



Leseprobe aus Brandt, Fietz, Hampel, Kaschowitz, Lazarevic, Reichert und Wolter, Methoden der empirischen Alter(n)sforschung, ISBN 978-3-7799-3752-4

© 2018 Beltz Juventa in der Verlagsgruppe Beltz, Weinheim Basel

<http://www.beltz.de/de/nc/verlagsgruppe-beltz/gesamtprogramm.html>

isbn=978-3-7799-3752-4

1 Methodische Herausforderungen in der empirischen Alter(n)sforschung

Martina Brandt, Monika Reichert, Jennifer Fietz,
Sarah Hampel, Judith Kaschowitz, Patrick Lazarevič
und Veronique Wolter

1.1 Warum dieses Buch?¹

Die Alters- und Alternsforschung ist nicht erst seit der Entdeckung des demografischen Wandels durch die Öffentlichkeit ein zentrales Feld in den Sozial- und Verhaltenswissenschaften: Will man mehr über die Sozialstruktur und deren Veränderungen wissen, ist Alter eine grundlegende Dimension; sind Lebens(ver)läufe im Fokus, so ist Altern immer Teil der Entwicklung. Darüber hinaus gewinnt die Alter(n)sforschung verstärkt auch in der angewandten Gerontologie an Relevanz, so z. B. bei der Evaluierung von Projekten. Alter(n) ist ein Querschnittsthema und betrifft alle Bereiche und Disziplinen der sozial- und verhaltenswissenschaftlichen Forschung. Umso erstaunlicher ist es, dass unseres Wissens kein aktuelles Handbuch über die zahlreichen spezifischen methodischen Herausforderungen und Probleme in der empirischen Sozialforschung zum Alter und Altern existiert (siehe aber z. B. Arbeitspapiere von Chandola/O’Shea 2013 und Kühn/Porst 1999 sowie das Buch von Motel-Klingebiel/Kelle 2002). Letztlich sollten alle ForscherInnen, die Daten Älterer aus Querschnitts-, Längsschnitts- oder Kohortenstudien verwenden oder Untersuchungen zur Evaluierung von spezifischen Maßnahmen für und mit älteren Menschen durchführen, auch diese methodischen Herausforderungen im Blick haben, da ihre Nicht-Beachtung Ergebnisse und Schlussfolgerungen substantiell verzerren kann.

In diesem einleitenden Kapitel werden zunächst grundsätzliche und spezifische methodische Herausforderungen der empirischen Alter(n)sforschung skizziert, die in den folgenden Kapiteln anhand ausgewählter Themen und praktischer Beispiele illustriert werden. Diese reichen von altersspezifischen Messfehlern (Kapitel 2) in quantitativen und qualitativen Befragungen schlecht

1 Wir danken Robert Balliet, Melanie Boekholt, Franziska Colak und Eva Mainusch für die tatkräftige Hilfe bei der Manuskripterstellung.

erreichbarer oder schwer befragbarer Älterer (Kapitel 3 und 4) bis hin zu besonderen Herausforderungen bei der Erfassung von Veränderungen in verschiedenen Lebensbereichen über den Lebenslauf (Kapitel 3) und bei der Evaluation von Wirkungen von Interventionen über die Zeit (Kapitel 5). Der Schwerpunkt dieser Kapitel liegt auf den Konsequenzen, die methodische Probleme für inhaltliche Schlussfolgerungen haben, und belegt damit gleichzeitig die praktische Relevanz für die empirische Sozialforschung. Im abschließenden Kapitel 6 geben die AutorInnen Einschätzungen und Empfehlungen zur Überwindung der überblicksartig und exemplarisch dargestellten methodischen Schwierigkeiten und umreißen den zukünftigen Forschungsbedarf in diesem Feld.

1.2 Methodische Herausforderungen im Überblick²

Eine Auswertung ist nur so gut wie die ihr zugrunde liegenden Informationen. Jede Methode und jedes Forschungsdesign hat dabei eigene Vorteile, Einschränkungen und Probleme. Dies gilt insbesondere für empirische Analysen in der Alter(n)sforschung und betrifft alle Schritte des Prozesses, von der Definition von „Alter“ und der Auswahl der Befragten über die Art der Erhebung, die Formulierung der Fragen und Antwortkategorien bis hin zur Analyseform und Interpretation. Unterschiedliche Forschungsmethoden und -designs beeinflussen einerseits die Aussagekraft der Ergebnisse an sich und andererseits die Vergleichbarkeit zwischen verschiedenen Erhebungen und Gruppen. Das bedeutet jedoch nicht, dass wir keine Aussagen treffen können. Es ist aber essenziell, darauf zu achten, welche Schlüsse auf welcher Basis zulässig sind. Es gilt darzulegen, mit welcher Form von Verzerrungen und methodischen Artefakten wir es zu tun haben (könnten), um die Ergebnisse realistisch zu interpretieren und einordnen zu können. Sind Schätzungen, beispielsweise in Bezug auf Gesundheit, eher „optimistisch“, da die Stichprobe aus vergleichsweise jungen, gesunden und aktiven Älteren besteht? Erscheint der Zusammenhang zwischen Gesundheit und sozialer Ungleichheit weniger deutlich („konservativ“ geschätzt), als er tatsächlich ist, da sozioökonomisch schlechter gestellte Personen aus ganz unterschiedlichen Gründen seltener Teil der Erhebung sind (u. a. unterschiedliche Erreichbarkeit und Antwortbereitschaft, Stichworte „Mittelstandsbias“ und „Survivorship Bias“, s. auch Brandt/Kaschowitz/Lazarevič 2016)? Sind Interventionsmaßnahmen deshalb erfolgreich, weil ohnehin jene älteren Menschen

2 Die überblickenden Ausführungen basieren teilweise auf dem Methodenkommentar zum Bundesteilhabebericht (Wacker et al. 2017, S. 49 ff.) und auf dem Kapitel „Gesundheit im Alter“ im Handbuch Gesundheitssoziologie (Brandt/Kaschowitz/Lazarevič 2016).

daran teilnehmen, die generell offener für eine Inanspruchnahme von Angeboten sind?

Es hat vor allem nichtmethodische Gründe, dass ein ausführlicher Methodenbericht inklusive umfassender Robustheitsprüfungen und kritischer Reflexion Hintergrund jeder Analyse sein sollte. Methodische Probleme können sich erheblich auf inhaltliche Schlüsse auswirken, z. B. wenn etwaige Ausfälle Befragter nicht zufällig sind. Dies trifft insbesondere für die Alter(n)sforschung zu, denn ältere und hochaltrige Personen in ihren diversen Lebenslagen sind in nahezu allen Untersuchungen eingeschränkt repräsentiert (s. z. B. auch Schnell 1991; 1997). Hinsichtlich der *Stichprobe* gilt unter anderem:

- bei Stichprobenziehungen sind „schwer erreichbare“ und „schwer befragbare“ Personen (z. B. in Institutionen Lebende oder solche mit schwerwiegenden Beeinträchtigungen) häufig unterrepräsentiert;
- die Administration des Fragebogens (telefonisch, schriftlich, computergestützt, verwendete Sprache, Visualisierungsformen etc.) hat Auswirkungen auf die Befragbarkeit unterschiedlicher Gruppen und schließt unter Umständen bestimmte Personenkreise (Menschen mit Demenz, Sinnesbeeinträchtigte, Nicht-Muttersprachliche etc.) systematisch aus;
- über Personengruppen, die quantitativ wenig vertreten sind (z. B. hochaltrige MigrantInnen), können keine statistisch verlässlichen Aussagen gemacht werden;

sowie bezüglich des *Antwortverhaltens*:

- kognitive Beeinträchtigungen (z. B. in Verbindung mit demenziellen Erkrankungen) erschweren systematisch die Gewinnung von Selbstberichten der Betroffenen; hier werden z. B. Proxy-Interviews und Beobachtungen eingesetzt, die wiederum gegebenenfalls nicht mit Selbstberichten vergleichbar sind.
- bei Selbstberichten kann das Antwortverhalten (gerade im Beisein einer/s Interviewers/in) oder einer anderen Person, z. B. durch soziale Erwünschtheit) objektive Tatsachen verzerren.

Unterschiedliche Verfahren zur Vermeidung und Einschätzung von Ausfällen (z. B. „Proxy“- und „Doorstep“-Interviews) bis hin zur Erhebung von Biomarkern und der nachgelagerten Verwendung von Analyse-Gewichten erlauben die Eingrenzung der Auswirkungen auf die Ergebnisse oder gar die Korrektur derselben, wenn hinreichende Details über die nichtrepräsentierte/n Gruppe/n bekannt sind. Unterschiedliche Designs (Quer- oder Längsschnitt, Trend oder Panel, komparativ) ermöglichen oder verhindern darüber hinaus bis zu einem gewissen Grad die Identifikation bzw. Trennung von Alters-, Kohorten- und

Periodeneffekten, die ja alle gleichermaßen zeitgebunden sind (z. B. Mayer/Huinink 1990). Die Integration einer Längsschnittperspektive ist ein wichtiger Schritt, gerade wenn es um die Beschreibung des gesellschaftlichen Wandels, die Beobachtung von Verläufen und die Aufdeckung von Zusammenhängen oder Kausalitäten (vorher – nachher) geht. Allerdings ergeben sich bei Studien über die Zeit, die Verläufe und Alternsprozesse in den Blick nehmen, wiederum einige spezifische Problemstellungen im Hinblick auf die *Stichprobe*:

- selbst wenn alle Personengruppen erreicht oder beobachtet werden könnten, sind die Schwerstbeeinträchtigten oder jene in einem sehr hohen Alter mit fortschreitender Zeit eher verstorben (selektive Mortalität, „Survivorship Bias“);
- Studien, die dieselben Personen über mehrere Zeitpunkte beobachten, leiden häufig unter „Panelmortalität“, d. h. bestimmte Befragte nehmen aus unterschiedlichen Gründen (z. B. schlechte Interviewerfahrung, Zeitmangel, Umzug) nicht über alle Erhebungen hinweg an diesen teil;

wie auch in Bezug auf das *Antwortverhalten*:

- retrospektive Informationen können durch Erinnerungsschwierigkeiten oder sogenanntes „Anchoring“ oder „Coloring“ (Interpretation aus Sicht der Gegenwart; siehe z. B. Havari/Mazzonna 2015) verzerrt sein.

Die Beobachtung einer Person über die Zeit und über Veränderungen kontextueller Bedingungen (z. B. Sozialpolitik, Gesundheitssystem) ermöglicht Aussagen darüber, ob eine zuvor gegebene Situation sich infolge veränderter Bedingungen wandelt – bestenfalls unter Berücksichtigung weiterer wichtiger Faktoren wie z. B. grundlegenden Eigenschaften der Person und ihrer sozio-ökonomischen Lage, um sogenannte „unbeobachtete Heterogenität“ zu minimieren. So existiert eine immer größere Zahl an nationalen, repräsentativ angelegten Längsschnitt- und Panel-Studien, die Menschen durch die zweite Lebenshälfte begleiten (z. B. Interdisziplinäre Längsschnittstudie des Erwachsenenalters (ILSE), Sattler et al. 2017; Deutscher Alterssurvey (DEAS), Mahne et al. 2017), und Studien, die besonders auf Gruppen wie Hochaltrige oder demenziell Erkrankte in quantitativen (z. B. Repräsentativbefragung NRW80+, Wagner et al. 2018; Bayerischer Demenz Survey, online unter www.izph.fau.de/projekte/bayerischer-demenz-survey) und qualitativen Settings (s. Kapitel 4 und 5) gerichtet sind.

Im Hinblick auf sozialpolitische Ziele und Konsequenzen lohnt unter Umständen auch ein Blick in die Nachbarländer und andere Wohlfahrtsstaaten. Die europäische oder globale Einordnung der deutschen Gegebenheiten in Bezug auf Lebenslagen verschiedener Kohorten und Altersgruppen ist ein wich-

tiges Forschungsfeld. Zur Anwendung kommen hier speziell auf Ältere gerichtete international vergleichende Panelstudien (z. B. Survey of Health, Ageing, and Retirement in Europe SHARE und Schwesterstudien; Börsch-Supan et al. 2013) oder nationale Paneldaten, die mit vergleichbaren Studien aus anderen Ländern analysiert werden können (z. B. Cross-National Equivalent File CNEF des Sozio-oekonomischen Panels SOEP; Frick et al. 2008). Solche Erhebungen widmen sich nicht nur gezielt der sogenannten „Panelpflege“, die die Interviewten zur weiteren Teilnahme an der Studie motivieren soll, sondern auch der Vergleichbarkeitsproblematik von der Studienanlage bis hin zu bereitgestellten statistischen Korrekturen. Die vielfältigen Verfahren, die zur Auswertung dieser komplexen Daten zur Anwendung kommen, erfordern wie auch die dazugehörigen Statistikprogramme viel statistisch-mathematische Einarbeitung und praktische Analyse-Erfahrung – zumal immer größere Datenmengen und mehr unterschiedliche Datenarten zu Verfügung stehen und Hand in Hand damit die Entwicklung neuer Verfahren rasant vorangeht.

Bei der Verwendung von qualitativ orientierten Methoden, und hier insbesondere bei qualitativen Interviews, sind zusätzliche Aspekte zu beachten. Bei der Stichprobenziehung handelt es sich meist um ein bewusstes theoretisches Sampling (Glaser/Strauss 1967), denn in aller Regel steht die Untersuchung von ausgewählten (Rand-)Gruppen im Vordergrund (z. B. von Armut oder von häuslicher Gewalt betroffene ältere Menschen, Menschen mit Demenz). Dies hat zur Folge, dass qualitative Untersuchungsansätze eher mit kleinen Stichproben arbeiten (können), wobei ihr Umfang und ihre Verteilungsmerkmale vor der eigentlichen Untersuchungsdurchführung nicht festgelegt werden. Vielmehr wird die Stichprobenbeziehung idealerweise dann beendet, wenn die sogenannte „Sättigung“ erreicht ist, d. h. wenn weitere Fälle keinen Erkenntnisgewinn mehr versprechen (Glaser/Strauss 1967). Zwar erheben qualitative Daten keinen Anspruch auf statistische Repräsentativität; durch das häufig angewandte Prinzip der Sättigung kann aber dennoch eine „moderate Generalisierung“ von Ergebnissen erreicht werden (Mayring 2007; vgl. auch Flick 2016). Hinzu kommt, dass gerade qualitative Interviews hohe Anforderungen an die kommunikative und soziale Kompetenz sowie an die Aufmerksamkeits- und Konzentrationsfähigkeit von InterviewerInnen stellen (Doering/Bortz 2016): Diese müssen sich nicht nur schnell auf die zu Befragenden und auf auftretende Probleme in der Interviewsituation einstellen, sondern bei leitfadengestützten Interviews ggf. gezielt nachfragen, ohne aber gleichzeitig das Antwortverhalten zu beeinflussen. Schließlich sind die Auswertungsverfahren, die bei qualitativer Forschung Anwendung finden, sehr aufwendig und anspruchsvoll. Bei der Transkription von qualitativen Interviews – an sich schon sehr arbeitsintensiv – fällt eine enorme Datenfülle an. Darüber hinaus erfordert die Auswertungstechnik, von der „Grounded Theory“ (Strauss/Corbin 1996) bis hin zur Inhaltsanalyse nach Mayring (2015; siehe auch Kuckartz 2010), um nur zwei Beispiele

zu nennen, viel Einarbeitung in die entsprechenden Verfahren. Auch die Durchführung von Gruppeninterviews bzw. Fokusgruppen (Merton 1987), ein ebenfalls qualitativ orientiertes Verfahren zur Ermittlung unterschiedlicher Perspektiven zu einem Sachverhalt, ist nicht voraussetzungslos und erfordert sorgfältige Planung, Durchführung und Auswertung. Es gilt hier vor allem, die Vor- und Nachteile gegenüber einem Einzelinterview abzuwägen. Jede Studie hat so ihre methodischen Besonderheiten, auf die entsprechend reagiert werden muss.

1.3 Die folgenden Kapitel und Beispiele

Dieses Buch soll dazu dienen, einige der oben skizzierten Schwierigkeiten unter Bezugnahme auf konkrete Beispiele zu identifizieren und praktische Lösungswege aufzuzeigen. Wir behandeln anhand unserer eigenen Forschung wichtige Aspekte der Konzipierung, Durchführung und Auswertung von Studien über die Methodengrenzen hinweg und geben so einen Einblick in die häufig nicht explizit beleuchteten Hintergründe der empirischen Alter(n)sforschung.

Im anschließenden Kapitel 2 beschäftigen wir uns mit altersspezifischem Antwortverhalten und untersuchen, inwieweit die Beantwortung der Frage nach dem subjektiven Gesundheitszustand vom Alter abhängig ist. Wir gehen dem in zwei Schritten nach: Zuerst stellen wir ein theoretisches Modell zu kognitiven Prozessen bei der Beantwortung der Frage nach dem allgemeinen Gesundheitszustand sowie zur statistischen Erklärung der Antwort vor. In einem zweiten Schritt werden ausgewählte empirische Befunde vorgestellt und in ihrer Bedeutung für die theoretischen Modelle diskutiert. Ziel des Beitrags ist es, zu einem besseren Verständnis des mit Abstand wichtigsten Gesundheitsindikators, dem selbstberichteten allgemeinen Gesundheitszustand, im Rahmen der empirischen Alter(n)sforschung beizutragen.

Die Möglichkeiten und Grenzen der quantitativen Alter(n)sforschung im Hinblick auf gering- bzw. unterrepräsentierte kleinere Subgruppen auszuloten, ist Thema in Kapitel 3. Hier werden die Unterstützungsnetzwerke älterer MigrantInnen im Zeitverlauf auf Basis des SOEP betrachtet. Unsere Betrachtung zeigt: Die repräsentative Erhebung sozialer Netzwerke und Unterstützungsleistungen für MigrantInnen verschiedenster Herkunft sowie die Analyse solcher Daten ist mit großen methodischen Herausforderungen verbunden und nicht zuletzt aufgrund des Anstiegs der Zahl von ZuwanderInnen ein spannendes und auch in Zukunft bedeutendes Forschungsfeld.

Im Weiteren werden methodische Problemstellungen und Lösungsansätze bei qualitativen Studien zu Menschen mit Demenz und ihren pflegenden Angehörigen vorgestellt. In Kapitel 4 geht es um praktische Hindernisse bei der Durchführung qualitativer Interviews im Setting häuslicher Pflege von Men-

schen mit Demenz. Sowohl die Befragung der Erkrankten selbst als auch die ihrer pflegenden Angehörigen stellen besondere Anforderungen an Forschende. Wir thematisieren zielgruppenspezifische Besonderheiten im Forschungsprozess, die von der Projektkonzeption bis zur Durchführung reichen, und geben Hinweise, wie Interviews in diesem Setting dennoch einen wichtigen Erkenntnisgewinn generieren können.

Kapitel 5 zur Evaluation von Interventionsprogrammen in einer alternden Gesellschaft beschäftigt sich eingehend mit der Methode der Fokusgruppeninterviews im Rahmen der Evaluation von Interventionsprogrammen in der Alter(n)sforschung. Obwohl Gruppendiskussionsverfahren nicht immer direkt den klassischen Interviewmethoden zugeordnet werden, sondern eher als Moderationstechnik zur Erhebung eines kollektiven Orientierungsrahmens gelten, nehmen sie in der qualitativen Interviewforschung einen bedeutenden Stellenwert ein. Ziel des Kapitels ist es, die Chancen, Herausforderungen und Grenzen von Fokusgruppeninterviews in Evaluationsprozessen in der aktuellen Literatur aufzuzeigen und anschließend anhand von Praxisbeispielen zu verdeutlichen.

Abschließend ziehen wir in Kapitel 6 ein Fazit über alle in den Kapiteln beleuchteten Forschungsprojekte hinweg, nehmen offene Forschungsfragen und Lösungsstrategien in den Blick und geben einen optimistischen Ausblick in die Zukunft der Alter(n)sforschung.

2 Ausgezeichnet, sehr gut, gut, mittelmäßig, schlecht

Theoretische Konzepte und empirische Befunde
zur Erhebung des allgemeinen Gesundheits-
zustandes bei älteren und alten Menschen

Patrick Lazarevič

- Es wird ein kognitives Modell des Antwortprozesses selbst eingeschätzter Gesundheit aufgestellt.
- Darauf aufbauend folgt ein empirisch umsetzbares Modell zur Grundlage der Gesundheitsbewertung.
- Die Modelle lassen sich zur Konzeption und Umsetzung empirischer Untersuchungen nutzen.
- Das Modell wird anhand ausgewählter empirischer Befunde und eigener Untersuchungen illustriert.
- Implikationen und Anknüpfungspunkte für die Forschungspraxis werden diskutiert.

2.1 Die Selbsteinschätzung der Gesundheit als „State of the Art“?

Gesundheit ist als Querschnittsthema ein zentraler Aspekt von empirischen Studien in vielen wissenschaftlichen Disziplinen, wie z. B. den Sozialwissenschaften, Politikwissenschaften, der Ökonomie und natürlich der Epidemiologie und Gerontologie. Gerade im Zuge der Bevölkerungsalterung beschäftigt sich eine stetig wachsende Anzahl quantitativ-empirischer Studien mit Fragen individueller und gesellschaftlicher Gesundheit (vgl. Hank/Brandt 2016). Dabei stehen oft entweder die Verbesserung von Gesundheit bzw. die Verlängerung gesunder Lebenszeit oder die Bedingungen und Auswirkungen gesundheitlicher Ungleichheit auf diverse andere Lebensbereiche im Mittelpunkt (vgl. z. B. Brandt/ Deindl/Hank 2012; Deindl/Brandt/Hank 2016; Hank/Deindl/Brandt 2013; Kaschowitz/Brandt 2017). Dies gilt umso mehr für die Alter(n)sforschung bzw. Gerontologie, da sich insbesondere Menschen im höheren Alter mit einer sich verschlechternden Gesundheit konfrontiert sehen. Darum ist es auch nicht

verwunderlich, dass große Altersstudien wie z. B. die Berliner Altersstudie (BASE), der DEAS oder SHARE immer auch diese Aspekte berücksichtigen. Die Ergebnisse der auf diesen Erhebungen basierenden Untersuchungen haben eine hohe gesellschaftliche Relevanz und dienen als wissenschaftliche Basis für entsprechende politische Maßnahmen wie die aktuelle deutsche Pflegereform (vgl. Brandt/Kaschowitz/Lazarevič 2016).

Analytischer Dreh- und Angelpunkt sämtlicher derartiger (sozial-)gerontologischer, (sozial-)epidemiologischer und artverwandter Studien ist die Definition und Messung von Gesundheit. Zur generischen Messung der Gesundheit, d. h. zur Messung der allgemeinen gesundheitlichen Verfassung (auch latente Gesundheit genannt) einer Person, gibt es eine Vielzahl umfassender Verfahren (vgl. Jylhä 2009). Die Methoden hierzu reichen von ausführlichen Fragebögen, wie dem Health Utilities Index (HUI-3) (vgl. Feeny et al. 1995), dem Cohen-Hoberman Inventory of Physical Symptoms (CHIPS) (vgl. Cohen 1988), dem Short Form Health Survey (SF-36) (vgl. Ware 1993), dem European Health Status Module (EHSM) (vgl. Robine/Jagger 2003), über physische Leistungstests, wie die Greifkraftmessung oder die Messung der Gehgeschwindigkeit (für einen Überblick vgl. Cooper et al. 2011), bis hin zu den sogenannten Biomarkern, also Laboruntersuchungen von Blut-, Speichel-, Urin- oder Haarproben (für einen Überblick vgl. Mayeux 2004; McDade/Williams/Snodgrass 2007). In vielen Untersuchungen insbesondere im Bereich der Surveyforschung stehen jedoch weder die nötigen zeitlichen noch finanziellen Ressourcen zur Verfügung, um eine solche umfassende Gesundheitsmessung zu gewährleisten. Zudem ist zu beachten, dass speziell in multithematischen Studien eine ausführliche Messung der Gesundheit durch eine längere Befragungsdauer und einen größeren Aufwand für die Befragten die sogenannte „Respondent Burden“ erhöhen kann, wodurch potenziell die Datenqualität leidet (vgl. Bradburn 1979; Sharp/Frankel 1983). Gleichzeitig können diese Methoden aber auch abschreckend für ForscherInnen sein, da diese weiterführende Kenntnisse statistischer Methoden (z. B. Faktorenanalysen oder Strukturgleichungsmodelle im Falle von Skalen) bzw. der Humanbiologie/-medizin (im Falle von Biomarkern) oder einen größeren Aufwand (z. B. bei der Kodierung des HUI-3) von den ForscherInnen und gegebenenfalls RezipientInnen erfordern.

Daher greifen sowohl Fragebogendesigner großer Surveys wie ForscherInnen als Alternative häufig auf direkte subjektive Selbsteinschätzungen der Befragten zurück, die gewöhnlich in Form einer Einzelfrage erhoben werden, z. B. durch Formulierungen wie „Würden Sie sagen, Ihr Gesundheitszustand ist ...?“ (Ausgezeichnet/Sehr gut/Gut/Mittelmäßig/Schlecht) (Fragebogen des SHARE 2015). Der Reiz der Messung der allgemeinen Gesundheit über diese Frage aus surveymethodologischer Sicht liegt darin, dass es sich um eine vergleichsweise einfache und kurze Frage handelt, zu der vermeintlich alle Befragten entweder eine Meinung haben oder gegebenenfalls schnell sich eine solche bilden können (vgl. Jylhä

2009). Hierdurch erscheint die Implementation dieser Frage – und damit einer allgemeinen Messung von Gesundheit – im Fragebogen eines Surveys genau wie ihre Verwendung durch ForscherInnen tendenziell unproblematisch. Entsprechend wird diese Frage – mit leichten Variationen – nicht nur in den oben genannten Alterssurveys, sondern auch in vielen allgemeinen gesundheitsbezogenen (z. B. Gesundheit in Deutschland aktuell (GEDA)) wie multithematischen Bevölkerungssurveys (z. B. im SOEP) national wie international standardmäßig verwendet, was eine generelle Vergleichbarkeit der Messung impliziert. Dabei wird dieser Indikator in der Datenanalyse üblicherweise entweder quasimetrisch (vgl. z. B. Leinonen 2002; Manderbacka/Lahelma/Martikainen 1998) verwendet oder vereinfacht in Kategorien wie „mindestens gute Gesundheit“ und „schlechter als gut“ (vgl. z. B. Manor/Matthews/Power 2000), wodurch er sich problemlos für eine Vielzahl statistischer Methoden eignet.

Empirisch lässt sich die Erhebung und Verwendung dieses auch als „Self-Rated Health“ (SRH) bezeichneten generischen Gesundheitsindikators mindestens dadurch begründen, dass sich in zahlreichen Studien immer wieder zeigen ließ, dass er einen konsistenten, unabhängigen und merklichen Beitrag zur Erklärung späterer Mortalität leistet (für einen Überblick vgl. Idler/Benyamini 1997). Aber auch auf theoretischer Ebene werden insbesondere vier Eigenschaften von SRH zur Begründung seiner Nützlichkeit herangezogen (vgl. Idler/Benyamini 1997; Benyamini 2011):

Erstens sind diese Bewertungen inklusiv in dem Sinne, dass Befragte bei der Bewertung sämtliche für sie relevanten Aspekte berücksichtigen können. Dies trifft nicht nur auf manifeste und bekannte Tatsachen wie Diagnosen oder das Vorliegen von Symptomen zu, die möglicherweise nicht abgefragt werden, sondern auch auf eher latente Eigenschaften wie die Schwere und Auswirkung von Krankheiten oder auch den Befragten unbewusste biomedizinische Prozesse wie Entzündungen, die nur über Biomarker erfasst werden könnten. Weiterhin handelt es sich *zweitens* bei diesen Gesundheitseinschätzungen um dynamische Bewertungen, die nicht nur den (objektiven) aktuellen Gesundheitszustand, sondern auch individuelle Gesundheitsentwicklungen und Veränderungen berücksichtigen. Es werden also auch kurz- und langfristige Änderungen in die Wertung miteinbezogen, die gegebenenfalls fundamentale Auswirkungen auf die Bewertung haben. *Drittens* beeinflusst die Einschätzung der eigenen Gesundheit vermutlich auch das Verhalten von Befragten, da ein schlechterer Gesundheitszustand z. B. die Motivation für präventives Gesundheitsverhalten wie physische Aktivität verringert. Zuletzt lässt sich *viertens* auch argumentieren, dass in die Selbsteinschätzung der Gesundheit auch Ressourcen wie der sozioökonomische Status, soziale Unterstützung und eine allgemeine Vitalität von Befragten einbezogen werden. Diese Ressourcen können ihrerseits einen eigenständigen Einfluss auf den Gesundheitsstatus, z. B. in Form einer besseren Bewältigung gesundheitlicher Rückschläge, haben.