

Isabella Hager

Statistik für die Pflegewissenschaft

Isabella Hager

Statistik für die Pflegewissenschaft

Teil 1: Grundbegriffe und wichtigste Testverfahren

2. Auflage

Isabella Hager

... ist diplomierte Gesundheits- und Krankenpflegerin, war unter anderem in der Palliativ-Pflege und Behindertenarbeit tätig, hat sich im Rahmen ihres Soziologiestudiums auf Statistik spezialisiert und ist heute Forschende, Beratende und Lehrende, unter anderem an der Universität Wien, der FH Campus Wien, bei prospect – Research & Solutions und beim Wiener Projektzentrum, Institut für Sozialforschung, Bildung, Integration und Kunst.

Mein Dank ergeht an

Eva Hellsberg, die die duftenden Korinthenkuchen aus dem Backrohr holt,
für ihr unermüdliches und unbarmherziges Wurmlesen

Elisabeth Brousek, die mich immer erinnert, was Wissenschaftlichkeit ausmacht

Martina Koller, für ihre tollen Ideen und Anregungen

Philipp Weinfurter, für seine Kreativität und Geduld

Anselm Eder, der mir den Weg gewiesen hat

Ulrike Grittner, für den letzten Schliff

© 2020 Isabella Hager

2. Auflage

Isabella Hager

Umschlaggestaltung, Illustration: Philipp Weinfurter

Lektorat, Korrektorat: Elisabeth Brousek, Eva Hellsberg, Birgit Schönfelder

Verlag: Buchschmiede von Dataform Media GmbH, Wien

ISBN: 978-3-99110-800-9 (Paperback)

ISBN: 978-3-99110-871-9 (Hardcover)

ISBN: 978-3-99110-872-6 (e-Book)

Printed in Austria

Das Werk, einschließlich seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung des Verlages und der Autorin unzulässig. Dies gilt insbesondere für die elektronische oder sonstige Vervielfältigung, Übersetzung, Verbreitung und öffentliche Zugänglichmachung.

In diesem Buch wird die geschlechtsneutrale Sprache mittels * verwendet.

...zu diesem Buch...

Jeder Mensch kann Statistik verstehen und anwenden! In diesem Buch werden Sie die wichtigsten Grundbegriffe und Verfahren der quantitativen Pflegeforschung am Beispiel von echten Daten und Musterbeispielen erlernen und verstehen.

Statistische Verfahren in der Pflegewissenschaft stehen vor einer großen Herausforderung: Die Datengewinnung ist in der Praxis meist sehr mühsam und zeitaufwändig. Es besteht die Herausforderung, trotz geringer Fallzahlen möglichst viel „aus den Daten herauszuholen“ und dabei alle wissenschaftlichen Kriterien zu erfüllen. Die praktische Anwendbarkeit sowie der unmittelbare Nutzen stehen dabei im Vordergrund.

Statistik ist ein Werkzeug, das Ihnen dazu dient, Ihre Arbeit zu beobachten und dadurch zu verbessern. Die Evaluation von Pflegeinterventionen schließt neben der Überprüfung von Gruppenunterschieden auch ein, dass Maßnahmen auf ihre Wirksamkeit untersucht werden. Statistische Methoden sind bestens dazu geeignet! Das vorliegende Buch dient als Gebrauchsanweisung für die Konzeption und Auswertung Ihrer Forschungsprojekte!

Isabella Hager, September 2020

Kennzeichnungen

Wichtiger Merksatz

Fragestellung

Beispiel

Aufgabe

Interpretation eines Ergebnisses

Technisch-methodische Bemerkung

Literatur: Ich empfehle die folgende Literatur zusätzlich zur Vertiefung.

Koller, M. (jeweils neueste Auflage): Statistik für Pflege- und andere Gesundheitsberufe, facultas 2018	perfekter Überblick
Bortz, J. & Döring, N. (jeweils neueste Auflage): Statistik für Sozialwissenschaftler. Heidelberg: Springer-Verlag.	die „Bibel“ - das „Alte Testament“
Bortz, J. & Döring, N. (jeweils neueste Auflage): Forschungsmethoden und Evaluation. Heidelberg: Springer-Verlag.	die „Bibel“ - das „Neue Testament“
Field, A. (jeweils neueste Auflage): Discovering Statistics using IBM SPSS. London, California, New Dehli, Singapore. SAGE Publications Ltd.	... and sex and drugs and rock´n roll !!
Rasch, B. et al. (2014): Quantitative Methoden Band 1, Einführung in die Statistik für Psychologen und Sozialwissenschaftler, Springer Verlag Berlin	anspruchsvoll
Hatzinger R., Nagel H. (2013): Statistik mit SPSS, Pearson-Studium.	gute SPSS-Anleitungen
Rauch G., Neumann K., Grittner U., Herrmann C., Kruppa J. (2019): Medizinische Statistik für dummies. Weinheim: Wiley-Vch Verlag GmbH &	übersichtlich und gut
Porst, R. (2011) Fragebogen: VS Verlag für Sozialwissenschaften, Springer Fachmedien Wiesbaden	
Flaker, V. & Schmid T. (Hrsg.) (2006): Von der Idee zur Forschungsarbeit. Forschen in Sozialarbeit und Sozialwissenschaft. Böhlau Verlag Wien	
Bühner, M. (2004): Einführung in die Test- und Fragebogenkonstruktion. München: Pearson-Studium.	
Petrie, A., Sabin C. (2009): Medical statistics at a Glance, Wiley Blackwell, Oxford, England.	
Jede Art von Handbuch für SPSS ist fürs praktische Arbeiten geeignet!	

Inhalt

1	Fragestellung und Operationalisierung	1
1.1	Was ist eigentlich empirische Sozialforschung?	1
1.2	Das Formulieren einer Fragestellung	2
1.3	Konzeption und Operationalisierung	4
1.4	Hypothese und Hypothesenmodell	7
1.5	Pflegediagnostik und Pflegedokumentation	9

2	Grundgesamtheit und Stichprobe	10
2.1	Repräsentativität – oder: Der Zufall ist der „Master of the Universe“	12
2.2	Konfidenzintervall und Schwankungsbreiten	16
2.3	Wie groß soll/muss eine Stichprobe sein?	17
2.4	Stichprobenarten	18

3	Fragebogenkonstruktion	21
3.1	Gütekriterien von Messinstrumenten	21
3.2	Sozialwissenschaftliche Befragungen	22
3.3	Der Aufbau des Fragebogens	23
3.4	Die Elf Gebote zur Fragebogenerstellung	24

4	Datenniveaus – kategorisieren und messen	25
4.1	Nominalskala (kategorial)	26
4.2	Ordinalskala oder Rangskala (kategorial)	26
4.3	Intervallskala (metrisch)	27
4.4	Rationalskala/Verhältnisskala/Absolutskala (metrisch)	27
4.5	Das Stufenmodell der Datenniveaus	28
4.6	Kriterien einer Antwortskala	29
4.7	Arten von Antwortskalen	30
4.8	Zur Verwendung einer „neutralen“ Mittelkategorie	30
4.9	Textantworten	31
4.10	Die Erhebung: Vom Fragebogen zur Datenmatrix	32

5	Beschreibung von Häufigkeiten	34
5.1	Häufigkeitstabelle	34
5.2	Mehrfachantworttabelle	36
5.3	Berechnen eines Zähl-Index aus einer Mehrfachantwort	37
5.4	Kreuztabelle	39

6	Statistische Kennwerte	41
6.1	Berechnung der statistischen Kennwerte	42
6.2	Anwendung und Interpretation statistischer Kennwerte	47
6.2.1	Kennwerte bei nominalem Messniveau	47
6.2.2	Kennwerte bei ordinalem Messniveau	48
6.2.3	Kennzahlen bei metrischem Messniveau (1)	49
6.2.4	Exkurs: Feinberechnung von Perzentilen	51
6.2.5	Kennzahlen bei metrischem Messniveau (2)	52
6.2.6	Boxplot – grafische Darstellung von Verteilungen	54
6.3	Verteilungsformen	56

7	Die Normalverteilung	58
7.1	Die empirische Normalverteilung	58
7.2	Wann ist eine empirische Variable normalverteilt?	60
7.3	Die theoretische Normalverteilung	62
7.4	Standardnormen normalverteilter Testwerte	63
7.5	Berechnung von Konfidenzintervallen	65

8	Prüfmaß und Wahrscheinlichkeit – die Logik von Signifikanztests	68
8.1	Die Vorgangsweise bei einem Signifikanztest	70
8.2	Die Durchführung eines Signifikanztests am Beispiel Chi-Quadrat-Test	71
8.3	Hypothesen testen – eine Übersicht über die Signifikanztests	77
8.4	Anwendungsregeln – Übersicht	78
8.5	Entscheidungsbaum – Check-List	79

9	Parametrische Testverfahren	80
9.1	t-Test für unabhängige Stichproben: Testen eines Mittelwertunterschiedes zwischen zwei Gruppen	80
9.1.1	Überprüfen der Normalverteilung mit dem Kolmogorov-Smirnov-Test	82
9.2	Einfaktorielle Varianzanalyse: Testen eines Mittelwertunterschiedes zwischen mehreren Gruppen	84
9.3	Mehrfaktorielle Varianzanalyse: Testen eines Mittelwertunterschiedes anhand mehrerer Einflussfaktoren	88
9.4	t-Test für abhängige Messungen: Testen eines Mittelwertunterschiedes zwischen zwei Messzeitpunkten	90
9.5	Einfaktorielle Varianzanalyse mit Messwiederholung: Testen eines Mittelwertunterschiedes zwischen mehreren Messzeitpunkten	95
9.6	Mehrfaktorielle Varianzanalyse bei Messwiederholung: Testen einer Mittelwertveränderung anhand mehrerer Einflussfaktoren	99

10	Parameterfreie Testverfahren	103
10.1	Mann-Whitney-U-Test: Testen eines Verteilungsunterschiedes zwischen zwei Gruppen	103
10.2	Kruskal-Wallis-Test: Testen eines Verteilungsunterschiedes zwischen mehreren Gruppen	106
10.3	Wilcoxon-Test: Testen einer Veränderung zwischen zwei Messzeitpunkten	109
10.3.1	Exkurs: Die Signifikanz in Abhängigkeit von der Fallzahl	112
10.4	Friedman-Test: Testen einer Veränderung zwischen mehreren Messzeitpunkten	113
<hr/>		
11	Korrelation – Stärke eines Zusammenhanges von zwei Merkmalen	116
11.1	Stärke des Zusammenhanges bei Kreuztabellen	117
11.2	Pearson-Korrelation	119
11.3	Reliabilitätsanalyse: Berechnen eines Summen-Index mithilfe der Korrelation	127
11.4	Spearman-Rang-Korrelation	132
<hr/>		
12	Zufall oder Nicht-Zufall? – Das ist hier die Frage!	
	Überlegungen zum Zufall und zur Signifikanz	134
12.1	„The Art of Fehler“ oder: Des widerspenstigen Zufalls Zähmung	137
12.2	Stichprobengröße und Signifikanz	138
12.3	Effektstärke	140
12.4	Teststärke und Poweranalyse	143
<hr/>		
13	Forschung in der Pflegepraxis	145
13.1	Verlaufsdarstellung (ohne wiederholte Interventionen)	145
13.2	Verlaufsdarstellung (mit wiederholten Interventionen)	147
13.3	Messen von Veränderungen	148
13.4	Abteilungsvergleich mittels Kennzahlen	151
13.5	Kleiner Ausblick auf multivariate Verfahren	154
<hr/>		
14	Stichwortverzeichnis	155
15	Literatur	156

1 Fragestellung und Operationalisierung

Zu Beginn müssen wir einige Grundbegriffe klären, um uns das „quantitative Denken“ anzueignen. Um eine quantitative Studie erfolgreich durchzuführen, sind bei der Vorbereitung der Durchführung folgende Arbeitsschritte notwendig:

- Formulieren der **Fragestellung**
- Definition und **Operationalisierung** der darin enthaltenen Begriffe und Messdimensionen
- Formulieren von **Hypothesen** und Erstellen eines Hypothesenmodells
- Definition der **Grundgesamtheit** und Auswahl der **Stichprobe**

1.1 Was ist eigentlich empirische Sozialforschung?

Wenn wir empirische Sozialforschung in einem **Kontinuum von Strukturiertheit und Vorwissen** sehen, dann steht auf einer Seite der qualitative und auf der anderen Seite der quantitative Forschungsansatz. Je nach Fragestellung wird die dafür geeignete Methode ausgewählt, oder es erfolgt eine Kombination beider Methoden, wie bei der triangulativen Forschung.

Beim qualitativen Forschen geht es darum, die Logik eines Einzelfalls zu verstehen und daraus allgemeine Schlüsse zu ziehen. Das quantitative Paradigma benötigt eine ausreichende Menge an untersuchten Objekten, um daraus Gesetzmäßigkeiten abzuleiten.

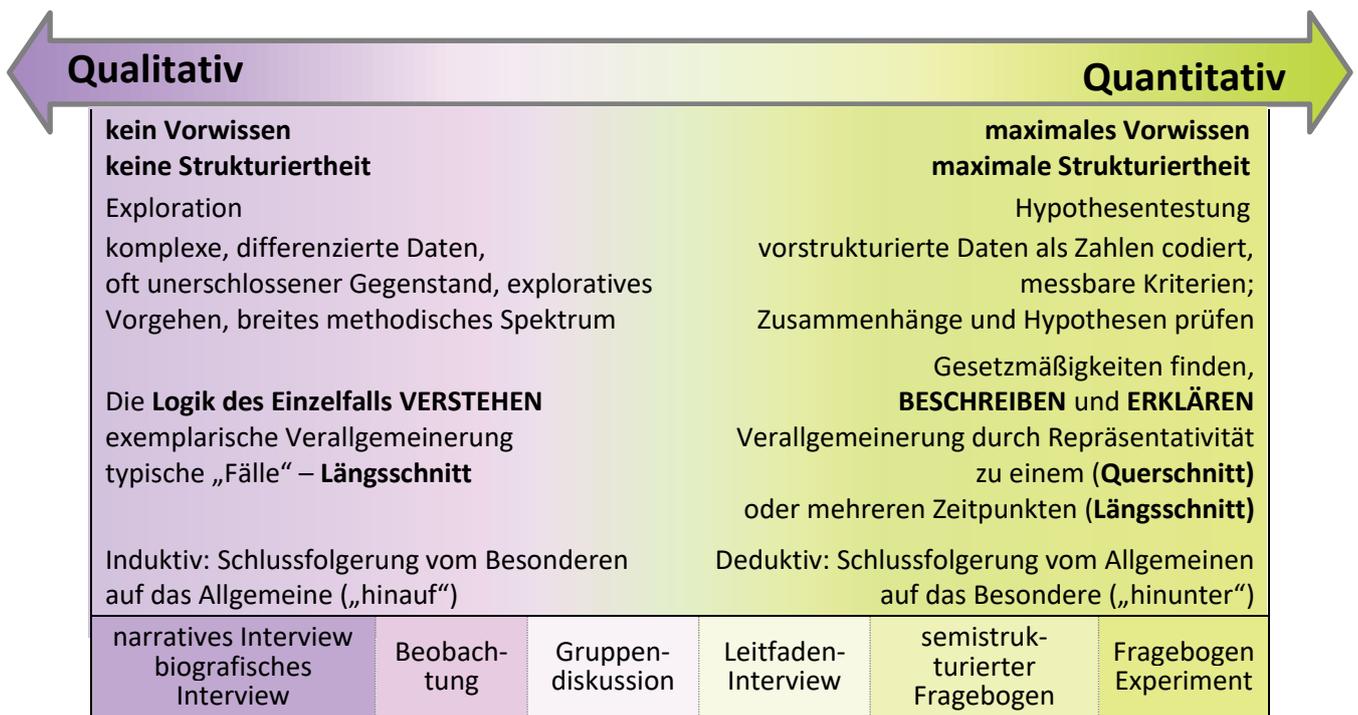


Abbildung 1: Überblick qualitative und quantitative Sozialforschung – Das Kontinuum der Datenerhebung

Die quantitative Forschung zeichnet sich dadurch aus, dass maximales Vorwissen bei maximaler Strukturiertheit benötigt wird. Antworten oder Eigenschaften von Menschen werden auf Messwerte oder Kategorien („Schubladen“) reduziert, die nur dann Sinn ergeben, wenn wir wissen, was wir da eigentlich messen. Das bedeutet: Beim quantitativen Forschen findet die meiste und wichtigste Arbeit **vor** der Erhebung statt: nämlich die Operationalisierung der Kriterien, die gemessen werden sollen und die Konstruktion des Messinstruments. Beim qualitativen Forschen kommt die meiste und wichtigste Arbeit erst **nach** der Erhebung: das qualitative Auswerten!

Der Durchführung einer quantitativen Studie muss daher eine umfangreiche Planung vorausgehen. Der erste und wichtigste Arbeitsschritt ist das Formulieren einer Fragestellung.

1.2 Das Formulieren einer Fragestellung

Am Anfang einer empirischen Forschungsarbeit steht die Entwicklung einer **Fragestellung**.

Eine wissenschaftliche Fragestellung muss systematisch formuliert und empirisch überprüfbar sein.

- „**Empirisch**“ heißt: „*erfahrungsgemäß*“: beobachtbar bzw. messbar anhand theoretisch formulierter Annahmen, definiert in Begriffen und Messdimensionen. Diese Annahmen (Hypothesen) werden anhand spezifischer sozialer Wirklichkeiten¹ überprüft.
- „**Systematisch**“ heißt: die Überprüfung erfolgt nach festgelegten Regeln (nach einer „Methode“) und ist daher intersubjektiv **nachvollziehbar**, im Unterschied zur Alltagsbeobachtung, die immer subjektiv bleibt. Die Systematik bezieht sich auf drei Grundkriterien der Wissenschaftlichkeit: Sie ist **strukturiert** nach einem geplanten Ablauf und einem festgelegten Design, sie ist **allgemein**, da sie auf einer Stichprobenauswahl beruht, die bestimmte Kriterien erfüllt (z.B. Repräsentativität) und sie ist **kontrolliert selektiv**, das heißt, es gibt festgelegte Kriterien darüber, was beobachtet bzw. gemessen wird.

Eine quantitative Fragestellung muss so formuliert sein, dass sie mit quantitativen Methoden beantwortbar ist. Das klingt trivial, ist aber oft ein schwieriges Unterfangen. Dabei müssen wir uns folgende Fragen stellen:

1.) WAS?	Was will ich eigentlich wissen?	→ Gefragt ist: O-RIGINALITÄT
2.) WARUM?	Wozu will ich das wissen? Wem nützen die Ergebnisse?	→ Gefragt ist: RE-LEVANZ
3.) WIE?	Was muss ich tun, um das herauszufinden? (Zugang, Erhebung, Design...)	→ Gefragt ist: SO-RGFALT

Abbildung 2: Die drei Fragen zur Entwicklung einer Fragestellung nach dem „ORESOSchema“²

Die erste Frage ist die schwerste und erfordert oftmals viel Geduld, bis sie beantwortet werden kann. Bei der zweiten Frage stellt sich heraus, ob die Studie realisiert wird – finden sich Geldgeber*innen, die für die Ergebnisse zahlen? – wenn ja, dann wird sie – sobald wir bei der dritten Frage angelangt sind – immer konkreter und auf bestimmte Gruppen, einen bestimmten Zeitraum, eine bestimmte Region eingegrenzt, um beantwortbar zu werden. Meist muss die erste Frage erneut angepasst werden, um die Fragestellung beantworten zu können. Dieser Kreislauf setzt sich so lange fort, bis die **Fragestellung systematisch formuliert und empirisch überprüfbar** ist.

Im **Entdeckungszusammenhang** wird das „Problem“ benannt, die Fragestellung formuliert und zeitlich, sachlich und räumlich eingegrenzt. Diese Eingrenzung ist essenziell, da immer nur Ausschnitte der sozialen Wirklichkeit empirisch erfasst werden können. Im Rahmen des **Begründungszusammenhangs** wird das Design geplant, das Messinstrument konstruiert und die Erhebung mit anschließender Analyse der gewonnenen Daten durchgeführt. Schließlich werden im **Verwertungszusammenhang** die Ergebnisse anwendungsorientiert dargestellt und interpretiert.

Gute Forschungsarbeit zeichnet sich dadurch aus, dass neue Fragen entstehen, an denen andere Forscher*innen weiter ansetzen können und eine ständige, reflexive Weiterentwicklung der Erkenntnisgewinnung stattfindet!

¹ Was ist **soziale Wirklichkeit**? Darunter verstehen wir Institutionen (z.B. Krankenhaus), Kommunikation, Interaktion, soziale, pflegerische, oder medizinische Merkmale (z.B. Geschlecht, Alter, Mobilität, Pflegestufe, Diagnose ...), mitgeteilte Einstellungen, Empfindungen, Erfahrungen, Emotionen, (z.B. Schmerzempfinden, Symptomempfinden...), beobachtetes oder mitgeteiltes Verhalten (z.B. gesundheitsrelevantes Verhalten, Inanspruchnahme von Vorsorgeuntersuchungen...) und vieles mehr.

Soziale Wirklichkeit ist also (wie der Wiener Kabarettist Gunkl es ausdrücken würde: "eh alles"), was wir Menschen geschaffen haben, miteinander teilen und das von unserem Handeln abhängig ist.

² Das „ORESOSchema“ wurde nach Anregung von Anselm Eder, meinem Statistikprofessor, entwickelt.

Daraus folgt eine **Anleitung zum Formulieren einer quantitativen Fragestellung**:

- 1.) Erfasse dein Erkenntnisinteresse in einem Satz! Was ist das genaue Problem, das du bearbeiten willst?
- 2.) Formuliere die Problemstellung – dabei wird dir die Vielschichtigkeit des Themas klar...
- 3.) Aus der/den Problemstellungen heraus kannst du deine konkrete Forschungsfrage entwickeln.
 - Stelle konkrete Fragen zum Problem! Teile dein Problem in einzelne Teilaspekte auf!
 - Mehrere einzelne Fragen sind zulässig! Aber: Finde dazu eine übergeordnete Fragestellung!
 - Die Fragestellung bestimmt den weiteren Weg der Forschung: den Ansatz, das Design, die Methoden, den Datenerhebungsmodus und die Vorgangsweise bei der Auswertung.

Woran du erkennst, dass deine Fragestellung wissenschaftlich ist – im Unterschied zur Alltagsfrage:

- Sie ist theoretisch und ihre Beantwortung leistet einen Beitrag zur Entwicklung des Themas
- Sie ist relevant, d.h. sie besitzt praktischen bzw. wissenschaftlichen Nutzen
...ist sie es nicht? Frag dich, warum!

Stolpersteine bei der Formulierung einer Fragestellung: Eine quantitative Fragestellung...

- ...muss einen wissenschaftlichen **Erkenntniswert** haben.
Wissenschaftliche Forschungsfragen sind nicht zu einfach, sondern haben eine gewisse Komplexität und müssen an einen Erkenntnisaspekt anschließen!

Was ist der häufigste Grund, warum Menschen in die Notfallambulanz kommen?	Inwiefern unterscheiden sich die Umstände des Notfallgeschehens zwischen Ambulanzen in städtischen oder ländlichen Regionen? 
--	--

- ...muss im Rahmen einer Studie **realistisch beantwortbar** sein.

Würden mehr Notfall-Patient*innen überleben, wenn das Personal verdoppelt würde?	In welchem Zusammenhang steht der Personalschlüssel mit der Überlebenschance nach einer Notaufnahme? 
--	--

- ...muss mit **quantitativen Methoden beantwortbar** sein (messen und zählen!)

Welche Bedeutung hat ein Notfall für die Patient*innen?	Unterscheiden sich die Gründe, warum Menschen in die Notfallambulanz kommen, nach soziodemografischen Merkmalen? 
---	--

- ...muss als Ziel die **Beschreibung** eines Phänomens haben – und nicht das Verstehen!
Quantitative Forschungsfragen zielen auf das Auffinden von Gesetzmäßigkeiten ab, wodurch die beobachteten Umstände erklärt werden können.

Warum kommen manche Menschen bei bestimmten Symptomen in die Notfallambulanz und andere nicht?	Beeinflusst der Schweregrad der Symptome die individuelle Inanspruchnahme der Notfallambulanz? 
--	--

- ...muss **begrenzt und präzise** sein! Was hilft? Eingrenzen!

Wie gut erholen sich Menschen nach Notfällen?	In welchem Zusammenhang steht das Alter mit der Verweildauer im Krankenhaus bei Patient*innen nach Kohlenmonoxidvergiftung? 
---	---

- ...darf **nur eine (Haupt-)Frage** beinhalten (Eingrenzen und notfalls die Fragestellung aufsplitten)!

Wie erleben alte Menschen einen Krankenhausaufenthalt, spielt dabei eine konstante Ansprechperson eine Rolle und in welchem Zusammenhang steht dies mit der Selbständigkeit und der subjektiven Lebensqualität zum Zeitpunkt der Entlassung?	Erreichen betagte Patient*innen auf einer Station mit Bezugspflegesystem zum Zeitpunkt ihrer Entlassung eine höhere Selbständigkeit als Patient*innen auf einer Station mit Funktionspflegesystem? 
--	--

Beispiel aus dem Pflegealltag: Patient*innen mit Hörsturz auf der HNO-Station am SMZO (Hiemetzberger et al 2014); Am Anfang bestimmt meist eine nicht überprüfbare, diffuse Alltagsfragestellung das grundlegende Interesse: Wie geht es den Patient*innen? Hilft ihnen unsere angebotene Therapeutic Touch-Behandlung? Nach Literaturrecherche, Feldforschung und Expert*inneninterviews wird eine wissenschaftliche Fragestellung formuliert: **Verbessert sich das Symptomempfinden von Patient*innen nach Hörsturz durch die Behandlung Therapeutic Touch im Vergleich zu jenen, die keine Behandlung in Anspruch genommen haben?**
Jetzt müssen wir „nur noch“ definieren, wie wir das „Symptomempfinden“ messen....

4 Fragestellung und Operationalisierung

Beispiel aus der Pflegeausbildung: In einer Pflegeakademie werden die Absolvent*innen regelmäßig mit einem Fragebogen über ihre pflegerische Handlungskompetenz (als Kriterium für die Ausbildungsqualität) befragt. Als neue Lehrmethode wird eine Lernstation für die Studierenden eingeführt. Die Alltagsfragestellung: „Bringt´s was?“
Wissenschaftliche Fragestellung: **Verändert sich die pflegerische Handlungskompetenz aus Sicht der Absolvent*innen nach der Teilnahme an der Lernstation?**
Fragt sich jetzt, wie ist „*pflegerische Handlungskompetenz*“ definiert...?

Aufgabe: Formulieren Sie eine **quantitative pflegewissenschaftliche Fragestellung!**
Beschreiben Sie kurz die Einrichtung/den Bereich, in der/in dem die Forschungsarbeit zur Beantwortung der Fragestellung durchgeführt werden soll. Warum ist Ihre Fragestellung relevant? Welche*r Geldgeber*in würde für die Ergebnisse bezahlen? Wem nützen die Ergebnisse?

1.3 Konzeption und Operationalisierung

Der nächste Arbeitsschritt ist die Operationalisierung. Wir erinnern uns: Beim quantitativen Forschen müssen wir vor der Erhebung bereits ganz genau wissen, was wir messen wollen:

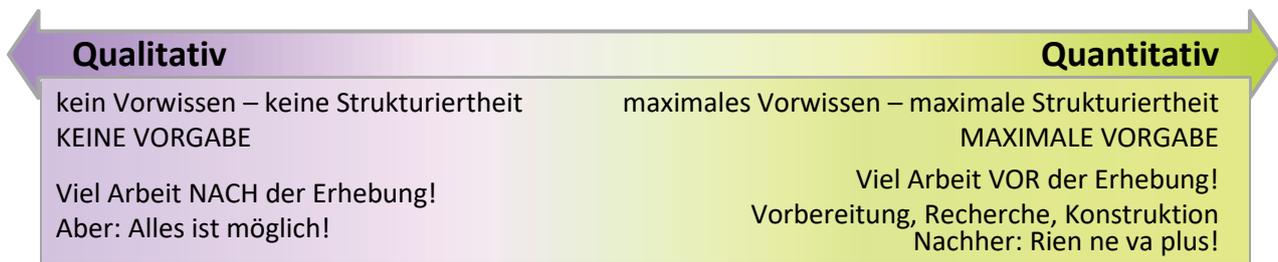


Abbildung 3: Arbeitsschwerpunkte bei qualitativer versus quantitativer Forschung

Diese Vorbereitung, das „Messbar-Machen“, sind konzeptionelle Definition und Operationalisierung der in der Fragestellung enthaltenen **Begriffe** und **Messdimensionen**. Dabei muss präzise definiert werden, was genau auf welche Weise gemessen werden soll.

Alle Begriffe/Messdimensionen in der Fragestellung müssen erst definiert und dann operationalisiert werden. Die Definition und die Operationalisierung der Inhalte einer Fragestellung bestimmen das Forschungsdesign und die Ergebnisse!

- **Konzeptionelle Definition:** Was bedeutet der Begriff / die Messdimension?
Welche theoretischen Aspekte / Konstrukte umfassen den Begriff / die Messdimension?
- **Operationalisierung:** Wie kann der Begriff / die Messdimension messbar gemacht werden?

Die Operationalisierung ist das UM-UND-AUF einer quantitativen Studie – sie bestimmt die Ergebnisse! Alles, was unpassend operationalisiert oder übersehen wurde, geht verloren! Um die relevanten Inhalte messbar zu machen, sind eine **theoretische Basis**, umfangreiche **Literaturrecherche** (Hat schon jemand an diesem Thema geforscht?), **explorative Interviews** mit Zielgruppenpersonen bzw. Expert*innen notwendig. Die Operationalisierung kann nur gelingen, wenn die Fragestellung und das Ziel präzise formuliert und die verwendeten Konzepte exakt definiert sind. Nur auf diese Weise kann sichergestellt werden, dass relevante Inhalte sinnvoll operationalisiert werden und nichts Wichtiges übersehen wird.

Wichtigste Kriterien für die Operationalisierung sind die Zweckmäßigkeit zur Beantwortung der Fragestellung sowie die Sicherstellung der Vergleichbarkeit mit anderen Studienergebnissen.

Beispiel: Konzeptionelle Definition des Begriffs „rückfällige Klient*innen“ auf einer Langzeittherapiestation für wiederholt Heroinabhängige: Die Fragestellung lautet: *Welche Faktoren beeinflussen die Rückfälligkeit von Klient*innen mit vollendetem Heroinentzug nach einer Langzeittherapie?*
→ *Rückfällige Klient*innen sind jene, die 6 Monate nach der Entlassung wieder harte Drogen konsumieren. Das sind: Heroin/Kokain/LSD in jeder Form sowie der Missbrauch von Substitutionsmedikamenten.*
Beachte: Bei einer Zeitspanne von 6 Monaten wird die Rückfallquote höher sein als bei 3 Monaten!

Beispiel Armut: Um im Rahmen einer länderübergreifenden Umfrage die latente Dimension „**Armut**“ messbar zu machen, so benötigen wir zuerst eine **konzeptionelle Definition von Armut**: „*Armut ist die mangelnde Befriedigung der Grundbedürfnisse nach Kleidung, Nahrung, Wohnung und Erhaltung des Lebens.*“

Hierzu können mehrere **Indikatoren** herangezogen werden:

1. Einkommen/Haushaltseinkommen/Geld, das monatlich zur Verfügung steht
2. Auskommen mit dem Einkommen/das Vorhandensein von Schulden
3. Auskommen müssen ohne ausreichend Nahrung
4. Auskommen müssen ohne entsprechende Kleidung
5. Wohnsituation, Wohnungsmängel
6. Betroffenheit von Arbeitslosigkeit
7. Betroffenheit von Analphabetismus
8. Soziale Absicherung im Falle von Arbeitslosigkeit, Krankheit und Alter
9. usw. usw. usw.

„**Armut**“ kann noch viel mehr sein! Für das, wie ich „**Armut**“ messe (= operationalisiere), gibt es keine „richtigen“ oder „falschen“ Kriterien, vielmehr werden inhaltliche Entscheidungen getroffen, was unter „**Armut**“ verstanden werden **soll**. Diese Entscheidung bestimmt alle späteren Behauptungen und Aussagen über die von Armut betroffenen oder nicht betroffenen Personen. Das heißt auch: Keine Feststellung, die sich auf konstruierte Wirklichkeit bezieht, ist „objektiv“, sondern einfach nur „definiert“.

Worauf zielt die Fragestellung ab? Liegt der Fokus auf.....

- ...der **Kaufkraft**, benötige ich das Einkommen und das Auskommen mit dem Einkommen (ad 1. und 2.).
- ...**Lebensstandard** oder **Lebensqualität**, fokussiere ich das Auskommen-Müssen ohne ausreichend Nahrung bzw. entsprechende Kleidung, Wohnsituation (ad 3., 4. und 5.).
- ...**Exklusion aus sozialen Systemen versus Inklusion in soziale Systeme**, brauche ich die Betroffenheit von Arbeitslosigkeit bzw. Analphabetismus und die soziale Sicherheit (ad 6., 7. und 8.).
- ...**Gesundheit**, verwende ich das Auskommen müssen ohne ausreichend Nahrung oder entsprechender Kleidung, die Wohnsituation und Wohnungsmängel und soziale Absicherung (ad 3., 4., 5., 8.).
- ...**xxx** brauche ich..... (ad 9. und so weiter).

Große Umfrageinstrumente, die auf Ländervergleiche abzielen, beinhalten etliche vielschichtige Fragestellungen, für die meist standardisierte Indikatoren herangezogen werden, wie beispielsweise Systeme sozialer Indikatoren.

In der Forschungspraxis werden Begriffe bzw. latente Dimensionen wie „**Armut**“, „**Mobilität**“, „**Selbständigkeit**“, „**Wohlbefinden**“, „**Pflegeabhängigkeit**“ etc. als Messdimension verstanden, die exakt definiert und mithilfe mehrerer Indikatoren erhoben wird. Was ist eine Messdimension?

Eine **Messdimension** besteht aus mehreren Einzelaspekten (Items), welche alle „das Gleiche“ messen.

Wichtig ist die inhaltliche und empirische **Eindimensionalität** einer Messdimension.

Inhaltlich: Die erhobenen Aspekte passen inhaltlich unter **eine** Überschrift (z.B. „**Selbstsicherheit**“)

Empirisch: Die Aspekte und hängen wechselseitig miteinander zusammen, was empirisch ermittelt wird.

Beispiel für eine Messdimension im Rahmen einer Untersuchung über die Gesundheit von Langzeitarbeitslosen: Die aus der Psychologie stammende Messdimension „**Selbstsicherheit**“ ergibt sich aus dem Grad der Zustimmung zu folgenden Aussagen:

		nie						immer					
		①	②	③	④	⑤	⑥						
a	Ich mag mich, so wie ich bin	①	②	③	④	⑤	⑥						
b	Ich habe Angst, dass andere schlecht über mich reden	①	②	③	④	⑤	⑥						
c	Ich traue mir viel zu	①	②	③	④	⑤	⑥						
d	Ich zweifle an meinen fachlichen Fähigkeiten	①	②	③	④	⑤	⑥						
e	Ich finde meist eine Lösung für schwierige Probleme	①	②	③	④	⑤	⑥						
f	Ich schäme mich wegen meines Aussehens oder meiner Figur	①	②	③	④	⑤	⑥						
g	Schwierigkeiten sehe ich gelassen entgegen	①	②	③	④	⑤	⑥						
h	Ich finde Mittel und Wege, mich gegen Widerstand durchzusetzen	①	②	③	④	⑤	⑥						

Quelle: Operationalisierung der Messdimension „**Selbstsicherheit**“

(eine Auswahl aus FABU – Fragebogen zur Arbeitslosigkeit und Beschäftigungsunsicherheit, Reick (2005) Uni Dortmund)

Bei komplexeren Fragestellungen werden die in der Fragestellung verwendeten Konzepte in mehrere (latente) Messdimensionen aufgesplittet, die wiederum in Subdimensionen zerlegt werden und anhand verschiedener (manifeste) Indikatoren gemessen werden können.

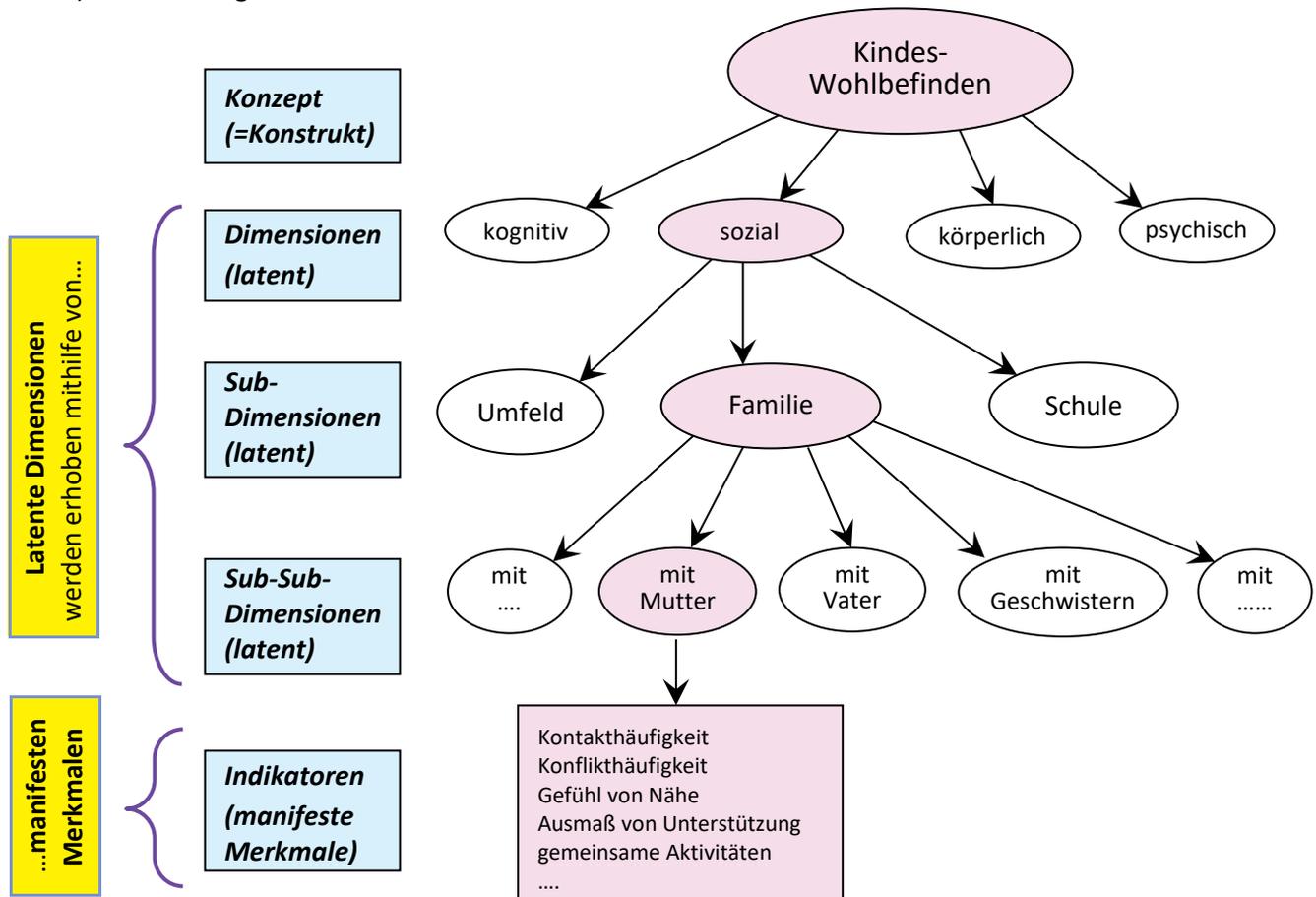


Abbildung 4: Beispiel einer Dimensionierung vom Konzept „Kindes-Wohlbefinden“ (Quelle: Moore & Theokas (2011))

Auch hier gilt: Es gibt keine „richtige“ oder „falsche“ Zusammenstellung der Items, vielmehr ist es eine auf inhaltlichen und methodischen Aspekten basierende Entscheidung, welche Items verwendet werden sollen.

Wir fragen also **nicht direkt** nach der latenten Dimension (etwa so: „Bitte schätzen Sie Ihre Selbstsicherheit auf der folgenden Skala ein.“), sondern: Als **Indikatoren** für die von uns gewählten Messdimensionen **bestimmen wir Teilaspekte in Form von formulierten Aussagen/Fragen**. Zu diesen Aussagen/Fragen erheben wir den Grad der Zustimmung bzw. die jeweilige Antwortmöglichkeit auf die Frage.

In diesem Zusammenhang wird auch von **Validität** (Gültigkeit) und **Reliabilität** (Zuverlässigkeit) von Messungen oder Messinstrumenten gesprochen. Die „**Konstruktvalidität**“ sagt aus, ob die ausgewählten Indikatoren das latente/theoretische Konstrukt (z.B. „Armut“, „Selbstwirksamkeit“) vollständig und passend messen (siehe auch Kapitel 3.1. auf Seite 21).

Um aus einer noch nicht getesteten „Fragebatterie“ (= Liste aus Items) die „besten“ (passendsten) auszuwählen, dienen Skalierungsverfahren:

- **Likert-Skala**: Sicherstellung möglichst großer Bandbreite der Antworten über die gesamte Antwortskala
- **Polaritätsprofil**: Gegensatzpaare von Begrifflichkeiten zur Messung von semantischen Bedeutungen, das sogenannte „semantisches Differenzial“
- **Guttman-Skala**: Konstruktion einer Skala mit steigender Zustimmung in einer theoretisch festgelegten und begründbaren Reihenfolge: Lässt sich die theoretische Abstufung auch empirisch beobachten?
- **Faktoranalyse**: Auffinden von latenten Subdimensionen innerhalb einer Batterie von Einzelaspekten
- **Reliabilitätsanalyse**: Testung der Güte der Eindimensionalität, d.h. alle Aspekte messen dasselbe.

Aufgabe: Definieren und operationalisieren Sie alle Begriffe und Messdimensionen in Ihrer pflegewissenschaftlichen Fragestellung!

1.4 Hypothese und Hypothesenmodell

Im besten Fall können die in der Fragestellung operationalisierten Begriffe und Messdimensionen in ein **Hypothesenmodell** integriert werden. Aber was ist eine Hypothese?

Eine **Hypothese** ist eine Vermutung / Behauptung über einen Sachverhalt, die überprüfbar ist.

Die Überprüfbarkeit ist der Unterschied zu Alltagshypothesen. Deshalb wird eine quantitative Hypothese niemals mit „Wie ist...“ beginnen, sondern ist in Form einer „Je-Desto“-Aussage formuliert.

Alltagshypothese	Wissenschaftliche Hypothese
„Wie ist die Lebensqualität von chronisch Kranken?“	„Die Betroffenheit von chronischer Krankheit geht einher mit einer geringeren Lebensqualität.“
„Die meisten Pflegekräfte sind an Weiterbildungen nicht interessiert.“	„Je höher die Qualifikation der Pflegekräfte, desto häufiger besuchen diese Weiterbildungsveranstaltungen.“
„Leuten, die lange arbeitslos sind, geht es schlecht.“	„Je länger die Dauer der Arbeitslosigkeit, desto negativer die Erfolgserwartung, einen Arbeitsplatz zu bekommen.“

Abbildung 5: Unterschied zwischen Alltagshypothesen und wissenschaftlichen Hypothesen

Hypothesen können verschiedene Sachverhalte betreffen, hier ein paar Beispiele:

Getestet werden Jugendliche in einer Wohngemeinschaft mit Milieuthherapie (Versuchsgruppe) und in einer WG ohne Milieuthherapie (Kontrollgruppe):

- **Unterschied:**
Jugendliche der Versuchsgruppe weisen eine höhere Konfliktfähigkeit auf als Jugendliche der Kontrollgruppe.
- **Veränderung:**
Nach der Teilnahme an der Therapie ist die Konfliktfähigkeit der Jugendlichen höher als vor der Teilnahme.
- **Zusammenhang:**
Je länger die Jugendlichen an der Therapie teilnehmen, desto höher ist deren Konfliktfähigkeit.

Eine Fragestellung besteht zumeist aus mehreren Einzelhypothesen, die im Optimalfall in einem **Hypothesenmodell** dargestellt werden können. Das Modell postuliert den theoretischen Hintergrund mit den wichtigsten Hypothesen der Fragestellung.

Wie könnte so ein Hypothesenmodell aussehen? Zum Beispiel so.....

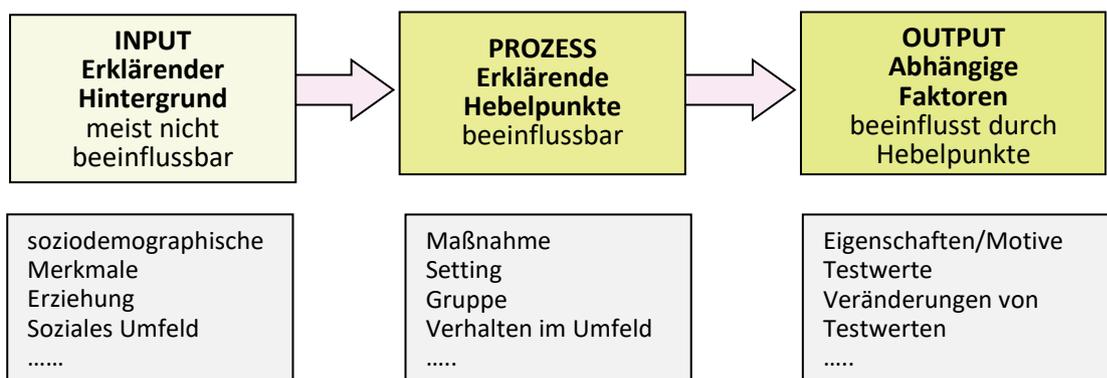


Abbildung 6: Schema eines möglichen Hypothesenmodells

8 Fragestellung und Operationalisierung

Befüllt mit einer Fragestellung aus der HBSC-Umfrage³ der WHO betreffend das Rauchverhalten von Schüler*innen kann das Modell so aussehen.....

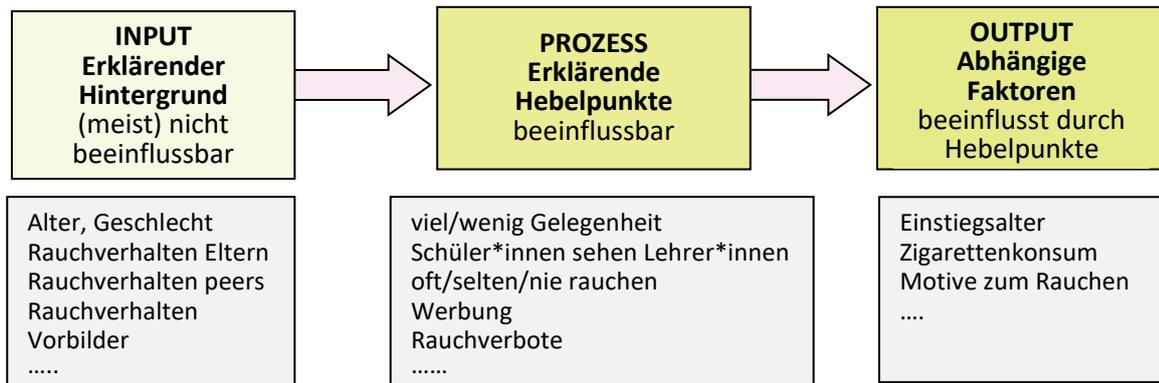
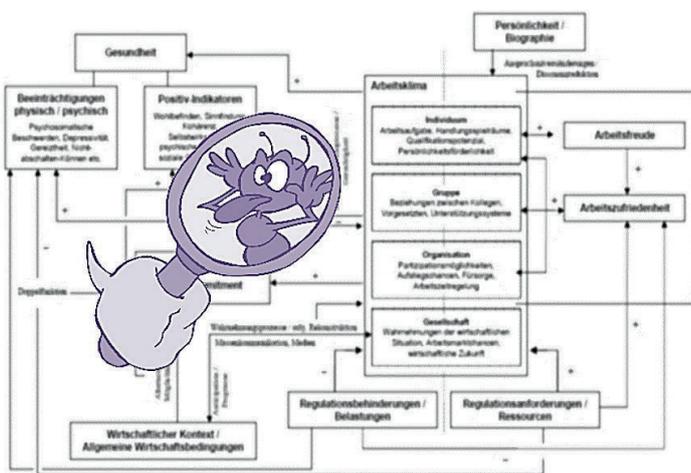


Abbildung 7: Hypothesenmodell Rauchverhalten von Jugendlichen (HBSC-Umfragen der WHO; alle 4 Jahre)



Bei komplexen Zusammenhängen können Hypothesenmodelle auch so aussehen.....

Abbildung 8: Schema eines Hypothesenmodells: Arbeitsklimaindex-Rahmenmodell (Quelle: Raml (2009), IFES-Institut)

...und meist sieht das Hypothesenmodell vor und nach der Studie ein bisschen anders aus als nach der Studie.....

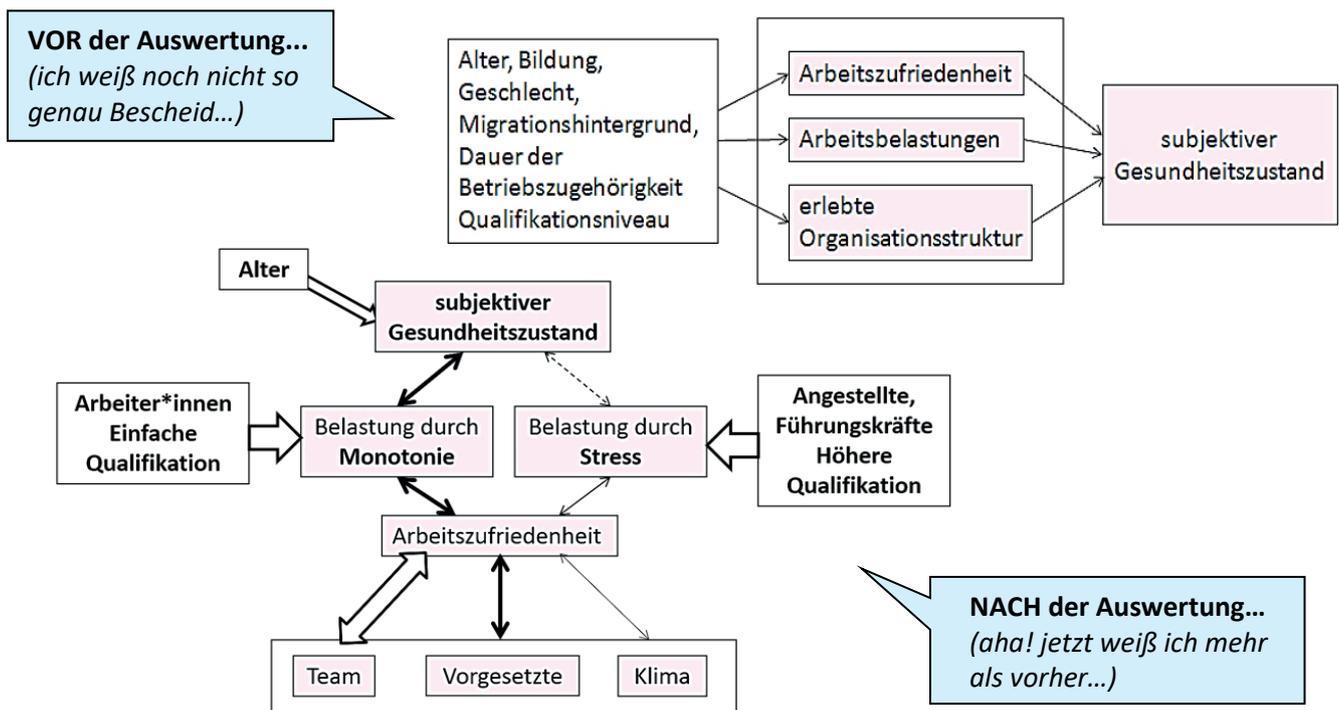


Abbildung 9: Das Hypothesenmodell vor und nach der Auswertung

³ Felder-Puig et al. (2019); siehe Health Behavior Of School aged Children: <https://www.ifgp.at/cdscontent/?contentid=10007.784109>

Aus einem Hypothesenmodell ergeben sich zahlreiche Hypothesen. Die **Fragestellung** besteht somit aus einer **Zusammenstellung von inhaltlich aufeinander abgestimmten Einzelhypothesen**.

Entscheidend bei der Operationalisierung ist: Die Forscher*innen legen die Kriterien fest, anhand derer gemessen wird. Dadurch entsteht ein vorkonstruierter Blick auf die soziale Wirklichkeit. Was wir Forscher*innen als relevant einstufen, wird erhoben und gemessen. Was nicht als relevant erachtet wird, wird auch nicht gemessen! Die Ergebnisse sind somit abhängig von den Indikatoren und Kategorien, die wir festlegen.

Aufgabe: Formulieren Sie drei zu Ihrer Fragestellung passende Hypothesen.

Welches Ergebnis vermuten Sie? Begründen Sie Ihre Vermutung und beschreiben Sie das fiktive Ergebnis.

Erstellen Sie dazu ein (kleines und einfaches) Hypothesenmodell zu Ihrer Fragestellung. Probieren Sie es!

1.5 Pflegediagnostik und Pflegedokumentation

Im Pflegealltag wird dokumentiert, was das Zeug hält. Aufnahme, Dienstübergabe, Entlassung – jede pflegerische Handlung soll nach Möglichkeit festgehalten werden. Dazu besteht sogar die gesetzliche Pflicht.⁴

Pflegebedarf (z.B. *Pflegestufen*), *Pflegeabhängigkeit*, *Pflegediagnosen*, Planung, Organisation, Durchführung, Kontrolle und die *Evaluation der pflegerischen Maßnahmen im Pflegeprozess* sollen dokumentiert werden.

Dazu entstanden viele Systeme, beispielsweise....

- **ICNP** – International Classification of Nursing
- **ENP** – European Nursing care Pathways
- **NANDA** - North American Nursing Diagnosis Association
- **NIC** (Pflegeinterventions-) und **NOC** (Pflegeergebnisklassifikation),
- **POP** – PraxisOrientierte Pflegediagnostik
- **ICF** – International Classification of Functioning, Disability and Health
- **Genogramm, biographischer Zeitbalken, Netzwerkkarte** und **Ecomap** (Pantucek 2005)
- **Sichtdiagnostik** (visuelle Ersteinschätzung)
- **ICD-10** (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme)
- **DSM-IV** (Diagnostic and Statistical Manual of Mental Disorders)
- ...viele spezialisierte **Assessmentinstrumente** für die Alten-, ambulante, palliative, psychiatrische, rehabilitative, aktivierende (...) Pflege mit vielen definierten, pflegewissenschaftlichen Messdimensionen....

Nicht nur pflegerische, auch soziale Aspekte – Biografie, Integration, individuelle Bedürfnisse, Ressourcen und Probleme⁵ – sind hier miteingeschlossen, die Gesundheit als Zustand des körperlichen, geistigen und sozialen Wohlergehens umfassen. Im Mittelpunkt steht die **Perspektive des*der Patient*in**, deren*dessen Situation vollständig abgebildet werden soll, nach Möglichkeit in einem gemeinsamen Aushandlungsprozess. Knapp die Hälfte der Pflegepersonen in Österreich verwenden 4 bis 6 Diagnoseinstrumente in der täglichen Arbeit, mehr als die Hälfte sind jedoch der Ansicht, dass dies für sie einen erheblichen zeitlichen Mehraufwand bedeutet.⁶ Das wundert nicht, sind doch Pflegekräfte aufgrund der psychischen und physischen Belastungen verstärkt dem Risiko von Burnout und anderen stressbedingte Erkrankungen ausgesetzt.⁷ Aber: Klassifikationssysteme sind der erste Schritt für die Entwicklung einer standardisierten Fachsprache, wie sie von Mediziner*innen oder Psycholog*innen schon lange angewandt wird. Die Vielfalt dieser Systeme birgt den Vorteil, dass wir für jeden Aspekt aus jedem Fachbereich eine Fülle von definierten Begriffen und Instrumenten zur Verfügung haben, die wir anpassen und zu Forschung und Evaluation verwenden können.

Also – wenn wir schon dokumentieren müssen, warum nutzen wir dies nicht zu unserem Vorteil? Wir können unsere Dokumentationen so gestalten, dass sie zu einer **regelmäßigen, gesamthaften Auswertung für eine (organisationsinterne) Evaluation der Qualität der Pflege (Monitoring)** geeignet sind.

Digitalisierung und IT-gestützte Dokumentationssysteme sind nicht aufzuhalten. Das bedeutet, dass wir zwar den Dokumentationsaufwand, dafür aber auch jederzeit abrufbare Längsschnittdaten zur Verfügung haben, die die Qualität und Wirksamkeit des innerprofessionellen Handelns belegen können.

⁴ In Österreich seit 1997 durch das GuKG, das Gesundheits- und Krankenpflegegesetz.

⁵ vgl. Behrens & Langer (2016): Das Dokumentationsmindestmaß in der stationären Altenpflege: zwischen Nutzlosigkeit und Schädlichkeit.

⁶ Haslinger (2017): Soziale Diagnostik versus Pflegediagnostik, Wien, Masterarbeit.

⁷ vgl. Janschitz et al. (2019): Neue Herausforderungen in der Pflege? in: SWS Rundschau S. 140-159.

Der erste, der begann, auf Basis von Dokumentationsdaten wissenschaftliche Fragestellungen zu bearbeiten, war Garfinkel (1967). Sein Ziel war es, Patient*innenkarrieren in einem psychiatrischen Krankenhaus nachzuzeichnen. Dabei stellte er fest, dass die Unterlagen schlecht ausgefüllt waren. Nach einem Jahr des Bemühens um vollständige Daten, wurde deutlich, dass das Klinikpersonal keinen Änderungsbedarf an den fehlenden Daten sah. Daraus schloss Garfinkel, dass es (gute?) organisatorische Gründe gibt, warum die Datengrundlagen nicht dazu geeignet sind, wissenschaftliche Evaluation zu betreiben.⁸

Ob eine Dokumentation „gut“ oder „schlecht“ ist, ist also erst die zweite Frage. Die erste Frage muss klären, ob das soziale System – welches die Dokumentation bedient und das von ihr bedient wird – wissenschaftliche Ergebnisse haben will. Sobald es sich dazu bekennt, sind wir nicht mehr aufzuhalten: dann wird im Pflegealltag geforscht, was das Zeug hält.⁹

Dieses Buch soll Sie unterstützen, in Ihrem Tätigkeitsfeld Forschungsideen zu entwickeln, im Zuge dessen sich ansehen, was andere auf diesem Gebiet bisher getan haben, und damit beginnen, Evaluationen durchzuführen, von denen Ihre Patient*innen und Ihre Profession profitieren. Im Kapitel 13 „Forschung in der Pflegepraxis“ ab Seite 145 werden einige Möglichkeiten aufgezeigt, wie das aussehen kann.

2 Grundgesamtheit und Stichprobe

Nachdem eine Fragestellung formuliert und die darin enthaltenen Begriffe und Messdimensionen definiert und operationalisiert sind, wird im nächsten Schritt die **Grundgesamtheit** (GG) bestimmt. Welche Personen (oder Elemente) betrifft meine Fragestellung?

Die **Grundgesamtheit** stellt die Menge aller Elemente (Personen, Texte, Interaktionen, Gerichtsakten, Regionen u.a.) dar, für welche die Fragestellung beantwortet werden bzw. für die die Ergebnisse gelten sollen. Die Grundgesamtheit muss **explizit definiert** und **eingegrenzt** werden.

Warum ist auch hier wieder eine genaue Definition erforderlich? Weil wir ganz genau wissen müssen, auf wen unsere Ergebnisse anwendbar sind und auf wen nicht.

Beispiel: Wenn die Mitarbeiter*innen eines Betriebes hinsichtlich ihrer Arbeitszufriedenheit und ihres Gesundheitszustandes befragt werden, dann ist die Grundgesamtheit einfach zu definieren, es sind **alle Mitarbeiter*innen des betreffenden Betriebes zum Befragungszeitpunkt**.

Beispiel: Wenn die Klient*innen einer Beratungseinrichtung befragt werden, dann sind **alle Klient*innen der betreffenden Einrichtung in einem bestimmten Zeitraum oder zu einem bestimmten Zeitpunkt** die Grundgesamtheit.

Wenn wir untersuchen wollen, *ob betagte Patient*innen auf einer Station mit Bezugspflegesystem zum Zeitpunkt ihrer Entlassung eine höhere Selbständigkeit aufweisen als Patient*innen auf einer Station mit Funktionspflegesystem*, dann müssen wir erst mal überlegen... Wichtig ist die Eingrenzung der Erhebung:

Räumlich: in ganz Österreich? nur in Wien? nur Niederösterreich? Welche Einrichtungen?

Sachlich: Alle Patient*innen? Welche Altersgruppen? Ausschluss von psychiatrischen Patient*innen?

Zeitlich: Einmalig? In welchem Zeitraum? Oder Panelstudie (immer dieselben Personen) zu mehreren Zeitpunkten?

Mögliche Grundgesamtheit: **Alle Patient*innen im AKH Wien, die über 80 Jahre sind und zwischen 1. 1. und 31.12. 2019 nach einem mehr als 3-wöchigen, aber nicht länger als 6-wöchigem Aufenthalt nach Hause entlassen wurden.**

Wenn wir wissen wollen, *ob die Befriedigung der Bedürfnisse von Patient*innen einer Notfallambulanz vom Personalschlüssel abhängen*, dann müssen wir uns dazu wieder mehrere Fragen stellen...

Räumlich: Welche Notfallambulanzen? In ganz Österreich? nur Wien? nur NÖ? Welche Krankenhäuser?

Sachlich: Alle Patient*innen? Nur jene, wo keine Reanimation durchgeführt wurde? Nur Erwachsene? Nur Patient*innen mit gravierenden Notfällen? Was sind „gravierende“ Notfälle? Wenn ein lebensbedrohlicher Zustand besteht? Woran genau machen wir das fest?

Zeitlich: Einmalige Befragung innerhalb einer bestimmten Zeitspanne? Welche Zeitspanne? Warum genau diese Zeitspanne? Rund um die Uhr? Nur tagsüber? Nur nachts? Nur wochentags?

Mögliche Grundgesamtheit: **Österreichweit alle Patient*innen ab dem 15. Lebensjahr, die am Tag X von 0 bis 24 Uhr in sämtlichen Notfallambulanzen (vollständige Liste) aufgenommen und behandelt wurden.**

⁸ Garfinkel (1967): „Good“ organizational reasons for „bad“ clinic records.

⁹ ...das muss ja noch lange nicht heißen, dass alle organisationsinternen Forschungen öffentlich gemacht werden!

Oft lässt sich eine sehr „ehrgeizig“ definierte Grundgesamtheit als Grundlage für eine Stichprobenziehung nicht realisieren, dazu müssten alle in Frage kommenden Einrichtungen motiviert sein, bei der Studie mitzuwirken. Meist sind aber einige dazu nicht bereit und Grundgesamtheit sowie Fragestellung müssen dementsprechend angepasst werden. Bei wissenschaftlichen Studien wird das Design an die herrschenden Gegebenheiten angepasst (Budget, Zeitrahmen, Zugang, Reichweite, Teilnahmebereitschaft, Realisierbarkeit...). Wenn die ausgewählten und teilnahmebereiten Einrichtungen „Sonderfälle“ darstellen (z.B. Spezialisierung auf Intoxikation, auf Unfälle o.ä.), bereitet die Verallgemeinerung – das vordergründige Ziel quantitativer Forschung – oft große Schwierigkeiten bzw. ist nur mit Einschränkungen möglich.

Die Grundgesamtheit ergibt sich daher oft aus rein praktischen Gesichtspunkten, nämlich dem Zugang bzw. den faktischen Gegebenheiten. Die Definition der Grundgesamtheit erfordert also eine konkret formulierte (eingegrenzte) Fragestellung, die Abklärung des Untersuchungszeitpunktes sowie die Definitionen der verwendeten Begriffe. Meist erfolgen diese Schritte (Formulierung der Fragestellung, Operationalisierung, Eingrenzung und Festlegung der Grundgesamtheit) in einem reflexiven Abstimmungsprozess.

Sobald die Grundgesamtheit präzise definiert ist, kann überlegt werden, wer von der Grundgesamtheit untersucht werden soll. Alle? Eine Teilmenge?

Bei einer **Vollerhebung = Totalerhebung** werden alle Elemente der Grundgesamtheit untersucht.
 Bei einer **Stichprobe** wird eine Teilmenge aus der Grundgesamtheit untersucht.

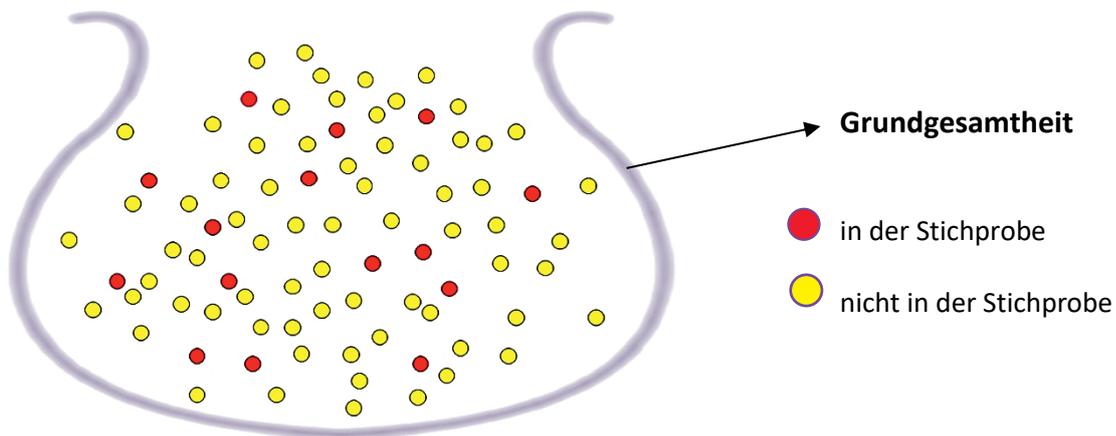


Abbildung 10: Grundgesamtheit und Stichprobe

Die folgende Grafik veranschaulicht die Überlegungen bei der **Stichprobenziehung**.

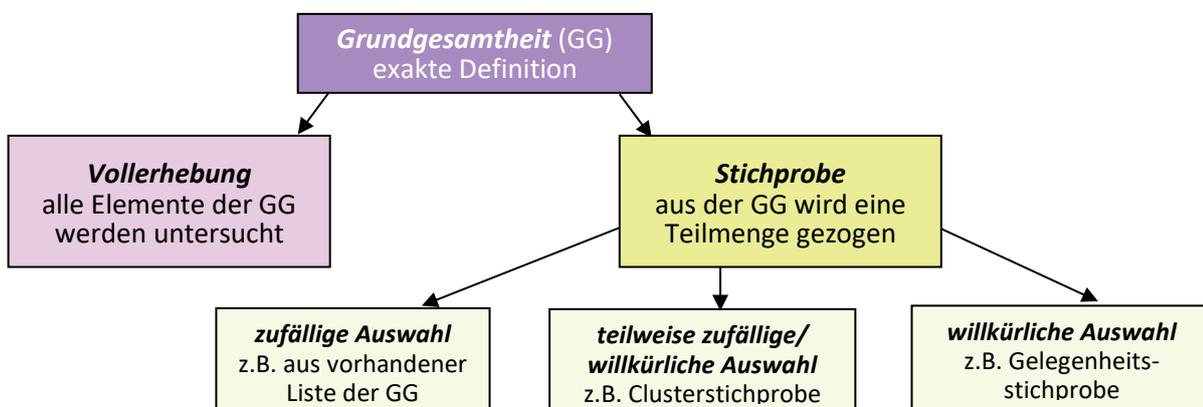


Abbildung 11: Grundgesamtheit und Stichprobe – Arten von Stichproben

Grundsätzlich gelten folgende Prioritäten: Eine **Totalerhebung** ist der Optimalfall, sie ist aber meist zu teuer oder nicht möglich. Wenn alle Elemente der Grundgesamtheit bekannt und in einer Liste vorhanden sind, dann ist eine **Zufallsstichprobe** (jede*r Zweite, jede*r Dritte, jede*r Zehnte auf der Liste – je nach Fallzahl und Budget) anzustreben. Wenn sowohl eine Totalerhebung als auch eine Zufallsauswahl nicht möglich sind, dann wird eine **willkürliche (=bewusste)** oder **teilweise willkürliche Auswahl** der untersuchten Einheiten/Personen getroffen.

Wir unterscheiden zwei Arten von Stichprobendesigns:

- Beim **repräsentativen Stichprobendesign** liegt der Schwerpunkt auf der **Beschreibung**. Hier ist es besonders wichtig, dass die Stichprobe die Grundgesamtheit in allen relevanten Merkmalen widerspiegelt, daher haben wir es meist mit einer sehr „bunten“ (heterogenen) Personengruppe zu tun und wollen (grob gesagt) wissen, wie es um diese Zielgruppe bestellt ist.

Ein typisches Beispiel für eine **repräsentative Studie** ist die Befragung des Pflegepersonals zu einem bestimmten Zeitpunkt in einem bestimmten Bereich/Tätigkeitsfeld/Bezirk zu Themen wie Weiterbildung oder Arbeitsbelastung.

- Das **experimentelle Design** hingegen zielt auf eine **Erklärung** ab. Im Fokus stehen hier die Bedingungen, unter denen der Versuch/die Testreihe stattfindet. Die untersuchten Gruppen sind meist homogen, um auf diese Weise „störende“ Einflüsse möglichst auszuschalten. Denn: Der gemessene, erzielte Effekt soll ausschließlich auf der Bedingung (z.B. Behandlungsmethode) beruhen und nicht durch andere („störende“) Umstände (z.B. unterschiedliches Alter der Proband*innen) (mit-)bedingt sein.

Ein Beispiel für ein **experimentelles Design** ist, wenn im Rahmen einer klinischen Interventionsstudie untersucht wird, ob Mukositis durch Lasertherapie erfolgreicher behandelt werden kann als ohne Lasertherapie.

Die folgende Abbildung stellt diese beiden Designs anhand ihrer wesentlichsten Kennzeichen gegenüber.

Repräsentativ – deskriptiv verwendet bei Umfragen		Experimentell – kausal verwendet bei Interventionsstudien	
Stichprobe repräsentiert Grundgesamtheit (in allen „relevanten“ Merkmalen) Üblicherweise: Geschlecht, Alter, Region, Bildung, Einkommen, Erwerbstätigkeit – möglichst viele! Begründen, warum ein Merkmal nicht relevant ist! Die Ergebnisse sind auf die Grundgesamtheit generalisierbar – Bedingung: Zufallsauswahl (Liste, Urne – alle Elemente mit gleicher Chance)		Stichprobenauswahl durch Variation des zu untersuchenden Kriteriums oder Zeitpunkts Die Varianz des Kriteriums unterstellt Kausalität. „Bedingungen“ (z.B. Behandlung, Film, Unterrichtsmethode...) Gruppen unter verschiedenen Bedingungen: Bedingung A – nicht A (Versuchs- und Kontrollgruppe) oder Bedingung A – B – C...	
Beschreiben: Status quo in der Grundgesamtheit		Erklärung: Was bringt die „Bedingung“?	
Heterogene Stichprobe (möglichst viele Eigenschaften wie in der Grundgesamtheit) Rückschluss auf größere Personengruppe möglich	Homogene Gruppen (möglichst viele Eigenschaften gleich) Einfluss der Bedingung soll „sauber“ bleiben, soll kontrolliert werden „experimentell“ = Randomisierung: Gruppenzuteilung per Zufall „quasi-experimentell“ = Parallelisierung: homogene Gruppe natürliche, vorgegebene Gruppeneinteilung		
Repräsentativität als Kontinuum Je größer und heterogener die Stichprobe, je „zufälliger“ die Auswahl, desto repräsentativer die Ergebnisse...	eindeutige, kontrollierte Bedingungen mit wenig störenden Einflüssen (intervenierende bzw. „störende“ Variable) Interne Validität – hohe Kontrolle (<i>Laborexperiment</i>) Externe Validität – hohe Generalisierbarkeit (<i>Feldexperiment</i>)		

Abbildung 12: Repräsentatives und experimentelles Stichprobendesign im Vergleich

2.1 Repräsentativität – oder: Der Zufall ist der „Master of the Universe“

Repräsentativ ist eine Stichprobe dann, wenn sie die Grundgesamtheit repräsentiert. Was heißt das?

Repräsentativ heißt, die Stichprobe ist ein **Miniaturabbild der Grundgesamtheit**, sie entspricht in allen relevanten Merkmalen der Grundgesamtheit. (Was „relevant“ ist, definiert der*die Forscher*in!) Rückschlüsse auf die Grundgesamtheit sind nur dann zulässig, wenn die Stichprobe ein Abbild der Grundgesamtheit ist.