

## **Transdisziplinäres Kolloquium „Zwischen Objektivität und Normativität: Wie kommt die Evaluation zu(m) Werten?“**

**Gießen, Freitag, 27.02.2015**

### **ERÖFFNUNG UND EINSTIEGSREFERATE**

Jan Hense (Universität Gießen) begrüßt die Teilnehmenden und geht in seinem einführenden Vortrag auf die Bedeutung von Werten in der Evaluation sowie verschiedene Problematiken im Zusammenhang mit dem Bewerten ein. Obwohl Werte und das Werten zumindest dem Wortsinn nach elementare Bestandteile der Evaluation sein müssten, zeigt er anhand von Beispielen aus konkreten Evaluationsberichten auf, dass Evaluationen sich oft auf deskriptive Befunde beschränken, den Bezugsrahmen für Bewertungen im Dunkeln lassen oder die eigentliche Bewertung an Dritte delegieren. Im Ausblick auf die Tagung wirft er die Frage auf, ob dies wirklich Defizite sind, oder legitime Praxis in bestimmten Evaluationskontexten.

#### Einstiegsreferat 1:

##### **Die Bedeutung der Evaluationsfragestellung für das Bewerten**

(Verena Friedrich, Universität Bern)

Verena Friedrich legt in ihrem Beitrag dar, dass die Bewertung (was ist gut/schlecht?) ebenso wie die Operationalisierung bereits in der Evaluationsfragestellung angelegt sein sollte, da auf dieser Grundlage die Kriterien und Indikatoren entwickelt werden, welche für die Messung und Bewertung benötigt werden. Sie veranschaulicht dies anhand eines konkreten Beispiels einer Fragestellung zur Veränderung der Luftqualität in öffentlichen Räumen nach Einführung des Nichtrauchergesetzes.

#### Einstiegsreferat 2:

##### **Verwendung von Bewertungskriterien in der Praxis: Eine empirische Untersuchung**

(Christine Heuer, Bundesamt für Gesundheit, Bern)

Christine Heuer bezieht sich auf den Input ihrer Vorgängerin, wenn Sie aus Basis ihrer empirischen Masterarbeit feststellt, dass Bewertungen und Bewertungskriterien bei von Schweizer Bundesämtern beauftragten Evaluationen in den Berichten nur teilweise nachvollziehbar und transparent sind.

Es werden viele interessierte Nachfragen gestellt und im Anschluss rege diskutiert. Ein Fazit aus der Diskussion ist, dass sich das Thema Bewertungskriterien durch die Aus- und Weiterbildung in der Evaluation ziehen sollte.

## **FALLBERICHTE AUS DER BERUFLICHEN BILDUNG**

### Fallbericht 1:

#### **Die Vergleichsperspektive als Triebfeder einer Bewertungskultur**

(Miguel del Estal, InterVal GmbH, Berlin)

Miguel del Estal fordert in seinem Beitrag einen „Fahrplan“ zu einer Bewertungskultur. Das Potential von Evaluationen werde ohne s.m.a.r.t.e Ziele für den Binnenvergleich oder eine Referenzbasis für einen Außenvergleich aus seiner Sicht nicht ausgeschöpft.

In der Diskussion wird unter anderem in Frage gestellt, ob Bewertungskriterien in jedem Fall politisch machbar sind: Geben die Gesetze eine Grundlage dafür überhaupt her?

### Fallbericht 2:

#### **Das Qualitätssiegel des MedienCampus Bayern als Beispiel zur Evaluation von Bildungsangeboten**

(Veronika Alz, Medien-campus München)

Veronika Alz beschreibt in ihrem Beitrag konkrete Schwierigkeiten bei der Bewertung im Rahmen des Qualitätssiegels, welche beispielsweise durch die hohe Komplexität oder Heterogenität der Evaluationsgegenstände, in Ermangelung einer Referenzbasis für den Binnen- oder Außenvergleich o.ä. hervorgerufen werden, sowie den Umgang mit dieser Problematik beim MedienCampus Bayern.

## **PARALLELSESSION A: WER BEWERTET LETZTENDLICH? EVALUATION ODER REZIPIENTEN?**

Dörte Schott führt als Moderatorin die Positionen der beiden folgenden Beiträge ein, in denen es darum geht, wer (Evaluierende oder Rezipienten?) an welchen Stellen Bewertungen vornehmen oder Werte vorgeben. Während Felix Albrecht in seinem Beitrag aufzeigt, wie er die Rezipienten anstatt von Evaluierenden der Lehrveranstaltungsevaluation die Bewertung vornehmen lässt, legt Christian Diller dar, dass im Feld der regionalen Siedlungsentwicklung in den USA die Evaluation am Ende eine Bewertungslücke füllt, welche von der Politik offen gelassen wird.

### Beitrag 1:

#### **Evaluation einer Lehrveranstaltung als Bewertungsgrundlage für verschiedene Interessensgruppen**

(Felix Albrecht, Karlsruher Institut für Technologie)

Felix Albrecht stellt in seinem Vortrag die Evaluation einer Lehrveranstaltung am KIT durch, die zur Semestermitte hin getätigt wird. Ziel dieser frühen Lehrevaluation ist, noch während des Semesters gegensteuern zu können, falls sich Problemstellungen ergeben sollten. Als

nutzenorientierte Evaluation soll sie allen Beteiligten bzw. Interessenten eine verwertbare Rückmeldung zur Lehrveranstaltung geben können. Da Hinweise auf Verbesserungsmöglichkeiten in den klassischen Standardverfahren zur Lehrevaluation aus seiner Sicht fehlen, schlägt er einen qualitativen Lösungsansatz vor: Auf der Grundlage eines Leitfadens führen Hilfskräfte (keine in den Lehrbetrieb involvierten Personen) Gruppendiskussionen mit den Studierenden der jeweiligen Lehrveranstaltung zur Bewertung der Ergebnisse durch, welche anschließend inhaltsanalytisch ausgewertet werden. Hieraus ergeben sich ganz konkrete Verbesserungsideen, welche anschließend vom Dozenten zur Kenntnis genommen und umgesetzt werden können. Die Evaluation kann von den Dozenten selbst angeregt und selbstevaluativ durchgeführt werden.

#### Beitrag 2:

### **Bewertung in der Evaluation regionaler Siedlungsentwicklung: Ein Vergleich Deutschland- USA**

(Christian Diller, Universität Gießen)

Christian Diller gibt zunächst einen Ein- und Überblick über die Raum- und Siedlungsplanung in Deutschland und den USA. Sein Beitrag beschäftigt sich mit dem Zusammenhang zwischen Planungsdichte und Evaluationstätigkeit.

In Deutschland wird die Stadt- und Raumplanung aufgrund der dichten Besiedlung im Gegensatz zur den USA viel stärker gesteuert; gleichzeitig gibt es kaum Evaluationen dieser. Anders ist es in den USA, wo die Stadt- und Raumplanung teilweise kaum gesteuert wird, aber eine große Anzahl von Evaluationen zu diesem Thema vorliegt, welche meist von Hochschulen durchgeführt werden. Die Evaluationskriterien werden hier im Rahmen von und durch die Evaluierenden (unter Einbezug der betroffenen Bevölkerung) entwickelt.

Er berichtet über den Versuch, einen Satz von Standardkriterien für die Stadt- und Raumplanung in Deutschland zu entwerfen, welcher in der Fachcommunity als Qualitätsstandard diskutiert wird.

Den beiden Beiträgen schloss sich eine rege Frage- und Diskussionsrunde an, in welcher es abschließend um die Frage ging, ob es Aufgabe der Evaluation sei, Lücken in den Bewertungskriterien der Politik zu schließen.

### **PARALLELSESSION B: WESSEN WERTE? WERTE(N) IN SOZIALEN UND KOMPLEXEN KONTEXTEN**

#### Beitrag 1:

### **Der Einsatz der FIWE-Methode im Kontext der externen Schulevaluation**

(Dirk Groß, Agentur für Qualitätssicherung, Evaluation und Selbständigkeit von Schulen, Bad Kreuznach)

Dirk Groß stellte ein Beispiel externer Evaluation in staatlichen Schulen in Rheinland-Pfalz vor. Das hierbei zum Einsatz kommende „wertepriorisierenden Evaluationsmodell“ (Beywl, Speer &

Kehr 2004) wurde darauf geprüft, inwiefern die Werte der Beteiligten und Betroffenen im Sinne des DeGEval-Standards „N5 Transparenz von Werten“ sinnvoll identifiziert und thematisiert werden können. Exemplarisch wurde dies aufgezeigt am Einsatz der FIWE-Methode‘ (Fakten-Interpretation-Werte-Empfehlungen; Beywl & Schepp-Winter (2000)).

Die anschließende Diskussion griff die folgenden Fragestellungen auf: Wodurch ist ein Wert definiert? Was ist ein Wert, was nicht? Wie kann ein Wert in der Praxis insbesondere von einer Interpretation unterschieden werden? Wie kann FIWE in den Arbeitsalltag von Schulen integriert werden? Wie ist das Verhältnis zwischen Aufwand und Nutzen? Worin besteht der Mehrwert von FIWE jenseits des psychologischen Momentes? Wie nutzen die Schulen die Ergebnisse und Empfehlungen von FIWE?

#### Beitrag 2:

#### **Developmental Evaluation: Was bedeutet die Evaluation von Innovationen in komplexen, dynamischen Umfeldern für das Bewerten**

(Maria Gutknecht-Gmeiner, Impulse – Evaluation und Organisationsberatung, Wien)

Maria Gutknecht-Gmeiner ging auf Aspekte der „Developmental Evaluation“ (dt. „evolutive Evaluation“, Michael Patton) ein – ein spezifisches Vorgehen für die Evaluation von innovativen, sich ständig anpassenden Maßnahmen unter veränderlichen Rahmenbedingungen. Im Zentrum ging sie der Frage nach, welchen Einfluss ein „evolutives“, entwicklungsorientiertes Vorgehen in der Evaluation auf das (Be)Werten hat. Anhand des Projektes „Technik zu den Mädchen bringen“ stellte sie dar, in welchen Bereichen die Prämissen für das Bewerten und der Bewertungsvorgang in der Developmental Evaluation sich von gängigen Vorstellungen von (Be)Werten in Evaluationen unterscheiden. Dabei ging sie auf typische Aspekte der Developmental Evaluation ein, wie die Instabilität und der geringe Ausdifferenzierungsgrad des evaluierten Programms, das Ziel der Entwicklungsunterstützung, die direkte Handlungsrelevanz, die intensive Zusammenarbeit mit den Programmumsetzenden sowie insgesamt die enge Verflechtung von Evaluation und Intervention.

Die anschließende Diskussion griff die folgenden Fragestellungen auf: Inwiefern beeinflussen die Werte des Evaluators das Ergebnis der Evaluation? Inwiefern ist ein solcher Einfluss legitim? Wie verhält sich ein solcher Einfluss zur Neutralität des Evaluators? Welche Rolle hat der Evaluator im Prozess des Bewertens, welche sollte er innehaben? Wie verhalten sich in diesem Kontext Objektivität und Normativität zueinander?

In der *übergreifenden* Diskussion beider Beiträge wurden folgende Aspekte angesprochen:

Verantwortung der Evaluation: Neben der Produktion von Daten sollte Evaluation immer auch antizipieren, wie die Ergebnisse genutzt werden, verstärkte Bemühungen zur Herstellung der Transparenz von Bewertungsprozessen notwendig, Kreis der Bewertenden: sollte grundsätzlich ausgeweitet werden und unterschiedliche Akteure umfassen (z.B. Schulleitungen, Zielgruppen, Trägerverbände, Kommunen).

Aktivitäten der DeGEval: Standard „N5 Transparenz von Werten“ sollte gestärkt werden, verstärkte Ansprache neuer Zielgruppen notwendig, insbesondere von Auftraggebern sowie Beteiligten/Betroffenen von Projekten/Programmen.

## **KEYNOTE: WIE KOMMT DIE EVALUATION ZU(M) WERTEN?**

(Wolfgang Beywl, Pädagogische Hochschule FHNW Brugg/Univision Köln)

In seiner abschließenden Keynote geht Wolfgang Beywl noch einmal grundsätzlich auf die Frage des Bewertens ein. Er stellt die These auf, dass Evaluation werten muss – und zwar systematisch. Micheal Scriven zitierend gibt er zudem eine Antwort auf die Frage, wer bewerten muss: die Evaluierenden.

Grundsätzlich unterscheidet er drei unterschiedliche Bedeutungen des Wortes „Wert“: a) Güte, Tauglichkeit, Bedeutsamkeit (worth/merit), b) sozialer Wert im normativen oder moralischen Sinne, c) statistischer Wert.

Die Bestimmung des Wertes eines Evaluationsgegenstandes (a), d.h. die Bewertung, sei zentrale und unverzichtbare Aufgabe jeder Evaluation. Kriterien seien dabei die Bezugsmaßstäbe für die Bewertung, und damit der Dreh- und Angelpunkt der Bewertung. Vier generische Kriteriendimensionen für Evaluationen, das heißt für die Bewertung einer Evaluation (Metaevaluation) sind Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit (vgl. DeGEval-Standards). Für Bewertungsakte in Evaluationen sind in der Regel operationalisierte Kriterien erforderlich.

Zum Umgang mit Werten im Vergleich zwischen Forschung einerseits und Evaluation andererseits stellt er die These auf, dass die Evaluation der Sozialforschung aktuell „um eine Nasenlänge voraus“ ist, da sie sich systematisch mit der Bewertungsfrage auseinandersetzt (statt sie zu vermeiden). Schließlich geht er noch auf generische Kriteriendimensionen für Gegenstände der Bildungsevaluation und Quellen für wertebasierte Kriterien in Gleichstellungsevaluationen ein.

Er schließt seinen Vortrag ab mit der Beantwortung zweier Fragen, welche aus den Reihen des Nachwuchsnetzwerkes, im Rahmen des Kolloquiums am Vortag, an ihn gerichtet worden sind.

Zum Abschluss des Transdisziplinären Kolloquiums dankt Jan Hense für die Gastgeber allen Anwesenden für die rege Beteiligung und freut sich auf weitere Diskussionen, die im Rahmen der Tagung nur angestoßen, jedoch nicht abgeschlossen werden konnten.