Prof. em. Dr. phil. habil. Johann Behrens

Der EbM+Theory Ansatz

Eine Würdigung seiner 6 Annahmen und seiner Risiken und Nebenwirkungen

Die Community der Versorgungsforschung bemüht sich um Evidence, um die Unsicherheit unserer alltäglichen Entscheidungen in Ökonomie, Politik, Behandlung bei Krankheit und in gesundheitsförderlicher Pflege zu verringern (Soziale Welt Behrens 2015). Holger Pfaff und Jochen Schmitt legten hierzu 2023 und 2021 einen „EbM+Theory Approach" vor. Wie ihr Schaubild (2023, S.3, Figure 1) zeigt, gehen sie davon aus, es gäbe zwei trennbare Gütehierarchien wissenschaftlicher Studien: eine für „theory-based knowledge" und eine für „evidence-based knowledge". „Scientific knowledge" sei das Resultat der Kombination beider unabhängiger Wissenshierarchien.

>> Den Beiträgen von Pfaff und Schmitt* kommt das Verdienst zu, Errungenschaften, die die Selbstreflexion der evidence-based Nursing, der evidence-based Medicine und andere Bewegungen von Praktikerinnen vor einem Vierteljahrhundert, also am Endes des letzten Jahrtausends erreicht und z. B. als „Methoden und Ethik der Pflegepraxis und der Versorgungsforschung" publiziert hatten (zusammengefasst in Behrens & Langer 2004 und 2022, Behrens 2019), ausschnittsweise zu übernehmen und für die Versorgungsforschung zu diskutieren. Einige wenige der zahlreichen Erkenntnisprobleme evidencegestützten** Handelns, die Ende des letzten Jahrtausends theoretisch zu bewältigen waren, sprechen sie wieder an. Sie übernehmen zumindest programmatisch die Abkehr von einem mechanistischem Zugang zur Handlungs- und Interventionsforschung und die Zuwendung zu einem systemtheoretischen Zugang, der die Welt nicht mehr nach dem Modell einer Uhr analysiert, deren Räderwerk dem Uhrmacher völlig offenliegt – wie dem allwissenden „Uhrmacher"-Gott bei Leibniz und Bourdieu (Behrens & Zimmermann 2017). Das ist zweifellos ein großes Verdienst.

Der vorgeschlagene „EbM+Theory Approach" basiert auf weitgehenden Annahmen, darunter der Annahme, es könne überhaupt ein „evidence-based knowledge" geben, das nicht „theory-based" sei. Der Vorschlag, die empirische Evidence** durch eine theoretische Evidence und weiter durch „Experience und Context-based knowledge", wie sie eminente Experten haben, zu einem EbM+Theory Approach zu ergänzen, beruht allerdings auf sechs Annahmen, die womöglich leider für Menschen nicht realisierbar sein können und daher mit einer anthropologisch begründbaren, also bescheiden gewordenen Erkenntnistheorie nicht konsistent sind. Das ist in diesem Aufsatz zu prüfen. (Die Hoffnung auf Eminenzen, die nicht an anthropologische Grenzen der Erkenntnis gebunden sind, wird selbstverständlich nie aussterben. Die Hoffnung ist zu schön.) Die Annahmen, die Pfaff und Schmitt für ihren Vorschlag eines EbM+Theory-Approaches brauchen, sind: 1) Es gäbe einen riesigen Bereich in der Krankenbehandlung und Gesundheitsförderung, in dem man aus (RCT)-Studien (Be)Handlungsentscheidungen direkt generalisierend ableiten könne – und daher die ganze Kritik, die die EbP-Bewegung seit dem vorigen Jahrtausend an dieser viel zu naiven Annahme über RCTs erarbeitete, ruhig ignorieren dürfe. Mit dieser ersten Annahme verfehlen sie die häufigste Standardsituation Evidencebasierten professionellen Handelns. Mit der zweite Annahme versuchen die Autoren die These von der Verbreitung „noncomplexer" Interventionen zu retten – wobei sie Luhmann als Zeugen in Anspruch nehmen: Es gäbe ein organisches, ein psychisches und ein soziales System, in das das „Healthcare"-System jeweils getrennt intervenieren könne, was man z. B. an der Wirkung „Drugs" und „surgical interventions" erkenne. Unabhängig davon, ob die Luhmann'sche Systemtheorie diese Annahme beinhaltet, verdient sie Erörterung, weil sie folgenreich verbreitet ist.

Die dritte Annahme scheint zu sein, bei Kollektiven sei das Generalisierungsproblem gering und man müsse keine zusätzliche interne Evidence der individuellen natürlichen oder juristischen Person aufbauen. Die vierte Annahme ist, RCTs bewältigen alle wesentlichen Verzerrungen externer Evidence. Die fünfte Annahme ist, es könne eine theoriefreie empirische Studie geben. Die sechste Annahme ist, es gäbe eminente Theorie-Experten, die die bloß empirische Evidence ergänzen könnten – optimalerweise gestützt auf eine „Generally Accepted Theory". Alle diese sechs Annahmen verdienen eine sorgfältige Erörterung – nicht nur aus Dankbarkeit gegenüber Pfaff und Schmitt, sondern auch wegen ihrer weiten Verbreitung und ihrer hochriskanten Folgen. Ein Fazit geht auf die Risiken und Nebenwirkungen des EbM+Theory Approaches ein und macht eine Prognose zur ewigen Überlebensdauer dieses Approachs.

Zusammenfassung

Holger Pfaff und Jochen Schmitt greifen in ihrem „EbM+Theory Approach" (2021 und 2023) Fragen der Evidencebasierung von Entscheidungen auf, die die evidenzbasierte Versorgungsforschung vor 30 Jahren bearbeitete. Ihr EbM+Theory Approach beruht mindestens auf sechs Annahmen, deren praktisch-empirische Haltbarkeit geprüft wird. Ein Fazit geht auf die Risiken und Nebenwirkungen des EbM+Theory Approachs ein und prognostiziert diesem Approach, da Eminenz basierend, ein ewiges Überleben.

Schlüsselwörter

Wissenschaftstheorie, anthropologisch bescheidene Erkenntnistheorie, evidence-basierte Versorgungsforschung, EbM+Theory Approach, Systemtheorie

Crossref/doi

http://doi.org/10.24945/MVF.02.25.1866-0533.2710

1. Verbreitet sei ein Bereich stabiler „noncomplexer" Interventionen, bei denen man das richtige Handeln direkt aus RCTs generalisieren könne.

Es gäbe tatsächlich einen Bereich pflegerischen, ärztlichen und therapeutischen Handelns, für den die „traditional form of standardized, burocratic knowledge production, review, and transfer

– – – – – –

* Für Holger Pfaff zum 14. März 2025 in Dankbarkeit.

** Der Autor verwendet den Begriff Evidence immer in der englischen Schreibweise, weil Evidence im Englischen Indizien-Beweis bedeutet. „Evidenz" im Deutschen heißt seit der Romantik ein Evidenzerlebnis: eine Offenbarung, die einem wie Schuppen von den Augen fällt und keines empirischen Beweises bedarf. Also das Gegenteil von Evidence.

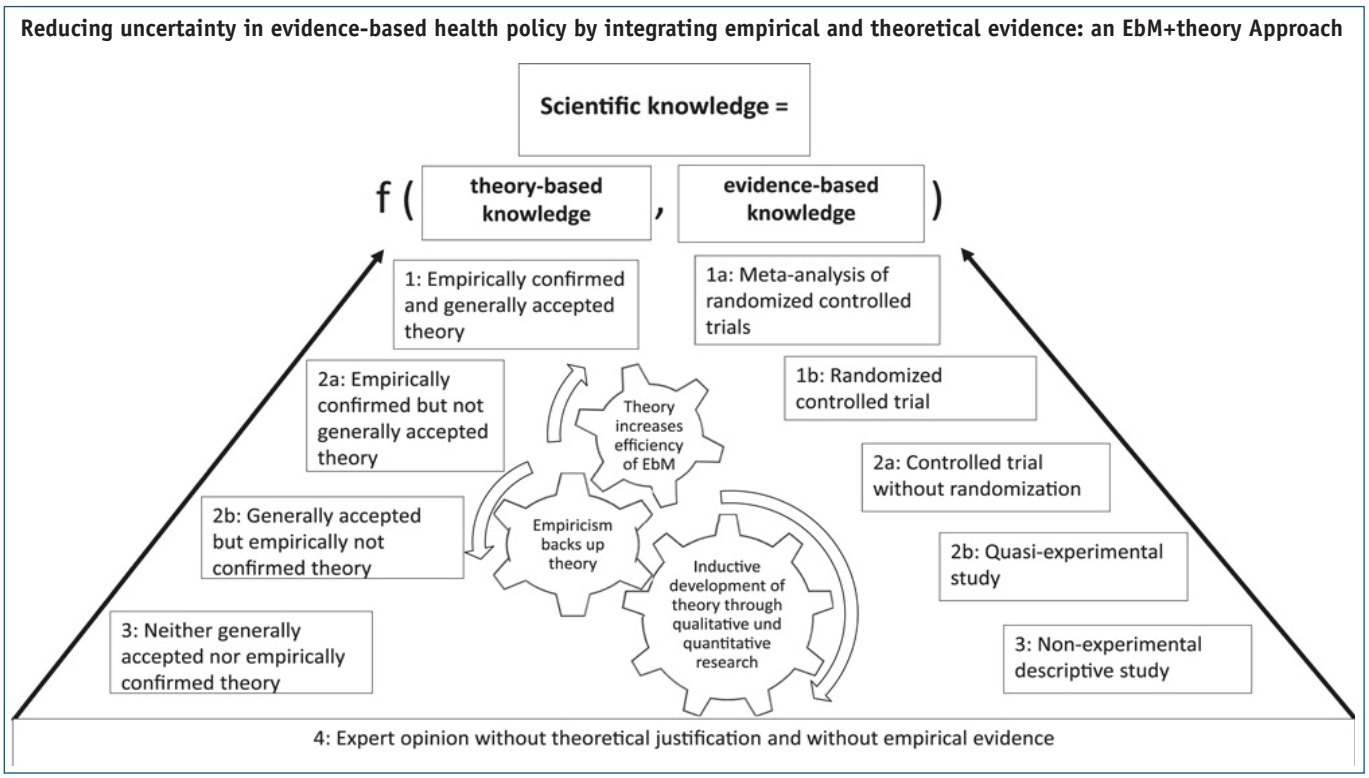


Abb. 1: Pyramide des wissenschaftlichen Wissens: Wissenschaftliches Wissen als Ergebnis der Kombination von evidencebasiertem und theoriebasiertem Wissen. Quelle: Pfaff H, Schmitt J, Reducing uncertainty in evidence-based health policy by integrating empirical and theoretical evidence: an EBM+Theory approach. J Eval Clin Pract 2023;1-15, hier: S. 3, Figure 1.

(knowledge processing) is suitable for stable environments, but not for unstable environments" (Pfaff & Schmitt 2021, S. 2). Einen solchen Bereich eines „stable environments" kann man sich ausdenken. In der Realität gibt es ihn höchstens ausnahmsweise.

Dass man aus Studien, und seien es RCTs, die richtige pflegerische, ärztliche und therapeutische Handlung im Einzelfall ableiten könne, diese Behauptung würde alle in diesen Behandlungen tatsächlich Praktizierenden weltweit äußerst überraschen. Für die Falsifikation dieser ersten Annahme gibt es einen systematischen Grund. Wer immer sich als Jugendlicher mit Statistik in der Schule beschäftigt hat, kennt den systematischen Grund: Unser Wissen aus Studien liegt weitgehend in der Gestalt von Häufigkeitsaussagen vor. Von 100 Personen, die eine bestimmte Behandlung nutzen, hilft fast nie allen diese Behandlung, sondern meist nur einem Teil, z. B. 76 von 100. Diese Häufigkeitsaussage wird gern im Konstrukt eines „average man" zusammengefasst. Von einer Häufigkeitsaussage kann man aber nie genügend sicher auf den Einzelfall schließen. Man kann im Vorhinein nie sicher sein, ob eine Person zu den 76 Personen gehört, die von der Behandlung einen großen Nutzen haben, oder zu den 24, bei denen die unerwünschten Wirkungen den Nutzen erheblich einschränken.

Für einen Viehzüchter-König mag diese Häufigkeitszahl hinreichend für die Behandlung seiner Herde erscheinen (siehe aber unten Annahme 3). In Public Health, Medizin, Pflege und Therapie behandeln wir aber den Einzelfall einer natürlichen oder juristischen Person (z. B. Kommune oder Land), nicht den konstruierten Durchschnittsfall. Dieses Dilemma wird auch nicht durch die Randomisierung im Zugang zu Interventions- oder Kontrollgruppe gelöst. Denn diese Randomisierung mag den Auswahl-Bias bewältigen und damit die unbeobachtete Heterogenität all der Faktoren (Kontexte, Ressourcen, Dispositionen, Handlungsweisen) ausgleichen, die zusammen mit unserer Intervention auf das Handlungsergebnis einwirken. Aber die Randomisierung gleicht diese unterschiedlichen Faktoren nur für die Gesamtgruppe aus, keinesfalls für eine einzelne natürliche oder juristische Person. Bereits diese einfache Tatsache widerlegt, dass das Ergebnis von Häufigkeitverteilungen – also im besten Fall wohldokumentierter externer Evidence aus Studien – allein für eine Handlungsentscheidung genügen kann. Es muss immer die interne Evidence der natürlichen oder juristischen Person aufgebaut werden. Schon deswegen gibt es das „stable environment", das Pfaff und Schmitt für große Bereiche der Versorgung voraussetzen, kaum.

Im Einzelfall einer natürlichen oder juristischen Person ist nahezu jede Handlung komplex und in einem „unstable environment". Das gilt auch, wie wir belegen werden, für Interventionen mit Pharmaka. Kontinuierlicher Aufbau interner Evidence ist unverzichtbar.

Fast jede Behandlungsentscheidung – ob Selbstbehandlung oder Behandlung anderer – ist eine Entscheidung unter Unsicherheit in einem durch diese Unsicherheit „unstable environment". Auch die größten und bestdokumentierten Studien heben diese Unsicherheit des Einzelfalls nicht auf. Es bleibt immer eine existenzielle Entscheidung unter Unsicherheit zu treffen. (Nur im Fall einer vorher nicht abzusehenden number needed to treat = 1, wenn alle Behandelten von einer Behandlung in gleicher Weise profitieren, wäre eine Entscheidung überflüssig.) Ende des letzten Jahrtausends ist daher die Standardsituation *Evidencebasierten professionellen Han-*

delns so definiert worden:

Medizin, die therapeutischen Wissenschaften und die Pflegewissenschaft sind alle – wie die beratenden Politik-, Wirtschafts- und Sozialwissenschaften (Behrens und Langer 2004 und 2022, Soziale Welt 2015) – **Handlungswissenschaften** mit einem spezifischen, von den kontemplativen Wissenschaften der Biologie und Soziologie u. a. trennbaren Gegenstand:

Ihr Gegenstand ist die zukunftsunsichere, aber vernünftige innovative Krisenentscheidung im jeweiligen einzigartigen Fall – unter Handlungsdruck und Begründungszwang gemeinsam mit den je einzigartigen Klienten.

Sie als „Künste" statt als Handlungswissenschaften zu bezeichnen, wie das Gross und Löffler (1997, S. 8) tun, macht ihre Praxis als „vernünftige" undiskutierbar, unkritisierbar, unerforschbar.

Diese Entscheidung soll keinesfalls durch Münzwurf gefällt werden, sondern gerade unter Reflexion aller Unsicherheit vernünftig begründbar sein. Deswegen lehnten wir Ende des vorigen Jahrtausends die Aufteilung in der „Theorie der Medizin" ab, die Forschung sei Wissenschaft, die Praxis mehr Kunst (Gross und Löffler 1997, S.8). Denn dann wäre eine Behandlungsentscheidung gar nicht mehr vernünftig kritisierbar und zu rechtfertigen. Sie würde beliebig. „Ut aliquid fiat"– also eine beliebige Handlung, um irgendetwas zu tun – ist aber in den Gesundheitswissenschaften zu Recht seit Jahrtausenden verpönt.

Wer muss sich in die Lage versetzen, diese existenzielle, nicht aus Studien ableitbare Entscheidung zu fällen?

Offenbar die, die die Folgen dieser Entscheidung tragen. Das sind nur bei grober Fahrlässigkeit oder Vorsatz die pflegerisch, ärztlich oder therapeutisch Behandelnden allein. Es sind in aller Regel die „Behandelten" oder „Beratenen". Deshalb müssen sie „interne Evidence" aufbauen können. Interne Evidence umfasst alles, was man nur von der zu behandelnden Person wissen kann. Sie wählt die für sie relevanten outcomes einer Behandlung, also jene Teilhabe am Leben , die sie selbstbestimmt unter Restriktionen erreichen will (vgl. die UN BRK, in Deutschland § 1 Sozialgesetzbuch IX). Nur diese natürliche oder juristische Person nimmt die Unbill und Schmerzen, aber auch Erleichterungen wahr. Nur von dieser Person sind individuelle Ressourcen und Kontexte zu erkennen. All diese höchstpersönlichen Eigenschaften und Informationen kann man aus keiner Studie erschließen – auch nicht, siehe unten, wenn man die „Diagnose" kennt. Nur von der Person selbst sind diese höchstpersönlichen Eigenschaften, die „interne Evidence" zu erfahren. Typischerweise kennt eine natürliche oder juristische Person ihre interne Evidence nicht, bevor sie krank, beratungsbedürftig oder pflegebedürftig wird. Man kann sie daher nicht wie „Präferenzen" in der ökonomischen Konsumtheorie abfragen. Denn Personen haben ‚Präferenzen' für oder gegen Mohnbrötchen oder Sahnetorte. Aber höchst selten, bevor sie krank, beratungs- oder pflegebedürftig werden, für oder gegen Amputationen oder nebenwirkungsreiche Eingriffe.

Die interne Evidence dafür müssen sich Personen in der Krise erst erarbeiten – manche Personen schätzen es, wenn sich ein Professionsmitglied für ihren Aufbau interner Evidence mit ihnen niedersetzt, mit dem sie alle Optionen und ihre Bedürfnisse offen besprechen und ihre interne Evidence aufbauen können (Behrens & Langer 2004, 2022, Behrens 2019). Wenn sie ihre interne Evidence aufgebaut haben, ist für sie eine existenzielle Entscheidung zumutbar. Der Aufbau interner Evidence ist also etwas ganz Anderes, als die Zustimmung (die Unterschrift) zu einer Entscheidung zu erlangen, die pflegerisch, ärztlich oder therapeutisch Behandelnde – fürsorglich bevormundend – für diese Person für richtig halten.

Der Aufbau interner Evidence für eine Entscheidung kann in der Realität nicht nur einmal, z. B. am Anfang einer Behandlung, erarbeitet werden. Da sich im Genesungsprozess oder in der Krise die Situation, also das entscheidungsrelevante „environment" ändert – manchmal täglich – ist es mit einer einmaligen Entscheidung nicht getan. Es muss immer wieder entschieden werden. Immer wieder muss interne Evidence als Entscheidungsgrundlage aufgebaut werden. Dazu ist fast nie viel Zeit.

Daher ist das „environment" fast jeder Behandlung oder Beratung doppelt „instable". Einerseits ändert sich die verfügbare externe Evidence. Andererseits und oft viel schneller ändert sich die Situation, in der entschieden werden muss, und damit die aufzubauende interne Evidence. Aus der internen Evidence ergeben sich regelmäßig neue Fragen an die verfügbare externe Evidence. Aber die externe Evidence reicht aus den genannten systematischen statistischen Gründen so gut wie nie ganz aus. Es bleibt immer eine Entscheidung unter Unsicherheit zu treffen. Von einem „stable environment" kann keine Rede sein. Man kann sich selbstverständlich die Utopie eines „stable environment" ausdenken. Vorzufinden ist sie damit noch nicht.

2. Annahme, man könne in Organismus und Psyche getrennt eingreifen

Mit einer zweiten, womöglich utopischen Unterstellung versuchen die Autoren ihre These von der Verbreitung noncomplexer Interventionen zu retten. Sie schreiben, es gäbe

„interventions in healthcare, if noncomplex, often intervene in one of these subsystems of the healthcare system, for example, the organisme (e.g., drugs and surgical intervention). In contrast, complex interventions occur in more then one subsystem, for example, in organisms and psychic systems"

(Pfaff and Schmitt 2023, S. 8).

Komplexe Interventionen sind bei ihnen also dadurch definiert, dass sie in Organismus und Psyche zugleich intervenieren. Gibt es denn Interventionen, die das nicht tun? Die Vorstellung der getrennten Intervention in Organismus und Psyche ist weitverbreitet, seit „Psychologen" das Leib-Seele-Thema von Theologie und Philosophie dazu nutzten, sich einen eigenständigen Arbeitsbereich neben der Medizin zu schaffen. In der Folge schienen Seelen ohne Körper und Körper ohne Seelen behandelbar. Diese Konstruktion verdient Klärung. In der Tat wäre es ein Traum nicht nur vieler Behandelnder, sondern auch ihrer Patienten, man könne Organe unabhängig von Personen behandeln, z. B. durch Medikamente oder Operationen. Wie bei modernen Autos könne man Teile aus Personen entnehmen und Organe und Gewebe in Ruhe reparieren, um sie dann runderneuert wieder einzusetzen oder durch neu gefertigte zu ersetzen. Leider (?) ist das weitgehend eine Utopie. Organe hängen an der Person und sind von ihr nicht dauerhaft erfolgreich

zu trennen. Organisches, psychisches und soziales System einer Person lassen sich nicht genügend dauerhaft getrennt behandeln. Eine Intervention in einen Organismus ohne Psyche ist so utopisch wie eine Intervention in eine Psyche ohne Körper. Unter Vollnarkose kann man operieren und Drugs injizieren, aber die Vollnarkose kann nie so lange währen, bis die Genesung völlig abgeschlossen ist. Nur handelnde Personen können völlig gesunden. Jedes Medikament wirkt auch psychisch, und jede Psychotherapie löst organische Reaktionen aus.

Sieben Wirkfaktoren einer Medikamentenverschreibung:

Von der Utopie, es ließen sich Organismen ohne Psyche behandeln, hat sich die Pharmazie seit Langem verabschieden müssen. Die Wirkung fast jedes Medikaments wird seitdem am Vergleich mit einem Placebo gemessen. Sie wird also mit der Wirkung der psychosozialen Beziehung ohne die zu untersuchende Substanz verglichen. Die psychosoziale Beziehung ist zweifellos ein Haupteffekt medizinischer Behandlung überhaupt, wenn nicht ihr größter Effekt. Eine Substanz muss schon sehr stark wirken, um den Placeboeffekt zu übertreffen.

Bei jeder verschriebenen oder empfohlenen Tablette wirken mindestens 7 Faktoren:

1. Eine schöne Ärztin oder pharmazeutisch-technische Assistentin
2. im weißen Kittel
3. sieht einem fest in die Augen
4. und spricht mit freundlicher Stimme
5. „Nehmen Sie das morgens, mittags und abends"
6. „und es wir Ihnen bald besser gehen"
7. Wirkt die Substanz.

Die Wirkung der Faktoren 1 – 6 kann so stark sein, dass der 7. Faktor, die Substanz des Medikaments, selbst wirkungslos oder sogar eher schädlich sein kann: Die Wirkung des 7. Faktors kann übertroffen werden von den 6 anderen Faktoren. Besonders interessant ist der 5. Faktor: Chronobiologisch mag es Unsinn sein, die gleiche Dosis morgens, mittags und abends zu nehmen, weil die Substanzen je nach Tageszeit unterschiedlich wirken. Aber das „morgens, mittags und abends" ruft Zeiten des (Tisch-)Gebets auf oder zumindest Zeiten verstärkter Selbstsorge in Pausen fürs Essen und Trinken. Darum liegt in Faktor 6 eine große rituelle Prophezeiung.

Der vorige Präsident des für die Zulassung von Medikamenten durch den Gemeinsamen Bundesausschuss (G-BA) bedeutenden IQWiG, Windeler, hat den Verfasser auf einem Kongress einmal gefragt, was er denn mache, wenn nicht eine schöne PTA oder Ärztin ihm in die Augen sähe, sondern ein griesgrämiger Arzt (Faktor 1). „Shit happens. Thats life" war die Antwort, in seinem Alter sei man Kummer gewohnt. Die systematisch korrekte Antwort ist selbstverständlich, dass die Schönheit der PTA oder der Ärztin nur einer von 7 Wirkfaktoren ist, dessen Mangel von den anderen 6 Wirkfaktoren kompensiert werden muss und zuweilen sogar kann.

Was die von Pfaff und Schmitt angeführten scheinbar nur organischen chirurgischen Eingriffe angeht, so gilt für sie: Es gibt wohl fast keine psychisch beeindruckendere Behandlung, als sich unters Messer zu legen. Die These, dass chirurgische Operationen nur in den Organismus, nicht aber in die Psyche eingriffen, müsste genauer belegt werden, als die Autoren das tun.

Nun könnte man entgegnen: Sind nicht alle diese begleitenden Wirkfaktoren einer Tabletten-Verschreibung, diese ganze unbeobachtete Heterogenität vielfältiger mitwirkender Faktoren durch die Randomisierung statistisch ausgeglichen, sodass nur die Wirkung der Tablettensubstanz rein als Ergebnis übrig bleibt? Ja, in der Häufigkeitsverteilung einer großen Studie vielleicht, die die externe Evidence bildet. Aber selbstverständlich nicht für den Einzelfall. Denn, wie oben begründet, man kann nicht von einer Häufigkeitsverteilung auf den Einzelfall schließen. Und: In der alltäglichen Praxis oder Beratung wird nie eine ganze Studienpopulation insgesamt behandelt, sondern immer eine spezifische natürliche oder juristische Person nach der anderen. Auf die scheinbare Ausnahme der Viehzüchter-Könige, die nur Interesse an der Herde, nicht am Einzelnen haben, kommen wir unter der dritten Annahme.

Gewährsmann Luhmann für getrennte Intervenierbarkeit in das organische, psychische und soziale Systeme?

Pfaff und Schmitt berufen sich für ihre Trennung des organischen, psychischen, sozialen und Maschinen-„healthcare subsystems" auf Luhmann (Pfaff & Schmitt 2023, S. 7). Daraus, dass es psychologisch oder ärztlich Behandelnde (Interaction), psychologische oder ärztliche Praxen und Kliniken (organization) und Berufsordnungen (society) gibt, scheinen sie zu schließen, dass tatsächlich der Mensch in getrennt zu behandelnde Organe und Psyche zerfiele. Das halte ich für eine zu weit gehende Biologisierung von Kommunikationssystemen. So weit würde ich keineswegs gehen, die Profession Pflege auch nicht und Luhmann wohl erst recht nicht. Luhmann (Luhmann 1975, 1981, 1993, Behrens 2019) untersucht die Ausdifferenzierung von Kommunikationssystemen. Kommunikationssysteme kann man sich grob als Zuständigkeiten für Fragen vergegenwärtigen. Personen (Menschen) gehen nicht als Personen in diesen Zuständigkeitssystemen auf, sondern bilden deren „unmarked space", von dem man nicht mehr weiß und nicht mehr wissen muss, als dass er nicht zum Kommunikationssystem gehört. So ist die Wissenschaft in modernen Gesellschaften für die Unterscheidung von wahr und nicht wahr zuständig, auch wenn die Wissenschaftlerinnen und Wissenschaftler sich als Personen noch für ganz andere Unterscheidungen interessieren mögen. Das (kapitalistische) Wirtschaftssystem ist auf die Unterscheidung Geld/mehr Geld ausgerichtet, auch wenn Manager und Erwerbstätige als Personen noch ganz andere Interessen haben mögen als nur „mehr Geld".

Richter und Anwälte, zuständig für die Unterscheidung Recht/Unrecht, können als Personen in ihrer Freizeit die Folgen ihrer Unterscheidungen tief bedauern, als Richter und Anwälte müssen sie sich an sie halten. Ein Richter, der auf seinem Richterstuhl das persönliche Interesse an mehr Geld verfolgt, ist korrupt und agiert nicht mehr im Kommunikationssystem Recht. Ebenso sind Wissenschaftlerinnen und Wissenschaftler, die das persönliche Interesse an mehr Geld oder an der Rechtfertigung des Rechts über die Unterscheidung wahr/nicht wahr stellen, kein Teil des Kommunikationssystems Wissenschaft mehr, auch wenn sie noch Gehalt aus einer Forschungseinrichtung beziehen und das Gehalt nicht zurücküberweisen. Psychologen sind in dem Sinne kommunikativ für das psychische System zuständig, indem sie Seelen ohne Körper behandeln. Sozialarbeiter und Soziologen sind kommunikativ für das soziale System zuständig, auch wenn die Abgrenzung dieser

Systeme schwerfällt. Und Ärztinnen und Ärzte, sind sie zuständig für das organische System? Sie sind bei Luhmann für „Krankheit" zuständig in der Unterscheidung zum unmarked space „ohne Befund". So gibt es für Luhmanns Mediziner Tausende von Krankheiten, aber nur eine Gesundheit, über die sich medizinisch nicht mehr sagen lässt, als dass (noch) kein Krankheitsbefund zu diagnostizieren ist. Gesunde sind für das Kommunikationssystem der Krankenbehandlung (noch) nicht als krank Diagnostizierte (Luhmann 1975,1981,1993). Daher gibt es für Luhmann eigentlich kein Gesundheitssystem, sondern nur ein System der Behandlung von Krankheiten. Ein Kommunikationssystem, das auf Gesundheit und Schönheit ausgerichtet ist und in das Personen täglich viel Mühe investieren, ist das – von Luhmann nicht untersuchte – Pflegesystem mit dem Code ‚gesundheitsförderlich gepflegt/ungepflegt' (siehe Behrens 2009 und 2019).

Das Problem des Kommunikationssystems der Krankenbehandlung liegt offensichtlich darin, dass Krankheiten zu behandeln so schwer von erkrankten Personen zu trennen ist. Das ist in anderen Kommunikationssystemen ganz anders. Ein Gerichtsurteil kann juristisch richtig oder falsch sein, unabhängig von der Person der Richtenden wie vor allem der Gerichteten. Eine Kapitalanlage kann erkennbar mehr Geld gebracht oder aber verloren haben, ohne dass man zu dieser Feststellung die Personen kennen muss, denen das Kapital gehört oder die für die Kapitalvermehrung beschäftigt waren. Aber im System der Behandlung von Krankheiten werden bekanntlich Personen behandelt, nicht Krankheiten. Eine Krankheit kann nicht geheilt oder auch nur bewältigt werden, ohne entscheidende Mitwirkung der Person. „Operation gelungen, Patient tot" gilt nicht als hinreichender Erfolgsindikator einer Krankenbehandlung. „Fiat justitia et pereat mundus" dagegen beweist einen Erfolg im Rechtssystem. Recht kann gültig gesprochen sein, ohne dass Beklagte und Kläger überleben. Aber Krankheiten sind nicht zu behandeln ohne die Personen, die erkrankten.

In Luhmanns riesigem Gesamtwerk finden sich nur drei kurze Schriften zum System der Krankenbehandlung, obwohl das Interesse des Publikums an diesem Thema – gemessen an Zitierungen der drei kurzen Aufsätze – enorm war und Luhmann selbst in etwa 90 Anmerkungen einige Male auf das System der Krankenbehandlung zurückkam (Der Co-Herausgeber von Luhmanns inzwischen fünf Bänden „Schriften zur Organisation", Ernst Lukas, hat alle diese 90 Anmerkungen und Fußnoten identifiziert, siehe Luhmann 2019–2022). Luhmann hat aber im Gegensatz zu allen anderen Kommunikationssystemen der Gesellschaft das System der Krankenbehandlung nie in einem Buch systematisch dargestellt. Das Buch „Die Krankenbehandlung der Gesellschaft" gibt es nicht, obwohl mehr Menschen darauf warteten als auf das Buch „Die Kunst der Gesellschaft". Womöglich war es bei dem System der Krankenbehandlung zu unplausibel, die Person im „unmarked space" des Kommunikationssystems zu verorten. Deswegen kam Luhmann auf das Thema trotz größtem Interesse des Publikums nie systematisch zurück. Das ist wenigstens unsere Vermutung.

Für ihre weitreichende Behauptung, es seien tatsächlich getrennte Interventionen in „organische, psychische und soziale healthcare-Subsysteme" (Pfaff & Schmitt 2023, S. 7) möglich, können sich Pfaff und Schmitt nicht auf Luhmann berufen, so viel ist sicher.

3. Die scheinbare Unsicherheitsreduktion der Viehzüchter-Könige: das Generalisierungsproblem

Aber erlaubt die externe Evidence aus Häufigkeitsverteilungen nicht wenigstens den Schluss auf die richtige Behandlung von Kollektiven, wenn der Schluss auch nicht auf Individuen möglich ist? Kann man bei der Behandlung von Kollektiven auf den Aufbau interner Evidence verzichten? In der Tat haben Besitzer großer Herden kaum den Anspruch, dem einzelnen Individuum gerecht werden zu müssen. Dass sie aus einer Häufigkeitsverteilung nicht auf das einzelne Individuum statistisch schließen können, muss Besitzer großer Herden nicht extrem beunruhigen. Vielmehr scheint eine epidemiologische Häufigkeitsverteilung genau ihrem Interesse an der ganzen Herde zu entsprechen und nicht, wie bei pflegerisch, ärztlich oder therapeutisch Behandelnden, an den Individuen. Daher reduzieren Häufigkeitsverteilungen die Unsicherheit bei Viehzüchtern viel mehr als bei pflegerisch, ärztlich oder therapeutisch Behandelnden. Spätestens mit dem Absolutismus der „Viehzüchter-Könige" Englands und Frankreichs war der Blick des Königs auf das Volk – beraten vom Arzt Quesnay – der Blick eines Viehzüchters, der sich um die Gesundheit seiner Herde, des Volkes, sorgt, um das meiste aus ihm herauszuholen. Der gute Hirte weidet seine Schafe und führt sie, wenn die Zeit gekommen ist, in die Schlacht(bank). Dazu scheint die Kenntnis von Häufigkeitsverteilungen zu reichen. Das gilt aber nur scheinbar. Denn ein methodische Problem bleibt, auch wenn man sich mit Häufigkeitsverteilungen zufriedengibt: Der Schluss von einem Studienergebnis (Häufigkeitsverteilung) auf eine nächste Situation, in der man zu handeln hat, bleibt herausfordernd. Generationen von Soziologinnen und Soziologen, Psychologinnen und Psychologen haben dieses Problem zu lösen versucht.

Es überrascht, dass Pfaff und Schmitt dies Problem gar nicht erwähnen. Dabei ist es quälend. Der große Cronbach (1982) hat sich jahrzehntelang heroisch bemüht, es zu lösen, und die Bedingungen formuliert, unter der ich von einer Feldstudie auf die Situation schließe, in der ich handeln muss: Das Ergebnis einer Feldstudie auf ein nächstes Feld zu übertragen ist nach Cronbach (1982) nur möglich, wenn man sich vom Vorliegen vierer Bedingungen überzeugt hat: Die Units (Probanden), Treatments (intervenierende Behandlungen), Observing Operations (Beobachtungen, Messungen) und das Setting (Kontext) müssen in beiden Feldern übereinstimmen (Cronbach 1982, S. 78 ff.).

Nach den Anfangsbuchstaben dieser vier notwendigen Übereinstimmungen spricht Cronbach von „UTOS". Wenn auch nur eine diese vier Übereinstimmungen nicht nachweislich erfüllt ist, ist nach Cronbach eine Übertragung nicht möglich. Aber wie kann man wissen, ob diese vier Übereinstimmungen gegeben sind? Offensichtlich kann man gar nicht erkennen, ob diese Bedingungen gegeben sind oder nicht, ohne die Situation, für die man Schlüsse ziehen will, neu und gründlich zu untersuchen, also die interne Evidence dieser Situation aufzubauen. (Soziale Welt 2015, Behrens 2019, S. 99). Folglich muss auch der Viehzüchter-König, der nur am Schicksal seiner Herde bzw. seines Volkes interessiert ist und nicht am einzelnen Individuum, interne Evidence aufbauen, sobald er von heute auf morgen schließen will oder von einer Herde auf eine andere.

Womöglich erwähnen Pfaff und Schmitt das Übertragungs- oder Generalisierungsproblem bei der Evidencebasierung gar nicht, weil

sie einen Deus ex Machina kennen, der das Generalisierungs- oder Übertragungsproblem für sie löst: den Theorieexperten (siehe Annahme 6).

4. Annahme: RCTs bewältigen alle wesentlichen Verzerrungen externer Evidence

Pfaff und Schmitt (2021 und 2022) sind erstaunlich unkritisch, was Randomized Controlled Trials in ihrer Hierarchie Evidencebasierten Wissens angeht (Pfaff und Schmitt 2023, S. 3 und Figure 1; 2021, S. 8, Figure 2). In ihrer Grafik gibt es eine eindimensionale Hierarchie der Evidence, in der RCTs ganz allein oben stehen. Darin hängen sie fast 30 Jahre hinter der Selbstreflexion der verschiedenen Netzwerke für Evidencebasierte Praxis und damit der Versorgungsforschung hinterher. Schon Ende des letzten Jahrtausends hatte es sich dort herumgesprochen, dass Randomisierung des Zugangs zur Interventions- und Kontrollgruppe nur eine der vielen Verzerrungsgefahren („Bias") von Studien bewältigt. Diese eine Verzerrung ist die Auswahl-Verzerrung, der sogenannte Selection-Bias. Die Auswahl-Verzerrung bewirkt, dass Interventions- und Kontrollgruppe nicht vergleichbar zusammengesetzt sind, vor allem auch, weil man die Heterogenität der Gruppen vorher nicht beobachten konnte. Randomisierung bewältigt bei genügend großer Gruppengröße vorher unbeobachtete Heterogenität. Aber die Auswahl-Verzerrung ist nur eine von vielen Verzerrungsgefahren.

All die vielen anderen Verzerrungsgefahren können durch Randomisierung allein nicht bewältigt werden. Wie in Kapitel 4 von Behrens & Langer schon 2004 und 2022 (dort S. 155–304) belegt , gibt es (leider) keine einzige Methode, die auch nur alle relevantesten Verzerrungsgefahren zugleich bewältigt. Daher kann es logisch auch keine eindimensionale Gütehierarchie der Studien geben, wie Pfaff und Schmitt sie 2021 und 2023 referieren, da jeder Studientyp nur bestimmte Verzerrungsgefahren bewältigt. Eine „Meta-Analysis of randomised controlled trials", bei Pfaff und Schmitt 2021 und 2023 angeblich die höchste Stufe des „evidence-based konwledge", kann in der Aussagekraft völlig irrelevant sein, wenn der Wirkungsmechanismus der Intervention nicht vergleichbar durchgeführt, die Repräsentativität der Grundgesamtheit unbekannt oder das Outcome nicht angemessen, also den Bedürfnissen nach selbstbestimmter Teilhabe i. S. der UNO und des § 1 Sozialgesetzbuch IX entsprechend bestimmt wurden. Viele RCTs sind völlig irrelevant und auch nicht übertragbar. Spätestens im 4. Semester erwarteten die Mitglieder der Netzwerke für evidencebasierte Praxis und Versorgungsforschung von allen Studierenden der verschiedenen Gesundheitswissenschaften, dass sie die Stärken und Schwächen jeder Studienmethode abwägen können. Beobachtungsstudien und repräsentative Panelstudien haben Stärken, die von realen RCTs nie zu erwarten sind, und umgekehrt. Jeder Studientyp bewältigt eine Verzerrungsgefahr am besten. Wer eine Studie unternimmt, muss sich entscheiden, welche Verzerrung er oder sie am meisten fürchtet. Und wer Studien für sich auswertet oder bewertet, muss sich auch über seine eigenen Prioritäten bei der Bewältigung von Verzerrungsgefahren klar werden. Das ist im Einzelnen in Kapitel 4 von Behrens und Langer 2004 und 2022 dargelegt. (Die Aussagekraft aller Studien beruht auf hermeneutisch-interpretativen (vulgo: qualitativen) Verfahren .

Die Erwartungen an die Studierenden der Gesundheits-, Pflege- und Therapiewissenschaften, die einzelnen Studienmethoden unterschiedlichen Verzerrungsgefahren zuordnen zu können, wurden nie enttäuscht. Im vierten Semester konnten fast alle Studierenden kompetent Stärken und Schwächen jeder Methode abwägen und wussten, dass statistische „quantitative" Studien nur Sinn haben auf der Basis interpretativ-hermeneutischer „qualitativer" Untersuchungen. .

Pfaff und Schmitt 2021 und 2023 kennen aber merkwürdigerweise keine anderen „High-Quality"-Studien zur Erzeugung externer Evidence, des „evidence-based knowledge", als Meta-Analysis of randomized controlled trials. Noch viel merkwürdiger: Sie kennen nur eine Hauptschwäche von RCTs: ihren Zeitbedarf. So optimistisch, was RCTs angeht, war bereits vor fast 30 Jahren niemand in den Netzwerken von EbM, EbN und EbP. Der Zeitbedarf ist zweifellos nur ein und nicht das größte Problem.

5. Annahme: es gäbe ein „Evidence based knowledge", das nicht theoriebasiert sei

Die Lösung sehen Pfaff und Schmitt 2021 und 2023 (Figure 1 bzw. 2) darin, dem „evidence-based knowledge" ein „theory-based knowledge" zur Seite zu stellen. Der Vorschlag basiert auf der überraschenden Unterstellung, es gäbe überhaupt ein „evidence-based knowledge", das nicht „theory-based" sei. Ohne eine Theorie ist, anthropologisch gesprochen, ein Mensch gar nicht fähig, irgendetwas zu sehen, zu beobachten, zu verstehen, zu erklären oder zu behandeln (Singer 2017, Behrens und Langer 2004 und 2022). Niemand hebt auch nur eine Hand zu einer geplanten Handlung, ohne eine (zumindest implizite) Theorie zu haben, was mit dieser Handlung bewirkt werden könne (Singer 2017, Behrens & Langer 2004, Behrens 2019). Die jeweilige Theorie kann sich als falsch erweisen und erweist sich oft als falsch. Dann berichtigen Menschen ihre theoretischen Erwartungen und bilden eine neue theoretische Erwartung, die sie ihren Handlungen zugrunde legen und die sich wieder als falsch erweisen können. Sie verfeinern ihre Theorien und werden immer seltener enttäuscht.

Nur eins ist den Menschen nicht gegeben: die Objektivität der Welt sich in Menschen direkt wie auf ein unbeschriebenes Blatt Papier einschreiben zu lassen. Das hatten Locke (1670) und Hume (1748) gehofft. Zuletzt haben Lenin (in der Kritik an Mach) und Stalin diese Vorstellung von Objektivität vertreten. Aber die Hoffnung auf eine Objektivität der Welt, die sich in uns wie auf ein unbeschriebenes Blatt einschreibt, ist vergebens. Wir nähern uns der Objektivität der Welt, auch dem gestirnten Himmel über uns, ausschließlich als Subjekte, die subjektiv-theoretische Erwartungen konstruieren und sie bei jeder Enttäuschung neu bilden. Das ist im Alltag nicht anders als in der Wissenschaft (Behrens & Langer 2004, 2022, Behrens 2019).

Nur Subjekte bilden Theorien und können ohne subjektive Theorien weder wahrnehmen noch handeln noch theoretische Erwartungen, die sich als falsch herausstellten, zu neuen Theoriekonstruktionen nutzen. Charles S. Peirce (2008) hat das vor fast 200 Jahren begründet. Rudolf Carnap (1966) hat Anfang des 20. Jahrhunderts als Letzter eine induktive Theoriebildung versucht. Als er scheiterte, verlegte er sich auf eine Theorie der Entscheidung unter Unsicherheit und Nichtwissen. Es sind Peirce, der späte Carnap (1966) nach dem Scheitern des Induktivismus und der Kunsthisto-

riker und Kriminalist Carlo Ginzburg, (1983), auf denen die – individuellen natürlichen und juristischen Personen verpflichtete, „Indizienbeweise" (Evidence) nutzende und interne Evidence in der Begegnung aufbauende – evidencebasierte Praxis meist gründet.

Wenn dem Menschen ein Wahrnehmen und Handeln ohne theoretische Erwartungen anthropologisch gar nicht möglich ist, ist völlig auszuschließen, dass ausgerechnet die komplizierten Handlungen zur Erzeugung von evidence-based knowledge ohne theoretische Erwartungen auskommen. Die Theorien können falsch sein und sind es oft. Aber theorieloses evidence-based knowledge kann es anthropologisch gar nicht geben. Daher kann theory-based knowledge keine Ergänzung von evidence-based knowledge sein, weil eine mangelhafte oder fruchtbare Theorie bereits Grundlage jeder Studie ist, auch wenn sie den Forschenden wenig bewusst ist und von ihnen daher nicht expliziert wird.

Die Fiktion eines theoriefreien evidence-based knowledge führt dazu – insbesondere wenn man solche Überzeugungen hat wie die geschilderten Annahmen 1 bis 5 –, dass die impliziten und expliziten Theorien jeder Studie gar nicht mehr wahrgenommen und kritisch analysiert werden. Das ist ein großes Risiko und eine erhebliche unerwünschte Wirkung. Die Ergebnisse des evidence-based knowledge erscheinen als „Daten und Fakten", ohne zu untersuchen, mit welchen theoretischen Erwartungen die „Fakten" gemacht (lateinisch facere) und jedes „Datum" gegeben (lateinisch dare) wurde.

6. Experten einer „Generally-accepted Theory" gegenüber „evidence-based knowledge" – „generally-accepted" Eminenzen, ihre Risiken und Nebenwirkungen

Sieht man sich die beiden Aufsätze von Pfaff und Schmitt genauer an, steckt hinter dem Vorschlag, „evidence-based knowledge" um „theory-based knowledge" zu ergänzen, auch gar nicht der Nachweis, es gäbe theoriefreies evidence-based knowledge. Vielmehr geht es ihnen explizit darum, „Experten" und ihre „generally accepted" Aussagen (expert opinions) gegen evidence based knowledge in Stellung zu bringen und zu verteidigen (Pfaff und Schmitt 2021). Systematisch deutlich wird das in ihrer oben zitierten Hierarchie des „theory-based knowledge". Als eigenständiges Merkmal neben „empirical confirmed" steht „generally accepted". Damit haben der Konsens und die Zitierkartelle der Ordinarien neben der empirischen Bestätigung ein eigenes Gewicht. Eine empirisch bestätigte, aber nicht generell von den Ordinarien und ähnlichen Experten akzeptierte Theorie darf nicht als Spitze des „theory-based knowledge" angesehen werden. Sie erweist sich dieses Platzes erst als würdig, wenn sie „generally accepted" wurde (Pfaff und Schmidt 2023 und 2021, S. 3.). Sei eine Theorie erst einmal auf diesem Spitzenplatz, könnten ihre Experten nahezu unverzüglich, ohne zeitraubende Studien der EbM, ohne zeitraubenden Aufbau interner Evidence und ohne weitere empirische Bestätigung die Politik beraten, die unter Zeit- und Entscheidungsdruck steht.

Um die Geschichte der eminenten Experten für die Bewegung der evidence-based Praxis zu verstehen, lohnt es sich, auf die Entstehung dieser Eb-Bewegungen in antiautoritären Studierendenbewegungen zurückzublicken. An der McMaster-Uni in Kanada forderten die Studierenden, ächzend unter der Menge des Lehr- und Prüfungsstoffs, nur noch solche Theorien lernen zu müssen, deren Evidence erwiesen sei. Professoren hatten in ihren Vorlesungen zu oft erklärt, die Evidence der Theorien kennten sie auch nicht. Sie wüssten nur, dass sie im Lehrstoff vorgeschrieben seien und die Behandlung nach diesen Theorien in Kanada vergütet würde. Die antiautoritären Studierenden dachten, wenn sie all diese Theorien nicht mehr pauken und nur noch Evidencebasierte Theorien lernen müssten, verringere sich der Prüfungsstoff erheblich. (Das war, wie wir heute wissen, falsch: Das problem-based Learning, das mit der Forderung nach Evidencebasierung aufkam, macht viel mehr Arbeit, als einfach im Frontalunterricht mitzuschreiben, welche Antwort auf eine Multiple-Choice-Prüfungsfrage angekreuzt werden muss (Them et al., 2002). Die Kritik der Studierenden an eminenten Experten richtete sich also gegen die von ihnen aufgerufenen mangelhaften Grundlagen. Eine Expert-Opinion, deren empirische Evidence nicht angebbar ihre Theorie begründete, sollte das Allerletzte sein. Die Vertrauenswürdigkeit einer Eminenz-Expertise sollte nicht mehr durch ihre Gehaltsklasse (C 4 bzw. W 3) oder durch ihre im Dienst verbrachten Jahre (= Erfahrung) bestimmt sein. Sondern allein durch die von ihr intersubjektiv nachprüfbar vorgelegten Belege. Für Cochrane Reviews ist es z. B. völlig unerheblich, in welcher Gehaltsklasse ihre Autoren arbeiten. Es zählt im Sinne von Habermas das Argument – nicht die Person, die das Argument vorträgt, ob sie nun eine Pflegeschülerin mit Hauptschulabschluss oder eine Chefärztin ist. Selbstverständlich haben Autorinnen und Autoren von Cochrane Reviews „Expertise" – aber ihre persönliche Meinung ist weitgehend irrelevant. Es überrascht nicht, dass die EbM-Bewegungen enorme narzisstische Kränkungen auslösten, als seien altgediente Chefärzt:innen nicht mehr besonders vertrauenswürdig. Eine Gegenbewegung war und ist nur zu verständlich. Heute geht es gar nicht mehr um den persönlichen Expertenstatus, sondern allein um die Expertise, auf der dieser Status nachprüfbar beruht.

Hier ist das von Pfaff und Schmitt 2023, S. 3 betonte Kriterium einer „Generally-accepted Theory" höchst problematisch.

In ihrer Pyramide theoretischen Wissens ist die „generally accepted theory" das Höchste an Vertrauenswürdigkeit, genau wie in ihrer Pyramide des „evidence-based knowledge" merkwürdigerweise die Meta-Analyse von RCTs. Die Gefährlichkeit des Kriteriums genereller Akzeptanz unter Theorieexperten ist vielfach offensichtlich geworden. So z. B. bei der Behandlung des Druckgeschwürs, unter dem fast alle Menschen, die bettlägerig sind, bald leiden – also viele Millionen Menschen. Die von Professuren und Experten „generally accepted theory" der modernen Behandlung und Vorbeugung des Druckgeschwürs („Dekubitus") war Föhnen und Eisen. Es entsprach der physiologischen Theorie, nach der Hautreizung abwechselnd durch Wärme (Föhnen) und Kälte (Eisen) zur besseren Durchblutung der Haut führe. Föhnen und Eisen war eine extrem zeitaufwändige und teure Behandlung. Dennoch wurde sie mit Liebe und Zuwendung praktiziert. So sehr entsprach sie der Theorie, der Erfahrung und der Intuition aller Experten, dass so gut wie niemand es ethisch verantwortbar oder sogar nötig fand, die Wirkung von Föhnen und Eisen an einer Kontrollgruppe zu messen, denen diese liebevolle und teure Behandlung mit Föhnen und Eisen versagt wurde. Dass sich nach Föhnen und Eisen das Druckgeschwür oft nicht zurückbildete oder sogar schlimmer wurde, führte fast

niemand auf Föhnen und Eisen zurück. Es schien in der Entwicklung dieser Erkrankung zu liegen. Fast kein Experte war davon auch nur im Mindesten überrascht. Es entsprach der Theorie, der Erfahrung und der Intuition. Bei mehr Dekubitus schien eher noch mehr Föhnen und Eisen nötig als weniger. Erst als mehr oder weniger zufällig doch ein Vergleich zwischen zwei Gruppen Dekubitus-Kranker mit und ohne Föhnen und Eisen stattfand – mit eher primitiven als raffinierten Methoden –, verbreitete sich allmählich die Erkenntnis, dass Föhnen und Eisen eine zwar liebevoll erbrachte, aber doch besonders quälende und schädigende Behandlung von Dekubitus war. Wem als Dekubitus-Erkranktem aus ökonomischem Geiz und Lieblosigkeit die Behandlung mit Föhnen und Eisen verweigert worden war, hatte Glück gehabt. Allmählich wurde Föhnen und Eisen mit derselben theoretischen Gewissheit, mit der diese Methode vorher für theoretisch alternativlos erklärt worden war, als Folter bezeichnet. Der Ausdruck Folter ist aber unangemessen. Denn Folter setzt Vorsatz voraus: Es werden vorsätzlich Menschen gequält. In diesem Fall liegt aber Schuld nicht im Vorsatz, sondern im grob fahrlässigen Vertrauen auf eine „generally accepted theory" (Behrens und Langer 2004 und 2022, Kapitel G1 und G2).

Aus welcher Zeit stammt die Trennung von Theory und empirischer Evidence?

Heute wissen wir, es kann anthropologisch keine empirische Evidence geben ohne Theorie (Behrens & Langer 2004, 2022, Singer 2017, Behrens 2019). Das war Jahrhunderte lang völlig anders. Denn es gab die Unterscheidung von menschlichem empirischem und göttlichem theoretischem Wissen. Wie man noch in Kants ausgesprochen witziger Schrift „Über den Streit der Fakultäten" (Kant 1995, S. 5–140) nachlesen kann, beruhten die höheren Fakultäten auf den Offenbarungen Gottes oder der von Gott eingesetzten Fürsten, während die niedere Fakultät, die „philosophische", auf menschlicher Empirie und Räsonnement beruhte. Privilegierten Zugang zum göttlichen Wissen vermittelten theologische Priesterweihen, die Erkenntniszugänge der Priester und später etwas eingeschränkt der Pfarrer waren definitiv von denen der Laien unterschieden. Privilegierten Zugang zum göttlich begnadeten Fürstenwissen vermittelte die fürstliche Berufung in die juristische Fakultät oder ins Richteramt. Für beide Fakultäten galt, „Du glaubst, was Dein Chef sagt" (so die nicht allzu freie Übersetzung von „cuius regio, eius religio"). Kant bemüht sich, die niedere Fakultät möglichst autonom von Ansprüchen der höheren Fakultäten zu halten. In der niederen philosophischen Fakultät soll allein menschliche Vernunft gelten. Akademische Weihen sollen in der niederen Fakultät keine Priesterweihen sein. „Daher schöpft der biblische Theolog (als zur oberen Fakultät gehörig) seine Lehren nicht aus der Vernunft, sondern aus der Bibel, der Rechtslehrer nicht aus dem Naturrecht, sondern aus dem Landrecht, der Arzneigelehrte seine ins Publicum gehende Heilmethode nicht aus der Physik des menschlichen Körpers, sondern aus der Medicinalordnung.... Daher müssen die oberen Fakultäten am meisten darauf bedacht sein, sich mit der unteren ja nicht in Mißheirat einzulassen, sondern sie fein weit von sich abzuhalten, damit das Ansehen ihrer Statute nicht durch die freien Vernünfteleien der letzteren Abbruch leide" (Kant 1995, S. 27 f., Randnote 23, fette Hervorhebungen von jb).

Um diese Autonomie der niederen philosophischen Fakultät zu erhalten, ist Kant zu weitgehenden Kompromissen bereit, um Überschneidungen zu regeln. Eine offensichtliche Überschneidung ist, dass die offenbarten Heiligen Schriften und Gesetzesbücher eben auch Texte sind. Und als Texte sind sie philologisch, also logisch-menschlicher Auslegung fähig, wie sie die niedere Fakultät betreibt. Kants Abgrenzungskompromiss ist, dass Pfarrer und Juristen, die sich bei der Textinterpretation menschlicher Philologie bedienen, ihre philologischen Interpretationen keinesfalls von der Kanzel oder im Gerichtssaal vortragen dürfen. Von der Kanzel und im Gerichtssaal soll Wissen nur als göttlich oder durch die von Gott begnadete Obrigkeit offenbartes Wissens verkündet werdet. Nur so ist zu vermeiden, dass die niedere Fakultät mit den höheren Fakultäten in Konflikt gerät und darüber einen Eingriff der Obrigkeit in ihre Autonomie provoziert. Am schwersten fällt Kant diese Abgrenzung bei der Medizin. Denn die Medizin gehört traditionell zu den höheren Fakultäten. Gleichzeitig beruhen aber ihre Methoden zum erheblichen Teil auf empirischer Naturbeobachtung der niedrigen philosophischen Fakultät und nicht nur auf obrigkeitlicher Offenbarung. Hier sucht Kant einen Formelkompromiss. Er behauptet, die von Gott begnadete Obrigkeit habe nur ein Interesse an der medizinischen „Polizey" („Medicinalordnung"). Details, mit der sich die Naturbeobachtung der philosophischen Fakultät befasse, kümmerten die Obrigkeit angeblich gar nicht. (Heute fühlt sich die Medizinische Fakultät in der „Missheirat" im Bett mit der alten niederen, der empirisch orientierten Fakultät wohl, zumindest am Platze.)

Mit dieser Aufteilung von göttlich-theoretischem und menschlich-empirischem Wissen hatte sich im 13. Jahrhundert schon der Theologe Albertus Magnus abgemüht, der danach 700 Jahre auf seine Heiligsprechung warten musste. Er kritisierte scharf, wenn in Vegetabilibus (Pflegewissenschaft der Landwirtschaft, also agrikulturelles Pflanzenwissen) theoretische Offenbarung der Heiligen Schriften gelten solle statt die menschliche empirische Evidence des Experiments:

„Ein Grundsatz, der mit der experimentellen Sinneswahrnehmung nicht übereinstimmt, ist in Wirklichkeit kein Grundsatz, sondern ein fundamentaler Fehler"

(Magnus, 1981, S. 144).

Besser kann man auch heute Methoden externer Evidence nicht begründen.

Heute sieht die bescheiden gewordene historische Anthropologie in jeder Empirie die Theorie. Die Theorie ist nicht mehr göttliches Offenbarungswissen. Theorie ist menschlich. Akademische Weihen sind für sie keine Priesterweihen mehr.

7. Fazit und Prognose

Den Beiträgen von Pfaff und Schmitt kommt das Verdienst zu, Errungenschaften, die die Selbstreflexion der evidence-based Praktizierenden vor einem Vierteljahrhundert, am Ende des letzten Jahrtausends erreicht und publiziert hatte, ausschnittsweise zu übernehmen. Dazu zählt die programmatische Abkehr von einem mechanistische Zugang und die Zuwendung zu einem systemtheoretischen, der die Welt nicht mehr nach dem Modell einer Uhr analysiert, deren Räderwerk dem Uhrmacher völlig offenliegt – wie einem allwissenden „Uhrmacher"-Gott bei Leibniz und Bourdieu. Auch ei-

nige wenige der zahlreichen Erkenntnisprobleme evidencegestützten Handelns, die Ende des letzten Jahrtausends theoretisch zu bewältigen waren, sprechen sie wieder an. Das ist zweifellos ein großes Verdienst. Ihr Lösungsvorschlag, die empirische Evidence durch eine theoretische „Justification" per „generally accepted theory" und weiter durch Experience und context-based knowledge zu einem EbM+Theory-Approach zu ergänzen, beruht allerdings auf 6 Annahmen, die leider für Menschen nicht realisierbar sind und daher mit einer anthropologisch begründbaren, also bescheiden gewordenen Erkenntnistheorie nicht vereinbar sind. Wie die Ausführungen von Pfaff und Schmitt belegten, führte ihre Unterscheidung von „theory-based Knowledge" und „evidence-based Knowledge" (Pfaff und Schmitt 2021, S. 6) bzw. sogar von „empirical and theoretical evidence" (Pfaff und Schmitt 2023, S. 1) dazu, extrem kritiklos eine seit 30 Jahren überwundene Version von „EbM" zu tradieren. Pfaff und Schmitt entlasten „EbM" völlig von der seit 30 Jahren herausfordernden Pflicht, bei jeder Studie deren implizite Theorien zu explizieren, zu analysieren und die meist nicht durch eine einzige Methode zu bewältigenden Verzerrungs-Risiken zu priorisieren: Pfaff und Schmitt stellen einfach eine davon unabhängige „theoretical evidence" daneben. Das Phänomen kennt man vom Hausbau. Anstatt ein altes Haus instand zu setzen und zu renovieren, wird einfach ein neuer, etwas wackeliger Anbau angefügt. Nach einiger Zeit fallen beide zusammen, das Haus und der Anbau. Zweifellos hätte eine Studie, die behauptet, theoriefrei empirische evidente Beiträge leisten und generalisieren zu können, mit Evidencebasierung nichts zu tun. Theoriefreie Wahrnehmung ist Menschen – im Unterschied zum allwissenden Gott – nicht möglich – auch wenn sich nicht wenige Menschen darauf berufen, durch eine persönliche Procura des allwissenden Gottes als Prokuristen Gottes an seiner Allwissenheit teilzuhaben. Die Bewegungen der Evidencebasierten Praxis entfalteten sich als theorieorientierte Bewältigungsversuche allfälliger Risiken verzerrter Wahrnehmungen und theoretischer Fehlschlüsse.

Prognose zur Ewigkeit und zu Risiken und Nebenwirkungen des „Ebm+Theory Approach"

Eine Prognose ist nicht schwer: Die Grenzen der Evidence-Basierung sind eine solche Zumutung, dass Menschen immer wieder ihre Hoffnung auf Theorie-Eminenzen setzen werden, die ihnen unter Berufung auf „Theorie" – oder auf ihnen exklusiv offenbarten Anteilen am Wissen des übermenschlich allmächtigen Gottes – die Begrenztheit menschlichen Wissens zu reduzieren versprechen. Solange es Menschen gibt, werden Eminenzen wohl nicht vergehen. Einige werden vielleicht in den Ruhestand gehen, andere vertrieben werden, aber immer wieder neue Eminenzen werden entstehen. Die Begrenzung menschlicher Erkenntnis ist einfach zu schwer auszuhalten. Sie zwingt zu existenziellen Entscheidungen unter Ungewissheit, Entscheidungen, die bei aller Ungewissheit als „vernünftig" den Mitmenschen zu begründen und zu verantworten sind. Diese Last an scheinbar eminente Experten abgeben zu können, wird immer eine höchst attraktive Versuchung bleiben. Was ist so schlimm an dieser so verständlichen Coping-Strategy? Dass die Berufung auf Theorie als Ergänzung und Relativierung von Evidence die Lernfähigkeit menschlicher Gesellschaften behindert und menschliches Leid verlängert, weil ihre zu naive Vorstellung von evidence-based knowledge den notwendigen Aufbau interner Evidence für überflüssig ansieht und sich stattdessen Eminenzen der „general-accepted theory" ausliefert.

Gerade am Umgang mit der Covid-19-Pandemie lässt sich die Differenz von externer Evidence und Entscheidung erkennen, die auf Pressekonferenzen oft legitimatorisch verdeckt wurde. Neben einem zuständigen, parlamentarisch gewählten Minister oder Amtsleiter stand meist schweigend eine Wissenschaftlerin oder

Literatur

Behrens J 2019. Theorie der Pflege und der Therapie. Bern usw.: Hogrefe

Behrens J & Langer G 2004. Evidence based Nursing. Vertrauensbildende Entzauberung der „Wissenschaft" Bern: Huber

Behrens J & Langer G 2022. Evidence based Nursing and Caring. Methoden und Ethik der Pflegepraxis und Versorgungsforschung V, Göttingen usw.: Hogrefe

Carnap R 1966, An introduction to the philosophy of science. New York: Basis Books

Cronbach L J 1982, Designing Evaluations of Educational and Social Programs. San Francisco: University Press

Ginzburg C 1983, Spurensicherung, Berlin: Wagenbach

Gross R & Löffler M, 1997. Prinzipien der Medizin: Eine Übersicht ihrer Grundlagen und Methoden, Berlin: Springer

Hume J [1748] 1993, Untersuchung über den menschlichen Verstand, Hamburg: Meiner

Kant I 1995, Der Streit der Facultäten, Köln: Könemann. Klemperer D unter Mitarbeit von Kuhn J und Robra B-P, www.corona-verstehen.de, Version 78 vom 9.3.2024

Locke J 1670, An Essay Concerning Human Understanding, London: T. Basset

Luhmann N 1975, Soziologische Aufklärung 2. Aufsätze zur Theorie der Gesellschaft, Opladen: Westdeutscher Verlag

Luhmann N 1981, Soziologische Aufklärung 3. Soziales System, Gesellschaft, Organisation, Opladen: Westdeutscher Verlag

Luhmann N 1993, Soziologische Aufklärung 5: Konstruktivistische Perspektiven, Opladen: Westdeutscher Verlag

Luhmann N 2019 -2022, Schriften zur Organisation. Herausgegeben von Ernst Lukas und Veronika Tacke, Bände 1–5, Wiesbaden: Springer VS

Magnus A 1981. Zum 700. Todestag [Auszüge der Schriften De Vegetabilibus, Ethica u. a.]. (Ausstellungskatalog des Historischen Archivs der Stadt Köln)

Peirce C S 2009, The logic of interdisciplinarity, Berlin: Akademie-Verlag

Pfaff H, Schmitt J 2023, Reducing uncertainty in evidence-based health policy by integrating empirical and theoretical evidence: an EbM+theory approach. J Eval Clin Pract. 2023;1-15. doi:10.1111/jep.13890

Pfaff H and Schmitt J (2021) The Organic Turn: Coping With Pandemic and Non-pandemic Challenges by Integrating Evidence-, Theory-, Experience- and Context-Based Knowledge in Advising Health Policy. Front. Public Health 9:727427. doi: 10.3389/fpubh.2021.727427.

Raspe H 2015, Standpunkt: „Der Nächste bitte". Anmerkungen zur Indikationsstellung in der klinischen Medizin. Z Evid Fortbild Qual Gesundhwes 109(1), 46–50

Singer W (2017) Zur Hirnforschung und Willensfreiheit. Frankfurter Rundschau 73(159), S. 27

Them, C.; Behrens J, 2003 Comparison of frontal teaching vs. problem-oriented learning at the school of healthcare and nursing: nursing neurological patients in: International Journal of Medical Informatics; Reference: IJB1911, 46–50.

ein Wissenschaftler, als seien externe Evidence und Entscheidung eins. Akademien empfahlen, was Politik dringend zu machen habe. Die Differenzierung zwischen (begrenzter) empirischer externer Evidence und parlamentarisch zu treffender Entscheidung schien eingeebnet, der Aufbau interner Evidence nicht mehr nötig. Bekanntlich war die verfügbare externe Evidence zu Beginn der Covid-19-Pandemie sehr begrenzt. Entscheidungen, die sich nachträglich als falsch herausstellen würden, waren weder zu vermeiden noch vorzuwerfen. Vorzuwerfen, weil leicht zu vermeiden, war das Versäumnis, von Anfang an methodisch gute Studien außer zur Wirkung auch zu „Risiken und Nebenwirkungen" der Entscheidungen zu unternehmen, z. B. bei stationär Pflegebedürftigen und jüngeren Schülern, aber auch vielen anderen Gruppen. Die externe und interne Evidence hätte sich schneller vermehren und damit Entscheidungen eine bessere Grundlage geben können. Für die nächste Pandemie, die sicher kommen wird, wäre einiges mehr an externer Evidence bereitgestellt worden. Positiv ist auf das erst alle 14 Tage, dann monatlich aktualisierte (inzwischen 78. Version) Buch von David Klemperer unter Mitarbeit von Joseph Kuhn und Bernt-Peter Robra hinzuweisen „www.corona-verstehen", das den jeweils neuesten Stand des Wissens zusammenfasste und sicher zu ergänzen ist.

Es war in den letzten 30 Jahren, seit der Bewegung der antiautoritären Studierenden der McMaster University, immer zu erwarten: Die Eminenzen schlagen zurück. Sie werden mit utopischen Hoffnungen empfangen werden und gefeiert. So gefeiert, werden nur wenige Eminenzen zur – in der Tat langwierigen und wenig glanzvollen – Arbeit zurückkehren, die Theorien und Methoden in den EbP-Studien externer Evidence kritisch zu analysieren und individuellen Klienten bei deren Aufbau interner Evidence zur Verfügung zu stehen. <<

The EbM+Theory Approach – an appreciation of its 6 assumptions and its risks and side effects

In their "EbM+Theory approach" (2021 and 2023), Holger Pfaff and Jochen Schmitt address questions of evidence-based decisions that evidence-based health services research worked on 30 years ago. Their EbM+Theory approach is based on at least six assumptions, the practical-empirical validity of which is examined. A conclusion deals with the risks and side effects of the EbM+Theory approach and predicts eternal survival for this approach, since it supports Eminence.

Keywords

Philosophy of science, anthropologically modest epistemology, evidence-based health services research, EbM+Theory approach, Systems Theory

Autorenerklärung

Der Autor erklärt, dass keine finanziellen Interessenkonflikte vorliegen.

Zitationshinweis

Behrens, J.: „Der EbM+Theory Ansatz – eine Würdigung seiner 6 Annahmen und seiner Risiken und Nebenwirkungen", in: „Monitor Versorgungsforschung" (02/25), S. 79–88. http://doi.org/10.24945/MVF.02.25.1866-0533.2710

Prof. em. Dr. phil. habil. Johann Behrens
war Gründungsdirektor des ersten Instituts für Gesundheits-, Therapie- und Pflegewissenschaften an einer öffentlichen Medizinischen Fakultät in den deutschsprachigen Ländern (D-A-CH), der Universität Halle-Wittenberg. Er war Vorstandsmitglied der auch in der Versorgungsforschung aktiven DFG-Sonderforschungsbereiche 186 (Bremen 1988–2000) und 580 (Halle/Jena 2000–2012) sowie der Dekanekonferenz Pflegewissenschaft. Zudem war er Gründungsmitglied des German Centers for Evidence based Nursing „sapere aude". Kontakt: Johann.Behrens@medizin.uni-halle.de

