



LEUPHANA
UNIVERSITÄT LÜNEBURG

Empirisch begründete Rückmeldungen zur Qualität von Unterricht

Marcus Pietsch



Wie viel ist ein Lehrer wert?

- 1) Lehrer sind maßgeblich für den Lernerfolg ihrer Schüler verantwortlich.



Wie viel ist ein Lehrer wert?

- 1) Lehrer sind maßgeblich für den Lernerfolg ihrer Schüler verantwortlich.
- 2) Lernzuwachs bei Schülern, die durch einen effektiven Lehrer, einem bei dem man viel lernt, unterrichtet werden, innerhalb von ca. 4 Jahren: ca. 0,20 bis 0,30 SD (ca. 1 zusätzliches Lernjahr) mehr als bei einem durchschnittlich effektiven Lehrer.



Wie viel ist ein Lehrer wert?

- 1) Lehrer sind maßgeblich für den Lernerfolg ihrer Schüler verantwortlich.
- 2) Lernzuwachs bei Schülern, die durch einen effektiven Lehrer, einem bei dem man viel lernt, unterrichtet werden, innerhalb von ca. 4 Jahren: ca. 0,20 bis 0,30 SD (ca. 1 zusätzliches Lernjahr) mehr als bei einem durchschnittlich effektiven Lehrer.
- 3) Lernzuwachs bei Schülern die durch einen ineffektiven Lehrer unterrichtet werden: ca. 0,20 bis 0,30 SD (ca. 1 zusätzliches Lernjahr) weniger als bei einem durchschnittlich effektiven Lehrer.

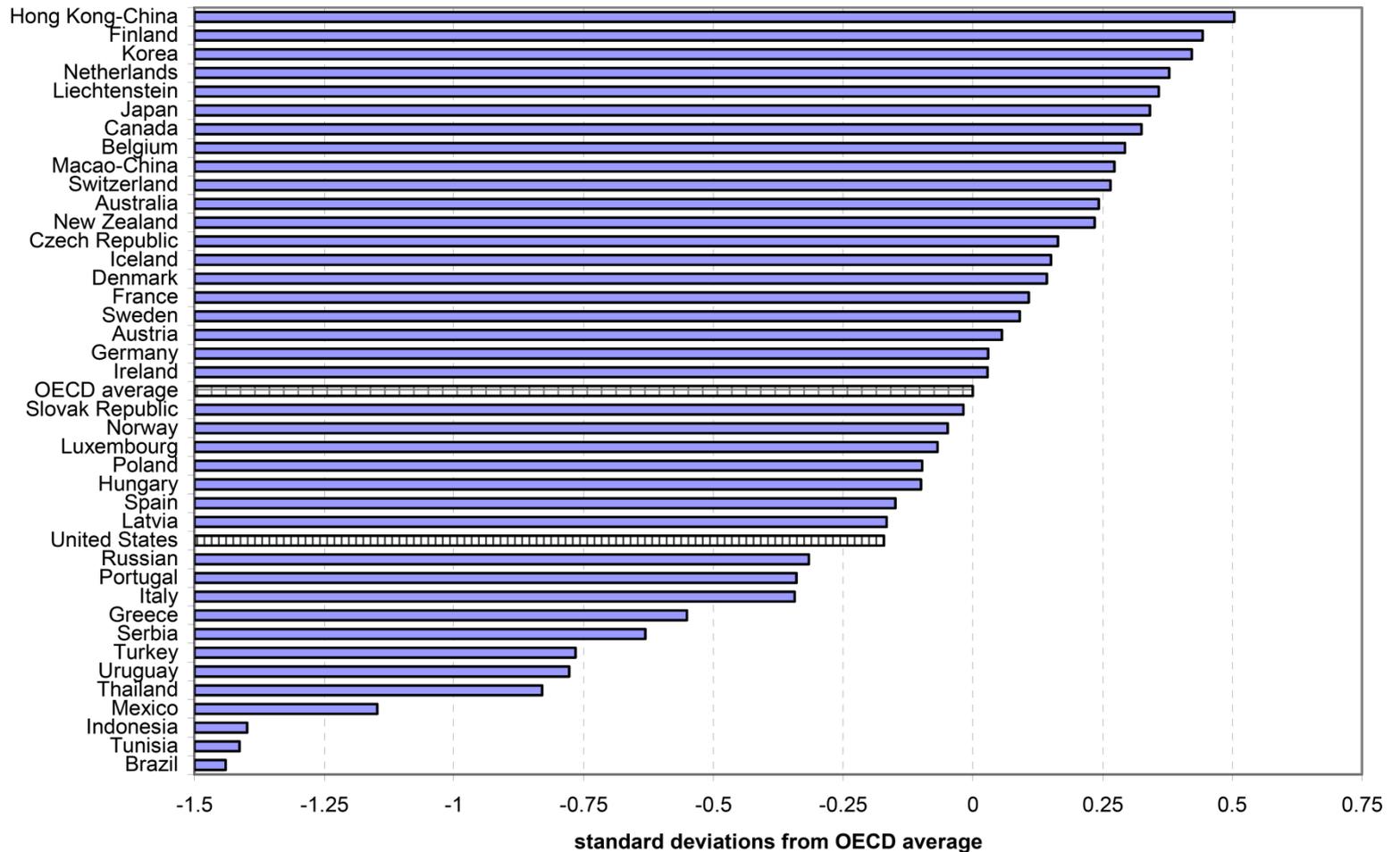


Wie viel ist ein Lehrer wert?

**Was würde passieren, wenn man die
ineffektiven Lehrkräfte entließe?**



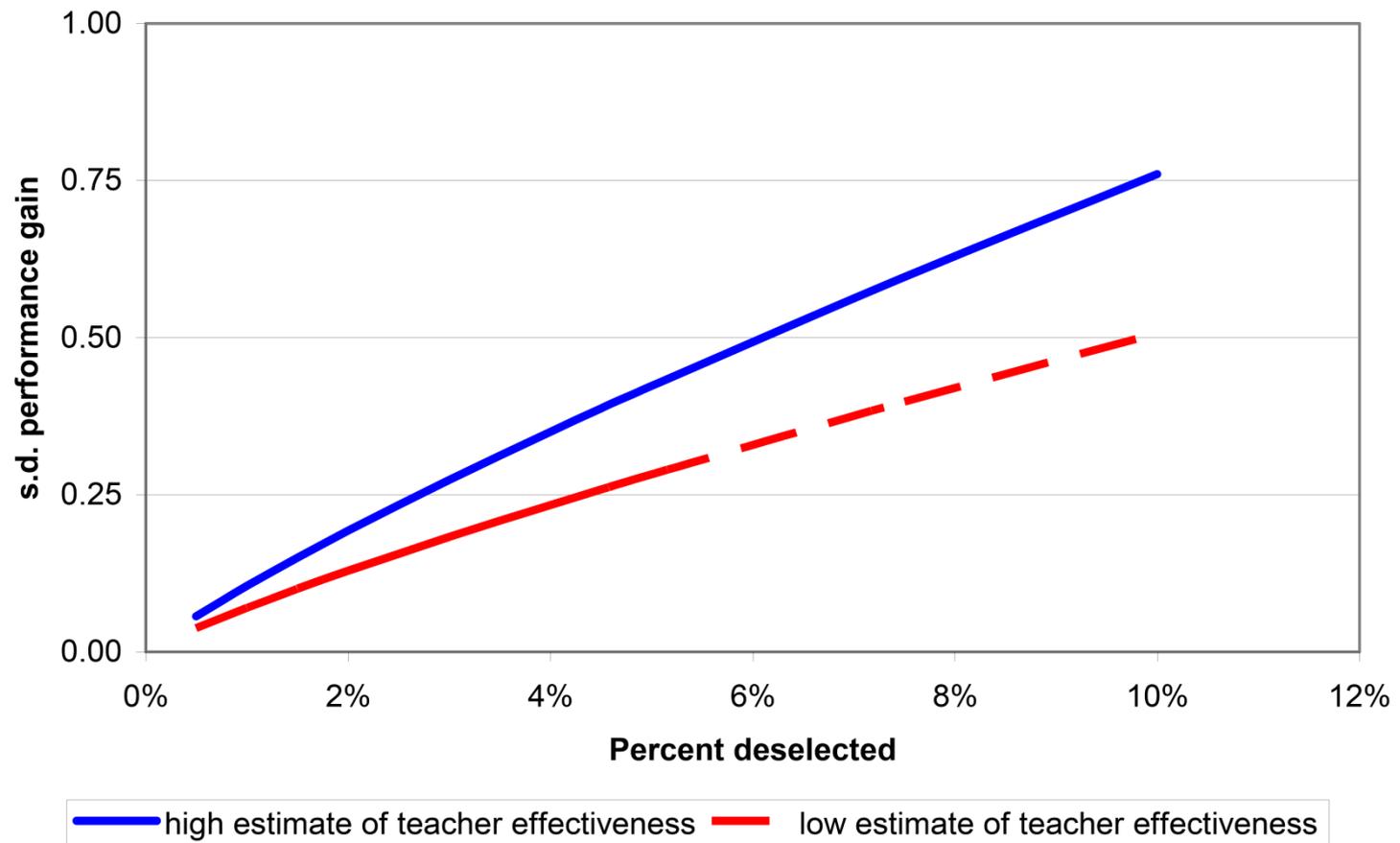
Unterschiede in PISA 2003 in SD





Wie viel ist ein Lehrer wert?

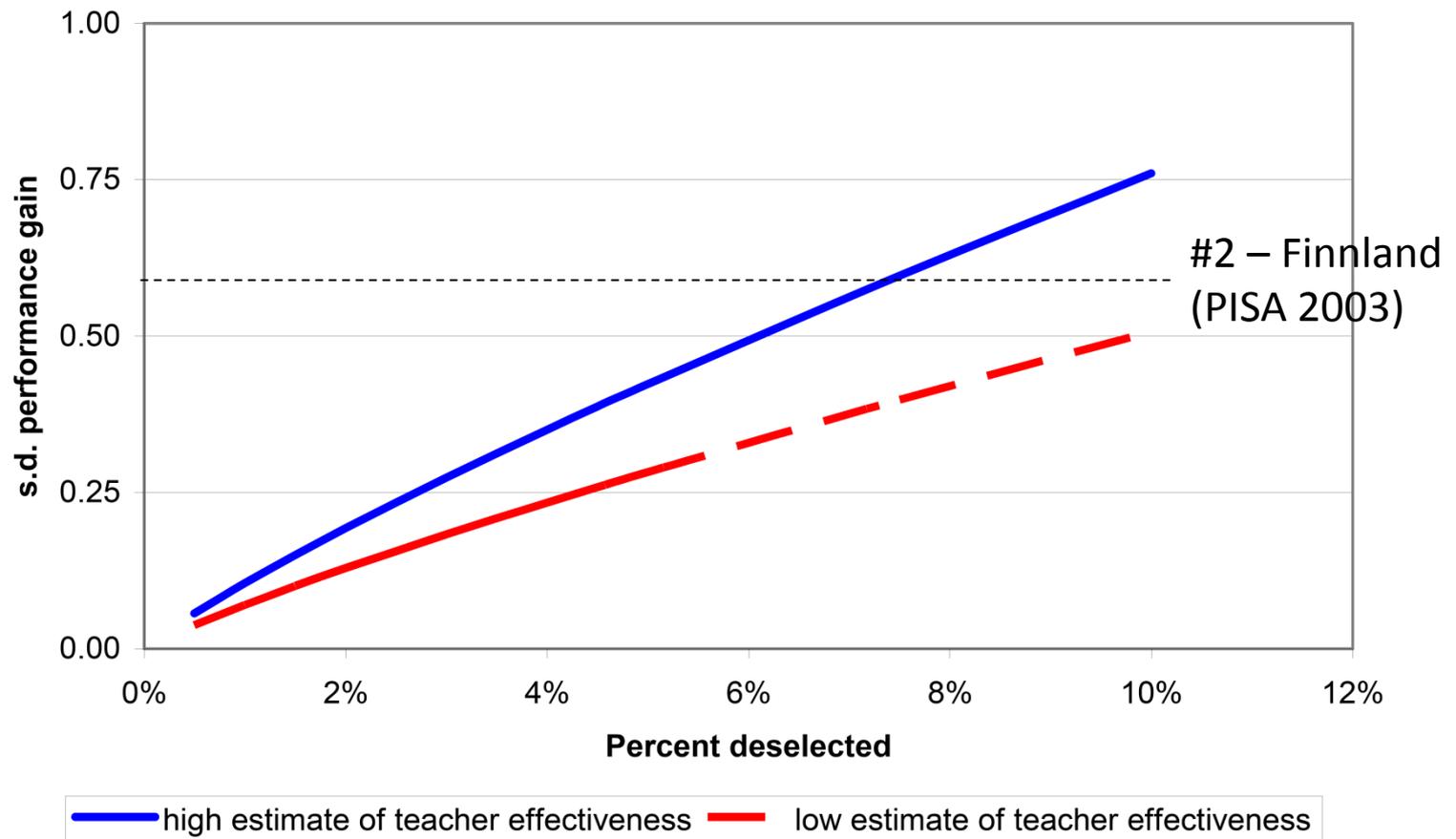
Figure 4. Alternative Estimates of Teacher Deselection and Student Achievement





Wie viel ist ein Lehrer wert?

Figure 4. Alternative Estimates of Teacher Deselection and Student Achievement



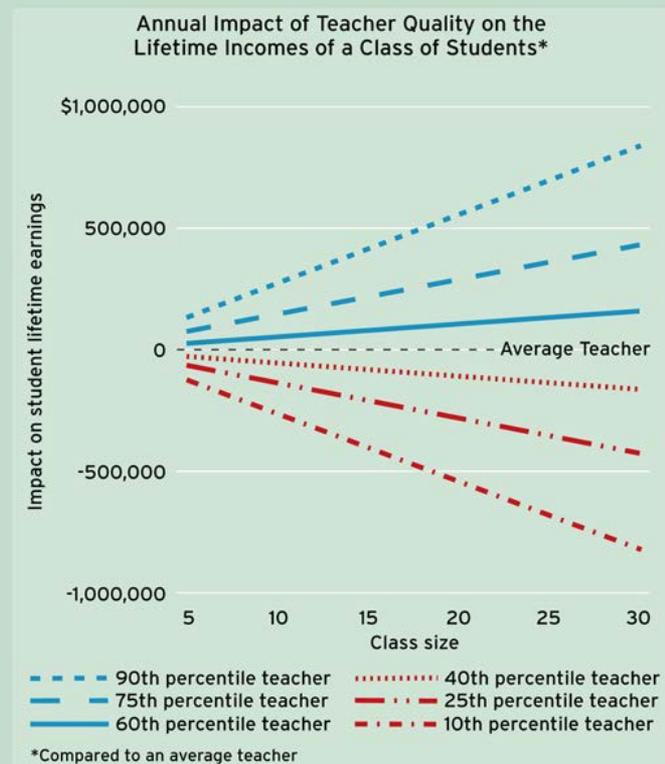


Der individuelle Nutzen dürfte groß sein!

Effective Teachers Raise Students' Earnings

(Figure 1)

The economic value of an effective teacher grows with larger classes, and the economic costs of having an ineffective teacher are substantial.



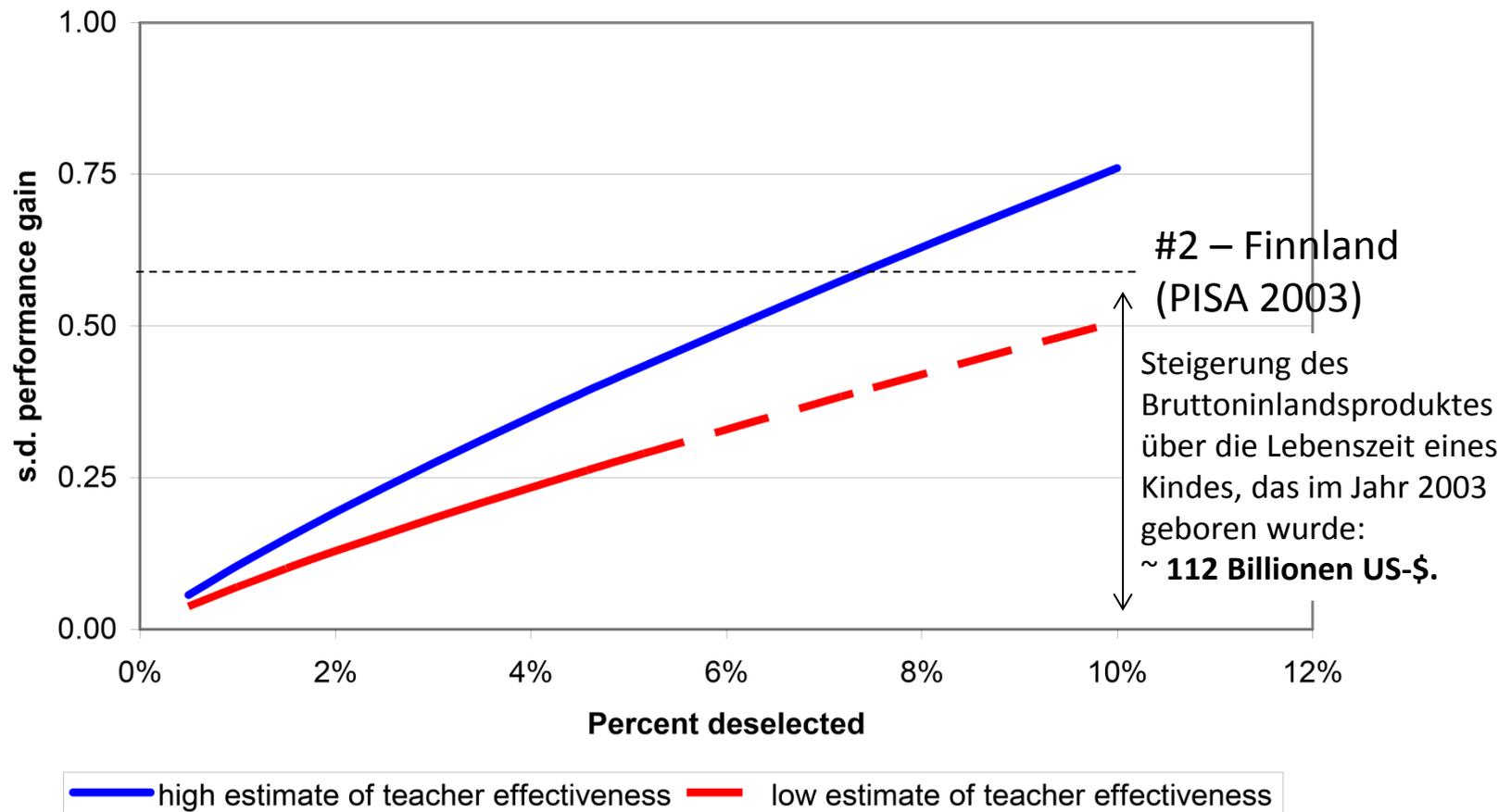
SOURCE: Authors' calculations

Würde man in den USA die 5% ineffektivsten Lehrkräfte entlassen und durch durchschnittlich effektive Lehrkräfte ersetzen, würden die Schülerinnen und Schüler in den durch sie unterrichteten Klassen im Lebensverlauf ca. 267.000 US-Dollar mehr verdienen.



Nicht zu reden von dem volkswirtschaftlichen Nutzen!

Figure 4. Alternative Estimates of Teacher Deselection and Student Achievement

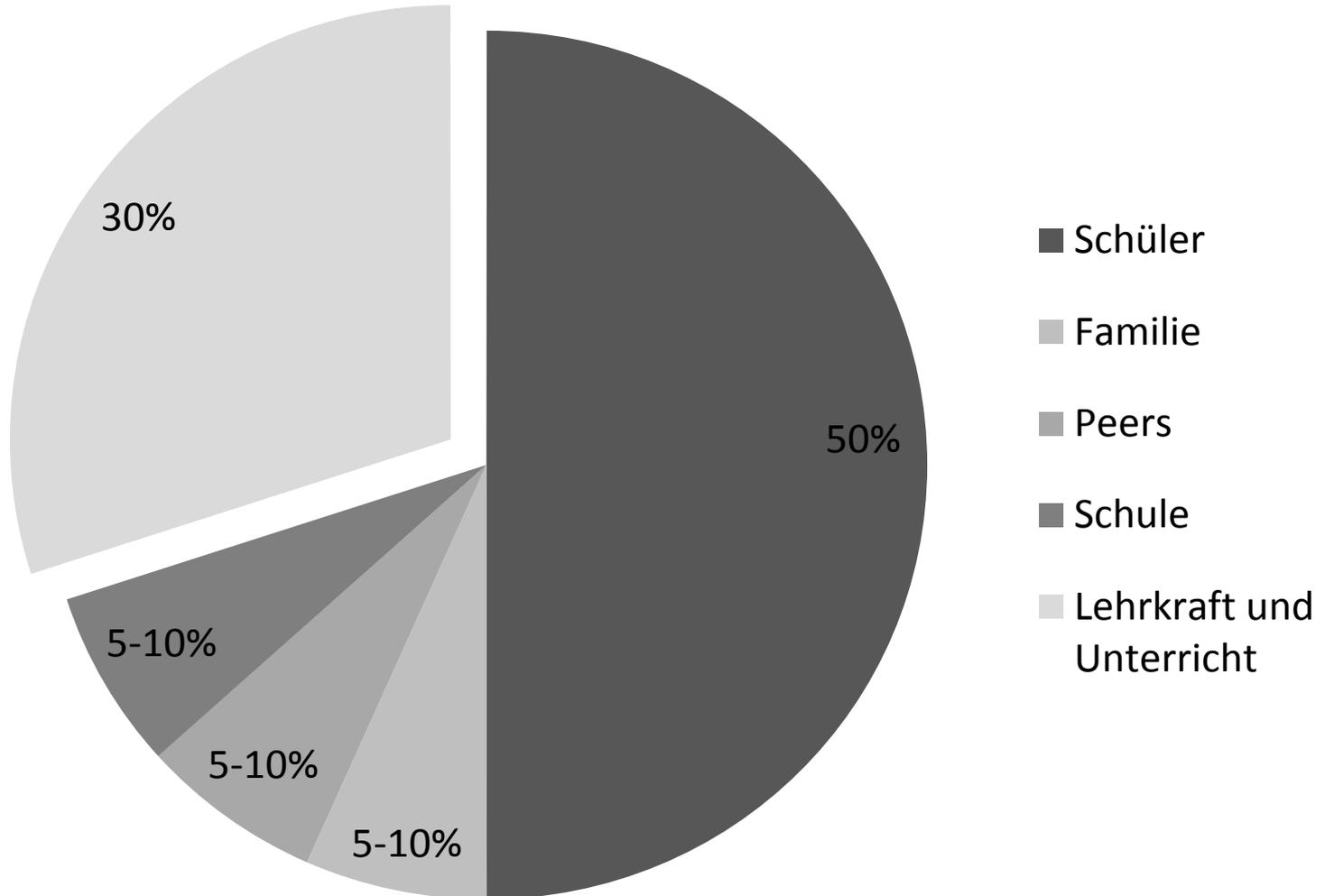




Wie wäre es stattdessen mit
einer Verbesserung des
Unterrichts?

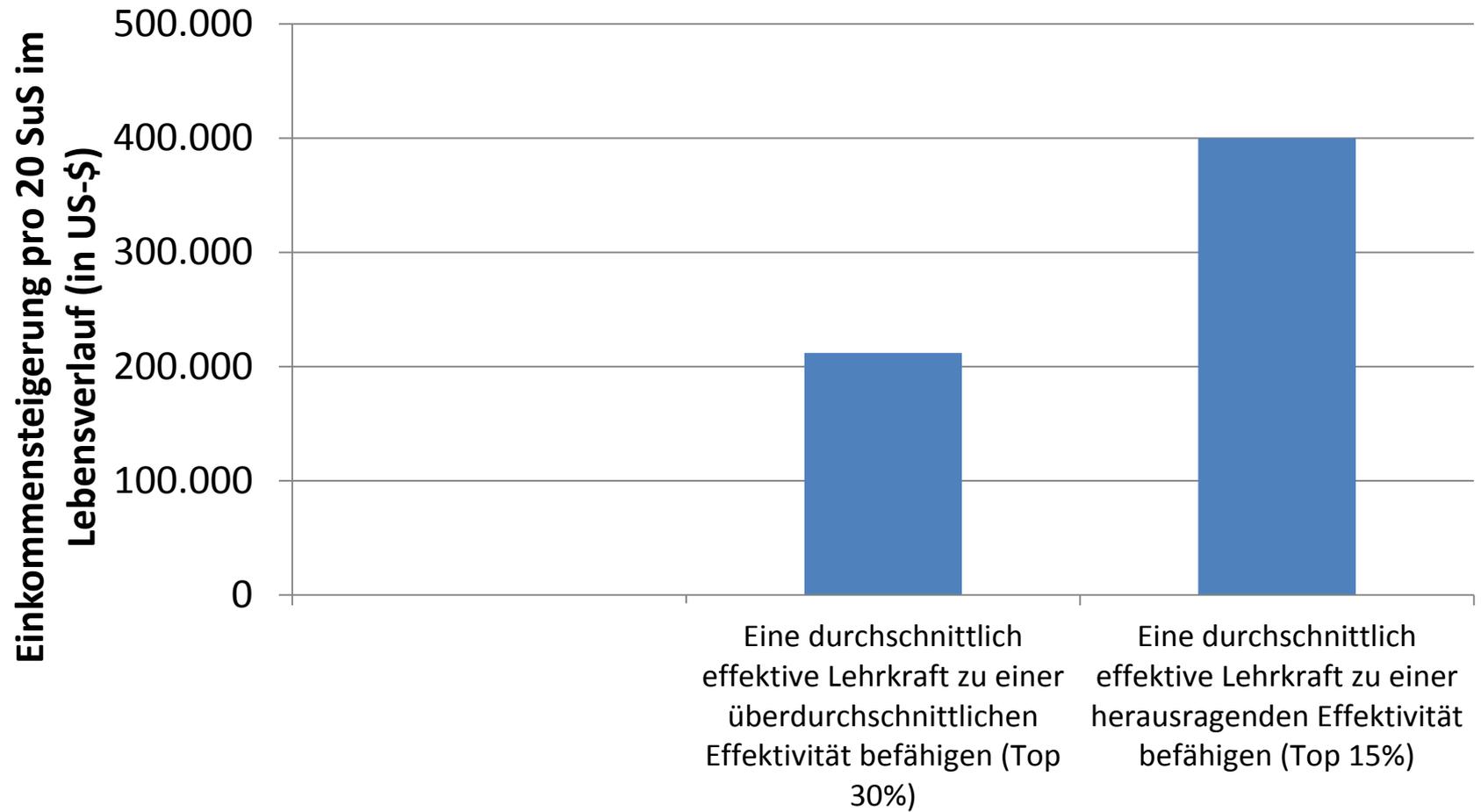


Erklärungen für Unterschiede in Lernzuwächsen



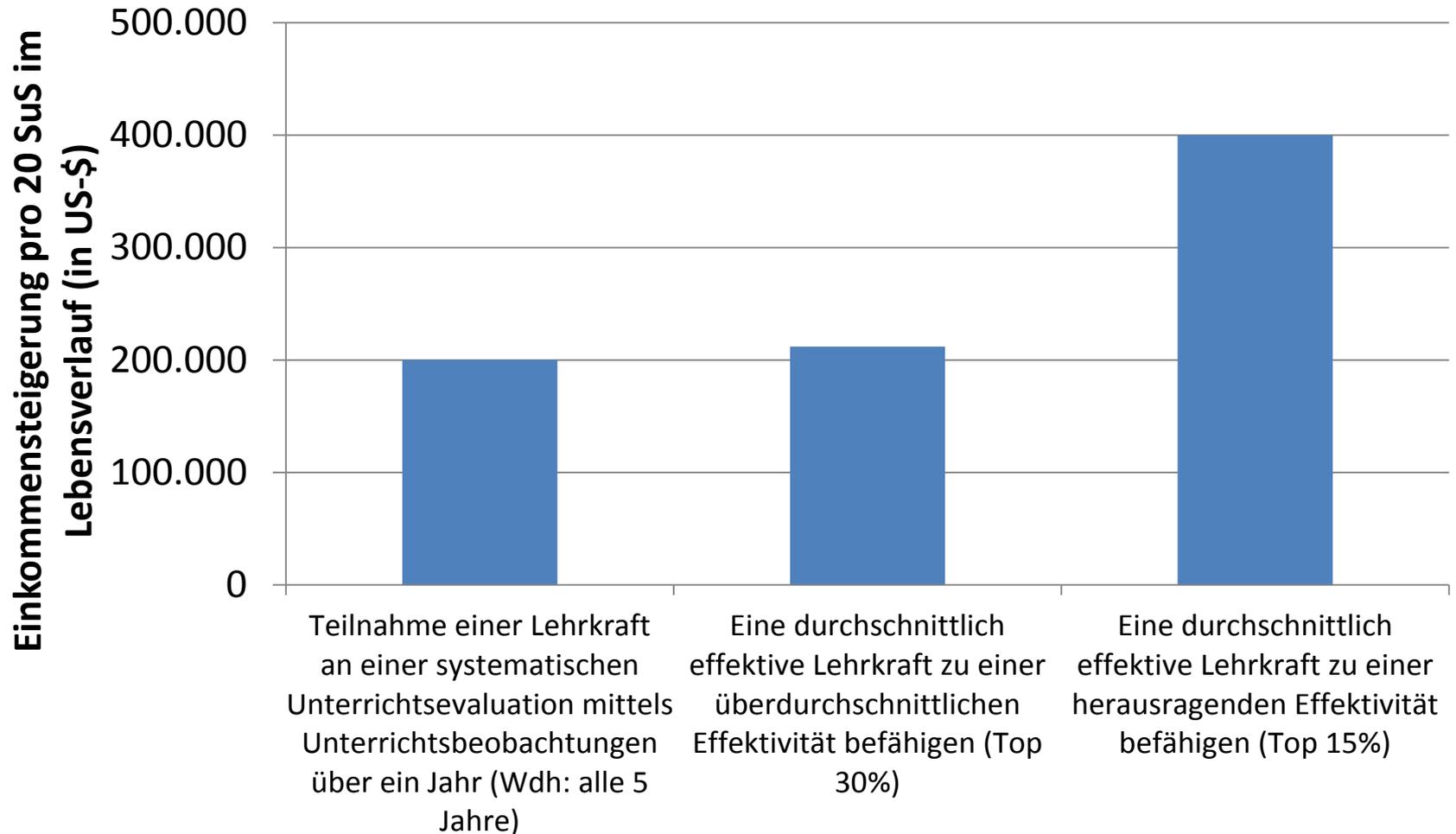


Wie viel ist Unterrichtsentwicklung wert?





Wie viel ist Unterrichtsentwicklung wert?



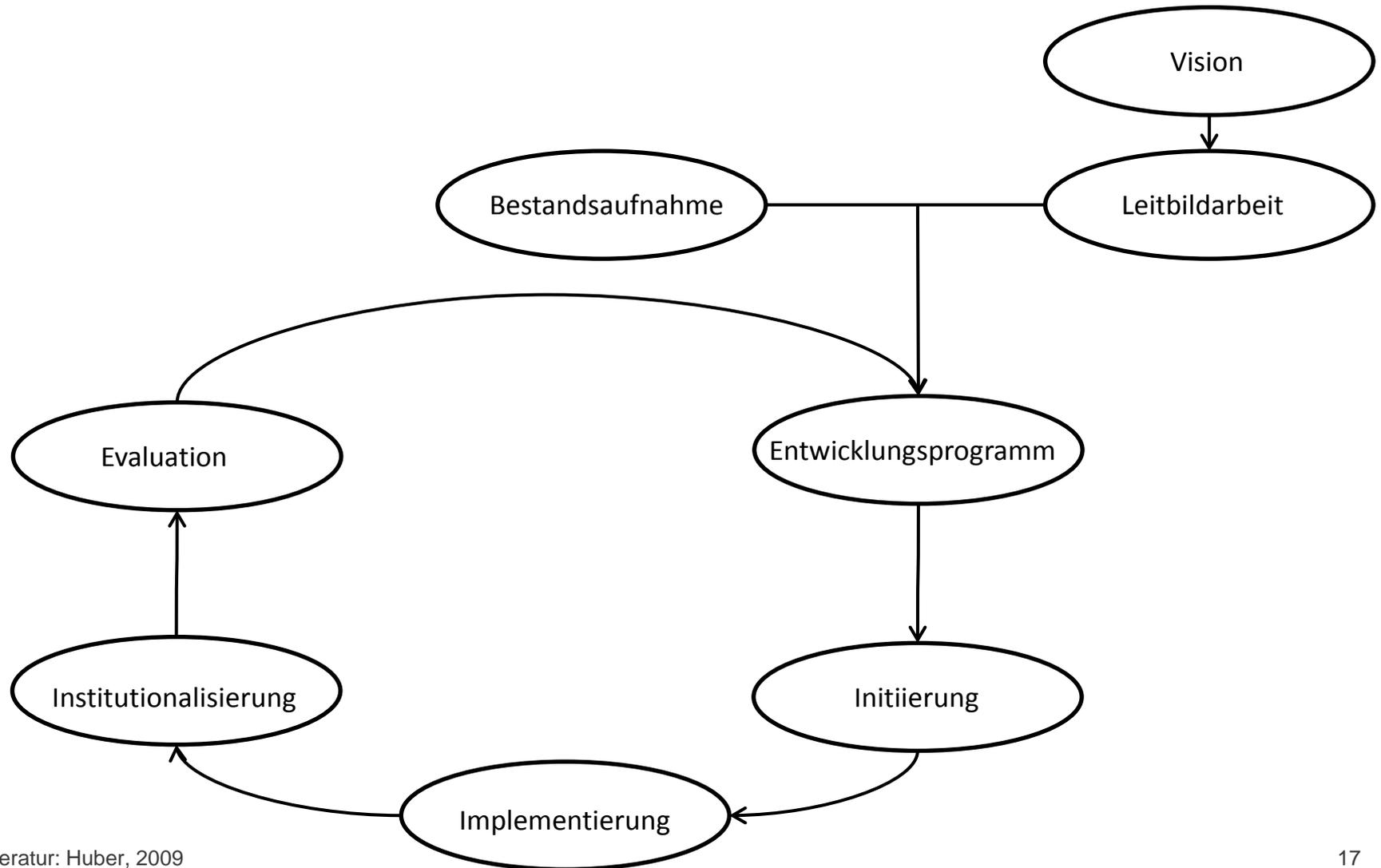


Was ist Unterrichtsentwicklung?

„Unter Unterrichtsentwicklung wollen wir die Gesamtheit der systematischen Anstrengungen verstehen, die darauf ausgerichtet sind die Unterrichtspraxis (...) zu optimieren. (...) **Das grundlegende Ziel der Unterrichtsentwicklung ist die Effektivierung des Lernens der Schüler/innen** in allen Dimensionen: Unterrichtsentwicklung ist **eine Aufgabe** nicht nur der individuellen, sondern auch **organisationalen Lernens**. Die Lehrkräfte einer Schule müssen sich **über ihre Vorstellung von Unterricht verständigen, die für die Realisierung notwendigen Schritte vereinbaren** und die **Kriterien definieren, anhand derer sie den Erfolg ihrer gemeinsamen Anstrengungen messen wollen.**“

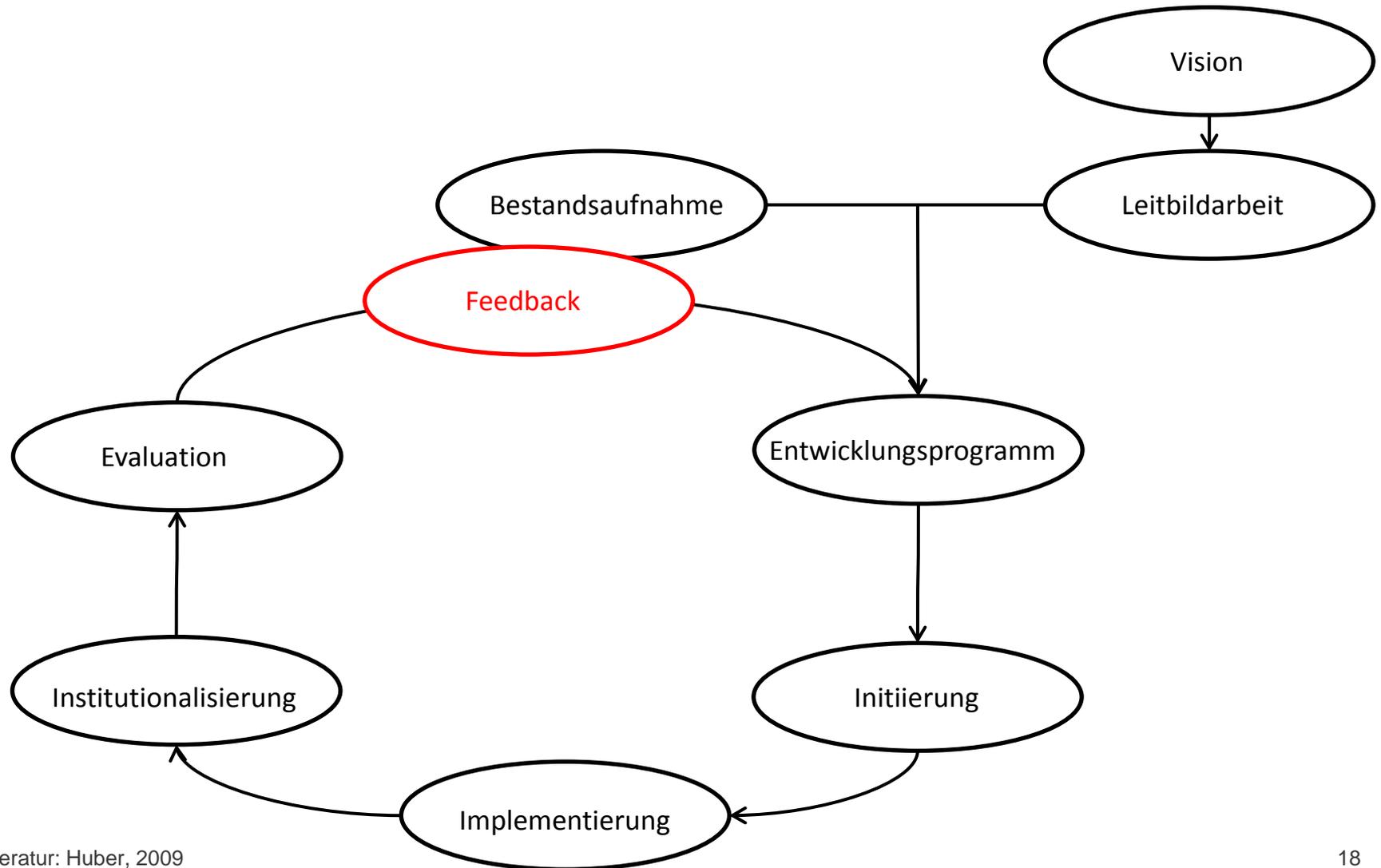


Kreislauf schulischer Qualitätssicherung





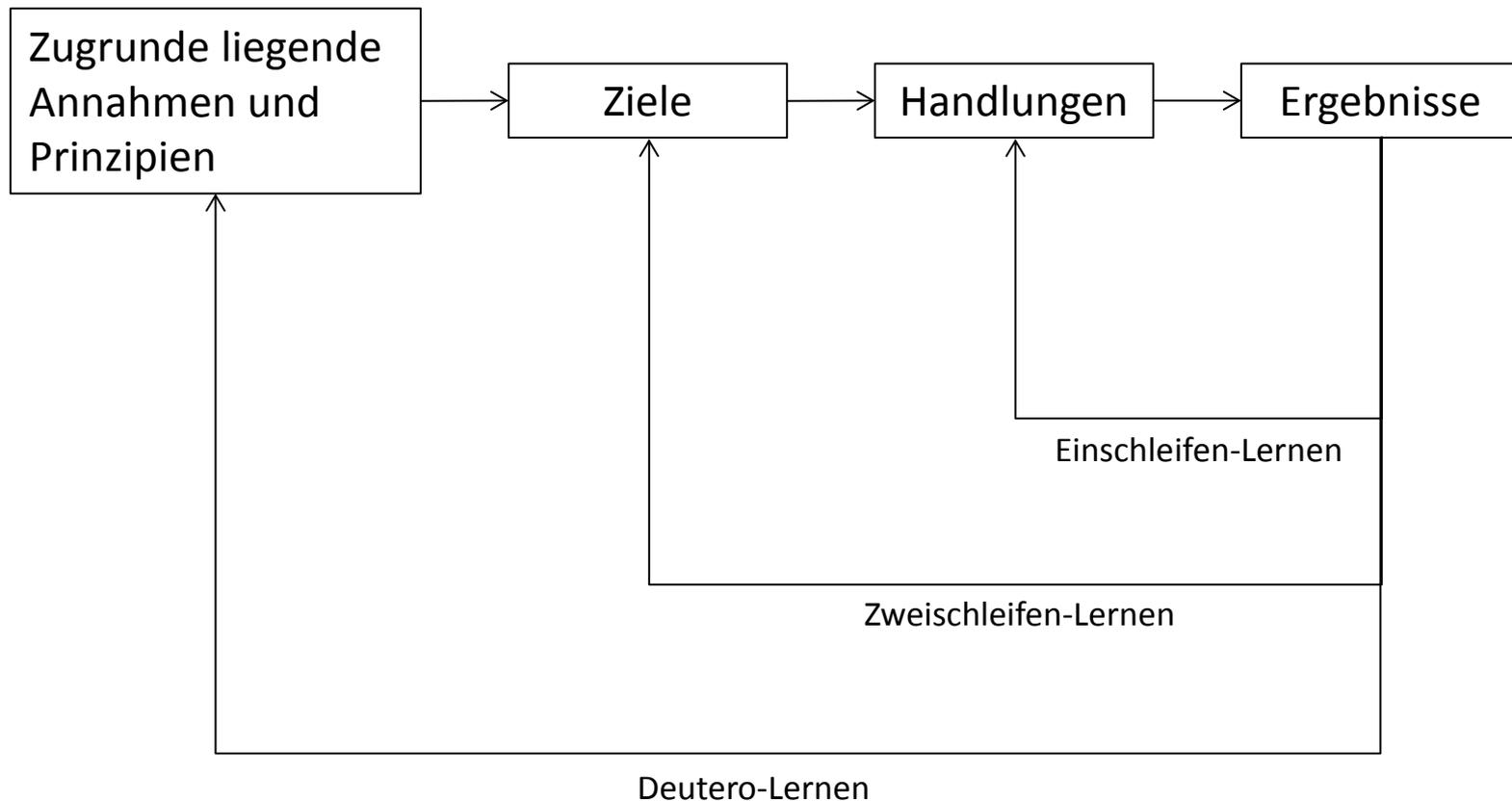
Kreislauf schulischer Qualitätssicherung





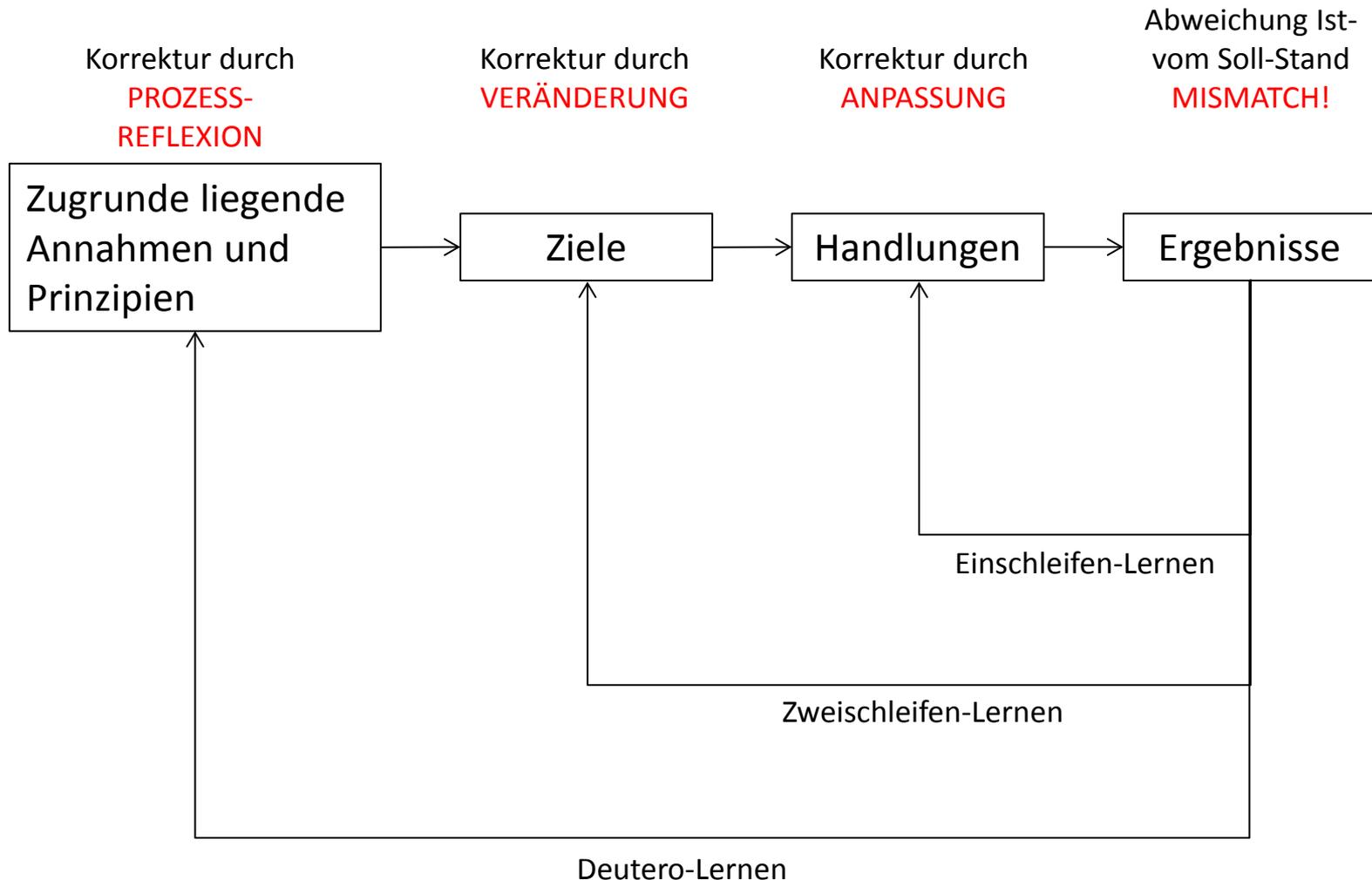
Organisationales Lernen durch Evaluation und Feedback

Abweichung Ist-
vom Soll-Stand
MISMATCH!





Organisationales Lernen durch Evaluation und Feedback





Was wissen wir über
wirksames Feedback?



Feedback aus Evaluationen soll evidenzbasierte Entscheidungen ermöglichen

„According to the rational ideal, evaluation would have a natural place in a clear-cut decision-making setting, where scientific methods and scientific knowledge are used to guide (...) decision making.“

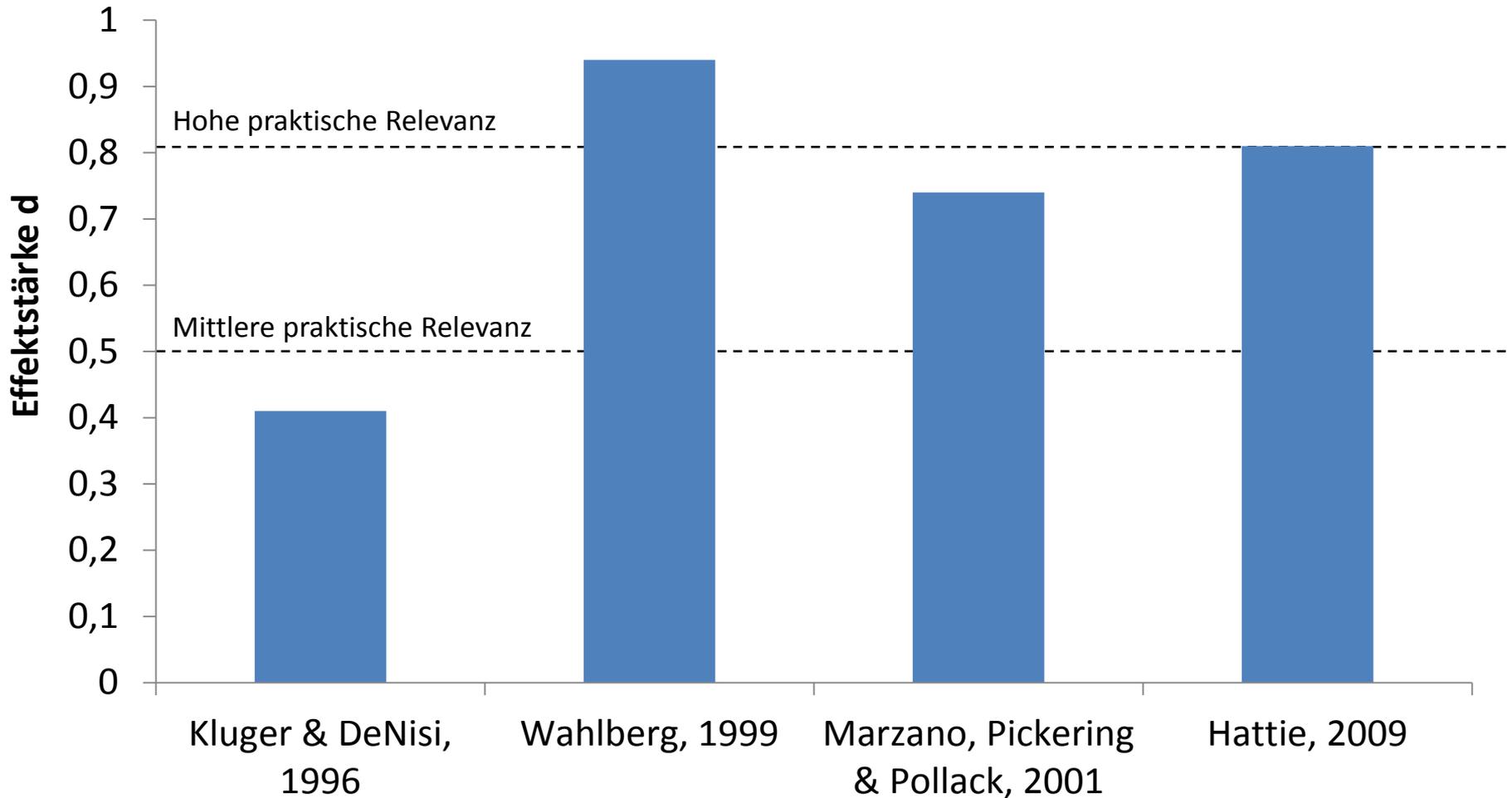


Der Entscheidungsprozess

- 1. Identifizierung eines Problems:** Bestimmung von Abweichungen zwischen Ist- und Soll-Stand eines Merkmals.
- 2. Diagnose des Problems:** Sammeln und Analysieren von Informationen, die die Art des Problems näher beschreiben.
- 3. Bestimmung von Alternativen:** Entwicklung möglichst aller Alternativen, mit deren Hilfe das Problem gelöst werden kann.
- 4. Überlegungen zu erwartbaren Konsequenzen anstellen:** Antizipieren der potenziellen Effekte der eruierten Alternativen.
- 5. Treffen einer Entscheidung:** Auswahl der besten Alternative, derjenigen, mit der das anzustrebende Ziele am besten zu erreichen ist.
- 6. Umsetzen der Entscheidung:** Umsetzen der Entscheidung, Implementierung von Veränderungen.

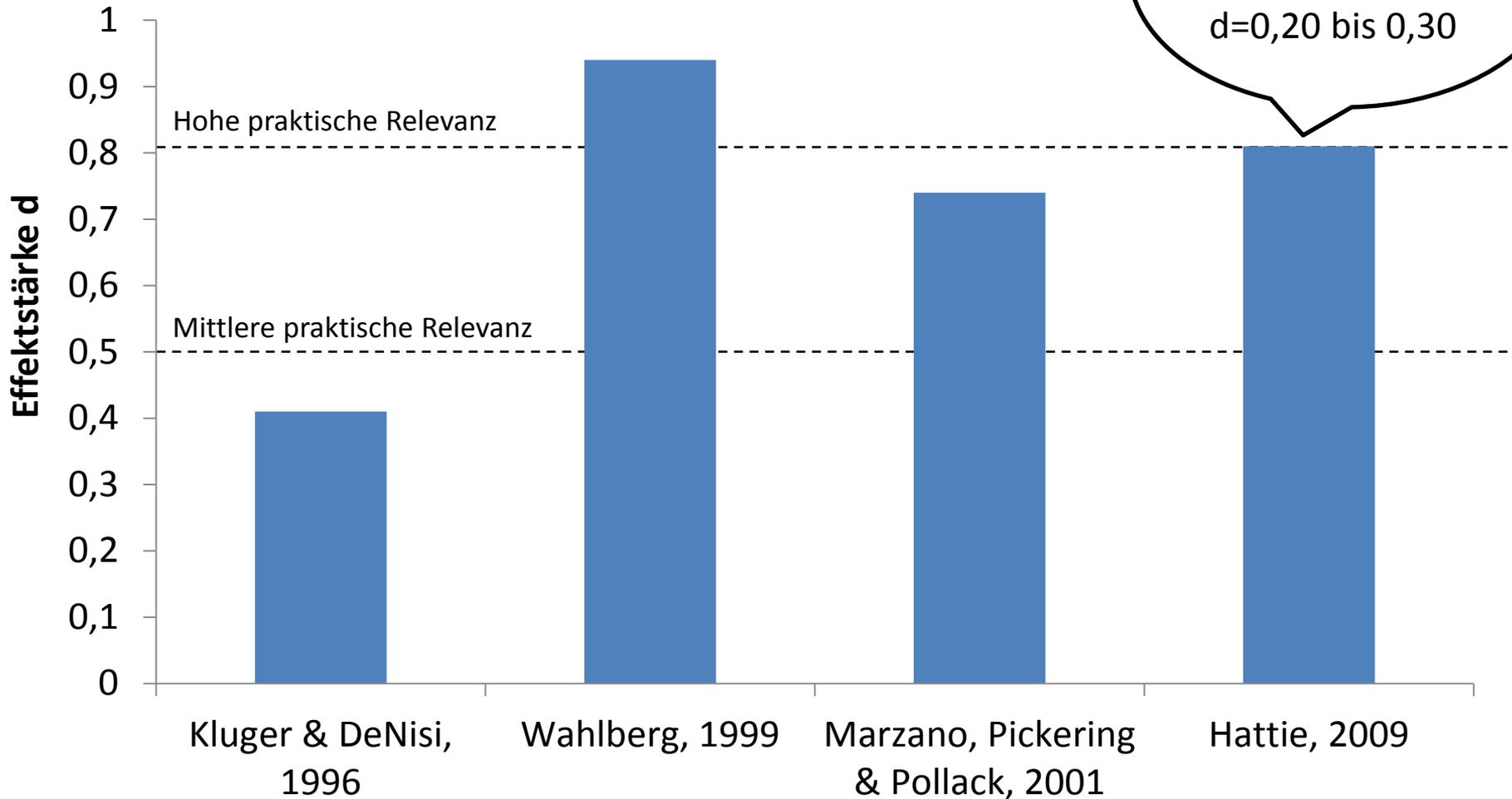


Gut gemachtes Feedback ist sehr wirksam!



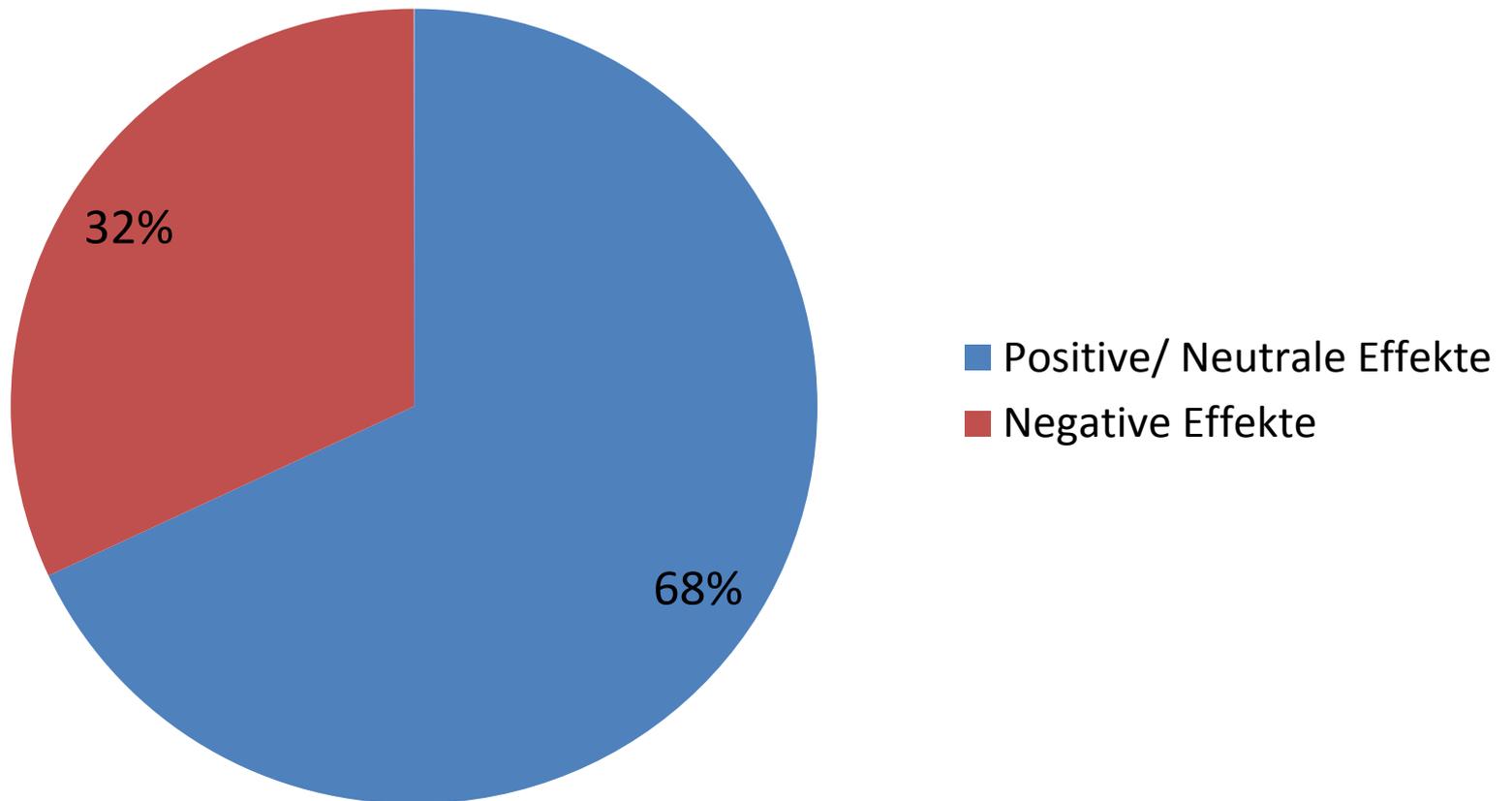


Gut gemachtes Feedback ist sehr wirksam!



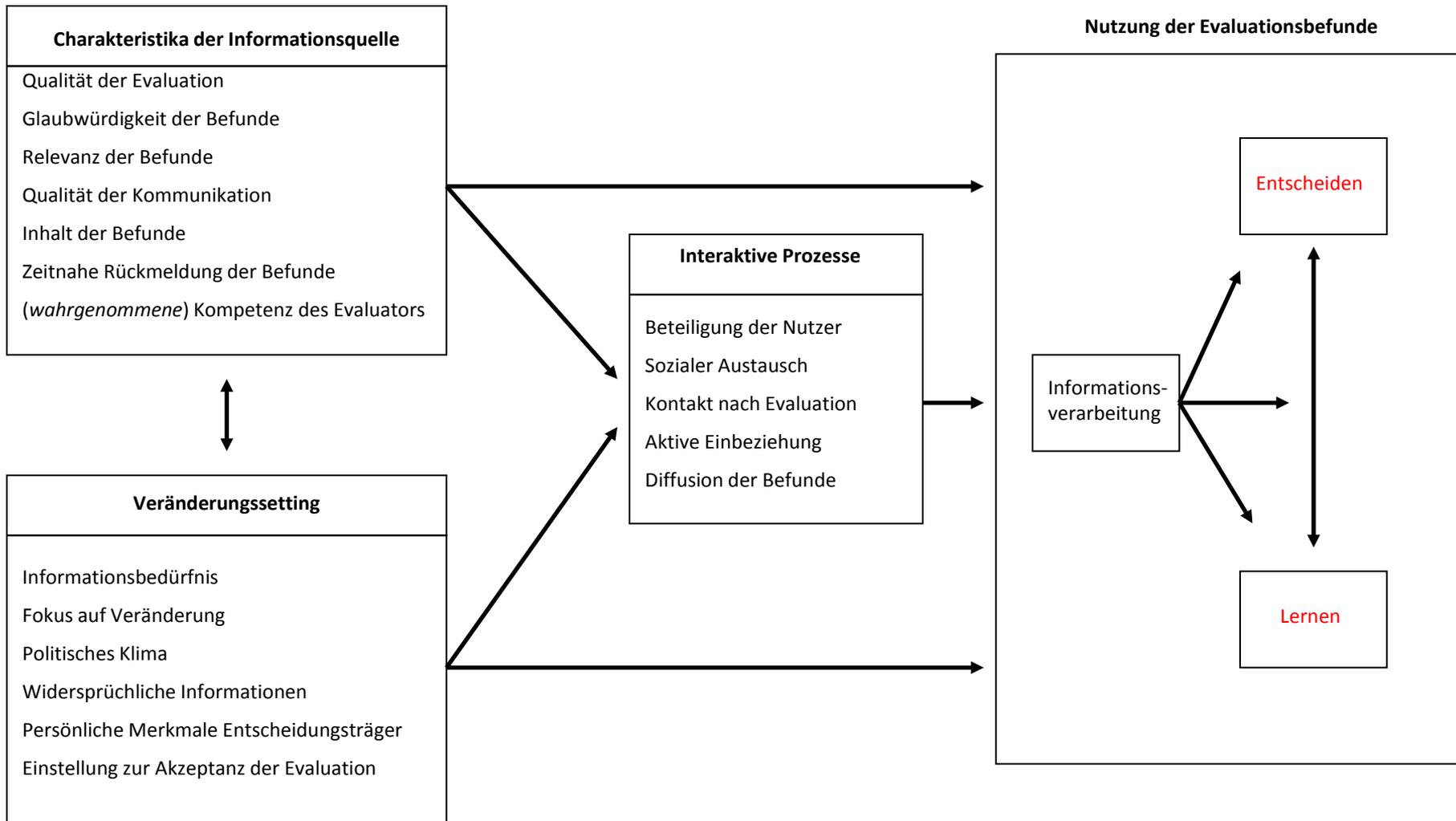


Es kann aber auch daneben gehen!





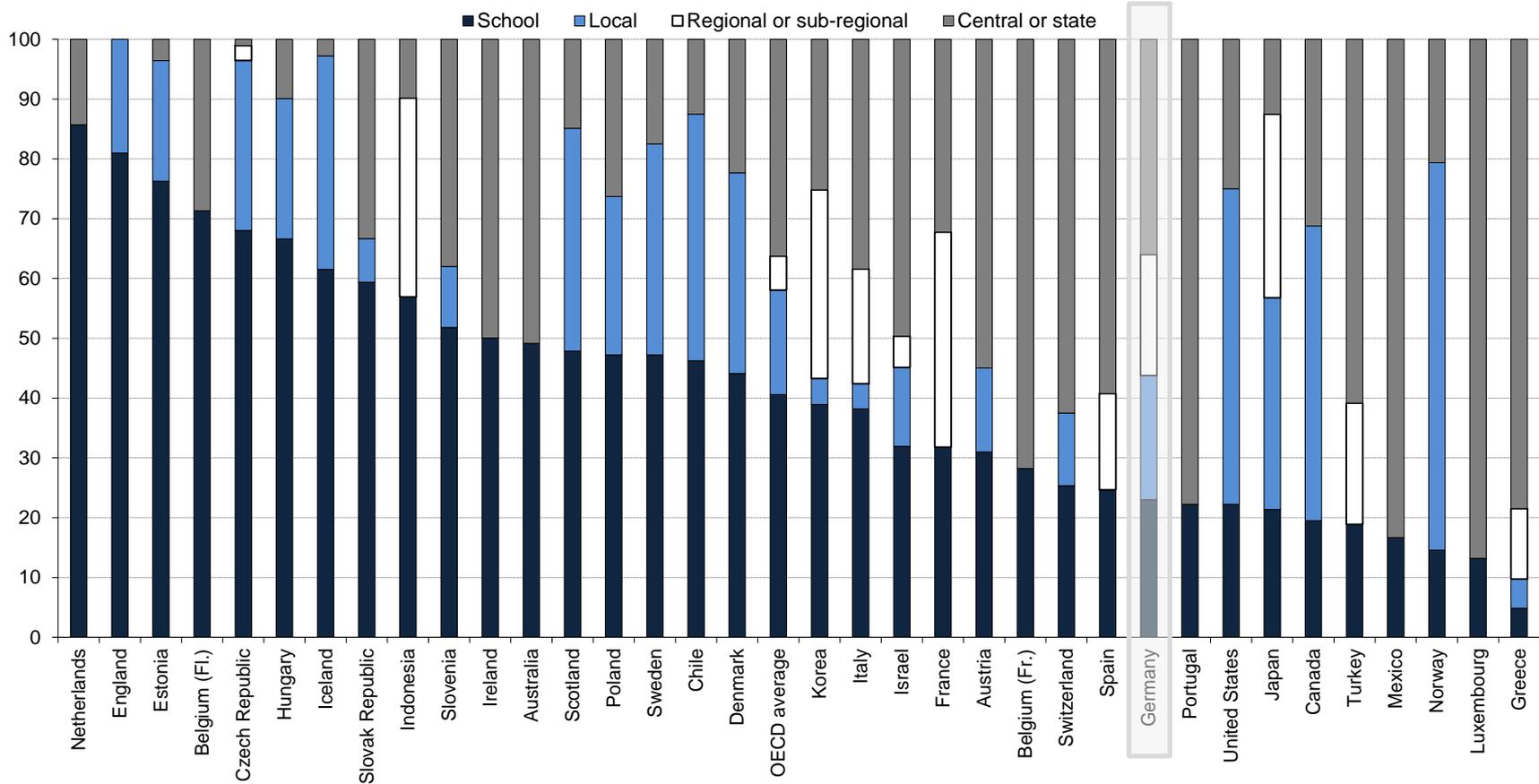
Feedback ermöglicht nur bedingte Entscheidungen!





Aber wie viel können Schulverantwortliche überhaupt beeinflussen?

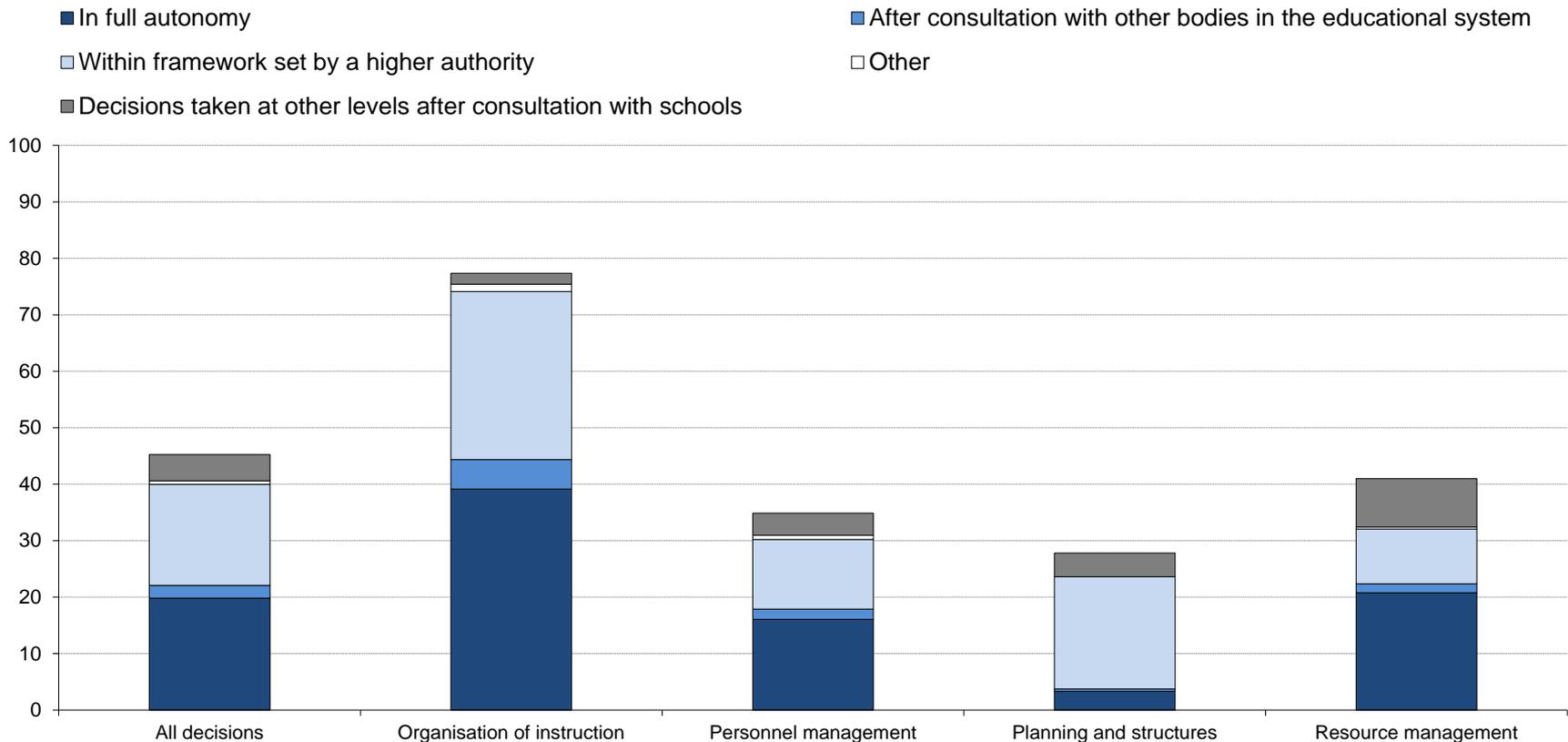
Chart D6.1 Percentage of decisions taken at each level of government in public lower secondary education (2011)





Jedoch: Der Unterricht liegt in der Hand der Schule!

Chart D6.4 Decisions taken at the school level in public lower secondary education in OECD countries, by mode of decision making and domain (2011)

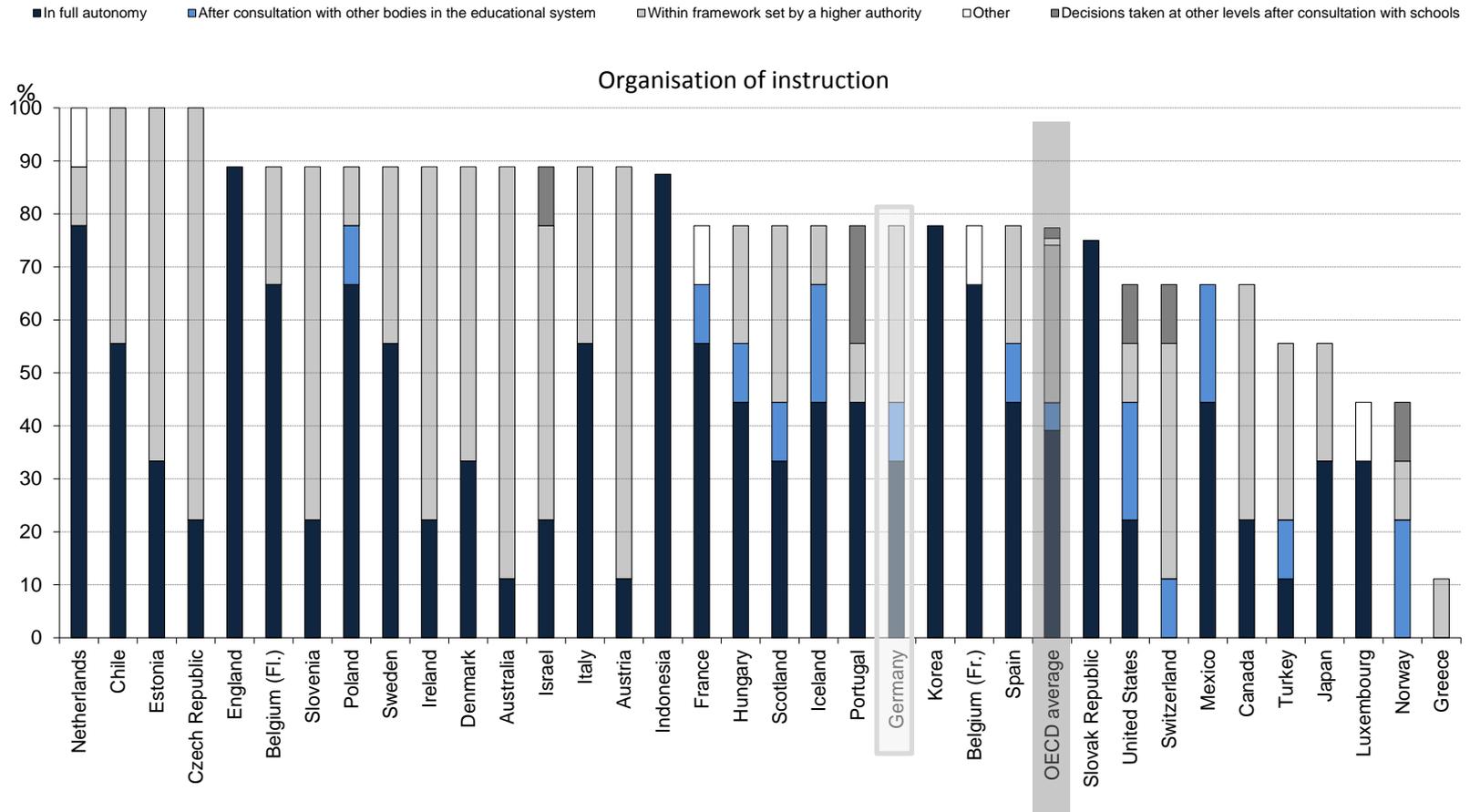


Source: OECD. Table D6.4a and D6.4b. See Annex 3 for notes (www.oecd.org/edu/eag2012).



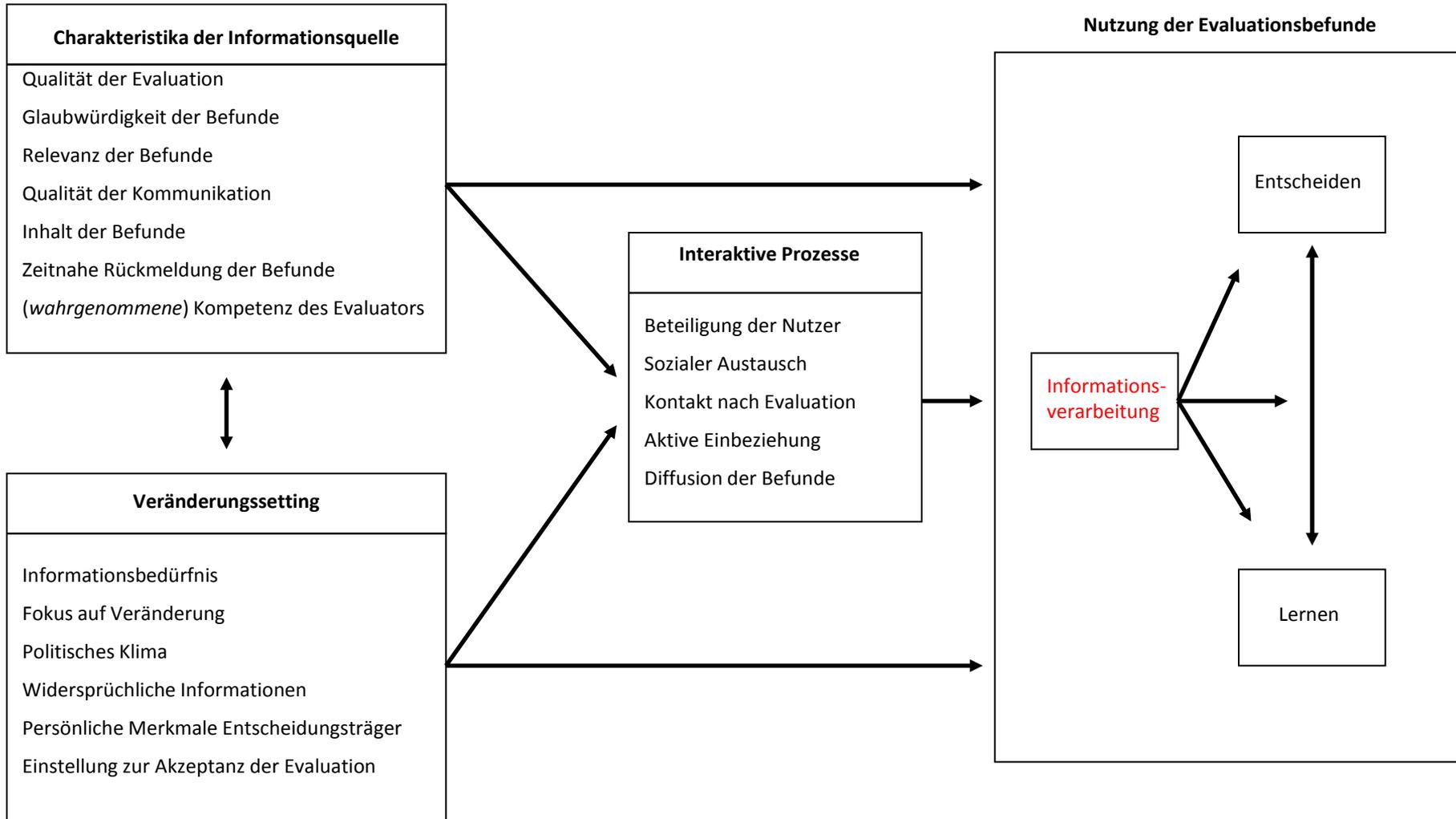
Deutschland liegt hier im OECD-Mittel

Chart D6.7. Decisions taken at the school level in public lower secondary education, by mode of decision making and domain (2011)





Feedback ermöglicht nur bedingte Entscheidungen!





Feedback wird intendiert genutzt...

Nutzeraktivität

Ergebnis

Einfluss auf
Informations-
verarbeitung

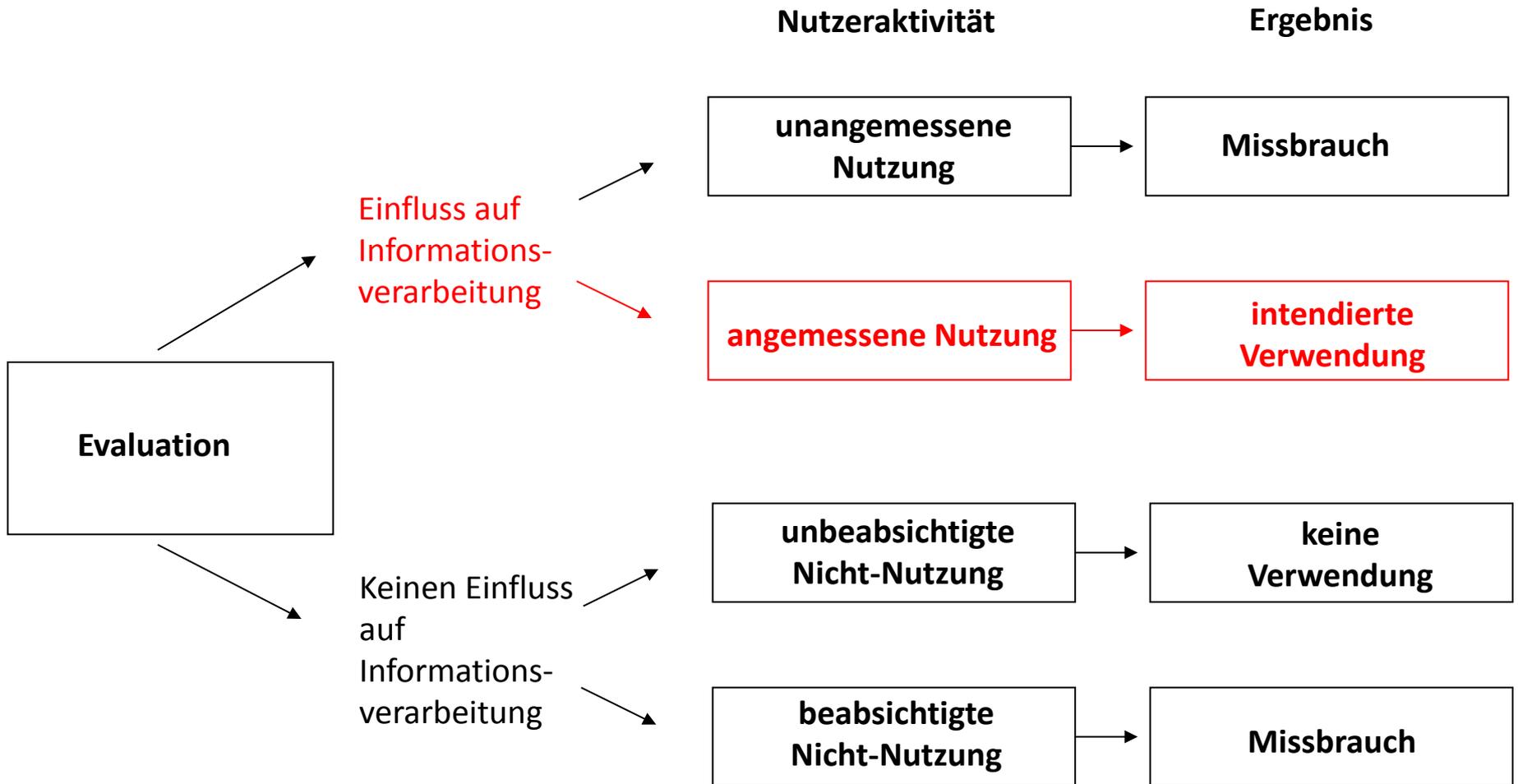
Evaluation

angemessene Nutzung

**intendierte
Verwendung**

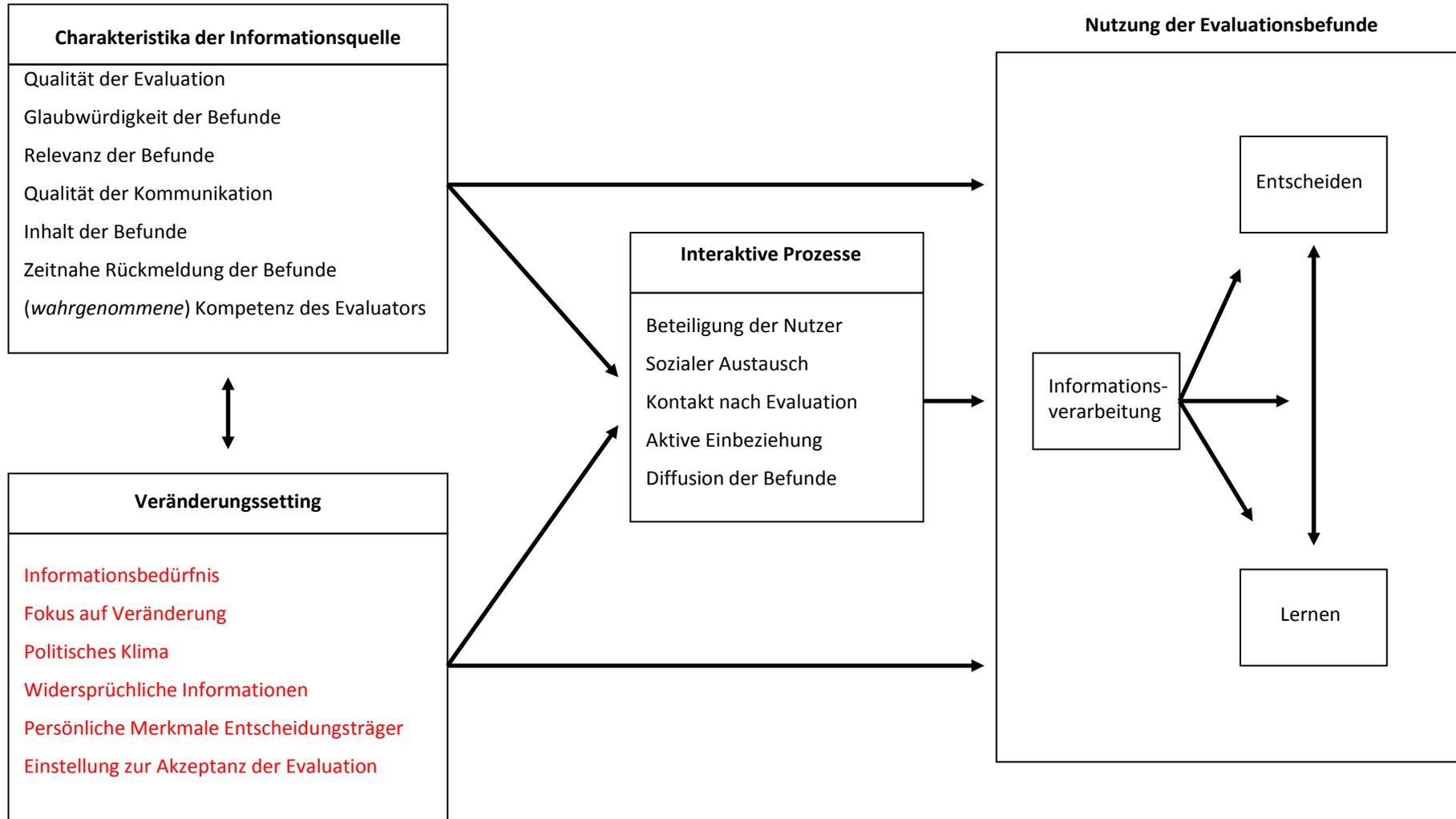


... oder auch nicht



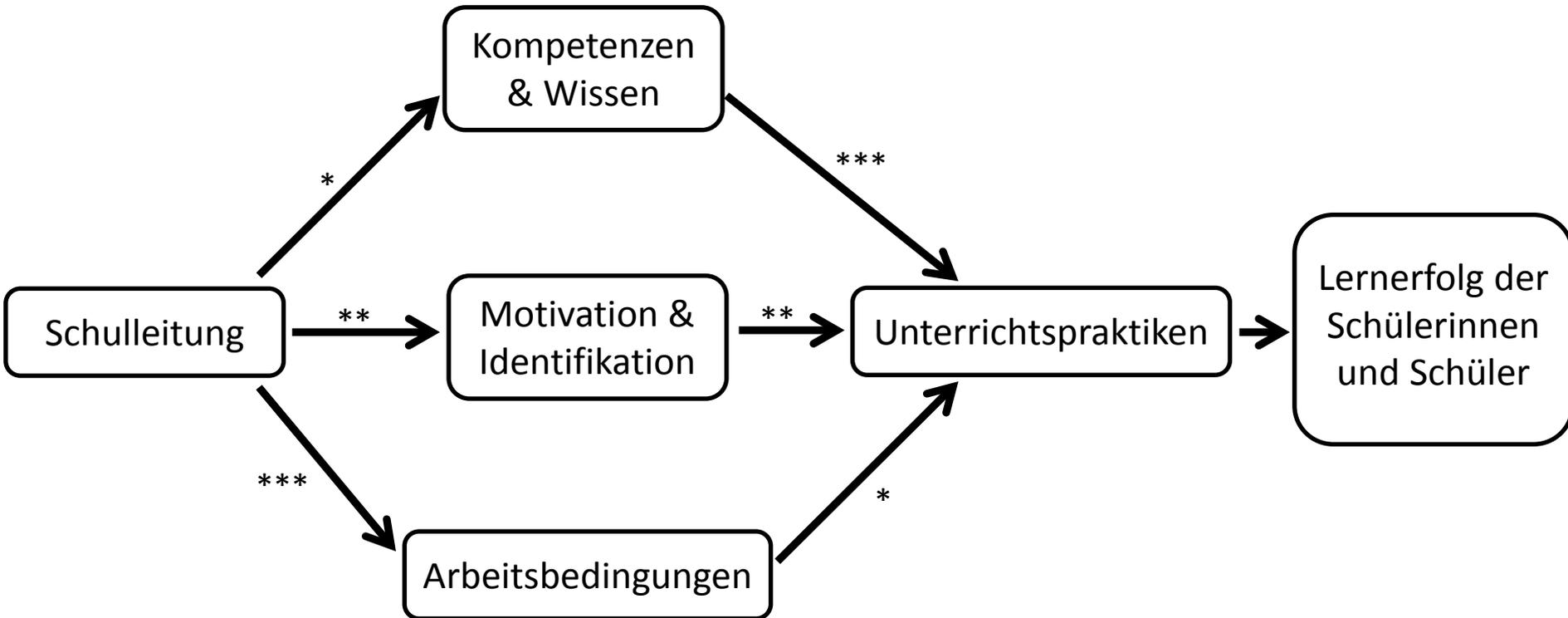


Feedback ermöglicht nur bedingte Entscheidungen!





Schulleitungen als zentrale Akteure



* = schwache Beeinflussung
** = mittlere Beeinflussung
*** = starke Beeinflussung



Was ist besonders wichtig?

Führungsdimensionen	Effektstärke		
	r	d	Odds Ratio
Schulleitung als (aktiver) Experte für Unterrichtsfragen	0,39	0,84	4,59
Aktive Teilnahme an Planung, Koordination und Evaluation von Unterricht und Curriculum	0,21	0,42	2,14
Setzung hoher Ziele und Erwartungen	0,21	0,42	2,14
Strategischer, am Ergebnis orientierter Einsatz von Ressourcen	0,15	0,31	1,75
Sicherung einer geregelten und Lehr-Lern-förderlichen Schulumwelt	0,13	0,27	1,63



Was ist besonders wichtig?

Ein Lernjahr im Sekundarbereich:
 $d=0,20$ bis $0,30$

Führungsdimensionen	Effekte		
	r	d	Odds Ratio
Schulleitung als (aktiver) Experte für Unterrichtsfragen	0,39	0,84	4,59
Aktive Teilnahme an Planung, Koordination und Evaluation von Unterricht und Curriculum	0,21	0,42	2,14
Setzung hoher Ziele und Erwartungen	0,21	0,42	2,14
Strategischer, am Ergebnis orientierter Einsatz von Ressourcen	0,15	0,31	1,75
Sicherung einer geregelten und Lehr-Lern-förderlichen Schulumwelt	0,13	0,27	1,63



Nutzen Schulleitungen ihre Möglichkeiten?

	Hamburg	GEW-TALIS	OECD-TALIS
Die Schulleiterin/ Der Schulleiter stellt sicher, dass die Arbeit der Lehrerinnen und Lehrer mit den Lehrzielen der Schule übereinstimmt.	66	31	76
Die Schulleiterin/ Der Schulleiter stellt sicher, dass Lehrerinnen und Lehrer über die Möglichkeiten informiert sind, ihr Wissen und ihre Fähigkeiten auf den neuesten Stand zu bringen.	72	44	70
Wenn eine Lehrerin/ ein Lehrer Probleme in ihrer/ seiner Klasse hat, ergreift die Schulleiterin/ der Schulleiter die Initiative, um darüber zu sprechen.	57	30	54
Die Schulleiterin/ Der Schulleiter gibt den Lehrerinnen und Lehrern Anregungen, wie sie ihren Unterricht verbessern können.	31	10	35
Die Schulleiterin/ Der Schulleiter oder jemand anderes aus dem Leitungsteam hospitiert im Unterricht.	26	5	34

Summe aus „eher oft“ und „sehr oft“ in Prozent

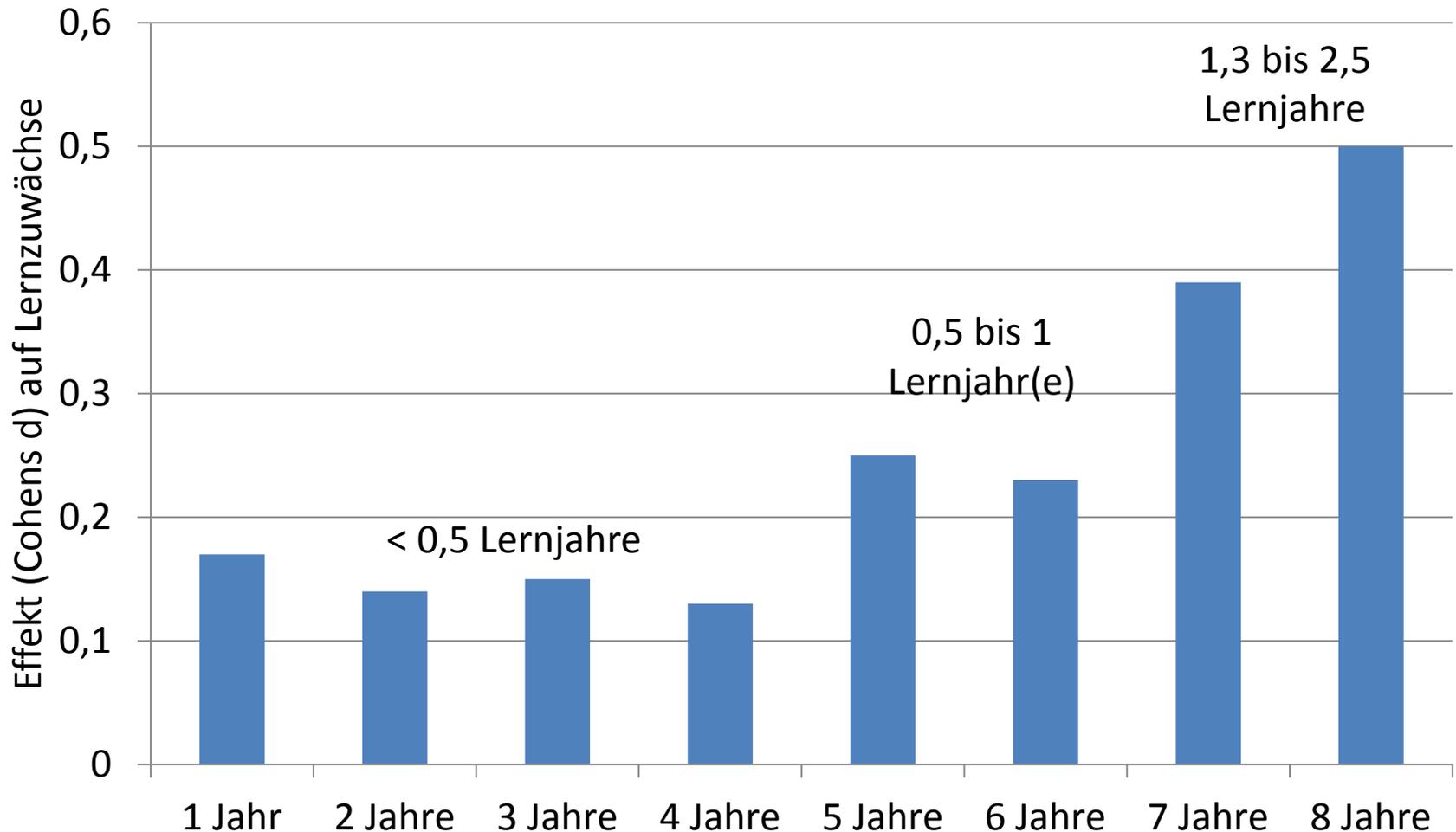


Daher: Strukturierte, theoriegeleitete Veränderungen

Entwicklung durch...	Schultyp	Anteil (in %)	sinnvolle Veränderungsstrategie
Systemische Steuerung	Sand Schools <ul style="list-style-type: none">- Schulen mit guten Schülerleistungen, die unter ihren Möglichkeiten bleiben- Schulen in Krisen/ Umbrüchen	ca. 5%	Externer Eingriff in die Schule; vor allem extreme Schwächen abstellen; Kurzfristziele setzen; für Veränderungen motivieren
Schul- und Schulnetzwerk-basierte Reformen	Seed Schools <ul style="list-style-type: none">- Besondere Kapazitäten zur Umsetzung eigener Visionen- Herausragende Führung	ca. 5%	Selbstgesteuerte Bottom-up-Entwicklung mit Critical Friends
	Brick Schools <ul style="list-style-type: none">- Bereitschaft für Veränderungen- Geringe Kapazität für eigene pädagogische Visionen	ca. 90%	Bereitstellung von Blaupausen für die Entwicklung; z.B. durch ganzheitliche (CSR) oder fachbezogene (Leseförderung) Maßnahmen; Begleitung der Entwicklung

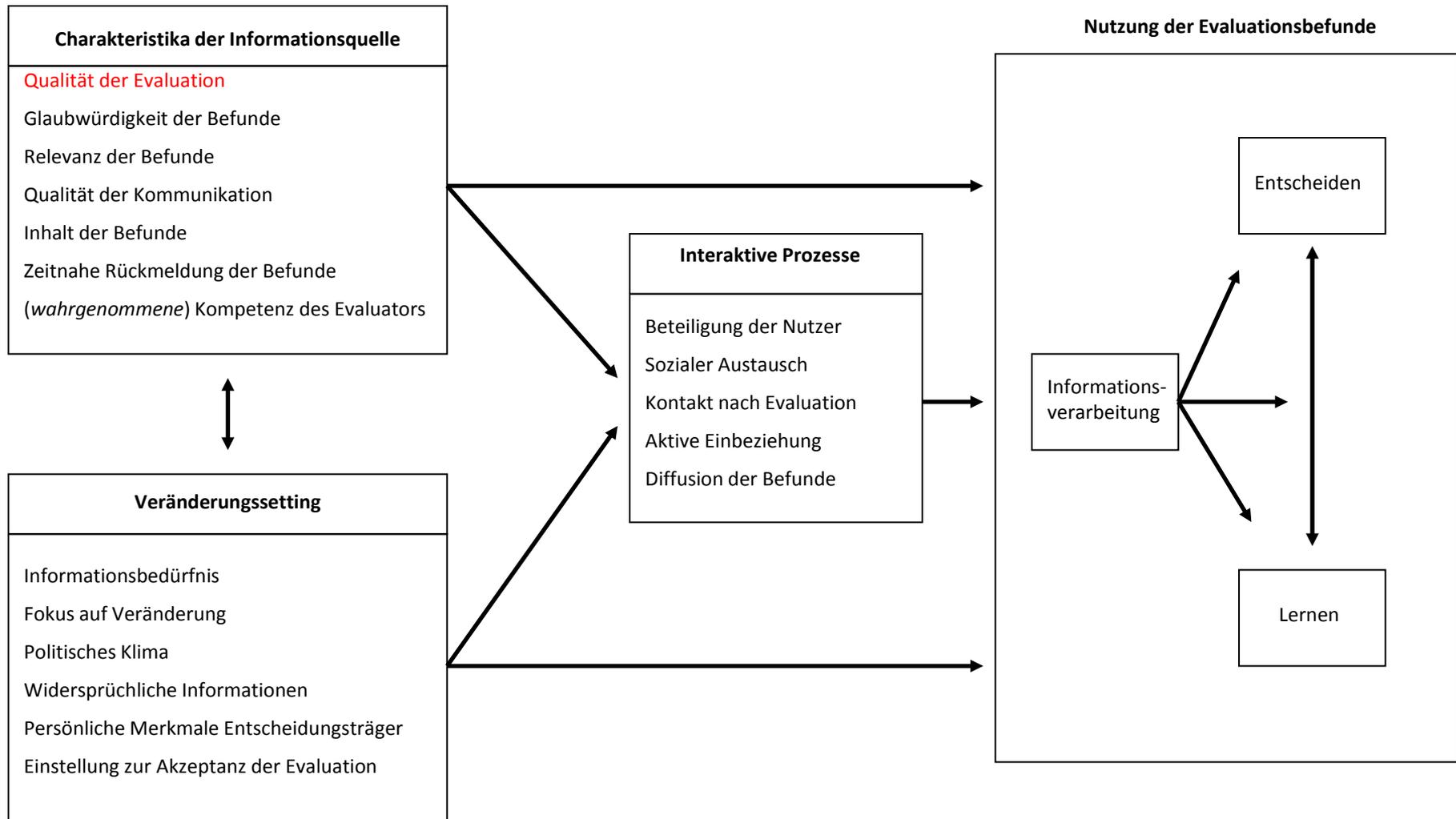


Diese Maßnahmen benötigen jedoch Zeit um wirksam zu werden



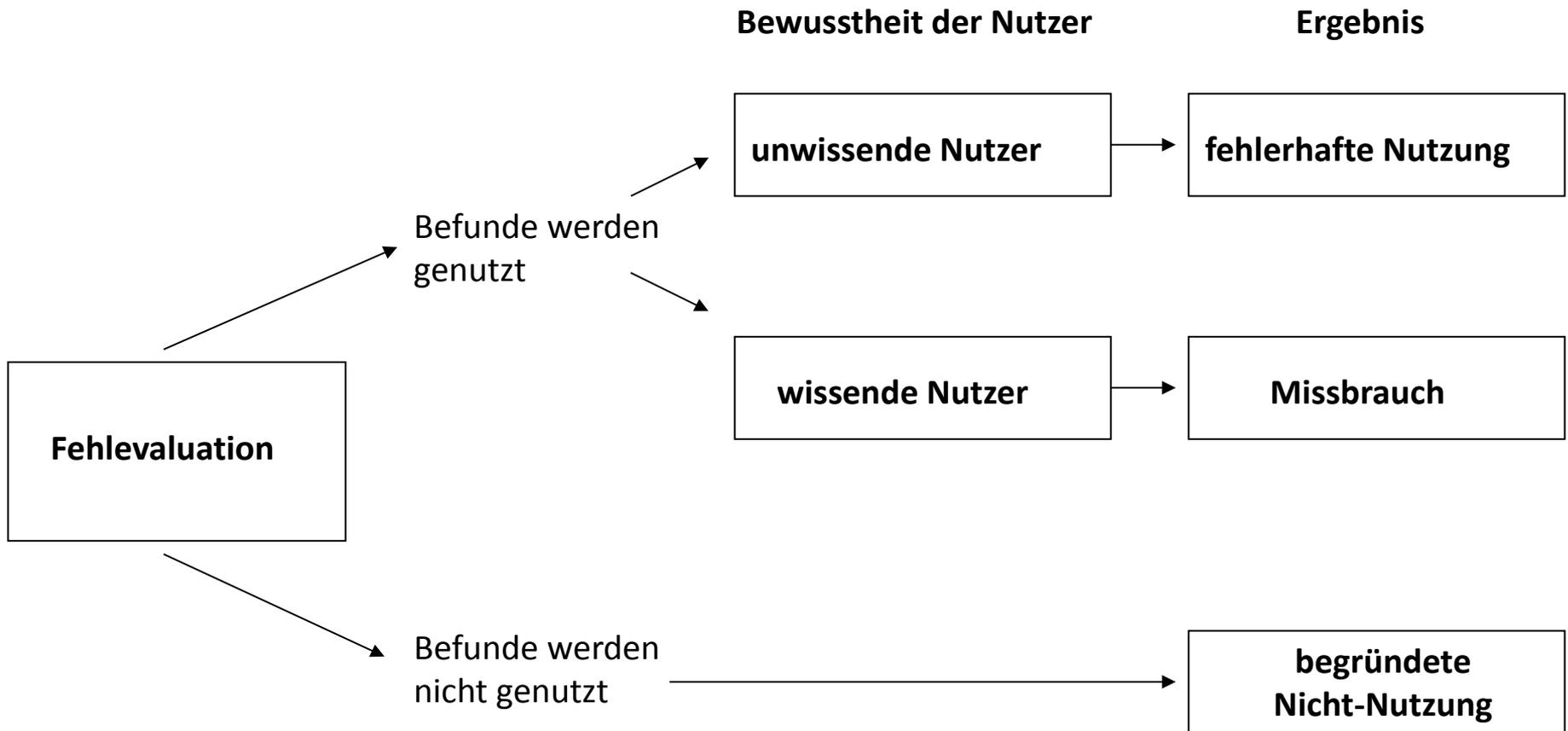


Bedingungsfaktoren für Feedback-Wirksamkeit



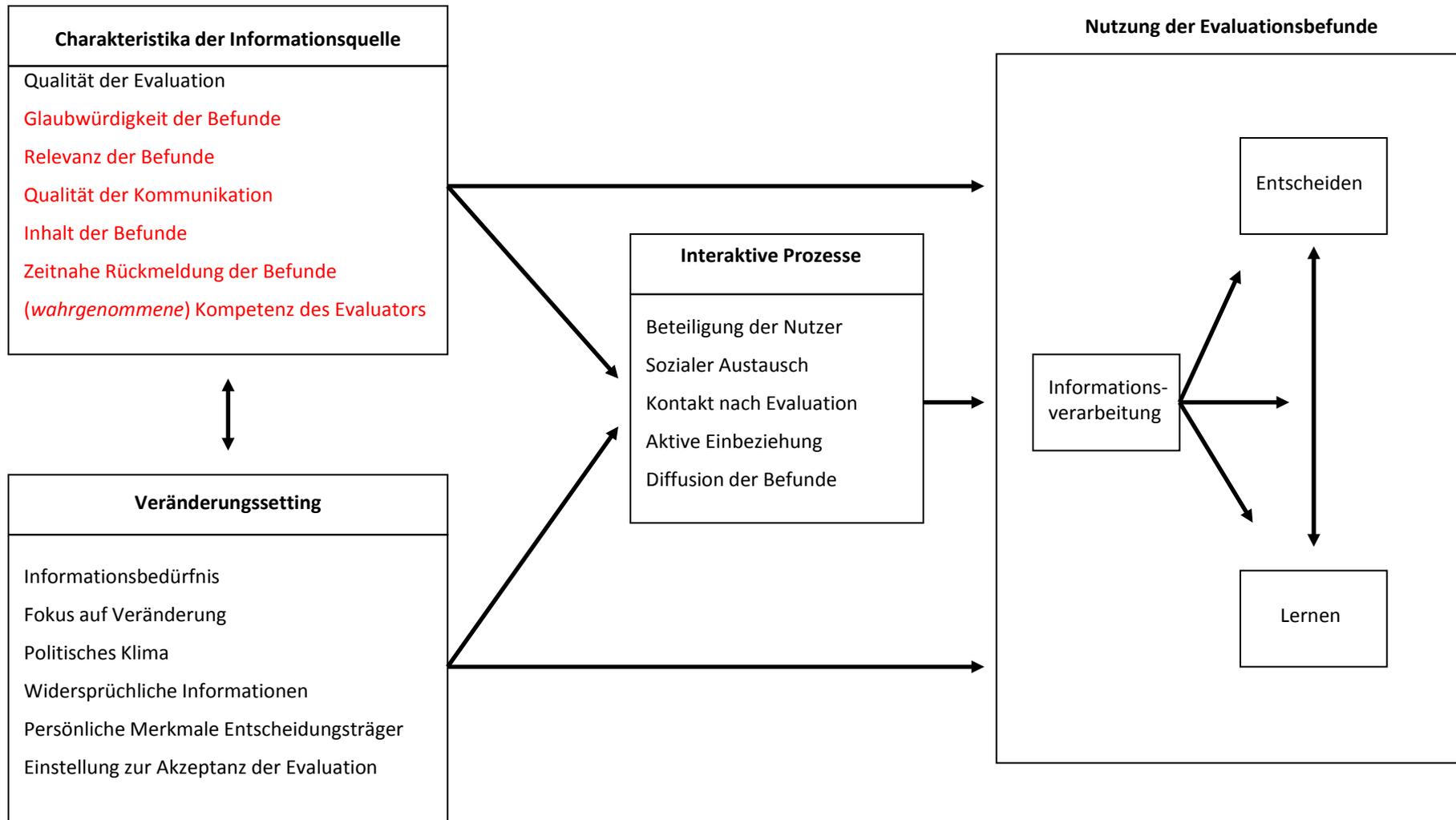


Ist eine Evaluation fehlerhaft, werden intendierte Wirkungen nicht erreicht





Bedingungsfaktoren für Feedback-Wirksamkeit





Welche Art von Feedback ist wirksam?

Variable	Anzahl der Meta- Analysen	Anzahl der Studien	Anzahl der Effekte	Effektstärke
Hinweise	3	89	129	1,1
Feedback	74	4,157	5,755	0,95
Bestätigung	1	19	19	0,94
Ziele und Feedback	8	640	121	0,46
Korrigierendes Feedback	25	1,149	1,040	0,37
Verspätetes versus Sofortiges	5	178	83	0,34
Belohnung	3	223	508	0,31
Bestrafung	1	89	210	0,2
Lob	11	388	4,410	0,14



Welche Art von Feedback ist wirksam?

Moderator	Effektstärke	Lernjahre
Richtiges Feedback		
Es ist korrekt	0,43	1,4-2,2
Es ist inkorrekt	0,25	0,8-1,3
Aufgabenfeedback über Veränderungen zu vorherigen Evaluation		
Ja	0,55	1,8-2,8
Nein	0,28	0,9-1,4
Demotivierendes Aufgabenfeedback		
Ja	-0,14	-0,7—0,5
Nein	0,33	1,1-1,7
Lobendes Feedback für die Aufgabe		
Ja	0,09	0,3-0,5
Nein	0,34	1,1-1,7
Häufigkeit des gegebenen Feedback		
Häufig	0,32	1,1-1,6
Selten	0,39	1,3-2,0
Komplexität der Aufgaben		
Sehr komplex	0,03	0,1-0,2
Nicht komplex	0,55	1,8-2,8
Zielsetzung		
Schwierige Ziele	0,51	1,7-2,6
Einfache, gib dein Bestes, Ziele	0,30	1,0-1,5
Bedrohung des Selbstwertgefühls		
Große Bedrohung	0,08	0,3-0,4
Kleine Bedrohung	0,47	1,6-2,4



Welche Feedbackkomponenten sind wirksam?

Moderator	Effektstärke	Lernjahre
Richtiges Feedback		
Es ist korrekt	0,43	1,4-2,2
Es ist inkorrekt	0,25	0,8-1,3
Aufgabenfeedback über Veränderungen zu vorherigen Evaluation		
Ja	0,55	1,8-2,8
Nein	0,28	0,9-1,4
Demotivierendes Aufgabenfeedback		
Ja	-0,14	-0,7—0,5
Nein	0,33	1,1-1,7
Lobendes Feedback für die Aufgabe		
Ja	0,09	0,3-0,5
Nein	0,34	1,1-1,7
Häufigkeit des gegebenen Feedback		
Häufig	0,32	1,1-1,6
Selten	0,39	1,3-2,0
Komplexität der Aufgaben		
Sehr komplex	0,03	0,1-0,2
Nicht komplex	0,55	1,8-2,8
Zielsetzung		
Schwierige Ziele	0,51	1,7-2,6
Einfache, gib dein Bestes, Ziele	0,30	1,0-1,5
Bedrohung des Selbstwertgefühls		
Große Bedrohung	0,08	0,3-0,4
Kleine Bedrohung	0,47	1,6-2,4



Was macht somit ein wirksames Feedback aus?

- Feedback benötigt Information dazu
 - welches Ziel grundsätzlich angestrebt wird (**Feed-Up**),
 - wie weit man auf dem Weg zur Zielerreichung bereits fortgeschritten ist (**Feed-Back**)
 - und welche Schritte auf dem Weg zur Zielerreichung als nächstes vollzogen werden sollten (**Feed-Forward**).
- Darüber hinaus
 - dürfen die Ziele nicht zu komplex und indifferent sein,
 - sind Hinweise (*cues*) nötig, was man machen muss um sich weiter zu entwickeln,
 - und muss es positiv und wertschätzend formuliert sein, um das Selbstwertgefühl nicht zu bedrohen.



Was wissen wir über
wirksamen Unterricht?



Erklärungen für Unterschiede in Lernzuwächsen

Was schadet?

Sitzenbleiben: $d = -.17$

Sommerferien: $d = -.09$

Was schadet und hilft nicht?

Offener Unterricht: $d = .01$

Jahrgangsübergreifender Unterricht:
 $d = .04$

Was hilft ein wenig?

Individualisiertes Lernen: $d = .22$

Störungsprävention: $d = .34$

Was hilft mehr?

Gutes Klassenmanagement: $d = .52$

Herausfordernde Ziele setzen: $d = .58$

Was hilft sehr?

Lehrerfeedback: $d = .73$

Klarheit der Instruktion: $d = .75$



Es zählt die unterrichtliche Tiefenstruktur!

<i>Teaching</i>	<i>d</i>	<i>Working conditions</i>	<i>d</i>
Quality of teaching	.77	Within-class grouping	.28
Reciprocal teaching	.74	Adding more finances	.23
Teacher-student relationships	.72	Reducing class size	.21
Providing feedback	.72	Ability grouping	.12
Teaching student self-verbalization	.67	Multi-grade/age classes	.04
Meta-cognition strategies	.67	Open vs. tradit. classes	.01
Direct instruction	.59	Summer vacation classes	-.09
Mastery learning	.57	Retention	-.16
<i>Average</i>	<i>.68</i>	<i>Average</i>	<i>.08</i>



Kernmerkmale wirksamen Unterrichts

	Helmke (2006)	Meyer (2004)
1	effiziente Klassenführung und Zeitnutzung	hoher Anteil echter Lernzeit
2	lernförderliches Unterrichtsklima	lernförderliches Klima
3	vielfältige Motivierung	klare Strukturierung
4	Strukturiertheit und Klarheit	inhaltliche Klarheit
5	Schülerorientierung, Unterstützung	sinnstiftendes Kommunizieren
6	angemessene Variation von Methoden und Sozialformen	angemessene Methodenvielfalt
7	Förderung aktiven, selbständigen Lernens	individuelles Fördern
8	Konsolidierung, Sicherung, Intelligentes Üben	intelligentes Üben
9	Wirkungs- und Kompetenzorientierung	transparente Leistungserwartungen
10	Passung	vorbereitete Umgebung

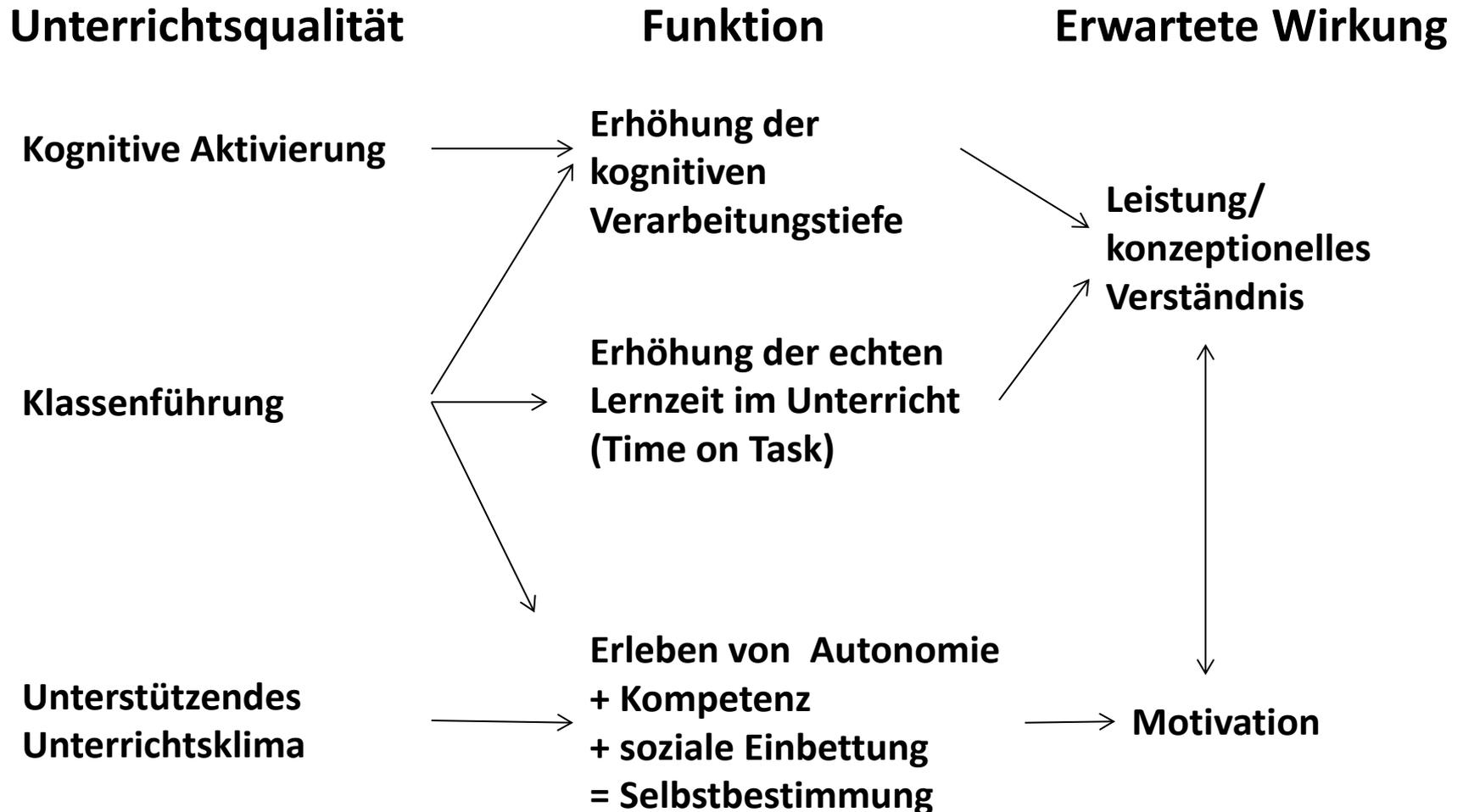


Verringerung der Komplexität!

- Der Wirkungsgrad einzelner Merkmale erfolgreichen Unterrichts ist in der Regel begrenzt.
- Einzelne Qualitätsmerkmale erfolgreichen Unterrichts bilden ein Netzwerk sich gegenseitig unterstützender Faktoren.
- In verschiedenen Studien ließen sich entsprechend Basisdimensionen finden, die voneinander nicht unabhängig sind:
 - Strukturierte, klare und störungspräventive Unterrichtsführung,
 - unterstützendes, schülerorientiertes Unterrichtsklima,
 - kognitive Aktivierung.



Basisdimensionen, ihre Funktionen und deren Wirkung





Alle Basisdimensionen sind wichtig und hierarchisch

„Es ist deutlich, dass guter Unterricht in allen drei Grunddimensionen ausgewiesen sein muss. Jede Dimension erfüllt im Hinblick auf die Leistungs- und Motivationsentwicklung der Schüler und die Sicherung der Arbeitsbedingungen im Klassenverband wesentliche Funktionen. Es ist daher falsch, beispielsweise Schülerorientierung und direktes, störungspräventives Verhalten des Lehrers gegeneinander auszuspielen. [...] Das eine ist die Grundvoraussetzung, auf der kognitiv aktivierende Instruktionsprozesse aufbauen müssen, um erfolgreiches fachliches Verstehen zu ermöglichen.“



Eine gute Voraussetzung für Unterrichtsstandards

- „Unterrichtsstandards beschreiben Qualitätsstufen realen Unterrichts im definierten Qualitätsbereich, die auf Grundlage (vergleichender) empirischer Forschung gewonnen und auf Grundlage einer Unterrichtstheorie gewichtet worden sind.“
- „Unterrichtsstandards sollten den gleichen drei Ansprüchen genügen wie die Bildungsstandards:
 - (1) Sie müssen an ein theoretisches Modell der Unterrichtsqualität angedockt werden.
 - (2) Sie sollten in sich gestuft dargestellt werden.
 - (3) Und sie sollten standardisiert sein, d.h. in geeichte regional, national oder international gültige Messskalen übertragen worden sein.“



Die Basisdimensionen sind abstufbar

Stufe 3 = kognitive Aktivierung

Stufe 2 = Unterstützendes Unterrichtsklima

Stufe 1 = Klassenführung



Die Basisdimensionen sind abstuftbar

Stufe 4 = Lernautonomie

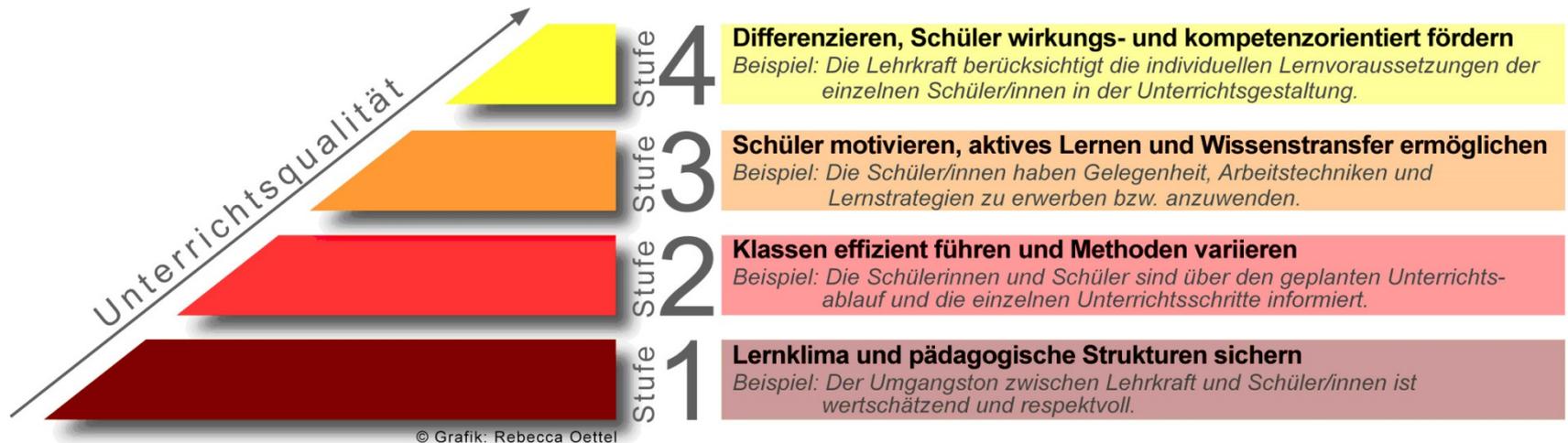
Stufe 3 = kognitive Aktivierung

Stufe 2 = Unterstützendes Unterrichtsklima

Stufe 1 = Klassenführung



...und lassen sich in ein praxisorientiertes Modell für Rückmeldungen überführen



Stufe 4 = Lernautonomie

Stufe 3 = kognitive Aktivierung

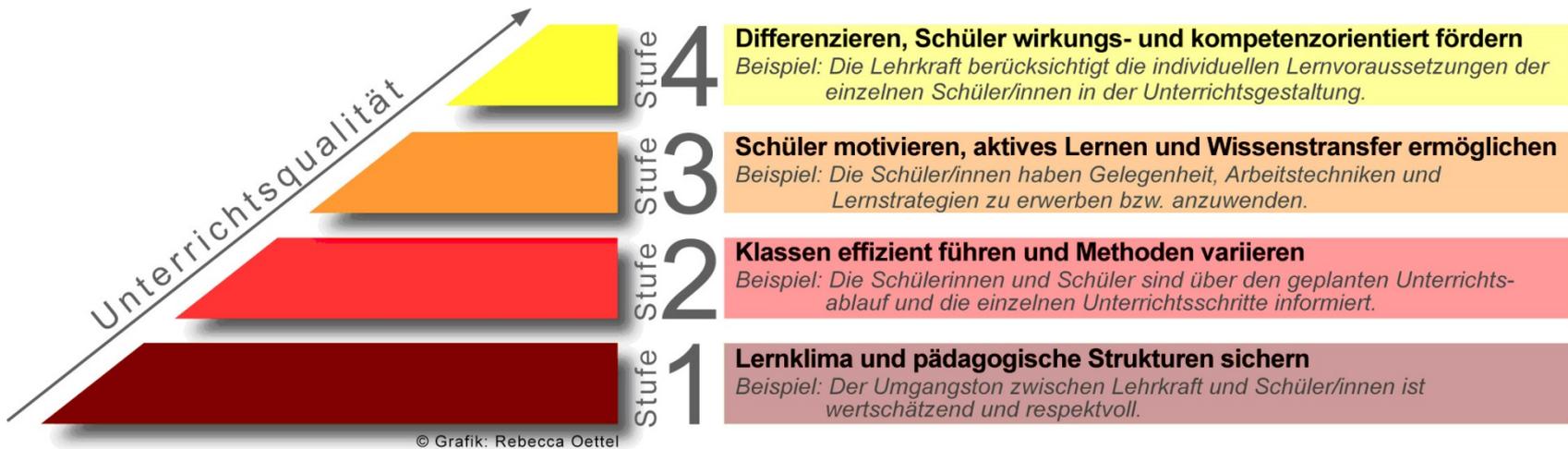
Stufe 2 = Unterstützende Unterrichtsklima

Stufe 1 = Klassenführung



Dies ermöglicht Feed-Back,

Wo stehe ich? Stufe 2.





...Feed-Up

Wo soll ich hin? Stufe 4.





...und Feed-Forward

Was muss ich als nächstes tun? Stufe 3.





Erhebung von
Unterrichtsqualität
mittels Beobachtung
- Ein Beispiel -



Unterrichtsbeobachtungen der Schulinspektion Hamburg

- **Ziel: Aussagen zur Qualität des Unterrichts auf Schulebene treffen.**
- **Vorgehen:**
 - Einsicht in bis zu 100 Unterrichtssequenzen an 2-3 Schultagen
 - Grundgesamtheit sind die Stunden an der Schule in einer Woche.
 - Lehrkräfte können häufiger, jedoch in verschiedenen Situationen, gesehen werden (Angebot-Nutzungs-Modell).
 - Mittelwerte können mit einer Präzision von ca. 90% berichtet werden.
 - Zufallsziehung von Unterrichtssequenzen vor dem Schulbesuch
 - 20-minütige Unterrichtssequenzen: Expertenrating durch Inspektoren und geschulte externe Beobachter
 - Zufallszuordnung von Inspektoren zu den Sequenzen
 - Je Schule ca. 10% Doppelbeobachtungen zur Qualitätssicherung



Über die Aussagekraft von Unterrichtsbeobachtungen

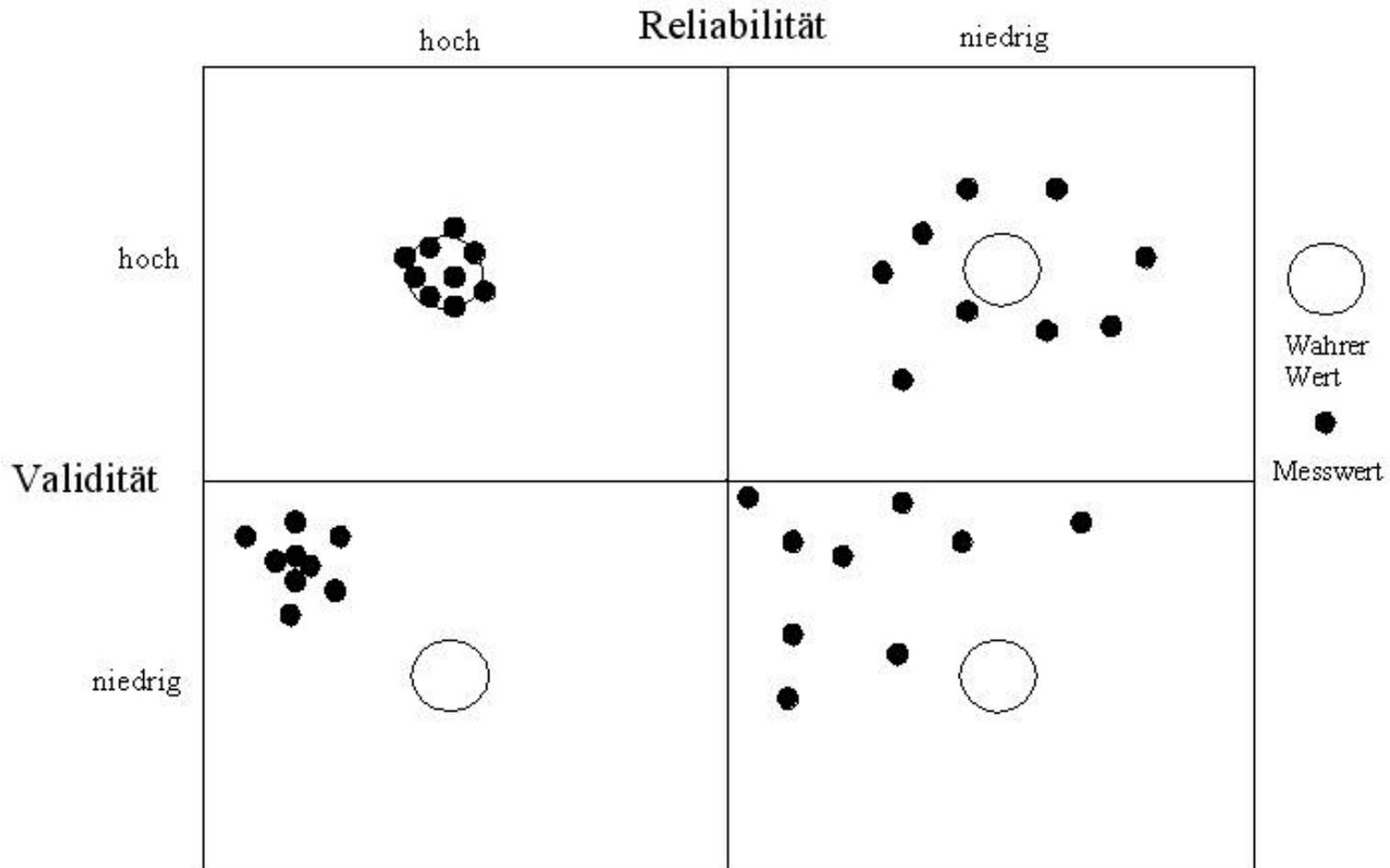


Was wissen wir über Unterrichtsbeobachtungen?

- Unterrichtsbeobachtungen ermöglichen hoch-inferente Einschätzungen zur Qualität von Unterricht.
- Die Korrelation mit Eigenaussagen von Lehrkräften und Einschätzungen durch Schüler liegen im Bereich von 0,30 bis 0,40. Dies liegt u. a. daran, dass
 - Beobachter mit Lehrkräften und Schülern eher übereinstimmen, wenn ein Merkmal einfach beobachtbar ist.
 - Schülerinnen und Schüler übergeneralisieren, also wenig differenzieren.
- Beobachterurteile ermöglichen vor allem eine neutrale Beschreibung und Bewertung des Unterrichts.



Beobachtungen müssen reliabel und valide sein!





Das Problem von Beobachtungsdauer und Gegenstand

Table 4

Comparison of the Reliability Estimates (ρ) for Different Combinations of Raters and Lessons for the Whole Lesson and the First 30 Minutes of a Lesson

Number of Lessons and Raters	Richness		Errors and Imprecision		Student Participation in Meaning Making and Reasoning	
	Whole Lesson	30 Minutes	Whole Lesson	30 Minutes	Whole Lesson	30 Minutes
One lesson						
1 rater	0.45	0.50	0.37	0.34	0.46	0.32
2 raters	0.58	0.59	0.50	0.46	0.59	0.41
3 raters	0.64	0.63	0.57	0.53	0.65	0.45
4 raters	0.67	0.65	0.61	0.57	0.68	0.48
Two lessons						
1 rater	0.59	0.65	0.51	0.49	0.63	0.49
2 raters	0.71	0.73	0.64	0.62	0.74	0.58
3 raters	0.76	0.77	0.71	0.68	0.79	0.62
4 raters	0.79	0.78	0.74	0.71	0.81	0.65
Three lessons						
1 rater	0.66	0.73	0.58	0.57	0.72	0.59
2 raters	0.77	0.80	0.71	0.70	0.81	0.68
3 raters	0.81	0.83	0.77	0.75	0.85	0.71
4 raters	0.84	0.84	0.80	0.78	0.87	0.73
Four lessons						
1 rater	0.69	0.77	0.63	0.63	0.77	0.66
2 raters	0.80	0.83	0.75	0.74	0.85	0.74
3 raters	0.84	0.86	0.81	0.79	0.88	0.77
4 raters	0.86	0.87	0.83	0.82	0.90	0.78



Das Problem von Beobachtungsdauer und Gegenstand

Table 4

Comparison of the Reliability Estimates (ρ) for Different Combinations of Raters and Lessons for the Whole Lesson and the First 30 Minutes of a Lesson

- Werden 15 Minuten statt einer ganzen Stunde beobachtet, werden in der Regel 80% der Beobachtungselemente unverändert bewertet.
- Die Unterschiede zwischen den Bewertungen eines Unterrichts durch mehrere Beobachter nehmen mit der Beobachtungsdauer zu.
- Problematisch sind vor allem Items, die auf Merkmale abzielen, die psychologische Konstrukte/ Prozesse messen sollen (z.B. kognitive Aktivierung).
- Hier ist es sinnvoll längere Sequenzen mit mehreren Personen gleichzeitig zu beobachten.

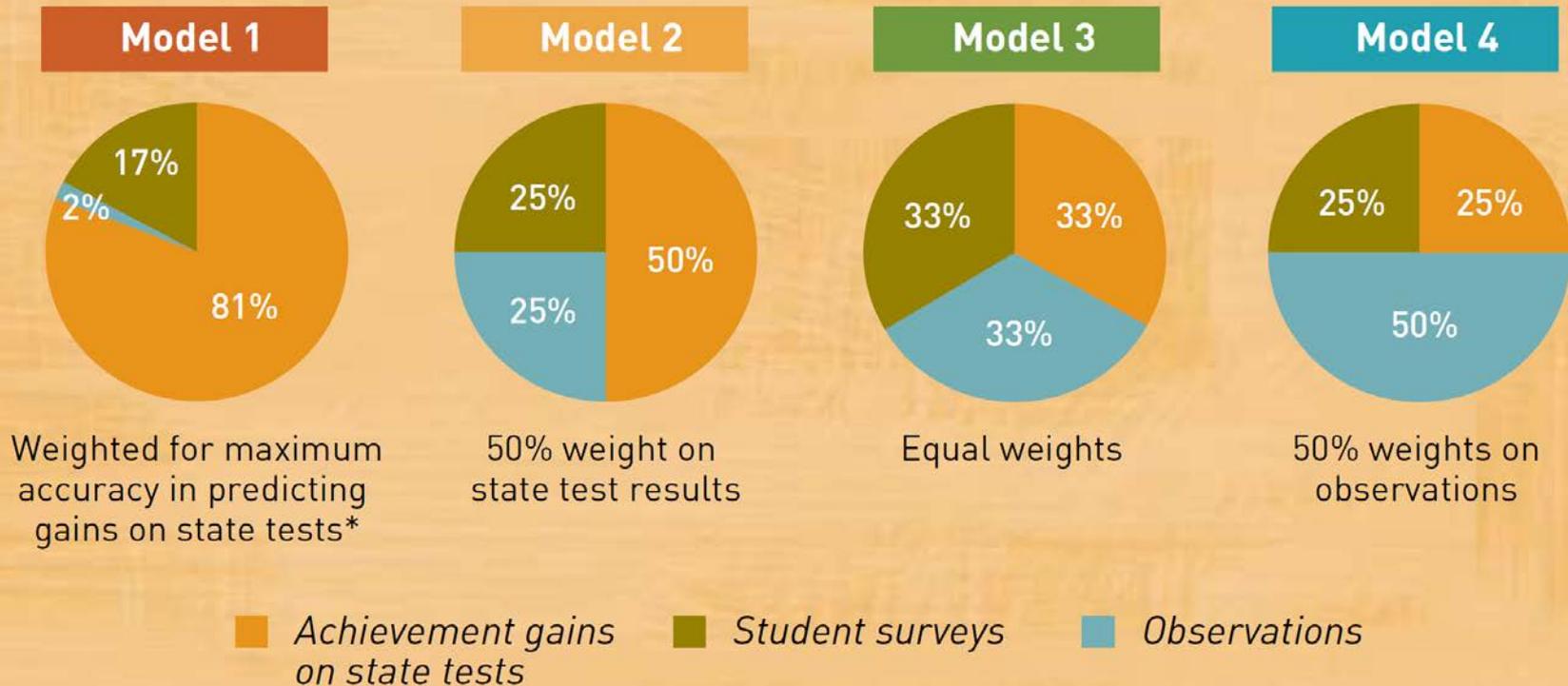
	4 raters	3 raters	2 raters	1 rater	4 raters	3 raters
Four lessons	0.64	0.64	0.60	0.70	0.67	0.75
1 rater	0.69	0.77	0.63	0.63	0.77	0.66
2 raters	0.80	0.83	0.75	0.74	0.85	0.74
3 raters	0.84	0.86	0.81	0.79	0.88	0.77
4 raters	0.86	0.87	0.83	0.82	0.90	0.78



Wie kann man den Problemen begegnen?

Figure 3

Four Ways to Weight

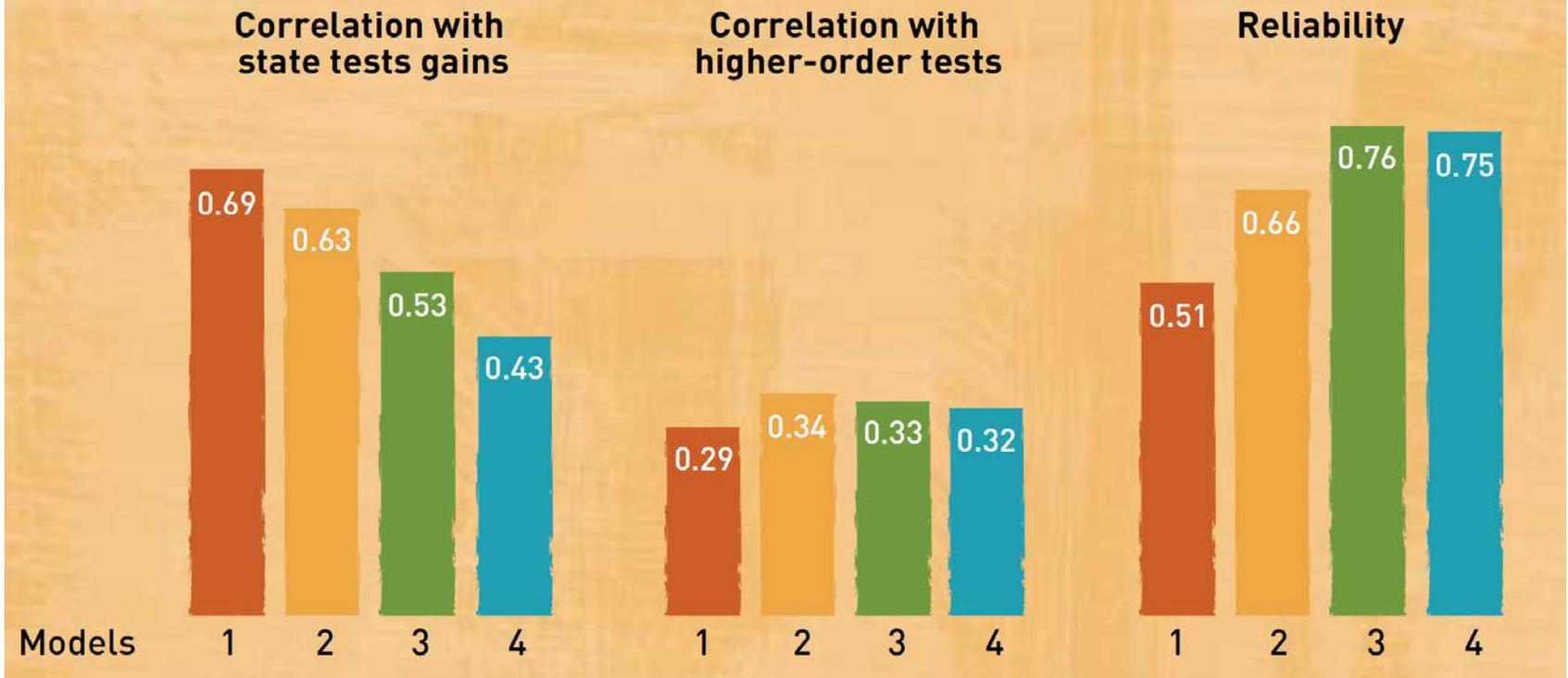




Wie kann man den Problemen begegnen?

Figure 4

Trade-Offs from Different Weighting Schemes *Middle School English Language Arts*

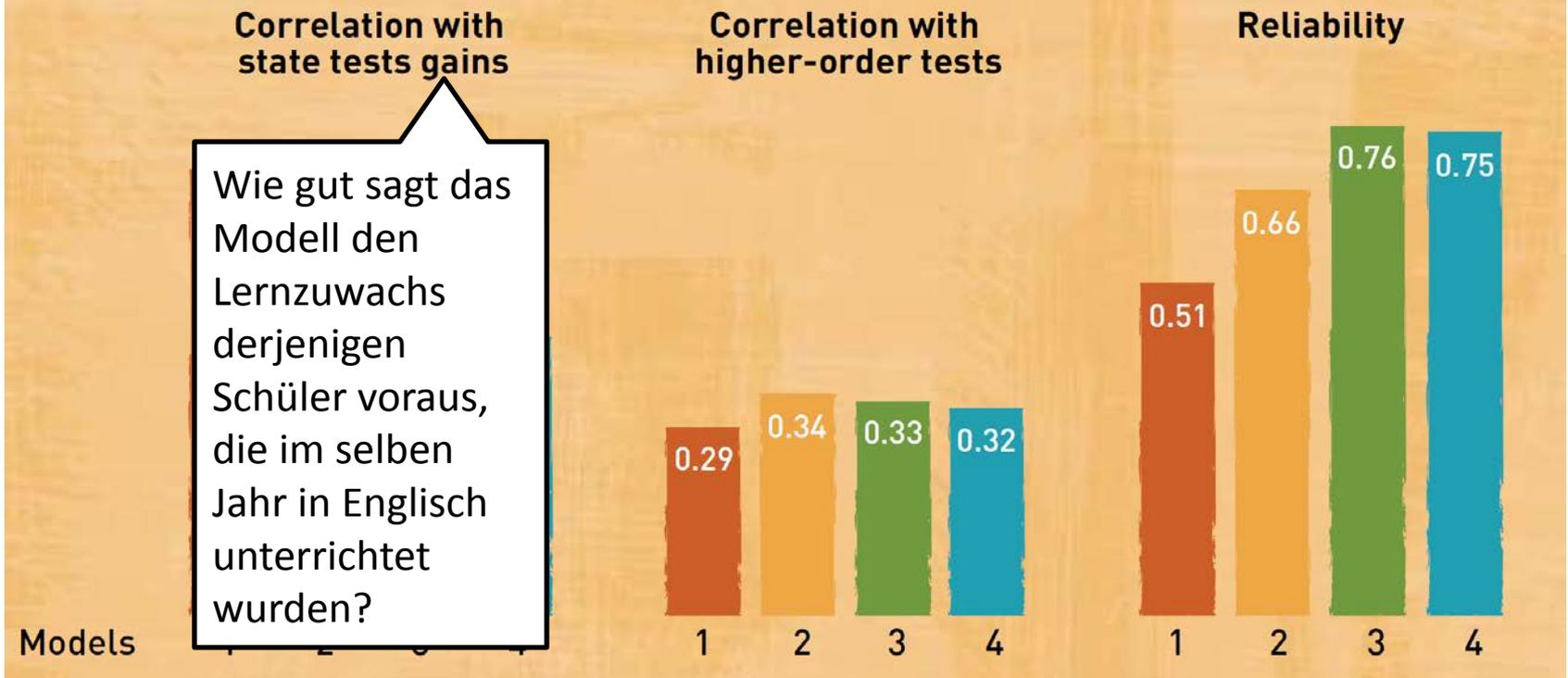




Wie kann man den Problemen begegnen?

Figure 4

Trade-Offs from Different Weighting Schemes *Middle School English Language Arts*

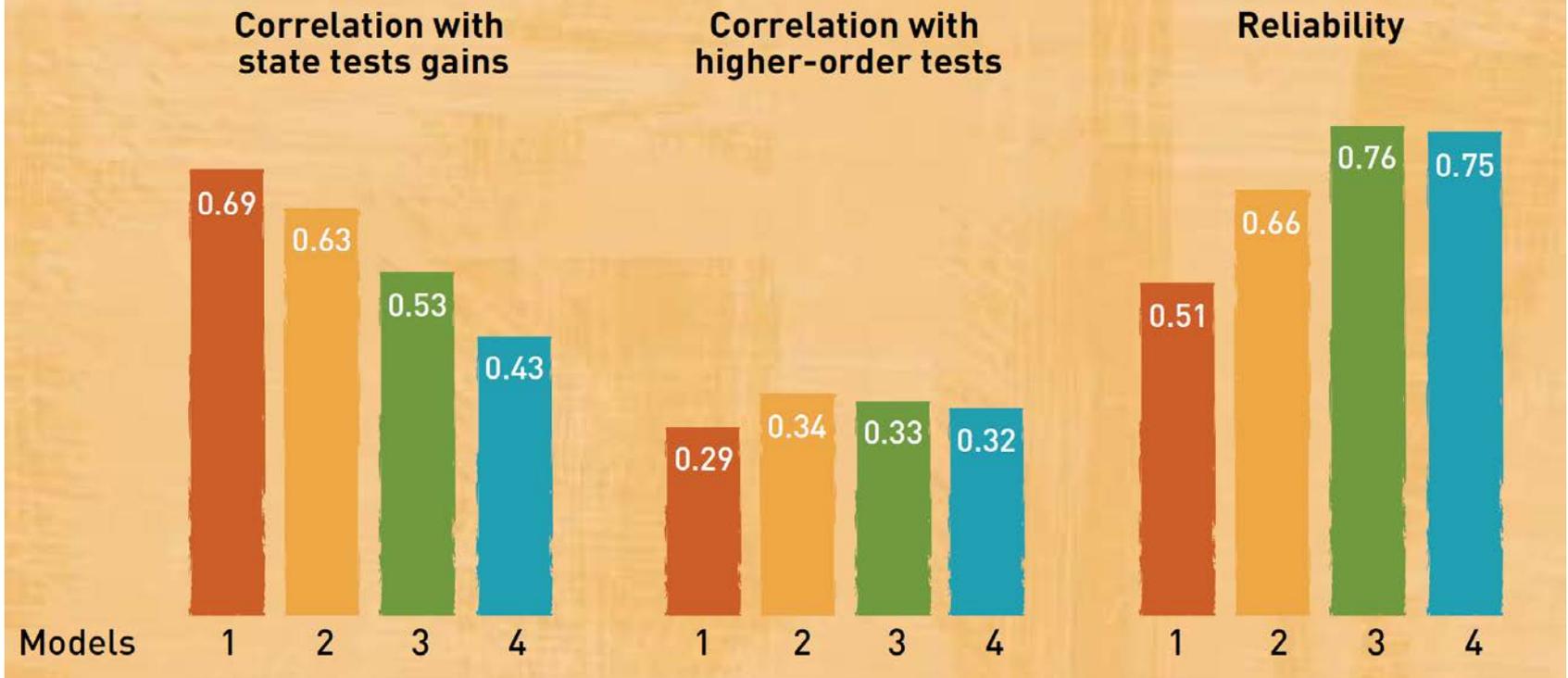




Wie kann man den Problemen begegnen?

Figure 4

Trade-Offs from Different Weighting Schemes *Middle School English Language Arts*

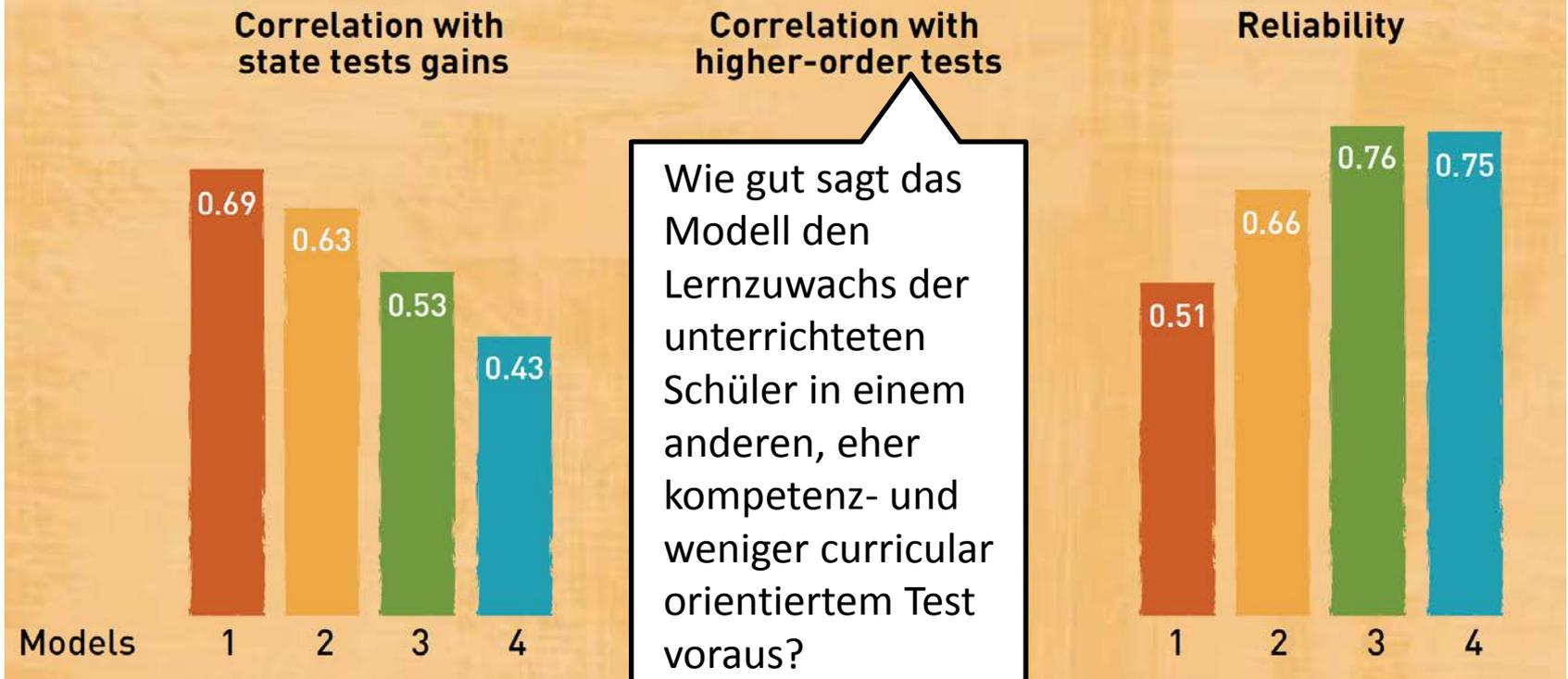




Wie kann man den Problemen begegnen?

Figure 4

Trade-Offs from Different Weighting Schemes *Middle School English Language Arts*

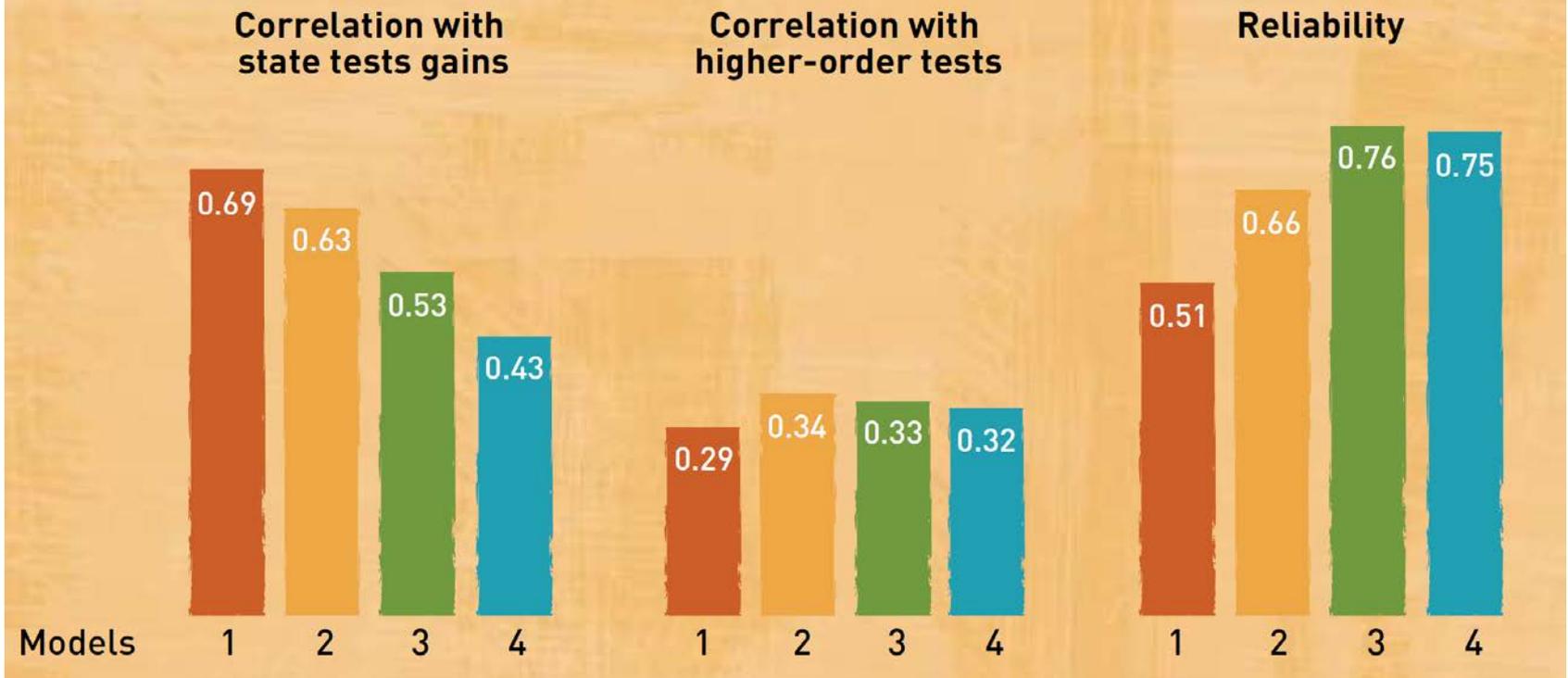




Wie kann man den Problemen begegnen?

Figure 4

Trade-Offs from Different Weighting Schemes *Middle School English Language Arts*

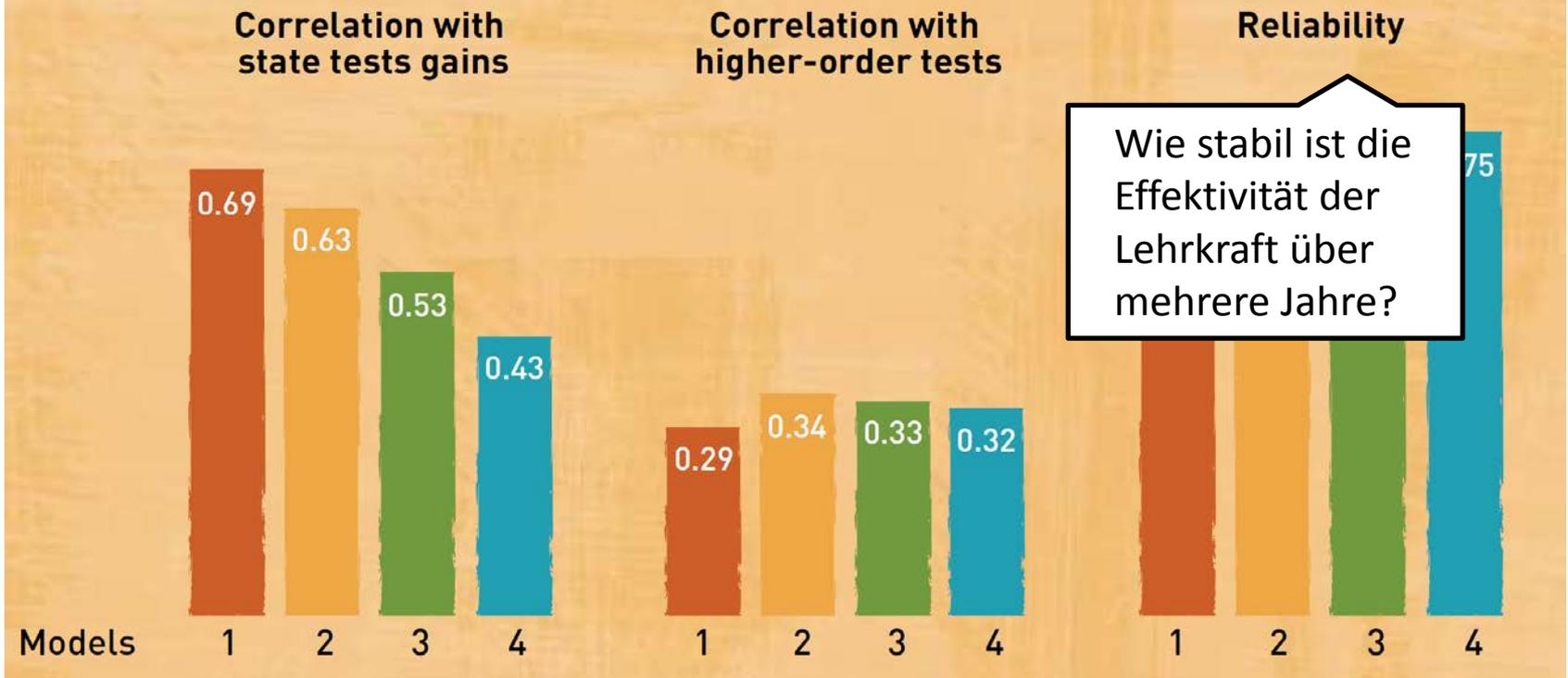




Wie kann man den Problemen begegnen?

Figure 4

Trade-Offs from Different Weighting Schemes *Middle School English Language Arts*



Wie stabil ist die Effektivität der Lehrkraft über mehrere Jahre?

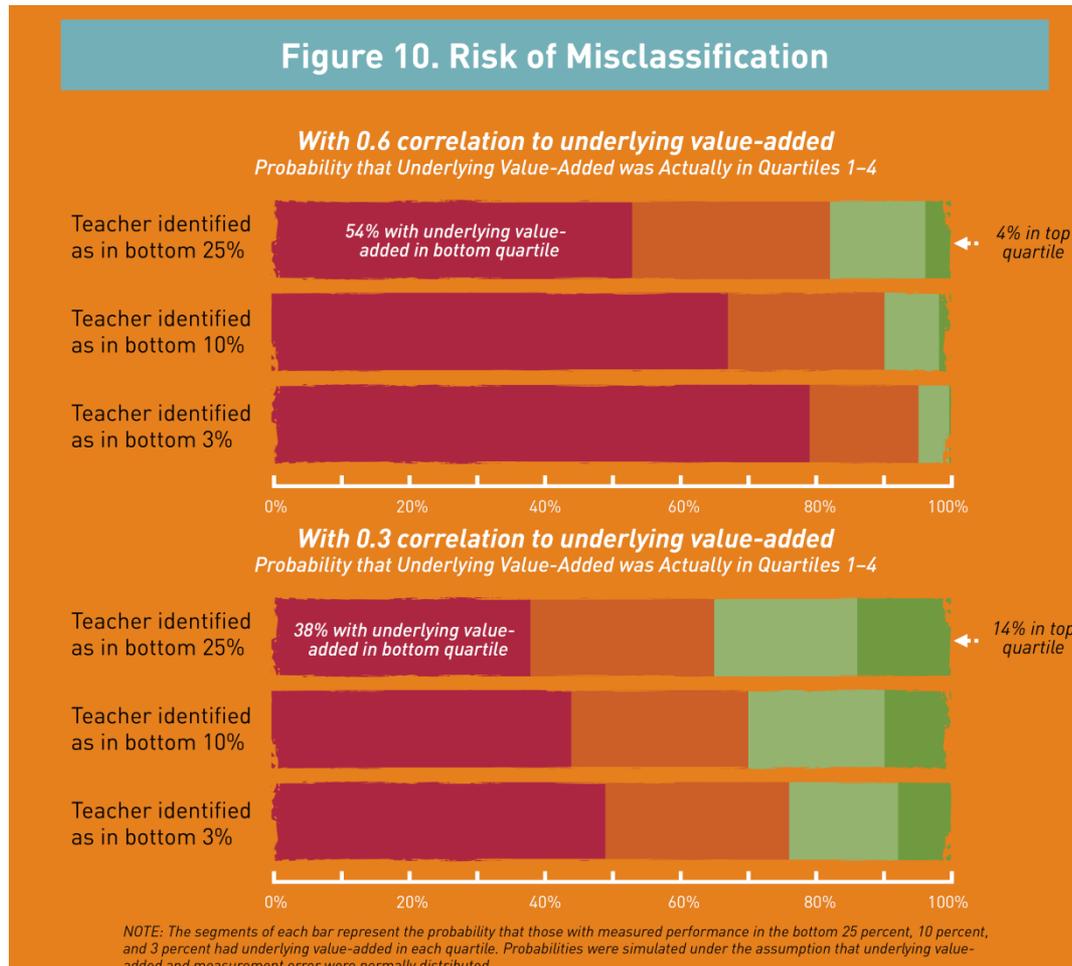


Was bedeutet das?

- Zur Einschätzung der Effektivität werden immer mehrere Datenquellen benötigt.
- Eine starke Gewichtung auf Lernzuwächse, gemittelt über alle Fächer und alle Klassen in einem Schuljahr, ermöglicht die beste Voraussage der Lernzuwächse in anderen Klassen im selben Schuljahr.
- Eine Gleichgewichtung von Beobachtungen, Schülereinschätzungen und Lernzuwächsen ermöglicht hingegen die beste Voraussage darüber, ob ähnliche Lernzuwächse im Unterricht einer Lehrkraft auch in Zukunft zu erwarten sind.



Aber: Die Unsicherheit bei Einzelbeobachtungen bleibt in der Regel bestehen!





Die Konsequenz

- Einzelbeobachtungen sind immer mit einer hohen Unsicherheit behaftet.
- Stichproben, mit vergleichsweise großem Umfang, ermöglichen belastbarere Aussagen, da dort das „Gesetz der großen Zahlen“ gilt:
 - Mit wachsendem Stichprobenumfang nähert sich das Mittel der empirischen Verteilung dem Mittel der wahren Verteilung an.



Welche Stichprobengrößen benötigt man?

Grundgesamtheit	Beispiel	+/- 5%	+/- 7%	+/- 10%
25 Stunden	~ Lehrkraft (hD) an einem GY in BW pro Woche	24	23	21
1000 Stunden	~ Lehrkraft (hD) an einem GY in BW pro Jahr	278	165	88
128 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Tag	97	78	56
640 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Woche	241	151	84
5120 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Jahr	358	189	95
512 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Tag	220	142	82
2560 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Woche	335	183	93
20480 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Jahr	378	195	96



Welche Stichprobengrößen benötigt man?

Grundgesamtheit	Beispiel	+/- 5%	+/- 7%	+/- 10%
25 Stunden	~ Lehrkraft (hD) an einem GY in BW pro Woche			1
1000 Stunden	~ Lehrkraft (hD) an einem GY in BW pro Jahr			8
128 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Tag			6
640 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Woche	241	151	84
5120 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Jahr	358	189	95
512 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Tag	220	142	82
2560 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Woche	335	183	93
20480 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Jahr	378	195	96

Der wahre Wert liegt im Bereich der Messung +/-5%.

Beispiel: MW = 3,00, wahrer Wert = 2,85 bis 3,15.



Welche Stichprobengrößen benötigt man?

Grundgesamtheit	Beispiel	+/- 5%	+/- 7%	+/- 10%
25 Stunden	~ Lehrkraft (hD) an einem GY in BW pro Woche	24	23	21
1000 Stunden	~ Lehrkraft (hD) an einem GY in BW pro Jahr	278	165	88
128 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Tag	97	78	56
640 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Woche	241	151	84
5120 Stunden	~ 2-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Jahr	358	189	95
512 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Tag	220	142	82
2560 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Woche	335	183	93
20480 Stunden	~ 8-zügige Schule, KS 5-12, 8 Stunden Unterricht pro Tag, Jahr	378	195	96



Praktisches Vorgehen

- Datenerhebung -



Stichprobenziehung

		VSK A	VSK B	1A	Fach	1B	Fach	2a	Fach
Mittwoch	1	Ek-Projekt	En-Projekt	***	Sport	***	M	***	M
	2	Ek-Frühstück	En-Frühstück	***	D	***	M	***	Sp. Engl./D
	3	Ek-Sport	En-Projekt	***	D	***	SU	***	D
	4	Wa-Kunst	En-Sport	***	s.u.	***	WP	***	D
	5	Ek-Vorlesen	En-Vorlesen	***	M	***	SU	***	Musik
	6	Ek-Betreuung	En-Betreuung	***	Betreuung	***	Betreuung	***	Betreuung
Donnerstag	1	Ek-Projekt	En-Projekt	***	D	***	D/Musik	***	Sport
	2	Ek-Frühstück	En-Frühstück	***	M	***	D/Musik	***	M
	3	Ek-Freispiel	En-Projekt	***	Engl.	***	Engl.	***	SU
	4	Ek-Projekt	En-Kleingrupp	***	D	***	M	***	SU
	5	Ek-Vorlesen	En-Vorlesen	***	Freiarb./ Sport	***	SU/ Sport	***	D
	6	Ek-Betreuung	En-Betreuung	***	Betreuung	***	Betreuung	***	Betreuung



Stichprobenziehung

	VSK A	VSK B	1A	Fach	1B	Fach	2a	Fach	
Mittwoch	1	Ek-Projekt I2.1	En-Projekt	***	I1.1 Sport	***	M	***	I1.2 M I2.1
	2	Ek-Frühstück	En-Frühstück	***	D	***	I1.1 M	***	Sp. Engl./D
	3	Ek-Sport	En-Projekt	***	D	***	SU	***	D
	4	Wa-Kunst	En-Sport	***	s.u.	***	WP	***	D
	5	Ek-Vorlesen	En-Vorlesen	***	M	***	SU	***	Musik
	6	Ek-Betreuung	En-Betreuung	***	Betreuung	***	Betreuung	***	Betreuung
Donnerstag	1	Ek-Projekt	En-Projekt	***	D	***	D/Musik	***	Sport
	2	Ek-Frühstück	En-Frühstück	***	M	***	D/Musik	***	M
	3	Ek-Freispiel	En-Projekt	***	Engl.	***	Engl.	***	SU
	4	Ek-Projekt	En-Kleingrupp	***	D	***	M	***	SU
	5	Ek-Vorlesen	En-Vorlesen	***	Freiarb./ Sport	***	SU/ Sport	***	D
	6	Ek-Betreuung	En-Betreuung	***	Betreuung	***	Betreuung	***	Betreuung



Beobachtungsplan

Besuchsblock	Zeit	Stunde	Klasse	Jahrgang	code	Lehrer	Fach	Raum	
		Inspektor 1							
1	8.00-8.45	1	VSK b	VSK	2000101	***	Projekt	Pavillon	
2	8.00-8.45	1	4a	4	2040101	***	Mathe	Wabe 2	
3	8.50-9.35	2	1a	1	2010102	***	Mathe	Wabe 4	
1	8.50-9.35	2	4b	4	2040202	***	Mathe	Wabe 2	
2	9.55-10.40	3	2a	2	2020103	***	Su	Wabe 3	
3	9.55-10.40	3	VSK a	VSK	2000103	***	Freispiel	Pavillon	
1	10.45-11.30	4	2b	2	2020204	***	Kunst	Wabe 3	
2	10.45-11.30	4	1a	1	2010104	***	Haf	Deutsch	
3	11.50-12.35	5	1b	1	2010205	***	Bor	Su/Sport	Wabe 4/S-Halle
1	11.50-12.35	5	4b	4	2040205	***	Deutsch	Wabe 2	
2	12.40-13.00	6	VSK b	VSK	2000103	***	Betreuung	Pavillon	
3	12.40-13.00	6	2b	2	2020206	***	Betreuung	Wabe 3	



Das Erhebungsinstrument

	++	+	-	--	0
28 Die individuellen Lernvoraussetzungen der einzelnen Schülerinnen und Schüler werden in der Unterrichtsgestaltung berücksichtigt.					
29 Die Schülerinnen und Schüler werden entsprechend ihrer individuellen Lernvoraussetzungen gefördert.					
16 Die Schülerinnen und Schüler erhalten gezielt die Möglichkeit, Lern- und Arbeitstechniken zu erwerben.					
04 Die Lehrkraft behält den Überblick über unterrichtsbezogene und unterrichtsfremde Aktivitäten der Schülerinnen und Schüler.					
05 Mit Störungen wird angemessen und effektiv umgegangen.					
17 Den Schülerinnen und Schülern wird die Möglichkeit gegeben, kommunikative Fähigkeiten (und Fertigkeiten) zu lernen bzw. anzuwenden.					
24 Die Reflexion eigener Lernprozesse ist Bestandteil des Unterrichts. ►					
09 Die Schülerinnen und Schüler gehen freundlich und rücksichtsvoll miteinander um.					
26 Die Schülerinnen und Schüler erhalten differenzierte Leistungsrückmeldungen.					
07 Dem Unterricht liegt eine klare Struktur zugrunde.					
02 Das Unterrichtstempo ist angemessen.					
15 Die Schülerinnen und Schüler werden angeregt/angeleitet, den Unterricht aktiv mitzugestalten, oder sie gestalten den Unterricht aktiv mit.					
03 Der Unterricht erfolgt auf Basis eines festen Regelsystems.					
25 Mit Schülerfehlern wird angemessen umgegangen.					
06 Die Arbeitsaufträge und Erklärungen sind angemessen, klar und präzise formuliert. ►					



Verringerung der Inferenz durch Kriterien

Erläuterungsbogen

01 Die Unterrichtszeit wird effektiv genutzt.	Die für die Unterrichtseinheit zur Verfügung stehende Unterrichtszeit ist auf die Unterrichtsinhalte bezogen. <ul style="list-style-type: none">• SuS sind einen großen Anteil der Unterrichtszeit mit dem Unterrichtsinhalt befasst.• Übergangsphasen als besonders kritische Punkte im Unterrichtsgeschehen sind nahtlos und schnell, so dass die Lehrdynamik beibehalten wird.• Das Unterrichtsziel bleibt fokussiert, indem Abschweifungen eingegrenzt werden.• Zeitkonsumenten wie beispielsweise Technikaufbau oder ein <u>Rearrangement</u> der Sitzordnung bleiben von marginaler Bedeutung.
02 Das Unterrichtstempo ist angemessen.	Das Unterrichtstempo ist so angelegt, dass die SuS dem Unterricht gut folgen können. <ul style="list-style-type: none">• Das grundlegende Unterrichtstempo ist zügig, wird jedoch an die Komplexität der Aufgaben (mehr Zeit für komplexere Aufgaben, einfachere Aufgaben werden schneller behandelt) und Sozialform (Unterrichtsformen mit leistungsheterogenen Kleingruppen benötigen i.d.R. mehr Zeit als lehrergeleiteter Unterricht) angepasst.• SuS haben ausreichend Zeit zum Nachdenken und Antworten (mindestens 3 Sekunden nach Fragen).• Unabhängig von der Unterrichtsmethode/Lernform gibt es weder SuS in großer Zahl, die früher mit Aufgaben fertig sind, noch SuS in großer Zahl, die Aufgaben nicht in der vorgesehenen Zeit bewältigen (Jahrgangs- und <u>Schulformspezifisch</u> beachten!).
03 Der Unterricht erfolgt auf Basis eines festen Regelsystems.	Ein festes Regelsystem ist für den Unterricht vorhanden und wird umgesetzt. <ul style="list-style-type: none">• Die vereinbarten Regeln, verstanden als allgemeine Standards des Verhaltens der SuS und in Bezug auf Unterrichtsabläufe, werden von den SuS beachtet, eingehalten und umgesetzt (z.B.: Es herrscht ein der Arbeitsform entsprechender Lärmpegel bzw. der Unterricht wird erst fortgesetzt, wenn Ruhe zum Arbeiten herrscht).• Die Organisation des Unterrichts wird über Rituale und Prozeduren gestützt (akustische und visuelle Signale, Symbole, Gestik und Körpersprache, Blickkontakte, Hausaufgabentafel, Ermahnungskarten, Klokette, etc.). <ul style="list-style-type: none">◦ <i>Regeln sind visualisiert.</i>

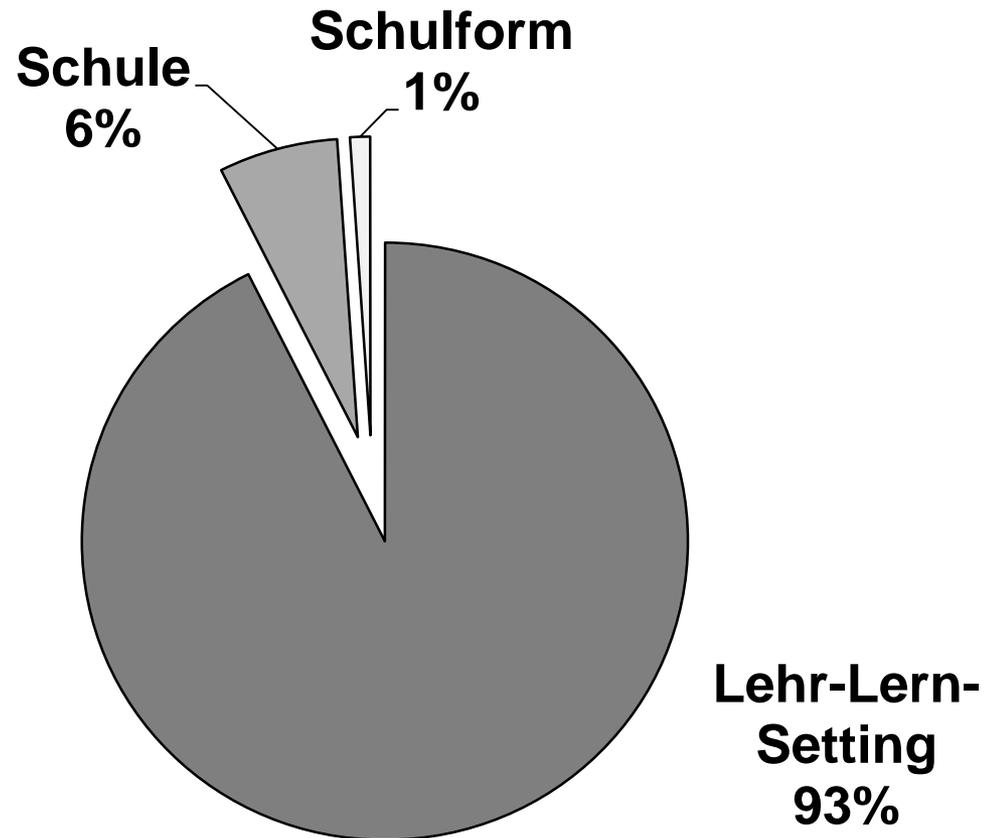


Praktisches Vorgehen

- Befunde -

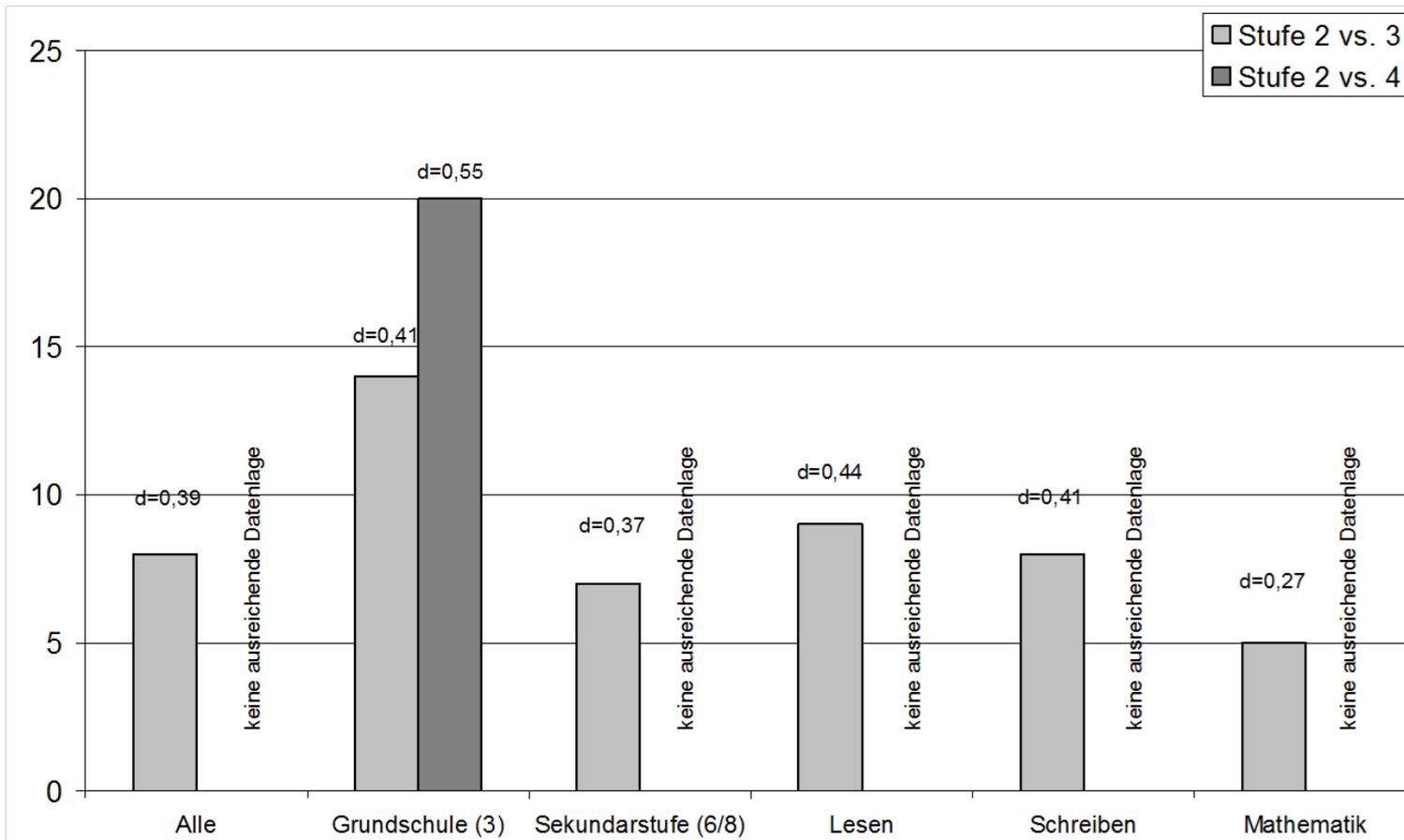


Die Schulform hat nahezu keinen Einfluss auf die Unterrichtsqualität



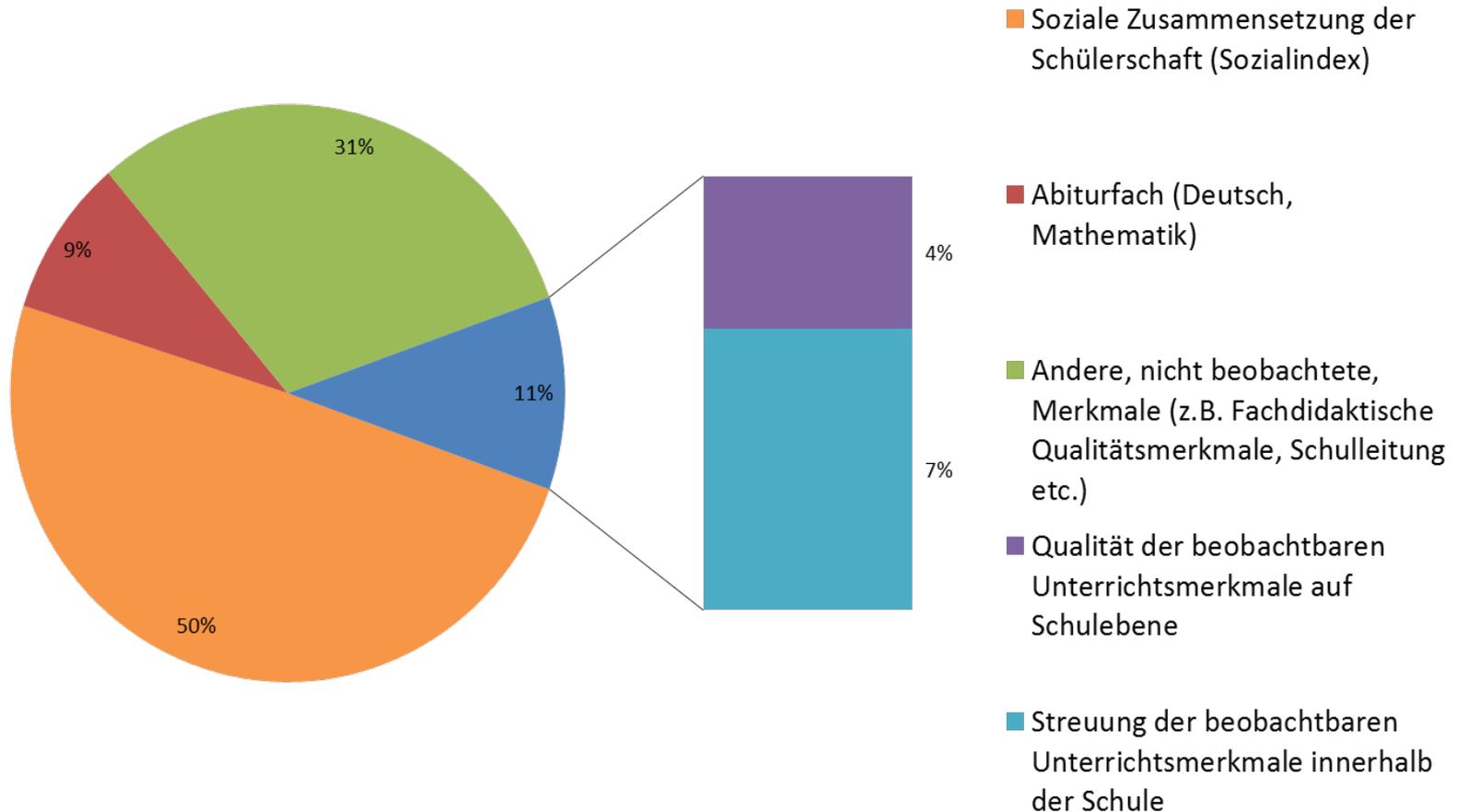


Die Beobachtungen messen wirksamen Unterricht





Guter Unterricht = hohe Gesamtqualität + geringe Unterschiede zwischen einzelnen Unterrichtssequenzen





Praktisches Vorgehen - Rückmeldung -



Rückmeldung der Schulinspektion Hamburg

Um die an Schulen wahrgenommene Unterrichtsqualität beurteilen zu können und Entwicklungsperspektiven aufzuzeigen, nutzt die Schulinspektion Hamburg ein vierfach abgestuftes Modell, das auf Basis empirischer Analysen entwickelt wurde.

[...]

Der Unterricht an Ihrer Schule hat in diesem Modell die Stufe 3 erreicht. D.h., generell gelingt es bereits gut das Lernklima und pädagogische Strukturen zu sichern, Klassen effizient zu führen und Methoden zu variieren sowie Schülerinnen und Schüler zu motivieren und ihnen aktives Lernen und den Transfer erworbenen Wissens zu ermöglichen. Entwicklungsperspektiven zeichnen sich hingegen mit Blick auf die innere Differenzierung im Unterricht sowie die wirkungs- und kompetenzorientierte Förderung der Schülerinnen und Schüler ab. Die Qualität des Unterrichts an ihrer Schule verhält sich dabei zwischen den einzelnen von uns eingesehenen Unterrichtssequenzen vergleichsweise heterogen.



Kerninformationen in den Rückmeldung

Mittelwert auf Schulebene

Der Unterricht an Ihrer Schule hat in diesem Modell die Stufe 3 erreicht.

Variationskoeffizient

Die

Qualität des Unterrichts an ihrer Schule verhält sich dabei zwischen den einzelnen von uns eingesehen Unterrichtssequenzen vergleichsweise heterogen.



Feed-back in den Rückmeldungen

D.h., generell gelingt es bereits gut das Lernklima und pädagogische Strukturen zu sichern, Klassen effizient zu führen und Methoden zu variieren sowie Schülerinnen und Schüler zu motivieren und ihnen aktives Lernen und den Transfer erworbenen Wissens zu ermöglichen.

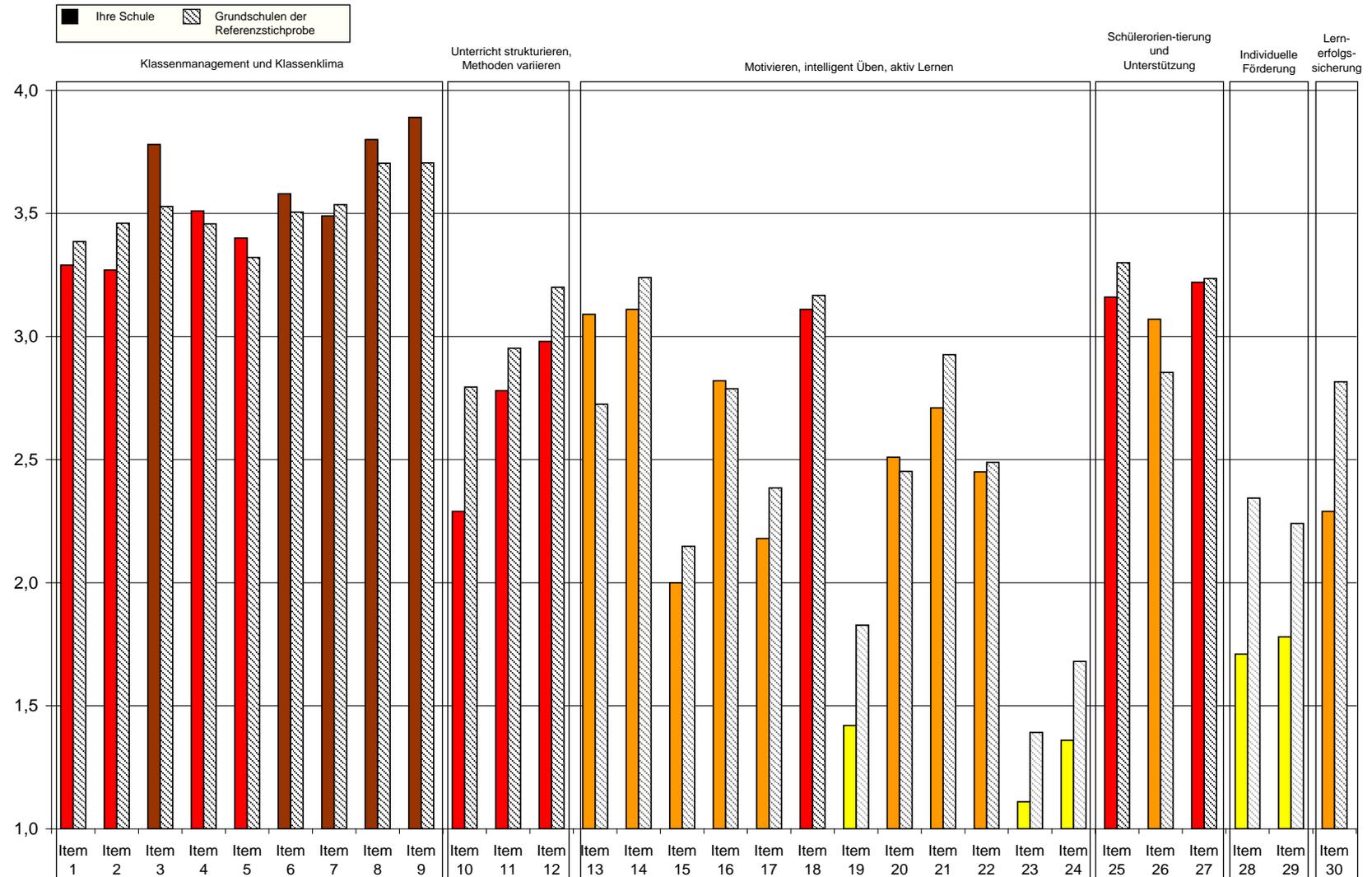


Feed-forward in den Rückmeldungen

Entwicklungsperspektiven zeichnen sich hingegen mit Blick auf die innere Differenzierung im Unterricht sowie die wirkungs- und kompetenzorientierte Förderung der Schülerinnen und Schüler ab.



Zusatzinformationen zur Detailanalyse





Sozialer Vergleich

Anteile des Unterrichts an Ihrer Schule auf den einzelnen Stufen
und Vergleichsgrößen in Prozent

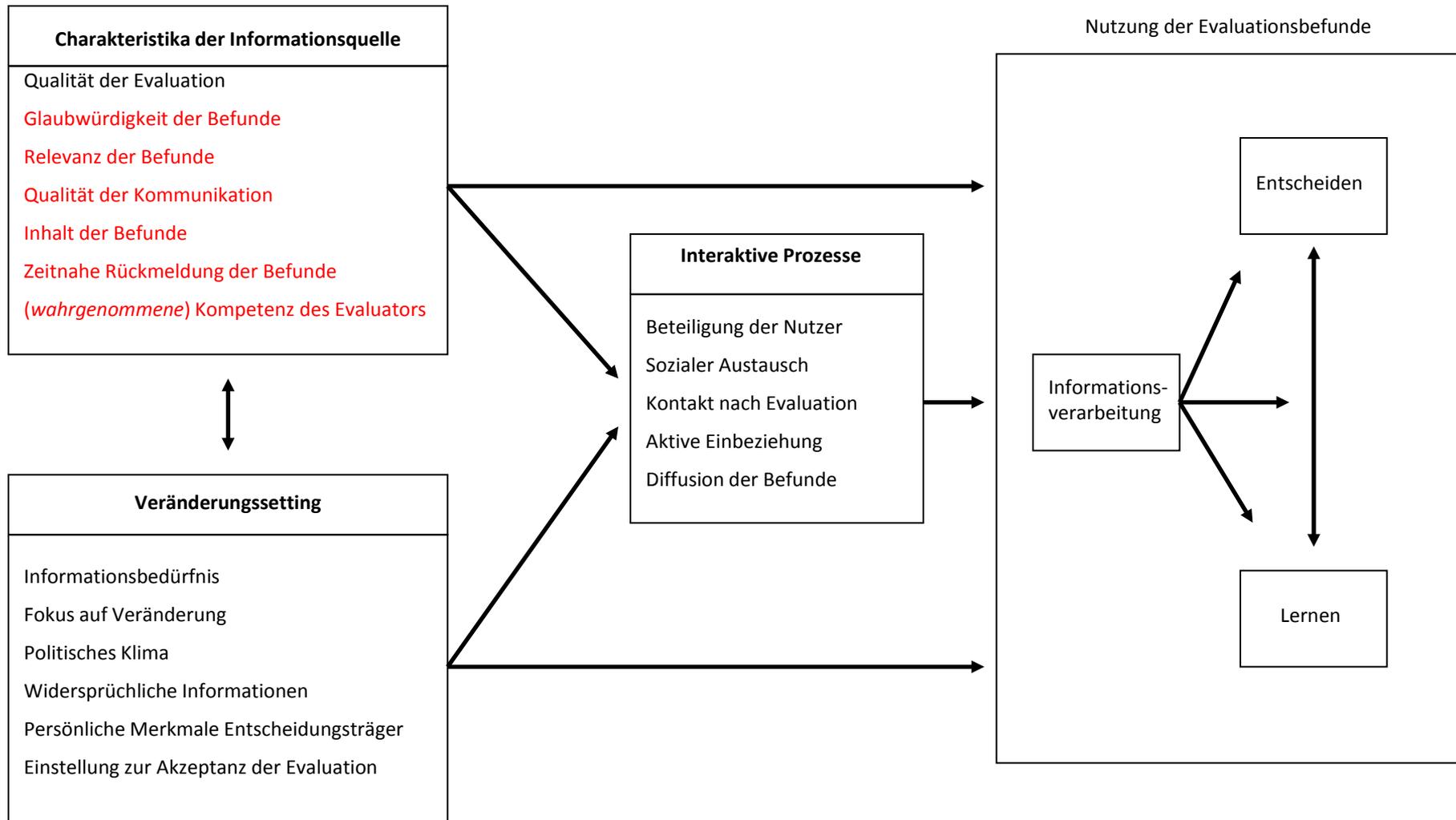




Und die Wirkung?

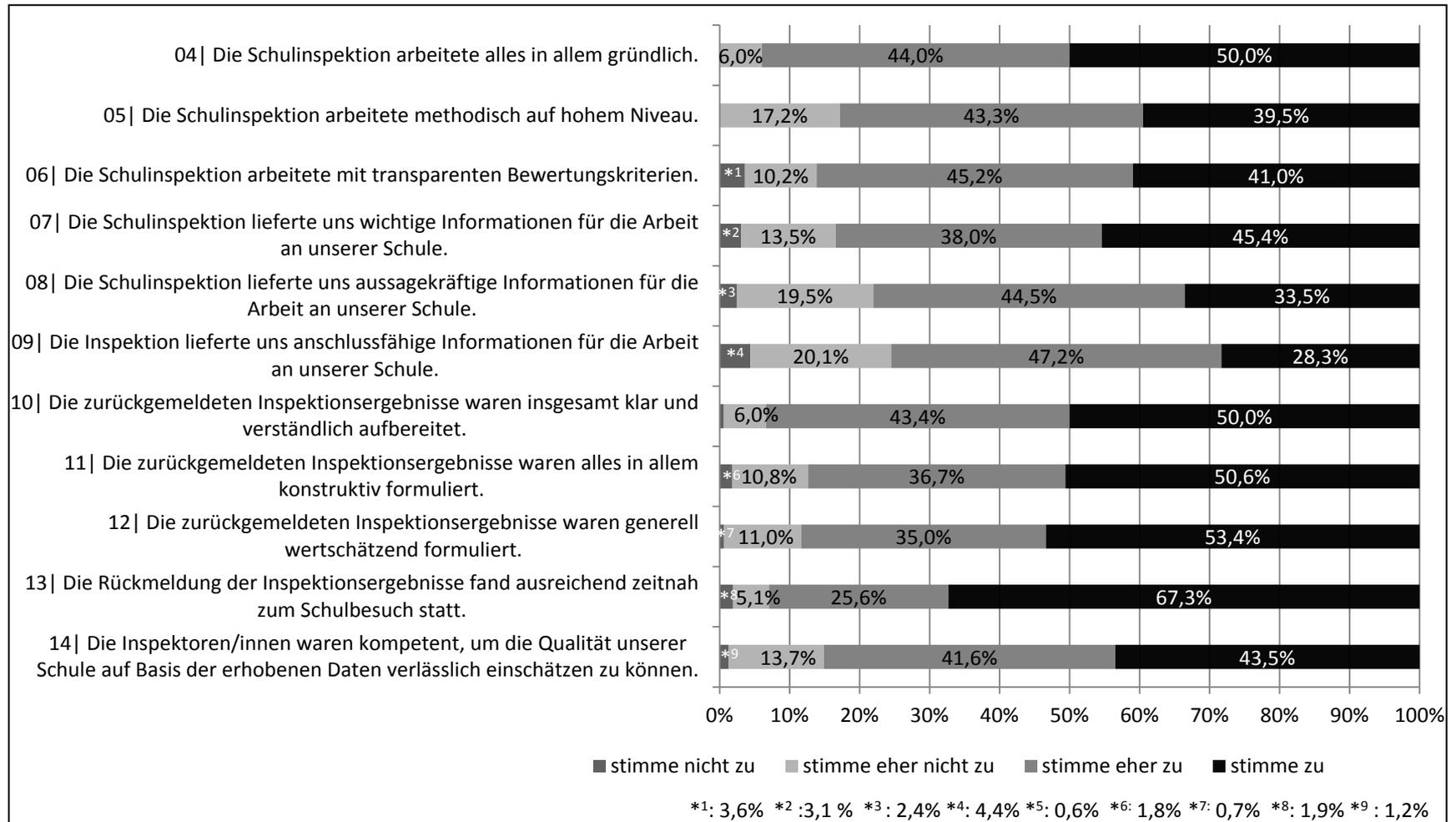


Bedingungsfaktoren für Feedback-Wirksamkeit



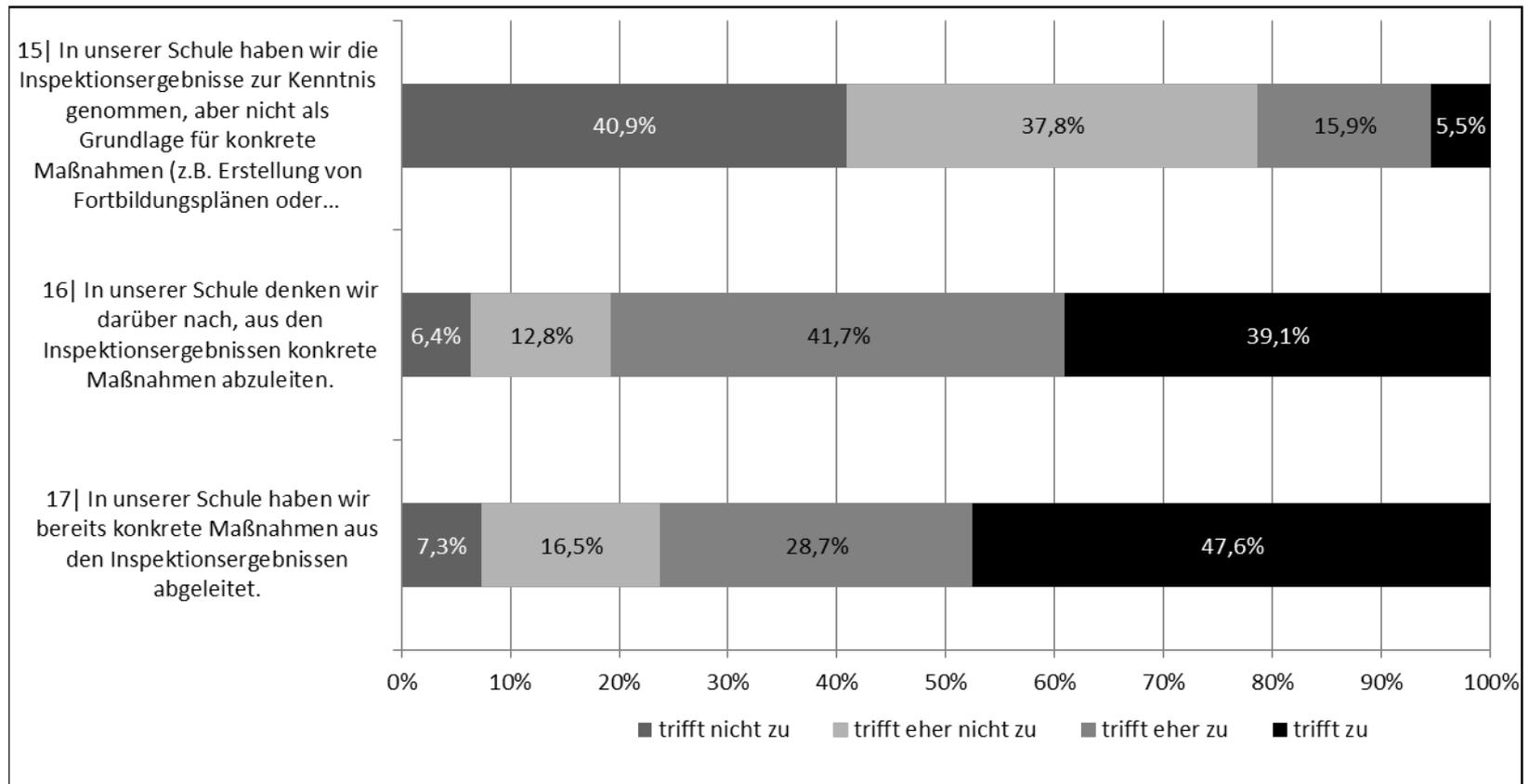


Kriterien für wirksame Rückmeldungen Großteils erfüllt





Konsequenz: Rund 80% der Schulen nutzen die Daten





Rückmeldung führt zu besseren Schülerleistungen

- **Es konnte gezeigt werden, dass an inspizierten Schulen**
 - die Abiturleistung der Schülerinnen und Schüler in Deutsch und Mathematik zwei Jahre nach einer Inspektion im Schnitt um 0,25 Punkte höher liegen als erwartet.
 - die Leistungen in Leistungstests (KESS) zwei Jahre nach einer Inspektion ebenfalls höher als erwartet sind. So beläuft sich der Effekt beim Lesen im Sekundarbereich auf rund $\frac{2}{3}$ eines Lernjahres.



Aus der ARD vom 26.02.2013

Bei der Qualität des Schulunterrichts kommt es auf den Lehrer an - das ist das Fazit des Jahresberichts der Hamburger Schulinspektion. "Entscheidend ist, ob man Herrn Fritz oder Frau Schmidt im Unterricht hat", sagte Schulsenator Ties Rabe (SPD) am Dienstag. Konkret heißt das: Schlechter Unterricht führt zu schlechteren Abi-Noten. Was banal klingt, ist von großer Tragweite für die Schüler.

Bei der Untersuchung sei es erstmals geglückt, einen Zusammenhang zwischen Unterrichtsqualität und Abitur-Noten nachzuweisen, sagte die Leiterin der Schulinspektion, Martina Diedrich. "Dagegen ist der institutionelle Einfluss der Schulform oder der einzelnen Schule minimal." Je nach Qualität des Unterrichts können die Abiturergebnisse der Schüler so um mehrere Punkte schwanken.

Nun wird diskutiert, wie der Unterricht besser werden kann, ohne die Strukturen groß umzubauen. Die Konsequenz aus diesem Ergebnis sieht Schulsenator Ties Rabe (SPD) darin, **Lehrer stärker zu unterstützen, zum Beispiel durch Fortbildungen oder durch Teamarbeit bei der Planung des Unterrichts.** Viele Lehrer wollen guten Unterricht machen, "aber es klappt einfach nicht", so Rabe. "Sich gegenseitig zu beraten und den Unterricht zu verbessern - das wird die große Aufgabe sein."



..... in 50 years a lot has changed in school

1964

What's the meaning of these marks?



2014

What's the meaning of these marks?

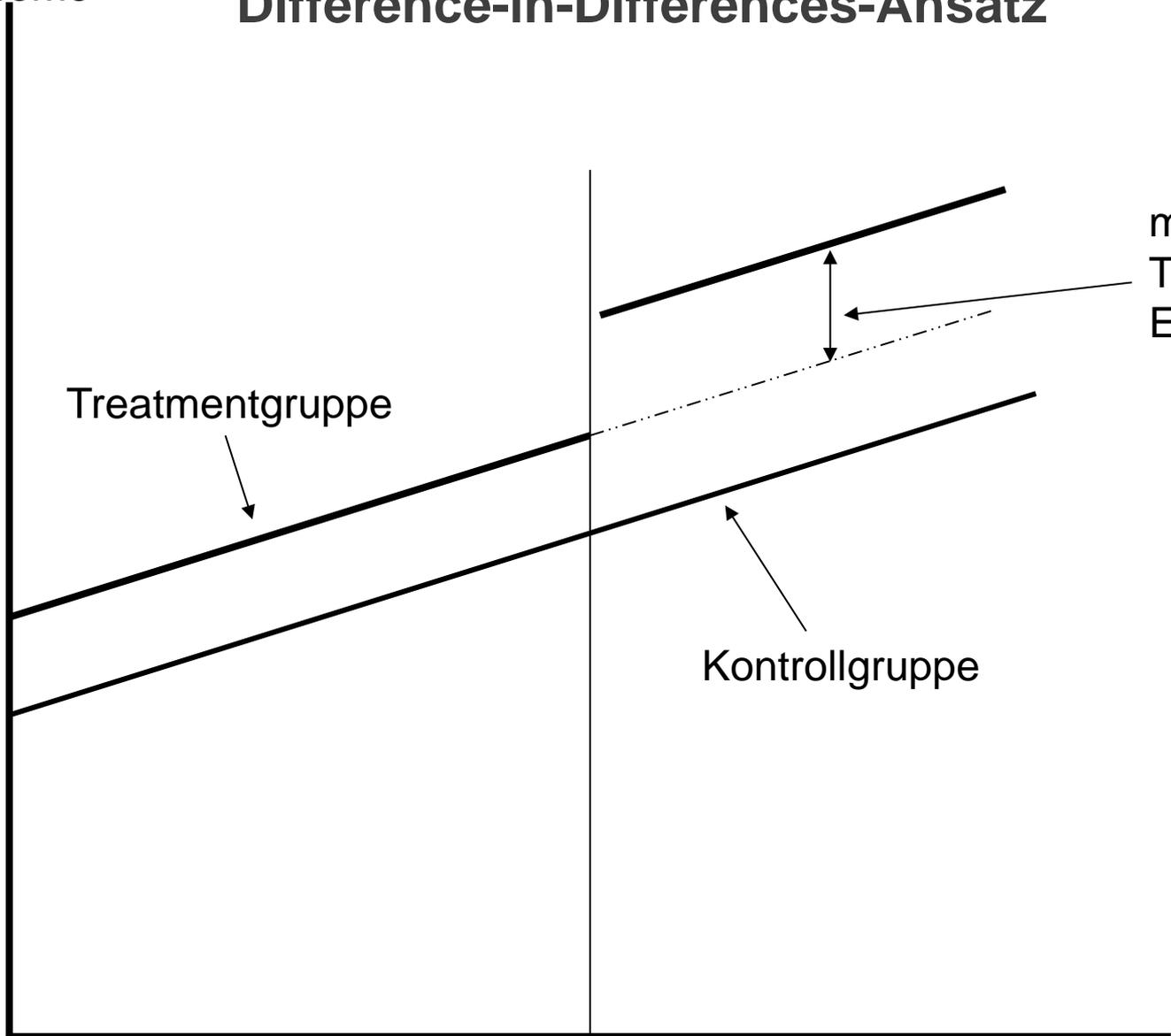


Vielen Dank!



Difference-in-Differences-Ansatz

Outcome



Treatmentgruppe

Kontrollgruppe

mittlerer
Treatment-
Effekt

Treatment

Zeit

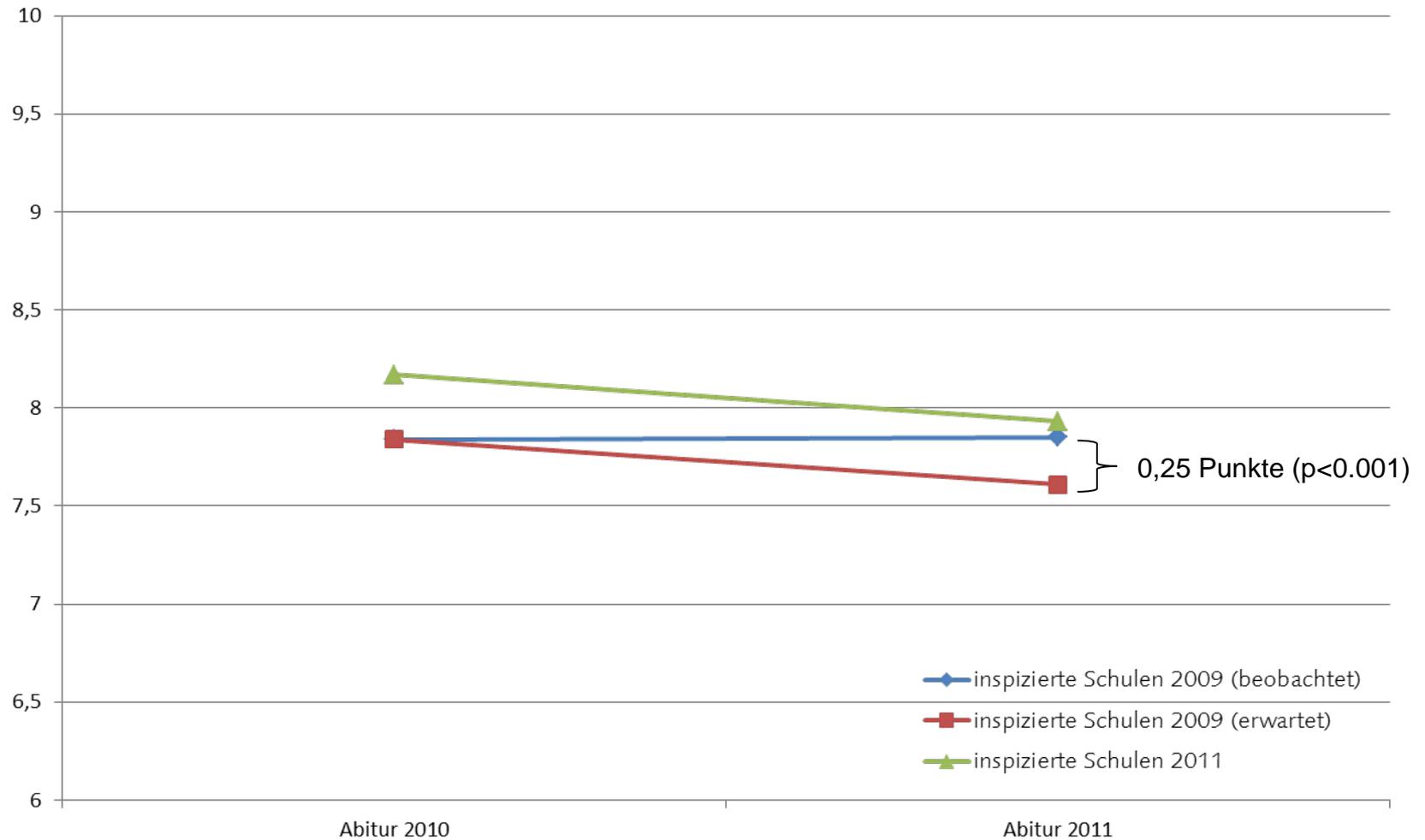


Studie 1: Hamburger Zentralabitur

- Rechtliche Regelungen, die die Verbindlichkeit in der Bewertung regeln, seit 2009 (APO-AH)
- Zentrale Aufgabenstellung seit dem Jahr 2010 in den Fächern Deutsch, Mathematik, 1. Fremdsprache
- *Cross-Section* Difference-in-Differences-Ansatz
- Treatmentgruppe: 21 Schulen aus 2009, SuS:
 - D $t_0=1164$, $t_1=1226$, Ma $t_0=869$, $t_1=795$
- Kontrollgruppe: 17 Schulen aus 2011, SuS:
 - D $t_0=878$, $t_1=894$, Ma $t_0=523$, $t_1=611$

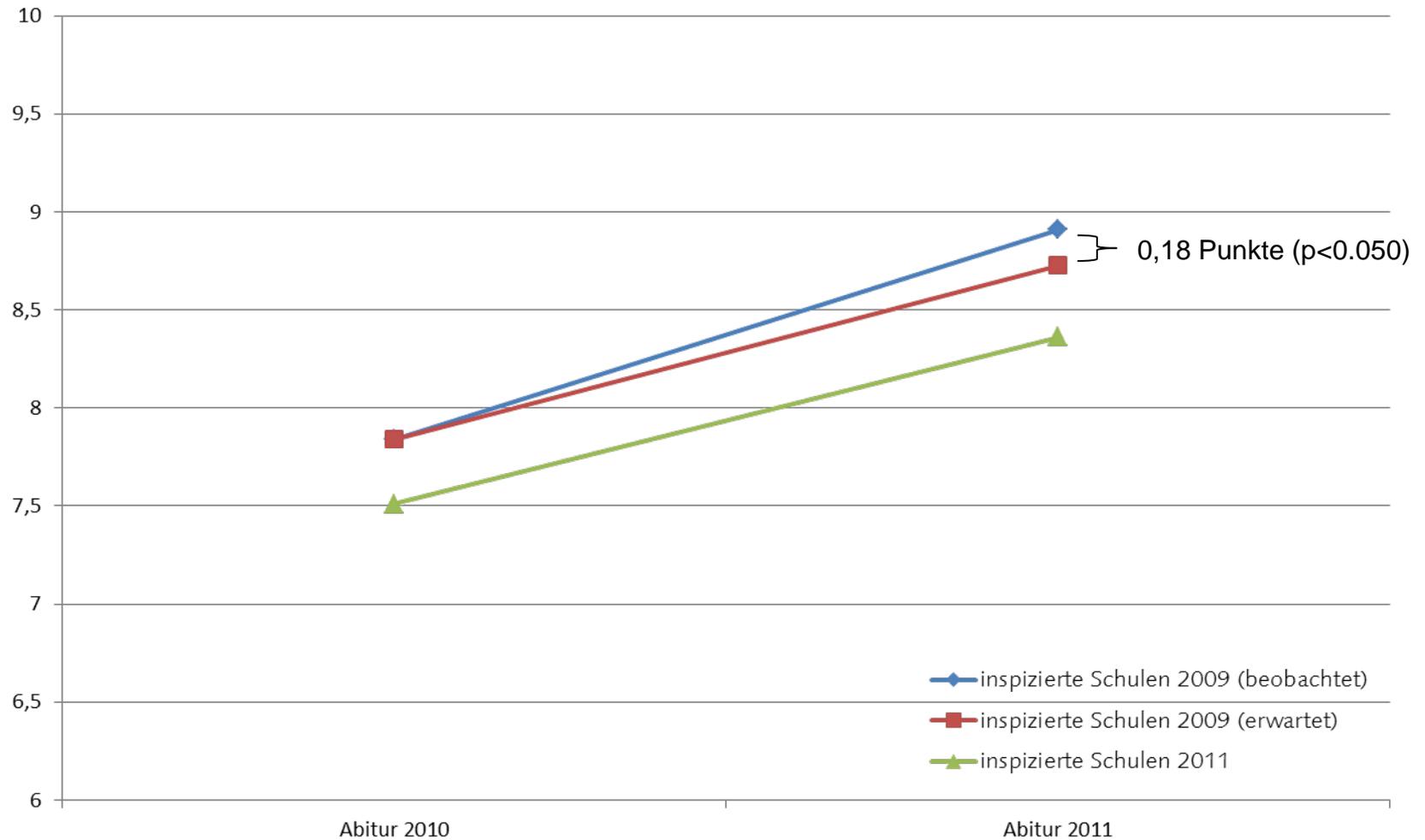


CS-DiD Zentralabitur-Deutsch





CS-DiD Zentralabitur-Mathematik



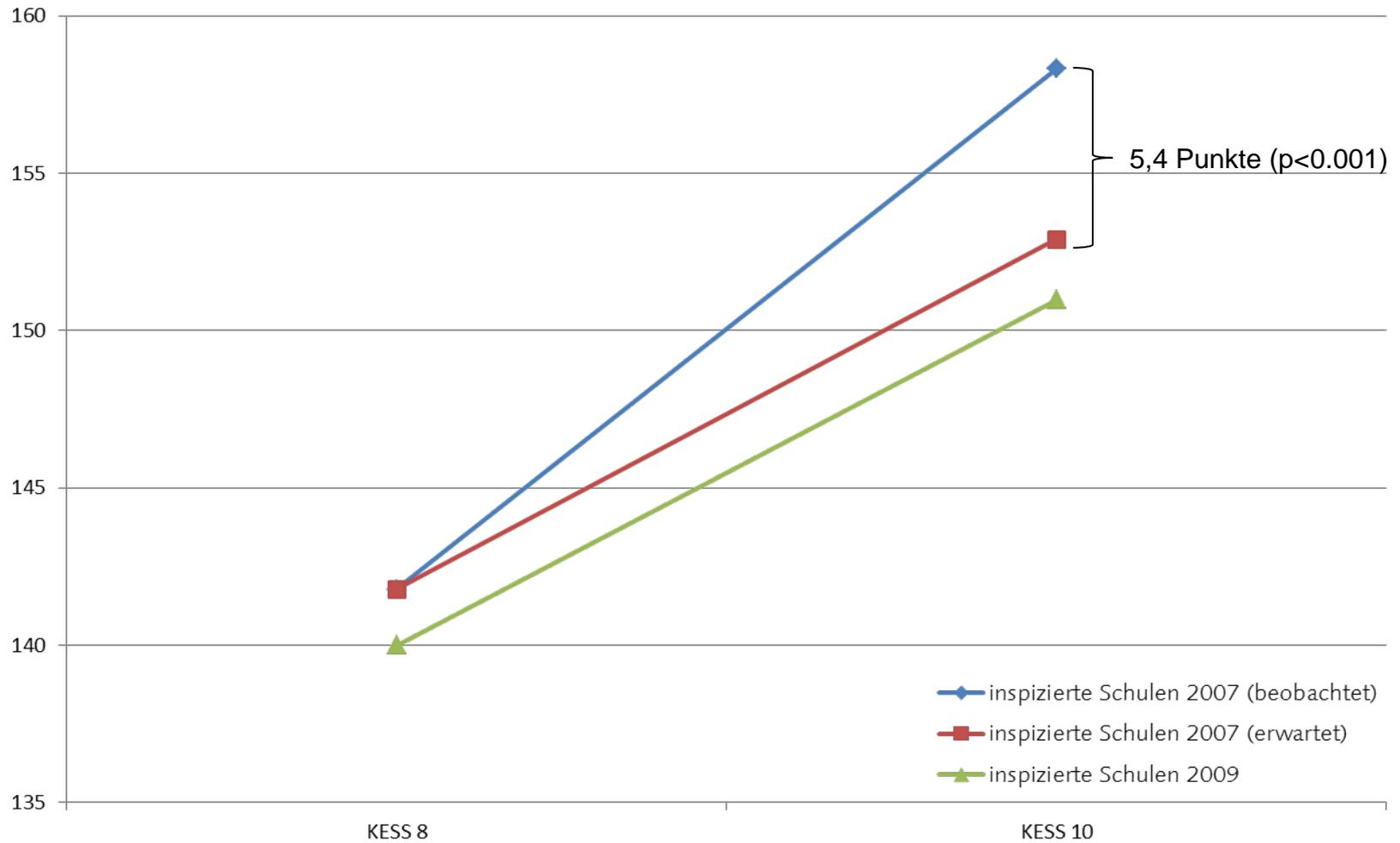


Studie 2: KESS-Studie

- KESS erfasst seit 2003 ca. alle 2 Jahre einen kompletten Schülerjahrgang längsschnittlich
- Zu allen Messzeitpunkten liegen Daten für Deutsch-Leseverständnis und Mathematik flächendeckend vor
- *No Mover* Difference-in-Differences-Ansatz
- Treatmentgruppe: 11 Schulen aus 2007, SuS:
 - D/L $t_0=681$, $t_1=681$, Ma $t_0=683$, $t_1=683$
- Kontrollgruppe: 21 Schulen aus 2009, SuS:
 - D/L $t_0=1209$, $t_1=1209$, Ma $t_0=1207$, $t_1=1207$

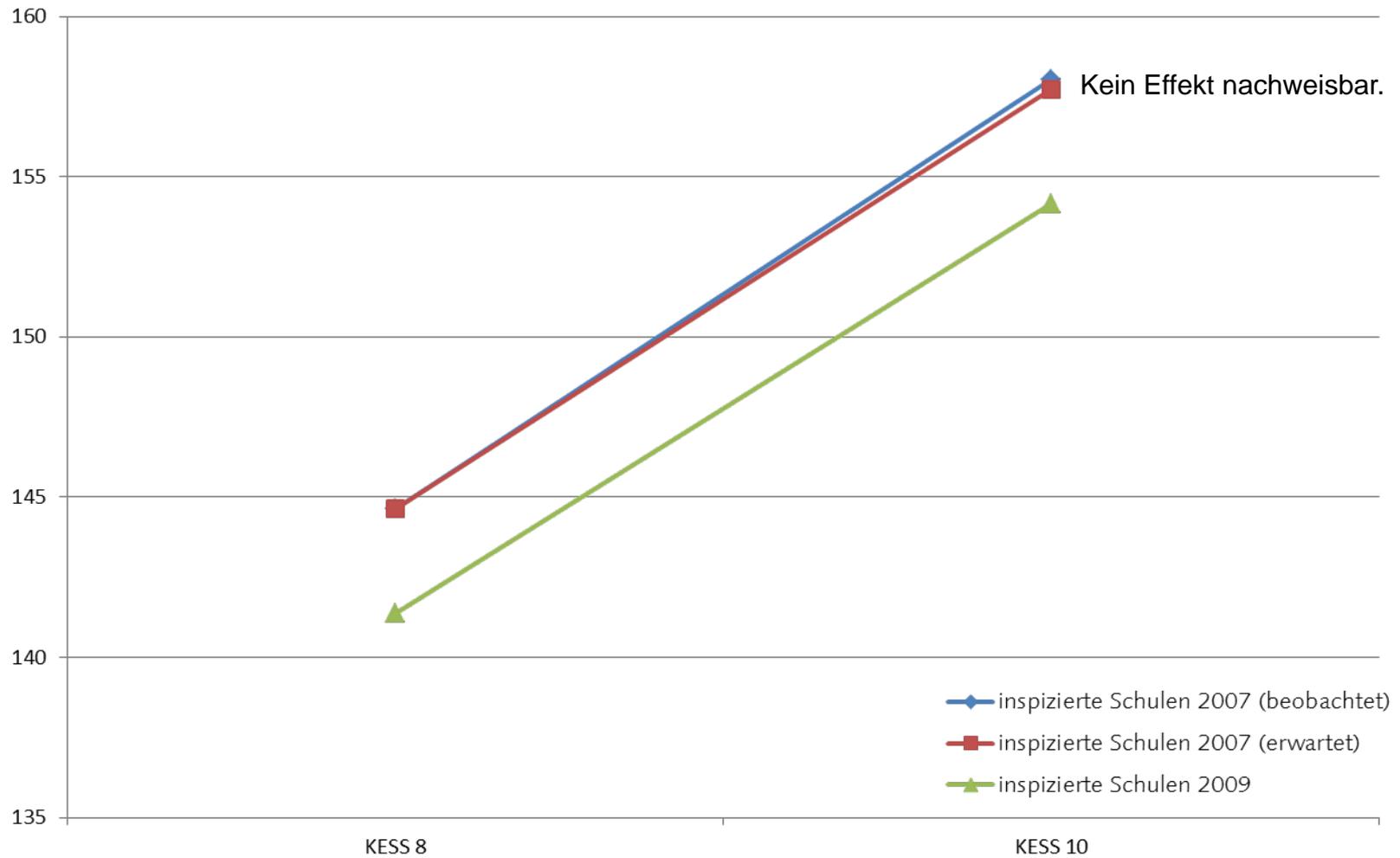


FE-DiD KESS-Leseverständnis



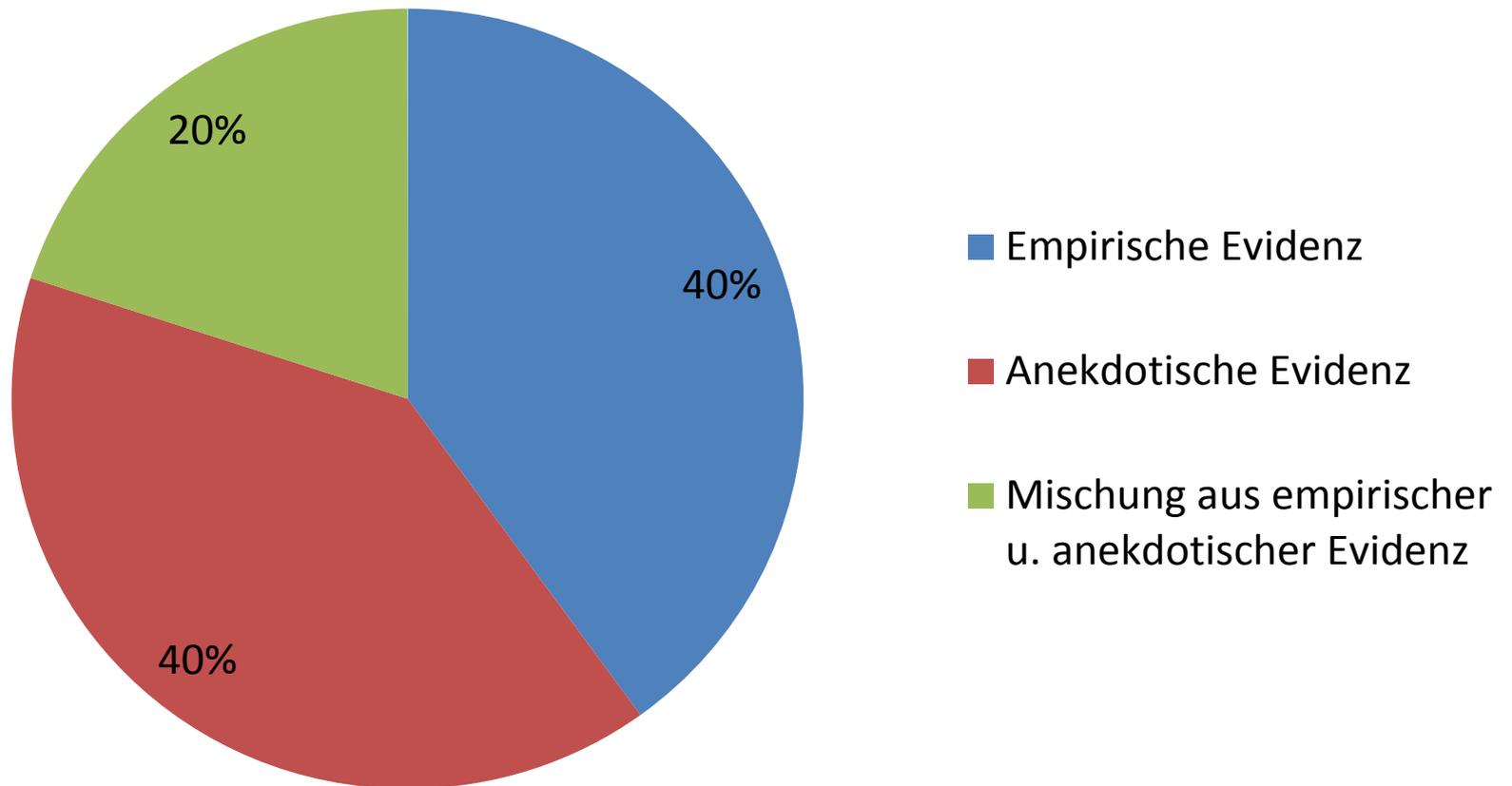


FE-DiD KESS-Mathematik





Welche Informationen werden an Schulen zur Entscheidungsfindung genutzt?





Warum ist empirische Evidenz wichtig?

