

# Generalisierung von Forschungsergebnissen – ein häufiges Problem empirischer Wissenschaften und seine Lösung im Aufbau interner Evidence

## *Natürlichkeit“ und „Generalisierbarkeit“ sozialwissenschaftlicher Feldexperimente: Verallgemeinerungen zu externer und interner Evidence*

Soziale Welt (SozW) 2015, S. 246 – 267

Johann.Behrens@Medizin.Uni-Halle.de

### Zur Prüfung der Befangenheit: Offenlegung derzeitiger Verpflichtungen

Haftender ehrenamtlicher Vorstand folgender gemeinnütziger eingetragener Vereine und Forschungsbereiche:

Pro Familia; Institut für Qualität in der Pflege; Institut für Supervision, Institutionsanalyse und Sozialforschung; Theodor Wille–Heinrich Diederichsen Preis „Fair Trade“; Pflegeforschungsverbund; Sektion Medizin– und Gesundheitssoziologie der Deutschen Gesellschaft für Soziologie; auslaufend: DFG–SFB 580



# Kölner Weisheit: Jede Jeck ist anders.

## Jedem nach seinen Bedürfnissen, jeder nach seinen Fähigkeiten.

*Warum dann nicht alle Schwierigkeiten der Generalisierung (über unterschiedliche Kontexte und Ziele hinweg) vermeiden, indem wir die Untersuchungen nur an unserem jeweils einzigartigen Klienten durchführen?* Windeler et al, Deutsches Ärzteblatt 195, S. 565–570 (2008)

Alles an sich auszuprobieren überstünde niemand.

*,Um-Zu-Kausalität‘ und ,Weil-Kausalität‘ (Husserl)*

Alltagsleben unterstellt Kausalität, wissenschaftliche Erfahrung verlangt Beweise.

*Es ist schon mal gutgegangen: Der Nächste bitte!*



# Generalisierung: „Theorie“ leitet „Anwendung“ an oder Anwendung fragt Theorie?

## Gliederung

- 1– Unterscheidung Generalisierung 1 und 2
- 2– Sozialwissenschaftliche Tradition der Generalisierung Natürlicher Feldexperimente: Cronbach  
UTOS und Pseudo-UTOS . Aporie: Externe Validität nicht ex ante zu haben
- 3– Drehe Fragerichtung um, Aufbau interner Evidence entlastet Anforderungen an externe  
Evidence
- 4– Warum angemessene outcomes, Umwelten nur individuell zu haben sind
  - A) „Participation“ WHO
  - B) Lebewesen reagieren nur auf Stimuli, denen sie eine Bedeutung erteilen (C.Peirce, Dewey, J.  
und Thure v. Uexküll
- 5) Anforderungen an die Darstellung externer Evidence
  - Ökonomie
- 6) Schlussfolgerungen



# Vorschlag: Zwei Stufen von Generalisierungen:

- 1) Generalisierung einer Erfahrung auf eine Population
- 2) Generalisierung auf den Einzelfall eines Klienten

Beide sind nur bei Poppers ‚All-Aussagen identisch‘, nicht bei Häufigkeitsaussagen (nicht bei ‚number needed to treat > 1‘)

*Für kontemplative Wissenschaften reicht Generalisierung 1, Handlungswissenschaften brauchen Generalisierung 2.*



# Medizin, Pflege- und therapeutische Wissenschaften, aber auch die *beratende* Ökonomie, Regional- und Bildungswissenschaft als „*Handlungswissenschaften*“

Medizin, die therapeutischen Wissenschaften und die Pflegewissenschaft sind Wissenschaften mit eigenen, von den kontemplativen Wissenschaften der Biologie und Soziologie u.a. trennbaren Gegenständen:

**Ihr Gegenstand ist die zukunftsunsichere, aber vernünftige innovative Krisenentscheidung im jeweiligen einzigartigen Fall**

–

**unter Handlungsdruck und Begründungszwang gemeinsam mit den je einzigartigen Klienten.**

Daher erscheint „Merksatz“ (R.Gross M. Löffler, Prinzipien der Medizin, Berlin 1997, S. 8) problematisch:

*„In der Medizin sind Wissenschaft, Kunst [falsch übersetzt ‚technä‘] und Handwerk untrennbar verbunden. Wenn auch die Forschungsergebnisse mehr wissenschaftlicher Natur sind, der Umgang mit den Kranken mehr eine Kunst, so handelt es sich dabei um Akzente.“*

Denn das Praktizieren als „Künste“ statt als „Handlungswissenschaften“ zu bezeichnen, macht die Praxis als „vernünftige“ undiskutierbar, unkritisierbar, unerforschbar.



# Berühmtester gescheiterter ? Lösungsvorschlag: Je natürlicher die Kontext-Bedingungen sind, umso leichter sind Ergebnisse von Feldexperimenten zu generalisieren, umso eher entspricht die interne der externen Validität??

Um die externe Validität zu erhöhen, sollen in natürlichen Designs „wahrgenommene Kontextbedingungen trotz Eingreifen der Forschenden nicht von realweltlichen Gegebenheiten abweichen“ Keuschnigg /Wolbring 2015

Schulenburgender **L.J. Cronbach** (1982 ff) mit UTOS und Pseudo\*UTOS:  
Wann sind (natürliche) Feldexperimente generalisierbar? Wenn in Feldstudie und Zielkontext übereinstimmen:

*U*nits = Probanden

*T*reatments = Interventionen

*O*bserving Operations = Beobachtungen, Messungen

*S*etting = Kontext

Weicht auch nur eine der vier Dimensionen vom Zielkontext ab, auf den man generalisieren will, so ist „*„diese Extrapolation nicht mehr vom Design selbst gedeckt“*“



# Probleme: a)UTOS ohne neue Studie gar nicht erkennbar b) Natürlichkeit (Spezifität) verringert Generalisierbarkeit

Ohne neue Studie kann ich gar nicht erkennen, ob alle U,T,O,S in der Situation, auf die ich verallgemeinern möchte, ähnlich sind, vor allem wegen des Problems unbeobachteter Heterogenität.

Je spezifischer eine Studie die U,T,O,S erfasste, umso weniger ist zu erwarten, dass ich die Ergebnisse auf eine andere Situation verallgemeinern kann.

Alle menschliche Wissenschaft ist ‚Indizienwissenschaft‘ i.S. des Historikers und Kriminologen Carlo Ginzburg.



# Vorschlag: Wir drehen die Fragerichtung um und beginnen bei der Kölner Weisheit statt bei der Suche nach grenzenloser externer Validität

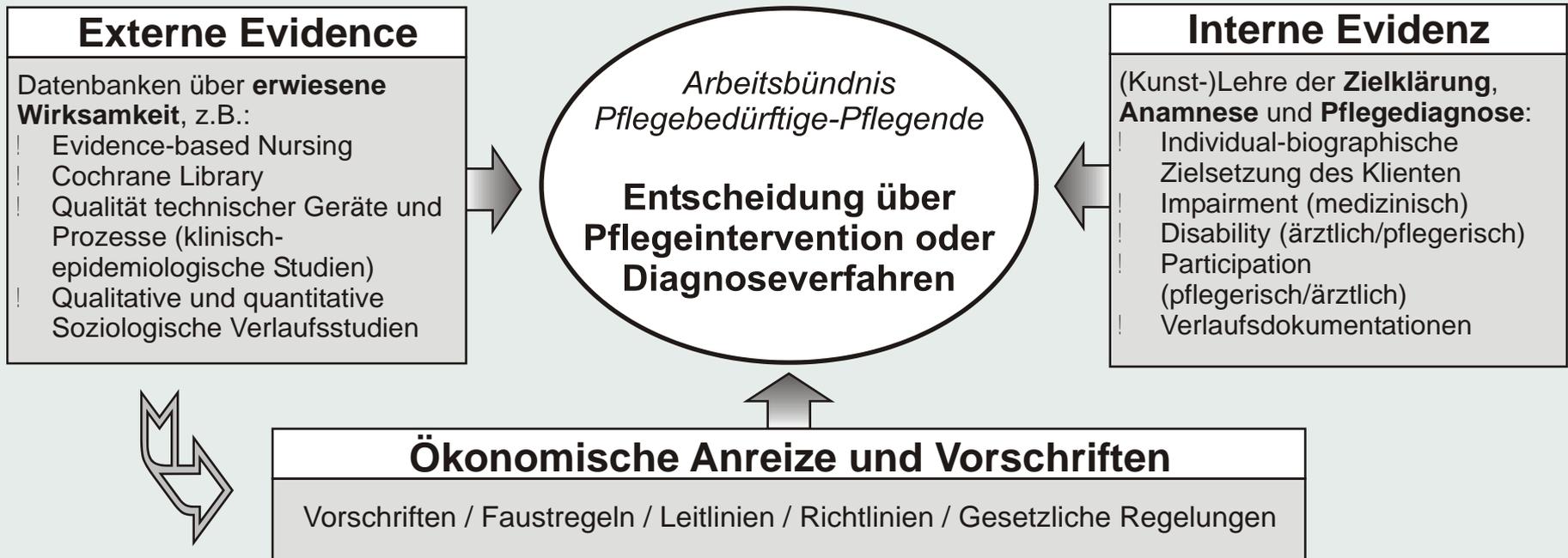
*Statt zentraler Entscheidungen über dezentrale Anwendungen –  
dezentrale Entscheidungen mit zentralen Informationen über externe Evidence  
und Begründungspflicht*

Entspricht *evidencebasierter Handlungspraxis*, zumindest im *Evidence-based Nursing and Caring*.

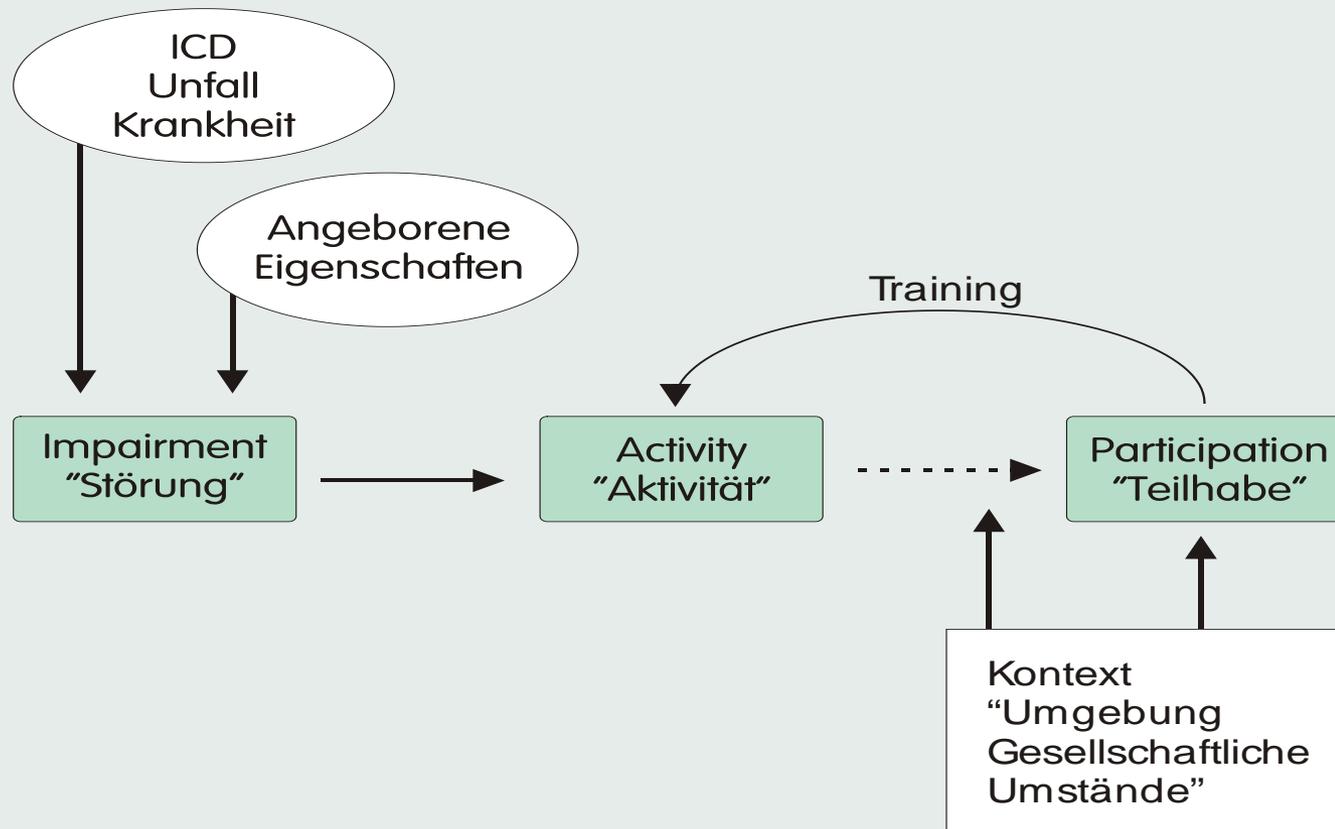
Pflege konnte noch nie von einer organischen Diagnose auf die angemessene Pflege einer Person schließen.

Evidence-based Nursing and Caring pflegte noch nie Organe, sondern immer Personen, noch nie ‚Körper ohne Seelen‘ oder ‚Seelen ohne Körper‘.





# Selbstbestimmung und Teilhabe als *individuelles* Ziel von Therapie und Pflege (§ 1 SGB IX): Jede Indikationsstellung setzt interne Evidence der Teilhabeziele voraus



# Selbstbestimmung und Teilhabe als *individuelles* Ziel von Therapie und Pflege (§ 1 SGB IX): Jede Indikationsstellung setzt interne Evidence der Teilhabeziele voraus

	Krank	Gesund
Körperfunktionen und Strukturen	-----x-----	-----
Aktivitäten	-----	-----x-----
Teilhabe („Partizipation“)	-----	-----x-----

Die drei Ebenen von „Gesundheit“ variieren relativ unabhängig voneinander.

Eine Person kann große selbstbestimmte Teilhabe erlangen, auch wenn Körperfunktionen und Körperstrukturen nicht mehr zu bessern sind und Aktivitäten auch kaum noch.

Daher definiert § 1 SGB IX als Behandlungsziele nicht „ordentlicher Zustand der Organe“ oder „selbstständig ausgeübte Aktivitäten“, sondern „Selbstbestimmung“ und „Teilhabe“, also selbst- und nicht fremdbestimmte Teilhabe am Leben der Gesellschaft. Teilhabe ist immer etwas Individuelles, über das nur die Person selber zu entscheiden hat.

*Beispielhafter Beleg:* Krebskranker Herr S. fragt Klinik: Kann ich unter den Nebenwirkungen der onkologischen Behandlung sicher noch meinen Kommissionsbericht zu Ende schreiben? Da das nicht sicher ist, entscheidet er sich gegen die Therapie und für die Teilhabe, den Bericht.

Das darf niemandem vorgeschrieben werden.

Daher ist in der Onkologie ein „(tumor-)progressionsfreies Leben“ noch kein patientenrelevanter Surrogat-Endpunkt. Aber auch „Lebensqualität“ = „Teilhabe“ ist immer etwas Individuelles, nicht für jeden sind Aktenaufräumen und ein Kommissionsbericht Lebensqualität.



# Uhren vs. Lebewesen, offene vs. geschlossene Systeme

## *Erst der innere „Interpretant“ (C. Peirce) eines Lebewesens prägt dem Stimulus eine Bedeutung auf, auf die ein Lebewesen reagiert*

Offenes System ‚Uhr‘ zweigliedrige Wirkungskette :

*gleicher Stimulus (,Ursache‘) – gleiche Reaktion (,Wirkung‘)* (Leibniz Gott als der große Uhrmacher)

Geschlossenes System ‚Tiere/Menschen‘ dreigliedrige Wirkungskette:

*gleicher Stimulus–innerer Interpretant– ungleiche Reaktion*

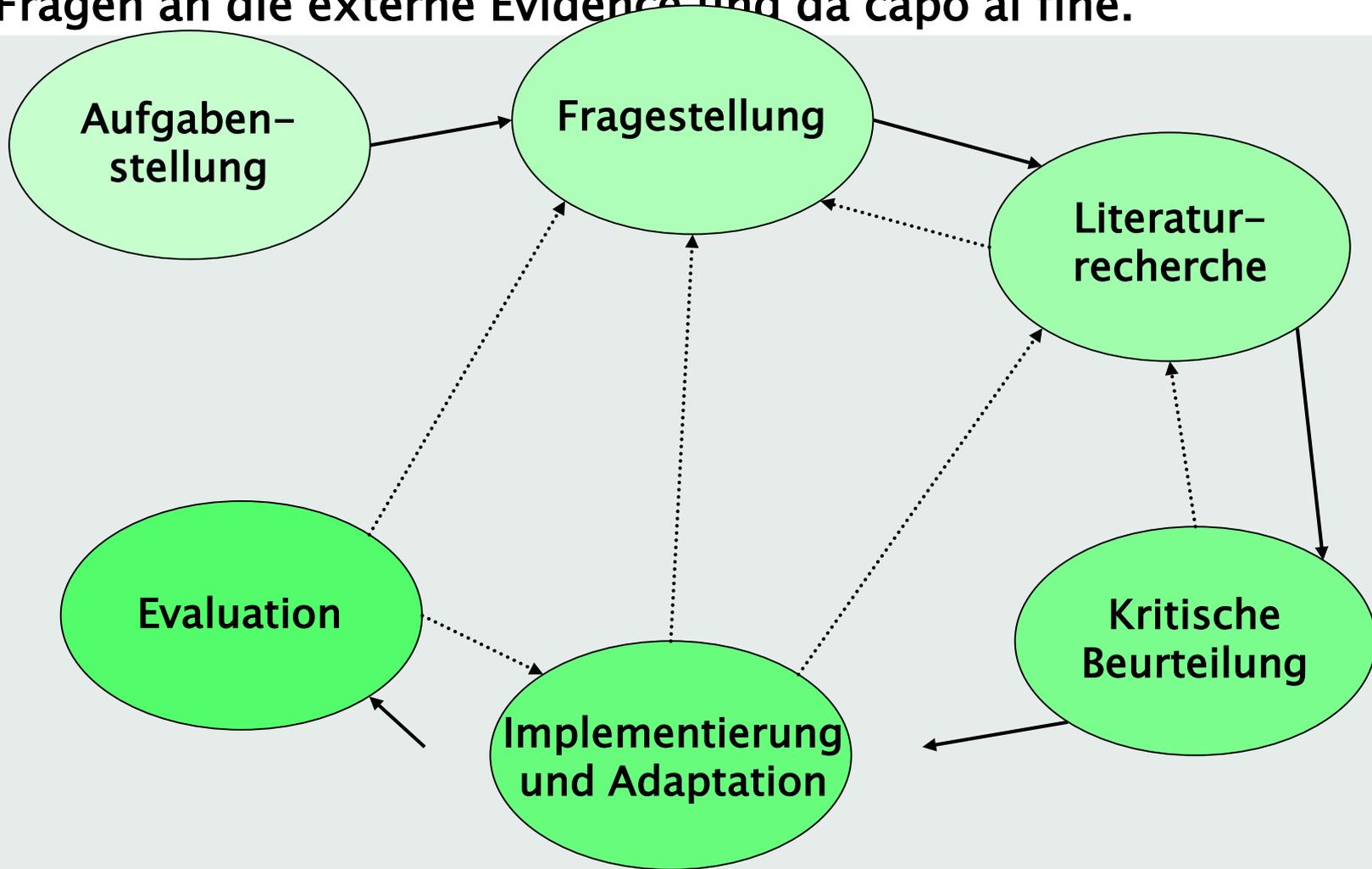
(später übernommen für die Biologie von Jakob v. Uexküll, von Norbert Wiener für die Kybernetik, von Umberto Eco 1976 für die Semiotik ausgebaut, für die Medizin von Thule v. Uexküll, für die Hirnforschung von Wulf Singer ,von Richter (1957) im Rattenexperiment gefunden).

Beispiel: ‚Verum–Effekt‘ und ‚Placebo–Effekt‘ wirken beide körperlich durch die Bedeutungserteilung des Interpretanten.

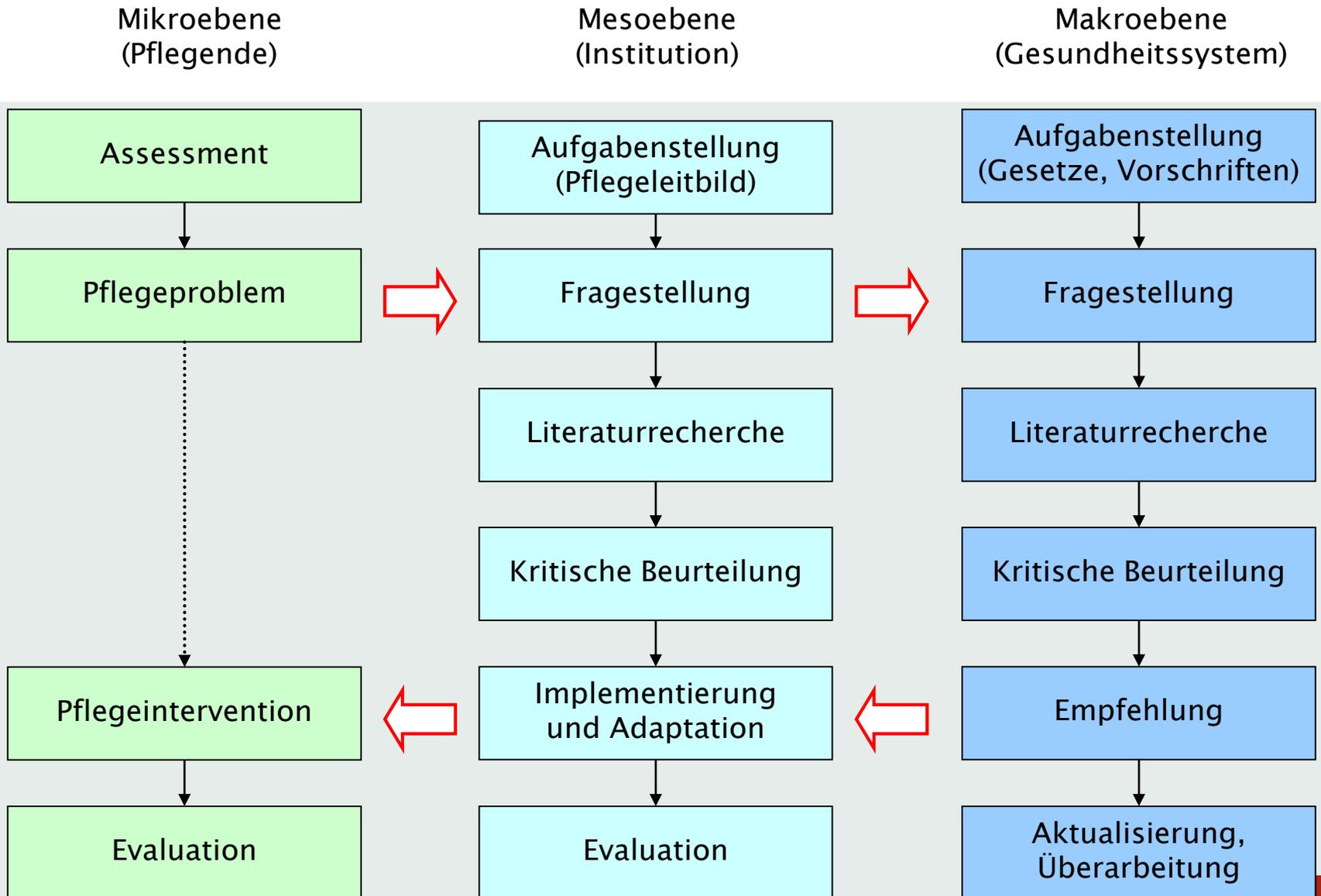
„Wenn die physikochemische Wirkung und die Wirkung durch Bedeutungserteilung in die gleiche Richtung zielen“ (Hontschik 2013, S. 216), dann hat eine Intervention (z.B. Medikament) eine überlegene Wirkung.



# Die Generalisierung 2 und 1 im Evidence-based Nursing: Von der internen Evidence des Einzelnen zu Fragen an die externe Evidence und da capo al fine.



# Anwendung der Methode EBN



# Was unterscheidet den Aufbau interner Evidence in der Begegnung (=„Indikationsstellung“) von Informationen zur Förderung von Health Literacy?

- 1) Die Professionsangehörigen teilen nicht (nur) etwas mit, was sie schon wissen. Im Gegenteil kennen sie das Wesentliche, nämlich die Teilhabeziele, die Ressourcen, die ‚Befunde‘, die Kontexte und die Wahrnehmungen ihrer Klienten nicht. Sie können sie nur von den Klienten selber erfahren.
- 2) Den Klienten wird nicht unterstellt, sie hätten bereits Klarheit über die eigenen Teilhabeziele, die Ressourcen, die Kontexte, ihre Wahrnehmungen und ‚*Präferenzen*‘. Im Gegenteil bedürfen sie oft des Gesprächs, um sich diese Klarheit über sich selbst zu verschaffen – in Kenntnis der Erfahrungen, die andere in vergleichbarer Situation machten (externe Evidence, „Optionen“).
- 3) ‚Interne Evidence‘ ist ein gemeinsames Produkt der Begegnung, keine Information, die der eine schon hat und der andere noch nicht.
- 4) Den Klienten wird nicht unterstellt, sie hätten bereits Präferenzen und wollten vor allem Informationen in einfacher oder sogar leichter Sprache. Im Gegenteil wird akzeptiert, dass sie viele Details gar nicht so genau wissen wollen und es ihnen weniger um Informationen als die Teilung der Angst geht und die Wiedergewinnung selbstbestimmter Teilhabe geht.
- 5) Es wird nicht unterstellt, Empathie reiche zum Fremdverstehen beim Aufbau interner Evidence.
- 6) Es wird nicht unterstellt, was der Verbraucherminister am 2.1.2017 verspricht: „Eine informierte Einwilligung schützt Patientinnen und Patienten vor falscher Behandlung ...“ . Im Gegenteil ist leider eine existentielle Entscheidung unter Ungewissheit nicht vermeidbar.



**Auch noch so wünschenswerte Ausdifferenzierung externer Evidence befreit nicht von der Notwendigkeit, interne Evidence aufzubauen.  
Das Modell von Gordon Guyatt und Holger Schünemann:**

**Evidence of prognosis and context for a given patient**

Estimates required for balancing benefit and downsides from interventions  
"can this evidence work in my patient?"

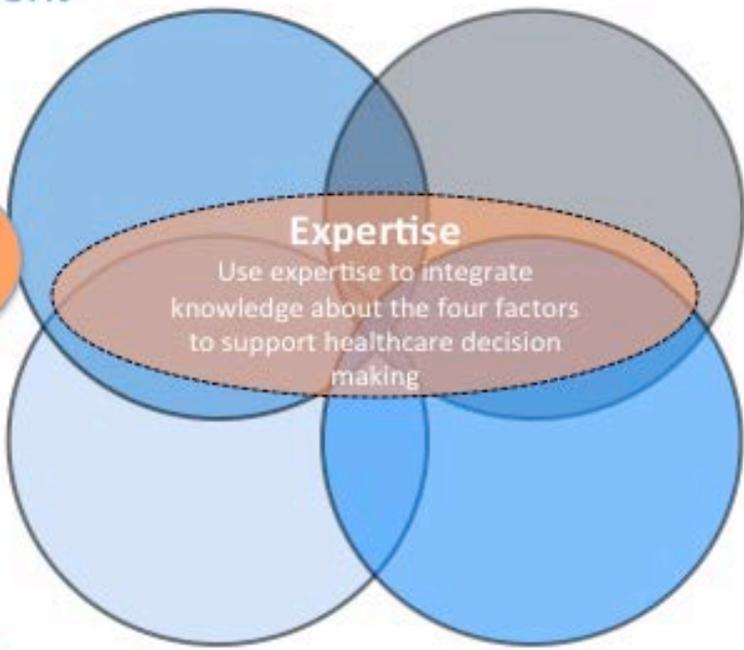
Use best evidence about appropriate implementation and integration strategies  
"what works in this setting"

**Evidence of implementation & integration**

**Evidence of patient's values & preferences**

For daily practice, the best evidence may come from the patient in front of you  
"elicit the patient's values and preferences"

Determine all important outcomes and effects on those – provides relative estimates of effect  
"how large are the effects in research studies"



**Evidence of intervention\* effects**

\*includes diagnostic, therapy and cost



# Aufbau interner Evidence in der Begegnung zweier „geschlossener Systeme“

Stimulus – Bedeutungserteilung durch Interpretanten des Klienten – Reaktion

Stimulus – Bedeutungserteilung durch Interpretanten des Professionmitglieds – Reaktion

Gemeinsames Produkt Interne Evidence.



# Für die Generalisierung 1 ist die Generalisierung 2 eine Entlastung: Grenzenlose externe Validität ist weder zu haben noch nötig.

Dass ich interne Evidence aufbauen muss, um Fragen an die externe Evidence formulieren zu können, entlastet die Generalisierung 1 von der Forderung, für jeden auftretenden Fall extern valide oder tatsächlich UTOS zu sein. Nie kann und nie muss von externer Evidence des Durchschnitts auf den Einzelfall geschlossen werden. Daher sind auch Pseudo\*UTOS geeignet, Anregungen für die Reflektion des Einzelfalls zu geben.



# Welche Darstellung externer Evidence braucht man zum Aufbau interner Evidence: *zusammenfassende „Noten“ oder ungewichtete Transparenz?*

Beispiele *Stiftung Warentest und GRADE*

- 1) Aggregationen nach Graden der Vertrauenswürdigkeit mit Punktabzügen haben die Probleme von Schulzeugnissen
- 2) RCTs bewältigen den Auswahl-Bias.  
Keine RCT kompensiert
  - die Wahl eines unpassenden individuellen outcome-Indikators
  - Fehler bei der Erfassung einer Intervention
  - Verzerrungen durch eine nicht bevölkerungsrepräsentative Grundgesamtheit.

Für jede dieser Verzerrungsrisiken halten mixed methods andere Methoden bereit.



# Ausweitung der Forderung nach Externer und Interner Evidence auf *beratende* Ökonomie, Bildungs- und Regionalforschung

„Jede Fördermaßnahme ab einer festzusetzenden Größenordnung sollte soweit möglich standardmäßig einer empirisch-kausalen Wirkungsanalyse unterzogen werden. Im Verlauf dieser Maßnahme sollte diese durch qualitative Interviews und Fallstudien ergänzt werden...Durch eine solche Kombination aus quantitativen und qualitativen Methoden kann häufig ein vertieftes Verständnis von Wirkungsmechanismen entstehen“ (*Wissenschaftlicher Beirat beim BMWi 2013, S. 18*).

"An important part of nudging is to collect evidence on which policies actually work as intended, before they are implemented on a larger scale. Ideally, the policies should be tested and evaluated in randomized field experiments". Laudatio auf den Nobelpreisträger für Wirtschaftswissenschaften 2017 - Richard Thaler.

*Einführung DRG über Nacht landesweit*

Kennedy in der Kuba-Krise: vermeide die Fehler von 1914

EZB nach der Finanzkrise: vermeide die Fehler der 1920er Jahre



# Neun Schlussfolgerung für Generalisierungen 1 und 2 in allen Handlungswissenschaften (morgen nachmittag zu ergänzen )

- 1) *Jede Jeck ist anders*–in Teilhabezielen, Interpretanten, Wahrnehmungen, Ressourcen, Umwelten und Umgebungen
- 2) *Interventionen* sollen explizit *jedem nach seinen Bedürfnissen* gerecht werden und *jeder* hat sie *nach seinen Fähigkeiten* zu finanzieren.
- 3) *RCTs wuden gerade nicht für den Einzelfall, sondern für den Durchschnittsfall erfunden* (RCTs Durchschnittliche Wirkung des gleichen Samens auf unterschiedlichen Böden)
- 4) *Aufbau interner Evidence ist unverzichtbar zur Nutzung externer Evidence*
- 5) Statt *zentrale Entscheidungen über dezentrale Anwendungen*  
*dezentrale Entscheidungen mit zentralen Informationen über externe Evidence*
- 6) Eine grenzenlose externe Validität von Häufigkeits–Ergebnissen ist ausgeschlossen. Aber: Durch Abstriche bei der internen Validität gewinnt man noch keine externe Validität.
- 7) Alles Handeln ist Entscheiden unter Unsicherheit und Ungewissheit. Das ist kein Grund, nicht zu handeln. „*Nihil nocere*“, „*Risikovermeidung*“, schadet oft.
- 8) *Warranted assetabilities (John Dewey, C. Peirce)*: Eine Entwicklung kann dann *vorläufig* als evidencebasiert kausal bewirkt gelten, wenn alle anderen Erklärungen und Verzerrungsmöglichkeiten mit mixed methods ausgeschlossen wurden (vgl. Carlo Ginzburgs ‚Indizienwissenschaft‘)
- 9) „Wer sich nicht an Standards hält, wird sich verantworten müssen“ ist falsch: Auch wer sich an Standards hält, wird sich dafür verantworten müssen, dass ein Standard für seinen Klienten zutrif.

Vielen Dank. *Möge Ihr Samen aufgehen und Ihr Weizen blühen.*

