

Zwischen Objektivität und Normativität: Wie kommt die Evaluation zu(m) Werten?

Transdisziplinäres Kolloquium

Freitag, 27.02.2015

Justus-Liebig-Universität Gießen

Arbeitskreis Aus- und Weiterbildung in der Evaluation
in der DeGEval

Arbeitskreis Berufliche Bildung in der DeGEval

Nachwuchsnetzwerk der DeGEval



Wo sind die Werte in der Evaluation?

Thematische Einführung in das Kolloquium

Prof. Dr. Jan Hense

Justus-Liebig-Universität Gießen

Professur für Hochschuldidaktik und Evaluation

jan.hense@psychol.uni-giessen.de



Freitag, 27.02.2015

Justus-Liebig-Universität Gießen

Evaluation kommt vom „wert sein“

lat. *valere*

→ frz. *valoir* → frz. *value* → frz. *évaluer* → frz. *évaluation*

→ engl. *evaluation*

→ dt. *Evaluation*

(Karbach, 2000)

Evaluation kommt vom „wert sein“

„Evaluation ist die systematische Untersuchung
des **Nutzens oder Wertes** eines Gegenstands“

(Standards der Evaluation, DeGEval, 2002)



“Bad is bad and good is good and it is the job of evaluators to decide which is which”

Michael Scriven, 1986



Heinrich Wottawa
Heike Thierau

Lehrbuch Evaluation

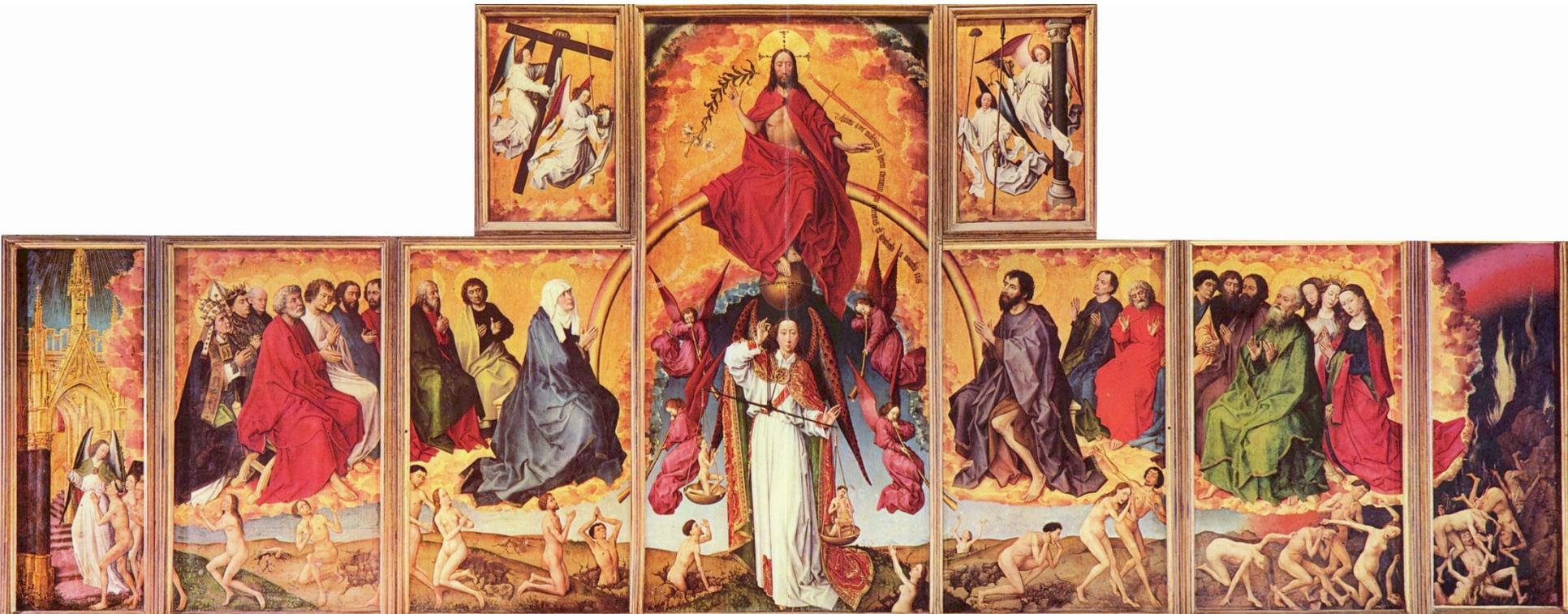
2., vollständig
überarbeitete
Auflage

Verlag
Hans Huber



Der Erzengel Michael
bei der
„ultimativen Evaluation“

Rogier van der Weyden
„Das jüngste Gericht“, 1443
Hôtel-Dieu, Beaune



Quelle: http://de.wikipedia.org/wiki/Datei:Rogier_van_der_Weyden_001.jpg

Welche Rolle spielen Werte und das Werten wirklich in der Evaluation?

Einige Beispiele

Gibt es überhaupt ein Werturteil?

Evaluating the Use of Computer Mediated Communication Resources by Trainee Educational Psychologists

1) How receptive were the trainees to the introduction of the conferencing system?

2) What was the pattern and content of the online interaction, how dependent was it on task characteristics, and how far did it differ from off-line interaction?

Tolmie et al. (2000). Evaluating the Use of Computer Mediated Communication Resources by Trainee Educational Psychologists. In Mogyey (ed.), *LTDI Evaluation Studies*, pp. 31-38

Womit wird verglichen?

Kriterien	Traditionelle Lehre		POL- Kurs Onkologie	
	M*	SD	M*	SD
Reaktionsebene				
– Akzeptanz der organisatorischen Rahmenbedingungen	4,17		3,43	
– Allgemeine Einstellung zu POL	3,96		3,81	
Lernebene				
– Wissen	3,99		3,44	
– Fähigkeiten	3,90		3,71	
– Motivation zu selbständigem Lernen	3,31		3,13	
– Erfahrungen mit kooperativer Arbeit	3,92		4,08	
Transfer				
– Transfererwartung	3,98		3,49	

*Wertebereich: 1 = niedrig bis 6 = hoch

(Bergmann, Eisfeldt & Lanario (2004). Messwiederholungspläne als Instrumente einer vergleichenden Evaluation. *Zeitschrift für Evaluation*, 1/04, S. 16)

Womit wird verglichen?

„Die multivariate Varianzanalyse weist signifikante Unterschiede zugunsten der traditionellen Lehrform aus. Die Effektunterschiede sind aber **klein** und in den Einzelvergleichen nur für die Skala Wissen ($p=0,001$, Eta Quadrat $0,181$) und die Transfererwartung ($p=0,004$, Eta Quadrat $0,133$) **signifikant**.

Diese Studie zeigt, dass die mit dem PÜL-Kurs Onkologie angestrebten **Ziele** in der evaluierten Version **noch nicht erreicht** sind und **erheblicher Verbesserungsbedarf** besteht.“

(Bergmann et al., 2004, S. 15-16)

Womit wird verglichen?

„In dieser Studie sind die Einschätzungen zu den Kriterien generell **nicht günstig**. Die Mittelwerte liegen für die traditionelle Lehre mit einer Ausnahme (Motivation zu selbständigem Lernen) knapp über dem Wert von 3,5 und zeigen somit eine **tendenzielle Zielerreichung** an. Für den POL-Kurs Onkologie **unterschreiten sogar** die Mittelwerte bei vier Kriterien den Wert von 3,5. **Schon** die Art der organisatorischen Implementierung dieses Kurses erreicht **wenig Akzeptanz**. Die zusätzliche Berücksichtigung der Streuungen macht klar, dass ein großer Teil der Studenten bei vier Kriterien mehr als die Hälfte die Zielerreichung nicht bestätigt. Damit ist **drastischer Verbesserungsbedarf** angezeigt.“

(Bergmann et al., 2004, S. 16)

Wer fällt das Werturteil?

3.4 6 Gesamtbewertung der BMBF-Förderlinie

Befragt nach der subjektiven Gesamtbewertung der BMBF-Förderlinie gibt die Mehrheit der befragten Projektverantwortlichen an, insgesamt sehr zufrieden/zufrieden mit der BMBF-Förderlinie zu sein (n=33).

(bmbf, 2013, Evaluierung der BMBF-Förderlinie „Hochschulforschung als Beitrag zur Professionalisierung der Hochschullehre“ – Abschlussbericht, S. 90)

Welche "Schulnote" (1-5) würden Sie der Dozentin / dem Dozenten als Veranstaltungsleiter/in geben?

sehr gut (1) mangelhaft (5)

Wer fällt das Werturteil?

The image shows the logo for 'Stiftung Warentest' (Stiftung Warentest) on the left, which consists of a grey rectangle with the text 'Stiftung Warentest' and a red rectangle with the word 'test' and a large white number '1'. To the right, on a grey background, are the test results: 'TESTSIEGER' (Test Winner), 'GUT (2,1)' (Good (2,1)), and a list of criteria: 'Bildqualität gut (2,2)' (Image quality good (2,2)), 'Handhabung gut (2,4)' (Handling good (2,4)), 'Vielseitigkeit sehr gut (1,2)' (Versatility very good (1,2)), and 'Stromverbrauch gut (2,0)' (Power consumption good (2,0)). At the bottom, it says 'Im Test 16 digitale Bilderrahmen' (In the test 16 digital picture frames) and 'Ausgabe 11/2010' (Issue 11/2010).

**Stiftung
Warentest**

test

TESTSIEGER

GUT (2,1)

Bildqualität gut (2,2)
Handhabung gut (2,4)
Vielseitigkeit sehr gut (1,2)
Stromverbrauch gut (2,0)

Im Test 16 digitale Bilderrahmen
Ausgabe 11/2010

(test 11/2010, S. 37-41)

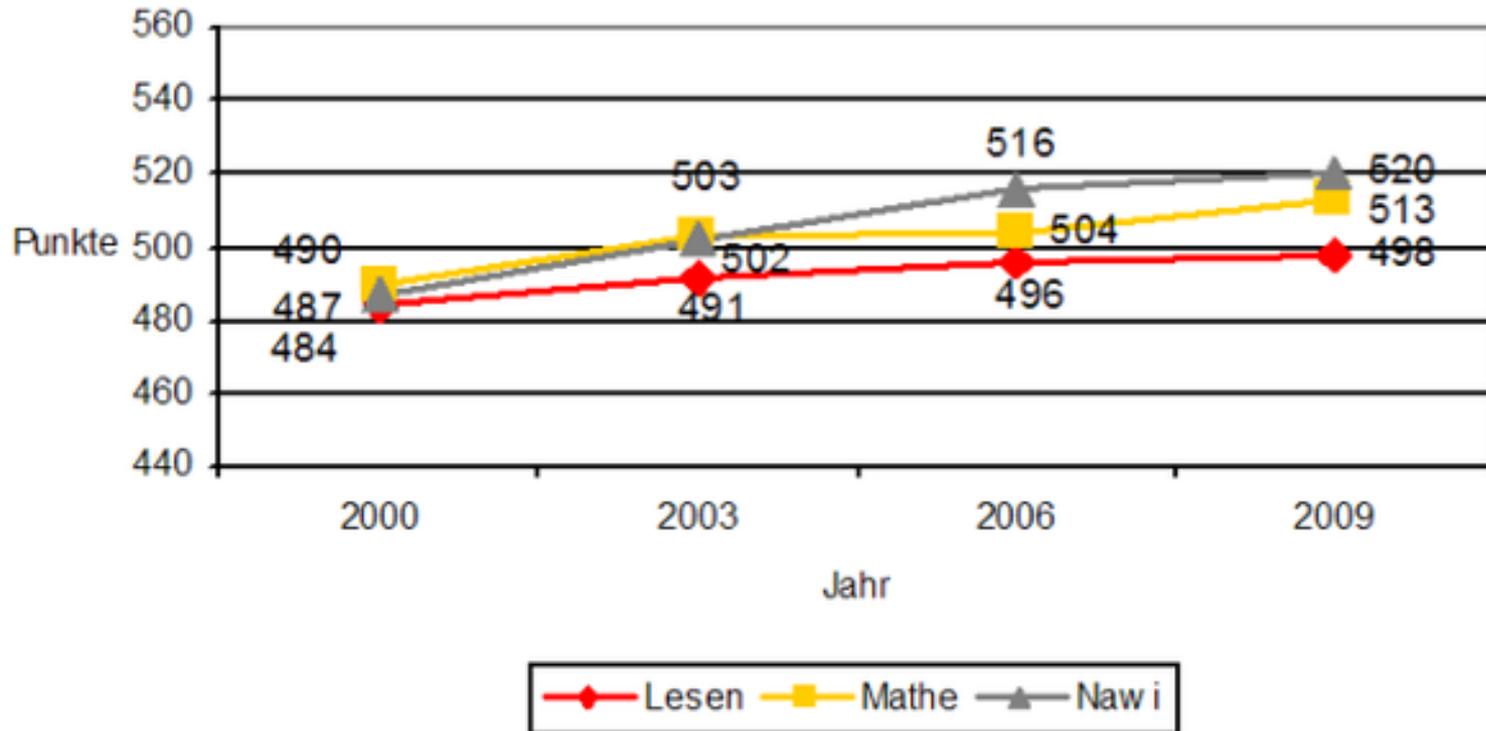
Wer fällt das Werturteil?

		Nur mit Netzteil						
	Gewichtung	1 Transcend PF830W	2 Aiptek Monet Deluxe P8V5	3 Intenso 10" Media-Creator	4 Samsung SPF-1000W	5 Alessio L.A.8 ¹⁾	6 Kodak Pulse	
Stromaufnahme in Euro ca.			70 bis 90	90 bis 129	140 bis 180		109 bis 160	117 bis 150
Herstellerpreis in Euro ca.		86	78	113	162	59	135	135
test -QUALITÄTSURTEIL	100%	GUT (2,1)	GUT (2,4)	GUT (2,4)	GUT (2,4)	GUT (2,5)	BEFRIEDIGEND (2,7)	BEFRIEDIGEND (2,8)
BILDQUALITÄT	55%	gut (2,2)	gut (2,3)	gut (2,0)	befried. (2,8)	gut (2,4)	gut (2,2)	befried. (2,9)
Sehtest / Messungen		+ / ○	○ / +++	+ / +	+ / +++	+ / +	+ / +	+ / +
Betrachtungswinkel / Reflektionsarmut		⊖ / +++	⊖ / +++	○ / +++	○ / -*)	⊖ / +	○ / +++	⊖ / -*)
HANDHABUNG	25%	gut (2,4)	befried. (2,7)	befried. (3,0)	sehr gut (1,5)	gut (2,5)	befried. (3,3)	gut (2,5)
Nutzwahl / Bedienung		○	○	○	+	+	⊖*)	○
Bedienung / Fotofunktionen		○ / +	○ / +	○ / ○	++ / +	+ / ○	+ / ○	○ / +
LEISTUNGSFÄHIGKEIT	15%	sehr gut (1,2)	gut (1,9)	befried. (2,7)	gut (1,7)	befried. (3,4)	befried. (3,1)	befried. (3,1)
ENERGIEVERBRAUCH	5%	gut (2,0)	befried. (3,1)	gut (2,4)	befried. (3,1)	gut (1,9)	befried. (2,7)	befried. (2,7)
Wartungsintervalle / Laufzeit		+ / Entfällt	○ / Entfällt	+ / Entfällt	○ / Entfällt	+ / Entfällt	○ / Entfällt	○ / Entfällt
PHYSISCHE MERKMALE								
Abmessungen (B x H x T) in cm		16,2 x 12,1	16,2 x 12,2	20,4 x 15,4	20,4 x 15,3	16,2 x 12,2	16,2 x 12,2	16,2 x 12,2
Verhältnis Bildbreite zu Bildhöhe		4:3	4:3	4:3	4:3	4:3	4:3	4:3

(test 11/2010, S. 37-41)

Wessen Werte sind maßgeblich?

Vergleich der PISA-Ergebnisse 2000 bis 2009 in den Kompetenzbereichen Lesen, Mathematik und Naturwissenschaften in PISA-Punkten



Es geht auch anders (und damit auch besser?)



Program Assessment of iKNOW Politics

December 2008

Summary of Key Findings

Indicators	Rating
------------	--------

Facilitating a global exchange of knowledge

Level of active usage of exchange features	 Low
Level of passive usage of exchange features	 Moderate – High
Frequency of usage of exchange features	 Low
Change to knowledge of members	 Moderate
Instances of exchanges contributing to women’s participation in politics	Identified

Facilitating a global exchange of resources

Level of usage of resource areas of the website	 High
Frequency of exchange of resources from website to members	 High
Frequency of exchange of resources from members to website	 Low
Frequency of exchange of resources from members to peers	 High

Ausblick auf das Programm

10:00 Einstiegsreferate

1. Die Bedeutung der Evaluationsfragestellung für das Bewerten (*Verena Friedrich*, Universität Bern)
2. Verwendung von Bewertungskriterien in der Praxis: eine empirische Untersuchung (*Christine Heuer*, Bundesamt für Gesundheit, Bern)

11:00 Kaffeepause

11:15 **Vergleichsperspektiven in der Evaluation:** Fallberichte aus der beruflichen Bildung

1. Die Vergleichsperspektive als Triebfeder einer Bewertungskultur (*Miguel del Estal*, InterVal GmbH, Berlin)
2. Das Qualitätssiegel des MedienCampus Bayern als Beispiel zur Evaluation von Bildungsangeboten (*Veronika Alz*, Medien-campus München)

12:45 *Mittagspause mit Imbiss (in Teilnahmegebühr enthalten)*

13:45 (A) **Wer bewertet letztendlich?**
Evaluation oder Rezipienten?

1. Evaluation einer Lehrveranstaltung als Bewertungsgrundlage für verschiedene Interessengruppen (*Felix Albrecht*, Karlsruher Institut für Technologie)
2. Bewertung in der Evaluation regionaler Siedlungsentwicklung: ein Vergleich Deutschland-USA (*Christian Diller*, Universität Gießen)

(B) **Wessen Werte?** Werte(n) in sozialen und komplexen Kontexten

1. Der Einsatz der FIWE-Methode im Kontext der externen Schulevaluation (*Dirk Groß*, Agentur für Qualitätssicherung, Evaluation und Selbstständigkeit von Schulen, Bad Kreuznach)
2. Developmental Evaluation: Was bedeutet die Evaluation von Innovationen in komplexen, dynamischen Umfeldern für das Bewerten? (*Maria Gutknecht-Gmeiner*, Impulse - Evaluation und Organisationsberatung, Wien)

15:15 Kaffeepause

15:30 Keynote „Wie kommt die Evaluation zu(m) Werten?“ und Resümee zur Tagung (*Wolfgang Beywl*, Pädagogische Hochschule/Fachhochschule Nordwestschweiz)