

Trends in der Auseinandersetzung und Nutzung von Rückmeldungen aus Vergleichsarbeiten

17. Jahrestagung der DeGEval- Gesellschaft für Evaluation e.V.,
gemeinsam mit der Schweizerischen Evaluationsgesellschaft SEVAL

Session C7: Evaluationsforschung und Forschung über Evaluation im
Kontext der Professionalisierung schulischen Handelns
12. September 2014, Pädagogische Hochschule Zürich

Dr. Jana Groß Ophoff & Prof. Dr. Ingmar Hosenfeld

Die empirische Wende im deutschen Bildungswesen


Theorie

Fragen

Methoden

Ergebnisse

Diskussion

- 2000ff.: Enttäuschende Ergebnisse in internationalen Schulleistungsstudien (u.a. TIMMS: Baumert et al., 2000; PISA: Baumert et al., 2001)
- 2002: 7 Handlungsfelder, u.a. Entwicklung und regelmäßige Evaluation von Bildungsstandards (KMK, 2002)
- 2003: Vergleichsarbeiten in einzelnen Bundesländern (RP, TH)
- 2004: Verabschiedung Bildungsstandards (KMK, 2004)
- 2004:  **Vergleichsarbeiten** in der Grundschule : Verbund aus 7 Bundesländern (Helmke 2004)
- 2006: Gesamtstrategie zum Bildungsmonitoring (KMK 2006)
- 2010: Kompetenzniveauumodelle auf Basis der Bildungsstandards (IQB, 2013)

Vergleichsarbeiten in der Grundschule

Theorie

Fragen

Methoden

Ergebnisse

Diskussion



VERA

- seit 2004: jährlich flächendeckend
- Universität Koblenz-Landau (LD)
- Institut z. Qualitätsentw. im Bildungswesen (IQB)
- Low-Stakes
- partizipativ

Lehrkräfte

- Testimplementation
- Korrektur der Aufgabenbearbeitung
- Dateneingabe
- Beschäftigung mit Feedback
- ggf. Ableitung von Veränderungen

Projektgruppe (LD)

Entwicklung von...

- Leistungstests (seit 2009: IQB)
- Begleitmaterialien
- Internet Portal & technischer Support
- Datenaufbereitung und Skalierung
- Feedback Leistungsergebnisse
- Begleitstudien



VERA

Beispiel-Rückmeldung

Theorie

Fragen

Methoden

Ergebnisse

Diskussion

Kompetenzniveaus (KN): Lesen – mit Texten und Medien umgehen

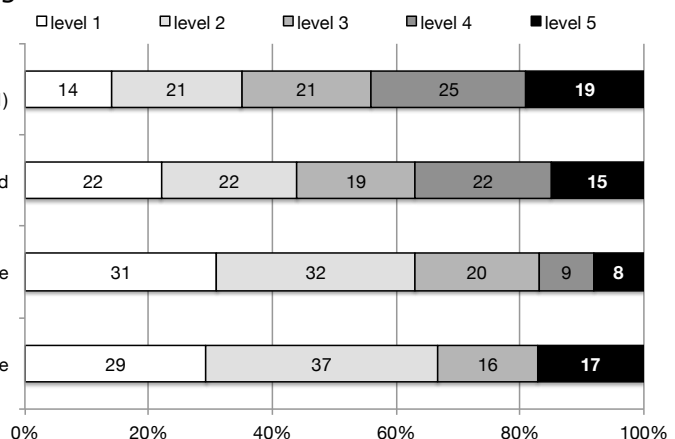
- 5: Optimalstandard
- 4: Regelstandard+
- 3: Regelstandard
- 2: Minimalstandard
- 1: unterhalb der Erwartung

(IQB, 2013)

Lesen		Fairer Vergleich (Referenzgruppe 1)
Schüler	Kompetenzniveau	
1 m	Niveau 1	
2 w	Niveau 2	
3 m	n.a.	
4 m	Niveau 5	
5 m ND	Niveau 1	
6 w	Niveau 1	
...		
Anmerkung. n.a.= nicht anwesend; m=männlich; w=weiblich; ND=Deutsch nicht dominante Sprache		

Kompetenzniveaus: Lesen (Klasse 3a)

Ein Schüler entspricht ca. 5 Prozent.





	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Baden-Württemberg						3	3	3		
Bayern										
Berlin										
Brandenburg	4*	4	4	4						
Bremen	4	4	4	3	3	3	3	3	3	3
Mecklenburg-Vorpommern	4	4	4*	3	3	3	3	3	3	3
Niedersachsen										3
Nordrhein-Westfalen	4	4		3						3
Rheinland-Pfalz	4	4	4	3	3	3	3	3	3	3
Saarland					3	3	3	3	3	3
Schleswig-Holstein	4	4	4*	3	3	3	3	3	3	3
N Schulen	6091	6487	2336	5960	10730	9966	10921	11064	8712	8738

Portal/Materialien: kontinuierliche Optimierung



Schematische Theorie zum Einfluss von Evaluationen

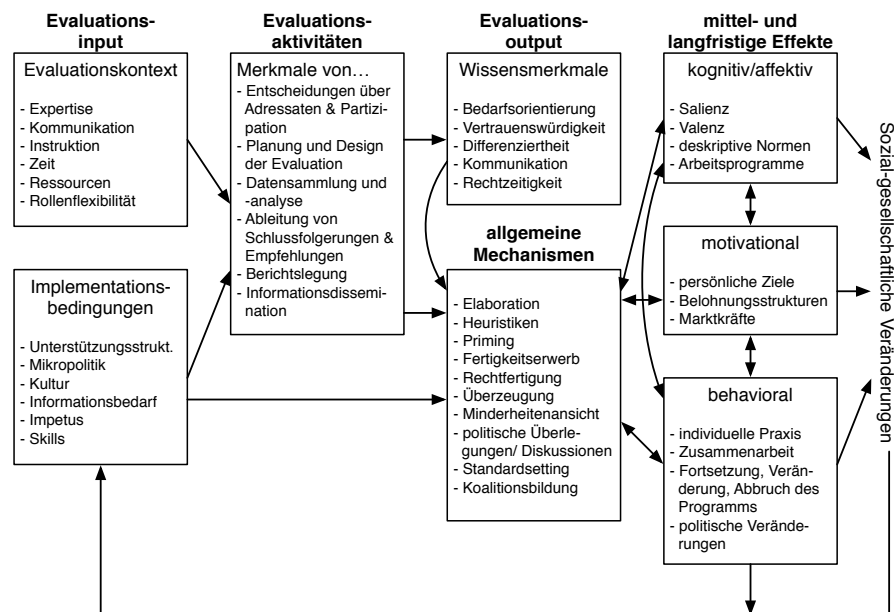
Theorie

Fragen

Methoden

Ergebnisse

Diskussion



(Mark & Heny, 2004; Cousins, 2003)



Stand der Forschung

Theorie

Fragen

Methoden

Ergebnisse

Diskussion

Teilnehmende Lehrkräfte...

➡...berichten intensive Auseinandersetzung und schätzen Feedback als verständlich und nützlich ein
(Bonsen et al., 2006; Grabensberger et al., 2008; Groß Ophoff, 2013; Kühle, 2010; Kühle & Peek, 2007)

➡...interessieren sich v.a. für Individualergebnisse
(Groß Ophoff, 2013; Maier, 2008; Nachtigall & Jantowski, 2007; von der Gathen, 2006)

➡... sind unsicher bei Interpretation und Nutzung von Feedback für Unterrichtsentwicklung (Groß Ophoff et al., 2006; Klug & Reh, 2000; Schneewind, 2006; Peek & Döbelstein, 2006; Tresch, 2007)

➡...berichten direkte Nutzung (Unterrichtsplanung und Gestaltung), aber nur selten innovative Maßnahmen
(Kühle, 2010; Groß Ophoff et al., 2006; Tresch, 2007)

Allgemein: Stagnierende oder rückläufige Trends
(Groß Ophoff, 2013; Nachtigall & Hellrung, 2013; Maier, 2009)

→ Aber: Längsschnittliche Studien sind selten.

Fragestellung

Theorie

Fragen

Methoden

Ergebnisse

Diskussion

Verändert sich die Beschäftigung mit Feedback aus Vergleichsarbeiten mit der Zeit und falls ja, wie?

Berichtete
Intensität der
Auseinandersetzung

Wahrgenommene
Verständlichkeit

Wahrgenommene
Nützlichkeit

	Gewöhnungs- Effekt	Neuigkeits- Effekt
	Gewöhnung reduziert kognitiven und behavioralen Load (Ajzen, 2002)	Neueinführung führt zu nicht nachhaltigem Aktivitätsanstieg (Clark, 1983; Kerres, 2001)
	↑	↓
	↑	↓
	↑	↓

Stichproben

Theorie

➔ Replikative Survey-Studie (6 Monate nach VERA)

Fragen

➔ web-basiert

Methoden

Ergebnisse



>50% Fragen unvollständig (AAPOR, 2008)

Diskussion

year	N _{Fragebögen}	Ausschluss	N _{bereinigt}	%Schulen
2004	2409	28.4%	1725	22%
2005	2233	27.8%	1613	17%
2006	1273	35.5%	821	25%
2007	1832	46.5%	981	14%
2008	3007	61.8%	1149	11%
2009	1531	56.2%	670	6%
2010	3835	66.5%	1289	10%
2011	2337	65.1%	816	7%
2012	842	60.9%	331	< 5.0%

Survey-Studie: Skalen zur Auseinandersetzung mit Feedback

Theorie

2004-2012:

Fragen

➔ Intensität: Wie intensiv haben Sie sich (...) beschäftigt?

Methoden

➔ Verständlichkeit: Wie verständlich fanden Sie die Rückmeldungen?

Ergebnisse

➔ Nützlichkeit: Wie nützlich fanden Sie die Rückmeldungen?

Diskussion

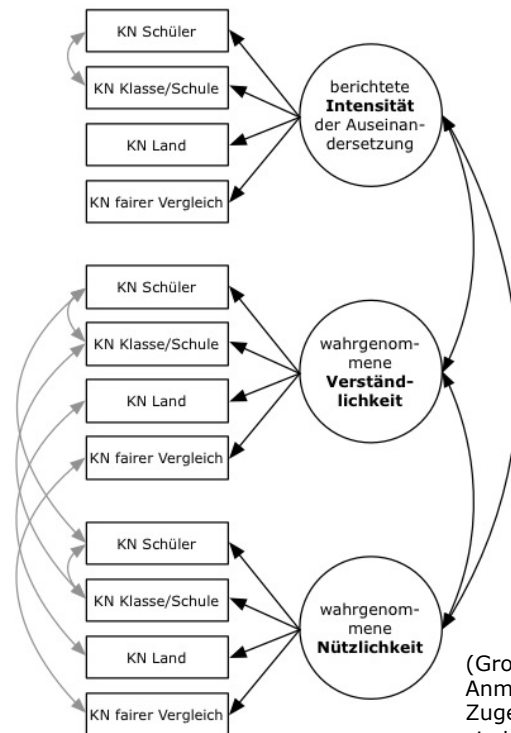
	gar nicht (1)	weniger (2)	eher (3)	sehr (4)
KN Schüler	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
KN Klasse/Schule	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
KN Land	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
KN im Vergleich zur Kontextgruppe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Anmerkung. KN = Kompetenzniveaus

Analyse: SEM (Mplus 6, Muthén & Muthen, 2010), Mehrgruppen-CFA für komplexe Stichproben

Analyse: SEM, Messmodell

- Background
- Questions
- Methods
- Results**
- Discussion



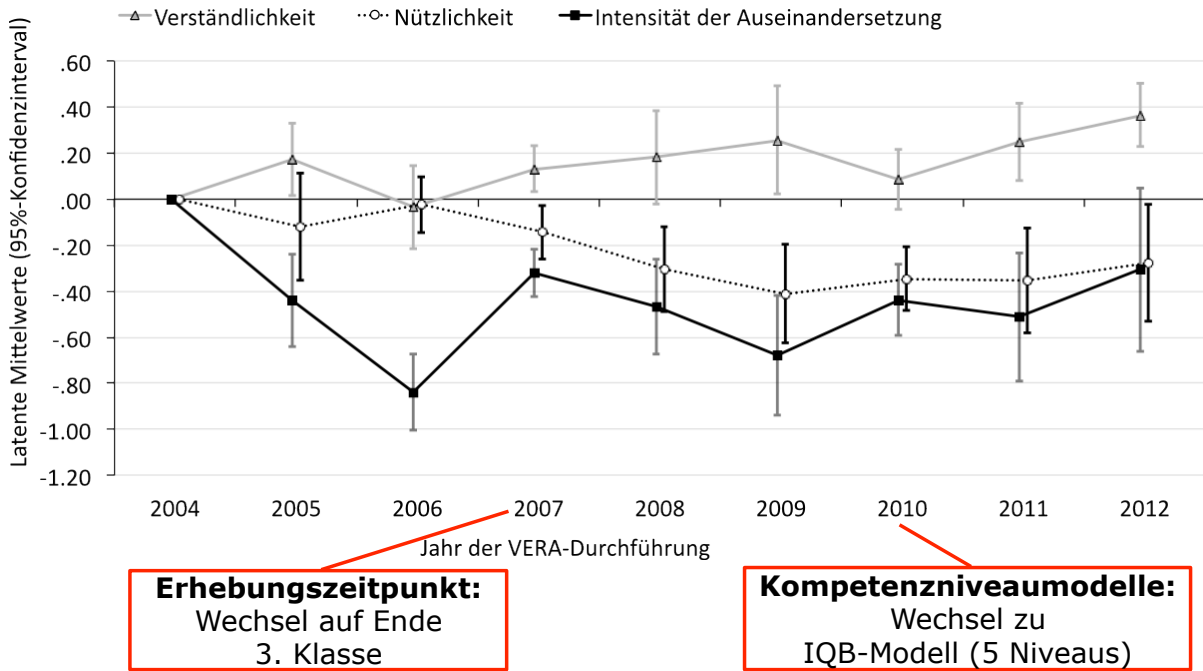
(Groß Ophoff, 2013)
 Anmerkung. KN = Kompetenzniveau
 Zugelassene Residualkorrelationen
 sind beispielhaft dargestellt.

Analyse: SEM, Messmodell

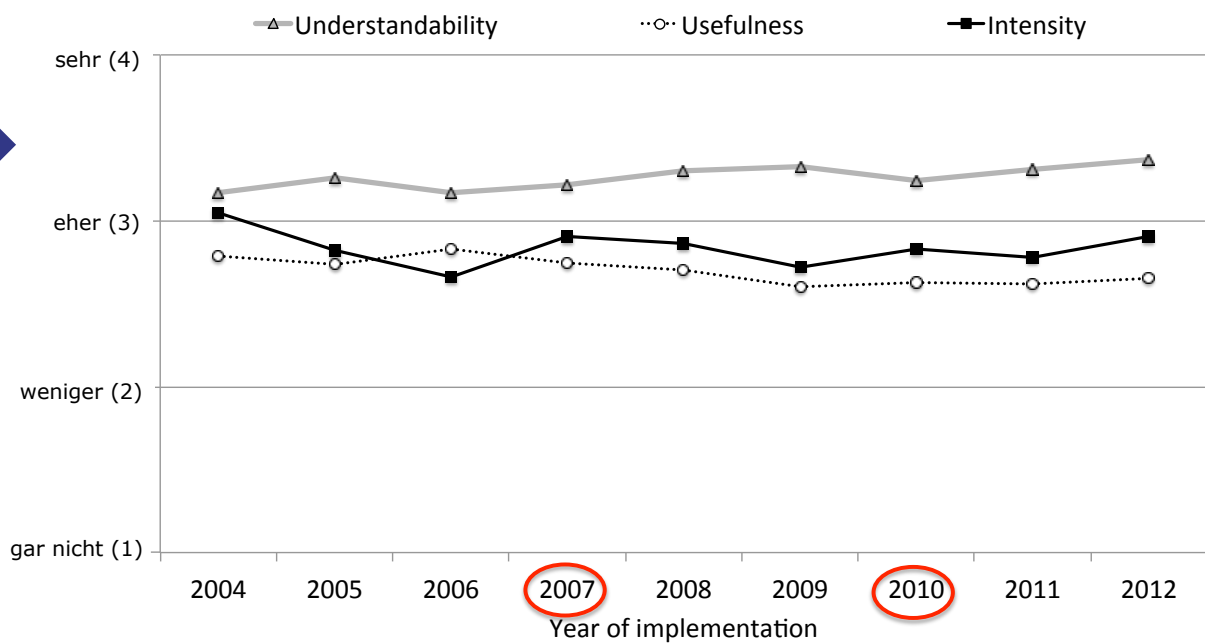
- Theorie
- Fragen
- Methoden
- Ergebnisse**
- Diskussion

	RMSEA (cutoff)	CFI	ΔCFI (cut-off)
2004	.056	.985	
2005	.048	.990	
2006	.054	.984	
2007	.055	.984	
2008	.048	.989	
2009	.062	.982	
2010	.052	.987	
2011	.050	.988	
2012	.046	.991	
Invarianz Hypothese			
2004-2012: Konfigurale Invarianz	.057 (.093)	.975	akzeptiert
2004-2012: Schwache faktorielle Inv.	.055	.973	.002 (.009) akzeptiert
2004-2012: Starke faktorielle Inv.	.054	.970	.003 (.008) akzeptiert
2004-2012: Strikte faktorielle Inv.	.059	.956	.014 (.009) zurückgew.

Analyse: SEM, Latente Mittelwerte



Analyse: Skalenmittelwerte (manifest)



Zusammenfassung

Theorie

Fragen

Methoden

Ergebnisse

Diskussion

Verändert sich die Beschäftigung mit Feedback aus Vergleichsarbeiten mit der Zeit und falls ja, wie?

Berichtete
Intensität der
Auseinandersetzung
Wahrgenommene
Verständlichkeit
Wahrgenommene
Nützlichkeit

Gewöhnungs- Effekt	Neuigkeits- Effekt
2004ff: steigender Trend	Rückgang bei Veränderungen
↑	↓
↑	
	↓



Diskussion

Theorie

Fragen

Methoden

Ergebnisse

Diskussion

Gewöhnungs- oder Neuigkeitseffekt? Beides!

- Veränderungsprozesse im Bildungswesen brauchen Zeit und werden durch Rahmenbedingungen beeinflusst (z.B. Rogers, 2003)
- Entscheidungen über Fortführung/Einstellung sollten nicht kurzfristig getroffen werden



Gemeinsam für Bildung

**10 Jahre VerA – das Ziel ist verfehlt.
Schulen brauchen Unterstützung statt Testeritis.**

Eine gute Schule ist Lern- und Lebensort.

Wie die Menschen in ihr zusammen leben und lernen, bestimmt ihre Qualität.

Schule lässt sich nicht reduzieren auf messbare Fachleistungen,

• sie ermöglicht Erfahrungen, die alle Kinder und Jugendlichen in ihrer Persönlichkeitsentwicklung voranbringen;

• VerA erfasst nur, was ist, und bringt kaum Hilfe, um Schule zu verbessern – vor allem fehlt es an Unterstützung für LehrerInnen in schwierigen Situationen.

• Der Aufwand für VerA ist hoch, verschlingt viel Geld und bindet Lern- und Arbeitszeit.

• Der Ertrag für Förderung ist gering, Ressourcen für pädagogische Vorhaben, Schul- und Unterrichtsentwicklung feh-



Trends in der Auseinandersetzung und Nutzung von Rückmeldungen aus Vergleichsarbeiten

17. Jahrestagung der DeGEval- Gesellschaft für Evaluation
e.V., gemeinsam mit der Schweizerischen
Evaluationsgesellschaft SEVAL

Session C7: Evaluationsforschung und Forschung über
Evaluation im Kontext der Professionalisierung schulischen
Handelns

12. September 2014, Pädagogische Hochschule Zürich

jana.grossophoff@ph-freiburg.de