Teilnehmer, wenn dieses an eine bekannte Bewertungsskala (z. B. Schulnoten) angelehnt wird. Das vergleichsweise geringe Risiko der Antwortverzerrung

durch das Antwortformat wurde somit zu Gunsten der Verständlichkeit bewusst eingegangen. Die empirische Werteverteilung des FLQM spricht für das gewählte Skalenformat: Beim FLQM war zwar eine leichte Verschiebung der Verteilung zum positiven Pol der Skala zu beobachten<sup>2</sup>, diese geht jedoch völlig konform mit der aktuellen Forschungsliteratur und scheint nicht auf einem systematischen Fehler auf Grund des Fragebogenformats zu basieren. Letzterer Möglichkeit könnte durch systematische Variationen des Antwortformats nachgegangen werden. Generelle Antworttendenzen einzelner Teilnehmer können, wie bei fast jedem Fragebogen, selbstverständlich nicht ausgeschlossen werden. Zumindest auf der Ebene von Gruppenvergleichen kann man jedoch davon ausgehen, dass sich diese möglichen Tendenzen ausgleichen und in der Gesamtstichprobe nicht zu systematischen Verzerrungen führen. Auf das Problem des sozial erwünschten Antwortverhaltens wird im Rahmen der Diskussion von Einschränkungen und Grenzen der Studien 1 und 2 weiter unten gesondert eingegangen (Abschnitt 4.6).

Das zur Likert-Skala alternative Antwortformat ist eine visuelle Analogskala (VAS) mit zwei Ankerpunkten an den Extremen. Diese Skalenart umfasst häufig einen Wertebereich von 0 (geringste Ausprägung) bis 100 (höchste Ausprägung); auf diesem Kontinuum können sodann nicht durch Intervalle beschränkte Antworten visuell markiert werden. Auf den ersten Blick scheint die VAS das einfachere Format zu sein. Es ist jedoch wesentlich schwieriger, Antworten auf dieser Skala zu interpretieren als auf einer Likert-Skala, da Antworten außerordentlich stark von der spezifischen Formulierung der Ankerpunkte abhängen. Zudem fehlen Belege dafür, dass eine VAS eine bessere Diskrimination der Merkmalsausprägung ermöglicht, als beispielsweise eine Likert-Skala mit sieben Punkten (Steiner & Norman, 2003). Hinzu kommt, dass vor allem ältere (aber durchaus auch jüngere) Probanden Probleme mit dem Verständnis und der Anwendung dieses Skalentyps berichten (z. B. Guyatt, Townsend, Berman & Keller, 1987). Die Likert-Skalierung erweist sich einer VAS trotz deren scheinbaren Einfachheit in vielen Punkten als überlegen, so dass bei der Auswahl des Antwortformats keine bessere Alternative zur Verfügung stand.

---

<sup>2</sup> Man beachte, dass zur Berechnung die von den Teilnehmern genannten Werte für Zufriedenheit und Gewichtung umgepolt werden. Hohe Werte indizieren im eigentlichen Fragebogen negative Ausprägungen, in der umgepolten, ausgewerteten Form jedoch positive Ausprägungen von Lebensqualität bzw. Zufriedenheit.

Ein letzter wichtiger methodologischer Einwand könnte die große Heterogenität der anteilig individuell konstruierten Items betreffen. Im Zuge der Bearbeitung wird ein maximal idiografisches System von Bereichen der Lebensqualität generiert, welches beispielsweise auch als Grundlage der Bildung von Gruppenmittelwerten dient. Die Werte, mit denen operiert wird, werden jedoch aus inhaltlich unterschiedlichen Items gebildet. Dieses streng idiografische Vorgehen fußt jedoch auf einem Modell der Lebenszufriedenheit, welches durch die verschiedenen Schritte des FLQM lediglich *expliziert* wird: Der FLQM separiert die einzelnen kognitiven Operationen, die nach den Annahmen des Modells von Campbell et al. (1976) der Beantwortung einer Frage zur Lebenszufriedenheit vorgenommen werden. Wenn man von der Gültigkeit dieses Modells ausgeht, dann werden auch alle globalen Zufriedenheitsurteile auf Basis interindividuell unterschiedlicher Bereichssysteme gebildet. Pavot und Diener (1993b) beispielsweise tragen dieser Annahme in ihrer Satisfaction with Life Scale (SWLS) Rechnung, die mit dem Anspruch konzipiert wurde, auf Grundlage individuell gewählter Werte oder Lebensbereiche und deren Gewichtung einen Zugang zur Lebenszufriedenheit zu gewähren. Dieses individuelle Bereichssystem wird jedoch nicht explizit aufgeschlüsselt, vielmehr sind die Fragen der SWLS so global formuliert, dass, so die Annahme der Autoren, der Befragte nach eigenem Gutdünken seine persönlich bedeutsamen Lebensbereiche integrieren und gewichten kann. Beispielsweise lautet das erste Item „*In most ways my life is close to my ideal*“ (Pavot & Diener, 1993b, S. 172) – der Befragte soll dann auf einer siebenstufigen Likert-Skala den Grad seiner Zustimmung angeben. Wenn man von der Zulässigkeit globaler Zufriedenheitsurteile ausgeht, spricht nichts dagegen, eine Offenlegung der Schritte, welche zu diesem Urteil führen, ebenfalls als Möglichkeit in Betracht zu ziehen – solange Vergleiche auf der Ebene des Produktsummenscores stattfinden und nicht auf spezifischen Bereichsebenen.

### **4.3 Verständlichkeit und Durchführbarkeit**

Die „Durchführbarkeit“ des Fragebogens hängt in erster Linie von pragmatischen Gesichtspunkten bei der Durchführung des FLQM ab, wie Bearbeitungsdauer, Akzeptanz oder Verständlichkeit. Aus Gründen des Studiendesigns werden die folgenden Ausführungen in erster Linie deskriptive Aspekte und die Perspektive der *Befragten* adressieren. Zu allen Aspekten der Durchführung und Auswertung, welche die *Anwender*

betreffen, können auf Grund der Tatsache, dass die gesamte Pilotstudie vom Testentwickler durchgeführt wurde, lediglich Befunde anderer Studien mit patientengenerierten, individualisierten Messinstrumenten übertragen werden.

Auf den ersten Blick erscheint der Zeitaufwand für die Anwendung des FLQM beträchtlich: Bearbeitungszeiten von fast 19 Minuten im Mittel und bis zu 46 Minuten insgesamt werfen die Frage auf, ob dieser Fragebogen für die Einbettung in einen klinischen Routineablauf geeignet ist. Eine Fokussierung auf Mittel- und Extremwerte verzerrt indes das Bild, denn die Hälfte der Teilnehmer konnte den Fragebogen in höchstens 15 Minuten bearbeiten, insgesamt ein Viertel sogar in höchstens 12 Minuten. Hinzu kommt, dass der FLQM in der Pilotstudie ohne zeitliche Restriktionen getestet wurde: Auch wenn Teilnehmer sehr weitläufig oder abschweifend antworteten, wurden sie in der Regel *nicht* auf die Ausgangsfrage zurückgeführt. Es gab daher zahlreiche Teilnehmer, die mehr als die Hälfte der gemessenen Zeit nicht mit der direkten Bearbeitung der jeweiligen Stufe des FLQM (Benennen, Bewerten und Gewichten von Bereichen) beschäftigt waren, sondern stattdessen unterschiedlichste Aspekte und Phasen ihres jetzigen oder früheren Lebens resümierten. Eine explorative Längsschnittuntersuchung mit dem FLQM (vgl. Abschnitt 3.4) zeichnet ein positiveres Bild vom zeitlichen Rahmen, der zur Bearbeitung des FLQM notwendig ist. Im Längsschnitt wurden die Teilnehmer sehr schnell wieder auf die jeweils anstehende Bearbeitungsstufe zurückgeführt, wenn sie zu stark davon abwichen. Die durchschnittliche Bearbeitungszeit lag hier bei 14 Minuten, der Median bei 13. Diese Werte geben vermutlich eine realistischere Einschätzung des Zeitaufwands und werden voraussichtlich in einer anschließend geplanten, stärker standardisierten Validierungsstudie repliziert werden können. Gleichzeitig gaben drei Teilnehmer im Rahmen begleitenden Meta-Fragebogens sogar an, der Fragebogen sei in Anbetracht des Themas zu kurz.

Die Abwägung zwischen den jeweils gegebenen ökonomischen und personellen Rahmenbedingungen einerseits und den Bedürfnissen der Befragten bzw. der Komplexität der Fragestellung andererseits erscheint sinnvoll, bevor eine endgültige Bewertung der Durchführungsdauer bzw. eine Anpassung des Fragebogens vorgenommen wird. Es darf angenommen werden, dass der größte Teil der Teilnehmer im reflexiven Denken nicht geübt ist und somit die Art der Aufgabenstellung selbst einen wesentlichen Teil der Schwierigkeiten bei der Bearbeitung bedingt – unabhängig vom konkreten Inhalt der gestellten Anforderung. Meiner Ansicht nach muss eine gewisse Mindestdauer der

Durchführung einkalkuliert werden, um den Teilnehmern die Gelegenheit zu geben, sich mit der ungewohnten kognitiven Anforderung vertraut zu machen und damit erst zu verwertbaren Ergebnissen zu gelangen. Zudem ist der FLQM in der Durchführung wesentlich kürzer als vergleichbare Fragebögen wie der SEIQoL (mittlere Durchführungsdauer bei Älteren 38 Minuten, Range: 15 – 105 Minuten; (Mountain et al., 2004) oder der PGI (Range bei Älteren: 45 – 80 Minuten; (Tully & Cantrill, 2000)). Auch für die Bearbeitung des in der Klinik stark verbreiteten SF-36 benötigen ältere Teilnehmern im Schnitt 34 Minuten (Range: 8 – 100 Minuten; (Mountain et al., 2004)). Bei allen drei genannten Fragebögen ist zudem die Auswertung zeitaufwändig und, mit Ausnahme des PGI, nur mit spezieller Software möglich.

In der Altersgruppe der Teilnehmer sollten einige Besonderheiten im Erhebungsmodus beachtet werden, um den Fragebogen „altersfair“ zu gestalten. Durch den Einsatz des Meta-Fragebogens in einer Teilstichprobe (n = 21) können einige dieser Gesichtspunkte nicht nur extern, sondern auch aus subjektiver Sicht diskutiert werden. Beide Perspektiven werden im Folgenden berücksichtigt.

Da visuelle und auditive Einschränkungen im Alter sehr verbreitet sind, wurde die Präsentationsform des Vorlesens gewählt in Kombination mit der Möglichkeit, selbst mitzulesen. Dadurch wird den Teilnehmern ermöglicht, je nach individueller Präferenz auf ein Darbietungsformat zu fokussieren. Diese Art der Darbietung stieß aus Teilnehmersicht auf große Zustimmung und kann deshalb befürwortet werden. Aufgrund möglicherweise geringerer Merkfähigkeit oder Konzentrationsschwäche sollten in älteren Populationen Fragebögen bzw. Erhebungssituationen generell möglichst kurz gehalten werden – nicht nur aus Gründen der Belastung für die jeweiligen Teilnehmer, sondern auch um möglichen Antwortverzerrungen durch Ermüdung vorzubeugen. Im Zuge der Bearbeitung des FLQM wurden den Teilnehmern daher sowohl die Instruktionen als auch die jeweiligen Antwortskalen für Bewertung und Gewichtung der Lebensbereiche in großer Schrift vorgelegt. Die visuelle Präsenz der relevanten Informationen stellt im Rahmen des FLQM einen wichtigen Anker dar, der kognitive Besonderheiten älterer Populationen gegebenenfalls ausgleichen kann. Die Notwendigkeit dieser Maßnahme spiegelt sich auch in der Einschätzung der Teilnehmer: Die Prädikate der einzelnen Skalenstufen (Wertelabels) wurden bis auf eine Ausnahme von allen Befragten als hilfreich empfunden. Auch empfand keiner der Befragten den Fragebogen insgesamt als schwierig oder, wie bereits erwähnt, als zu lang.

Wichtige Punkte einer altersangemessenen Erhebung sind selbstverständlich auch die inhaltliche Verständlichkeit und Nachvollziehbarkeit der Instruktionen, noch vor Präsentationsformat und Länge (vgl. Fleischmann & Oswald, 2001; Gunzelmann & Oswald, 2005). Auch hier zeichnen die Befragten ein positives Bild des FLQM: Alle empfanden die Instruktion als gut verständlich und der überwiegende Teil gab an, mit den Bewertungs- und Gewichtungsskalen gut zurecht zu kommen. Eine so positive Bilanz der Verständlichkeit war nicht als selbstverständlich zu erwarten, kam es doch während der Datenerhebung wiederholt zu Verständnisproblemen bezüglich der Antwortskalen zur PGCMS, einem Fragebogen der explizit für ältere Personen entwickelt wurde (vgl. Lawton, 1975). Demgegenüber hatten die Teilnehmer keine beobachtbaren oder berichteten Probleme mit dem Antwortformat des FLQM, was sich auch in der positiven Einschätzung der verwendeten Skalen und Skalenformate spiegelt. Die Bedeutung der guten Verständlichkeit für die Beurteilung der Qualität des FLQM kann nicht hoch genug eingeschätzt werden. Zusammenfassend erschienen die Akzeptanz des FLQM sowie Verständnis und Kooperationsbereitschaft seitens der Teilnehmer gut. Untermauert wird dieser subjektive Eindruck während der Datenerhebung durch den begleitenden Meta-Fragebogen.

Die Anregungsliste als wesentlicher Bestandteil des Fragebogens wurde von den befragten Teilnehmern bis auf eine Ausnahme als hilfreich bei der Bearbeitung des Fragebogens gewertet. Gleichzeitig gaben sieben der 21 Befragten an, zumindest „teilweise“ in der Auswahl der ihnen wichtigen Lebensbereiche durch die Anregungsliste beeinflusst worden zu sein. Unterschiede in der Abhängigkeit der Nennungen von der Anregungsliste zwischen den aus eigener Sicht beeinflussten und nicht beeinflussten Teilnehmern sind statistisch nicht signifikant. Dennoch bleibt hier Diskussionsbedarf: Auch eine unintendierte Vorgabe der Urteilsdimensionen durch den Fragebogen im Sinne einer *Auswahl* würde dessen Anspruch unterminieren, *individuelle* Bereichspräferenzen widerzuspiegeln. Andererseits wurde bei der Zusammenstellung der Beispiele sorgfältig darauf geachtet, eine vielseitige und breit gefächerte Auswahl zur Verfügung zu stellen. Dadurch soll eine variationsreiche individuelle Zusammenstellung unterstützt werden, die unterschiedlichste Bereiche des Lebens umfasst. So liegt auch das mittlere Verhältnis von Übereinstimmungen mit der Anregungsliste und Gesamtzahl der Nennungen relativ hoch: Etwa jede zweite individuelle Nennung findet sich in vergleichbarer Form auf der Anregungsliste. Ein gewisser Grad an Übereinstimmung mit der Liste ist jedoch nicht

überraschend, denn es befinden sich eben zahlreiche Bereiche darunter, die von vielen Menschen als bedeutsam angesehen werden und teilweise gerade unter diesem Gesichtspunkt in die Liste aufgenommen wurden (vgl. Abschnitte 2.3.1.3 und 3.2.2). Die bloße Übereinstimmung einer individuellen Wahl mit der Liste ist also nicht grundsätzlich problematisch, ebenso wenig die Wahrnehmung, dass die Beispiele die eigene Auswahl *teilweise* beeinflusst haben. Wäre die Liste jedoch von mehr Teilnehmern als Vorgabe empfunden worden oder der Übereinstimmungsquotient deutlich über 0,5 gestiegen, stünde das gesamte Konzept des FLQM in Frage. Vermutlich wäre es in einem dieser Fälle sinnvoller, mit vorgegebenen Dimensionen zu arbeiten. Weiter unten wird der Grad der Übereinstimmung mit der Anregungsliste im Kontext der Inhalte der individuellen Nennungen detailliert diskutiert.

Die Kommentare der Studienteilnehmer – zu den Items 1, 2, 3, 9 und 10 des Meta-Fragebogens – geben Anregungen zur Optimierung des FLQM sowie weitere Hinweise auf mögliche Schwierigkeiten. Aus mehreren Äußerungen kann geschlossen werden, dass weniger die Form des Fragebogens problematisch für die Befragten war, als vielmehr die Komplexität der dahinter stehenden Aufgabe „in so kurzer Zeit Rechenschaft abzulegen über die eigenen Bedürfnisse und das Leben“, wie Teilnehmer 51 es ausdrückte. Die Reduktion komplexer Facetten des Lebens auf wenige Stichworte und die Vergabe von Punktwerten, die diese Komplexität angemessen widerspiegeln, scheint trotz des guten Verständnisses eine Herausforderung für zahlreiche Teilnehmer gewesen zu sein. Auch die Anmerkungen zur Verständlichkeit des Fragebogens für andere Personen in ähnlichen Lebenslagen (Frage 9) reflektieren das Bewusstsein der Teilnehmer für die vergleichsweise hohe Komplexität der Anforderung: Alle Teilnehmer die (unaufgefordert) diese Frage kommentierten waren der Ansicht, der „intellektuelle Hintergrund“ oder die „geistige Regeksamkeit“ seien entscheidende Voraussetzung für ein Verständnis des Fragebogens (vgl. Tabelle 22). Diese Befunde bestärken aus der Teilnehmerperspektive die Einschränkung der Studienpopulation auf Personen ohne kognitive Defizite. Im Falle einer erfolgreichen Validierung des FLQM an einer größeren, zielgruppenrepräsentativen Stichprobe wäre es erstrebenswert, Modifikationsmöglichkeiten und Anpassungen des FLQM für Personen mit zumindest beginnenden oder leichten kognitiven Beeinträchtigungen zu explorieren. Möglicherweise wird eine individualisierte Messung der Lebensqualität in dieser Population aber nur eingeschränkt möglich sein. Außerdem werden restriktivere Vorgaben bzw. eine stärkere Eingrenzung der Auswahlbereiche notwendig sein, um den Betroffenen

eine Bewältigung der kognitiven Anforderung zu ermöglichen. Denkbar wäre auch eine weitere Vereinfachung der Gewichtung auf eine dreistufige Skalen (z. B. „sehr wichtig“, „nicht so wichtig“, „egal“).

In Bezug auf die *Befragten* sprechen die Erfahrungen bei der Fragebogenvorgabe und die empirischen Ergebnisse für eine gute Durchführbarkeit des FLQM. Zu klären bleibt die Frage, wie potenzielle *Anwender* die Vorgabe und Auswertung des FLQM bewerten. Dazu gibt es bislang keine empirischen Erkenntnisse, da sämtliche Befragungen im Rahmen der Pilotstudie vom Entwickler des Fragebogens durchgeführt wurden. Im Rahmen einer größeren Validierungsstudie sollte dieser wichtige Aspekt parallel quantitativ und qualitativ erkundet werden. In Anbetracht der positiven Resonanz seitens der Studienteilnehmer steht zu hoffen, dass Anwender dem FLQM die gleiche Akzeptanz entgegenbringen.

Die Kosten für einen Einsatz des FLQM sind relativ gering anzusetzen: Abgesehen von Beschaffungskosten und Personal für die Durchführung genügt zur Auswertung einfachste Statistiksoftware – notfalls kann die Auswertung problemlos per Hand durchgeführt werden. Für eine volle Ausschöpfung des Informationsgehalts des FLQM ist eine Speicherung der quantitativen wie auch qualitativen Informationen über einen längeren Zeitraum sinnvoll. Denkbar ist die Erstellung einer einfachen, EXEL™-kompatible Eingabemaske für die genannten Lebensbereiche sowie ihre Bewertungen und Gewichte.

Zusammengenommen sind die Einschätzungen des FLQM auf dem Meta-Fragebogen seitens der Teilnehmer sehr ermutigend. Sowohl der Umfang als auch das Format des Fragebogens stieß auf überwiegende Zustimmung. Die Instruktionen sowie die Antwortskalen scheinen aus Sicht der befragten älteren Menschen angemessen und die Auswahlliste wurde insgesamt gut akzeptiert und als unterstützend wahrgenommen. Durchführungsdauer und Kosten bewegen sich in einem gut vertretbaren Rahmen. Der Einsatz des FLQM kann somit unter Gesichtspunkten der Durchführbarkeit und Akzeptanz auf Grundlage der Befunde der Pilotstudie als Messinstrument der Lebensqualität älterer, multimorbider Personen empfohlen werden.

#### **4.4 Reliabilität und Objektivität des FLQM**

Neben der Validität sind Objektivität und Reliabilität die beiden wesentlichen klassischen Gütekriterien eines Fragebogens. Die Auswertungsobjektivität des FLQM kann als



gesichert angesehen werden, da die Auswertung in einer einfachen mathematischen Verknüpfungsfunktion ohne Interpretationsspielräume besteht. Zu Durchführungs- und Interpretationsobjektivität können im Rahmen von Studie 2 jedoch keine Angaben gemacht werden, da der Fragebogen nicht von unterschiedlichen Anwendern verwendet wurde. Für eine Bestimmung der Interpretationsobjektivität ist zudem das Ziel der Erhebung von entscheidender Bedeutung. Unzweifelhaft wird bei einfachen Vergleichen über die Zeit oder über Personen und Gruppen jeder Anwender zum gleichen Ergebnis gelangen; mögliche Implikationen der jeweiligen Befunde können aber nur mit populations-spezifischen Normen objektiv abgeleitet werden. Ob die Entwicklung solcher Normwerte bei einem so hochgradig subjektiven Konstrukt und der großen Heterogenität der Zielgruppe überhaupt sinnvoll ist und welche Anwendungsfelder für normbezogene Interpretationen sich möglicherweise bieten, steht zur Debatte. Gerade die Durchführungsobjektivität ist möglicherweise ein besonders sensibler Bereich, denn die von den Teilnehmern im Verlauf des Fragebogens verlangten Informationen sind sehr privater Natur. Es kann angenommen werden, dass der Grad der Offenheit, mit dem der FLQM bearbeitet wird, in hohem Maße von interpersonalen Faktoren und Sympathie abhängig ist. Daher sollte der Überprüfung und Sicherstellung dieses Gütemerkmals in einer zukünftigen Untersuchung besonderes Augenmerk gelten.

Eine Reliabilitätsprüfung im herkömmlichen Sinne verbietet sich auf Grund der Grundannahmen bezüglich des untersuchten Konstruktes. Lebensqualität ist in der Operationalisierung des FLQM ein hochgradig fluktuierendes, heterogenes Merkmal. Ein einfacher *retest* ist in Anbetracht der Fluktuation nicht sinnvoll, *split-half* Reliabilität oder interne Konsistenz lassen sich wegen der großen Heterogenität der individuell konstruierten Items nicht bestimmen. Erschwerend kommt hinzu, dass die Anzahl der Items in Abhängigkeit von den individuell benannten Dimensionen stark schwankt. Eine Möglichkeit, tatsächlich Hinweise auf die Reliabilität des FLQM zu erhalten, würde in einem relativ aufwändigen längsschnittlichen Design bestehen: Parallel zu einer erneuten Vorgabe des FLQM müssten Variablen mit erhoben werden, die einen nachgewiesenen oder theoretisch begründbaren Einfluss auf die Ausprägung von Lebensqualität (gemessen mit dem FLQM) haben. Denkbar sind beispielsweise Maße der Multimedikation, des Schmerzes, subjektiver Gesundheit, kritischer Lebensereignisse, globaler Lebenszufriedenheit oder Symptomstärke und Diagnosen. Der Grad der gemeinsamen Variabilität über die Zeit und die Kovarianz der Veränderungen auf den einzelnen Maßen

erlauben Rückschlüsse auf die Sensitivität des FLQM hinsichtlich Stabilität oder Veränderung. Man sollte in diesem Zusammenhang jedoch vermutlich eher von *responsiveness* als von Reliabilität des Instruments sprechen – also dem Grad, zu dem der FLQM auf tatsächliche Veränderungen angemessen reagiert (vgl. Fitzpatrick et al., 1998).

Die Ergebnisse einer explorativen Überprüfung der Reproduzierbarkeit durch wiederholte Messung sechs Monate nach der Baseline-Erhebung lassen sich in diesem Kontext als positive Signale werten. In dieser zweiten Erhebung wurden neben dem FLQM und der globalen Lebenszufriedenheit auch bedeutsame Ereignisse im Leben der Teilnehmer in offener Fragestellung erfasst. Der Zusammenhang zwischen FLQM-Gesamtwert und globaler Lebenszufriedenheit war ähnlich hoch positiv wie in der Pilotstudie, allerdings waren beide Variablen zeitlich relativ stabil. Eine systematische Überprüfung differentieller Effekte der subjektiv bedeutsamen Ereignisse zwischen den Erhebungszeitpunkten konnte leider nicht durchgeführt werden, da die Fallzahl derjenigen, die *keine* Ereignisse berichteten zu niedrig für einen sinnvollen Gruppenvergleich war. Eine zukünftige, größere Studie sollte in jedem Fall diesen Forschungsstrang weiter verfolgen.

#### **4.5 Inhaltsvalidität und querschnittliche Unterschiede subjektiver Konstruktionen von Lebensqualität im Alter**

Sowohl in der qualitativen Vorstudie (Studie 1) zur Entwicklung des FLQM, als auch in der Pilotstudie selbst (Studie 2) wurden Informationen zur subjektiven Konstruktion der eigenen Lebensqualität der Teilnehmer zugänglich, die unter anderem Hinweise auf die Inhaltsvalidität des FLQM liefern können. In Studie 1 geschah wurden diese inhaltlichen Informationen mittels expliziter Fragestellungen im Rahmen eines qualitativen Interviews gewonnen, während sie Studie 2 als Teilprodukt der Bearbeitung des FLQM angesehen werden können. In Studie 1 war der Zeitrahmen noch offener gehalten als in Studie 2, zudem waren die Teilnehmer dort aufgefordert, alle Aspekte zu benennen, die ihnen im Rahmen der jeweiligen Frage wichtig erschienen, während in Studie 2 eine bewusst auferlegte Beschränkung der Antworten auf „*in etwa fünf Bereiche, die am meisten damit zu tun haben, wie zufrieden [der Teilnehmer mit seinem Leben ist]*“ stattfand. Dementsprechend fand sich eine deutlich größere Auswahl von Lebensbereichen je Teilnehmer in Studie 1. Diese Unterschiede in der Reichhaltigkeit der individuellen

Kategoriensysteme sollten jedoch nicht interpretiert werden, denn Studie 1 diente im Wesentlichen dazu, das Feld der individuellen Definitionen von Lebensqualität und Wohlbefinden empirisch zu erschließen und einen Überblick über die Bandbreite möglicher Elemente der individuellen Systeme zu gewinnen. Die unterschiedlichen Arten von Fragestellungen und das im Vergleich zu Studie 2 noch niedrigere Abstraktionsniveau der Kategorien in Studie 1 beeinflussen maßgeblich die Anzahl der individuellen Nennungen und der abgeleiteten Kategorien. Ein ähnliches Vorgehen wurde z. B. auch in den Studien von Bowling et al. (2003) und Richard et al. (2005) gewählt (vgl. Tabelle 4).

Der Vergleich mit der internationalen Forschungsliteratur zu diesem Thema bestätigt die Ergebnisse der Vorstudie: Auch ältere Menschen berichten eine große Fülle bedeutsamer Lebensbereiche, die zwar möglicherweise andere Schwerpunkte aufweisen als die Präferenzprofile jüngerer Kohorten, jedoch an Heterogenität diesen in keiner Weise nachstehen. Es war ein Anliegen der Vorstudie, diese Vielfalt zu replizieren und im Abgleich mit der Forschungsliteratur eine empirisch begründete Auswahl von Beispieldimensionen für die Instruktion des FLQM zu ermöglichen. Die große Bandbreite von Bereichsnennungen in der relativ kleinen Stichprobe aus Studie 1 ( $N = 8$ ) belegt auf eindrucksvolle Weise, wie diskussionswürdig (und möglicherweise auch fragwürdig) eine universelle Festlegung von Bewertungsdimensionen der Lebensqualität in der Population mehrfach erkrankter Älterer ist. Bestätigung erfährt das aus den Interviews im Rahmen von Studie 1 abgeleitete Kategoriensystem durch die qualitativen Daten aus Studie 2. Zwar wurden für die inhaltliche Auswertung von Studie 2 aus pragmatischen Gesichtspunkten einige Kategorien zusammengelegt bzw. aufgrund fehlender Belegung gestrichen, doch konnte die große Vielfalt bedeutsamer Bereiche des Lebens klar repliziert werden. Selbst unter dem Vorbehalt der Fusion von Kategorien bzw. der partiellen Anpassung des deduktiven Kategoriensystems, wurden 19 Kategorien bedeutsamer Lebensbereiche aus den Nennungen abgeleitet. Diese Zahl ist, wie oben bereits angesprochen, im Kontext des Abstraktionsniveaus der Kategorien zu interpretieren. Dieses war in Studien 1 und 2 bewusst niedrig angesetzt, um das Kategoriensystem möglichst lebensnah und differenziert zu halten. Diese Prämisse galt zum einen in Hinblick auf die Auswahlliste, die aus dem Kategoriensystem abgeleitet werden sollte. Zum anderen lässt sich so die erlebte Komplexität von scheinbar homogenen Lebensbereichen abbilden, wie beispielsweise „Gesundheit“. Die Bewertung eines solchen Bereiches wird umso ambivalenter und weniger valide sein, je komplexer die Struktur der einzelnen Facette in sich selbst vom

Teilnehmer wahrgenommen wird. Mehrere Äußerungen von Teilnehmern im Rahmen der Befragung mit dem begleitenden Meta-Fragebogen unterstützen den gewählten Ansatz: So bemerkte beispielsweise ein Teilnehmer, dass es *„bei den einzelnen [Bereichen] [...] manchmal schwierig [ist], einen einzelnen Punktwert zu vergeben, da Facetten sehr unterschiedlich sein können“*. Eine zu starke Reduktion der idiografischen Lebensbereiche auf übergeordnete Hauptkategorien führt so möglicherweise zu einem Informationsverlust und Verzerrungen der Bewertung. Mit den Hierarchien einzelner Lebensbereiche und ihrer Facetten beschäftigt sich beispielsweise eine Arbeit von Oishi und Diener (2001).

Die Entwicklung der individuellen Bereichssysteme mit dem Alter, d. h. der subjektiven Konzepte von Lebensqualität, kann mit dem vorliegenden Datenmaterial nicht nachvollzogen werden. Trotzdem werfen querschnittliche Vergleiche der beiden Altersgruppen und des Zusammenhangs zwischen Alter und Facettenreichtum der Konzepte von Lebensqualität etwas Licht auf diese komplexe Entwicklungsthematik. Die Ergebnisse von Studie 2 legen nahe, dass sich die Vielfältigkeit der bedeutsamen Lebensbereiche mit zunehmendem Alter verringert: Hochbetagte (über 80 Jahre) nannten im Schnitt einen Lebensbereich weniger als die jungen Alten (65 bis 79 Jahre). Außerdem wiesen das Alter und die Anzahl der im Rahmen des FLQM benannten Lebensbereiche einen hoch signifikanten negativen Zusammenhang auf. Eine differenzierte Auswertung der genannten Lebensbereiche über alle Studienteilnehmer hinsichtlich der relativen Häufigkeit der Nennungen gibt weitere Aufklärung: Hochbetagte nannten lediglich sieben Kategorien mit einer relativen Häufigkeit von mindestens 0,2 – die jüngeren Alten fast doppelt so viele, nämlich 13 Kategorien. Das bedeutet, dass die älteren Teilnehmer nicht nur durchschnittlich weniger Kategorien bzw. Bereiche benannten, sondern auch deutlich weniger Bereiche mit nennenswerter Häufigkeit genannt wurden. Die Schwelle der relativen Häufigkeit von 0,2 wurde willkürlich gewählt. Sie entspricht bei den jungen Alten mindestens sechs, bei den alten Alten mindestens drei Nennungen. Daraus kann in den Grenzen dieser Pilotstudie auf eine Verengung und zunehmende Dedifferenzierung der subjektiven Struktur des Konstruktes „Lebensqualität“ im sehr hohen Alter geschlossen werden. Die Zahl der regelmäßig genannten Bereiche verringert sich, d. h. der „Pool“, aus dem geschöpft wird, verkleinert sich.

Mit einem etwas anderen Schwerpunkt wurde die Verengung der Bandbreite von bedeutsamen Lebensbereichen wurde der Ebene des *investment*, der Kraftinvestition in die Zielerreichung in bestimmten Lebensbereichen, im Rahmen der Berliner Altersstudie

untersucht. Die Teilnehmer sollten für zehn ausgewählte Lebensbereiche das Ausmaß ihrer persönlichen Investition angeben. Auch in diesem Forschungskontext konnte eine Reduktion der Bereiche beobachtet werden, in welche die Teilnehmer in stärkerem Maße investieren (Ebner, Freund & Baltes, 2006; Riediger & Freund, 2006). Ähnliche Ergebnisse zeigten die Analysen zum Facettenreichtum der Selbstdefinition innerhalb der BASE: Je stärker gesundheitlich eingeschränkt ältere Teilnehmer waren, desto weniger und inhaltlich weniger reichhaltige Selbst-definierende Bereiche benannten sie (Freund & Smith, 1999). Die Untersuchung der Balance von Gewinnen und Verlusten im Alter stützt diese Beobachtung: Mit dem Alter entwickelt sich die Bilanz der objektiven Gewinne und Verluste („*gains and losses*“) zunehmend negativ, einhergehend mit weniger adaptiven Funktionsprofilen im höheren Alter (Smith, 2003; Smith & Baltes, 1997; vgl. Abschnitt 1.2.1). Die Tatsache, dass zunehmend gesundheitliche, psychische und soziale *Einschränkungen* gegenüber den *Möglichkeiten* überwiegen, würde die Reduktion der wahrgenommenen Quellen von Lebensqualität und Wohlbefinden erklären. Gleichzeitig spiegelt sich diese Einschränkung der Grundlagen von Lebensqualität eben nicht in einer verminderten erlebten Lebensqualität wieder – weder im FLQM, noch in der Frage nach der allgemeinen Lebenszufriedenheit. Die Verringerung der Möglichkeiten bzw. der Bandbreite sinn- und zufriedenheitsstiftender Erfahrungen muss also keinen Rückgang der wahrgenommenen Lebensqualität bedingen. Solche Ergebnisse aus inhaltlich verwandten Forschungsprojekten lassen die Befunde dieser Arbeit plausibel erscheinen: Mit zunehmendem Alter scheint sich die Bandbreite der bedeutsamen Lebensbereiche zu verkleinern – jedoch ohne dass ein ursächlicher Einfluss auf die wahrgenommene Lebensqualität angenommen werden kann.

Die bisherigen Ergebnisse und Folgerungen lassen sich gut vor dem Hintergrund von Modellen der Anpassung im Alter interpretieren: Sowohl im Rahmen des SOK-Modells (Baltes & Baltes, 1990a) als auch des Prozessmodells der Anpassung von Brandstädter (z. B. Brandstädter & Greve, 1994b) werden adaptive Strategien der Verarbeitung bzw. Bewältigung zunächst aversiver individueller Entwicklungen beschrieben. Die Strategien der Selektion, Optimierung und Kompensation, welche eher auf die konkrete Verhaltensebene Bezug nehmen, wie auch ein akkomodativer Verarbeitungsmodus, welcher primär auf die kognitiv-emotionale Bewertung von Entwicklungsverläufen Bezug nimmt, ermöglichen einen in Hinblick auf Wohlbefinden und Zufriedenheit förderlichen Umgang mit objektiv ungünstigen Bedingungen (vgl. Ebner

et al., 2006). Dabei greifen verschiedene Mechanismen ineinander: Einschränkungen des Handlungsspielraums können in ihrer Bedeutung für das subjektive Wohlbefinden durch eine bewusste Selektion von Optionen mit größerer Handlungskompetenz oder positiverer Bewertung abgemildert werden – dies würde die Verringerung der Bandbreite bedeutsamer Lebensbereiche in der Studienstichprobe erklären. Gleichzeitig kann eine akkomodative Verarbeitung Einschränkungen und nachlassende Fähigkeiten in den bereits selektierten Bereichen ausgleichen, was die auch in der älteren Teilstichprobe hohen Ausprägungen auf dem FLQM und der globalen Zufriedenheit plausibel machen würde.

Die beschriebenen Studienergebnisse fügen sich in das geläufige Bild des „Paradox’ des subjektiven Wohlbefindens“ ein und sprechen daher für die Inhaltsvalidität des FLQM. Trotz der mit dem Alter im Schnitt zunehmenden Einschränkungen auf gesundheitlicher, kognitiv-psychischer und sozialer Ebene und trotz der Verkleinerung des individuellen Systems bedeutsamer Lebensbereiche besteht unvermindert die Möglichkeit, dieses Leben als erfüllt und erfüllend wahrzunehmen. Der Nutzen des FLQM könnte unter anderem darin liegen, nicht nur den reinen Bewertungsaspekt, d. h. die quantitative Seite, des Paradox’ des subjektiven Wohlbefindens zu betrachten, sondern simultan auch relativ unkompliziert Informationen über die *inhaltlichen Grundlagen* der subjektiven Zufriedenheitsbewertungen zugänglich zu machen. Informationen über Stabilität und Veränderungen hinsichtlich der individuell bedeutsamen Lebensbereiche sowie ihrer relativen Bedeutsamkeiten erweitern nicht nur den Blick auf die Prozesse, welche dem Entstehen subjektiver Zufriedenheit zu Grunde liegen. Sie bieten vielmehr eine verbreiterte Basis für individuelle, ressourcenorientierte Diagnostik und Interventionsplanung in der geriatrischen psychosozialen und medizinischen-pflegerischen Versorgung.

#### **4.6 Methodische Einschränkungen und Grenzen der Interpretierbarkeit**

Das gemischte Design der vorgestellten Studien mit qualitativen und quantitativen Anteilen bietet viele Möglichkeiten, birgt jedoch auch Risiken. Qualitative Daten sind mindestens genauso anfällig für Verzerrungen durch Interviewereffekte und Antworttendenzen wie quantitative Daten (Flick, 1995, 2002; Schwarz, 1999). Da in einer Interviewsituation notwendigerweise sämtliche Informationen direkt an den Interviewer gegeben werden, sind Interviewdaten für bestimmte Verzerrungen möglicherweise sogar anfälliger: Tendenzen zum sozial erwünschten Antwortverhalten oder bewusste

Verzerrungen bei Fragen zu sehr intimen und/oder schambesetzten Themen müssen bei der Bewertung und Interpretation der Antworten einkalkuliert werden (Diekmann, 2007). Das gilt beispielsweise für die Bereiche Sexualität und Inkontinenz (vgl. Ahnis, 2005).

Die vorgestellte Gesamtstudie (Studien 1 und 2) ist eine Pilotuntersuchung, mit allen Beschränkungen, denen eine Studie dieser Größenordnung hinsichtlich der statistischen Auswertung unterliegt. Die Stichprobengröße der Hauptstudie (Studie 2) von  $N = 44$  – und die entsprechenden Größen von Teilpopulationen wie Altersgruppen oder Geschlechter – verbietet die Anwendung zahlreicher Auswertungsverfahren und erschwert die Ableitung allgemeingültiger Schlüsse aus den Daten. Insbesondere beim Vergleich von Teilpopulationen, d. h. Altersgruppen, Geschlecht und Bildungstand, müssen die Ergebnisse vorsichtig interpretiert werden und sollten lediglich als Hinweise auf die Richtung von Zusammenhängen oder Unterschieden verstanden werden.

Problematischer als die geringe Stichprobengröße indes ist die Art der Stichprobe: Die Teilnehmer wurden mehr oder weniger unsystematisch als Gelegenheitsstichprobe aus einer sehr stark vorselektierten empirischen Population gezogen. Selektionseffekte speisen sich vor allem aus dem Umstand, dass die Teilnehmer von sich aus Kontakt zum Studienleiter aufnehmen, also ein relativ großes Maß an Eigeninitiative zeigen mussten. Es handelt es sich also vermutlich um Personen, die Interesse an sozialen Kontakten haben, eine relativ große Kommunikationsbereitschaft besitzen und über die Offenheit verfügen, im Gespräch sehr private Vorstellungen, Bedürfnisse und Wünsche zu äußern. Diese Eigenschaften haben möglicherweise die genannten Lebensbereiche nicht unwesentlich geprägt. Unter Umständen stehen das Interesse und die Offenheit der Teilnehmer auch in Zusammenhang mit der Reichhaltigkeit der individuellen Kategoriensysteme. Es ist denkbar, dass Personen, die sich noch stärker aus dem Leben und der sozialen Interaktion zurückziehen, eine geringere Bandbreite bedeutsamer Bereiche nennen würden (vgl. Cumming & Henry, 1961). Darüber hinaus sprechen Offenheit und Interesse für eine aktive Auseinandersetzung mit dem eigenen Leben und damit möglicherweise für eine affirmative, positive Grundhaltung, die sich in einer positiven Verzerrung von Zufriedenheit und Lebensqualität in der Stichprobe äußern könnte.

Insofern es sich um Teilnehmer aus den Pflegediensten oder der Praxis für physikalische Therapie handelt, wurden diese Personen zudem im Vorfeld bereits vom eingebundenen Personal der jeweiligen Einrichtung ausgewählt – auch dies stellt eine Verzerrung dar. Vermutlich wurden diejenigen Klienten eher ausgewählt, denen die

jeweiligen Pflegedienstleitungen oder Therapeuten eine größere soziale Kompetenz und einen besseren kognitiven Status zugeschrieben. Auch hier könnte man Verzerrungen des Kategoriensystems sowie des FLQM-Gesamtscores vermuten. Obwohl Alters- und Geschlechterverteilungen annähernd der deutschen Gesamtpopulation entsprechen, gilt dies nicht für die Population der mehrfach erkrankten älteren Menschen (vgl. Robert-Koch-Institut, 2003): Insbesondere der Anteil der Hochbetagten müsste in der Stichprobe deutlich höher liegen, um annähernd repräsentativ zu sein. Allerdings war eine populationsrepräsentative Erhebung auch kein Anliegen dieser Pilotstudie. Vielmehr war ein wichtiges Anliegen, alle Altersgruppen mit in die Untersuchung einzubinden und einen möglichst breiten Altersbereich abzudecken. Auch hinsichtlich der Häufigkeiten bestimmter Erkrankungen bzw. Beschwerdeprofile kann nicht von Repräsentativität ausgegangen werden. In diesem Kontext stellen sich epidemiologische Probleme: Einerseits ist das Sampling einer in Hinblick auf Erkrankungen repräsentativen Stichprobe methodisch sehr aufwändig – es ist eine genaue klinische Abklärung aller Komorbiditäten erforderlich – , andererseits erschwert die relativ schlechte Datenlage zu Prävalenzen von Krankheiten und Krankheitskombinationen (in bestimmten Teilstichproben und der Gesamtbevölkerung) die konkrete Bestimmung der abzubildenden Verteilungen in der Grundgesamtheit (vgl. Abschnitt 1.2.4).

Zur Prüfung der Konstruktvalidität des FLQM wurden Annahmen über Zusammenhänge mit verschiedenen Skalen aufgestellt – auch hier mussten, wie bereits an gegebener Stelle erwähnt, Beschränkungen durch die Stichprobengröße hingenommen werden (vgl. Abschnitt 4.1). So wäre zur Überprüfung der Vorhersagekraft der unterschiedlichen Variablen und Skalen eigentlich eine Regressionsanalyse das Mittel der Wahl gewesen. Anhand der Korrelationsmatrix konnten die Einzelzusammenhänge und die durch die einzelnen Prädiktoren aufgeklärte Varianz des FLQM abgeschätzt werden. Eine gleichzeitige Analyse der jeweiligen Anteile an der gesamten Varianzaufklärung war jedoch aufgrund der Stichprobengröße nicht sinnvoll möglich. Die gegebene Stichprobe hätte lediglich die Einbeziehung von zwei Variablen in die Regressionsanalyse erlaubt, so dass durch die Vorselektion das Ergebnis bereits wesentlich determiniert gewesen wäre. Die Bestimmung der unabhängigen Varianz des FLQM wird somit nur innerhalb einer deutlich umfangreicheren Stichprobe erfolgen können.

Bei der Auswertung der mit dem FLQM assoziierten qualitativen Daten, nämlich den von den Teilnehmern benannten Lebensbereichen, machte sich die relativ kleine



Stichprobengröße besonders bemerkbar: Da zahlreiche Kategorien lediglich mit einer Häufigkeit unter fünf belegt waren, musste auf dieser Ebene fast gänzlich auf konventionelle Auswertungstechniken zum quantitativen Gruppenvergleich verzichtet werden. Die dargestellten Ergebnisse sind als *Hinweise* auf Unterschiede zwischen den verschiedenen Teilpopulationen bezüglich inhaltlicher Schwerpunkte der Bereichsnennungen zu verstehen. Ähnliches gilt für die Vergleiche zur Übereinstimmung der individuellen Nennungen mit den Beispielen aus der Anregungsliste/Instruktion. Mit einer größeren Zahl von Teilnehmern wäre es beispielsweise möglich, genauer nachzuvollziehen, welche differentiellen Merkmale als Prädiktoren einer besonders starken Anlehnung des individuellen Kategoriensystems an die Vorgaben angesehen werden können.

Ganz allgemein gilt für die qualitative Auswertung der Bereichsnennungen die Einschränkung, welcher jedwede Kategorisierung von Texten unterliegt: Die Zuordnung der Inhalte zu den Kategorien des Systems bzw. die aus den Daten generierten Kategorien sind teilweise ambivalent. So sind u. a. andere Klassenbezeichnungen denkbar und einige Äußerungen hätten inhaltlich mehreren verschiedenen Klassen zugeordnet werden können. Abhängig von der gewählten Abstraktionsebene der Kategorien können Klassengrößen sehr unterschiedlich ausfallen. Diese Umstände sind ein der deduktiven wie induktiven Kategorisierung inhärentes Problem, wenn es um die Systematisierung verbal vermittelter Inhalte geht (Flick, 2002; Mayring, 2003). Da die qualitativ-inhaltliche Auswertung jedoch nur ein sekundäres Ziel dieser Arbeit war, wurde dieser Umstand in Kauf genommen. Im Idealfall sollten, wie in der qualitativen Vorstudie zur Entwicklung des FLQM (Studie 1), zwei unabhängige Kodierer die Kategorisierungen vornehmen und Konsens bezüglich der Zuordnungen von Nennungen zu Kategorien erlangen. In Studie 2 wurde aufgrund des Pilotcharakters der Untersuchung auf dieses sehr viel aufwändigere Vorgehen verzichtet. Zukünftige inhaltliche Analysen der Bereichsnennungen sollten jedoch auf Urteilen mehrerer Kodierer fußen, um die Reliabilität der Klassifikation zu erhöhen.

#### **4.7 Zusammenfassung und abschließende Bewertung**

In dieser Forschungsarbeit wurde theoriebasiert ein Fragebogen zur Lebensqualität älterer, multimorbider Menschen (FLQM) entwickelt. Die Operationalisierung von „Lebensqualität“ im FLQM lehnt sich an das Modell der Lebenszufriedenheit von

Campbell et al. (1976) an – für die eigene Lebensqualität bedeutsame Bereiche werden von den Befragten selbst generiert, bewertet und gewichtet. Ähnliche, wenngleich nicht explizit theoriebasierte Ansätze wurden in den letzten Jahren z. B. von Ruta et al. (1994) oder Joyce et al. (2003) entwickelt. Obwohl sie mehr oder weniger erfolgreich bei jüngeren Patienten eingesetzt werden können, sind diese Ansätze aus verschiedenen Gründen zur Erfassung der Lebensqualität älterer Menschen mit multiplen Krankheitsgeschehen verwendet wenig geeignet: Sie sind eher defizitorientiert, bereiten vielen älteren Menschen Verständnisprobleme und die Bewertung und Gewichtung der individuell generierten Lebensbereiche ist sehr komplex und unübersichtlich. Zudem sind die Verfahren sehr zeitaufwändig. Eine Weiterentwicklung des „individualisierten“ Ansatzes für ältere, multimorbide Menschen scheint also geboten.

Multimorbide ältere Menschen sind hinsichtlich ihrer internalen kognitiven und affektiven Prozesse per se keine Sonderklasse Älterer: Ein Erhebungsinstrument für multimorbide Menschen sollte genauso valide für jüngere oder nicht mehrfach erkrankte Personen sein. Multimorbide Ältere bzw. für sie konstruierte Messinstrumente, wie der FLQM, sind gewissermaßen „Teilmengen“ der Population älterer Menschen bzw. altersspezifischer Erhebungsinstrumente. Diese Relation impliziert jedoch gleichzeitig, dass nicht automatisch jedes Messinstrument für ältere Menschen auch für diese besondere Population geeignet sein muss. Ihre Merkmale sind eine große Heterogenität von Symptom- und Krankheitsmustern und dementsprechend eine große Variationsbreite individueller Grenzen und Möglichkeiten der Lebensgestaltung. Diese „geballte Vielfalt“ hebt die Teilpopulation Multimorbider aus der Gesamtheit älterer Menschen hervor. Sie sollte reflektiert werden in einer heterogenen Erhebungsmethodik, welche die vielfältigen Lebensweisen und –welten dieser Menschen angemessen spiegeln kann. Vor diesem Hintergrund wurde der FLQM als individualisierter Fragebogen zur Lebensqualität multimorbider älterer Menschen entwickelt.

Die im Rahmen einer explorativen Konstruktvalidierung gefundenen Zusammenhänge mit parallel zum FLQM erhobenen Fragebögen zeigen die erwarteten Richtungen, fast ausnahmslos mit der theoretisch zu vermutenden Stärke. Da der FLQM sehr eng an ein Prozessmodell zur globalen Lebenszufriedenheit angelehnt ist, verwundert es nicht, dass der engste Zusammenhang zwischen dem FLQM und einer Frage zur allgemeinen Lebenszufriedenheit besteht. Hohe und mittlere Zusammenhänge mit globaler Lebenszufriedenheit, allgemeinem subjektiven Wohlbefinden und positivem Affekt weisen

auf die Gültigkeit der mit dem FLQM gewonnenen Daten hin. Den Beziehungen des FLQM zu *state*- und *trait*-Aspekten des affektiven Erlebens sollte jedoch nachgegangen werden, um eine Konfundierung mit aktuellem positivem Affekt ausschließen zu können. Es zeigte sich keine systematische Abhängigkeit der Messwerte von externen Variablen wie Alter, Geschlecht oder Bildungsstand. Auch die Anzahl der selbst berichteten Erkrankungen als Index der Krankheitsbelastung ging nicht mit einer schlechteren subjektiven Lebensqualität einher, jedoch spielt die subjektive Gesundheits- und Funktionseinschätzung anscheinend eine stärkere Rolle für die subjektive Lebenszufriedenheit als ursprünglich angenommen. Die Ergebnisse der Konstruktvalidierung ermutigen dazu, den Fragebogen FLQM weiter zu verfeinern und an einer größeren Stichprobe multimorbider älterer Menschen endgültig zu validieren.

Der Fragebogen besitzt maximale Inhaltsvalidität, da die Befragten selbst das Konstrukt explizieren, welches Inhalt des Fragebogens ist. Dadurch wird gewährleistet, dass ausschließlich Items im Fragebogen verwendet werden, die tatsächlich der subjektiven Konstruktion von „Lebensqualität“ entsprechen. Diese sehr starke Annahme setzt voraus, dass die Befragten die Instruktion – und damit die Intention – des FLQM korrekt interpretieren. Andernfalls entstünden schwer kalkulierbare Verzerrungen des Antwortverhaltens. Bei der Konstruktion des FLQM wurde daher besonderer Wert auf eine sorgfältige und präzise Frageformulierung gelegt. Zudem wurde etwa der Hälfte der Teilnehmer ein Fragebogen zum inhaltlichen Verständnis des FLQM und zur Akzeptanz des Fragebogenformats vorgelegt. Die Ergebnisse dieser Befragung sind ebenfalls positiv in Hinblick auf die Verwendbarkeit des FLQM bei älteren, multimorbiden Menschen zu interpretieren. Die Teilnehmer attestieren eine gute Verständlichkeit der Instruktionen und äußern sich positiv über das gewählte Darbietungsformat und die verwendeten Antwortformate.

Sowohl aus Sicht der Teilnehmer als auch von der Konzeption des FLQM kommt der Fragebogen nicht für eine Verwendung bei Personen mit kognitiven Einschränkungen in Frage. Eine Adaption des FLQM für Personen mit *beginnenden* kognitiven Beeinträchtigungen ist denkbar, müsste jedoch in noch stärkerem Maße auf ihre Verständlichkeit und Validität überprüft werden als die derzeitige Version des Fragebogens.

Im Rahmen des FLQM werden über die individuelle Generierung von bedeutsamen Lebensbereichen inhaltliche Informationen zur subjektiven Konstruktion von

„Lebensqualität“ gewonnen. Während in der Stichprobe keine wesentlichen Differenzen zwischen den Geschlechtern beobachtet werden konnten, unterschieden sich alte (80 Jahre und älter) und junge Alte (65-79 Jahre) in der Reichhaltigkeit der Bereichsnennungen. Ältere Menschen nannten nicht nur im Schnitt weniger bedeutsame Bereiche, sondern auch ein eingeschränktes Spektrum von Bereichskategorien mit einer größeren Häufigkeit. Dies deutet auf eine Verengung der Bandbreite zufriedenheitsstiftender Lebensbereiche im höheren Alter hin. Gleichzeitig berichteten die älteren Teilnehmer jedoch keine geringere Lebensqualität (gemessen mit dem FLQM) oder globale Lebenszufriedenheit. Unabhängig davon schätzten die Teilnehmer insgesamt ihre körperliche und funktionelle Gesundheit unterdurchschnittlich schlecht ein. Diese Befunde können im Lichte des SOK-Modells erfolgreichen Alterns von Baltes und Baltes (1990a) oder des Prozessmodells der Anpassung im Alter von Brandtstädter (Brandtstädter & Greve, 1994b) erklärt werden. Es ist möglich auch „objektiv“ widrige Umstände, wie schlechte Gesundheit, körperliche Einschränkungen oder soziale Isolation zu kompensieren, indem eine Anpassung der individuellen Ziele, Werte und Standards vorgenommen wird. Diese oft unbewusst ablaufenden Prozesse der Adaptation des alternden Selbst greifen nach den Ergebnissen dieser Studie anscheinend auch in der objektiv stark benachteiligten Gruppe mehrfach körperlich erkrankter älterer Menschen bis ins hohe Alter. Da es nicht primärer Inhalt dieser Studie war, einer solchen Fragestellung nachzugehen, wäre es wünschenswert, eine weitere, systematische Studie der individuellen Entwicklung der subjektiven *inhaltlichen* Konstruktion von Lebensqualität anzusetzen. Möglicherweise gelingt es so, sich der Auflösung des „Zufriedenheitsparadox“ inhaltlich anzunähern.

Sowohl Ansatzpunkte für Interventionen als auch die Evaluation deren Nutzen müssen die Perspektive der Betroffenen eng mit einbeziehen, um eine patientenzentrierte Optimierung der Situation multimorbider älterer Menschen zu ermöglichen. Dabei darf der Fokus nicht einseitig auf einer verallgemeinerten oder normorientierten Qualität der Bedingungen und Fähigkeiten liegen. Vielmehr sollte die Qualität des Erlebens gerade bei objektiv stark benachteiligten Personen in den Vordergrund treten, um auf individueller therapeutischer, pflegerischer und psychosozialer Ebene die Förderung und den Erhalt von Lebensqualität und Wohlbefinden zu unterstützen. Auch diesem Personenkreis mit einem deutlichen Übergewicht der „Verluste“ über die „Gewinne“ scheint, gemessen am subjektiven Wohlbefinden und der subjektiven Lebensqualität, noch ein „erfolgreiches Altern“ möglich zu sein.