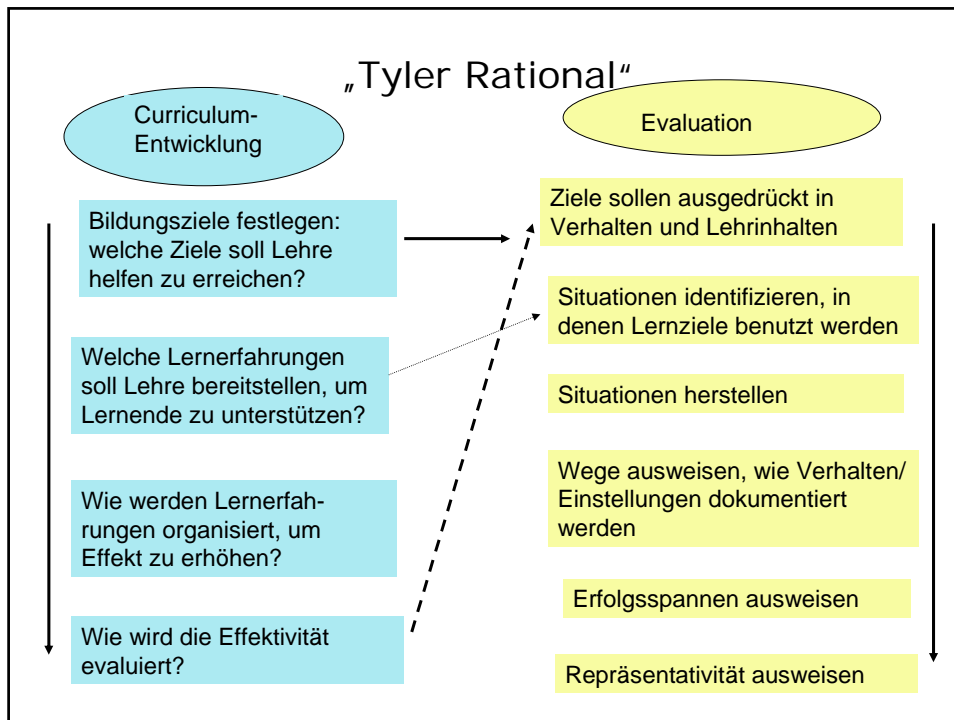


Evaluationsforschung

1.12.2005

Plan für heute

- Ziel basierte Evaluation: Reste
- Experimentelle und quasiexperimentelle Designs
- Deren Stärken und Schwächen
- Deren Nutzen für die bisher vorgestellten Projekte



Die Stärken

- Ziele liefern Bewertungsmaßstab
- Ziele können gleichzeitig für Programm- und Evaluationssteuerung genutzt werden
- hohe Praxistauglichkeit

Die Schwächen

- Wie ist entscheidbar, dass „richtige“ Ziele festgelegt werden?
- Ziele haben viele Funktionen
- Ziele sind veränderlich
- Wie ist feststellbar, dass das Erreichen der Ziele auch Aufschluss darüber gibt, dass das Programm für die Zielerreichung verantwortlich ist?

Experimente und Quasi-Experimente – typische Fragestellungen

- Hat tatsächlich das Programm die Effekte/Veränderungen ausgelöst?
- Welchen Unterschied macht das Programm?
- Welche **Wirkung** hat das Programm?
 - Programm als Experiment
 - Grundprinzip: Vergleichen

Grundbestandteile des Experiments

X	„Treatment“, Maßnahme Intervention	verstanden als Ursache bzw. unabhängige Variable
M (O)	Messung von Beobachtungsmerkmal en, die durch X verändert werden sollen (Observation)	Verstanden als Wirkung bzw. als abhängige Variable

Prä-experimentelle Designs

X	M	„one-shot case study design“	
M ₁	X	M ₂	„one group pretest- posttest-Design“
X	M ₁	----- M ₂	Statischer Gruppenvergleich (Vergleichsgruppen sind nicht äquivalent)

Einwände gegen Validität

- „history“ und „maturation“: Veränderungen sind erklärbar durch andere Ursachen als X, z.B. andere externe Einflüsse, Reifungsprozesse
- Selektion: unterschiedliche Ausgangsbedingungen zwischen Experimental- und Kontrollgruppe
- Regressionseffekte
- Ausfallquoten
- Veränderungen im Messinstrument
- Testeffekte

„Salomon Four Design“ – als „echtes“ Experiment

- Hinzu kommt **R**: Randomisierung, d.h. die zufallsgesteuerte Zuweisung zu Experimental- und Kontrollgruppe

G1: R	M ₁	X	M ₂
G2: R	M ₃		M ₄
G3: R		X	M ₅
G4: R			M ₆

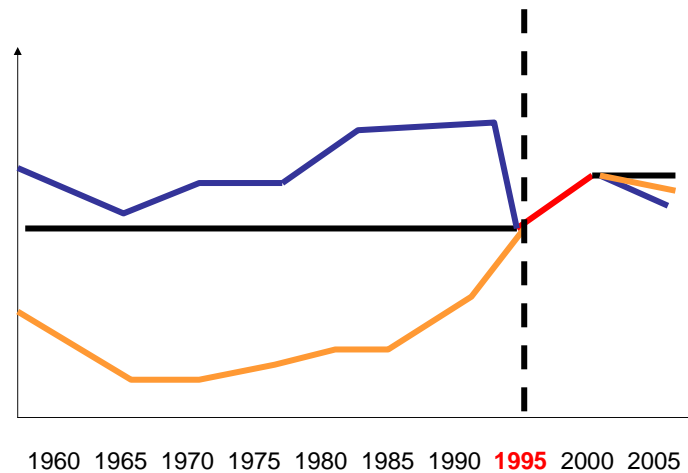
Probleme der Randomisierung

- Ethische
- Praktische:
 - „Flächendeckende“ Programme
 - Kontroll- und Experimentalgruppe haben Kontakte untereinander
- Randomisierung kann in Konflikt zum Programm stehen

Alternativen durch Quasi-experimentelle Designs

- Wenn nicht zwischen randomisierten Kontroll- und Experimentalgruppen verglichen werden kann, müssen alternative Vergleichsmöglichkeiten geschaffen werden
- Z.B. Zeitreihenanalyse:
M₁ M₂ M₃ X M₄ M₅ M₆
- „interrupted time-series design“
M₁ X M₂ ... M₃ ... M₄ X M₅ ... M₆

Absolventenquoten: Studienreform 1995



Stärken der (quasi-) experimentellen Designs

- Gelten als der „Königsweg“
- Hohe Akzeptanz in der Wissenschaft (interne Validität)
- Eindeutige Zuweisung von Effekten
- Nähe zu Strategien aus dem Alltag

Schwächen (quasi-) experimenteller Designs

- „Black box evaluation“
- Keine Hinweise dazu, wie Treatment- und Beobachtungsvariablen bestimmt und operationalisiert werden können.
- Kontrolle von allen möglichen Einflüssen steht im Widerspruch zu Programmen, die mit Menschen arbeiten

Anwendungsmöglichkeiten?

- Beispiel: Pflege
- Beispiel: Krankenhaus
- Beispiel: Equal
- Beispiel: mediengestützte Lernumgebungen
- Sonstige ...

Zum nächsten Mal lesen:

- Rogers, Patricia J., 2000: Program Theory. Not Wether Programs Work, But How They Work. In: Stufflebeam et al., Evaluation Models, S. 209-232.