

# **Kriterien-Kriterienpunkte-Indikatoren**

– ihr Stellenwert im  
inklusiv-nutzungsfokussierten Forschungsmodus –  
entschlüsselt aus evaluationswissenschaftlicher Sicht

Workshop: Ansätze für die Nachhaltigkeitsbewertung <sup>durch und</sup> von Forschung

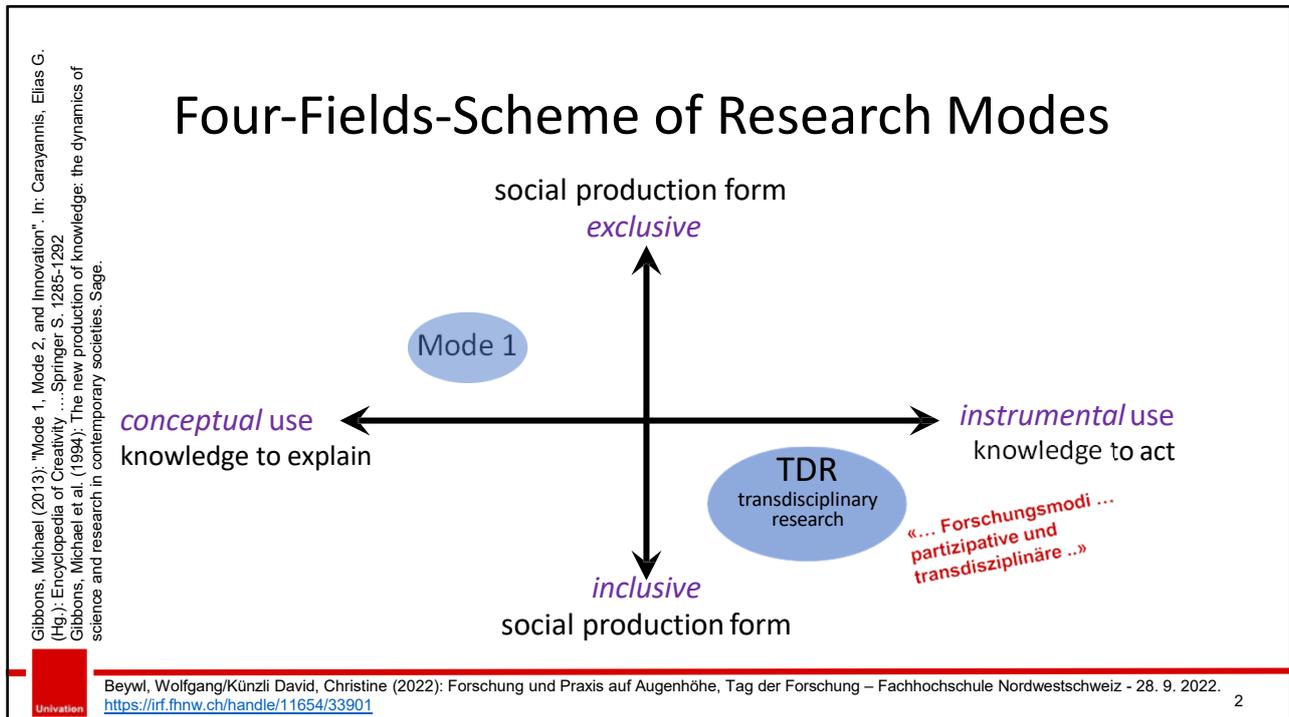
11.11.2022 - online



Der Beitrag behandelt Kriterien und Indikatoren. Inspiriert ist er durch den Ausschreibungstext zu diesem Workshop. Sie sehen hier den Workshop-Titel in **roter** Schriftfarbe nach links gekippt. Es folgen weitere solcher Zitate aus der **Ausschreibung**.

Hintergrund der Argumentation sind evaluationstheoretische Überlegungen. Eine Ausgangsthese ist, dass hinter den Zielkonflikten oft auch Wertkonflikte liegen. Diese materialisieren sich in Kriterien, die miteinander konkurrieren können. Ziele sind eine wichtige Quelle für Kriterien. Es gibt aber auch andere (vgl. Folie 10).

Mit dem ergänzten »durch und« möchte ich auf zwei Sichtweisen hinweisen: **einerseits**, dass die Bewertung von Nachhaltigkeit zum Beispiel politischer Maßnahmen **durch** Forschung – aktiv – erfolgt, dass sie **andererseits selbst** – passiv - bewertet wird, also Bewertung **von** Forschung, u.a. bezüglich ihrer Nachhaltigkeit. **Beide** Linien werden in meinem Input angesprochen. Sie sind in der Transdisziplinären Forschung komplex verwickelt. Dies soll ein Stück entwirrt werden.



Die Ausschreibung zum Workshop fokussiert auf «partizipative und transdisziplinäre» Forschungsmodi. Dies erinnert an die Unterscheidung zwischen Mode1 und Mode-2 von Gibbons und anderen. Beide Modi sind hier in einem Viererfeld verortet:

- In der **Horizontalen** steht links der **konzeptionelle** Nutzen. Die Grundlagenforschung arbeitet darauf hin, möglichst allgemeingültiges Erklärungswissen zu generieren. Demgegenüber steht der **instrumentelle** Nutzen, der sich im enger gekoppelten forschungsinformierten Handeln von Politik und Gesellschaft realisiert.
- In der **Vertikalen** finden Sie zwei unterschiedliche soziale Produktionsweisen gegenübergestellt: oben die **exklusive**, in der Forschende Einflussnahmen Dritter auf Fragestellungen und Forschungsprozess ausschließen; unten die **inklusive**, in der sie gesellschaftliche Stakeholder in solche Entscheidungen paritätisch einbeziehen.

**Heute** geht es um die die Mode-2 oder »transdisziplinäre Forschung«.

## Definition of evaluation (abbreviated)

- scientific endeavour and professional service, which describes and evaluates *evaluands*;
- guided by purposes and evaluation questions, clarified collaboratively by *stakeholders* ...
- ... to obtain reliable information for the descriptive task, evaluation uses empirical methods;
- for the *valuation* task – a uniqueness – it uses clarified *values* rooted *criteria* (and *threshold points*) and transparent methodical valuation procedures



Nun zur **Evaluation**. Dies ist die Wissenschaft, die sich mit systematischer Bewertung beschäftigt. Hier einige Definitionsbestandteile:

- es geht um das Beschreiben und Bewerten von Evaluationsgegenständen
- mit ausgewiesenen Verwendungszwecken und Evaluationsfragestellungen, abgestimmt mit Stakeholdern
- um »**beschreibend**« Informationen zu gewinnen, werden empirische Methoden eingesetzt
- um systematisch zu **bewerten**, müssen werte- bzw. interessenbasierte Kriterien ausgewiesen und transparente Bewertungsverfahren eingesetzt werden.

Sie sehen: auch Evaluation ist transdisziplinär.

## Kriterium = wertegeladen

Im Wortstamm bezogen auf altgriechisch *κρίτης* (der Richter). Kriterium (altgr. *κρίτηριον*) „meint Beurteilungsgesichtspunkt für den Richter“ (Kluge, 488).

- *Kriterien* basieren auf *Werten*:  
differierende Werte → verschiedene Kriterien oder Gewichtungen von Kriterien
- werden in der Evaluation genutzt, um darauf gestützt eine *Bewertung* zum *Evaluationsgegenstand* (*evaluand*) abzugeben (->*Bewertungssynthese*);



Um die Bedeutung von **Kriterium** zu erfassen, hilft die Etymologie weiter. Ursprünglich «Beurteilungsgesichtspunkt für den Richter». Kriterien basieren auf sozialen Werten. Wie diese konkurrieren sie miteinander. Ihnen wird unterschiedliches Gewicht zugemessen. Es ist Aufgabe der Evaluation, dies zu explizieren. Kriterien sind unverzichtbar für den Akt der Bewertung. Dafür gibt es elaborierte Verfahren bis hin zur abschließenden **Bewertungssynthese**, auf die ich heute nicht eingehe. Wichtig: Kriterien sind wertegeladen.

## Kriteriendimension

Ein *Evaluationsgegenstand* kann und muss i. d. R. bezüglich mehrerer Gesichtspunkte bewertet werden. Auf diese *Kriteriendimensionen* gründen Bewertende ihre Urteile. Die Dimensionen variieren zwischen *Stakeholdern* nach Priorität/Gewichtung.

Beispiel Auto	Beispiel Weiterbildung
Breite Einsatzmöglichkeiten	Bedarfspassung
Technologiestand	Wissenschaftliche Aktualität
Sicherheit, aktiv/passiv	Schutz persönlicher Rechte
Design	Ästhetik
...	...
Leistungsstärke	Trainingsnutzen
Wertstabilität	TransfERNutzen
Umweltverträglichkeit	Bildungswert
Wirtschaftlichkeit	Wirtschaftlichkeit

«Nachhaltigkeit»

besteht aus mehreren, teils konfligierenden Kriteriendimensionen - die sich je nach Kontext, Stakeholdern usw. verschieden mischen

Kriterien können sehr kleinteilig formuliert werden. Solange sie jedoch eine Art Oberbegriff darstellen, bezeichnen wir sie als »**Kriteriendimensionen**«. Beispiele sehen Sie hier für die Evaluationsgegenstände »Auto« und »Weiterbildung«. «**Nachhaltigkeit**» steht logisch noch darüber: Sie besteht aus mehreren Kriteriendimensionen. Wenn Sie die Ziele für nachhaltige Entwicklung ansehen, so haben sie dort mindestens 17 recht umfassende Kriteriendimensionen.

## Indikator = neutral & messbefähigend

*Indikatoren* von lateinisch INDICARE (auf etwas zeigen/etwas zeigen) sind ‚Anzeiger‘. Sie sind - *grundsätzlich unvollständige* - Anhaltspunkte für das Vorhandensein eines Sachverhaltes, dessen Vorliegen nicht unmittelbar beobachtbar ist.

Indikatoren sind unverzichtbar für das Messen komplexer Sachverhalte, wie sie typisch sind in Evaluation und TDR.

- Oft sind mehrere Indikatoren nötig, die für einen Sachverhalt ‚einspringen‘ (Indikatoren-Triangulation)
- Oft ist es nötig, Indikatoren für ein und den selben Sachverhalt mittels zweier/mehrerer Methoden/Instrumente zu ‚beobachten‘ (Triangulation der Datenerhebungsmethoden/-instrumente).



Nun folgt ein gewaltiger Sprung: Um empirisch basierte Aussagen über den Grad der Erfüllung gemäss Kriterien machen zu können, müssen diese operationalisiert werden. Dies geschieht mithilfe von **Indikatoren**. Auch hier hilft die Etymologie: Indikatoren sind Anzeiger, um etwas den Sinnen zugänglich zu machen, was nicht unmittelbar erfassbar ist. Z.B. »Armut« als SDG1. Das erforderliche Operationalisieren ist ein in den Natur- wie in den Sozialwissenschaften verbreitetes Verfahren. Unten rechts sehen Sie ein Indikatorstäbchen mit einer Farbskala. Indikatoren sind also messbefähigend und – und das verblüfft vielleicht manche von Ihnen – **neutral**; im Gegensatz zum wertegeladenen Kriterium.

## Kriterienpunkt(e)\* = wertgeladen & messbefähigend

(auch Schwellenwert, Zielwert, Zielgrösse ....)

= Referenz für die *Bewertung*.

### A auf nach oben offener Erfolgsskala

Negative Bewertung



### B Erfolgsspanne begrenzend

Negative Bewertung



\* Diese und andere Fachbegriffe der Evaluation aus dem kontrollierten Online7-Vokabular: [https://eval-wiki.org/glossar/Eval-Wiki:Glossar\\_der\\_Evaluation](https://eval-wiki.org/glossar/Eval-Wiki:Glossar_der_Evaluation) 7 sind in dieser Präsentation durch *Kursivsetzung* gekennzeichnet.

Um eine Bewertung vornehmen zu können müssen **Kriterienpunkte** festgelegt werden. Beim Indikatorstäbchen geschieht dies durch Deklaration eines oder mehrerer benachbarter Farbfelder als wünschenswert. Genauso werden in der Evaluation – auch hier selbstverständlich **vor** der Datenauswertung – Kriterienpunkte definiert. Diese können sich entweder auf einer nach oben offenen Erfolgsskala befinden oder eine Erfolgsspanne links und rechts begrenzen.

Mir liegt nicht daran, dass sie jetzt alles im Detail erfassen. Sie können es ja nachlesen. Wichtig ist mir, dass Sie die **Logik** von systematischen Bewertungsprozeduren, wie in der Evaluationswissenschaft ausgearbeitet sind, in Grundlinien nachvollziehen können.

## TDR as (I.) evaluation and (II.) evaluand

- I. TDR evaluates (political) programs, services, interventions, products, institutions etc. TDR could adapt evaluation methodologies and tools to do it highly systematically on the background of a broad set of scholarly literature
  - TDR inevitably includes valuation procedures  
*(first order valuation)*
- II. TDR as an evaluative mode of research can be evaluated. The procedures and criteria dimensions developed by evaluation science to evaluate itself to be adapted
  - meta-evaluation as a sound approach to valuate TDR-processes, products etc.  
*(second order valuation)*

«was sie betrachten»

«Beiträge von Forschung»

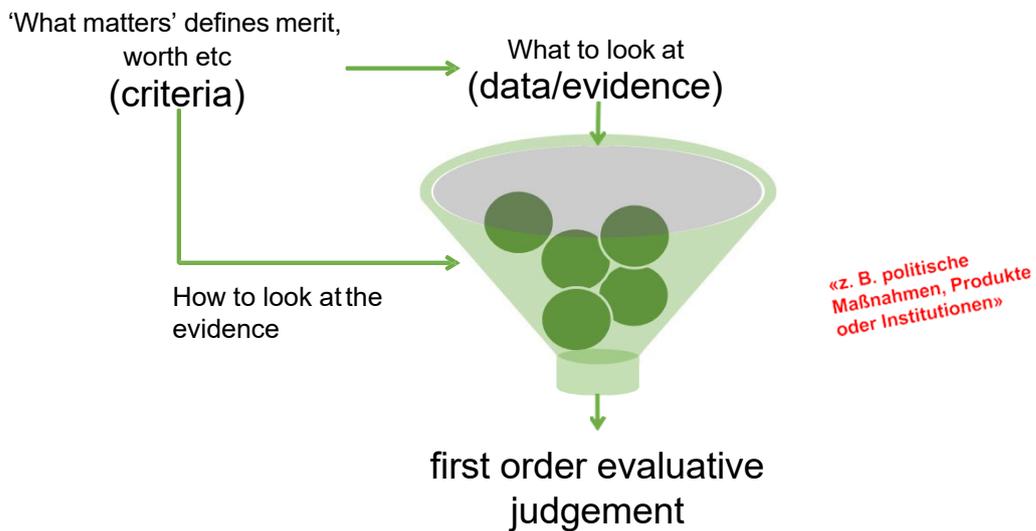


Zurück zum erweiterten Titel dieses Workshops:  
Nachhaltigkeitsbewertung durch /// und von Forschung:

**Fall I** ist, dass transdisziplinäre Forschung politische Programme usw. bewertet, also als »Evaluation« fungiert. In der Ausschreibung finden Sie diese Perspektive I. mit »was sie betrachten«. Es liegt nahe, dass transdisziplinäre Forschung hierfür Evaluationswissen nutzt.

**Fall II** ist, dass die »Beiträge der Forschung« für die Gesellschaft bewertet werden. Hier ist der Evaluationsgegenstand die transdisziplinäre Forschung selbst.

# I. How TDR evaluates its evaluands – procedure



Sie sehen hier, wie im Fall einer evaluativen **Bewertung I. Ordnung** Kriterien mit Daten kombiniert werden, also systematisch evaluiert wird.

# I. How TDR identifies criteria

## Pre-ordinate criteria

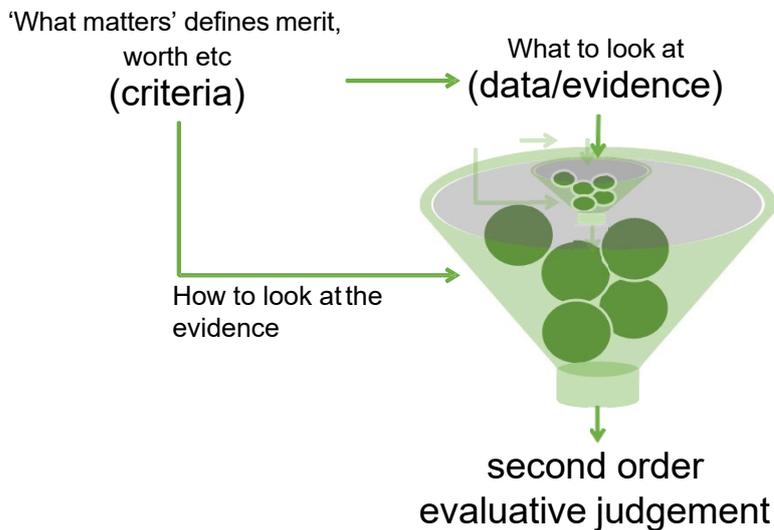
values as substantial for the mindset of humankind

		Normative perspectives				
		Consequentialist	Deontological		Virtue-based	
<b>Responsive criteria</b> Determining what's important for each stakeholder group	Respect for:	Consequences <i>Maximum benefits</i>	Duty <i>Meeting obligations</i>	Rights <i>Minimum protection</i>	Equity <i>Being fair or equitable</i>	Ethic of care <i>Relationships &amp; context</i>
	Interest group 1					
	Interest group 2		Dimensions of value			
	Interest group 3					

nach Roorda, Mathea Bendino Shulamith (2020): Developing defensible criteria for public sector evaluations. Melbourne: The University of Melbourne.  
<https://minerva-access.unimelb.edu.au/handle/11343/239232> (ihr missverständliches "prescriptive" durch "pre-ordinate", "descriptive" durch "responsive" ersetzt 10

Ich möchte Ihnen ein Finde-Schema für Bewertungskriterien zeigen. Es ist durch verschiedene philosophische Traditionen und partizipative sozialwissenschaftliche Ansätze begründet. Bei Mathea Roorda finden Sie auch eine erste Version eines Leitfadens.

## II. TDR as evaluand of meta-evaluation



Wenn es um **Bewertung II. Ordnung** geht, also die Frage, wie nachhaltig transdisziplinäre **Forschung** ist, dann wird es komplex. Dies ist hier mit «kleiner Trichter (Evaluation I. Ordnung) im großen Trichter» (Evaluation II. Ordnung) angedeutet.

## II. Criteria Dimensions for TDR as evaluand

### 4 + 1 for transdisciplinary research

**W - Wealth of Insights** – for the scientific community, for progress of disciplines, conceptual knowledge related to theories ... (Erkenntnisreichtum)

**U - Utility** – relevant & timely information for society/ stakeholders and concern for consequences and influence ... (*Nützlichkeit*)

**P - Propriety** -values to be acknowledged & fairly prioritized, human rights, transparency, inclusiveness, fiscal responsibility ... (*Fairness*)

**A - Accuracy** – concerning analysis of context, validity of data, sound procedures for interpretations/conclusions ... (*Genauigkeit*)

**F - Feasibility** – conclusive project management, resource-efficient, contextual viability .... (*Durchführbarkeit*)



Hierfür gibt es aus der **Meta-Evaluation**, also der Evaluation von Evaluationen, vier bewährte Kriteriendimensionen: Nützlichkeit, Fairness, Genauigkeit und Durchführbarkeit. Da es sich in Ihren Fällen oft um transdisziplinäre **Forschung** handelt, wäre als fünfte Kriteriendimensionen die des **Erkenntnisreichtums** hinzuzufügen.

# Zusammenfassung

TDR (wie auch Evaluation) ist inklusiv angelegt und zielt auf instrumentellen Nutzen für die Gesellschaft.

Evaluationswissenschaft hat ein theoretisch-begriffliches System für systematisches, empirisch basiertes Bewerten entwickelt.

I. TDR, durch die wie durch Evaluation Programme etc. bewertet werden, kann sich dieses Systems bedienen. Dies betrifft auch die Identifikation von responsiven und im Anthropozän hoch relevanten -vorausgesetzten (pre-ordinate) Kriterien.

II. Die Evaluation von TDR kann auf das bewährte Vierer-Set der Kriteriendimensionen für die *Meta-Evaluation* zurückgreifen, ergänzt um «Erkenntnisreichtum» als relevant für Forschung.



Abschließend nochmals der Argumentationsgang meines Inputs. Ich bin überzeugt: Wenn Sie sich mit den Grundlagen evaluationswissenschaftlicher Bewertung befassen, werden Sie die eine oder andere Frage, die sich jetzt noch stellen, produktiv auflösen.